中华医学会
第二十九次全国医学信息学术会议

主办单位： 中华医学会　中华医学会医学信息学分会

承办单位： 河南省医学会

协办单位： 开封市人民医院

河南开封

2023年11月9-11日

论文汇编

中华医学会第二十九次全国医学信息学术会议

论文汇编

2023 年 11 月 河南 开封

目 录

论文发言

OR-001 基于机器学习算法的老年慢性冠心病合并糖尿病预测模型构建研究…… 谈军涛,何雨芯,赵文龙 1
OR-002 基于深度学习和情境信息的在线健康社区科普文章推荐算法研究…… 钟玉骏,付延旭,姚艺倍等 9
OR-003 “互联网+医疗健康”的发展与应用研究…… 张钰珠 17
OR-004 基于互联网构建线上线下一体化服务平台…… 刘文容 21
OR-005 基于指标评价的电子病历隐私泄露风险评估…… 王雅晴,安新颖,耿瑞利 24
OR-006 基于机器学习模型公平性及可解释性的肺动脉高压分级诊断研究…… 段一凡,王瑞琦,王浩林等 29
OR-007 促进高血压患者主动参与临床实践的决策辅助工具开发 …… 耿劲松,张怡,马亚强等 37
OR-008 医院信息部门基于低代码平台的系统敏捷开发实践 …… 谢君,蓝灵,何传奇等 42
OR-009 Using the Baidu index to understand Chinese interest in common ocular diseases…… 纪之琳,于琦 51
OR-010 基于机器学习的生物医学论文关联数据可重用性检测研究 .. 周佳茵,钱庆,吴思竹等 61
OR-011 基于融合矩阵的文本相似度计算实现检索结果聚类 …… 赵悦阳 64
OR-012 中国医学卫生健康自主知识体系：创新视域和建设路径 .. 李晓瑛,任慧玲,李军莲等 73
OR-013 基于博弈论组合赋权的医院科研能力云模型评价…… 王琰,杨爽,王雅晴等 77
OR-014 基于 Lasso-SFO-BP 的医药专利价值预测研究…… 殷策,张云秋 81
OR-015 基于深度学习模型的医学文献主题识别与演化路径分析研究…… 王德超,李爱花,李永洁等 88
OR-016 基于真实世界数据的心脏疾病知识库的构建及应用探讨 …… 印玥,胡广伟 96
OR-017 新媒体环境下我国居民健康素养的提升研究与探讨 …… 文怡力,黄友良 103
OR-018 基于 CT 生成的 MRI 与常规 MRI 之间的一致性定量评估…… 李兆同,吴辉群 109
OR-019 ChatGPT 在公共卫生和健康领域的应用思考：以基层医疗卫生服务能力提升为例…… 曾华堂,梁万年,柯夏童等 115
OR-020 基于大数据的静脉血栓栓塞症主动预警预测系统与全程智能管理 …… 曾国军 120
OR-021 基于摄像头的深度学习两阶段训练法测量心率和心率变异性 …… 兰蓝 120
OR-022 结合可解释机器学习方法的新型高性能列线图：基于 MIMI-IV 预测入住 ICU 30 天内中风患者的死亡风险…… 陈浩然,李姣,王敏等 125
OR-023 基于隐私计算的健康医疗科研数据协作网络…… 王昌然,张莺耀,薛乐乐等 133
OR-024 中国居民昼夜节律时空分布及节律扰乱与压力的关系——基于微博活跃度数据…… 李顺利,向菲 136
OR-025 A Random Forest Algorithm-Based Prediction Model for Moderate to Severe Acute Postoperative Pain After Orthopedic Surgery Under General Anesthesia …… 师高翔,王琪,刘格良等 148

OR-026 基于因果表征学习的多中心结直肠癌预后预测模型 …… 秦园炳,田雨,周天舒等 168
OR-027 湖南省三级医院智慧管理的现状与对策…… 莫芳惠,罗爱静,欧阳威等 174
OR-028 智慧医院多院区发展的挑战与对策…… 左秀然,瞿朗,杨国良 181
OR-029 数智化综合运营管理模型构建在医院智慧管理体系建设中的
实践与思考…… 王韫秀,赵玉虹 184

壁报交流

PO-001 基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统的设计 …… 唐跃涛 191
PO-002 上海市某三级甲等专科医院提高住院病案首页数据质量实践 .. 马傲,葛小玲,马晴等 197
PO-003 数字疗法在心理健康领域的发展态势及分析建议…… 李晋芳,谢文照 200
PO-004 数字医疗在肿瘤管理中的应用与展望…… 陈珉惺 204
PO-005 医院网络安全现状及建设内容…… 吴天赐 208
PO-006 浅谈“互联网+医疗”模式下医学信息安全防护与隐私保护…… 郑健,贾晋 210
PO-007 院科两级自助统计方法研究与系统实现…… 桑磊 213
PO-008 泛癌智能应用平台设计…… 刘彦霞,渠鸿竹,方向东 218
PO-009 三级医院评审数据质量监管系统开发体会…… 蔡磊 222
PO-010 医院检验标本混采系统的设计与实现…… 龙凤舞,胡建中,冯嵩等 227
PO-011 中国老年人卫生健康政策主题演化研究…… 刘帅兵,张云秋 232
PO-012 基于三医一体化统筹的数据标准规范实践与思考…… 刘阳,康亚西,陈光焰等 242
PO-013 基于生物信息学的妊娠期糖尿病特征基因挖掘…… 喻慧心,赵玉虹,常青 247
PO-014 人工智能在医学科技查新工作中应用模式和方法探讨 …… 王博文,谢玉谦,吴晓炯等 262
PO-015 国防科技重点领域发展态势情报分析模型构建研究 …… 胡丽娜,郝继英,张谦 267
PO-016 基于 PubMedBert + TextRNN 的肿瘤文献分类模型构建研究 .. 张颖,李晓瑛,刘懿等 274
PO-017 新医科背景下医学信息专业虚拟教研室建设的思考与探索 .. 常青,喻慧心,鲁兴胜等 278
PO-018 数智时代医学信息管理类本科人才核心素养培养框架研究 ——以重庆
医科大学信息管理与信息系统专业为例…… 侯筱蓉,赵文龙,贺向前等 281
PO-019 智能医学工程专业医校企协同全过程育人体系构建与探索 …… 姚蓓,吴辉群 284
PO-020 基于全蛋白质组关联分析探究治疗炎症性肠病的
新型药物靶点…… 郝佳炜,陈淼然,宁宇鑫等 287
PO-021 脑卒中患者的神经系统损伤对预后影响的研究…… 王敏,李姣,王梓阳 289
PO-022 基于因果发现的卒中死亡早期预测模型…… 王梓阳,杨林,王敏等 293
PO-023 基于区块链的医患电子健康数据交换体系的研究…… 胡佳,杨麟,李颜等 298
PO-024 基于 MRI 多个解剖参数预测术后尿失禁的人工智能方法 …… 李佳坤,沈百荣 306
PO-025 基于图神经网络的癫痫脑电发作预测研究…… 石雨菲 307
PO-026 元宇宙医学视角下自闭症儿童康复训练模型构建的探索 …… 钟小云,王康龙,吕艳华 313
PO-027 基于多任务学习融合外部知识的临床实体识别…… 程铭,谭健 316
PO-028 面向智慧医院临床数据中心的多源异构数据融合平台研究与实践 …… 刘阳 321
PO-029 基于区块链的医疗电子病历大数据共享方案…… 刘曙霞,杨斯博,黄文波等 325
PO-030 临床专病数据库的建设现状与展望…… 王耀国,周毅,龙思哲等 327
PO-031 基于上皮细胞-间充质转化特征基因的胃癌预后
风险特征研究…… 王琪,张一凡,李晨龙等 332

PO-032 自然语言处理技术在在线健康社区研究中的
应用现状：系统综述…… 王哲轩,舒泳搏,罗爱静 342
PO-033 基于生物信息学方法筛选原发性转移肾透明细胞癌关键基因及
潜在治疗药物…… 莫状,耿杰,刘格良等 353
PO-034 贵阳市健康医疗大数据治理平台建设实践…… 李挺 359
PO-035 空间转录组技术发展及在肿瘤研究中的应用…… 王国梁,方向东,渠鸿竹 365
PO-036 医院微博的对话沟通能力分析…… 铉婧雯,张韦 369
PO-037 基于扎根理论的在线问诊服务质量研究：服务提供方的视角 …… 刘菁菁 374
PO-038 大型综合医院入院服务中心的信息化建设实践与应用创新 …… 陈启岳 380
PO-039 基于医院信息平台智慧病房的建设…… 曾丽娟 383
PO-040 基于 PACS/RIS 的数字影像服务系统的设计及应用 …… 赵治远 389
PO-041 电子病历档案系统在医院病案标准化管理中的
管理效果、管理规范情况探索…… 王刚 393
PO-042 面向社区居民的数字健康能力评估体系研究…… 谭政,曹锦丹,赵新蓉等 397
PO-043 OBE-CDIO 理念下面向医学信息人才创新实践能力培养的
教学模式优化路径研究…… 张新颖,沈丽宁,张治国 408

书面交流

PU-001 DRG 支付制度背景下医生的非预期行为及有效应对 …… 颜新宇,常维夫,罗爱静 419
PU-002 电子病历应用水平分级机遇与挑战…… 张红坤 440
PU-003 基于 SWOT 分析法浅析电子病历系统应用水平分级实践中的问题与对策…… 邵伟 441
PU-004 医保保障结算清单数据质量优化研究…… 王佳莹 444
PU-005 医院高质量发展下的手术操作编码体系现状与需求分析 …… 刘立煌,黄锋,张霞等 449
PU-006 基于互联网医院平台电子处方流转的云诊室服务…… 陈文豪,褚震宇 451
PU-007 电子病案智能化在健康中国建设中的应用与前景…… 周益权 459
PU-008 某医院 2022 年住院病案终末质量分析 …… 岳馨培 460
PU-009 电子病历风险防范策略研究…… 董一超 464
PU-010 基于五级电子病历的互联网医院系统信息化建设与实践 …… 瞿泽虹,张佳 465
PU-011 浅谈住院电子病历质控系统的应用和体会…… 陈丹 467
PU-012 病案无纸化管理实践及思考…… 王崇梁 470
PU-013 中文医学本体数据库建设现状与发展趋势分析…… 戴维,赵文龙 472
PU-014 智慧医疗背景下全椒县医院信息系统一体化建设研究 …… 陈旭 474
PU-015 关于“互联网+医疗健康”下的全民电子健康记录的思考…… 尚小平 477
PU-016 门诊电子病历建设的国际经验与中国实践…… 吴倩,陈瑶,卓丽军等 480
PU-017 基于人工智能的病案首页智能编码技术研究与应用 …… 温丹,党雅娜 483
PU-018 基于病案首页系统的数据挖掘研究…… 杨林朋 484
PU-019 互联网医疗患者满意度测量量表研究：系统综述…… 杜怡菲,谷雨 486
PU-020 基于手机 APP 延续性护理对心脏机械瓣膜置换术后患者
抗凝治疗的影响…… 丁娜娜,王晓华 492
PU-021 互联网+下健康医疗的融合与发展…… 李振浩,于琦,龚万里等 497
PU-022 基于元分析的用户数字健康素养影响因素研究…… 姚艺倍,曹锦丹,钟玉骏等 499

PU-023 “互联网＋医疗健康”新业态…… 张梦超 499
PU-024 健康管理特征视角下移动健康用户类型及其对移动健康满意度的影响研究…… 王桐,雷健波,王伟 501
PU-025 “互联网+医疗健康”背景下我国互联网医院管理模式及应用现状分析…… 刘尊尊 506
PU-026 关于云胶片服务在医院中的应用与现状分析…… 杨恺熠,俞刚,赵永根等 509
PU-027 优化营商环境背景下互联网医院发展研究—基于 SWOT-PEST 模型分析 …… 严恺,于琦 512
PU-028 基于实证数据的互联网医院线上线下融合模式的价值评估 …… 张萍,乔一鸣 516
PU-029 在线慢性病社区用户的健康信息需求研究——以百度贴吧高血压吧数据为例…… 赵梦圆,应峻 521
PU-030 以患者为中心的互联网医院平台建设实践与思考…… 万宁,刘阳 527
PU-031 数字医疗新风口 互联网医院发展按下“加速键” …… 谢琳艺,何慧敏,翁梦晴等 532
PU-032 医院数字化转型的探索与实践……王建磊,杜建芳, 张晓皓等 533
PU-033 基于太赫兹时域光谱及数据库的中药鉴别现状浅析 …… 彭树岩 534
PU-034 基于太赫兹时域光谱及数据库的中药鉴别现状浅析 …… 彭树岩 534
PU-035 从政策方向看关键技术与“互联网+医疗健康”新业态的形成…… 王叶 537
PU-036 医疗卫生领域商用密码安全评估与建设的实践与思考 …… 刘阳 543
PU-037 大数据时代下医学信息安全与隐私保护…… 龚万里,于琦,侯雅琪等 546
PU-038 零信任技术在医院信息化应用及展望…… 彭浩明,仇成凤 549
PU-039 病案管理中患者隐私的泄露风险及保护措施…… 赵敏 552
PU-040 医院信息安全和隐私保护的政策研究…… 黎颜 554
PU-041 内外网文件安全交换管理系统…… 冯东建 557
PU-042 新冠疫苗犹豫及其影响因素：系统综述…… 李晨露,吴泰来 560
PU-043 医院信息安全与患者隐私保护研究…… 陆杨,邓成生 564
PU-044 实验室生物安全风险及信息应对策略的探讨…… 刘婵桢 566
PU-045 大数据时代的医疗信息隐私保护管理 SWOT 分析 …… 侯梦薇 569
PU-046 医疗数字化背景下的信息安全建设探析…… 敬力源,邹小红 573
PU-047 医院数据隐私保护模型建立及应用实践…… 乔阳阳,韩雪峰 Han Xuefeng 576
PU-048 给医院管理者的几点网络安全建议…… 韩雪峰 579
PU-049 大数据时代下医学信息安全与保护…… 彭宇 581
PU-050 医院信息系统数据资产分类分级探索与实践…… 王鹏,马现 583
PU-051 我国个人健康信息保护现状研究…… 袁艺峰 588
PU-052 医疗信息系统在生殖中心临床决策上的支持与应用 …… 栗晔 590
PU-053 DIP 背景下基于改良波士顿矩阵模型大型公立医院专科发展对策研究…… 罗慧中,程楠 592
PU-054 医院信息系统与临床决策辅助 CDSS 应用的管理规范和应用效果…… 王秉慧 597
PU-055 基于 CDR 的信息闭环管理助力医院高质量发展…… 朱骄锋 599
PU-056 医疗信息系统临床评价提升问题初探…… 孙勇 601
PU-057 医学信息系统助力临床决策…… 智鹏 603
PU-058 数智化协同模式下的全医嘱质效双控决策支持平台构建 …… 徐蕾 604
PU-059 基于 PDCA 循环法的临床决策支持系统应用进展…… 解亚楼,黄心莹 607

PU-060 问题驱动的静脉血栓栓塞症智能临床决策支持体系构建 …… 任晓霞,黄桃 609
PU-061 基于物联网技术的手术医疗行为管控系统设计与思考 …… 刘阳 612
PU-062 放射报告分配流程优化…… 黄娟,屈健宁 615
PU-063 2014-2018 年河南省某三甲医院超长住院日患者临床特征分析 …… 隋美丽 618
PU-064 基于淋巴结阳性率等多因素预测完全切除的 III-N2 期非小细胞肺癌
术后放疗的生存期的机器模型的开发和验证…… 郭强,魏益平 621
PU-065 人事报工管理系统在医院精益化管理中的应用…… 朱骄锋 623
PU-066 日间手术模式下的信息化建设方向探究…… 曹靖霄,刘燕 624
PU-067 HIS 存储设备升级及性能优化 …… 韦吉程 631
PU-068 智能化自助入出院便民服务的创新研究与应用…… 黄盛明,王际强、骆振龙等 633
PU-069 医院信息化项目建设与管理…… 尹彭,刘寅 636
PU-070 新时代党建引领下医院防统方系统的运用与管理…… 唐鑫洁 639
PU-071 基于文献计量学的医院人力资源管理与信息化建设 …… 林静 641
PU-072 微服务架构在医院信息系统建设中的应用…… 赵霞,姚佳璇,曾艺等 645
PU-073 多院区医院互联网出口整合改造与安全建设探索…… 宫晨曦,郝晓伟,许锋波等 649
PU-074 以互联网“云平台”构建基层高效便捷超声质控体系 …… 杨希,李雪,刘潇等 655
PU-075 加速推进医联体双向转诊建设的研究与探讨…… 曾婷 658
PU-076 经济不发达地区三甲医院 信息化建设与项目管理 …… 孙盛轶,姜永峰 662
PU-077 医疗高值耗材自动计费的应用研究…… 周逸帆,周敏,倪剑 665
PU-078 基于 PMC 指数模型的卫生健康信息化建设政策评价分析 … 鲁志威,王剑莉,贺莲等 668
PU-079 国家医疗保障局公开通报欺诈骗保典型案件分析及防范对策探讨 …… 王宇彤 675
PU-080 基于 SWOT 方法分析中医药在防治传染病中的
发展契机 ——以新冠疫情防治为例 …… 薛丹阳,于琦,岳玉婷 678
PU-081 基于集聚度和 GIS 技术的湖北省卫生资源配置公平性评价…… 高梦婷,汤梓菲,袁磊 684
PU-082 数字健康应用现状与挑战…… 肖月,邱英鹏,游茂等 687
PU-083 我国基层数字健康发展现况及未来展望…… 肖月,邱英鹏,游茂 691
PU-084 一体化电子病历系统在临床应用的研究…… 高卫光 693
PU-085 手术操作编码标准体系的国际经验与启示…… 刘立煌,黄锋,张霞等 697
PU-086 贵阳市健康医疗大数据标准制定研究…… 邵建华 700
PU-087 高通量测序在结直肠癌的研究进展…… 赵思琪,方向东,渠鸿竹等 705
PU-088 “双一流”背景下临床医学 ESI 学科预测模型构建及
提升路径研究…… 李日,李文臣,徐行等 707
PU-089 气管镜介入技术的研究趋势及热点的 VOSviewer 聚类分析…… 徐行,李日,伦志军 714
PU-090 “互联网+教育”视域下知识图谱在医学教育的研究进展…… 徐瑾,卢洁,王亚军 723
PU-091 基于生物信息学筛选 IgA 肾病 Hub 基因…… 陈凯 729
PU-092 基于定量转录组与免疫浸润分析探究线虫感染对衰老的
影响机制…… 张力中,于琦,陈浩然等 733
PU-093 国内外知情同意能力研究热点的共词聚类分析…… 朱彦霞,华南,吴玉彬 734
PU-094 基于 GEO 和 TCGA 数据库探究 肝癌早期诊断标志物 …… 陈凯 744
PU-095 体力活动对炎症性肠病风险的影响和 BMI 的介导作用：孟德尔
随机化研究…… 赵宇慧,李红霞,贾林霖等 747
PU-096 指南制订对文献检索的要求与实践…… 丁文婧 757

PU-097 基于生物信息学分析方法探索霍奇金淋巴瘤一线治疗药物的优势与不足 …· 张昊军 760
PU-098 基于知识发现工具 Bitola 探索自闭症关联药理物质 …… 王康龙,钟小云,吕艳华 760
PU-099 关于清宫痹症医案用药规律的研究…… 吕艳华,李琦,梁傲雪等 761
PU-100 近十年全球老年多重慢病研究趋势及热点演化分析 …… 王洁,冯佳,余丹等 764
PU-101 基于 Web of Science 可视化分析 ChatGPT 的应用研究热点…… 李娟 770
PU-102 基于 citespace 的复杂网络在中医药研究中应用分析 …… 潘晔,赵移畛 779
PU-103 高校图书馆中文图书纸电同步采访关键因素分析…… 邓岩彬,刘强,刘玮 783
PU-104 Venetoclax 耐药机制研究热点与趋势分析——基于 CiteSpace
文献计量视角…… 吴利俊,王倩,应峻 786
PU-105 基于本体的糖尿病知识图谱构建方法研究…… 满坚平,唐圣晟,杨瑞等 793
PU-106 基于疾病网络的血友病并发症挖掘与关联发现…… 郜杨芳,咎彭 797
PU-107 基于知识发现工具 BITOLA 与 Arrowsmith 探索叶酸辅助
治疗自闭症的潜在机制…… 钟小云,王康龙,吕艳华 806
PU-108 医学语料库研究及其临床应用…… 赵浚植,戴维,赵文龙 809
PU-109 基于语义三元组和网络药理学的抑郁症药物知识发现 …… 郜杨芳,苏鑫 810
PU-110 新型冠状病毒肺炎医学影像学研究热点与研究前沿的
知识图谱分析…… 姜海,苑旸,潘希娟 822
PU-111 基于知识图谱的高血压及肾病共病挖掘…… 洪怡敏,刘金林,张晗 829
PU-112 生物信息学分析非小细胞肺癌差异表达关键基因的研究 …… 齐荣暄,于琦 836
PU-113 融合多源数据与方法的疾病-基因-药物知识发现模型 …… 郜杨芳,华国旻 844
PU-114 基于生物医学文献和图挖掘算法发现糖尿病-糖尿病肾病
复杂病因网络及路径分析…… 胡拯涌,段一凡,李佳明等 857
PU-115 基于用户画像的情报跟踪服务医学科研探讨…… 李海生 860
PU-116 Faculty Opinions 最新评分的多维数据源下生物医学科技
论文评价指标研究…… 许丹 863
PU-117 从无效审查程序看中国生物医药领域专利质量的
关键影响因素…… 单治易,关陟昊,王奎芳等 869
PU-118 河北省医学情报事业发展现状及对医学科技成果评价的
构想…… 张秀菊,翟俊霞,刘玉华等 876
PU-119 中国护士社交媒体依赖影响因素及作用机制研究：基于湖南省的
横断面调查…… 康春良,谢文照 879
PU-120 新形势下医学科技查新助推基层科研的 对策 …… 刘玉华,李向黎,张秀菊等 887
PU-121 通过质量和疗效一致性评价药品的供应可视化分析——基于
山东省医疗机构药品使用数据…… 孙雅群,宋燕 890
PU-122 基于 BERTopic 与 Citespace 的中医药领域研究热点与
主题演化分析…… 张文婷,王安莉 897
PU-123 基于网络药理学探讨定喘宜肾宜肾汤对肾透明细胞癌的
作用机制…… 李渊,于琦,齐荣暄等 904
PU-124 热点主题探测方法应用现状及在医学情报领域的实践 …… 徐源 912
PU-125 文献计量视角下国家谈判药品落地政策分析…… 强滟 918
PU-126 基于 CiteSpace 的指纹识别技术发展应用趋势分析 …… 何阿妹 919
PU-127 医学信息多维服务系统集群建设实践及研究…… 段永璇,甄天民,宋燕 924

PU-128 基本心理需要理论下微博用户信息行为的人格特质预测研究 …… 刘帅兵,张云秋 925
PU-129 融合多源数据特征及语义关系的创新前沿识别…… 陈逸菲,李永洁,孙轶楠等 933
PU-130 2021 年度上海市区级医院科研竞争力分析 …… 李娜 937
PU-131 中医药治疗原发性肝癌的用药规律及研究热点分析 …… 秦三利 945
PU-132 基于第三方评价数据的国家临床医学研究中心遴选方法构建与
实证研究…… 牟燕,何有琴 948
PU-133 基于 t-SNE 降维算法和 K-means 聚类的 HIV 感染者/AIDS 患者群体
在线健康社区信息需求分析…… 钟华,陈红,夏松 963
PU-134 大数据时代情报跟踪服务对医学科技的影响…… 王刚 967
PU-135 不同角色预设下的人工智能模型文献筛选…… 刘宇扬,伍丽群,田倩男等 969
PU-136 基于医疗社交媒体用户评论主题及情感的药物推荐策略研究——
以高血压为例…… 郜杨芳,魏宇虹,吴娟等 973
PU-137 浅谈医院图书馆在助力临床及科研发展服务中的作用 …… 顾显伟,何丽 973
PU-138 医学图书馆医学情报服务发展现况及趋势…… 杨山石,沙小苹,王常颖等 976
PU-139 中华优秀传统文化体验提升医学生人文素养实践研究 …… 刘强 981
PU-140 数字环境下医学图书馆推进特色读者服务的思考…… 周楠,王瑶,王维岩等 983
PU-141 河南省洛阳正骨医院（河南省骨科医院）科技论文产出
计量分析…… 刘瑞霞,丰锐,古想花等 986
PU-142 叙事医学视角下医院图书馆阅读推广服务模式研究 …… 秦方,侯胜超,胡鸿 990
PU-143 智慧图书馆视角下新一代图书馆服务平台的研究现状和思考 …… 冯研 993
PU-144 基于 DEA-Malmquist 的高校图书馆电子资源采购效率评价 …… 郭雪梅 996
PU-145 改进物元的开放存取资源质量综合评价模型…… 杨筠,孙艳,宋蕊等 1002
PU-146 NSTL 开放获取文献资源保障研究 …… 王勖,任慧玲,王巍洁等 1008
PU-147 蒙医药文献数字化平台建设的现状研究…… 赵宏亮 1017
PU-148 网络时代如何有效提升图书馆信息化建设水平…… 罗燕玲 1020
PU-149 医学信息素养教育理论与实践…… 刘文峰 1024
PU-150 Hotspots and research trends relating to infectious-disease-specific
health literacy in the past decade: a bibliometric study …… 严波,陈晃荣,谢瑶等 1026
PU-151 智慧医疗背景下医护人员信息素养探究…… 古林林,王琳凯 1026
PU-152 医学院校图书馆多元信息素养教育服务体系建设研究 …… 张帆 1028
PU-153 医院信息化人才队伍建设研究及对策…… 王文峰,陈贝贝,李砚 1030
PU-154 医学信息素养教育理论实践…… 吴含 1031
PU-155 慢性病用户健康信息搜寻行为的影响因素探索…… 赵梦圆,应峻 1035
PU-156 高校医学图书馆与医学生信息素养教育的思考…… 李华 1041
PU-157 数智时代医学生信息素养教育的内涵与培养路径初探 …… 刘靓靓,聂小莺,胡波 1042
PU-158 医工融合的创新创业教育生态研究——以重庆医科大学
智慧医学学科群建设为例…… 王佳,赵文龙,贺向前 1052
PU-159 智能化时代背景下医学信息素养教育研究…… 崔宇鹏 1054
PU-160 大数据视角下国内医学信息人才培养研究热点及需求研究 …… 周跃 1056
PU-161 我国医药信息管理研究现状、热点与趋势——基于 SATI 的可视化分析…… 喻真 1059
PU-162 “新医科”背景下的医学生数字素养培育与创新思考 …… 江帅卿,马敬东,秦选斌等 1064
PU-163 数据科学在医疗保健发展中的作用…… 张超 1067

PU-164 基于 CAS 理论的图书情报专业硕士创新实践能力提升体系构建…… 杨梦柯,沈丽宁 1071
PU-165 基于数字科研培训平台本科生深度参与暑期科研项目的
探索研究…… 梁宏锌,周海榆,刘盛波等 1071
PU-166 “智慧医疗”背景下医院信息化复合型人才培养探究 …… 韩明泽,彭晨辉 1072
PU-167 基于 OBE-CDIO 理念的医学信息人才实践教学体系构建 …. 符凌嘉,张治国,沈丽宁 1074
PU-168 医学信息学未来的发展与展望…… 李新龙 1080
PU-169 智能医学发展与思考…… 卜庆录 1083
PU-170 医疗信息教育与人才培养…… 谢添 1084
PU-171 浅析医学信息学发展和教育…… 孙楠 1087
PU-172 基于现状的基层全科医生医学信息继续教育发展对策研究 …… 孙翔飞,于琦 1090
PU-173 人工智能 GPT 技术在智慧医院应用中的机遇与思考…… 马青,孟啸 1093
PU-174 基于 5W1H 分析法的元宇宙智慧养老服务研究…… 阳杉,王安莉 1095
PU-175 心房颤动导管消融手术患者管理中的机器学习：系统综述 …… 陈炜,罗爱静,辛梓睿 1098
PU-176 基于 ChatGPT 的医生诊疗 AI 助手系统的设计与实现 …… 殷鸣 1099
PU-177 基于机器学习的药物组合协同预测方法综述…… 徐梦蝶,王俊杰,周海航等 1104
PU-178 人工智能技术在手术器械清点中的应用研究…… 靳超 1110
PU-179 移动医疗用于分娩镇痛管理的研究…… 刘新元,尚小平,杨林朋等 1113
PU-180 虚拟诊室在慢阻肺防治中的应用与展望…… 胡莉娟,胡译心,张静等 1114
PU-181 5G 技术在智慧医疗领域中的应用探讨…… 张靓 1117
PU-182 基于深度学习的心血管病风险预测模型研究…… 洪石陈,王逸嘉,丁祎等 1121
PU-183 基于 5G 技术的医院双域无线专网建设与应用 …… 左锋,陈利佳,胡冰等 1126
PU-184 基于物联网技术的智慧病房设计与实现…… 郝晓伟,许锋波,卫新乐等 1130
PU-185 面向儿科医学试题的参考解析推荐方法…… 王娟,侯丽,孙月萍等 1136
PU-186 基于新一代 5G 技术的智慧医院建设研究与应用 …… 许锋波,卫新乐,郝晓伟等 1140
PU-187 基因预测多发性硬化症的发生会增加患严重 COVID-19 的
风险…… 胡景喜,王茜,郝佳炜等 1143
PU-188 基于人工智能的 PI-RADS 自动评分在前列腺癌 MRI 诊断中作用
与展望…… 康亚西,王忠尧,刘阳 1148
PU-189 人工智能在麻醉学中的应用…… 闫丽璇 1152
PU-190 人工智能在医疗领域的应用与思考…… 周枫 1154
PU-191 医疗云应用的前景及优势…… 詹汉强 1157
PU-192 基于 RSM 框架的生存预测模型 …… 唐朝飞 1161
PU-193 基于边缘计算架构的医疗设备智慧物联管理平台的构建及应用 …… 陆靓亮,金伟 1164
PU-194 通过转化基因组学确定多发性硬化易感性的蛋白靶标 …… 王茜,王学琴,吕毅康等 1169
PU-195 代谢相关基因作为预后生物标志物的泛癌分析…… 黄泰,刘格良,莫状等 1172
PU-196 医疗信息化新需求下的桌面虚拟化…… 王旭 1183
PU-197 5G 物联网下生命体征智能采集系统设计及实现…… 彭洋,仇成凤 1187
PU-198 生成式人工智能在智慧医疗方面的实现方法研究…… 张萍,乔一鸣 1191
PU-199 浅谈智慧医院管理中检验信息化…… 赵旭涛 1193
PU-200 信息化在检验结果自动审核中的应用…… 赵旭涛 1195

PU-201 Immunotherapeutic response and prognosis prediction in lung adenocarcinoma based on lncRNA signatures of activated CD4+ memory T cells …… 刘盛波,周海榆,李洽轩等 1196
PU-202 智慧医院 5G 专网模型构建与应用 …… 任晓强,方志远 1197
PU-203 基于新一代物联网技术的急诊绿色通道时间采集系统的设计 …… 李晓莉 1201
PU-204 5G+WiFi6 医院物联网安全的探索与研究 …… 史乐萌 1204
PU-205 智慧医院物联网建设的应用…… 付宇,郭伟 1204
PU-206 基于机器学习的乳腺癌治疗靶标生物活性预测…… 唐圣晟 1207
PU-207 基于深度学习的外展神经自动分割成像初步研究…… 孙钰茗,杨汉丰 1213
PU-208 基于深度学习面神经磁共振图像自动分割…… 曾祥鑫,杨汉丰 1219
PU-209 健康医疗大数据支持下的脑胶质瘤诊治科研初探…… 李旭 1223
PU-210 基于深度学习的眼底图像辅助诊断…… 邵家辉 1229
PU-211 基于 5G 的全院心电网络建设与应用效果 …… 高标,张旸 1234
PU-212 人工智能在肺癌研究中的应用…… 王昕玥,渠鸿竹,方向东 1237
PU-213 基于人工智能的免疫检查点抑制剂反应预测模型：现状与前景…… 陈永森,渠鸿竹,方向东 1240
PU-214 智能医学在推进社区医院建设中的应用…… 时文睿,方向东,渠鸿竹 1242
PU-215 基于病人诊断与检验数据缺失的智慧诊疗…… 张海翔,陈香梓,李培培 1244
PU-216 医学知识大模型语料数据集构建研究 …… 吕婷钰,李晓瑛,刘宇炀等 1249
PU-217 基于机器学习的成人严重细菌感染临床时序预警模型:一项回顾性研究…… 刘剑,裘云庆 1253
PU-218 基于混合云架构的医院信息系统安全外联及管控实践 …… 沈亮,周敏 1255
PU-219 医疗行业云桌面系统技术架构与应用场景的探讨…… 何斌 1261
PU-220 基于深度学习的血清类型分类研究…… 张志国,吴红,叶庆 1266
PU-221 人工智能在血液学诊断中的应用与展望…… 张立娟,方向东,渠鸿竹等 1266
PU-222 医院远程介入手术协作平台的设计与实现…… 林志刚,陈明 1267
PU-223 基于 HIS 系统的基础医嘱数据整理与思考 …… 罗扬,刘征 1273
PU-224 基于主题模型的青少年网络行为公众关注热点研究——以知乎平台“青少年网络行为”相关话题为例 …… 刘灵涛,鲍俊,黄锦泉等 1275
PU-225 ICU 患者医院感染现况调查研究…… 张亚锋,刘新元,尚小平等 1281
PU-226 神经退行性疾病：生活习惯与智慧医疗…… 毕成,沈百荣 1281
PU-227 “健康中国”背景下健康医疗大数据资源的应用与治理 …… 侯雅琪,于琦,龚万里 1283
PU-228 基于医院综合信息集成平台的 DRGs 管理系统的分析…… 郭佳 1285
PU-229 医疗保险对农村中老年人健康状况和生活满意度的影响 --基于 CHARLS 队列研究 …… 路亚柯,刘新奎,赵艳婷等 1288
PU-230 大数据时代医院统计如何为医院管理服务…… 赵鹏飞 1293
PU-231 浅谈医院信息网络安全建设与管理…… 张红霞,周麒,薛五洋 1295
PU-232 基于可解释性机器学习的慢性心力衰竭合并肺部感染患者的生存分析…… 申采玉 1297
PU-233 肾透明细胞癌微环境预后相关基因的组学分析…… 李卓阳,贺培凤,于琦 1306
PU-234 基于在线健康社区的老年健康信息需求研究——以肺癌、高血压为例…… 张楚,郑颖,黄锦泉等 1306

PU-235 县域三甲医院的数据治理实践 …… 田梅 1314
PU-236 在线问诊平台用户情感影响因素及知识图谱构建研究 …… 杨爽,王琰,王雅晴等 1316
PU-237 大数据技术在中医药传承与创新中的应用…… 黄尤江 1322
PU-238 基于选举多方安全计算实现的临床研究联邦学习框架 …… 陈廷寅,冯嵩 1325
PU-239 基于大数据技术的医院资源规划系统应用与思考…… 冯标斌 1332
PU-240 新时代数据治理技术促进智慧医疗发展…… 黄虹 1336
PU-241 基于泛在物联网感知的大数据双碳管理及决策系统建设方案研究 …… 段涛,邹小红 1342
PU-242 基于 ETL 的心肺血管疾病专病数据库建设与实践…… 窦一峰,单淑香,庞东旭等 1348
PU-243 大数据时代背景下医院信息管理创新发展的工作讨论 …… 潘传红 1349
PU-244 多模态医疗健康数据分类与应用…… 葛令乔,渠鸿竹,方向东 1352
PU-245 欧洲 ELIXIR 生物医学数据工具服务平台体系
实践与启示…… 刘盛宇,吴思竹,胡拯涌等 1355
PU-246 基于人工智能的多组学分析技术助力精准医学…… 郑光敏,渠鸿竹,张昭军等 1360
PU-247 基于研究者视角的生物医学转录组数据库发展现状和建议 .. 李晨龙,王琪,刘格良等 1371
PU-248 健康医疗大数据的应用与治理…… 温丹,党雅娜 1375
PU-249 大数据在驱动智能精准多学科协作中的实践…… 郭涵怡,杨谨菲,左秀然 1376
PU-250 基于条件对抗生成网络的数字病理图像标准化方法 …… 柴振新,田雨,邓博洋等 1382
PU-251 “健商”培养对高校学生体育锻炼的影响研究
——以江西中医药大学为例…… 邵哲 1383
PU-252 基于知识图谱的全科未分化疾病小样本辅助推理方法研究 .. 郑奕骁,田雨,刘强华等 1391
PU-253 体数据的自由投影与平行切割体素绘制实现…… 李心舟,李晨,季达峰 1392
PU-254 硝苯地平的多中心临床疗效评价及患者概况研究…… 王婷婷,赵文龙 1396
PU-255 公立医院绩效考核背景下我国远程医疗的发展契机与思考 …… 王硕果 1397
PU-256 区域医疗中心模式下远程云服务平台的构建与探索 …… 王娜,褚震宇,周潞荣等 1400
PU-257 我国远程医疗发展现状分析…… 秦超雨 1401
PU-258 基于数字化平台的智慧健康社区建设案例分析-以深圳市为例 …… 赵靓 1403
PU-259 基于逻辑组学及药物表观精准对接技术筛选新型铁死亡诱导药物治疗
c-MYC/Bcl-2 重排的高级别 B 细胞淋巴瘤及基础实验验证 …… 李泓毅 1404
PU-260 智能医学的理念创新和范式突破的研究…… 智鹏 1404
PU-261 MTOR and type 1 diabetes risk: a Mendelian
randomized study …… 王灿,陈雨欣,郑蕾蕾等 1405
PU-262 全球电子素养的现状及影响因素：一项 meta 分析 …… 钟华,刘茜文,陈红 1411
PU-263 创建高级别电子病历助力医疗质量管理实践与创新 …… 侯彦东 1412
PU-264 基于价值医疗视角下医后付服务应用研究…… 汪滔,彭杰 1414
PU-265 基于先进医院信息系统的智慧医院建设…… 崔建国 1417
PU-266 影像检查智慧服务流程优化实践…… 汪伟,郭晓宙,王海星等 1420
PU-267 信息化在医院后勤管理的应用分析…… 周麒,张红霞 1424
PU-268 智慧医院建设…… 李若斌 1427
PU-269 基于大数据的科室耗材可视化闭环管理实践…… 周继光,李咏梅,方秀云等 1431
PU-270 浅析以党建引领智慧医院建设实践与思考…… 刘琦 1435
PU-271 业务驱动医疗服务项目收费实行研究…… 郭建军,徐挺玉,黄忠秋等 1440
PU-272 基于患者视角的智慧医院建设研究综述…… 李欣悦 1443

PU-273 新冠肺炎疫情防控常态化下如何提升智慧医疗服务 …… 李欣悦 1445
PU-274 公立医院连续性医疗服务体系的信息化路径研究…… 左秀然,彭明毅,杨国良 1448
PU-275 医院智能化建设及其在智慧医疗发展中的意义…… 郑健,薛志刚 1450
PU-276 基于内控设计的非医疗服务项目收费平台设计与应用 …… 陈云,刘剑锋,梁诗弋 1454
PU-277 一院多区可追溯智慧系统助力医疗物资精细化管理的
应用实践…… 张司露,李慧慧,郭金鸽 1456
PU-278 门诊慢特病跨省直接结算实践与思考…… 张继民 1457
PU-279 基于 RFID 技术构建固定资产全生命周期管理体系
实现全流程智能化管理…… 金迟 1460
PU-280 DRG 支付模式下医院精细化管理的思路与愿景 ……张远林,蔡文君, 卢熙等 1465
PU-281 基于微信公众号平台的公立医院病案复印申请与邮寄流程创新研究 …… 侯娜娜 1469
PU-282 智能医疗背景下医患交流的类型、特征与优化策略 …… 范美玉,赖昕 1472
PU-283 病例智能分组与规范化诊疗质控管理系统的实践与探索 .. 秦金华,田晓琼,欧成川等 1475
PU-284 科学与技术交叉视角下智能医学领域研究热点与技术预测研究 …… 殷策,张云秋 1483
PU-285 构建适合我国国情的医疗 AI 技术评估体系的探讨 …… 何达,李嫣然,苏锦英 1492
PU-286 基于患者特异性 CCTA 的左前降支多尺度建模
——0D-1D 耦合模型与 0D-3D 耦合模型的比较…… 刘杉峰,李伍各,吴辉群 1494
PU-287 深度学习在冠脉 CTA 影像中的应用研究进展 …… 李瑾,孟艺菲,周彤彤等 1499
PU-288 智能医学的发展与应用…… 吕倩 1500
PU-289 医疗数据创新延伸——“数据胶囊”的建设思考…… 杨雪 1503
PU-290 运用 PDCA 循环提升儿科门诊辅助诊断决策支持系统应用效果…… 尚于娟,葛小玲 1505
PU-291 智能医学应用现状和创新发展…… 丁立旻 1506
PU-292 基于“互联网+”背景下国内外慢病管理研究热点的
比较分析…… 孙雪超,鲁兴胜,赵玉虹 1508
PU-293 合理用药系统在医院中的应用与管理…… 周千惠,彭晨辉 1512
PU-294 智慧医疗安全防护之道——安全“脱管”不如安全托管 …… 苏子仁,刘阳,曹港华 1516
PU-295 从查新项目统计分析广东省医疗卫生领域人工智能应用现状 …… 李翠霞,郑利荣 1521
PU-296 医院信息化建设存在的问题及思考…… 王春杰 1524
PU-297 基于 BERT 模型的乳腺癌影像学报告中 TNM 分期实体识别
预测方法…… 唐现策,许延杰,燕青等 1528

摘要交流

ZY-001 浅谈智慧医院背景下的病案数据质量提升…… 魏梦迪 1533
ZY-002 全结构化电子病历的发展及困境…… 王奥钰 1533
ZY-003 浅谈系统方法在现在化病案管理的应用…… 李芳 1533
ZY-004 基于电子健康卡的急诊预检分诊管理系统的设计…… 唐跃涛 1534
ZY-005 基于病案无纸化管理的管理实践…… 雷保仓 1535
ZY-006 ICU 患者医院感染直接经济损失评价…… 刘新元,尚小平,杨林朋等 1535
ZY-007 DRGs 实施背景下公立医院高质量发展的思考…… 韩慧欣 1536
ZY-008 关于麻醉临床信息系统启用手术 ICD 编码的思考…… 任田 1536
ZY-009 论医学信息标准研究与应用…… 智鹏 1537

ZY-010 中国大学生对突发公共卫生事件健康信息需求的
定性研究…… 黄丹蓝,赖雨婷,周旭毓等 1537
ZY-011 基于表型网络的白化病共表型分析与预测研究…… 郜杨芳,蔡月琴 1538
ZY-012 我国短视频应用研究现状和热点分析…… 孙伟伟 1538
ZY-013 知识产权在创新企业中的发展趋势…… 张宁雅 1539
ZY-014 国内外医学院校图书馆研究文献的可视化分析…… 蔡汉虹 1539
ZY-015 浅谈元宇宙与图书馆的融合共生…… 马子龙,王馨 1540
ZY-016 努力建设一流的生命医学图书馆…… 王俊 1540
ZY-017 医院信息化数据平台的应用分析探究…… 付宇 1541
ZY-018 智慧运营管理助力医院高质量发展 …… 唐宇 1541
ZY-019 移动医疗在化疗患者管理骨髓抑制中应用的范围综述 …… 张琳,王晓华,杨莉 1543
ZY-020 “智慧医院”下的医疗装备全生命周期管理探讨…… 张利中 1543
ZY-021 智慧医疗、智慧服务、智慧管理 三位一体智慧医院建设
——智慧医疗在医院高质量发展中的应用…… 郝心歌 1544
ZY-022 山西省肿瘤医院规范化诊疗系统的探索与应用…… 高翔,田园 1544
ZY-023 复合闭环下临床用血管理系统的构建…… 康宁 1546
ZY-024 基于物联网的智慧医疗信息化关键技术探究…… 汤友喜 1547
ZY-025 智慧医院建设中的信息技术应用与挑战分析…… 许赞 1548
ZY-026 关于移动报账系统在医院的应用…… 李腾,刘珂,姜博 1550
ZY-027 智慧医疗对患者就诊，医护工作的优势分析…… 马欣钰 1552
ZY-028 统一预约系统助力传染病医院提升就医体验…… 张现辉 1552
ZY-029 论述智能医学的应用现状与创新发展…… 张嘉成 1554
ZY-030 医院智能随访系统的探讨…… 张骏 1554
ZY-031 人工智能在医疗信息服务中的应用研究…… 朱佳棋 1555
ZY-032 健康管理中心智能导检方案的研究与应用…… 熊佳佳 1555
ZY-033 肺腺癌预后相关转录组分子免疫分型及其特征综合分析 …… 谢戴鹏,尤培蒙,周海榆 1557

论文发言

OR-001

基于机器学习算法的老年慢性冠心病合并糖尿病预测模型构建研究

谈军涛[1]，何雨芯[2]，赵文龙[3,4△]
1.重庆医科大学附属巴南医院运营管理办公室
2.重庆医科大学附属巴南医院医务科
3.重庆医科大学医学信息学院
4.重庆医科大学医学数据研究院

1 引言

近年来，随着我国老龄化的加剧，慢性冠心病的临床发病率越来越高，已成为威胁老年人健康的重大疾病[1, 2]。糖尿病是一组代谢性紊乱性疾病，老年慢性冠心病患者一旦合并糖尿病，其生活质量和预后都将大幅下降[3]。研究表明在糖尿病患者中，76.4%的人存在至少一种并发症[4]。与没有糖尿病的成年人相比，患有糖尿病的成年人发生心血管并发症的风险增加了 2-4 倍，并且随着血糖控制的恶化，这种风险也会增加。糖尿病与成人死亡率增加 75%有关，心血管并发症占全因死亡率的很大一部分[5]。

近年来，研究逐渐开始探索老年慢性冠心病合并糖尿病的相关影响因素，并及时预测患者的病情[6-8]。鉴于此，本研究将利用 9 种主流机器学习算法探索老年冠心病患者合并糖尿病的影响因素并构建预测模型，研究结果可为重点人群糖尿病的早期预防、护理干预与指导提供参考依据。

2 资料与方法

2.1 数据来源

用于本研究老年慢性冠心病合并糖尿病预测模型构建的数据来源于重庆医科大学医学数据研究院平台。重庆医科大学医学数据研究院平台集成了 7 家附属三级医院的电子病历数据，数据量大，代表性好。抽取 2015 年 1 月-2021 年 12 月期间 7 家医疗机构中确诊为慢性冠心病的老年住院患者作为研究对象。此外我们还收集了 A 医院 2021 年 1 月-2023 年 8 月期间符合研究标准的老年慢性冠心病患者数据。

2.2 纳排标准

纳入标准：（1）年龄≥65 周岁；（2）符合慢性冠心病（包括稳定型心绞痛、缺血性心肌病和隐匿性冠心病）诊断标准。

排除标准：（1）恶性心律失常、肺栓塞、肺源性心脏病、先天性心脏病、非缺血性心肌病及其他心脏瓣膜病等非冠心病导致的糖尿病者；（2）需要器械辅助治疗的冠心病糖尿病患者；（3）合并肾脏、肝脏等其他重要脏器功能不全者；（4）合并严重感染者；（5）合并神经系统疾病者；（6）合并恶性肿瘤，合并免疫系统、血液系统及甲状腺疾病；（7）并发精神疾病的患者。

2.3 研究分组

建模组和内部验证组：纳入 2015 年 1 月-2019 年 12 月期间符合研究标准的老年慢性冠心病患者，并按照 7：3 的比例随机划分为建模组和内部验证组。最终建模组纳入研究样本 7816 例，内部验证组纳入研究样本 3350 例。

时间外部验证组：纳入 2020 年 1 月-2021 年 12 月期间符合研究标准的老年慢性冠心病患者作为时间外部验证组。最终时间外部验证组纳入研究样本 1763 例。

空间外部验证组：同时收集 A 医院 2021 年 1 月-2023 年 8 月期间符合研究标准的老年慢性冠心病患者作为空间外部验证组。最终空间外部验证组纳入研究样本 3445 例。

2.4 研究指标

在既往研究的基础上，选择了 36 个预测老年急性冠心病合并糖尿病的潜在危险因素，即年龄、性别、吸烟史、饮酒史、高血压、心力衰竭、高脂血症、慢性胃炎、舒张压、γ 谷氨酰基转

移酶、白细胞计数、中性粒细胞与淋巴细胞比率（NLR）、淋巴细胞与单核细胞比率（LMR）、血小板与淋巴细胞比率（PLR）、血肌酐、总胆红素、直接胆红素、全血钾、全血钙、尿酸、尿素、总蛋白、白蛋白、碱性磷酸酶、大型血小板比率、平均红细胞容积、红细胞分布宽度、血小板分布宽度、血红蛋白浓度、红细胞计数、丙氨酸氨基转移酶、门冬氨酸氨基转移酶、甘油三酯、总胆固醇、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇。

2.5 统计学方法

本研究采用 SPSS22.0 和 R 4.0.2 软件进行数据统计分析。计量资料服从正态分布的采用 $\bar{x}\pm s$ 表示，组间比较采用 t 检验；不服从正态分布的采用中位数和四分位数间距[M(P25~P75)]表示，组间比较采用 Mann-Whitney U 检验。计数资料采用百分比表示，组间比较采用 χ^2 检验。缺失数据采用蒙特卡洛多重插补法进行数据插补。

首先进行单因素分析，接着将单因素有统计学差异的指标纳入 Lasso 回归筛选最终的建模指标。构建模型应用了包括逻辑回归（LR）、随机森林（RF）、决策树（DT）、支持向量机（SVM）、极端梯度提升（XGBoost）、人工神经网络（ANN）、K 近邻算法（KNN）、轻量梯度提升机（lightGBM）以及朴素贝叶斯（NB）在内的 9 种主流机器学习算法。模型评价指标包括 ROC 曲线下面积 AUC 值、灵敏度、特异度、阳性预测率、阴性预测率、约登指数、F1 score 等。所有统计分析均为双侧，统计学显著性设定为 $P < 0.05$。

3 结果

3.1 一般结果

建模组糖尿病组和非糖尿病组在年龄（$z = -2.616$, $P=0.009$）、高血压（$\chi^2 = 468.192$, $P<0.001$）、高脂血症（$\chi^2=56.16$, $P<0.001$）、舒张压（$z=-3.254$, $P=0.001$）、白细胞计数（$z=-7.704$, $P<0.001$）等指标上差异存在统计学意义（$P<0.05$）；慢性胃炎、肌酐、尿酸、碱性磷酸酶、血小板分布宽度、血红蛋白浓度、总胆固醇以及低密度脂蛋白胆固醇等指标不存在统计学差异（$P>0.05$）。其余情况详见表 1。

表 1 建模组糖尿病组和非糖尿病组基线对比

指标	非糖尿病组(5685)	糖尿病组(N=2131)	统计值	P 值
年龄[岁, $M(P_{25}, P_{75})$]	75.00(70.00,81.00)	75.00(70.00,81.00)	-2.616	0.009
性别(男)[例(%)]	2409(42.46)	843(39.34)	6.26	0.012
吸烟史[例(%)]	1609(28.36)	521(24.31)	12.873	<0.001
饮酒史[例(%)]	1250(22.03)	410(19.13)	7.832	0.005
高血压[例(%)]	2837(50.01)	1653(77.13)	468.192	<0.001
心力衰竭[例(%)]	1481(26.11)	650(30.33)	14.002	<0.001
高脂血症[例(%)]	580(10.22)	351(16.38)	56.16	<0.001
慢性胃炎[例(%)]	768(13.54)	264(12.32)	2.016	0.156
舒张压[mmHg, $M(P_{25}, P_{75})$]	78.00(70.00,87.00)	78.00(70.00,85.00)	-3.254	0.001
γ 谷氨酰基转移酶[IU/L, $M(P_{25}, P_{75})$]	24.00(16.00,42.00)	25.40(17.00,43.00)	-2.725	0.006
白细胞计数[$\times10^9$/L, $M(P_{25}, P_{75})$]	6.16(5.00,7.80)	6.58(5.36,8.22)	-7.704	<0.001
NLR	3.25(2.19,5.46)	3.10(2.19,4.99)	-2.129	0.033
LMR	3.46(2.24,5.15)	3.92(2.62,5.62)	-7.274	<0.001
PLR	139.76(102.44,198.92)	123.50(91.97,172.64)	-9.492	<0.001
肌酐[umol/L, $M(P_{25}, P_{75})$]	69.67(56.80,86.90)	69.50(54.30,88.83)	-0.503	0.615
总胆红素[umol/l, $M(P_{25}, P_{75})$]	11.00(8.20,14.60)	10.20(7.70,13.82)	-4.809	<0.001
直接胆红素[umol/l, $M(P_{25}, P_{75})$]	3.90(2.80,5.50)	3.80(2.70,5.30)	-3.191	0.001
全血钾[mmol/L, $M(P_{25}, P_{75})$]	3.94(3.66,4.23)	4.02(3.71,4.31)	-5.492	<0.001
全血钙[mmol/L, $M(P_{25}, P_{75})$]	2.23(2.13,2.32)	2.24(2.14,2.34)	-2.706	0.007
尿酸[umol/L, $M(P_{25}, P_{75})$]	324.70(263.30,398.00)	323.80(264.20,399.05)	-0.272	0.786
尿素氮[mmol/L, $M(P_{25}, P_{75})$]	5.95(4.76,7.56)	6.21(5.01,7.99)	-5.526	<0.001

总蛋白[g/L, $M(P_{25}, P_{75})$]	67.63(62.90,72.44)	69.20(64.60,73.98)	-8.55	<0.001
白蛋白[g/L, $M(P_{25}, P_{75})$]	39.60(36.50,42.60)	40.10(37.05,43.10)	-4.152	<0.001
碱性磷酸酶[IU/L, $M(P_{25}, P_{75})$]	74.00(61.00,90.00)	73.00(59.90,91.00)	-1.24	0.215
大型血小板比率	33.10(26.40,40.10)	34.20(27.50,41.17)	-4.565	<0.001
平均红细胞容积[fL, $M(P_{25}, P_{75})$]	92.60(89.20,95.80)	90.90(87.80,94.10)	-12.536	<0.001
红细胞分布宽度[%, $M(P_{25}, P_{75})$]	13.50(12.90,14.30)	13.40(12.90,14.10)	-2.527	0.012
血小板分布宽度[%, $M(P_{25}, P_{75})$]	15.50(12.80,16.50)	15.50(12.80,16.50)	-1.028	0.304
血红蛋白浓度[g/L, $M(P_{25}, P_{75})$]	128.00(116.00,139.00)	128.00(116.00,139.00)	-0.563	0.573
红细胞计数[$\times10^{12}$/L, $M(P_{25}, P_{75})$]	4.20(3.83,4.57)	4.24(3.86,4.60)	-2.363	0.018
丙氨酸氨基转移酶[IU/L, $M(P_{25}, P_{75})$]	16.55(12.00,25.00)	17.85(13.00,26.83)	-4.809	<0.001
门冬氨酸氨基转移酶[IU/L, $M(P_{25}, P_{75})$]	21.40(17.66,28.00)	20.00(16.00,26.00)	-8.994	<0.001
甘油三酯[mmol/L, $M(P_{25}, P_{75})$]	1.15(0.85,1.62)	1.43(1.02,1.99)	-15.729	<0.001
总胆固醇[mmol/L, $M(P_{25}, P_{75})$]	4.25(3.53,5.01)	4.21(3.47,5.03)	-1.309	0.191
高密度脂蛋白胆固醇[mmol/L, $M(P_{25}, P_{75})$]	1.24(1.02,1.49)	1.11(0.92,1.35)	-13.539	<0.001
低密度脂蛋白胆固醇[mmol/L, $M(P_{25}, P_{75})$]	2.30(1.76,2.90)	2.28(1.73,2.91)	-0.882	0.378

3.2 建模指标筛选

将表 1 中经单因素分析具有统计学差异的指标进行 Lasso 回归。Lasso 回归采用 10 重交叉验证方法选择合适的惩罚项 λ。结果显示最小化均方误差（Mean Squared Error,MSE）的 λ 为 0.0008235759，λ+1se 为 0.01144355。图 1 呈现了随着 log（λ）的增加 MSE 值的变化。Lasso 回归在 λ+1se 处共筛选出了以下 14 个重要指标：高血压、高脂血症、舒张压、白细胞计数、LMR、PLR、钾、尿素、总蛋白、大型血小板比率、平均红细胞容积、红细胞分布宽度、甘油三酯和高密度脂蛋白胆固醇。

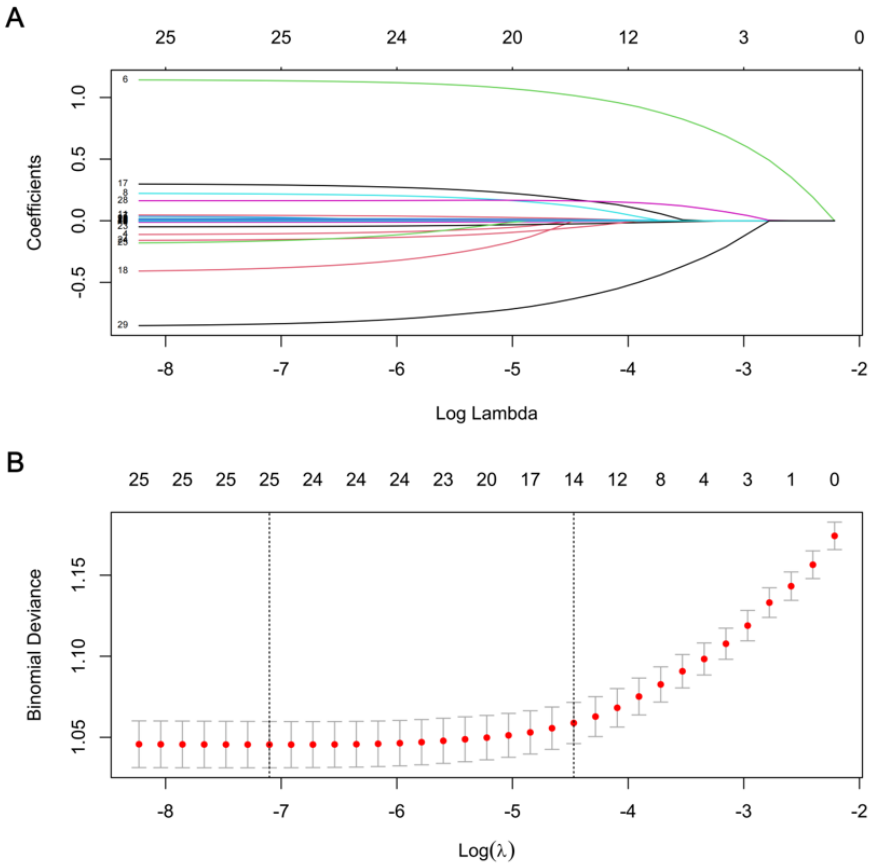


图 1 LASSO 回归指标筛选。A 显示了不同 log（λ）对应的指标数量；B 显示了 LASSO 回归调谐参数选择的十折交叉验证的结果。

3.3 模型性能对比

将最终筛选出的 14 个指标分别代入 9 种机器学习算法中构建糖尿病预测模型。模型效能评价指标包括 ROC 曲线下面积 AUC 值、敏感度、特异度、阳性预测率、阴性预测率、约登指数、F1 score 等。

图 2 展示了 9 种机器学习算法构建的糖尿病预测模型在建模组中的 ROC 曲线。表 2 详细罗列了 9 种机器学习算法构建的糖尿病预测模型的各个预测效能评价指标对比结果。结果表明 XGBoost 构建的糖尿病预测模型具有最高的 ROC 曲线下面积 AUC 值 0.896 (95%CI: 0.888-0.903)、灵敏度 0.816 (95%CI: 0.799-0.832)、特异度 0.812 (95%CI: 0.802-0.822)、阳性预测率 0.621 (95%CI: 0.603-0.639)、阴性预测率 0.921 (95%CI: 0.914-0.929)、F1 score 为 0.705，约登指数为 0.628。共有三种机器学习算法（XGBoost、SVM 和 lightGBM）构建的预测模型的 ROC 曲线下面积 AUC 值超过了 0.8。DT 构建的预测模型性能最差，ROC 曲线下面积 AUC 值仅为 0.667 (95%CI: 0.655-0.680)。

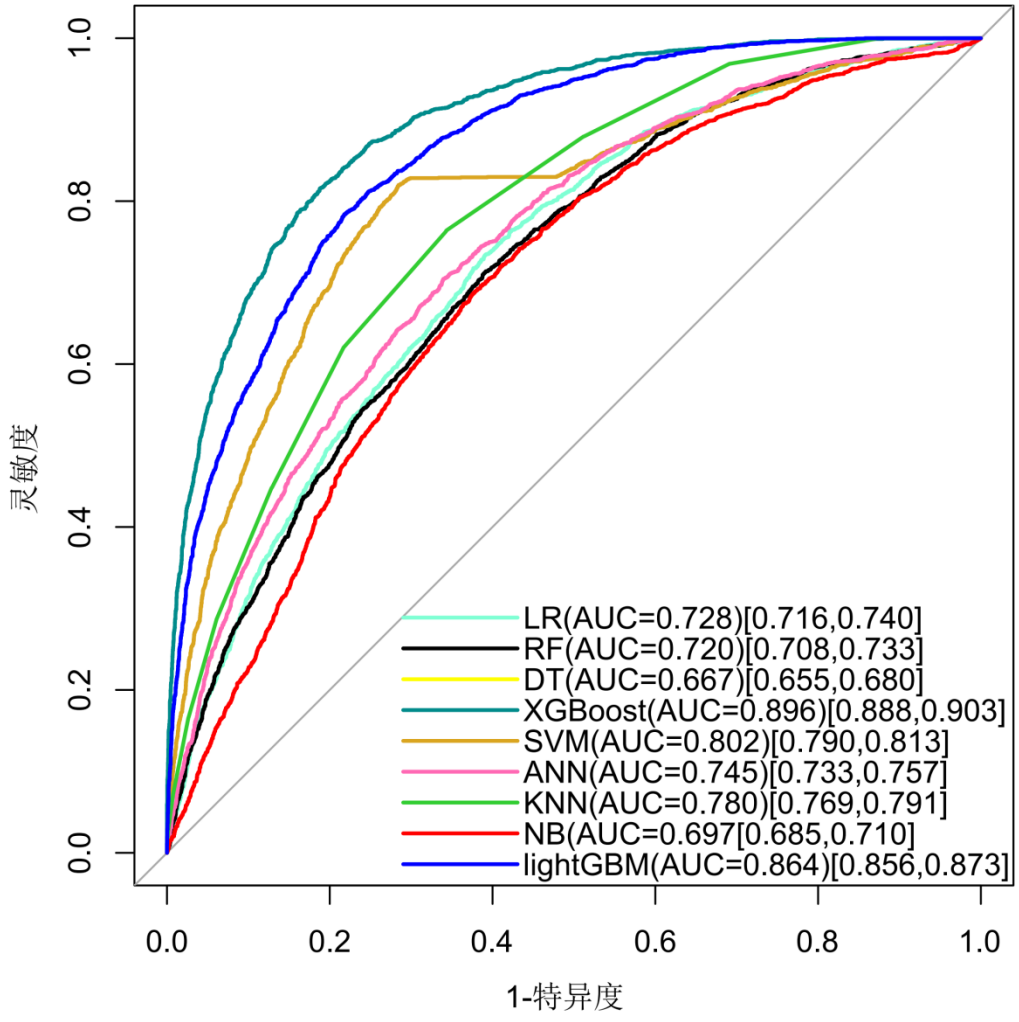


图 2 9 种预测模型的 ROC 曲线

表 2 9 种模型预测性能评价指标对比

模型	AUC (95%CI)	灵敏度 (95%CI)	特异度 (95%CI)	阳性预测率 (95%CI)	阴性预测率 (95%CI)	*F1* score	约登指数
LR	0.728 (0.716,0.740)	0.734 (0.715,0.752)	0.609 (0.597,0.622)	0.415 (0.399,0.431)	0.858 (0.847,0.869)	0.530	0.343
RF	0.720 (0.708,0.733)	0.714 (0.695,0.733)	0.608 (0.596,0.621)	0.408 (0.392,0.424)	0.849 (0.838,0.860)	0.519	0.322
DT	0.667 (0.655,0.680)	0.771 (0.754,0.789)	0.500 (0.487,0.513)	0.368 (0.354,0.382)	0.853 (0.841,0.865)	0.498	0.271
XGBoost	0.896 (0.888,0.903)	0.816 (0.799,0.832)	0.812 (0.802,0.822)	0.621 (0.603,0.639)	0.921 (0.914,0.929)	0.705	0.628
SVM	0.802 (0.790,0.813)	0.820 (0.804,0.836)	0.715 (0.704,0.727)	0.521 (0.504,0.538)	0.913 (0.905,0.921)	0.637	0.535
ANN	0.745 (0.733,0.757)	0.705 (0.685,0.724)	0.660 (0.648,0.673)	0.439 (0.423,0.456)	0.855 (0.845,0.866)	0.541	0.365
KNN	0.780 (0.769,0.791)	0.764 (0.746,0.782)	0.657 (0.6540.669)	0.457 (0.440,0.473)	0.880 (0.871,0.890)	0.572	0.421
NB	0.697 (0.685,0.710)	0.687 (0.668,0.707)	0.624 (0.611,0.636)	0.408 (0.392,0.424)	0.841 (0.830,0.852)	0.512	0.311
lightGBM	0.864 (0.856,0.873)	0.784 (0.767,0.801)	0.782 (0.771,0.793)	0.576 (0.558,0.594)	0.906 (0.897,0.914)	0.664	0.566

3.4 模型临床实用性分析

此外，我们绘制了 DCA 决策曲线来评估各个模型的临床实用性。其中纵轴表示临床净获益，横轴表示阳性阈值概率；蓝线横线表示所有老年慢性冠心病患者都判定未合并糖尿病，均不给予任何干预措施，临床净获益为 0；绿色斜线表示所有老年慢性冠心病患者都判定为合并糖尿病，都给予对应的干预措施，临床净获益为一条斜率为负的反斜线。结果表明 XGBoost 构建的预测模型的整体临床实用性仍然是最好的，反观 DT 模型和 NB 模型在阈值概率为 50%时，临床净获益就已经为 0，表示模型的临床实用性相对较差。其余情况详见图 3。

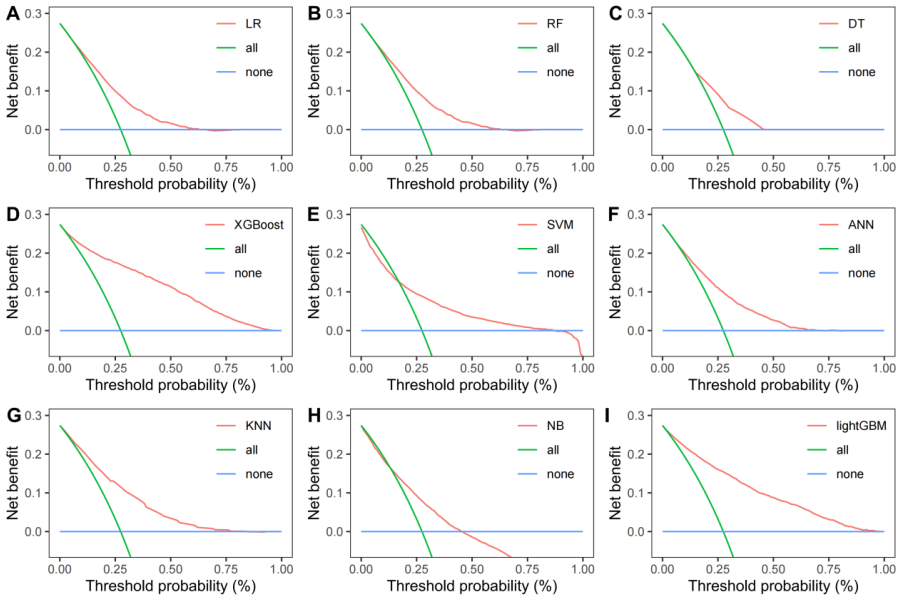


图 3 9 种预测模型的 DCA 曲线。(A) LR 的 DCA 曲线；(B) RF 的 DCA 曲线；(C) DT 的 DCA 曲线；(D) XGBoost 的 DCA 曲线；(E) SVM 的 DCA 曲线；(F) ANN 的 DCA 曲线；(G) KNN 的 DCA 曲线；(H) NB 的 DCA 曲线；(I) lightGBM 的 DCA 曲线。

3.5 风险因素重要性分析

接着我们利用 SHAP 算法对最优 XGBoost 模型进行解释，根据解释结果得到对纳入模型的风险因素对于老年慢性冠心病患者合并糖尿病的影响和相对重要性排名。其中风险因素重要性排名是根据训练集中全部患者的各个特征对应的 SHAP value 的绝对值总和由高到低进行排序，需要指出的是，每个风险因素的重要性都是在与其他风险因素的交互作用下得到的。图 4 展示了基于 SHAP 分析得到的对于老年慢性冠心病患者合并糖尿病风险升高最重要的 14 个风险因素。

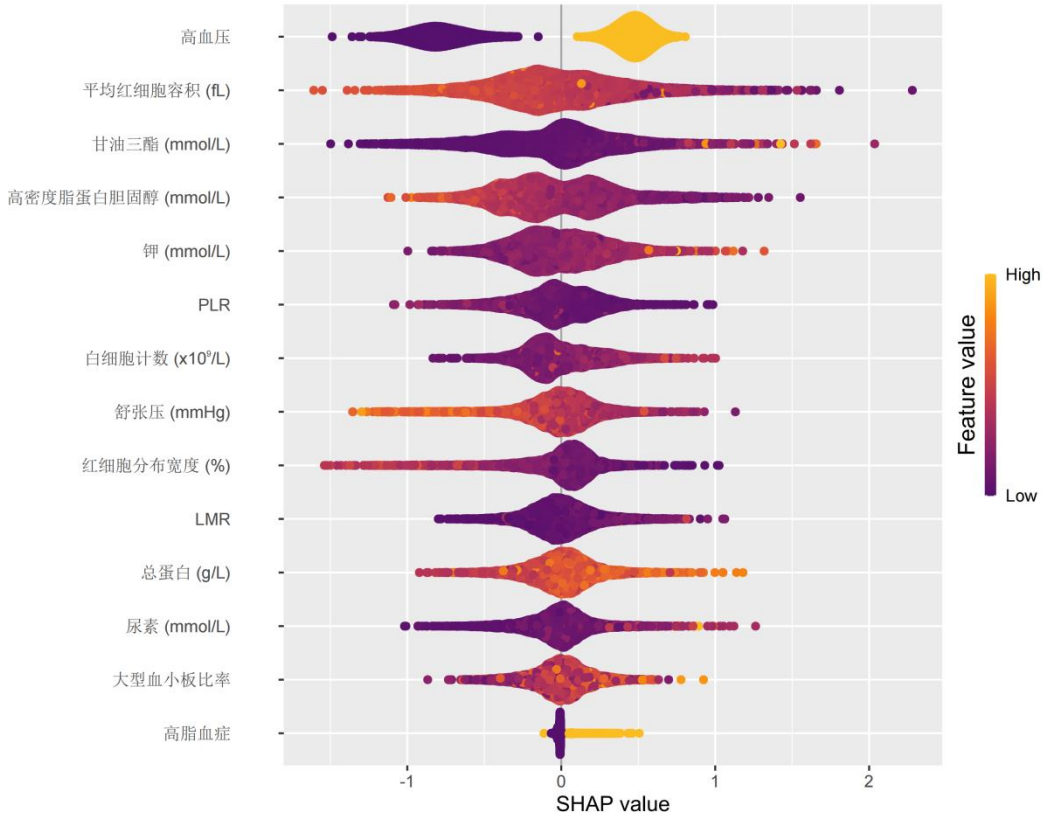


图 4 基于 SHAP 算法得到的特征重要性排名

图 4 中每一个圆点都代表着一个患者的某个特征值对应的 SHAP value，圆点的横坐标由 SHAP value 决定；纵坐标由该点对应的特征类别决定。SHAP value 的绝对值越大表示对应患者的特征值对预测老年慢性冠心病患者合并糖尿病风险的结果产生的影响越大，该值为正时表示对预测结果产生正向影响（增加糖尿病合并风险），反之为负向影响（降低糖尿病合并风险）。点的颜色表示该患者的风险因素的值的大小（点颜色越黄值越大；颜色越紫值越小）。因此，可根据图中各风险因素对应的点的分布和颜色判断风险因素对于老年慢性冠心病患者合并糖尿病风险的影响。例如，对于高血压，绝大部分黄色（表示该患者合并有高血压）样本点位于纵轴的右半侧，对应的 SHAP value 为正，即增加患者合并糖尿病的风险；紫色（表示该患者未合并有高血压）样本点大多位于纵轴左半侧，对应的 SHAP value 为负，即降低患者合并糖尿病的风险。因此可得出合并有高血压可能会增加老年慢性冠心病患者合并糖尿病的风险。

3.6 个性化风险因素分析

为了更好地将 XGBoost 模型应用于单个老年慢性冠心病患者，本文以建模组中心糖尿病的平均发生率作为基线概率，将患者风险因素对应的 SHAP 值映射为增加或降低合并糖尿病的概率。

图 5 展示了建模组中一个合并有糖尿病的老年慢性冠心病患者的个性化风险因素分析结果。患者合并糖尿病的基线风险为 0.273。患者全部风险因素的值和患者的基线风险共同决定了患者的预测结果。如分析结果所示，合并有高血压、平均红细胞容积=61.3fL、高密度脂蛋白胆固醇

=0.76mmol/L、舒张压=88mmHg、总蛋白=62.1g/L、LMR=7.818、大型血小板比率=32.64 以及尿素=9.38 mmol/L 分别增加了该患者 0.079、0.213、0.107、0.072、0.124、0.093、0.065 和 0.12 的合并糖尿病的风险；红细胞分布宽度=17.2%和钾=3.43mmol/L 分别降低了该患者 0.22 和 0.05 的合并糖尿病的风险。

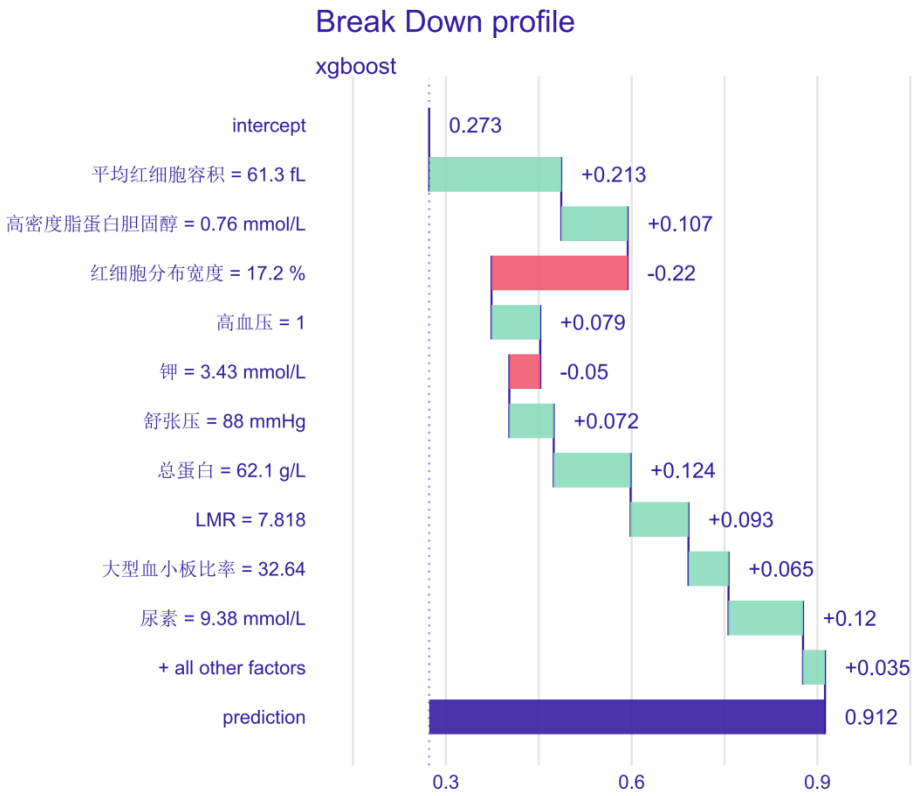


图 5 基于 SHAP 算法的个性化风险因素分析示例（合并糖尿病）

4 结语

本研究基于海量电子病历数据，结合 9 种主流机器学习算法（LR、RF、DT、XGBoost、SVM、ANN、KNN、NB 和 lightGBM）探索了老年慢性冠心病患者合并糖尿病的影响因素，并基于影响因素构建了预测模型。结果显示高血压、高脂血症、舒张压、白细胞计数、LMR、PLR、钾、尿素、总蛋白、大型血小板比率、平均红细胞容积、红细胞分布宽度、甘油三酯和高密度脂蛋白胆固醇是老年慢性冠心病合并糖尿病的重要影响因素。基于以上影响因素分别构建了 9 种机器学习算法预测模型，结果显示 XGBoost 算法构建的模型效能最好。基于 SHAP 算法提出的个性化风险因素分析方法应用于建模组中合并有糖尿病的老年慢性冠心病患者，结果显示模型的预测结果与患者实际的临床结局一致，风险因素分析结果与临床经验相符。

OR-002

基于深度学习和情境信息的在线健康社区科普文章推荐算法研究减轻脑出血后损伤机制

钟玉骏 付延旭 姚艺倍 蔡林 邹男男 曹锦丹
吉林大学公共卫生学院

2016 年，《“健康中国 2030”规划纲要》提出将健康摆在优先发展的战略地位[1]。此政策体现了国家对于公民健康的重视，国家的重视继而带动了公民健康意识的不断提升。2018 年，国务院办公厅印发《关于促进“互联网+医疗健康”发展的意见》，提出要健全“互联网+医疗健康”服务体系，完善“互联网+医疗健康”支撑体系[2]。医疗健康搭乘互联网带来的便捷性推动了公民通过互联网获取健康信息的积极性。在 2023 年，中国互联网络信息中心发布的第 51 次《中国互联网络发展状况统计报告》显示截至 12 月，我国互联网医疗用户规模达 3.63 亿[3]。这体现出互联网平台成为用户获取健康信息的重要平台。在线健康社区又是互联网医疗的重要组成部分，这使得其成为用户获取健康信息的重要渠道，也是反映用户健康信息需求，实现用户健康科普的重要平台。随着在线健康社区用户数量的飞速式增长，社区内的话题帖不计其数，使得信息过载问题出现。信息过载影响用户的个体决策和个体主观状态[4]。为了在一定程度上解决这个问题，可将个性化推荐算法应用于在线健康社区，根据用户的兴趣向用户推荐符合其需求的健康信息。这一方法在《“十四五”全民健康信息化规划》[5]中也有相应的体现。但目前面向在线健康社区的个性化推荐算法研究，多集中于传统的推荐算法，在数据稀疏、数据利用不充分、推荐准确度优化等问题上仍可进一步缓解和完善。而基于深度学习的个性化推荐算法有非线性转化、深层特征学习和高弹性以及可用性等优势[6]。这使得深度学习可以较方便高效地利用用户和健康文章信息，通过挖掘数据的潜在特性，获取用户和健康文章的特征描述，进而提升推荐算法的整体性能[7-9]。这有助于构建的推荐算法学习到用户和健康信息之间更加复杂的特征关联信息，在缓解数据稀疏问题的基础上进而提高推荐的准确率。并且在在线健康环境下，在个性化推荐算法中融合情境信息可更好的获取用户提问中未涉及的健康信息，作为用户偏好和行为的补充和完善，以便为用户形成更好的推荐结果。因此，针对在线健康社区的个性化推荐算法服务，探索基于情境信息和深度学习模型的在线健康社区个性化推荐算法，对于更好地满足用户健康信息需求和实现用户健康科普具有重要意义。

1 国内外相关研究

从个性化推荐应用领域来说，研究多集中于旅游[10]、音乐[11]、电影[12]、新闻[13]、电子商务[14]等领域。在线健康社区不仅有在线社区具备的信息过载的共性问题，而且还有其特殊性：一方面用户的健康信息需求涉及疾病诊断、治疗方法、药物副作用、预防措施等各个方面，呈现多样化和复杂化；另一方面健康信息具备高度的专业性，导致用户在阅读健康信息时很难判断是否与自身健康信息需求相符。针对以上问题，学者们开始将个性化推荐算法应用于在线健康社区。与其他领域研究相较而言，在线健康领域中个性化推荐算法的现有研究较少，且多集中于传统推荐算法、融合 LDA 主题模型与传统推荐算法所构建的个性化推荐算法。但传统的推荐算法未充分挖掘和使用用户和物品信息，造成推荐结果不精准的问题。

在情境信息应用研究中，学者们[15-17]提出了不同应用环境下的情境感知推荐框架，通过在信息推荐函数和传统推荐算法的基础上融入情境信息，以弥补和扩充用户模型，提高个性化推荐算法的推荐准确度。

深度学习（Deep learning）是一种机器学习方法，它基于人工神经网络进行模型构建和学习，通过多层次的神经元对数据进行分层抽象和特征提取，从而实现对复杂非线性模式的学习和识别。然而，现有研究对于融合深度学习技术与情境建模进行健康信息推荐的研究内容较少。但通过其他领域基于深度学习模型的研究发现，基于深度学习模型构建个性化推荐算法一方面可通过构建多层神经网络，有效的学习输入数据的深层特征；另一方面神经网络强大的非线性转换能力，有助于模型学习到用户和健康信息之间更加复杂的特征关联信息；另外，深度学习模型可有效地融入各类粗糙的情境数据，进而更好地获取情境信息的隐式表示。这些都可有效提高个性化推荐算法的准确率。

综上，在线健康环境下，探索基于情境信息和深度学习的个性化推荐算法，可更加精准地分析用户的行为和偏好，有效提高个性化推荐算法的准确度，以便为用户精准的推荐健康信息，进而促进用户科学决策和营造良好的主观状态。故本研究为了进一步满足用户多样化、复杂化和专业化的健康信息需求，帮助用户更有效地获取所需的健康信息，构建了基于深度学习模型和情境信息的个性化推荐算法。

2 基于深度学习和情境信息的在线健康社区个性化信息推荐算法研究

基于深度学习模型的个性化推荐算法构建步骤为：①数据处理，将用户数据和健康文章数据处理为神经网络可理解的形式，具体表现为利用词嵌入技术分别将用户数据和健康文章数据映射为向量；②模型设计，设计神经网络模型，将离散的文字数据映射为向量，具体表现为将用户向量和健康文章向量分别输入深度神经网络和卷积神经网络，得到用户和电影的向量表示，提取用户特征和健康文章特征，并计算二者的余弦相似度；③配置训练参数并完成训练，提取并保存训练后的数据特征，具体表现为用该相似度和用户真实评分的均方差作为该回归模型的损失函数，并展开模型训练和优化。

2.1 数据处理

2.1.1 基于深度学习模型的科普文章向量表示

本研究通过使用 Word2vec 模型将科普文章中的词转变为 512 维的词向量，使用医学词典分词后，文章的标题则指示文章的主要特征，将标题中的向量求平均，得到标题的总体向量表示。

但由于健康文章的分词结果出现长短不一的情况，需要对文章内容进行填充和截断，选取出现频率前五十的词作为文章内容的特征，其中不足 50 词的文章以 0 向量填充。则健康文章表示为一个特征向量和五十个 512 维的内容向量。

本研究将内容向量作为输入，使用特征向量作为标签，对卷积神经网络进行训练，使得模型能够从文章内容中提取健康信息的特征表示。

2.1.2 用户数据向量表示

将用户数据分词后，患者病例概要则指示患者想要得到的信息，即用户兴趣。将用户和患者的对话作为情境信息，将医生建议，用户想要得到的帮助和疾病描述作为训练的输入数据。同时将医生回答信息和患者回答信息分别截断为出现频率最高的 20 词和 10 词，不足的以 0 向量填充。

2.2 基于深度学习模型和情境信息的个性化推荐算法设计

本研究对于基于深度学习模型和情境信息的个性化推荐算法设计，主要划分为以下几部分。

首先，在线健康社区的健康科普文章数据方面，通过卷积神经网络进行健康科普文章的特征提取，其包含一个用来接收 50*512 维的输入的输入层；三个卷积层，其每个卷积层由两个 Conv1D 层、归一化层、一个最大池化层构成；最后连接一个全连接层和一个输出层。

其次，在用户数据方面，由于在线健康社区用户的健康信息需求复杂多样，单纯依靠分析用户的提问，无法准确的获取用户的健康状况，进而导致用户需求和兴趣定位不准确。而将用户和医生之间的对话作为情景信息纳入研究范围，可完善和补充用户特征，更加精准的定位用户兴趣和需求。故本研究中的用户数据由用户对疾病的描述、医生的建议和医患问答数据等构成，其中医患问答数据则作为患者健康需求的情境信息。接下来，将用户数据输入深度神经网络进行用户特征的提取，深度神经网络首先通过展层（Flatten）将每个输入展平为一维向量，再通过连接层（Concatenate）将多个输入进行连接，隐藏层由 10 个全连接层构成。

最后，通过用户特征向量和文章特征向量计算二者之间的余弦相似度矩阵，对于每个用户，选取相似度最高的 N 篇文章进行推荐（图 1）。

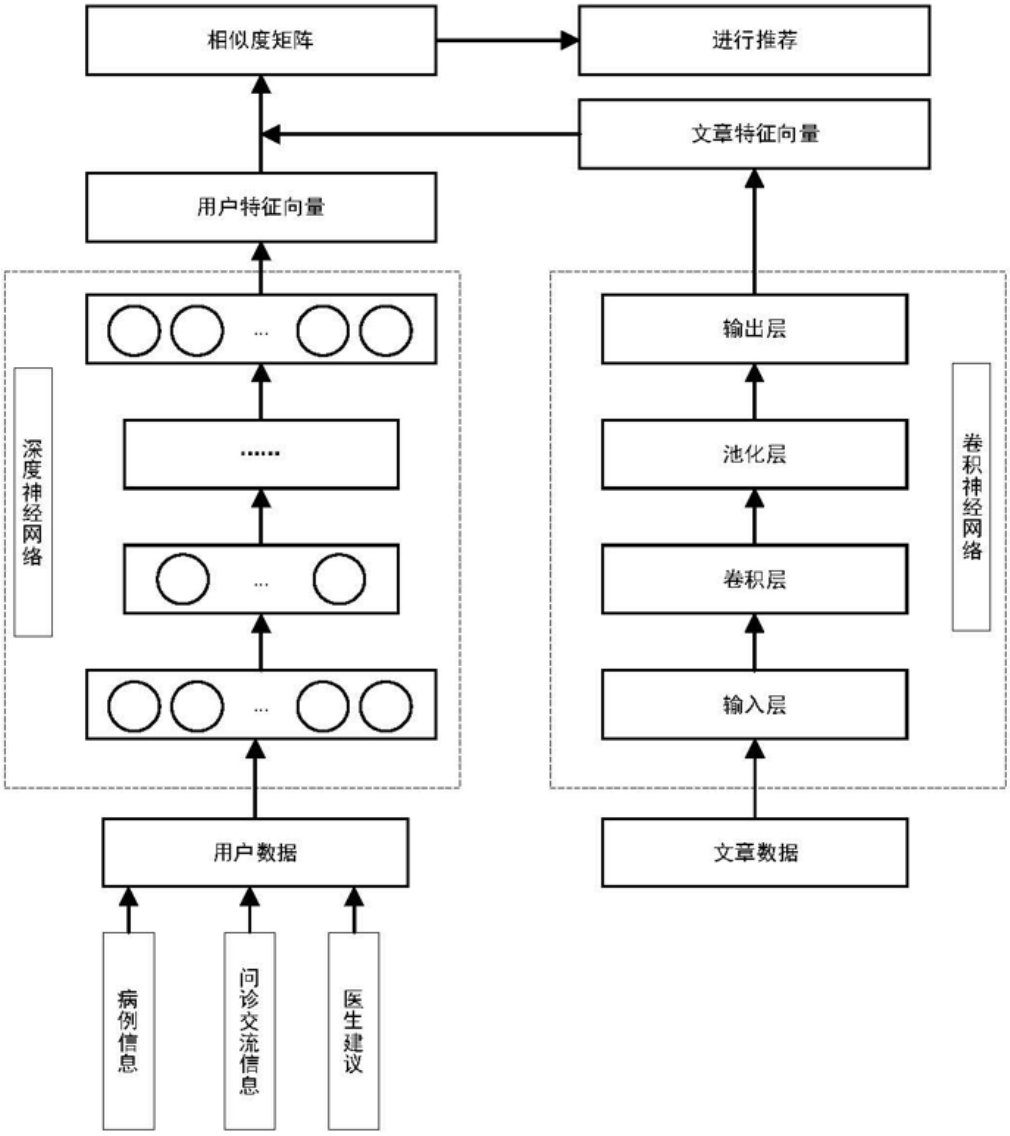


图 1 基于深度学习模型和情境信息的个性化推荐模型设计

2.3 模型训练与优化

对于健康科普文章数据：设置整个模型的损失函数为均方误差（Mean squared error），优化器为 Adam，学习率（Learning rate）为 0.001，评估指标包括均方误差、平均绝对误差（Mean absolute error）和 R 平方值（R squared）。在训练模型时，使用 GPU 加速，设置批大小(batch size)为 512，训练轮数(epochs)为 400。最后，使用训练好的模型对输入的健康文章向量进行特征提取，并将提取的特征用于预测输出值。训练集选用健康文章内容，标签使用从健康文章内容中提取的特征向量。训练结束时，评价指标分别为 mse:

0.0856、 mae: 0.2314、r_squared: 0.5492。

对于用户数据，模型的损失函数为均方误差，优化器为 Adam，学习率为 0.0001，评估指标包括均方误差、平均绝对误差和 R 平方值。在训练模型时，使用 GPU 加速，批大小为 200，训练轮数为 800。训练结束时，评价指标分别为 mse: 0.1251 - mae: 0.2631 - r_squared: 0.6276。

3 基于深度学习模型和情境信息的在线健康社区个性化信息推荐算法实证研究

3.1 数据来源

好大夫在线、丁香医生、39 健康网等均是我国规模较大、使用较为广泛的在线健康社区平台。其中“好大夫在线（www.haodf.com）”拥有 10090 家医院的 905486 位大夫，并成功就诊患者 4022267 位，一方面更能保证用户的访问量，另一方面其医生的专业性更强。

故本研究选取了“好大夫在线”在线健康社区作为实验数据来源，采用 python 爬虫技术分别从“专家文章版块”爬取专家文章 15070 篇，从“在线问诊版块”爬取用户问诊信息 3000 条。

其中，专家文章包括标题和内容两部分。在本研究中，考虑了用来补充和扩展用户兴趣和需求建模的情境信息，将代表用户信息的问诊信息分为三个部分：患者填写的病例信息、患者与医生的交流问诊信息、医生给出的问诊建议。

3.2 数据预处理

本研究使用 python 的 jieba 分词工具，添加清华大学开放中文词库[18]中的医学词库作为词典，并且整合哈工大停用词表、四川大学停用词表、百度停用词表、中文停用词表对文章信息和用户数据进行分词。同时将数据存为 json 格式，以便后期实验的调用。

在本研究中，采用了 word2vec 模型对词向量进行嵌入处理，使用清华大学开放中文词库[18]中的医学词库对 CMCQA 中国医学会话问答数据集[19]进行分词，作为训练集使用 CBOW 模型对 word2vec 模型进行训练。

3.2.1 健康科普文章信息的预处理 文章信息主要包括标题（title）字段和内容（content）字段，分词后其内容为包含单词的列表（表 1），对其词频进行统计，其中内容选取词频排名前 50 的单词作为其表示。

表 1 分词后的部分文章信息

Index	Title	Content
7998707163	["冠心病", "要放", "支架"]	["冠心病", "冠状动脉", "供血", "心肌", "耗氧",……]
7998733005	["腰椎间盘", "危害"]	["腰椎间盘", "危害", "腰", "腰椎间盘", "累计", "坐骨神经",……]
7998762540	["情绪", "激动", "引发", "冠心病", "患者", "猝死"]	["垂体后叶素", "冠状动脉痉挛", "情绪", "激动", "时", "肾上腺", "释放",……]
……	……	……

通过word2vec模型将其转换为512维的词向量，其中文章内容词向量个数不足50的，以 0 向量填充。一个词语的向量表示被初始化为 0 向量表示这个词语在模型的训练过程中没有得到充分的学习和表达。这时，这个词语的向量不具有意义的语义信息，也不会对模型的性能产生有效的影响。经过处理的数据如图 2 所示。

	seg_title	seg_content
7998707163	[[1.4200243, -0.60254604, -0.030644275, -0.727...	[[1.4200243, -0.60254604, -0.030644275, -0.727...
7998733005	[[1.1814094, -0.1442893, -1.9671038, 0.6222889...	[[1.1814094, -0.1442893, -1.9671038, 0.6222889...
7998762540	[[0.14521427, 0.82865787, 1.2940876, 1.0615965...	[[0.30485484, -1.9542263, 2.1369848, 2.2025783...
7998768163	[[-0.19918038, -1.4747806, -1.3657712, 1.62028...	[[-0.09310274, 0.17545569, -1.1545945, -0.0589...
7998788453	[[0.9422169, -1.0848218, -1.3294256, -0.030442...	[[0.75591063, 0.13235383, -0.14606842, 0.55396...
...	...	...
9392065669	[[0.7349509, 1.4016879, -2.4082692, 0.01085776...	[[0.67510325, -0.9958852, -0.07067924, 0.79031...
9392065682	[[0.61592805, 0.10249819, -1.2233915, -0.96503...	[[-0.21497156, 0.29963806, 0.6670732, -0.03962...
9392068218	[[0.7349509, 1.4016879, -2.4082692, 0.01085776...	[[0.67510325, -0.9958852, -0.07067924, 0.79031...
9392068223	[[0.7349509, 1.4016879, -2.4082692, 0.01085776...	[[-0.24222851, 1.1934345, 0.6740354, -0.579781...
9392068363	[[0.7349509, 1.4016879, -2.4082692, 0.01085776...	[[0.75591063, 0.13235383, -0.14606842, 0.55396...

15070 rows × 2 columns

图 2 文章的向量化表示

3.2.2 用户信息的预处理 用户信息包括三部分内容，分别为病例信息，问诊交流信息和医生建议。其中，病例信息包括病例概要、疾病描述、身高体重、患病时长、过敏史、既往病史、希望得到的帮助等；医生建议信息为处置建议；问诊交流信息由病人问诊列表和医生问诊列表构成。对用户信息的预处理包含以下三项：

一是由于问诊交流信息可能存在空白，为了避免此种情况的干扰，本研究选取了医生问诊信息大于 5 的数据进行建模，获得用户信息共计 839 条；

二是由于患者并未填充所有的病历信息，故本研究在其部分属性的取值为空。查看患者数据的分布中取值为空的数据所占比例如下图（图 3）所示；

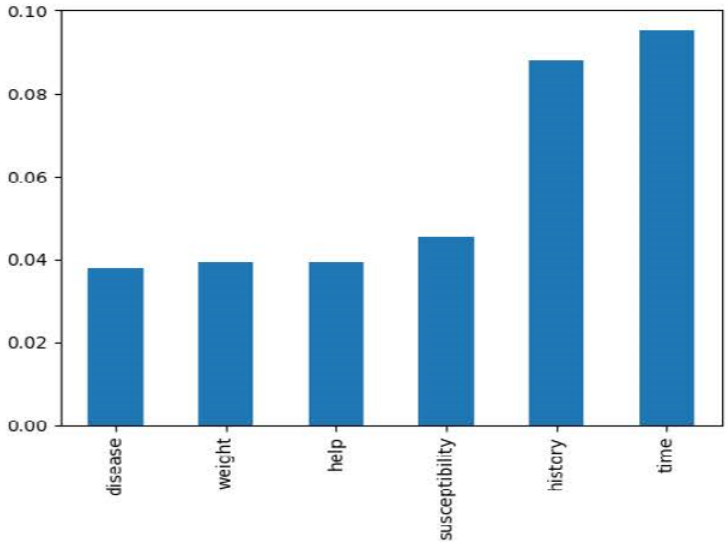
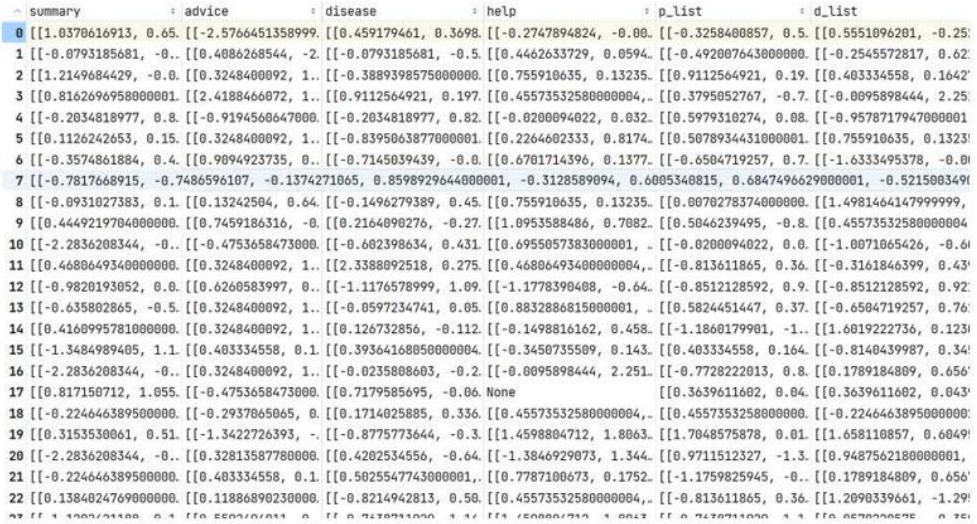


图 3 空值的分布

三是由于模型未涉及时空概念，所以本研究中的数据去除了过敏史、既往病史以及患病时长等属性；

四是筛选的属性列表为病例概要疾病描述、希望得到的帮助、处置建议、病人的问诊列表以及医生问诊列表。将病人问诊列表及医生问诊列表作为情境信息输入模型。

通过分词与词向量的嵌入操作，用户数据的向量化表示见图 4。



	summary	advice	disease	help	p_list	d_list
0	[[1.0378616913, 0.65.	[[-2.5766451358999.	[[0.459179461, 0.3698.	[[-0.2747894824, -0.00.	[[-0.3258408857, 0.5.	[[0.5551096201, -0.25
1	[[-0.0793185681, -0..	[[0.4086268544, -2.	[[-0.0793185681, -0.5.	[[0.4462633729, 0.0594…	[[-0.492007643000000.	[[-0.2545572817, 0.62
2	[[1.2149684429, -0.0.	[[0.3248400092, 1..	[[-0.3889398575000000.	[[0.755910635, 0.13235.	[[0.9112564921, 0.19.	[[0.403334558, 0.1642
3	[[0.8162696958000001.	[[2.4188466072, 1..	[[0.9112564921, 0.197.	[[0.45573532580000004,..	[[0.3795052767, -0.7.	[[-0.0895898444, 2.25
4	[[-0.2034818977, 0.8.	[[-0.9194560647000.	[[-0.2034818977, 0.82.	[[-0.0200094022, 0.032.	[[0.5979310274, 0.08.	[[-0.9578717947000001
5	[[0.1126242653, 0.15.	[[0.3248400092, 1..	[[-0.8395063877000001.	[[0.2264602333, 0.8174…	[[0.5078934431000001.	[[0.755910635, 0.1323
6	[[-0.3574861884, 0.4.	[[0.9094923735, 0..	[[-0.7145039439, -0.0.	[[0.6701714396, 0.1377.	[[-0.6504719257, 0.7.	[[-1.6333495378, -0.0
7	[[-0.7817668915, -0.7486596107, -0.1374271065, 0.8598929644000001, -0.3128589094, 0.6005348815, 0.6847496629000001, -0.521500349(					
8	[[-0.0931027383, 0.1.	[[0.13242504, 0.64.	[[-0.1496279389, 0.45.	[[0.755910635, 0.13235.	[[0.0070278374000000.	[[1.4981464147999999,
9	[[0.4449219704000000.	[[0.7459186316, -0.	[[0.2164090276, -0.27.	[[1.0953588486, 0.7082…	[[0.5046239495, -0.8.	[[0.45573532580000004
10	[[-2.2836208344, -0..	[[-0.4753658473000.	[[-0.602398634, 0.431.	[[0.6955057383000001, ..	[[-0.0200094022, 0.0.	[[-1.0071065426, -0.6
11	[[0.4680649340000000.	[[0.3248400092, 1..	[[2.3388092518, 0.275.	[[0.46806493400000004,..	[[-0.813611865, 0.36.	[[-0.3161846399, 0.43
12	[[-0.9820193052, 0.0.	[[0.6260583997, 0..	[[-1.1176578999, 1.09.	[[-1.1778390408, -0.64…	[[-0.8512128592, 0.9.	[[-0.8512128592, 0.92
13	[[-0.635802865, -0.5.	[[0.3248400092, 1..	[[-0.0597234741, 0.05.	[[0.8832886815000001, ..	[[0.5824451447, 0.37.	[[-0.6504719257, 0.76
14	[[0.4160995781000000.	[[0.3248400092, 1..	[[0.126732856, -0.112.	[[-0.1498816162, 0.458…	[[-1.1860179901, -1..	[[1.6019222736, 0.123
15	[[-1.3484989405, 1.1.	[[0.403334558, 0.1.	[[0.39364168050000004.	[[-0.3450735509, 0.143…	[[0.403334558, 0.164.	[[-0.8140439987, 0.34
16	[[-2.2836208344, -0..	[[0.3248400092, 1..	[[-0.0235808603, -0.2.	[[-0.0095898444, 2.251…	[[-0.7728222013, 0.8.	[[0.1789184809, 0.656
17	[[0.817150712, 1.055.	[[-0.4753658473000.	[[0.7179585695, -0.06.	None	[[0.3639611602, 0.04.	[[0.3639611602, 0.043
18	[[-0.224646389500000.	[[-0.2937865065, 0.	[[0.1714025885, 0.336.	[[0.4557353258000000 4,..	[[0.4557353258000000.	[[-0.2246463895000000
19	[[0.3153530061, 0.51.	[[-1.3422726393, -.	[[-0.8775773644, -0.3.	[[1.4598804712, 1.8063…	[[1.7048575878, 0.01.	[[1.658110857, 0.6049
20	[[-2.2836208344, -0..	[[0.32813587780000.	[[0.4202534556, -0.64.	[[-1.3846929073, 1.344…	[[0.9711512327, -1.3.	[[0.948756218000000 1,
21	[[-0.224646389500000.	[[0.403334558, 0.1.	[[0.5025547743000001,.	[[0.7787100673, 0.1752…	[[-1.1759825945, -0..	[[0.1789184809, 0.656
22	[[0.1384024769000000.	[[0.11886890230000.	[[-0.8214942813, 0.50.	[[0.45573532580000004,..	[[-0.813611865, 0.36.	[[1.2090339661, -1.29

图 4 用户数据的向量化表示

3.3 实验环境

本研究设计的推荐算法基于深度学习模型，使用了谷歌公司提供的人工智能学习平台 TensorFlow。TensorFlow 是一个功能强大、灵活性高并且具有广泛应用的机器学习框架。它提供了丰富的功能和工具，使得用户可以方便地构建、训练和部署各种机器学习模型。具体实验环境如下表 2 所示。

表 2 实验环境

实验环境	配置
操作系统	Ubuntu 20.04.5 LTS
GPU	NVIDIA Tesla P100
内存	13GB
硬盘	73GB
开发环境	Python 3.7.12

3.4 评价指标

本文采用余弦相似度和推荐结果评分作为推荐模型的评价指标。

3.4.1 相似度度量

由于神经网络输出的均是 512 维的用户及文章的特征向量表示，因此可用相似度来度量，主要有余弦相似度（Cosine Similarity）和欧几里得距离（Euclidean Distance）。

其中余弦相似度计算公式如 3-1：

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \times \sqrt{\sum_{i=1}^{n}(B_i)^2}} \quad (3\text{-}1)$$

通过召回不同数量的文章，计算其平均余弦相似度，来评价模型的性能。

3.4.2 推荐结果评分

对于提取的用户特征向量和文章特征向量，用户特征向量和文章特征向量所对应的十个单词的相同个数即为推荐结果的评分。

具体计算方法如公式 3-2 所示：

$$score = \frac{\sum_{i=1}^{N} card(C_i)}{N} \quad (3\text{-}2)$$

其中 N 为召回的文章数，集合 C 定义为$C_i = \{c | c \in A,\ c \in B_i\}$，其中集合 A 为用户向量的相似单词，集合 B 为文章向量的相似单词。card(C)定义为集合 A 与集合 B 所拥有的相同单词的数量。

3.5 实验方法

由于目前基于在线健康社区的个性化信息推荐算法多使用 TF-IDF 统计方法与 LDA 主题模型结合进行特征的提取，从而进行推荐。因此本文将所搭建的基于深度学习和情境信息的推荐模型与使用 TF-IDF 统计方法和 LDA 主题模型的个性化推荐算法进行特征提取的推荐模型进行比较，通过推荐文章与用户的语义相似度以及推荐文章的准确率对模型进行评估。

3.6 实验结果

实验结果如下表 3 和表 4 所示，分别召回 5，10，15 和 20 篇文章，计算召回文章与用户之间的余弦相似度和推荐结果评分。

表 3 用户间余弦相似度

模型	Top5	Top10	Top15	Top20	平均值
融合 TF-IDF 和 LDA 主题模型的个性化推荐算法	0.3389	0.3141	0.2973	0.2857	0.3090
基于深度学习和情境信息的个性化推荐算法	0.5443	0.5221	0.5075	0.4964	0.5176
相对提升比例	60.61%	66.22%	70.70%	73.75%	67.51%

由表 3 可知，基于深度学习模型和情境信息的个性化推荐算法相较于融合 TF-IDF 和 LDA 主题模型的个性化推荐算法，提取的文章特征与用户特征之间具有更高的余弦相似度。基于 Word2vec 模型所嵌入的向量，越高的余弦相似度意味着更高的语义相似度。对于更多文章的召回，基于深度学习和情境信息的个性化推荐算法也有着相对的稳定性。

表 4 推荐结果评分

模型	Top5	Top10	Top15	Top20	平均值
融合TF-IDF和LDA 主题模型的个性化推荐算法	1.3089	1.0792	0.9261	0.8277	1.0355
基于深度学习和情境信息的个性化推荐算法	1.8913	1.6843	1.5356	1.4312	1.6356
相对提升比例	44.50%	56.07%	65.81%	72.91%	57.95%

由表 4 可知，基于深度学习模型和情境信息的个性化推荐算法所提取的特征向量有着更高的推荐结果评分，即对于每个推荐结果有更多的单词与用户相同。相比之下，在在线健康社区的应用情境下，基于深度学习和情境信息的个性化推荐算法相对于融合 TF-IDF 和 LDA 主题模型的个性化推荐算法均有一定的性能提升，且在更多文章的召回上提升明显。

4 结语

本研究构建的在线健康社区应用场景下，基于深度学习模型和情境信息的个性化健康信息推荐算法在文章召回和推荐结果评分方面都优于其对比模型，并且在数量较大的文章召回情况下表现更加稳定，使得在线健康社区的个性化推荐更加准确，为之后的在线健康社区个性化推荐研究提供了参考。但本研究忽略了用户兴趣随着季节变换和病情的发展产生漂移的问题，在后续研究中可考虑时间因素对用户健康兴趣的影响。

OR-003

“互联网+医疗健康”的发展与应用研究

张钰珠[①] 刘新奎[②]
①郑州大学公共卫生学院
②郑州大学第一附属医院

“互联网+医疗健康”模式是指以互联网为基础、以现代化信息技术为依托，与传统医疗健康相结合，形成的一种具有现代化特征的新型医疗健康管理模式。党中央、国务院高度重视“互联网+医疗健康”工作。习近平总书记曾经指出，要推进“互联网+医疗”等，让百姓少跑腿，让数据多跑路，不断提升公共服务均等化、普惠化、便捷化水平。我国对医疗健康管理始终高度重视[1]，当前正值“十四五”期间，“互联网+医疗健康”模式的创新，为我国医疗健康事业的发展提供了技术保障。

1“互联网+医疗健康”的发展优势

1.1 国家政策的支持

2018 年 4 月，国务院办公厅印发了《关于促进“互联网+医疗健康”发展的意见》；2018 年 7 月，国家卫生健康委员会、国家中医药管理局联合发布了《关于深入开展“互联网+医疗健康”便民惠民活动的通知》；2018 年 9 月，国家卫生健康委员会国家中医药管理局联合发布了《互联网医院管理办法（试行）》《互联网诊疗管理办法（试行）》《远程医疗服务管理规范（试行）》等 3 个文件：2020 年 12 月，国家卫生健康委员会、国家医疗保障局、国家中医药管理局研究制定了《关于深入推进“互联网+医疗健康”“五个一”服务行动的通知》：2022 年 11 月，国家卫生健康委员会、国家中医药管理局、国家疾控局联合发布了《“十四五”全民健康信息化规划》；2022 年 12 月，中共中央、国务院印发了《扩大内需战略规划纲要（2022-2035 年）》等，这些文件都从不同方面提出了加快互联网与医疗健康的融合，进一步深化了“互联网+医疗健康”服务体系。

1.2 先进技术的支撑

5G 技术、大数据、云计算、人工智能、传感器等的发展，成为“互联网+医疗健康”的有效辅助和有力支撑，患者可以实现网络预约、异地就诊医疗保险信息即时结算；大数据可以完成患者电子医疗档案的建立、信息收集与存储、数据挖掘与分析医生可以通过互联网完成大量患者医疗数据的分析消除与患者的时间损耗与距离障碍；在同级医疗机构的检查结果互认的情况下，医疗系统和平台的信息传递与共享，将进一步节省医疗资源，减轻患者负担。

2 国内研究现状

互联网已经开始与医疗健康管理相互融合，我国能够从宏观角度颁布政策，对这一融合予以支持。互联网时代的到来，切实缓解了我国医疗卫生资源分布不均衡的问题，实现了医疗资源的流动，在解决医患矛盾方面，也起到了积极的促进作用。目前国内针对“互联网+医疗健康”模式的研究，主要从政策、技术以及应用模式三个方面展开。

曾理、杨雷等[2]研究指出“互联网+”这一概念在 2015 年出现在我国政府报告中，在此之后，国家针对“互联网+”开展了相关行动，针对“互联网+医疗模式”开展了 11 项行动，对该模式进行推广，随后国家针对这一模式颁布了多项政策，如《关于促进和规范健康医疗大数据应用发展的指导意见》《关于促进移动互联网健康有序发展的意见》等，在指导意见中明确提出要大力推动该模式的发展，为医疗健康管理服务探索出新的模式，促使医疗行业呈现出新业态。李琴兰[3]研究认为，在医疗政策方面，我国对“互联网+”这一模式十分支持，比如《医师执业注册管理办法》，其中对执业注册进行规定，从 2016 年起，开始实施电子证照，电子证照的出现，实现了对医师的动态化统一管理，管理的效率也有了切实的提高。我国各个地方也能够积极响应国家的号召，依据国家政策结合地方的发展现状，制定符合地域特色的“互联网+医疗健康”模式。

在技术方面，王会笑[4]研究表明，技术是发展与创新不竭的动力，“互联网+医疗健康”模式属于信息化建设的一种，其依托于技术发展。近几年，我国在信息安全技术、卫生信息技术、医

疗机构信息化等方面均取得了良好的成绩，这为新模式的发展奠定了技术基础。现代化信息技术，如云计算、大数据、人工智能等，这些技术与医疗卫生行业相互结合，为居民构建医疗健康档案，能够实现对患者健康情况的跟踪。王政、王萍[5]等认为移动互联网技术与医疗事业的结合，将医疗服务从线下转移到线上，患者就医以及医护人员管理均变得更加便捷。

在应用模式方面，孙越、游茂等[6]研究指出，“互联网+医疗健康”模式是医疗工作在模式方面的创新，通过云平台，患者足不出户就能够完成健康方面咨询和选择适合的医生等。在分级诊疗管理制度方面，也可以采用这一模式，逐步实践基层首诊、急慢诊分收、上下联动的医疗方案，促使医疗资源逐渐均衡。

3“互联网+医疗健康”面临的问题

3.1 医疗卫生行业的数据标准化问题

医疗卫生行业分工精细，每个医疗卫生机构内部分科细致，各个分科相近却有着明显的差异，实行标准化、规范化管理难度很大。“数据烟囱”“数据孤岛”一直是不同区域医疗机构之间实现信息互通共享的难题，数据标准、编码、医疗术语如果无法统一，信息无法互认，则难以实现数据间的交互和共享，这必将造成不同业务系统间的数据无法有效整合。

3.2 数据安全问题

随着“互联网+医疗健康”的推进，人们使用互联网问诊咨询越来越普遍，尤其在疫情期间，一些慢性病患者更倾向于在线问诊、在线开药快递到家这种方式。由于问诊、咨询等大多需要实名认证、上传病例等相关信息，这样就存在个人信息泄露、医疗数据外流等安全风险和隐患。同时随着智能可穿戴设备的使用，人们在实时监测身体健康的同时，也容易造成信息被不当收集、过度利用。可见数字医疗缺乏统一有效的标准，监管力度不足，威胁着医疗数据的安全。

3.3 缺乏完善的准入和监管标准

目前，“互联网+医疗健康”的准入标准、从业规范标准和相关的监督管理标准并不完善，对于开设互联网服务的医院资质、开展互联网医疗服务的平台，以及加入互联网平台开展诊疗咨询服务的医师的资格等标准的制定不完善，同时对于互联网平台与线下医院合作开展的“互联网+医疗健康工作，存在责任划分、收费标准等不明确的问题。

3.4 医疗机构参与互联网诊疗能力意愿和必要性问题

互联网医院必须依托实体医院。大医院信息化程度高，有大量的人员、设备及资金支持，具备自我开发互联网诊疗亚台的能力，与大医院相比基层医疗机构在医疗机构建设和参与互联网医疗方面，普遍缺乏能力和资源，很难在现阶段建立起互联网医院体系。由于有些大医院患者很多，医务人员的时间和精力有限，不愿再去开展互联网诊疗但是患者对大医院专家的需求很大，尤其是外地患者，希望能够通过互联网诊疗得到权威医师的诊治。

4“互联网+医疗健康”的发展建议

4.1 加强信息技术建设，促进模式融合

“互联网+”具有智能连接、智能体验等多种功能可以改善我国医疗资源不均衡、医疗机构之间信息不共享等问题。在该模式中，需要加强信息技术建设，促进“互联网+”与医疗健康管理之间的融合[7]。当前云计算、大数据、人工智能等技术的应用，能够提高医疗数据的管理水平。应该注意的是，在该模式中仍然需要以医疗健康为重心，基于患者需求开展相关的任务，通过这一新型医疗服务模式创造出更多的价值。

4.2 建立标准规范体系，加强行业监管

医疗产业是一个复杂的产业，在医疗产业链中，涉及了支付方、服务方、院方等多个主体，不同主体内部还设有更加具体的产业链。以服务方为例，内部还有多个层级的服务主体；在支付方中，也会具体涉及医疗保险与商业保险等等。因此，国家以及地方卫生部门应该建立标准的规范体系，实现对各个主体的统一管理，在“互联网+医疗健康管理”模式中，各个主体都能够受到监督与管理,以此促进行业规范发展，互联网与医疗事业也能够实现共赢[8]。

4.3 健全信息安全建设，保障模式安全

在信息安全建设方面，国家以及地方医院应该制定专项的政策与法律，明确医疗方面的隐私信息范围，加强对数据采集方、数据持有方等管理，对信息泄露事件予以处罚，以此对不法分子加以约束[9]。除此之外，在技术方面，还需要不断优化，采用最新的安全防护技术以此保障患者的信息安全。

4.4 完善配套政策，促使行业良好发展

完善“互联网+”模式的配套政策，有助于促使行业的良好发展。首先，应该建立健全的服务管理体系，制定“互联网+医疗健康”模式的管理制度，规范线上医疗健康管理的相关服务规范；其次，开展互联网医疗教育体系，使高校能够培养出互联网医疗相关的专业人才；第三，对执业注册以及电子证照进行统一的管理，在国内搭建资质统一管理平台，如医师、药师、护上等均需要通过系统的认证，以此提高医护人员流动过程中的管理水平[10]。

5 结论

“互联网+医疗健康”的兴起为传统医疗行业的发展与转型提供了新的方向，也深入挖掘了医疗健康管理行业的潜力，提高了医疗服务的便捷性，因此“互联网+医疗健康”是医疗事业发展的必然趋势。从当前“互联网+医疗健康”的发展现状来看，在运行机制等方面存在诸多的问题，这些问题均需要未来发展过程中不断地改善与解决，才能够促使医疗健康事业长效稳定地发展，为国民提供更加便捷、全面、高效的医疗健康管理服务。

OR-004

基于互联网构建线上线下一体化服务平台

刘文容[1] 段涛* 赵娅[2] 曾丽娟[2]
川北医学院附属医院

一、基本信息

（一）功能定位，建设目标

2018 年，国务院办公厅印发《关于促进 " 互联网 + 医疗健康 " 发展的意见》，提出允许医疗机构开展部分常见病、慢性病复诊等互联网医疗服务，为 " 互联网 + 医疗健康 " 明确了发展方向。川北医学院附属医院互联网医院主要功能定位为实体医院强有力的支撑。基于川北医学院附属医院互联网医院微信小程序， 以院内外融合一体化为建设目标，通过与院内信息系统(HIS\LIS\PACS）的对接，用户借助移动智能终端设备实现包括预约挂号、智能分诊导诊、在线复诊、线上医嘱、病历查询、报告查询、病案复印配送、药品配送、慢病管理、健康教育等在内的诊前、诊中、诊后一站式就医全流程，努力打通线上、线下一体化服务，搭建医患沟通新模式，提供便捷操作途径，为患者提供智慧服务，为管理者提供智慧管理，最终实现构建智慧医院的目标。

（二）发展历程

2019 年 积极探索互联网+医疗服务模式，布局建设互联网医院

2020 年 1 月 29 日 初步建成基于微信公众号的在线问诊平台，实现针对新型冠状病毒的免费咨询

2021 年 1 月 6 日 正式取得互联网医院牌照成为南充市首个获得互联网诊疗资质的三甲医院

2021 年 3 月 1 日 将线上问诊平台迁移至微信小程序，正式开始互联网诊疗活动

2022 年至今 不断完善门特专区、名医讲堂、中草药开具、心理卫生咨询、义诊专区等功能

（三）服务场景

我院互联网医院开展功能包含：线上问诊、复诊开方、心理卫生咨询、门特专区、新冠病毒咨询及新冠预防处方自助开药、用药咨询、检查预约、病历邮寄、药品配送、健康宣教、护理服务等；

1.线上问诊。提供图文问诊、电话问诊和视频问诊三种方式，患者在互联网医院输入自己的相关身份信息，建立档案，然后选择医生，根据医生的上线情况选择问诊方式（因医院管控，暂时关闭电话问诊和视频问诊功能）。图文问诊支持文本、图片、语音等方式，患者可将自己的检查结果、病历等图片发送给医生，患者挂号后，医生会收到挂号提醒， 利用空闲时间回复患者，解答患者疑问。医生可在线调阅患者在本院线上、线下的既往病历、检查检验结果及历史医嘱。

2.处方开具及配送。系统根据患者就诊记录判断是否为复诊患者，针对复诊患者，根据病情需要医生可以通过互联网医院为患者开立处方，也可直接调用历史处方。医生开具处方进行电子签名后，系统根据预先设置的合理用药规则进行前置审方或由药师干预审方，审方通过后将会推送消息提醒患者进行在线支付并选择药品寄送方式（到院自取或快递配送），可以将药品配送到家，免去了患者舟车劳顿到医院开药。我院互联网医院现在提供多种常见病慢性病药品，可以满足大多数患者的用药需求，还创立了中草药处方开立功能，可为患者开立草药、饮片和颗粒剂等中药，患者也可选择代煎，由代煎公司将中草药煎煮后直接邮寄给患者。

3.检查预约。医生可以为患者开立检查、检验申请，患者完成线上交费后，可以在微信公众号预约合适的检查时间，在所预约的时间凭电子健康卡直接到院内报到检查，检查完成后还能在互联网医院查询检查检验结果。形成线上线下一体化，院内院外全打通的模式。

4.心理卫生咨询。患者可在心理卫生咨询板块针对抑郁和焦虑进行免费自测并将自测结果发送给医生，也可通过电话咨询的方式由医生对患者心理进行干预指导，精心呵护患者心理健康。

5.用药咨询。针对患者面临的用药问题，上线药学门诊，设立用药咨询专区由专业药师提供综合性用药咨询和用药指导服务。患者完成线上处方支付后，即可直接点击按钮进入用药咨询药师

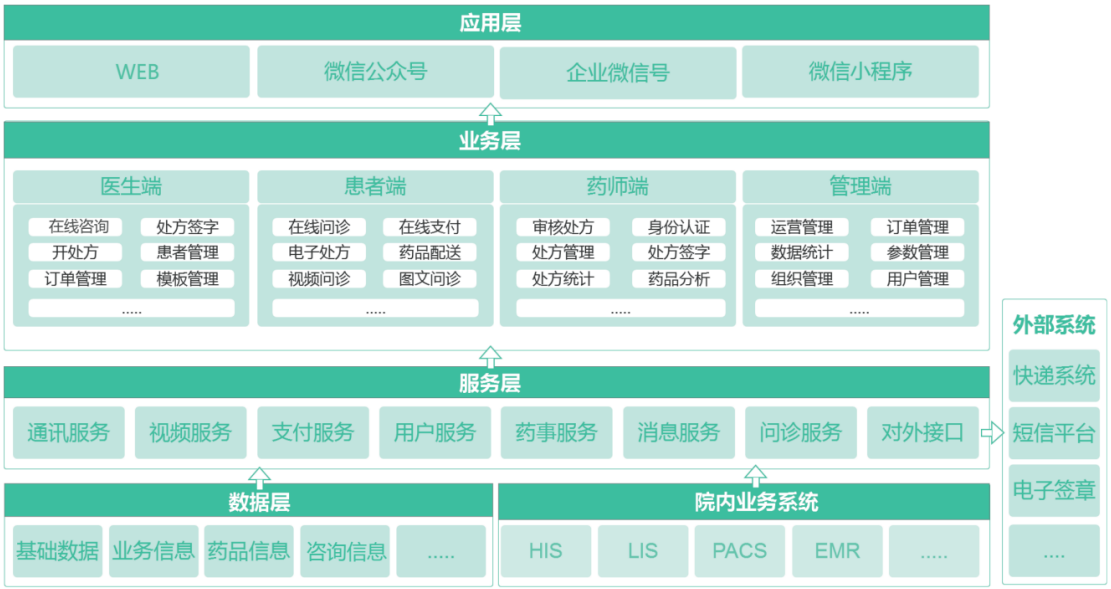
选择页。同时，为预防错误用药，根据患者处方精准化推送电子药品说明书及合理用药指导信息。

6.病历邮寄。对于有病历复印需求的患者，医院提供在线病历复印精准服务，患者通过互联网医院提交相关身份证明，医院对患者身份进行审核通过，按患者要求将复印好的病历为患者快递到家。

7.互联网 + 护理服务。是把优质护理服务延伸到社区到家庭的主要方式之一。“互联网 + 护理服务”开展以来，已经为多名不便来院做护理的病患送去了优质的护理服务。为保证院外护理服务人员安全，该院特配备安全记录仪，可以做到实时监控，全程留痕、sos 一键报警、GPS 实时定位等功能。

二、建设实践

（一）总体架构



（二）关键技术

高可用

1、系统支持 DNS 负载均衡，代理服务器(nginx,haproxy)负载均衡，并采用线程隔离(hystrix)，进程隔离(微服务)，读写隔离(redis，db)，动静隔离技术(静态资源和动态内容隔离)。

2、网络层限流(nginx),应用层限流(tomcat,undertow)，服务层限流(系统内部权限)

3、系统内的服务功能采用超时降级，故障降级，限流降级，超时重试处理机制。

高并发

1、采用分布式缓存(redis)，http 缓存(强缓存，协商缓存)

2、采用线程池，数据库连接池，数据库读写分离，多线程处理，异步处理，消息队列技术。

高扩展

1、对业务进行抽象建模，业务数据与业务逻辑解耦，系统各部件解耦，用于满足医院多元个性化需求。

数据安全

1、对于接口数据，请求和响应报文均采用多重加密算法进行传输，可以有效阻止第三方非法攻击和对数据信息进行篡改，同时可以对隐私敏感信息提供保护，降低了第三方恶意欺诈的机会。

2、对于数据库数据，一些重要数据采用加密存储方式，可以有效保护数据资源，使其免受网络攻击者的侵害，同时也可以对隐私敏感信息提供安全保护。

视频对话

1、使用 CDN 加速，保证医患之间无障碍沟通

2、使用混流录制功能，多端合流录制并保存医患沟通的内容，对于后续问题追溯提供便利条件。

3、保证视频文件存储时长满足国家要求标准。

电话沟通

1、使用号码隐私保护功能，保证医患电话号码的安全隐私，杜绝医患之间的侵扰。

三、效果评价

（一）运行效果

时间	线上接诊人次（人次）	参与线上接诊医生数（人）	参与线上审方药师数（人）	线上问诊人次占门诊总人次百分比
2020 年	10018	82	5	0.66%
2021年	17180	259	10	0.88%
2022年	25972	294	21	1.29%

（二）创新性及亮点

1.带 CA 的电子处方。慢病患者在线复诊后，医生根据患者当前病情为患者开具电子处方，医生开具的所有处方均进行电子签章，电子签章均进行 CA 认证。带 CA 的电子处方保证了信息的安全性、真实性、可靠性和完整性，不仅是对医生身份的真实性保证，也是一个权威的、公正的体现。

2.合理用药前置审方。利用合理用药系统结合人工审核，将原本后置的审方流程“前置”。医师开出处方后，系统立即自动预审，预审通过后，方可付费取药；预审没有通过的处方，可提交到审方药师工作端，审方药师根据系统中患者相关信息（包括诊断、病程记录、手术记录、检验检查等）快速审核处方，通过数字赋能，实现了“软件+专业审方药师审核”智慧处方前置审核模式，为患者用药安全设立“双保险”，进一步推动医院临床用药更具安全性、规范性与合理性。

3.AI 预问诊。通过模拟人类医生的问诊思路与患者进行多轮人机对话，推荐就诊科室，收集汇总患者的症状、体征等疾病信息，并自动生成结构化的电子微病历，一键发送给医生。

（三）结论

基于“互联网+医疗服务”的一体化智慧服务体系建设，实现了线上线下医疗服务的深度融合，以及对患者诊前、诊中、诊后全流程的统一管理，患者可以随时随地的向医生发起图文、视频、电话问诊。通过对医生和药师 CA 资格认证，医生及药师可在线有效完成开处方、开检查、开检验、审方等，线上对接药房，患者可选择药品寄送，极大便利患者就医。随着患者病历、检查检验报告等信息与院内接通，在线也能确保看病的及时性、有效性，有效优化了就医流程，提高医护工作效率，简化工作方式，降低不良事件发生率，改善医患关系，让就医体验不再“麻烦”。

四、未来规划

随着医院后续发展需要，我院计划推出处方流转、双向转诊、互联网医保支付、慢病管理等功能或服务，逐步完善基础就医及其拓展服务。并有序规划互联网医院建设运营工作，拓展“互联网+医疗健康”的服务内涵。

川北医学院附属医院的一体化智慧服务体系是为广大患者提供便利和医疗服务的重要实践。为进一步提升患者就医体验、提升患者就医满意度、提高医疗服务质量，医院将持续优化一体化智慧服务体系功能及服务，提高易用性及可靠性，逐步构建“以患者为中心，以信息技术为载体，以医疗服务质量为原则”的全方位、全生命周期的智慧医疗服务体系。

OR-005

基于指标评价的电子病历隐私泄露风险评估

王雅晴[1],安新颖[1],耿瑞利[2]
1.中国医学科学院医学信息研究所
2.郑州大学信息管理学院

1 引言

以医疗信息和数据共享为基础的开放电子病历未来将成主流趋势，此类服务利用电子设备保存和管理数字化的健康信息和医疗记录，通过跨院病历调阅实现数据的传输和重现，是主张促成医疗及科学数据在不受其拥有单位的限制下可以在更大范围内被利用的业务合作与共享方式。在此过程中，病历内涵将由医院内部信息丰富为跨行业的健康数据[1]。然而公开病历信息对患者个人隐私安全形成挑战。“互联网＋医疗”背景下的隐私泄露风险包括信息内容、开放程度、获取方式等多个维度[2]，电子病历与上述三种因素均有关联，其包含及公开的信息种类和数量极大影响了患者隐私泄露程度，本研究主要针对电子病历文本数据属性进行风险评估。当前基于区块链技术的隐私保护方法层出不穷，云服务平台对数据安全的保证使得医疗机构容许开放部分数据用于大数据挖掘与知识发现[3]。此过程中数据层、应用层和显示层均利用部分技术方法[4]，例如定时释放加密 ID-TRE 算法、DHT 等[5]对隐私数据进行脱敏处理，实现在不同阶段(数据生成、数据存储和数据处理)形成全生命周期的隐私保护方案[6]。即使通过上述方法对开放数据进行脱敏处理，然而敏感信息仍易通过数据间融合关联等方法被推断出来，因此为防止此情形下的隐私信息泄露，本研究着重对已发布数据中引发隐私泄露风险的数据类型进行切实的量化评估和调查。

2 文献回顾

2.1 隐私泄露风险评估方法

作为识别风险发生的潜在可能性和评估最终造成的影响的主要手段，风险评估在明确特定风险和指导风险应对方面十分重要。隐私泄露风险评估由于其功能性和实际应用性强而在网络服务和数据发布的各个场景下应用广泛。部分研究通过文献调研法或德尔菲法建立风险因素的指标体系[7]，使用层次分析法(AHP)、网络分析法(ANP)或熵权法计算指标权重[8]，构造分析模型量化隐私泄露风险[9]。风险等级划分是评价过程中的重要步骤，对隐私泄露行为使用博弈论进行风险量化，构建定义数据发布与使用双方的收益函数和博弈模型[10]。部分研究使用概率估计作为风险测度方法[11]，如利用公共数据组合变量产生的案例数来估计可能发生的身份暴露在整个数据集合中所占的比例[12]。多引入数学模型例如利用五元联系数对不同阶段的隐私评估指标进行处理[13]以及利用区块链构建不完全信息静态贝叶斯博弈模型[14]。

结合专业所学内容，本研究选择指标评价法对开放电子病历隐私泄露风险进行评估。以往常用的评价指标体系构建方法有内容分析法[15]、假设推理法[16]、阶段分析法[17]、系统和解释性审查[18]等其他概念框架开发方法。确定指标属性权重[19]是构建评价体系的重要部分，现有研究通常将赋权方法区分为主观和客观两个维度。国内外学者已针对模糊信息熵开展了大量的研究[20]。信息熵倾向于关注指标结构信息表达的重要性而忽略与实际应用之间的联系[21]。单纯使用离差法则存在同样的问题[22]。以往研究表明，确定属性权重应考虑不同决策对象在实际应用中的独立性[23]，在部分情境下属性权重应充分考虑评价对象的差异从而赋予不同的权重[24]。以往研究多基于模糊综合评价[25]、矢量评价[26]、TOPSIS[27]等模型对不同方案进行测量从而实现评估。TOPSIS 方法是常用的基于数据的评价模型[28]，传统 TOPSIS 方法具有一定弊端[29]，其将所有指标等同看待，在实际的问题中不同的因子重要程度不同。

2.2 医疗开放数据隐私保护现状

医疗数据[30]保护是指通过加密、访问控制等一定的技术手段使医疗数据不泄露[31]，有利于患者敏感数据的安全访问[32]。当前国内对电子病历的保护大部分基于区块链技术，采用椭圆曲线密码

体制等方法，确保系统具备良好的鲁棒性及数据安全性[33]以实现各医疗机构间可靠的数据共享[34]。

电子病历信息平台是医疗信息与电子技术融合发展的产物[35]，是医疗机构内部使用的采集、存储和访问患者病历数据的医疗信息处理服务平台[36]。当前环境下电子病历的共享并非完全开放，其往往存在于医疗机构之间，然而在实际应用中，仍有部分病历信息以不同形式公开和传播，电子病历其较为边缘化功能的实现仍能够导致部分隐私信息的泄露，例如各种基于电子病历的移动 APP 层出不穷，其主要功能集中在远程辅助诊疗[37]、医疗机构信息管理[38]、知识服务等，作为潜在的电子病历隐私泄露途径，对该情境下的风险进行评估具有一定的指导意义。

3 研究设计

3.1 数据处理

各项数据类型指标均为描述型数据，在此采用五分制法将其按照隐私范畴分别进行模糊数值转化，通过赋值得到每条病历数据在各个数据类型指标上的具体指标值，详细规则见表 1。

表 1 病历数据模糊数值转化规则

模糊值	标识数据	半标识数据	敏感数据
0.2	涉及 3 条以上	涉及 3 条及以上（相互关联后可以识别隐私主体）	涉及 2 条及以上（与识别隐私主体信息同时出现）
0.4	涉及 1-3 条	涉及 2 条（相互关联后可以识别隐私主体）	涉及 1 条（与识别隐私主体信息同时出现）
0.6	/	涉及 3 条及以上（相互关联后不能识别隐私主体）	涉及 2 条及以上（未识别隐私主体）
0.8	/	涉及 2 条（相互关联后不能识别隐私主体）	涉及 1 条（未识别隐私主体）
0.99	完全不涉及	完全不涉及或涉及 1 条	完全不涉及

在隐私范畴设定上，能够直接识别隐私主体的数据是标识数据；无法直接识别隐私主体，但与其他来源的半标识数据关联后可以间接识别隐私主体的数据是半标识数据；隐私主体不希望公开的，若公开可能会给其带来人格尊严或财产方面损失的数据是敏感数据[39]，每条病历数据均根据上述规则针对各个指标进行赋值，分值越高，隐私信息处理程度越高，隐私泄露的风险越低，因此各个指标均为逆向指标。

3.2 隐私泄露风险评估方法

3.2.1 权重计算方法

本文采用最小二乘法确定综合权重。根据指标体系模型建立对应量表体系，以电子问卷形式通过网络发布的偶遇抽样方法面向所有群体进行调查。对结果数据求均值并进行标准化后得出指标对应的主观权重。

熵权法利用信息熵的数学概念计算信息效用值，指标的分散程度与计算所得的信息熵呈反比，熵值越小，其包含的信息量越大，重要性越强，赋予的权重则越大。假设评价对象个数为 m，评价体系包含指标数量为 n，评价对象在各个指标项上的原始得分矩阵为：

$$R = (r_{ij})_{m*n}$$

熵权法计算权重步骤如下：首先将得分矩阵列指标除以该列的最大值，得到标准化矩阵：

$$S = (s_{ij})_{m*n}$$

归一化处理得到标准矩阵：

$$P = (p_{ij})_{m*n}$$

信息熵计算公式：

$$e_j = \frac{-1}{lnm} * \sum_{i=1}^{m} p_{ij} \ln p_{ij} \quad (j = 1,2,\ldots,n) \tag{1}$$

根据信息熵 e_j 确定指标熵权，计算公式为：

$$\mu_j = \frac{(1-e_j)}{\sum_{j=1}^{n}(1-e_j)} \tag{2}$$

本文采用最小二乘法确定综合权重。设由主观调查法和熵权法确定的权重分别为 Φ_j 和 μ_j，基于最小二乘法得到的综合权重为 ω_j，目标函数为：

$$\min f(\omega_1, \omega_2, \ldots, \omega_n) = \sum_{i=1}^{m}\sum_{j=1}^{n}\{[(\varphi_j - \omega_j)\, s_{ij}]^2 + [(\mu_j - \omega_j)\, s_{ij}]^2\} \tag{3}$$

约束条件为：

$$\sum_{x=1}^{n} \omega_x = 1, \quad \omega_x \geq 0$$

用拉格朗日乘数法求解此模型，得到综合权重向量。

3.2.2 优化 TOPSIS 模型

TOPSIS 模型因其能够充分利用原始数据信息的特质，能够满足本研究的需求。具体计算步骤如下：将所有指标均正向化处理得到正向指标[40]，使用极值法处理正向指标数据，得出标准化后的数据矩阵:

$$Z = (z_{ij})_{m*n}$$

确定所有指标值的最大值和最小值，得到对应正理想解矩阵 Z^+和负理想解矩阵 Z^-。利用空间距离计算公式求得每条病历各指标值与正理想解之间加权后的距离：

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \omega_j (Z_j^+ - z_{ij})^2} \tag{4}$$

同样，病历各指标值与负理想解之间加权后的距离公式为：

$$D_i^- = \sqrt{\sum_{j=1}^{n} \omega_j (Z_j^- - z_{ij})^2} \tag{5}$$

利用与正、负理想解的加权距离计算相对贴近度作为指标综合得分[41]，计算公式为：

$$E = \frac{D_i^-}{D_i^+ + D_i^-} \tag{6}$$

3.3 评价体系构建与权重计算

3.3.1 指标体系

运用文献法对初步构建的指标体系进行修改和完善，建立引发开放电子病历隐私泄露风险的数据类型指标体系见表 2。

表 2 开放电子病历数据类型体系划分

数据隐私范畴	数据隐私种类	数据隐私条目	数据的隐私信息示例
标识数据	个人基本信息	C1 自然属性	基因、肖像等
		C2 社会属性	身份证号、手机号、病案号、医保账号等
半标识数据		C3 自然属性	性别、年龄、出生日期、身高、体重等
		C4 社会属性	姓名、户口、国籍、民族、出生地、婚育史、宗教信仰、文化水平、职业等
敏感数据		C5 行踪数据	床号、病室、医院名、生活史（居住地以及疫水毒物接触史）、入院时间、出院时间等
	个人体检信息	C6 生理数据	人体生理指标、生理周期等
		C7 体格检查数据	皮肤黏膜、全身浅表淋巴结、头部及其器官、颈部、胸部、肺部、心脏、血管、腹部、肛门直肠外生殖器、脊柱四肢等一般情况
		C8 实验室和影像检查数据	医学影像、检查胶片、病理切片等
	个人诊疗信息	C9 入院数据	既往史，患者入院疾病表现、患者主诉、入院诊断结果、病程等
		C10 出院数据	出院结果诊断、患者出院情况等
		C11 医疗操作数据	治疗记录、处理记录、医嘱等
	遗传基因信息	C12 遗传数据	家族病史、遗传病史等

该体系借鉴“健康信息”的概念[42]，将电子病历涉及到的所有数据分为个人基本信息、个人体检信息、个人诊疗信息和遗传基因信息四大类。在隐私范畴设定上划分为标识数据、半标识数据和敏感数据。

3.3.2 问卷调查法计算主观权重

根据归纳出的数据类型建立开放电子病历隐私泄露风险评价模型如图 1 所示。

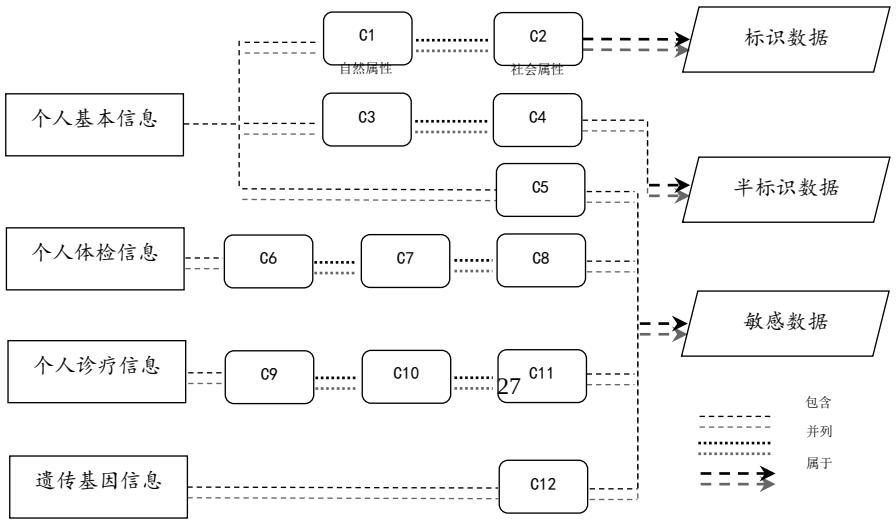


图 1 开放电子病历隐私泄露风险评价模型

主客观权重常用的组合方法有加权综合优化法、最小二乘法、乘数合成归一法等。以往研究曾提出判别主客观权重的 3 个标准：赋权模式的合理性、综合权重的可解释性、对评价对象有效性[43]。基于上述原则，本研究采用最小二乘法进行综合权重的计算。根据主客观权重数据，按照属性权重算法，使用拉格朗日乘数法进行求解，根据公式（3）构造拉格朗日函数：

$$L(\omega_1, \omega_2, \ldots, \omega_n; \lambda) = \sum_{i=1}^{m} \sum_{j=1}^{n} \{[(\varphi_j - \omega_j) s_{ij}]^2 + [(\mu_j - \omega_j) s_{ij}]^2\} + 4\lambda \left(\sum_{x=1}^{n} \omega_x - 1\right) \quad (m = 21, n = 12) \tag{7}$$

其中 Φ_j 和 μ_j 分别为已知的主观权重和客观权重，分别对公式（7）中函数 L 的 12 个自变量和拉格朗日因子参数 λ 求偏导数，得到有关 13 个变量的 13 元线性方程组，求解方程组得到最终综合权重具体得分见表 3。

表 3 各数据类型指标主观权重、客观权重及综合权重

准则层	指标层	单位	正逆性	主观权重	客观权重	综合权重
个人基本信息	C1 自然属性	无量纲	逆	0.115	0.078	0.097
	C2 社会属性	无量纲	逆	0.112	0.081	0.096
	C3 自然属性	无量纲	逆	0.041	0.078	0.060
	C4 社会属性	无量纲	逆	0.051	0.080	0.066
	C5 行踪数据	无量纲	逆	0.067	0.088	0.078
个人体检信息	C6 生理数据	无量纲	逆	0.059	0.090	0.075
	C7 体格检查数据	无量纲	逆	0.070	0.087	0.079
	C8 实验室和影像检查数据	无量纲	逆	0.073	0.086	0.079
个人诊疗信息	C9 入院数据	无量纲	逆	0.101	0.082	0.091
	C10 出院数据	无量纲	逆	0.101	0.085	0.093
	C11 医疗操作数据	无量纲	逆	0.102	0.083	0.092
遗传基因信息	C12 遗传数据	无量纲	逆	0.108	0.082	0.095

4 结果讨论

通过所调查病历在具体指标上的得分表现，探究针对不同指标即不同类型隐私数据保护采取的具体措施及存在的问题。

4.1 各指标得分具体情况

标识数据作为识别隐私主体的关键信息，现有开放病历对此类数据均作出较为完善的处理，意味着在开放电子病历中完全不涉及此类数据。在调查范围内，2.5%的数据未充分处理标识数据致使泄露。绝大多数病历均公开 3 条以上的半标识数据，然而一般情况下公开的数据之间相互关联未能精准识别隐私主体。

敏感信息作为患者不愿进行公开的信息，然而其部分内容在病历数据使用范畴下具备学习和利用的价值，在一定程度上很难完全隐去，但其中某些信息在此情景下并未发挥实质性作用，部分病历在检查报告单、现病史及入院记录中暴露出所在地区、就诊医院名称及就诊时间等行踪数

据。生理数据和体格检查数据对临床判断具备较高价值，在未识别隐私主体的情况下公开此类信息不会造成严重的隐私泄露风险。

4.2 各指标综合得分及得分范围对比

数据综合得分在[0.068，0.765]区间内，均值为 0.138，总体隐私泄露风险程度较低，但仍存在个别隐私泄露风险较高的病例信息被公开的情况。各指标间均值差异较小，然而其得分范围差异较大。标识数据指标间得分范围基本相同，且相较其余类型指标区间长度最小，说明其在不同科室来源的病历间得分差异较小。行踪数据和体检数据得分范围最大，诊疗数据次之。社会属性的半标识数据与标识数据性质相同，得分范围较小；自然属性的半标识数据和遗传数据得分范围处于中等水平。

5 结语

通过上述分析，针对未来开放电子病历数据保护处理本研究给出以下建议。首先应对不同格式信息进行差异处理，除文本信息外，对包括照片、化验单及影像检查报告等在内的图片格式数据应进行二次处理，避免因忽略其中隐含的其他属性数据而导致隐私的泄露。行踪数据作为敏感数据其同时具有半标识数据的性质，应重点处理出现在入院数据，例如现病史、既往史和病程中的潜在行踪数据，一般包括就诊医院名称及具体时间以及通过简单计算可得出具体时间的信息。由于医疗服务的进步和完善，电子病历中倾向于详细记录患者职业等社会身份信息及生活表现，这种趋势给电子病历的开放带来更大的敏感信息泄露风险，然而针对知识类 APP 此类信息参考价值相对较小，软件服务供应商应在满足功能需求的前提下最大程度隐匿此类信息。

OR-006

基于机器学习模型公平性及可解释性的肺动脉高压分级诊断研究

段一凡[1] 王瑞琦[2] 钱庆[1]* 王浩林[2]*
1. 中国医学科学院医学信息研究所
2. 重庆医科大学医学信息学院

1.引言

肺动脉高压（PH）是肺动脉压升高，导致血流动力学异常的病理生理状态[1]，PH 是一种潜在致命性疾病。2009 年欧洲心脏病学会颁布的 PH 诊治指南中，将 PH 分为 5 大类型，动脉性 PH、先天性心脏病相关肺动脉高压（CHD-PAH）、肺部疾病或缺氧 PH、慢性栓塞 PH 以及未知因素所致 PH[2]，不同类型的 PH 临床表现各异，但 PH 的典型症状表现为呼吸困难、疲劳等，临床症状缺乏特异性，因此 PH 诊断具有挑战性。2021 年中华医学会制订了《中国肺动脉高压诊断与治疗指南（2021 版）》，指南中将 PH 诊断结果分为两类分别是高度或中度、低度[3]。因此早期诊断 PH 分级对于防止并且恶化，降低死亡率具有重要意义。

近年来，数据挖掘技术的成熟为基于数据驱动的方法预测患者 PH 提供了研究方向，现有研究[4-6]多集中在基于超声心动图、电子病历预测患者是否患有 PH,对于患者 PH 分级预测的研究较少，同时针对 PH 分级结果预测背后的模型公平性及可解释性问题的研究也相对薄弱。

在医疗研究领域，公平性问题主要考察算法是否公平对待不同的亚组，如 Cirillo 的研究表明，不考虑个体之间的性别差异将产生歧视性结果[7]，可解释性主要考察机器学习模型的预测结果

是否可为临床诊断提供诊断结果的解释性信息，常用的模型解释工具包括 SHAP（Shapley Additive exPlanations）、LIME（Local Interpretable Model-Agnostic Expalnations）等，目前随着临床决策支持系统（CDSS）逐渐成熟，机器学习模型的公平性及可解释性问题逐渐引起了更多的关注。

本文将从机器学习的算法着手，分析不同算法在 PH 分级诊断预测中的模型，并讨论机器学习辅助下的模型公平性及可解释性问题，探讨机器学习在 PH 分级诊断预测场景中的应用策略。

2.资料与方法

2.1 研究对象

本研究数据来源于重庆地区七家医院 2010-2021 年确诊为 PH 患者电子病历数据资料。原始数据样本由 10853 名患者组成，结局变量为患者的 PH 程度（中度或重度=1，一般=0），其中中度及重度 PH 患者 1863 例（17.2%），一般 PH 患者 8990 例（82.8%）。人口学统计变量如（性别，年龄等）不完整的病历记录被排除，最终共 5161 名患者数据纳入后续分析。本研究获得了重庆医科大学伦理委员会的批准。

2.2 数据预处理

通过机器学习随机抽样方法将数据按 7:3 比例随机分为训练集（一般=3137，中度及重度=475）和测试集（一般=1356，中度及重度=193）。训练集用以模型构建，测试集用以模型验证和评估，变量筛选、数据填充、数据不平衡处理等预处理操作均在训练集上完成，防止出现数据泄露等问题。

首先变量筛选上，缺失值超过 20%以及方差为 0 或接近 0 的特征被剔除，最终纳入 40 个指标（包含基线特征、实验室检查结果、诊断等）。之后，数据插补利用“Miss Forest”方法，Miss Forest 是一种非参数的缺失值填补方法，可以有效处理包含连续变量以及分类变量的缺失值。最后，基于少数类样本合成过采样技术（SMOTE）进行数据平衡，其可以根据少数类属性特征随机产生相似样本，有效平衡少数类和多数类之间数量，减少数据倾斜，使训练数据集最终平衡为 3137 个 PH 中度及重度和 3137 个 PH 一般的个体。

2.3 模型构建与验证

目前有研究利用不同机器学习模型进行 PH 诊断预测研究，因此本文首先选用了 7 个使用较多的经典机器学习模型进行训练，分别是新型梯度提升框架（Light Gradient Boosting Machine, LightGBM）、随机森林（Random Forest, RF）、梯度提升机（Gradient Boosting , GB）、支持向量机（support vector machine，SVM）、K-最临近（k-nearest neighbor algorithm，KNN）、决策树（Decision Tree, DT）、极度随机树（Extra Trees Classifier, Extra Trees）。同时根据算法的原理，对模型的参数进行调节，找出最优的模型参数，这里利用的是 optuna 库进行自动化调节，optuna 是一个使用 python 编写的超参数调节框架。训练数据集用于拟合模型并调整模型参数，测试数据集用于评估模型。本文通过区分度、公平性指标、临床实用性 3 个方面来评价并选择最佳预测模

型：绘制 ROC 曲线，获取基于混淆矩阵的各类指标判定模型的预测效能；为了量化机器学习模型在临床决策的公平性，本文通过几个公平性指标进行评估，包括漏诊率（FNR, 假阴性）、误诊率（FPR, 假阳性）、人口均等比（Demographic parity ratio）以及均衡优势比（Equalized odds ratio）。漏诊率表示实际为阳性的人中，被诊断为阴性的概率；误诊率表示实际为阴性的人中，被诊断为阳性的概率。

$$FNR = \frac{FN}{(TP+FN)} \quad (1)$$

$$FPR = \frac{FP}{(TN+FP)} \quad (2)$$

尽管 FNR 和 FPR 可以观测那些亚组被机器学习模型所区分，但他们仍无法全面的量化机器学习模型的公平性，因此，本研究增加人口均等比和均衡优势比[8-9]进一步度量模型的公平性效应，使用性别作为敏感属性。人口均等比（Demographic parity ratio）表示性别群体中最小和最大亚组选择率的比率，高人口均等比表明对于性别这一敏感属性，预测结果受到敏感属性不同的影响更小。均衡优势比（Equalized odds ratio）定义为召回率和 FPR 比率之间的较小值，区别于 FPR，均衡优势比可以观测机器学习模型是否在不同敏感属性亚组件产生相等的召回率和 FPR。以上所有公平性度量指标的取值范围为 0-1，无偏模型应达到接近 1 的公平性得分。

制决策分析曲线（decision curve analysis, DCA）,根据患者行干预性治疗后的获益情况评估预测模型的临床实用价值。最后，通过 SHAP 分析各特征对预测的贡献程度，即 Shapley 值，从而分析影响 PF 发病程度的危险因素。

3.结果

3.1 模型的构建与验证

7 个模型的分类性能见表 1，ROC 曲线结果见图 1 所示。在所有 7 个模型中，RF、GB 和 LGBM 产生了最好的区分性能（AUC=0.82），而且 RF 的召回率更高达到 0.86，召回率高在临床决策支持系统中的开发更为实用，同时 RF 在公平性指标的表现也较其他模型更显著，其中人口均等比为 0.91、均衡优势比为 0.81（本研究中该两个公平性指标均以“性别”变量作为敏感属性）。最后，DCA 曲线显示 7 种模型实现了不同阈值下的相对于无治疗计划及全部干预的净效益。综合比较后，认为随机森林算法为最佳的预测模型（准确度=0.83，精确度=0.86，召回率=0.86，F1 值=0.84）。

进一步的，本文基于随机森林模型分析了每一类性别、保险类型亚组的其余两个公平性指标的效应，即漏诊率和误诊率（见图 3）。具体来说，与女性患者相比，男性被漏诊的可能性高 3.1%，被误诊的可能性高 1.4%；城镇职工医疗保险类型的患者漏诊率最高（47.2%），全自费类型的患者误诊率最高（13.1%）。

表 1 七种机器学习模型在 PH 分级预测中的性能和公平性

模型	模型预测性能评价指标				公平性指标	
	准确度	精确度	召回率	F1 值	人口均等比	均衡优势比
RF	**0.83**	**0.86**	**0.86**	0.84	0.91	**0.81**
SVM	0.54	0.84	0.54	0.61	0.92	0.80
GB	0.80	**0.86**	0.80	0.82	0.89	0.76
LGBM	0.84	**0.86**	0.84	**0.85**	0.82	0.53
DT	0.77	0.83	0.77	0.79	0.70	0.67
Extra Trees	0.74	0.83	0.74	0.77	0.76	0.62
KNN	0.64	0.80	0.64	0.70	0.82	0.62

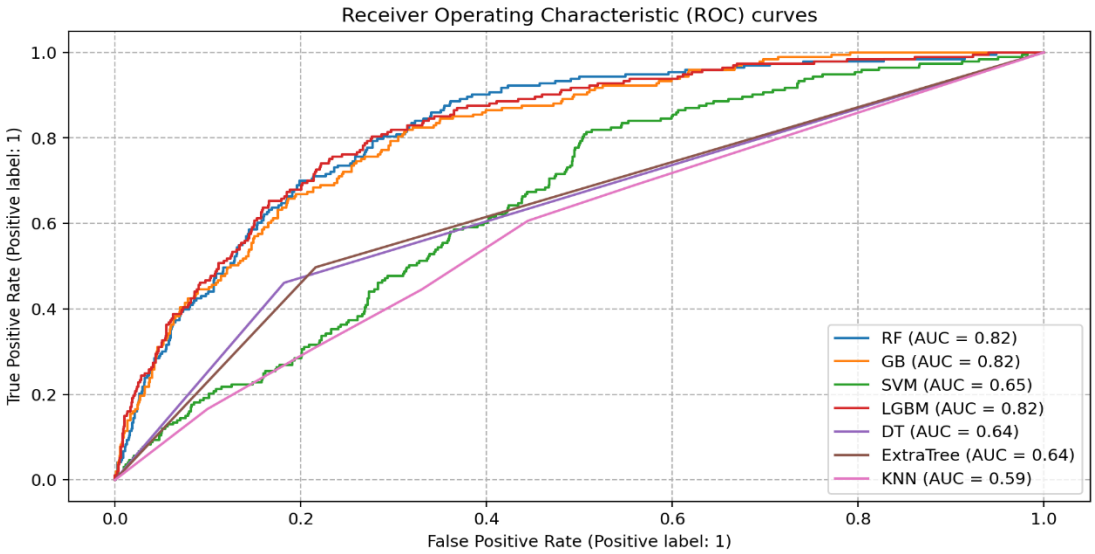


图 1 七种模型在测试集的 ROC 曲线

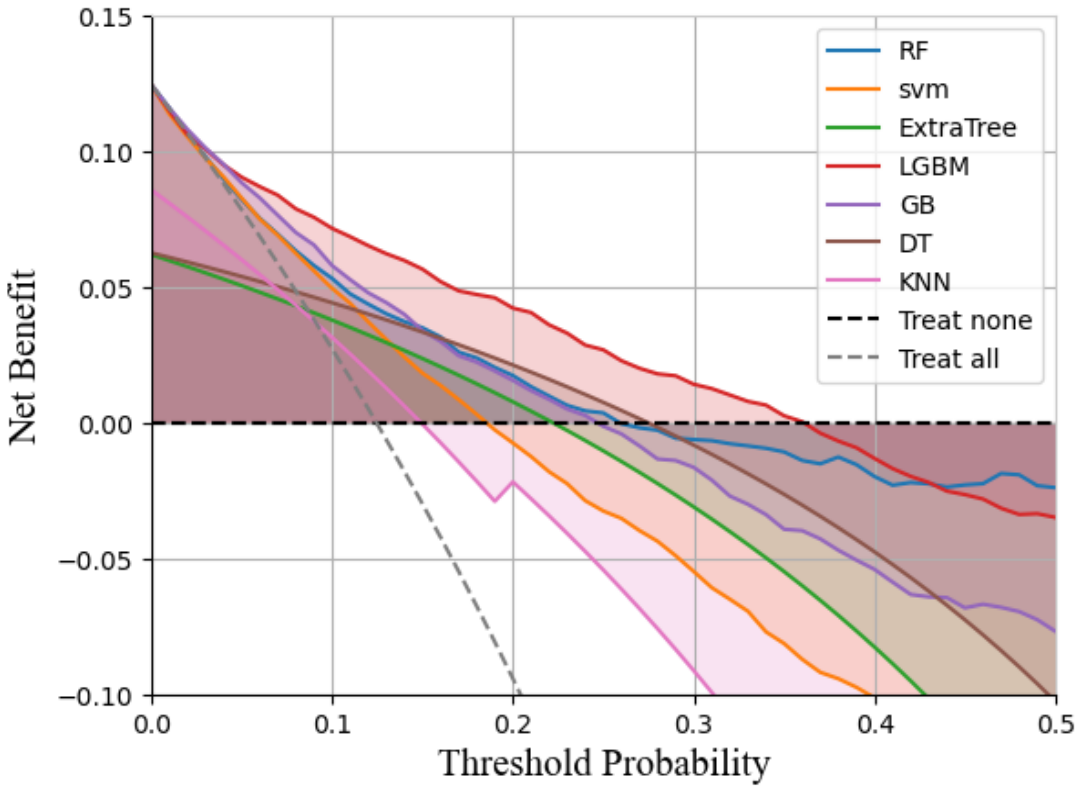


图 2 七种模型的决策分析曲线（DCA）

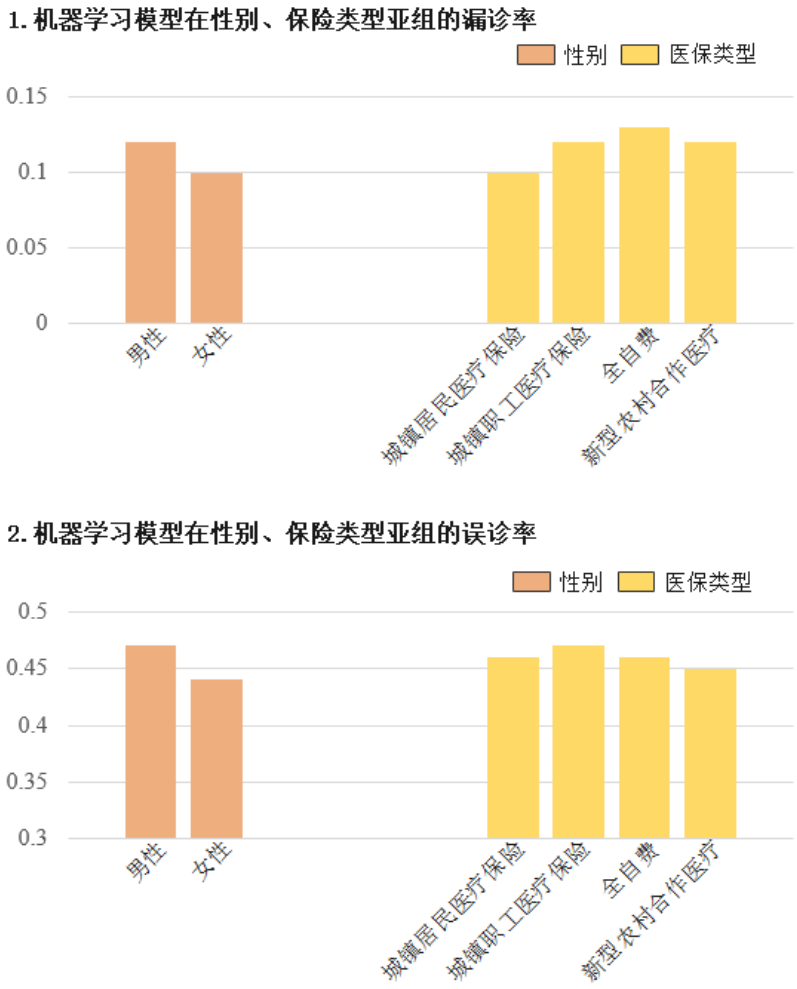


图 3 基于随机森林模型的性别、保险各亚组漏诊和误诊率

3.2 模型解释

SHAP 摘要图结果显示，PH 中度及重度的发病危险因素排序前十名为凝血酶时间（TT）、是否室间隔缺损（VSD）、肌酐水平（Creatinine）、中性粒细胞百分比（PMN）、年龄（Age）、凝血酶原时间（PT）、总钙水平（TC）、是否先天性三尖瓣关闭不全（CTVR）、是否患有肺炎（PN）、是否诊断为呼吸衰竭（RF）。

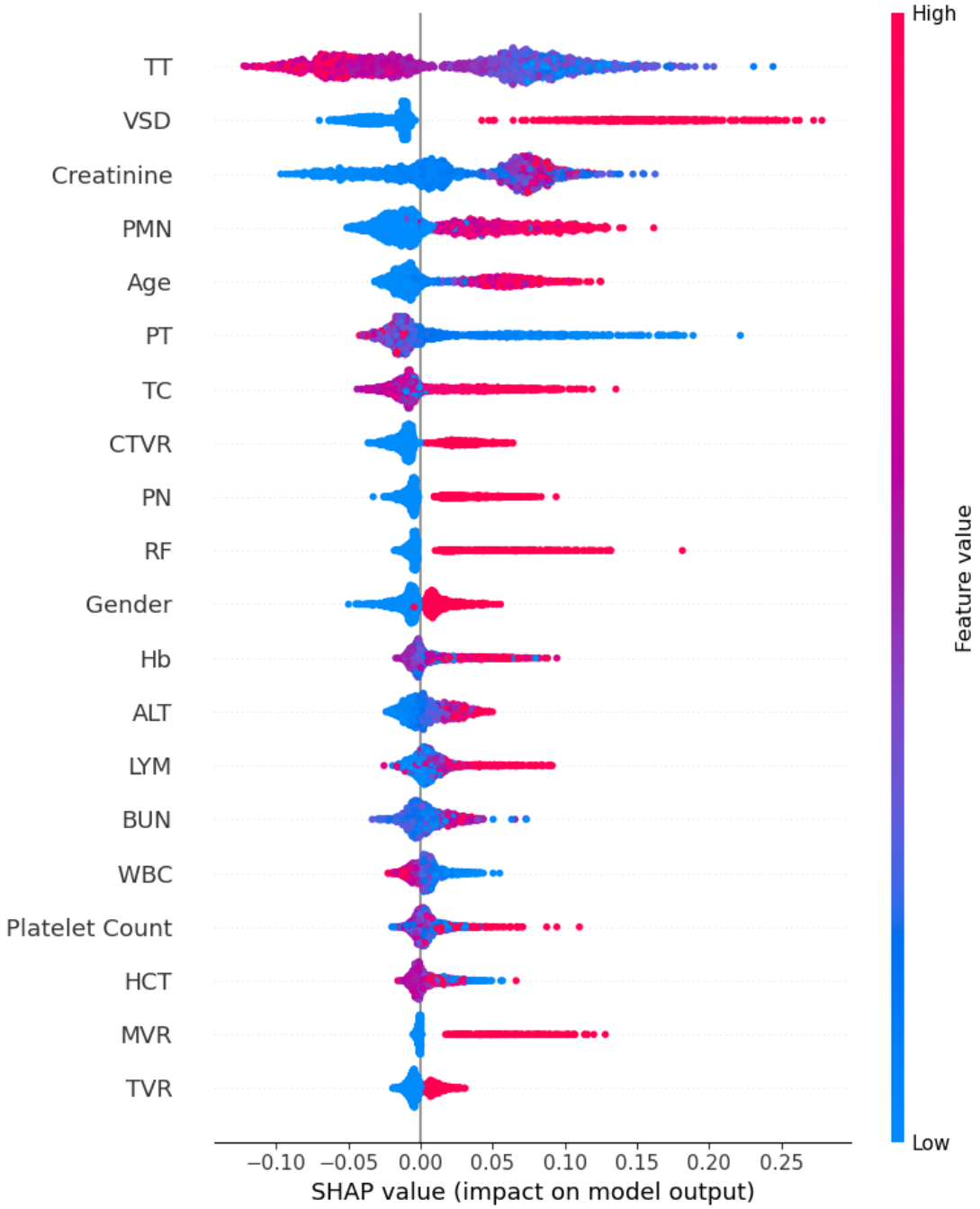


图 4 SHAP 摘要图

3.3 系统实现

本研究基于模型性能、公平性指标分析结果选择 RF 作为适用 PH 分级预测的机器学习模型，并实现用于 PF 患病程度分级预测的系统，所开发的系统可以通过输入模型所需的特征指标值，将自动预测患者患病结局（见图 5）。

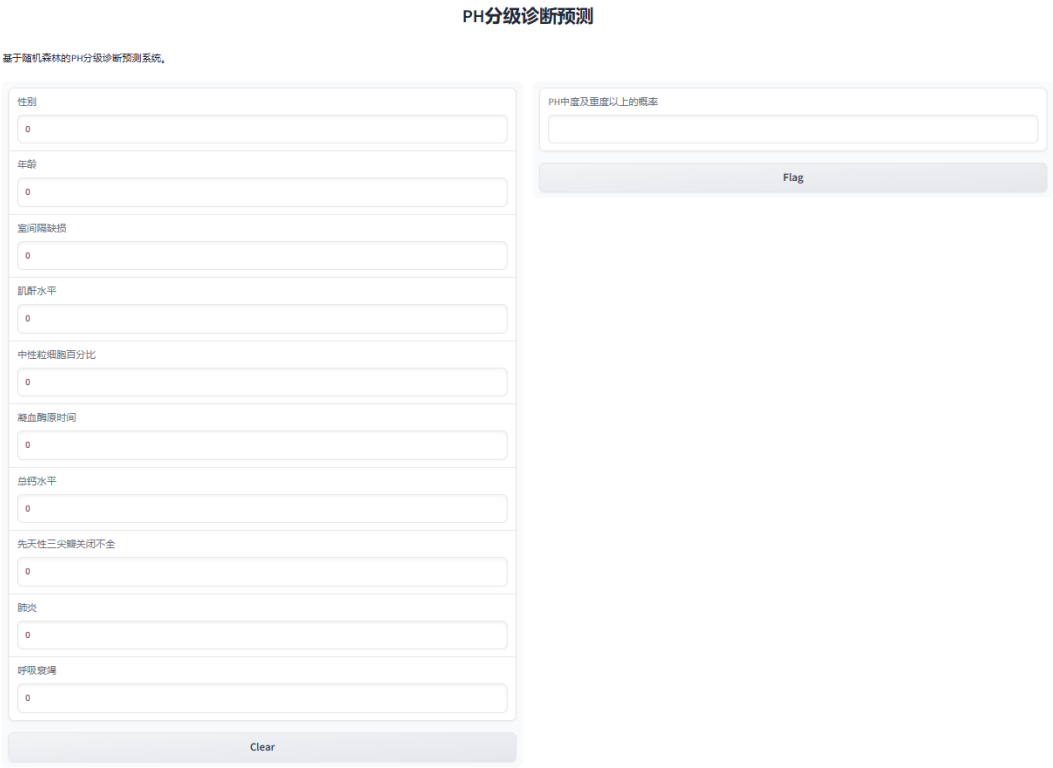


图 5 在线预测系统示例

4.讨论

本研究结果表明，利用机器学习算法并基于电子病历信息预测 PH 患病程度的是基本可行的，而且在真实世界研究中，PH 患病的漏诊和误诊会导致患者延迟接受治疗或医疗管理不当，同时机器学习模型的诊断偏差会在不同亚组间表现不同。因此本研究区别于先前 PH 预测模型的相关研究[10-12]，更聚焦于模型临床适用性以及公平性问题的分析。

研究使用 7 种机器学习算法来构建 PF 患病程度预测模型。尽管不同算法在不同任务中有各自优势，但通过构建以及利用区分度、公平性、临床适用性等指标评估后发现，随机森林模型表现出最高的综合性能，因此本研究最终选择使用随机森林算法构建模型预测患者 PH 患病程度。

研究发现，男性的误诊率以及漏诊率均高于女性亚组，城镇职工医疗保险类型亚组的漏诊率最高，全自费保险类型亚组的误诊率最高，机器学习模型选择性地在不同亚组上存在不公平性问题。根据流行病学结果，年龄大于 65 岁的人群患病率更高，因此对于高龄男性这类人群，预测模型可能容易对其产生双重危险。

综上所述，本研究基于随机森林机器学习算法建立模型来预测肺动脉高压的患病程度。该模型具有良好的预测精度和临床实用性，同时该模型在不同亚组的预测公平性上表现良好，同时本研究基于 SHAP 可解释工具分析结果指出了是否有室间隔缺损、年龄、中性粒细胞占比等风险因素，进一步基于该模型开发了在线预测工具，可供相关临床医生进行辅助决策使用。

OR-007

促进高血压患者主动参与临床实践的决策辅助工具开发

张怡，马亚强，冒杨洋，陈家然，谢同玲，耿劲松
南通大学医学院医学信息学系

1 引言

高血压是我国患病人数最多的慢性病之一，是心脑血管疾病死亡的最重要的危险因素。《“健康中国 2030”规划纲要》将高血压的防控作为慢性病防控的重要目标之一。从 2004 年到 2018 年，我国成人高血压患者的高血压知晓率从 30.8%上升至 38.3%，治疗率从 25.9%上升至 34.6%，控制率从 7.1%上升至 12.0%[1]。可见，虽然高血压的知晓率、治疗率和控制率虽然有所提升，但整体水平仍有待提升。循证医学是循证实践的医学过程，它将最佳的研究证据与医生的临床专业技能、患者独特的价值观以及决策环境相结合。循证临床实践有助于改善患者的临床结局[2]。患者参与、医患共同决策已被视为衡量医疗服务质量的一个重要标志。然而，如何将复杂的信息有效地传达给患者，这对临床医生来说是挑战[3]。循证临床实践还面临着如何有效融合患者价值观和意愿的难题。本文开发促进高血压患者参与临床实践的决策辅助工具，提供有关高血压的防治的信息以及降压药的多维度证据，帮助患者做出体现其价值观和意愿的临床决策。

2 研究方法

2.1 决策辅助工具的需求分析

系统功能的需求分析是决策辅助工具需要实现的具体功能，通过界定和描述系统中的功能，为后续的需求开发、设计和测试工作提供依据。临床决策辅助工具的总体需求就是帮助患者获取高血压及并发症的信息，辅助患者参与降压治疗的循证临床实践。因此，本系统的功能需求主要包括如下内容：工具简介及患者临床信息管理、疾病信息管理、心血管病风险评估管理、降压药的多维价值评估和循证决策管理。工具使用的结果反映了患者的药物治疗偏好，医生在了解患者偏好的基础上制定符合患者价值观的治疗方案。同时，管理员对工具中的证据进行定期更新，以满足高血压患者的信息需求。

2.2 工具开发的相关技术

工具以网站的形式呈现，整体开发采用 MVVM 框架。网页基于 B/S 模式，B/S 模式对客户端性能要求很低，对用户系统没有任何影响，具有成本低和安全性高的优点。通过前后端分离的方式，前端采用 Vue.js 进行数据绑定和页面设计，采用 Element UI 进行页面的布局渲染。Vue.js 能够高效地进行视图维护和数据管理，提高开发代码的效率。Element UI 作为设计和交互非常优秀的前端 UI 框架，组件多样，样式丰富且代码稳定。后端基于 Java 语言，采用 Redis 数据缓存技术。Java 语言相较于 C++语言更加简单，操作性较强，降低了使用者的操作难度，且能够适应不同的操作平台。Redis 数据库可使用数据类型丰富，具有运行性能高、持久性好且配置简单的特点，可

应用于各种不同的场景。通过 MySQL 数据库存储研究证据。MySQL 是一个关系型数据库管理系统，数据直观且易于操作。

3 决策辅助工具的设计与实现

3.1 工具的总体设计

通过工具需求分析确定了工具开发的总体目标和总体框架结构（图 1），完成了工具的设计。

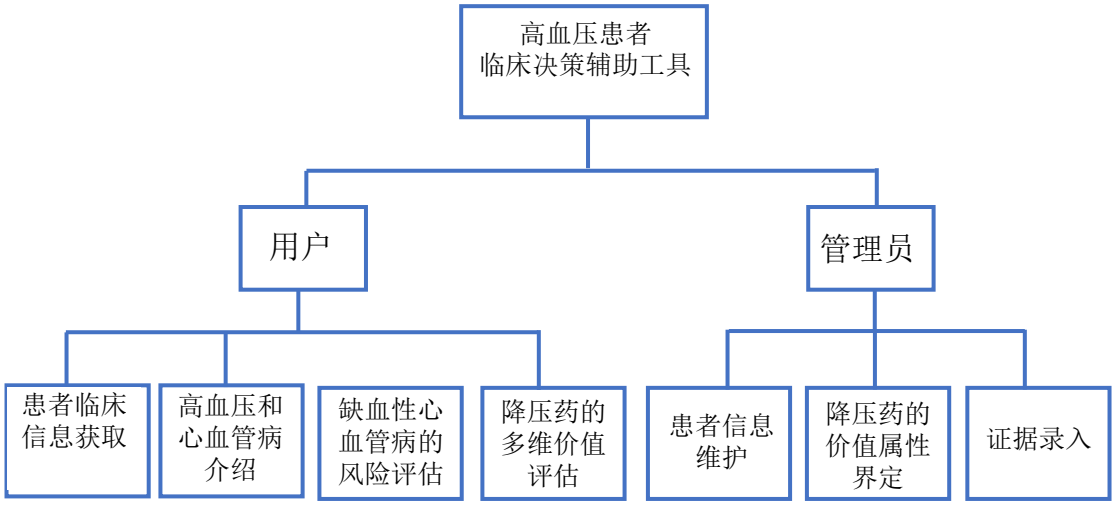


图 1 工具的总体框架

工具预期实现如下功能，分别是：（1）患者临床信息获取：让患者了解高血压及决策辅助工具，获取患者的临床基本特征；（2）高血压和心血管病的介绍：呈现高血压和心血管病的临床知识[4,5]，提升对高血压和心血管病的科学认知；（3）缺血性心血管病发病风险评估：结合我国人群缺血性心血管病的预测模型，形成缺血性心血管病的发病风险评估工具；同时根据指南提供预防或延缓心血管病发生的循证建议，利于患者更好地了解疾病的预后和配合医生进行病情控制；（4）降压药的多维价值评估和循证决策：患者通过查阅降压药的适应证和注意事项，以了解适宜的药品类别；工具提供体现降压药多维属性的介绍和体现多维价值的证据，让患者对降压药的价值进行循证的量化评估，以真正实现医患共同决策和循证临床决策。除上述核心功能，还有登录注册界面等基础功能设计。工具使用者为普通用户和系统管理员。普通用户通过步骤提示查看相应模块的内容，完成工具的使用。系统管理员在后台进行信息维护操作，主要包括患者信息、降压药的价值属性以及研究证据。

3.2 关键技术设计与实现

（1）工具的决策支持模块

工具以模块化的方式为高血压患者参与临床实践提供决策支持，逐步引导患者获取高血压、心血管病及降压药的相关信息，最终实现体现其价值观的决策（图 2）。工具提供了高血压诊断和

分类标准、心血管危险分层标准和心血管病评估功能，便于提升患者对高血压及心血管并发症的认知，加强血压的自我管理，从而改善高血压的控制水平并降低心血管病的发生率。工具通过降压药的证据呈现，帮助患者获取高质量的证据，促进患者在知晓证据的基础上主动参与临床决策。患者对降压药多维价值的循证评估，体现了“以患者为中心”的循证决策理念。

图 2 工具的决策支持模块

（2）引言模块

工具在引言模块对高血压的患病率、危害和疾病负担予以描述，并对工具的使用作简要介绍。在此模块中匿名获取患者的临床基本特征，主要包括病程（初次就诊患者初次发现血压异常的时间，或已确诊为高血压的时间）、目前血压控制情况以及是否患有并发症（图 3）。

图 3 患者的临床信息获取模块

（3）缺血性心血管病的风险评估模块

工具用表格清晰地展示高血压的诊断依据、分类标准以及心血管危险的分层标准，让患者结合自身的收缩压、舒张压、患病史和危险因素进一步了解病情。心血管病是高血压患者常见的并发症，因此基于我国缺血性心血管病 10 年发病危险度评估表[6]开发了心血管病的风险评估模块。患者可以通过输入性别、年龄、身高、体重、是否吸烟以及患有糖尿病、总胆固醇值、收缩压和舒张压的数值，获知未来 10 年患有缺血性心血管病的风险，从而形成科学血压控制的意识（图 4）。之后，可进一步浏览有关高血压和心血管病干预的循证建议[5,7,8]。

图 4 缺血性心血管病的风险评估模块

（4）降压药的多维价值评估模块

降压药的多维价值评估模块是工具的核心，该模块首先介绍了常用降压药物的适应证和注意事项。工具呈现的药物类别包括利尿剂、β 受体阻滞剂、钙通道阻滞剂、血管紧张素II受体拮抗剂、血管紧张素转换酶抑制剂、血管紧张素受体脑啡肽酶抑制剂。患者在对降压药物有了初步的认识后，选择是否愿意接受药物治疗，结合自身病情选择想要了解的药物类型。

之后，可进入该类药物的多维价值评估界面。该界面呈现了降压药在 10 个属性上的研究证据，包括血压控制效果、心血管事件的风险、低血压事件的发生率、高钾血症的发生率、对肾功能的影响、延长的生命年、对生活质量的影响、体重的变化、自付费用、给药频次以及单次用药剂量。患者结合系统中提供证据对降压药的各维度属性进行评分。系统记录并保存分值，以线性加权的方式量化药物的综合价值（图 5）。

图 5 降压药的多维价值评估模块

4 讨论

4.1 决策辅助工具有助于加强高血压患者的健康管理认知

心脑血管病的防治是《健康中国行动（2019-2030）》的重大专项行动的关键环节，高血压是心脑血管病防治中的重中之重。因此，以高血压防治为抓手，将有效推动我国心脑血管疾病防治进程。“互联网+”的高血压患者健康管理是信息技术发展的必然产物，是传统高血压管理方案的拓展，在提高患者的高血压知晓率和控制率，缩短就诊时间和降低再就诊率等方面具有一定的优势[9,10]。然而，目前高血压健康管理的信息技术研究聚焦远程医疗、基于移动端的用药提醒和管理、智能可穿戴设备等领域[11]。本文借助互联网技术开发的决策支持工具，基于循证临床实践指南提供患者高血压和心血管疾病的临床知识，提升患者对高血压的健康管理认知水平，加强血压的自我管理。

4.2 决策辅助工具有助于患者主动参与临床实践

患者对自身疾病和临床治疗方案了解得越多，在决策时就可以和医生进行更深入的交流[3]。医患之间可以通过信息共享加强沟通，促进患者更好地参与治疗决策，从而制定最佳且最符合患者期望的治疗方案，增强患者治疗依从性并更好地改善预后[12]。近年，计算机与信息技术发展迅速，大数据与人工智能已深度融入医学领域，计算机化的决策支持技术在临床诊疗过程中发挥着越来越重要的作用。然而目前高血压的医患共同决策模式和决策支持工具仍处于探索阶段[13]。循证医学强调以患者为中心，以个体化治疗为首要任务。本研究开发的决策辅助工具以促进患者主动参与临床实践为目标，对证据中的关键信息进行提取，以患者及家属易于理解的方式来呈现，使其及时获得完整和准确的医学信息，从而更加有效地参与临床决策。

4.3 降压药多维价值评估模块有助于实现医患共同决策

研究表明多数患者渴望参与医患共同决策[14]。然而，如果医患之间存在信息不对称，则难以有效实现医患之间的平等对话。患者作为信息的接受者，在未充分了解治疗方案的情况下做出的决策具有不确定性，会导致较差的治疗依从性。高血压患者对治疗方案的信息需求是多维度的，在选择治疗方案时不仅仅关注治疗获益，还会关注副作用、适宜性和可负担性等诸多因素[15,16]。因此，本研究开发的降压药多维价值评估模块，整合了降压药不同维度的高质量证据，以患者易于理解的方式展示，以便患者知晓降压药的多维价值属性。患者结合降压药的研究证据对药物的价值进行客观评估，做出体现其意愿的决策。降压药的价值评估过程体现了患者对药物各维度属性的偏好，循证的评估结果能够帮助医生更好地了解患者的价值观，促进更加科学和高效的医疗决策。

5 结语

促进高血压患者参与临床实践的决策辅助工具，能够拓宽患者的高血压和心血管疾病知识，提升患者对于疾病的认知度。缺血性心血管病的风险评估模块让患者了解不良预后结局的危险因素以

及循证的干预措施，有助于患者形成科学的血压控制意识。降压药的多维价值评估模块提供了体现药品有效性、安全性、可负担性和便捷性的研究证据，让患者在知晓证据的基础上参与降压药的多维价值评估。评估结果使得临床医生更好地了解患者的价值观和意愿，通过医患共同决策制定符合患者偏好的治疗方案，从而提升患者的治疗依从性并改善预后。

OR-008

医院信息部门基于低代码平台的系统敏捷开发实践

谢君①② 蓝灵③ 何传奇③ 袁璐② 王增① 石锐①
① 四川大学华西医院信息中心
② 三亚市人民医院/四川大学华西三亚医院
③ 四川大学华西医院膳食中心

1 引言

2022 年 8 月，由于管理及业务服务形态发生变化，医院患者及职工原由食堂按需供应，需立即转变为统一供应所有在院人员的一日三餐，供餐压力剧增。医院膳食中心将供餐调整为食堂与第三方餐饮公司共同供应的混合模式，并建立相应统一订餐流程。如何快速响应业务需求变化，采用信息化手段进行数据的规范填报与高效统计、提高膳食供应效率，成为当务之急。

2 现状分析

膳食中心采用共享在线 Excel 的方式进行科室数据的快速收集。通过锁定单元行来控制数据变更、取消文件共享方式来启停填报业务，利用 SUM 等统计公式实现数据汇总。但随着时间的推移，因患者入出院、转院导致加退餐、患者从缓冲病区到普通病房的转移等频繁发生，膳食中心需统计出各套餐的变化量，形成对比报表发给公司，供配餐时调整。统计是所有环节中工作量最大、最影响工作效率的环节，运用 Excel 统计的方式还存在着诸多不便。

（1）统计自动化程度有限，简单重复工作量大

为了各科室查找填写方便，膳食中心将填报表分成患者、职工两大类，再根据配送地点，细分成十多个相同格式的填报表（Excel 文件），每个填报表中再用一个表单代表一个科室，由各科室专人填报。每天填写完毕后，膳食中心逐个表单手工拷贝数据并汇总。

（2）文档管控功能不够，数据易受误操作影响

由于同一在线文档有多个科室的填报员在同时操作，存在科室间误删误改的情况。手机上的复制粘贴、行列增删等操作，经常导致单元格格式变化、合计数据移位，求和公式不准确。管理人员的手工统计，也易产生误操作，最终导致统计不准确，引发矛盾和投诉。

（1）业务逻辑功能不够，无法灵活应对业务变化

通过取消在线表格的共享以确保发送供餐公司版本一致性，其后一旦发生加餐、退餐、转科，科室需另行填写加餐、退餐表，管理人员需比对并重新计算最终预订量，再联系供餐公司调整。若科室填写数据有误，如某菜品出现退餐量大于之前订单的情况，膳食中心需联系科室核实更正。Excel 不方便处理该类业务逻辑，管理人员需不断手动查阅、核算和沟通。

综上，在配餐压力大、工作要求高的背景下，取代原有统计方式，用信息化手段进行准确高效的数据统计势在必行。因此，医院决定，在开发模式采用 SCRUM 敏捷模式加快交付速度，在开发工具上采用低代码平台提升开发效率。

3 SCRUM 敏捷模式运用

3.1 SCRUM 敏捷模式介绍

敏捷是一种通过创造变化和响应变化在不确定和混乱的环境中取得成功的能力。而敏捷软件开发是基于敏捷宣言定义的价值观和原则的一系列方法和实践的总称。敏捷开发以用户的需求进化为核心，采用迭代、循序渐进的方法进行软件开发。

Scrum 是迭代式增量软件开发过程，是敏捷方法论中的重要框架之一，通常用于敏捷软件开发。Scrum 包括了一系列实践和预定义角色的过程骨架。敏捷与 SCRUM 的区别是，敏捷是思想、原则，而 SCRUM 是敏捷思想指导下的实现框架、方法。

3.2 SCRUM 敏捷模式主要要素在本项目中的运用

SCRUM 框架主要包含要素有 3 个角色、3 个工件、5 个事件、5 个价值观。本项目中，3 个角色，即产品负责人(Product Owner)、开发团队(Development Team)、Scrum Master 分别为信息科主管、信息科工程师 1 名（全职）+膳食科管理人员 2 名（兼职）、信息科主管。3 个工件(Artifact)即产品待办列表(Product Backlog)、迭代待办列表(Sprint Backlog)、潜在可发布产品增量(Potential Shippable Product Increment)，均采用“需求功能管理表.xlsx”统一管理。5 个事件，即迭代，按 2 至 3 天为周期进行；迭代计划会议(Sprint Planning Meeting)及迭代回顾会议采用了远程会议；每日站立会议(Daily Scrum Meeting)及迭代评审会议在膳食中心办公室进行。

3.3 SCRUM 在本次项目中的迭代推进情况

大约 2 天左右为一个冲刺（Sprint），即产品增量发布周期，不断推陈出新。项目分为 4 个冲刺，见下图。需求沟通、新模块开发、已上线模块 Bug 修复、操作体验改善等工序根据需要混合进行。先完成基础设置与职工业务版块，再以职工业务为模板，开发患者业务，最终形成较为完善的系统，之后的主要工作为操作体验改进。

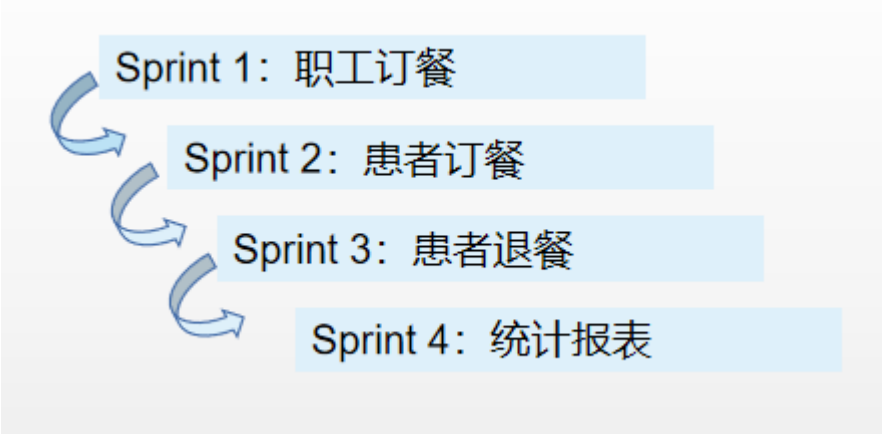


图 1 项目冲刺（Sprint、划分产品增量发布周期）划分

4 低代码平台开发工具运用

4.1 低代码平台介绍

低代码开发平台，是指编写较少量代码、主要靠可视化组件配置的方式生成自定义应用程序的快速开发平台。低代码平台一般采用“所见即所得”的开发模式，通过拖拽组件即可设计界面，再结合相应的逻辑设置，就可以完成界面设计。界面之间可以像在 Office 中画流程图一样，设置依赖关系，实现界面间联动，最终像搭积木一样完成系统的搭建。

目前国内外有大量低代码平台产品，整体市场增长率逐年上升，生态已日趋完善。低代码平台具有丰富的组件库，能高效完成大多数 MIS 系统（Management Information System，管理信息系统）常用的输入控制、查询、表格和图形化展示、界面联动等功能开发，在诸多行业已有运用案例。

对于本订餐统计系统，业务形态的紧急变化决定了采用迭代式开发模式，使用快速开发工具，边开发、边上线、边完善。而其典型的 MIS 系统 CRUD 数据操作特性，使得该系统非常适合基于低代码平台进行开发。

4.2 基于低代码平台的功能设计

4.2.1 患者职工预订与退餐

科室人员使用手机微信端填报，选择患者或职工订餐业务，即可进入各自科室的填报表。填写科室患者住院号、姓名、床号及早、中、午餐（含普食、特殊饮食等）数量，或职工姓名、工号及三餐数量，点击保存即完成填报。

对提交时间点进行判断，超过则不能提交相应餐次。早餐须在用餐前一日 21:00 之前提交，午餐须在当日 10:00 之前提交，晚餐须在当日 16:00 之前提交；退餐数量设有限制，大于已提交过的订餐总数量时无法录入，以减少无谓的人工核对与沟通时间。

4.2.2 订单数据版本控制

通过系统管理端的数据“确认”操作来代替 Excel 中数据行锁定的行为，以此来控制该数据能否被修改。数据初始被提交后默认为“待确认”状态，可随时修改；管理人员在后台确认后，变为“已确认”状态，此时意味着将这些订单发送给供餐公司，不允许被修改。科室若需修改预订量，可以在原填报表中像操作 Excel 一样在新的一行填入数据即可，或进入退餐业务进行登记。每种登记均需膳食中心确认，以保证对数据变化的把控。

4.2.3 每日预订启停管理

在系统管理端对订餐业务按用餐日设立“启停”操作，代替原方式中对各在线 Excel 文件的共享与取消共享操作，控制该用餐日的数据能否填报。在当日用餐数据不再存在预订、退餐等修改时，管理人员可停掉该用餐日填报业务，对数据定版，以防止科室误填。待管理人员开启新一日

业务后，科室可以进入系统填报相应日期订单。

4.2.4 统计报表

从不同维度设计数据统计模块，包括职工及患者各自的订餐退餐流水（即所有原始填报记录）、最终订餐数据（订退餐合并统计后的结果）等表格，并可按日期、人员、科室、配餐地点等统计汇总，并支持导出 Excel 格式用于记账。

4.3 基于低代码平台的迭代开发

根据选用的产品不同，其内部开发步骤、说法略有差异。本项目选用某国产低代码平台以云服务方式进行开发。

（1）划分业务页面

在工作台中，将系统创建为患者、职工、基础设置三大业务版块。在患者及职工的业务内，再创建就餐填报、退餐、预订启停、统计汇总等多个页面，作为子业务版块。基础设置中创建了部门设置、用户设置等页面。

（2）基于每个业务创建其窗口迁移流程图

在每个业务页面中创建其窗口迁移流程图，以约束每个窗口在不同条件点击不同按钮后应迁移的目标窗口。下图为患者订餐窗口迁移流程图。

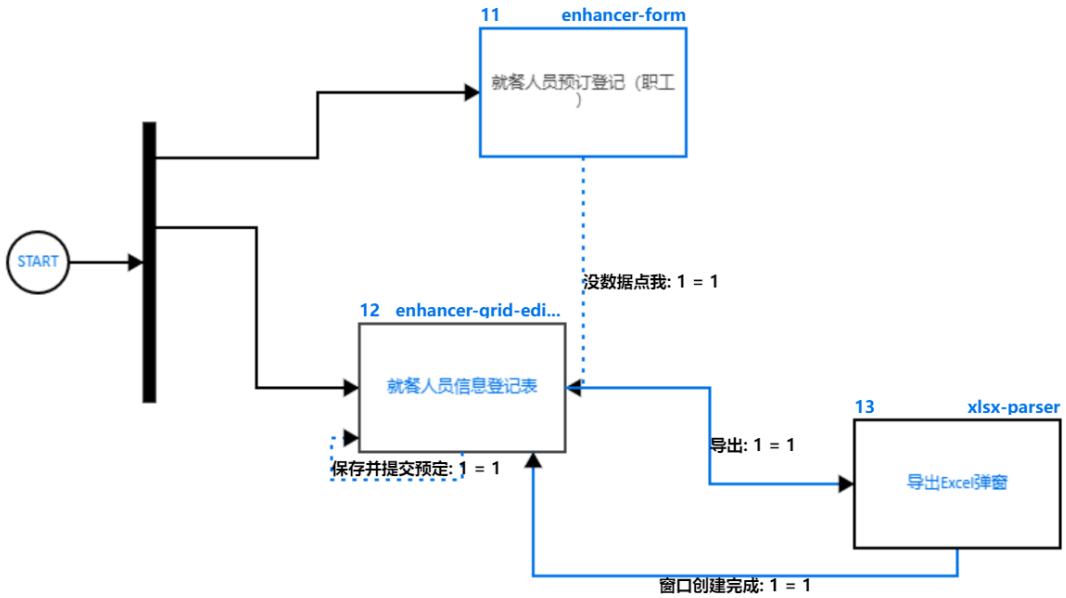


图 2 患者订餐窗口迁移流程图

（3）基于每个窗口进行组件设计

进入每一个窗口，拖拽相应组件进行界面设计。对组件设置对应事件、数据源后，组件可按后台配置的条件自动获取数据并刷新。同时，仅需对桌面端、移动端进行开关选择，就能让该窗口自动以手机端或 PC 端风格显示。见下图。

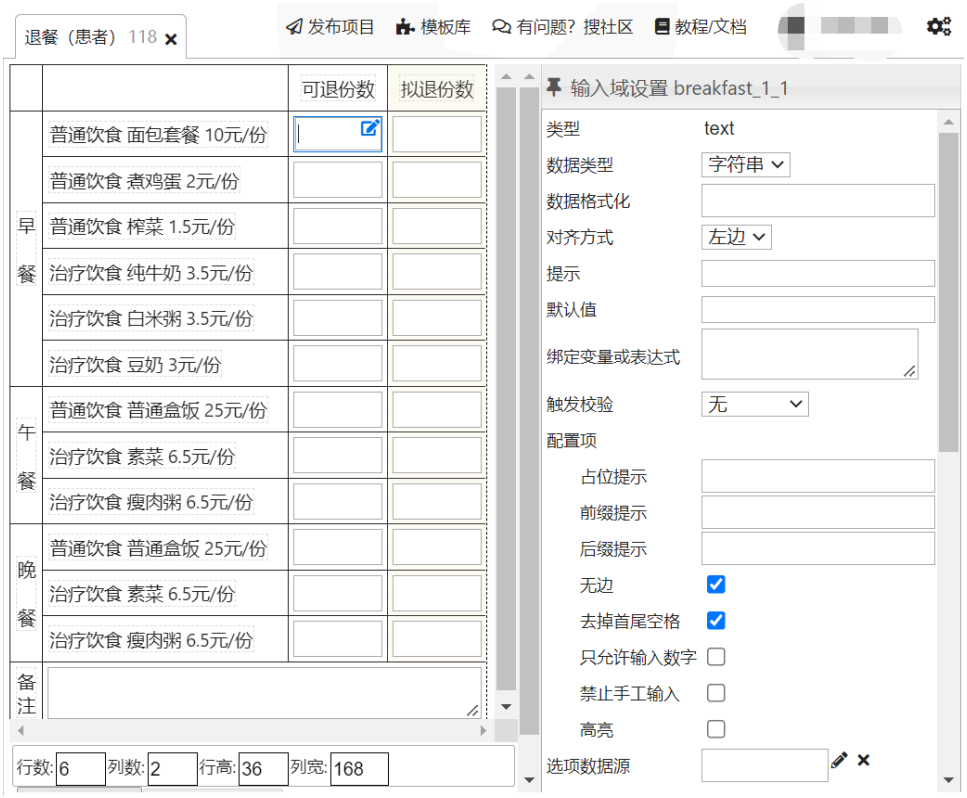


图 3 患者退餐窗口设计

4.4 迭代开发、一键发布

在研发过程中，需求沟通、新模块开发、已上线模块 Bug 修复、操作体验改善等工序根据需要混合进行。先完成基础设置与职工业务版块，再以职工业务为模板，开发患者业务，最终形成较为完善的系统，之后的主要工作为操作体验改进。低代码平台，覆盖应用开发从设计到发布的全生命周期，可以在开发后快速将应用程序发布到本地或云服务器。本系统选用低代码平台生态中的云服务，一键发布，系统立即投入使用。图 3 为系统部分界面。

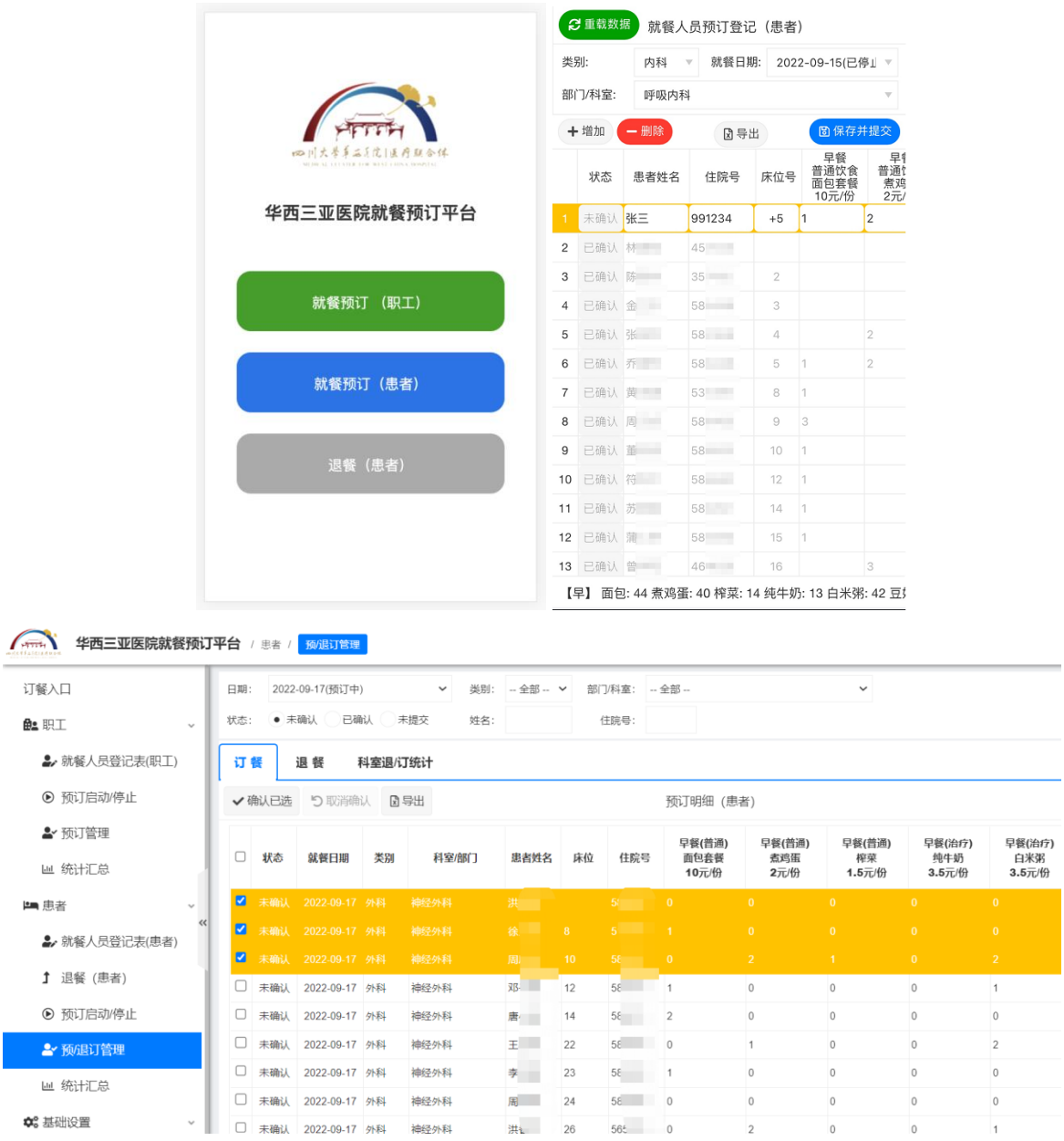


图 4 系统部分界面

5 建设成效

本项目共投入了 3 名项目人员，1 名膳食中心管理人员及 1 名信息中心主管利用碎片化时间沟通需求、1 名信息中心工程师全工作时间进行开发。系统分为基础设置、职工业务、患者业务三大版块，含 12 个页面、对应 12 个窗口迁移流程图，设计出 36 个界面联动窗口、对应 36 个功能点。

系统建成 2 个月内，共处理了约 6.4 万餐次的订餐。膳食中心每天数据统计工作由人工转为系统自动化，原本在统计上需要花费的时间从 2.5 小时缩短为约 30 分钟（主要为后台数据确认、对个别重复数据进行纠错及科室沟通等），将管理人员从重复易错的简单手工劳动中解出来，极大

提高了膳食中心工作效率，进而提高了医院后勤管理效率。得益于数据的准确性和可追溯性，因统计错误等带来的矛盾与投诉也大幅减少，患者和临床的满意度得到较大提升。

6 讨论

6.1 SCRUM 敏捷模式在医院运用的价值与适用场景

（1）用户满意度高

敏捷作为互联网产品开发的典型方法论，其迭代、递增式交付、允许试错、持续迭代中完善等思想，对于医院信息部门具有非常重要的指导意义。相比瀑布模型而言，敏捷思想使系统开发以不同于传统瀑布模型的串行工作，以小团队、多里程碑的冲刺，加快了交付，提高了用户满意度。

（2）加强项目管理

敏捷思想能加强医院信息化团队的项目管理意识。SCRUM 适合信息部门以几人小组为单位进行开发、管理，也有助于信息团队了解系统全生命周期，从而更加顺利地推进乙方项目。敏捷允许试错、在迭代中完善，通过阶段性交付，提前看效果，即能做到让好的项目提前成功，不好的项目尽早失败。

（3）适用场景

敏捷思想更适合在一些业务需求不明或变化大（尤其用户在初期都不知道需求）的中小型项目，克服资源有限、人员有限、时间有限的困难，做到快交付、短周期、低成本。

6.2 低代码开发工具在医院运用的价值与适用场景

6.2.1 提高了开发效率

通过对表单、界面、业务流程、逻辑规则等可视化建模，快速完成了业务系统的设计。由于平台具备功能强大的成熟组件库，具有二维化多行录入、联动计算、下拉列表条件过滤、多级联动过滤选择等复杂的输入控制功能组件，很好满足了数据录入和表格化显示需求。各组件还可自动适配移动终端。通过配置完成界面及功能设计，可一键部署，立即投入使用。

6.2.2 缩短了开发周期

该系统在低代码模式下，需求分析花费 2 人日，开发测试共花费 5 个人日，部署因采用了平台的自动发布功能未花费工时；全代码（即全部通过手工编码）模式下，在同一编程水平前提下，开发过程的人日通过专家经验法（FP、Function Point）进行估算，部署按 1 人日计算，如下表。

业务	页面		功能点	功能点数	全代码模式下估算人日	低代码模式下实际人日	全代码模式下估算人日（模块汇总）	低代码模式下实际人日（模块汇总）
基础设置	1	首页	页面迁移	1	0.5	0.1	3.5	0.5
	2	部门科室管理	组织机构树、查询、科室设置	3	2	0.2		
	3	用户管理	查询、增加用户、角色设置、密码修改	4	1	0.2		
职工业务	4	订餐（手机版）	预订登记、查询、导出	3	2	0.2	6	0.5
	5	预订启停	预订启停、增加新预订	2	1	0.1		
	6	预订管理	订餐明细	1	1	0.1		
	7	统计汇总	查询、按科室汇总、按职工汇总	3	2	0.1		
患者业务	8	订餐（手机版）	预订登记、查询、切换科室、导出	4	3	0.7	10.5	4
	9	退餐（手机版）	退餐登记、查询、表格化显示	3	2	1.2		
	10	预订启停	预订启停、增加新预订	2	0.5	0.1		
	11	预退订管理	订餐明细、退餐明细、订退餐量对比统计（早餐、午餐、晚餐）	5	2	1		
	12	统计汇总	查询、按科室汇总、按患者汇总、订餐流水、退餐流水	5	3	1		
总计							20	5

表 1 根据功能点的人日数估算对比

与全代码模式相比，除两者需求分析花费时间相同，低代码平台在开发、部署上花费时间均显著降低，如下图。

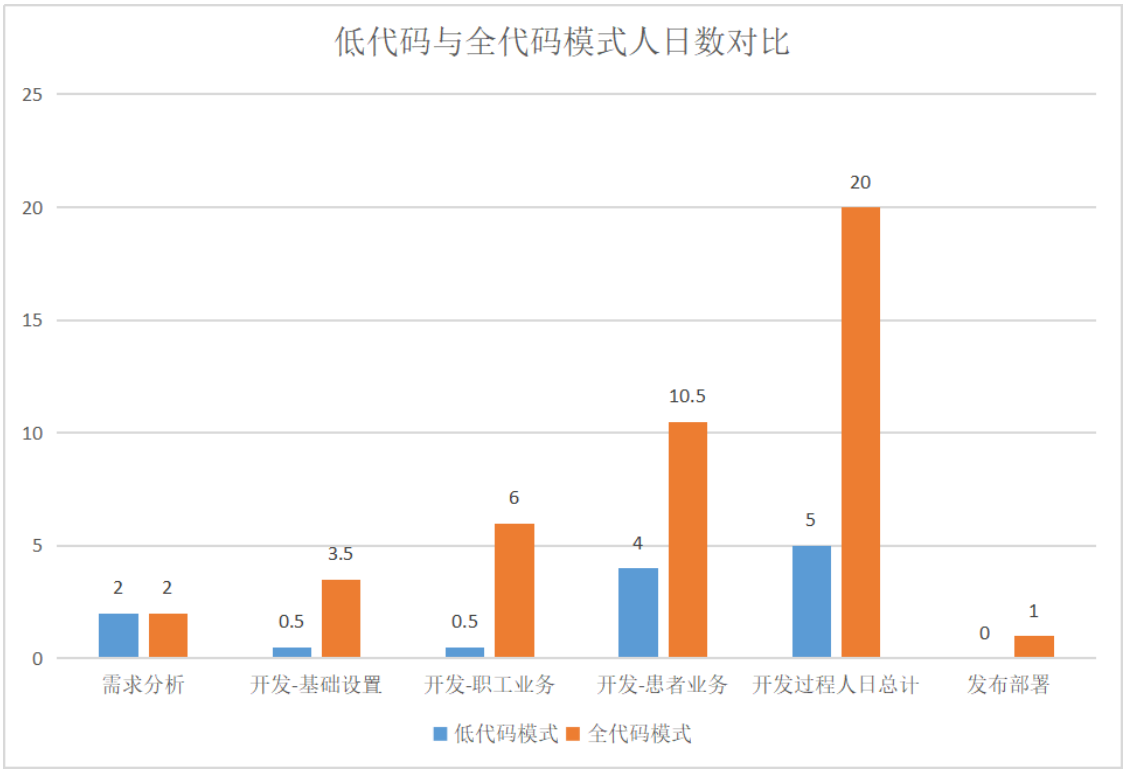


图 4 低代码与全代码模式下各软件过程工时对比

6.2.3 提高开发质量

得益于低代码平台大量成熟的组件和仅依靠配置实现“编程”的方式，系统具有相比传统全代

码开发方式更低的 Bug 发生率。

该系统在低代码模式下，运行过程中共发现 7 个 Bug。在全代码模式下，假设 FP:Bug（FP 与 Bug 数量比例）在团队为较高编程水平时为 3:1，即平均 3 个功能点出现 1 个 Bug 时，此时预测 Bug 数为 12 个；依此类推，对比发现与全代码模式相比，低代码平台所开发的系统 Bug 数量显著降低。如下图。

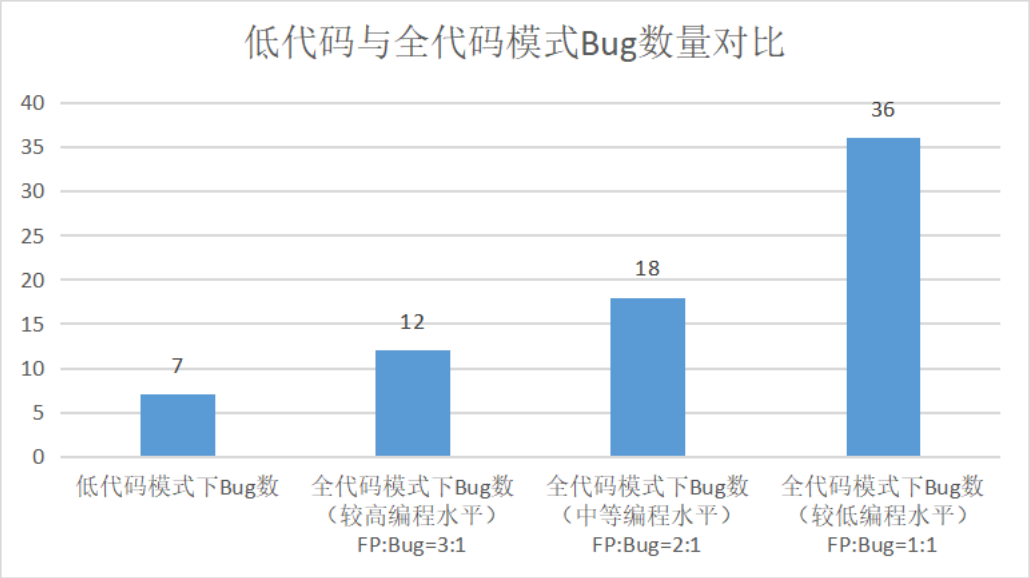


图 5 低代码与全代码模式下 Bug 数量对比

6.2.4 局限性及适用场景

在此次系统开发过程中，发现低代码平台亦有一些局限性，总结如下。

（1）低代码平台对技术有一定的要求

虽然在低代码平台中通过拖拽和配置即可完成系统界面与逻辑，但要把系统需求转化为配置性设置并熟练操作亦不是易事。尤其要实现一些较为复杂的数据计算、页面元素多控件联动效果时，需要用户具备一定的 IT 技术思维，部分甚至需编写少量代码才能实现。本项目所用低代码平台更适合信息技术部门运用，业务部门人员直接研发还有较大难度。

（2）低代码平台在调试上有一定的难度

由于低代码平台主要依靠高度集成化的组件，配合属性及事件设置实现功能开发，大多数逻辑并未以代码形式体现。系统开发过程中遇到 Bug 时，可能需进行断点调试以查找原因、或对某些逻辑进行全体调查、整体修改。在低代码平台中，若是开发层面的 Bug，进入各配置界面逐一修改，使得工作效率降低，此种方式也可能导致修改点的遗漏。若 Bug 是组件导致，代码开发环境的缺失可能导致比全代码模式花更多时间定位原因。

因此，低代码并不是一个能适用于所有人员或所有场景的万能工具。运用低代码平台进行开发

最高效的方式是采用业务人员提供需求、技术人员根据需求进行开发的合作模式。一个项目是否通过低代码平台开发，应结合系统功能、系统规模、后期扩展性等因素综合考虑。同时，低代码平台如何走出“技术人员看不上、业务人员学不会”的局面，降低编程门槛是一个焦点。易用性、易调试性、以及对快速修复平台组件缺陷的能力是低代码平台产品选型应考虑的因素。

7 结语

医院信息部门运用 SCRUM 敏捷模式及低代码工具进行开发和项目推进，是当前面对临床及外部需求快速变化的一种有效方式，可以做到以快速反应应对快速变化。深刻理解软件开发中的敏捷模式并正确地加以运用，熟练掌握低代码平台等快速开发工具，并对这些模式适用场景正确地判断与运用，是信息部门需要解决并不断探索的课题。

OR-009

Using the Baidu index to understand Chinese interest in common ocular diseases

JI Zhilin[1] YU Qi*[2]
1.School of Public Health, Shanxi Medical University
2.School of Management, Shanxi Medical University

China has the largest number of ocular diseases in the world, with more than 1 billion people suffering from refractive errors (about 700 million), dry eye (about 360 million), and cataracts (180 million), without considering account overlapping morbidity[1]. According to the World Health Organization, there are at least 2.2 billion blind and visually impaired people in the world, and at least 1 billion of them are due to untreated ocular diseases, so early detection and treatment are crucial for eye disease patients[2]. Common ocular diseases include refractive error, cataract, dry eye, macular diseases, glaucoma, and hordeolum (stye). Among them, refractive errors include myopia, hyperopia and astigmatism, to which almost half of all strength disorders are attributed. The cause of uncorrected astigmatism is reflected in a mismatch between the refractive power of the eye and the axial length of the eye, which results in a misalignment of the focus on the retina and the formation of an unclear image[3]. Cataract is a clouding of the clear lens in the eye that causes less

light to enter the eye, resulting in vision loss[4]. Dry eye disease (DED) is a multifactorial disorder, induced by decreased tear production from the lacrimal gland, increased tear evaporation, or poor tear quality[1]. Macular degeneration is a visual disorder caused by an abnormality in the retinal portion of the eye and is a major source of vision impairment[5]. Glaucoma is a degenerative optic neuropathy stemming from damage and progressive gradual loss of retinal ganglion cell (RGC) axons[6]. Hordeolum (stye) is an acute purulent localized swelling of the eyelid often associated with an eyelash follicle, Zeis gland, or Moll gland obstruction and infection[7]. Ocular diseases can affect people's daily work and life[8], and even lead to the occurrence of bad moods such as depression or anxiety, which can seriously affect the quality of life[9]. Because of the low treatment penetration of ophthalmic diseases and the significant economic burden on patients and society through direct treatment costs and indirect costs in terms of lost work productivity[10], there has been an increasing interest in research on the intervention and prevention of common ophthalmic diseases.

Internet data has been widely used as a new source of data for predicting and monitoring the prevalence of diseases. According to the 50th Statistical Report on the Development of the Internet in China released by the China Internet Network Information Center, as of June 2022, the Internet penetration rate reached 74.4%, and the number of Internet users reached 1.051 billion, of which the number of search engine Internet users reached 770 million, accounting for 77.8% of the total number of Internet users. In China, Baidu has become the dominant search engine, with 89.1% market coverage. Baidu as a commonly used search engine, has become the first choice for searching information, so the Baidu index contains a large number of health information search behaviors.

Internet search indexes can directly or indirectly reflect the behavior and psychology of Internet users, previous studies on socio-economic areas try to analyze search indexes and predict the connotative relationships between objects[11]. Thus, the Baidu search index can be used to investigate online health condition information search behavior and reflect the real-world health needs of Chinese Internet users. In addition, considering that screen time has been found to be an important risk factor for ocular diseases, we hypothesized that the population using Baidu to search for eye disease terms is highly overlapping with those with ocular diseases in mainland China. Therefore, the aim of this study was to conduct a secondary analysis of the Baidu index data to understand the level of public interest in ocular diseases and the trend of media attention. These results may inform the development of better interventions and treatments for patients with ocular diseases and explore the value of Internet search data in monitoring online information-seeking behavior.

2 Methods

2.1 Data sources

Baidu Index is a big data sharing platform constructed by massive user behavior information, which can show the search trend of keywords, insight into the change of user's demand, monitor the trend of media public opinion, and locate user characteristics. This study used data from Baidu Index[12], including Search Index (Based on the search volume of netizens in Baidu, using keywords as statistical objects, scientifically analyzing and calculating the weighted search frequency of each keyword in Baidu web search.) and Media Index (based on Baidu's intelligent distribution and recommended content data.) Meanwhile, this study obtains and calculates the annual netizen search rate (search index/ annual number of netizens) from the Statistical Report on Internet Development in that year.

The keywords "refractive error" "cataract" "dry eye" "macular diseases" "glaucoma" "hordeolum (stye)" were searched through the demand map of Baidu Index platform, and the weekly keyword demand map was collected in December 2022. Three words for non-ocular diseases were excluded: "what does refractive error mean" "dry eye symptoms" and "cataract symptoms". Search indexes and media indexes were obtained for each keyword from January 1, 2013 to December 1, 2022 for 10 full years, respectively. At the same time, due to the limitations of the Baidu index tool and the need for analysis, this study records the search index and media index data in the smallest unit of week, and aggregates them to yearly as the basis for subsequent data analysis.

2.2 Statistical methods

In order to understand the trend of public and media attention to common ocular diseases, this study firstly added up the search indexes of six keywords in each year to obtain the annual search index. Secondly, one-way variance (ANOVA) was used to analyze the differences between the annual search index and the annual media index for each keyword, and Pearson correlation analysis was used to analyze the correlation of the search indexes and year. Finally, after plotting the scatter plots and regression lines of the search rates of Internet users in each year, covariance analysis was performed to test the statistical differences of the slopes of the regression lines among each group. $P < 0.05$ (two-tailed) was considered statistically significant. Microsoft Office Excel 2019 and SPSS version 20.0 were used to draw figures, and all statistics analyses were performed with SPSS.

3 Results

3.1 Trends in the search index

Figure 1 showed that the sum of annual search indices for each keyword over the past 10 years shows an increasing trend and is positively correlated with the year. (Pearson's correlation = 0.808, P < 0.001). Keywords' search indices of refractive error (Pearson's correlation = 0.808, P < 0.001), dry eye (Pearson's correlation = 0.808, P < 0.001)、macular diseases (Pearson's correlation = 0.808, P < 0.001) and hordeolum (stye) (Pearson's correlation = 0.808, P < 0.001) were positively correlated with the year，while the results between cataracts (Pearson's correlation = 0.808, P < 0.001) and glaucoma (Pearson's correlation = 0.808, P < 0.001) with the year were not significant. In terms of search index growth and the growth rate of search index, the absolute increase of hordeolum (stye) search indices was the highest (2,234,728, 289%)，followed by dry eye (736,218, 169%). Table 1 shown the changes of each search index over the ten years and their correlation with years.

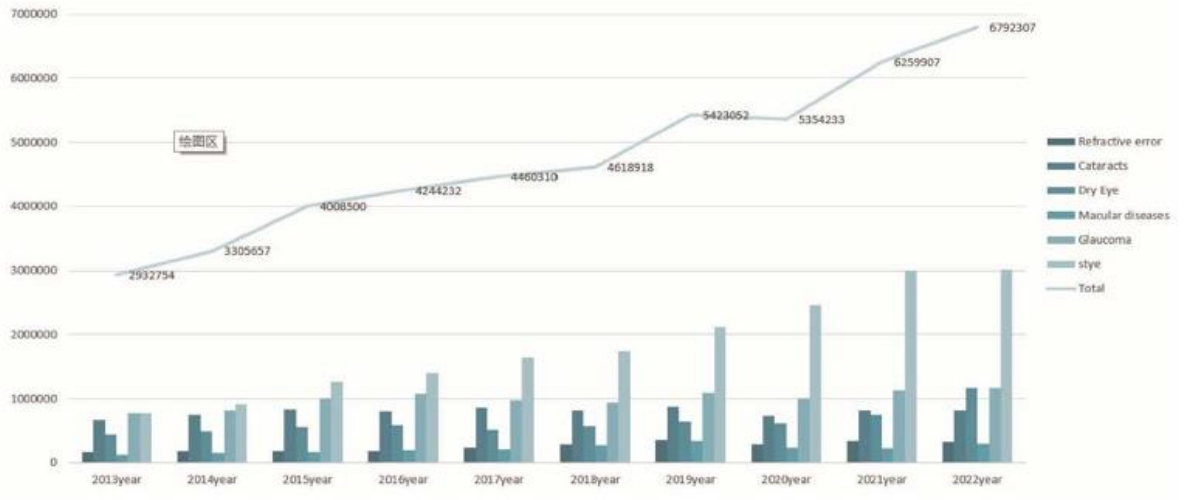


Figure 1 Changes in the annual search index.

Table 1 Basic situation and correlation analysis of search index from 2013 to 2022.

Keywords	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Increment	Increment rate(%)	Correlation coefficient	*P*
Refractive errors	162652	176778	183578	185557	232257	283483	352417	290314	344070	327031	164379	101	0.8	< 0.001
Cataracts	662415	751074	826581	809116	857531	816047	873216	743486	820026	818999	156584	24	−0.43	0.907
Dry Eye	435994	494184	560799	582266	523955	573937	647424	615021	751179	1172212	736218	169	0.801	< 0.001
Macular diseases	125238	151725	162224	193752	214374	273451	341513	236560	215911	299507	174269	139	0.817	< 0.001
Glaucoma	772327	816605	1004026	1076679	983515	933667	1085931	1008194	1123741	1165702	393375	51	−0.215	0.55
Stye	774128	915291	1271292	1396862	1648678	1738333	2122551	2460658	3004980	3008856	2234728	289	0.981	< 0.001
Total	2932754	3305657	4008500	4244232	4460310	4618918	5423052	5354233	6259907	6792307	3859553	131	0.808	< 0.001

Using the least-significant difference method, we found that there was a statistical difference between the search index of refractive error and cataracts, refractive error and glaucoma, refractive error and hordeolum (stye) ($P < 0.001$) and between refractive error and dry eye ($P < 0.001$). There was a statistical difference between the search index of cataracts and glaucoma, cataracts and hordeolum (stye) ($P < 0.001$) and between dry eye and macular diseases ($P = 0.009$), dry eye and hordeolum (stye) ($P < 0.001$). And there was a statistical difference between the search index of macular diseases and glaucoma ($P< 0.001$). However, there was no statistical difference between refractive error and dry eye ($P = 0.834$), and between cataracts and dry eye ($P = 0.297$). The search index for hordeolum (stye) was consistently higher than the other five ocular diseases, but in April 2022 the search index for dry eye surpassed that of hordeolum (stye). The specific results were shown in Table 2.

Table 2 Multiple comparisons between keywords (search index).

Keywords		*P*	95%CI	
Refractive error	Cataracts	< 0.001	-852832.69	-235238.11
	Dry Eye	0.016	-690680.69	-73086.11
	Macular diseases	0.834	-276409.09	341185.49
	Glaucoma	< 0.001	-1052022.29	-434427.71
	Stye	< 0.001	-1889146.49	-1271551.91
Cataracts	Dry Eye	0.297	-146645.29	470949.29
	Macular diseases	< 0.001	267626.31	885220.89
	Glaucoma	0.201	-507986.89	109607.69
	Stye	< 0.001	-1345111.09	-727516.51
Dry Eye	Macular diseases	0.009	105474.31	723068.89
	Glaucoma	0.023	-670138.89	-52544.31
	Stye	< 0.001	-1507263.09	-889668.51
Macular diseases	Glaucoma	< 0.001	-1084410.49	-466815.91
	Stye	< 0.001	-1921534.69	-1303940.11
Glaucoma	Stye	< 0.001	-1145921.49	-528326.91

Figure 2 showed that the search rate of Internet users has been on an upward trend in the past 10 years, and the regression slope of the six keywords was all greater than 0. The results of the covariance analysis showed that there was a statistically significant difference in the slope of the linear regression between the different groups ($F = 29.807$, $P < 0.001$).

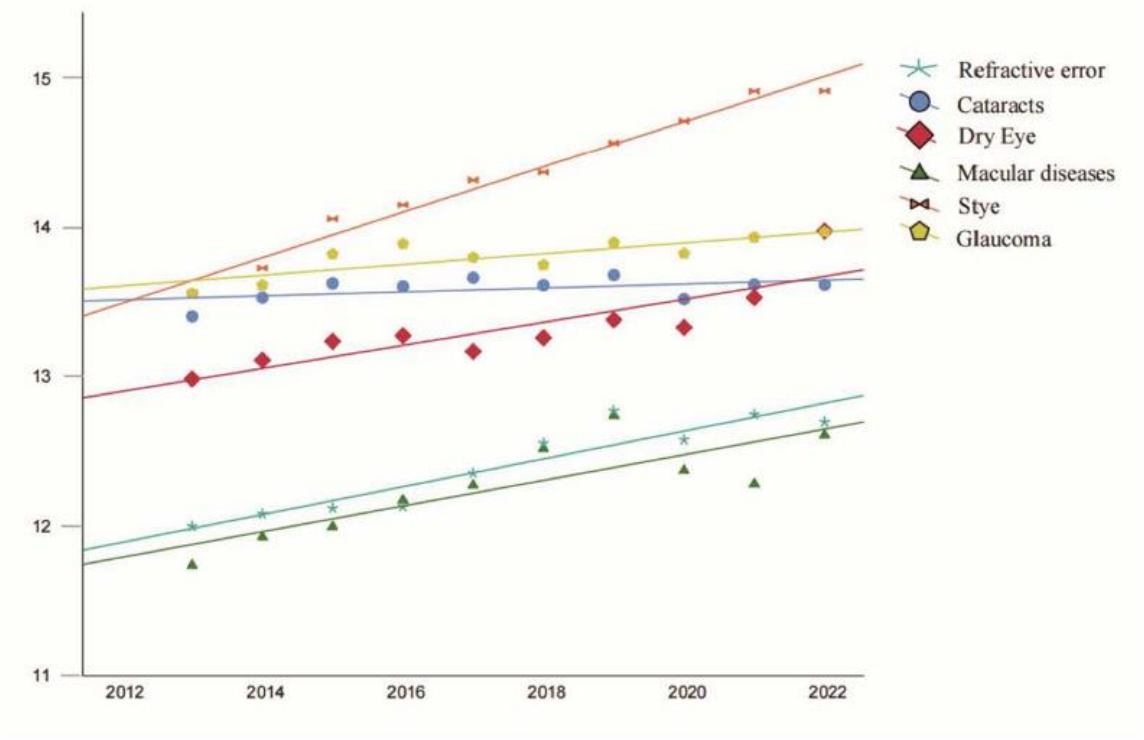


Figure 2 Scatter plot of annual netizen search rate.

3.2 Trends in the media index

Similar to the trend of the search index for keywords, the media index has shown an upward trend over the six years① (Pearson's correlation = 0.99, P < 0.001). Among them, the media index for dry eye rose the fastest and the Table 3 showed the changes in the media index for each keyword over the nine years. The media index of refractive error was statistically different from that of cataracts (P = 0.003), dry eye (P < 0.001). There was a statistical difference between the media index of cataracts and macular diseases (P = 0.021), macular diseases and hordeolum (stye) (P = 0.043). The media index of dry eye was statistically different from that of macular diseases (P = 0.002), glaucoma (P = 0.028), and hordeolum (stye) (P = 0.004). The relationship between other keywords were shown in Table 4.

Table 3 Basic situation of media index from 2017 to 2022.

Keywords	2017	2018	2019	2020	2021	2022
Refractive error	29210	139504	650068	231897	257279	198774
Cataracts	5492346	4274587	4650979	11813915	8385962	9556350
Dry Eye	276823	4780977	10631097	22500550	8464287	11440579
Macular diseases	27506	633663	311284	8353372	847670	1753210
Glaucoma	491267	3830055	5657900	8021241	3214837	6179009
Stye	178513	2493547	2704404	6945992	1856234	2002060
Media index	6495665	16152333	24605732	57866967	23026269	31129982

Table 4 Multiple comparisons between keywords (media index).

Keywords		*P*	95%CI	
Refractive error	Cataracts	0.003	-11623005.29	-2599463.71
	Dry Eye	< 0.001	-13943034.29	-4919492.71
	Macular diseases	0.438	-6248432.96	2775108.63
	Glaucoma	0.06	-8826366.96	197174.63
	Stye	0.277	-6957440.46	2066101.13
Cataracts	Dry Eye	0.302	-6831799.79	2191741.79
	Macular diseases	0.021	862801.54	9886343.13
	Glaucoma	0.215	-1715132.46	7308409.13
	Stye	0.043	153794.04	9177335.63
Dry Eye	Macular diseases	0.002	3182830.54	12206372.13
	Glaucoma	0.028	604896.54	9628438.13
	Stye	0.004	2473823.04	11497364.63
Macular diseases	Glaucoma	0.252	-7089704.79	1933836.79
	Stye	0.75	-5220778.29	3802763.29

4 Discussion

This study showed that public concern about ocular diseases has increased over the last 10 years, but there are differences between the concerns of different types of ocular diseases. Hordeolum (stye) and dry eye received significantly more attention than macular diseases, refractive error, glaucoma, and cataracts, with the growth rate of the search index for hordeolum (stye) more than twice that of the second place.

While all of the keywords showed an upward trend, the upward trend in the stye search index was more pronounced than for the other five keywords. This may be due to the fact that the pathogenesis of stye involves bacterial infections, hygiene problems, and other ocular diseases that result in bacterial invasion due to immunocompromise[7], and the etiology of the disease is closely related to living with the public. The web search data shows the behavior and psychology of Internet users[11], and Baidu as the hottest search tool in mainland China, covers a large number of public searching for information on health behaviors. Thus, the Baidu has been successfully applied in studies of COVID-19 and traditional infectious diseases outbreak surveillance and prediction, cancer investigations, and online attention to urological diseases, ocular diseases, and allergic diseases[13-15].

For the media index section, the results showed that the media pays more attention to dry eye than other keywords. This indicated that the media has pushed and reported more information about dry eye to the public in the past six years. Overall, the media attention of ocular diseases was on the rise. Dry eye can be caused by a variety of reasons, and the development of the disease is closely related to the external environment. The condition may be aggravated by bad eye habits, improper medication or diet, and improper eye cleaning[16]. Therefore, media attention to the dry eye could increase efforts to improve the dissemination of health information on prevention and treatment and improve patients' ability to self-manage. Ocular diseases pose a huge socio-economic and public health burden in China[3-6, 17], but fewer epidemiologic studies have been conducted on populations representative of the entire country. In particular, the Internet search data has rarely been used in studies related to

public eye problems in China. With the widespread use of the Internet and the improvement of citizens' health literacy, the public is increasingly relying on search engines for health information, and a number of online digital disease surveillance tools have been explored in recent years[13-15]. Search engines, as a query tool, can improve disease control by providing sensitive disease information prior to the diagnostic report, and the massive amount of data contained therein has a wide range of applications in the medical field, serving as a supplement and extension of existing clinical and epidemiological data.

5 Conclusion

The Internet search index and media index for six common ocular diseases have maintained continuous growth from 2013 to 2022. The Baidu Index can be used to track Chinese Internet users' online behavior and interest in ocular diseases. This can help to better identify and explore disease prevalence, patient literacy, and use of online services. Internet search trend data can be used as an exploratory tool to better understand patient characteristics and preferences, and provide a scientific basis for the prevention and treatment of ocular diseases among Chinese people.

This study also has several limitations. First, the Baidu index may have a sampling bias, as it only represents a sample of the population, which may be skewed towards higher socio-economic status and better educated segments. Second, we only focused on the attention of Baidu search engine users about the six common ocular diseases, without considering the public attention of other search engines or social media and the public's attention about more ocular diseases, which can only reflect some of the public attention about common ocular diseases. Therefore, to ensure data richness, future studies should explore including various search engines or social media for analysis. Simultaneously, the data should be thoroughly mined in order to control the impact of confounding factors on the study results and make the results more objective.

OR-010

基于机器学习的生物医学论文关联数据可重用性检测研究

周佳茵 段一凡 胡拯涌 唐明坤 钱庆* 吴思竹*
中国医学科学院医学信息研究所

研究背景

随着各种高通量组学技术的快速发展，全球每年产生的生物医学数据总量已达 EB 级，海量产出的科学数据，成为新时代科学研究、科技创新乃至国家战略的重要资源。由于新一代生物医学数据具有极端高通量、多来源性和异质性的问题，如何在海量数据中选择可信任的数据以及如何快速分析数据的有用性成为关键。研究调查显示，以生物学、医学为代表的自然学科领域科研人员数据重用往往是因为他们学科领域的数据具有一定的不可复制性[1-2]，若不进行数据重用，则会加大研究的难度，并且新获得的数据也很有可能受到外界质疑，重用科学数据可以规避一定的风险。

科学数据重用是指在科学研究过程中，使用原有数据对一个新的问题进行研究[3]。重用已有的科学数据不仅可以减少科研人员重复获取数据所消耗的时间、资源和精力，提高科研效率，还可以将科学数据再次应用到不同的研究目的中，使其得到更全面、更充分的利用，挖掘数据的潜在价值，提高数据的利用效率。然而，不同的数据可重用程度不同，例如在论文中展示的单一生化结果图与公共数据库中的大规模基因序列数据集在可用性和可重用程度上都有一定的差距，因此，进行可重用性的检测与分类具有较为实际的意义。

目前国内外对于数据可重用性的研究方向主要集中在对数据重用的动机、数据来源与获取方式、相关伦理与政策等方面，鲜有研究针对科学文献中的科学数据进行可重用性的检测，本研究提出使用自动化方法，抽取文献中的数据可用性声明，并使用随机森林、朴素贝叶斯、支持向量机、卷积神经网络、Bert 等机器学习方法实现在大规模生物医学文献中的科学数据可重用性检测。

研究目标

使用自动化方法，抽取文献中的数据可用性声明，并使用随机森林、朴素贝叶斯、支持向量机、卷积神经网络等机器学习方法实现在大规模生物医学文献中的科学数据可重用性检测。

研究方法

一、数据下载与抽取：

研究选取美国国立卫生研究院国家医学图书馆提供的免费的生物医学和生命科学期刊文献全文数据库 PubMed Central®（PMC），选取建库以来至 2022.9.13 的 802358 篇具有标注的 XML 格式的生物医学文献全文作为主要数据源。为避免单一数据库的偏倚性，本研究添加 53573 篇来源于欧洲 PMC 的全文文献（截至 2023.7.1）作为补充数据，去重后共得到 813191 篇全文文献，并使用自动化脚本抽取文献中的 DOI 号、标题、发表时间、数据可用性声明（Data Availability Statement，DAS）等信息。

二、DAS 数据分类与标注：

结合五大出版社的数据共享政策，将数据可用性程度分为 5 类：

1）数据完全不公开：例如：NONE reported

2）需通过联系作者或者向数据仓储提出申请获得：例如：Access requests can be submitted to Prof. J.M. Wardlaw.

3）数据已在文中或在附件中提供：例如：The data generated in this study are available within the article and its Supplementary data files.

4）数据部分包含在文中，部分公开/需要请求获取，或只公布数据存储机构：例如：All data and code compiled for this study are available on Github.（无 URL）

5）数据完全公开于公共数据仓储：例如：The Data for this study have been deposited in GEO with accession code: GSE 114234

通过人工标注的方式，将抽取出的 DAS 进行上述 5 类标签的标注。

三、DAS 数据预处理

1）删除与数据可用性无关、过长或只有标点的数据

2）删除停用词

3）使用 nltk 分词

4）使用 CountVectorizer()、Berttokenizer()等方法向量化

四、模型应用：

选用机器学习模型朴素贝叶斯（NB）、支持向量机（SVM）、逻辑回归（LR）、随机森林

（RF）、AdaBoost，KNN，深度学习模型 TextCNN、FastText、BioBert、BioBert-TextCNN，进行可重用性检测与分类。

研究结果

一、数据描述性分析

855892 篇文献中，仅有 14690 篇文献包含 DAS，大部分为 2016 年之后发表的文献，其中包含 DAS 最多的期刊是 Cell、nature 和 science，可见作为五大出版商旗下较为知名的期刊，其对数据共享要求较为严格。

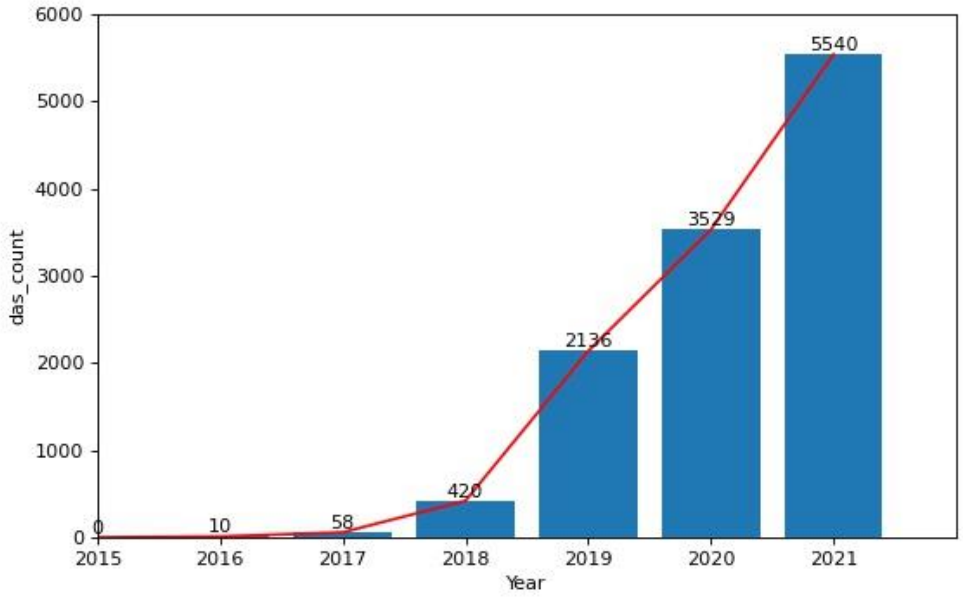


图 1 含有 DAS 论文发表年份统计

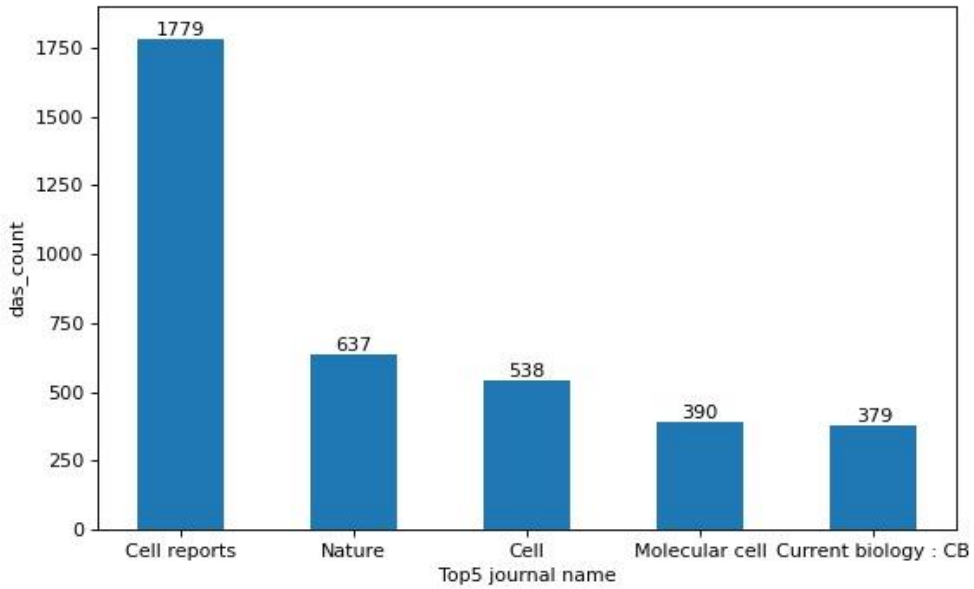


图 2 含有 DAS 论文所属期刊 TOP5

经过筛选和标注的数据量为 12471 条，其中 60%的数据为完全公开类，提示生物医学文献数据大部分存储在数据仓储平台，共享程度以及可重用行较高。

二、模型性能

模型结果显示，BioBert-TextCNN 的性能最好，达到了 0.881 的准确率和 0.884 的召回率，传统机器学习模型中，SVC 的性能最优，达到了 0.829 的准确率以及 0.813 的召回率且 SVC 的训练时间和成本远低于 BioBert-TextCNN。

研究结论

总体而言，生物医学文献中的科学数据可用性程度较高，超过一半的文献选用公共数据仓储存放研究数据，或是明确标注公开的数据集且提供 URL、登录号等数据唯一标识符。此外，本文提出的自动化流程，仅需标注数据可用性声明的开头和结尾即可自动提取以及分类，可以节省人力，实现大规模的文献数据可用性声明抽取与分类。最后，可用性检测不仅可以帮助期刊检测文献数据的公开程度还可以作为数据质量、数据影响力评价重要的量化指标。

OR-011

基于融合矩阵的文本相似度计算实现检索结果聚类

赵悦阳[1] 崔雷[2]
1.中国医科大学附属盛京医院图书馆
2. 中国医科大学医学健康管理学院

1 引言

文本聚类根据语义相似性将文本分组[1]，是网络索引、文本摘要、内容挖掘和信息检索等领域重要的研究手段[2, 3]。对 PubMed 检索结果进行文本聚类，能帮助科研人员迅速了解检索主题下涵盖的子方向，发现不同生物医学研究课题之间的潜在联系。针对 PubMed 检索结果的聚类研究，其核心是如何对高维且稀疏的医学文本实现语义表示，通过文本的相似性度量，实现医学文本的语义解析，提升聚类性能。

文本的语义表示对文本聚类效果的影响至关重要。在传统的聚类技术中，文档被表示为向量空间模型（Vector Space Model，VSM）[4]。向量是通过独热向量（One-Hot）形式或者词频-逆文档频率（Term Frequency-Inverse Document Frequency，TF-IDF）[5]模型表示文本。但是，由于特征词权重向量的特点，在训练时特征空间稀疏，特征的维数较高，导致模型性能偏低。与传统的文本表示方法相比，杰卡德（Jaccard）相似性算法[6]特别适合于应用于高维且稀疏度过高的数据[7]。

基于矩阵表示的语义相似性计算通过“词-上下文”矩阵动态地构建语义连接[8]。这种方法因其出现时间早、算法丰富，被广泛使用。上下文关系可以是词所在的文档、词与邻近词或目标词的关系，使用的上下文不同，则构成的矩阵稀疏程度也不同，其中“词-文档”矩阵最稀疏，计算效果也最差。

综上，本文基于 Jaccard 相似性算法，通过混合多种“词-上下文”矩阵，包括文档和词语之间，不同文档之间以及不同词语之间的连接，提出了融合矩阵的构建方法。通常来说，特征矩阵可以作为聚类分析的输入，该特征矩阵的每一行表示样本的特征值。为了达到更好的聚类效果，每个样本的特征值具有很好的区分性。

针对检索结果聚类的要求，本文满足了以下几点：

（1）用合适的模型表达文档分布。计算的融合矩阵不仅保留原始的邻近信息，还保留三个输入相似性矩阵所隐含的高阶关系。

（2）医学文本需要降维处理。基于矩阵表示的方法将每个节点向量化，生成一个低维表示，这种低维空间有利于对数据的进一步处理。

（3）为聚类生成标签。利用词-文档共现矩阵分别对文档包含的关键词进行计数并排序，为每个类簇生成聚类标签，方便用户了解每一类别的含义。

（4）在聚类之前，聚类的数目是未知的。为数据集确定一个合理的聚类个数很难，只能通过聚类算法来确定。

2 相关研究

文本聚类主要有两个研究方向：通过改进聚类算法提高聚类质量；通过提高文本的语义表示来提升聚类效果。

常用的文本聚类算法包括：基于划分的、基于密度的和基于层次的。K-means 是一种广泛使用的划分聚类算法，其中一个聚类中心与该聚类的其他数据点之间的距离平方和被最小化，以获得

给定数据集的最佳数据划分[9]。MiniBatch K-means [10]是标准 K-means 的变体，是处理大数据集的算法。Agglomerative 是一种自下而上的层次聚类方法[11]。BIRCH[12]也是一种流行的层次聚类。几种常用聚类算法的基本原理、主要应用、优势与不足等情况详见表 1。

表 1 常见的聚类算法比较

聚类技术	流行算法	基本原理	主要应用	优势	不足
层次聚类技术 Hierarchical clustering techniques	Agglomerative BIRCH[12]	通过使用启发式“自顶向下”拆分或“自底向上”合并技术生成聚类树（或树状图）	主题发现、员工/客户分组、软件聚类、绩效评价、主题演化、智能问答、舆情分析	定义距离比较容易、自由；不需要预先指定集群的数目；发现层次间的关系；可以发现任意形状的簇.	计算复杂度高，运行慢；对异常值敏感.
划分聚类技术 Partitional clustering techniques	K-means[9] Fuzzy C-Means Expectation-Maximization	将数据集划分为指定数目，通过对聚类中心的迭代重置，直到“聚类内部点足够近，类间点足够远”的目标效果，完成样本集的最优得分	主题划分、数据挖掘、模式识别、金融风控、主题发现	容易实施、灵活，运行快；可扩展性较好；对凸形簇效果好；容易解释.	算法依赖于数据；依赖初始条件；得不到全局最优解；需要预先指定集群的数量.
基于密度的聚类技术 Density-based clustering techniques	DBSCAN[13]	算法基于“簇”和“噪声”的直观概念，根据样本的紧密程度，把密集区域当做一个一个的聚类簇。	弹幕分析、热点话题分析、主题抽取、舆情分析、网络热点和媒体事件监测	不需要预先指定集群的数目；对噪声不敏感；擅长找到离群点.	算法参数复杂，对结果影响大；高维数据聚类有困难

文本的语义表示主要从表面文本相似度和语义相似度两方面进行研究[8]。表面文本相似度直接计算原始文本的字符串序列或字符组合的匹配程度或距离，来衡量相似度，其研究历史较长，且原理简单、易于实现。其中，N-Gram、Jaccard 系数、Dice 系数、Overlap 系数用来计算字符串序列或字符组合的匹配程度，比较样本集合之间的相似和差异。基于向量空间模型的方法则是通过空间距离上的相似度表达语义相似度。常用余弦距离、欧式距离将术语表示成向量。但是这种方法为了方便运算，只简单地将文本处理成向量，并不能含有语义信息。

Jaccard系数用于比较样本集之间的相似性和差异性，是计算机领域检验文本相似性的常用方法[14]。由于集合元素互不相似，用于文本相似度计算时不考虑词在文本中出现的频率。有学者基于词频比改进了Jaccard相似度算法[15]。医学文本的特征是不同主题之间的关键词基本不同，而同类主题的关键词相似度较大,决定了利用Jaccard系数来判断医学文本间的相似度是不错的选择[16]。

基于矩阵分布的模型是文本语义表示的另一个方向。它针对词语发生的上下文关联构建矩阵。上下文关系可以是词所在的文档、词与邻近词或目标词的关系，使用的上下文不同，则构成的矩阵稀疏程度也不同，其中“词-文档”矩阵最稀疏，计算效果也最差。矩阵里的数值除了使用词与上下文的共现次数表示以外，许多学者还会使用TF-IDF或取对数，对元素值实现加权或平滑。最后使用奇异值分解或非负矩阵分解等技术将原始的“词-上下文”矩阵从高维稀疏向量压缩为低维稠密向量[17]。

综上，本文提出利用 Jaccard 系数和 TF-IDF 构建融合矩阵的方法，在保留原始的邻近信息同时，还保留三个输入相似性矩阵所隐含的高阶关系，以解决在医学文本语义表示方面的不足，在对 PubMed 检索结果进行聚类分析时，提升聚类质量。

3 研究框架与方法

首先，使用 KeyBERT[18]抽取样本的重要关键词，给数据降维。然后，构建融合矩阵，分别训练 K-means，MiniBatch K-Means，Agglomerative 和 BIRCH 聚类算法，并以这四种算法作为基础训练集成矩阵（Stack）。对 PubMed 检索结果文献集做聚类分组，生成每一类别的特征词语，随后评价聚类结果和类别标签。研究框架见图 1。

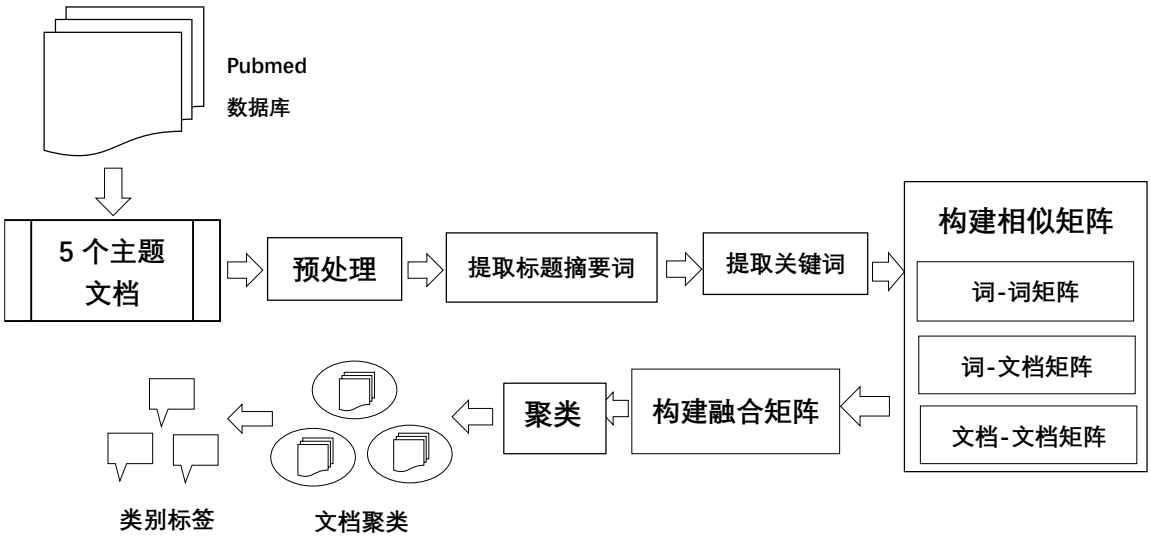


图 1　研究框架

3.1 数据获取

任意以 5 个 MeSH 主题词检索 PubMed 数据库，选取的主题词是：Aneurysm, Infected；Asthenopia；Glomerulonephritis, Membranoproliferative；Pelvic Floor Disorders；Restless Legs Syndrome。分别下载检索结果的标题和摘要, 并将其混合，作为数据集。

3.2 抽取文档核心词，生成 topN 关键词

应用预训练的 KeyBERT 模型抽取样本的 topN 关键词。KeyBERT 是非常基本但功能强大的关键词提取方法。它使用 BERT 嵌入模型提取词/短语，使用简单余弦相似度查找与文档本身最相似的关键词短语。可免除去停用词的步骤[18]。

3.3 构建相似性融合矩阵

- 构建词-词相似度矩阵

设关键词列表为$W = \{w_1, \dots, w_n\}$，文档集合为$D = \{d_1, \dots, d_m\}$。利用 Jaccard 算法计算词和词的相似性。首先计算词语w 的 word-in-doc 列表$L_d^w = \{L_{d_1}^{w_1}, \dots, L_{d_j}^{w_i}, \dots, L_{d_m}^{w_n}\}$，$L_{d_j}^{w_i} = idx_{d_j}$表示包含词

w_i的文档d_j的索引值为idx_{d_j}。利用 Jaccard 公式：$J(A,B)=\frac{|A\cap B|}{|A\cup B|}$，将$L_d^w$作为输入，计算两词之间的相似性分数。最终生成词-词相似性矩阵W_{n*n}，其中$W_{ij}=J(L_d^{w_i},L_d^{w_j})$。

- 构建词-文档相似性矩阵

词-文档相似性矩阵可以衡量词语和文档的相似性，利用 tf-idf 算法计算词与文档的相似性。首先根据上一步得到的L_d^w构建词-文档 0-1 共现矩阵occ_WD_{n*m}，其中$occ_WD_{i*j}=1$表示词语w_i存在文档d_j中。对于$occ_WD_{i*j}=1$的位置，根据 tf-idf 公式：

$$TF-IDF(w_i,d_j)=tf(w_i,d_j)*idf(w_i,d_j) \quad (1)$$

$$tf(w_i,d_j)=\frac{n_{w_i,d_j}}{\sum_k n_{k,d_j}} \quad (2)$$

$$idf(w_i,d_j)=\log\frac{|D|}{1+|D_{w_i}|} \quad (3)$$

其中，n_{w_i,d_j}指词w_i在文档d_j中出现的次数，$\sum_k n_{w_k,d_j}$指文档d_j中所有关键词的数量，$|D|$指文档总数，$|D_{w_i}|$指包含词w_i的文档数量。

通过公式(1)计算词语和文档的相似性分数，生成一个词-文档相似性矩阵tWD_{n*m}，其中当$occ_WD_{i*j}=1$时，$tWD_{i*j}=TF-IDF(w_i,d_j)$；当$occ_WD_{i*j}=0$时，$tWD_{i*j}=0$。

- 构建文档-文档相似性矩阵

利用 Jaccard 算法构建文档-文档相似性矩阵，与词-词相似性矩阵类似。首先，计算文档d的 doc-has-word 列表$L_w^d=\{L_{w_1}^{d_1},\dots,L_{w_i}^{d_j},\dots,L_{w_n}^{d_m}\}$，其中$L_{w_i}^{d_j}=idx_{w_i}$，表示存在于文档$d_j$的词语$w_i$的索引值。将$L_w^d$作为 Jaccard 公式的输入，计算两文档间的相似性分数。最终生成一个文档-文档相似性矩阵D_{m*m}，其中$D_{ij}=J(L_w^{d_i},L_w^{d_j})$。

- 构建融合矩阵

根据前面得到的tWD_{n*m}和D_{m*m}矩阵，生成融合矩阵MD_{m*m}。定义如下：

$$MD_{ij}=simi(d_i,d_j)=D_{ij}+\frac{1}{n}\sum_{k=1}^{n}tWD_{ki}*tWD_{kj} \quad (4)$$

其中，tWD_{ki}表示第 k 个关键词与第 i 个文档的相似性分数，tWD_{kj}表示第 k 个关键词与第 j 个文档的相似性分数，n 表示关键词数量。$\frac{1}{n}\sum_{k=1}^{n}tWD_{ki}*tWD_{kj}$可理解为文档 i 和 j 之间的相似性增益分数。

融合矩阵公式揭示了：两个文档可以通过它们共有的关键词与文档自身的相似性分数的求和平均来提升它们的相似性增益。求和后的平均是为了防止某些文档的相似性增益过大而进行的归一化处理。

3.4 训练集成聚类模型

根据MD相似矩阵分别训练 K-means，MiniBatch K-Means，Agglomerative 和 BIRCH 聚类算

法，将结果作为特征，利用 K-means 训练。具体步骤为：

训练四种聚类以后，每个聚类结果为一个相似性矩阵Res_M；

利用聚类算法 k-means 对若干个Res_M矩阵分别进行加权聚类，然后将这些聚类结果进行平均得到一致相似度矩阵Con_M；

最后基于Con_M矩阵，用 K-means 方法获得聚类集成结果。

3.5 聚类评价

使用聚类纯度、F1 分数和调整兰德系数评价聚类效果：

聚类纯度将聚类的正确样本数除以样本总数，也称为聚类准确度，类似于分类任务中的准确率。公式如下：

$$P=(\Omega,\mathbb{C})=\frac{1}{N}\sum_{k}\max_{j}|w_k\cap c_j| \tag{5}$$

F1 分数将聚类视为决策过程，当且仅当两个文档相似时，分组到同一聚类中。

Positive：TP 将两个相似的文档分组为一个簇（相同-相同）

TN 将两个不同的文档放入不同的集群（不同-不同）

Negative：FP 将两个不同的文档分组到同一个集群中（不同-相同）

FN 将两个相似的文档分组到不同的集群中（相同-不同）

精确率$P=\frac{TP}{TP+FP}$

召回率$R=\frac{TP}{TP+FN}$

F1 分数是精确率 P 和召回率 R 的调和平均值，可以准确地评价聚类算法的性能。

F1 分数定义为：F1=2 (P×R)/(P+R)

调整兰德系数（ARI）在 1985 年提出，假设模型是随机分布的，每一类和类簇上的节点数目是固定的。计算公式如下：

$$ARI=\frac{2\times(TP\cdot TN-FN\cdot FP)}{(TP+FN)(FN+TN)+(TP+FP)(FP+TN)} \tag{6}$$

3.6 生成聚类特征词语

聚类以后输出聚类标签$C=\{c_1,\dots c_i,\dots,c_m\}$，其中$c_i\in\{0,\dots,k\}$，$k$表示聚类的簇数量。针对每个簇，可利用词-文档共现矩阵occ_WD_{n*m}，分别对文档包含的关键词进行计数并排序，取前n个关键词，作为每个聚类簇生成的特征词语。

然后从这n个词中，以高频前 5 的词作为聚类标签词语。

本次实验通过 Python 3 编程来实现。

4 实验结果与分析

数据集包括 5 种类别和 7682 个文档。在聚类阶段选择聚类数目有很多标准[19]，需根据研究目

标来确定。本次研究将聚类数目设定为 2-10。表 2 是在设定不同聚类数目前提下表现最好的降维方法和算法。可看出，聚类数为 5 时，Agglomerative 算法的聚类表现最好。聚类数从 2 到 5 递增时，聚类表现越来越好，而从 6 到 10 则越来越差。表 3 是在聚类数目为 5 时不同聚类算法的聚类效果。可以看到，t-SNE 降维后 Agglomerative 的聚类结果最好。

表 2　不同聚类数目设定下表现好的聚类算法

聚类算法	降维方法	聚类个数	purity	ari	F1
Kmeans	t-SNE	2	0.63	0.476	0.657
Agglomerative	t-SNE	3	0.84	0.792	0.853
Agglomerative	t-SNE	4	0.924	0.939	0.955
Agglomerative	t-SNE	5	0.983	0.972	0.979
Agglomerative	t-SNE	6	0.983	0.818	0.859
Agglomerative	t-SNE	7	0.983	0.675	0.739
Agglomerative	t-SNE	8	0.983	0.602	0.672
Agglomerative	t-SNE	9	0.983	0.542	0.617
Agglomerative	t-SNE	10	0.983	0.479	0.574

表 3 基于融合矩阵和共现矩阵的聚类算法结果(k=5)

矩阵构建	聚类算法	降维方法	聚类数目	purity	ari	F1
融合矩阵	Kmeans	PCA	5	0.889	0.855	0.891
		t-SNE	5	0.908	0.824	0.867
	MiniBatchKMeans	PCA	5	0.9	0.761	0.818
		t-SNE	5	0.917	0.789	0.84
	Agglomerative	PCA	5	0.885	0.866	0.899
		t-SNE	5	0.983	0.972	0.979
	Birch	PCA	5	0.907	0.881	0.911
		t-SNE	5	0.927	0.834	0.874
	Stack	t-SNE	5	0.751	0.548	0.658
共现矩阵	K-means	PCA	5	0.889	0.855	0.89
		t-SNE	5	0.907	0.823	0.866
	MiniBatch K-Means	PCA	5	0.899	0.761	0.817
		t-SNE	5	0.919	0.794	0.844
	Agglomerative	PCA	5	0.917	0.902	0.927
		t-SNE	5	0.926	0.782	0.834
	Birch	PCA	5	0.903	0.779	0.832
		t-SNE	5	0.926	0.782	0.834
	Stack	t-SNE	5	0.79	0.532	0.638

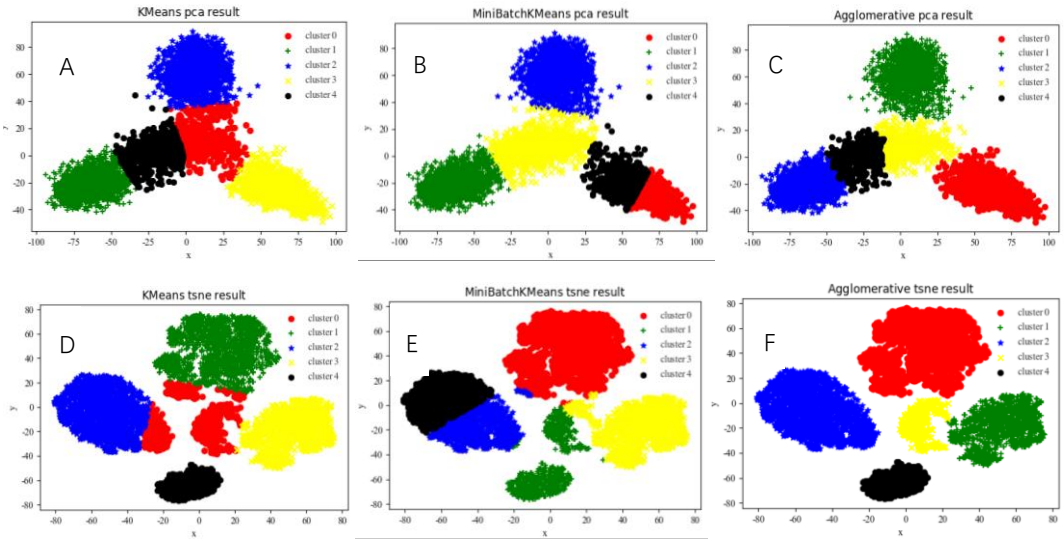
将基于融合矩阵与共现矩阵的结果比较（表 3），Agglomerative 算法基于 PCA 降维后共现矩阵效果好，但是 t-SNE 降维后共现矩阵效果远远不如融合矩阵。Birch 算法同样如此。基于融合矩阵的集成算法（Stack）综合看略好于共现矩阵的结果。但是集成算法的结果并没有单独一种聚类

算法有优势。说明融合矩阵的设计可以免去集成聚类的繁琐步骤，又能提高聚类效能。

从聚类效果图（图 2）可直观看出经 t-SNE 降维后的 Agglomerative 算法（图 2F），5 个类别的颜色没有重叠或分布不均，类内的点足够近，类间的点足够远，证明文献集得到了很好的分组。

将聚类后类簇数量与原始文档数做比较，簇 1 和簇 2 的数量与原始文档接近；而簇 3，簇 4 和簇 0 可看出基于融合矩阵的数量与原始文档数更接近，基于共现矩阵的簇 4 和簇 0 则相差很多，见表 4。

本文开展了预实验，分别使用 Cosine、Dice 和 Jaccard 相似性构建融合矩阵，结果显示，使用 Jaccard 训练 Agglomerative 算法的结果最好。2016 年，Mu 等人[20]构建了 CEDL 框架，将词-词，词-文档和文档-文档矩阵共同嵌入到框架内，同时完成文档聚类和标签生成。这种构建嵌入框架的方式是通过计算余弦相似度。本文构建的融合矩阵是通过 Jaccard 计算的文本相似度，对于有重复词语的文本，应用余弦计算的话，这类词的相似度会发生变化，而 Jaccard 计算的重复词的相似度则不变[21]。这是 Jaccard 的优势，而且在本次实验中也证实了，比共现矩阵得到的聚类效果好。



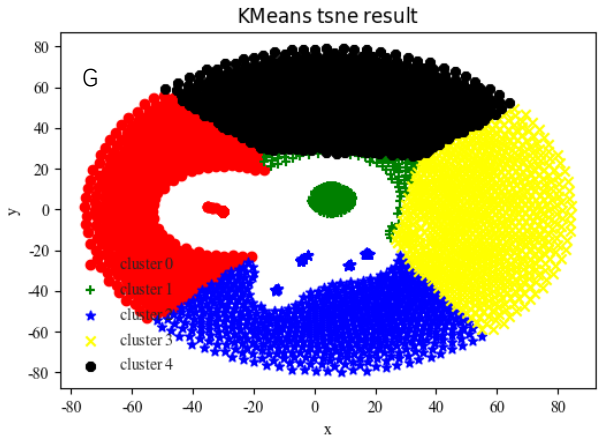


图 2 基于不同降维算法的聚类算法效果图

（注：A-C 是 PCA 降维后聚类算法效果图；D-F 是 t-SNE 降维后聚类算法效果图；G 是 t-SNE 降维后 stack 集成聚类算法效果图）

表 4 Agglomerative 算法聚类后类簇数量与原始文档数比较

从PubMed数据库下载的每个主题的文档数						
文档名	Aneurysm_Infected.txt	Asthenopia.txt	Glomerulonephritis_Membranoproliferative.txt	Pelvic_Floor_Disorders.txt	Restless_Legs_Syndrome.txt	总数
数量	2308	470	1638	697	2569	7682
基于不同的矩阵特征，使用Agglomerative算法聚类后各个簇的文档数						
簇ID	2	3	1	4	0	总数
基于共现矩阵	2275	681	1601	1561	1564	7682
基于融合矩阵	2275	583	1601	681	2542	7682

将融合矩阵生成的每类 top30 词语归纳。通过 Agglomerative 算法生成的特征词只出现在自己类里，其他三类中没有出现，充分满足聚类标签表现类别含义要求。而其它三种聚类得到的特征词，都存在高频词在其它类中出现的情况，比如 aneurysm，aortic。

对于每个簇，从 top30 里按出现频次由高到低，选择前 5 个短语作为标签词语，并与下载文献时选取的 MeSH 主题词比较，见表 5。Asthenopia 是视疲劳，标签词语是：visual，symptom，work，symptoms 和 ocular，组合起来是眼睛工作带来的症状，与视疲劳接近；Glomerulonephritis_Membranoproliferative 是肾小球肾炎_膜增生，标签词语是：nephritis，life，proliferative，deposit 和 renal，说的是肾脏某些物质增殖产生沉淀导致肾炎对生命有影响，也有增生肾炎的意思。其它 3 类标签词语基本与主题词含义相同。

综上，基于融合矩阵的聚类效果整体比基于文档共现矩阵的好，文献集得到了很好分组；基于融合矩阵训练集成聚类时，聚类效能反而下降，说明融合矩阵的设计可免去集成聚类的繁琐步骤，用简单的 Agglomerative 算法就能提高聚类效能；同时提取的描述类别的高频词也能很好的区别类别含义，所以融合矩阵的设计对于学术文本聚类描述的任务是有效的。

本文仍存在值得探讨的地方，数据集仅使用 PubMed 数据库，应该将融合矩阵的构建扩展到更多类型的文本资料库，验证聚类效果。

本文提取的聚类描述关键词的语义粒度较粗，只能反映查询结果内主题间的关系。根据细粒度文本表示概念[22]，开展知识单元细粒度层面的分析，例如论文中的研究范畴、方法、数据、指标、指标值等信息提取，为医学数据的智能分析提供重要的方法支撑。

表 5 标签词语与 MeSH 主题词比较

Cluster number	Label words（5个）	MeSH（1个）
2	aneurysm；case；aortic；report；aorta	Aneurysm_Infected
3	visual；symptom；work；symptoms；ocular	Asthenopia
1	nephritis；life；proliferative；deposit；renal	Glomerulonephritis_Membranoproliferative
4	floor；pelvic；pelvic floor；women；disorder	Pelvic_Floor_Disorders
0	syndrome；legs；restless leg；restless；restless legs	Restless_Legs_Syndrome

5 结语

本文探究了基于构建融合矩阵的文本相似度计算实现文本聚类的方法，主要探讨了四个问题：①如何构建有效的融合矩阵；②基于融合矩阵的聚类方法效果是否有所提升；③提出基于融合矩阵集成聚类方法，探讨其聚类效果；④生成聚类标签，描述聚类结果。

基于融合矩阵的聚类很好地提升了聚类效果，清晰地将文档集归类，聚类的同时提取的描述类别的高频词也能很好的区别类别含义，所以融合矩阵的设计对于学术文本聚类描述的任务是有效的。

今后，通过本文的研究方法可以将 PubMed 的检索结果聚类，通过提供关键词简明扼要地描述检索结果，反映特定查询下的内容分布。在此基础之上，通过细粒度分析，从文本中识别出研究范畴、研究方法、实验数据和评价指标及取值等知识，作为新型医学图书馆提供的新项目，面向知识层次的精细化服务，提升图书馆的服务水平。

OR-012

中国医学卫生健康自主知识体系：创新视域和建设路径

李晓瑛、任慧玲*、李军莲、李亚子、刘懿、冀玉静、邓盼盼、夏光辉、刘辉*
中国医学科学院医学信息研究所

1. 引言

2021 年 5 月，习近平总书记要求：“完善国家创新体系，加快建设科技强国，实现高水平科技自立自强。”党的二十大报告已将“实现高水平科技自立自强”纳入 2035 年我国发展的总体目标。科技创新是大国博弈的主战场，加快科技创新是实现高质量发展的关键。科技创新的源头是基础研究，夯实基础研究是实现高水平科技自立自强的必然要求。因此，在当今新一轮科技革命和产业革命突飞猛进、科学研究范式深刻变革的新时期，要加强基础前沿探索和技术突破，努力构建中国特色、中国风格、中国气派、中国自主的学科体系、学术体系、话语体系、知识体系。

医学卫生健康事业关乎我国国家安全、经济发展、社会文明和人民福祉，其重要意义涵盖了“面向世界科技前沿、面向经济主战场、面向国家重大需求、面向人民生命健康”的核心内容[1]。构建中国医学卫生健康自主知识体系，凝练我国医学卫生健康实践中具有主体性、原创性的理论观点，既能够指导我国医学卫生健康工作、推进健康中国建设，又助于增强中国在国际社会的话语主动权、展现新时代的中国精神。作为我国最高医学研究机构和最高医学教育机构，中国医学科学院以引领我国医学科技教育发展和维护人民健康为己任，联合全国科学技术名词审定委员会，探索中国医学卫生健康自主知识体系框架和建设路径，为应有之义。

2. 创新视域

在“大医学、大卫生、大健康”的健康中国背景下，中国医学科学院院长王辰院士对医学、卫生、健康产生了一些新的想法和认识。医学是为恢复、维护、增强人的健康而发展出的知识、技术、艺术、学术体系[2]。医学是多学，自然科学和技术、社会科学和方法、人文学科和文化是医学的三大基础学科。医学是人学，从“促防诊控治康”（健康促进、预防、诊断、控制、治疗、康复）“六域“、 “语药械食居环”（语言、药物、刀械、饮食、起居、环境）“六宝”，施行对人、人群、人类的健康和生命的全面照护，依照护对象的不同分为临床医学、群医学两大类。卫生是依据医学及相关学科的原理，为恢复、维护、增强人的健康所采取的行动。公共卫生以群医学和相关学科原理为基础，为恢复、维护、增强人群健康所采取的综合社会行动[3,4]。健康是人的一种良好生命状态，包括生理健康、心理健康、人际健康及环境健康四个方面，关注从个体到群体、到种群、到人类、到众生、到生态的健康问题，应和了中国“天人合一”思想和“One Health”理念。医学卫生健康已成为重大、主流、核心的社会事业，肩负着国家、民族、人类的期望。

3. 框架体系建设研究

3.1 资料与方法

国内外已公布的学科目录、标准规范、术语体系等相关资源（表 1），为研究构建中国医学卫生健康自主知识体系框架提供了借鉴。人类系统化的知识主要以学科专业的形式出现，故而学科目录、学科分类代码等对学科专业的划分依据，可用于知识组织和知识体系构建。术语体系对某一领域主要概念及概念间语义关系进行概括梳理，已成为领域知识组织的重要工具。本研究在参考国内外具有较高影响力的学科目录、标准规范、术语体系，以及国际著名医学院校学系专业设置（如美国哈佛大学、约翰霍普金斯大学、斯坦福大学，中国北京大学、复旦大学、中南大学等）的基础上，以“医学为学科、卫生为行动、健康为目标”的总体思想为指导，采用比较综合法，经多次讨论修改完善，初步提出中国医学卫生健康自主知识体系框架。

表 1. 主要借鉴的已有国内外资源

资源分类	资源名称	资源概况
学科目录	美国学科专业分类目录 Classification of Instructional Programs, CIP	由美国教育部国家教育统计中心负责编制维护，一级学科共 50 个[5]。医学相关学科主要分布在“26. 生物学与生物医学科学”、“51. 健康职业与相关临床科学”、 “42. 心理学”、“30. 交叉学科”和“14. 工程”5 个学科群，共有 63 个二级学科 413 个专业。
	研究生教育学科专业目录（2022 年）	由中国国务院学位委员会、教育部发布，涉及 14 个学科门类，共 181 个一级学科[6]。其中，“10. 医学”涉及 17 个一级学科，与医学相关的交叉学科包括“0712. 科学技术史”、“0831 生物医学工程”。
标准规范	学科分类与代码 现行版本 GB/T 13745-2009	适用于我国宏观管理、科技统计的学科分类原则、学科分类依据与编码方法[7]，共有 62 个一级学科、676 个二级学科及 2382 个三级学科。其中，医学相关的学科主要集中在“310. 基础医学”、“320. 临床医学”、“330. 预防医学与卫生学”、“340. 军事医学与特种医学”及“350. 药学”。
术语体系	医学主题词表 Medical Subject Headings，MeSH	由美国国立医学图书馆编制，收录解剖学、有机体、疾病、化学物质和药物、技术与设备、心理学和精神病学、现象和过程等 16 个大类、3 万条主题词[8]。
	系统化临床医学术语集 Systematized	由国际卫生术语标准开发组织维护发布[9]，收录了临床信息常用的疾病、临床发现、

Nomenclature of Medicine -- Clinical Terms, SNOMED CT	解剖结构、有机体、物质、药品、物理对象、标本等 19 个大类、36 万条概念[10]，以及这些概念间 264 种语义关系类型[11]。
肿瘤学叙词表 NCI thesaurus，NCIt	由美国国立癌症研究所构建并维护，涵盖临床医学、转化医学、基础医学、公共卫生等专题 17 万条概念[12]，包括疾病、基因、有机体、药品等[13]。
美国国立医学图书馆分类法 National Library of Medicine Classification, NLMC	美国国立医学图书馆负责编制维护，采用列举式类目体系，学科类目包括“Q 临床前科学”和“W 医学相关学科”，分成 36 个大类，按字母顺序排列。字母后的数字从小到大排列，无层级结构[14]。以“QS 人体解剖学”为例，“QS1-132 解剖学”，“QS 504-532 组织学”，“QS 604-681 胚胎学”。
中文医学主题词表 Chinese Medical Subject Headings，CMeSH	中国医学科学院医学信息研究所负责编制维护，整合了中译版 MeSH 及《中国中医药学主题词表》，收录了 30 个大类的 3.8 万条主题词[15]。
中文临床医学术语体系	中国医学科学院医学信息研究所自主构建，覆盖疾病、解剖部位、临床表现、检查检验、手术操作、化学药品等 14 个维度、20 余万条临床重要概念[16]。
中国图书馆分类法·医学专业分类表	中国医学科学院医学信息研究所负责编制，以当代医学学科分类为依据，兼顾中国传统医学，由预防医学、医学和药学三部分组成，类目 5000 余条[17]。

3.2 初步建设的框架体系

初步建设的中国医学卫生健康自主知识体系框架包括医学、卫生、健康三部分（图 1），一级类目 8 个、二级类目 83 个。特别提出的是，该体系暂未涉及中医药学相关知识，后续将在中医药学术语发展基础上进一步扩展整合。

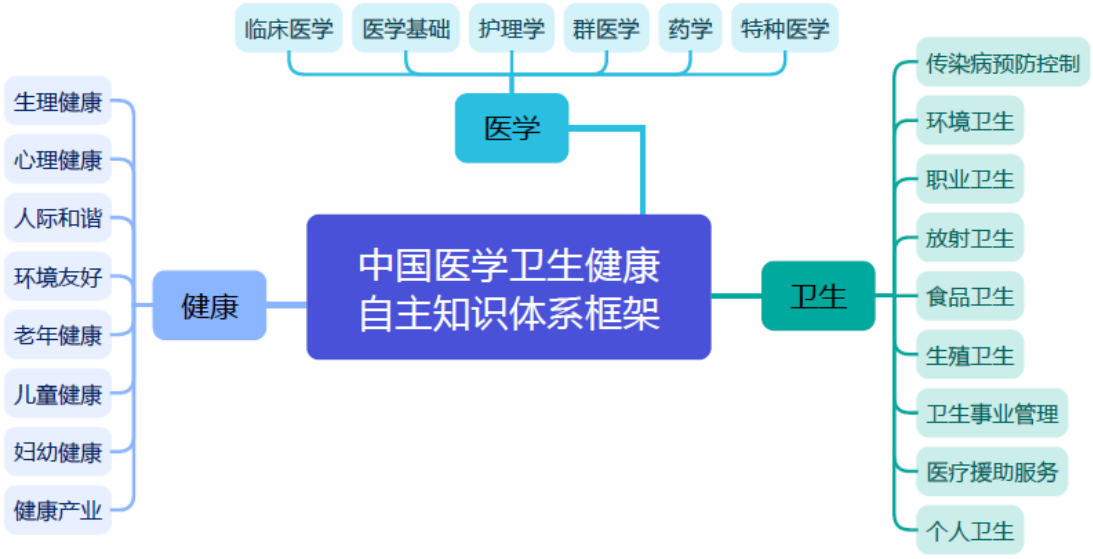


图 1. 中国医学卫生健康自主知识体系框架（初稿）

医学部分一级类目主要包括临床医学、医学基础、护理学、群医学、药学和特种医学。其中，临床医学侧重于描述临床常见疾病，具体又分为感染、肿瘤、肌骨骼疾病、消化系统疾病、口颌疾病、呼吸道疾病、耳鼻咽喉疾病、神经与精神疾病、眼疾病、泌尿生殖系统疾病、心血管疾病、血液和淋巴系统疾病、先天性遗传性新生儿疾病和畸形、营养和代谢性疾病、内分泌系统疾病、免疫系统疾病、环境因素诱发疾病、职业病、化学诱导疾病、创伤和损伤，以及人体的病理状态、体征和症状、医学技术与设备。医学是一个与众多学科密切相关的交叉学科，涉及人体形态学、人体生理学、生物化学、医学分子生物学、生物物理学、病理学、医学微生物学、医学寄生虫学、医学免疫学、医学遗传学、实验动物学；此外，还包括与医学相关的人类三大基础学科，即自然科学和技术、社会科学和方法、人文学科和文化。护理学涉及护理理论、护理技能以及护理伦理等，相关知识包涵基础护理、临床护理和社区护理。群医学是医学的新概念、观念、领域、学科，重点关注人群、种群、人类，追求群体、种群乃至人类和众生健康效益最大化，涵盖流行病学、营养学、社会医学、卫生检验学、卫生工程学、卫生经济学、卫生统计学、卫生监督学、健康教育与健康促进。药学亦是一门以药物为研究对象的综合学科，包括药物化学、药物分析、微生物与生化药学、生药学、药剂学、药理学、毒理学、药品和药事管理。特种医学研究在特殊环境条件下从业或从事其它活动的所有人群特有的卫生保健需求，解决在实践中涉及到的各种特殊医学问题，包括航空航天医学、航海潜水医学、高原医学、应激医学、军事医学、运动医学、放射医学和法医学。

第二部分的一级类目为卫生，二级类目主要包括传染病预防控制、环境卫生、职业卫生、放射卫生、食品卫生、生殖卫生、卫生事业管理、医疗援助服务和个人卫生。其中，传染病预防控制指研究特定人群中疾病、健康状况的分布及其决定因素，并研究防治疾病及促进健康的策略和措施的学科。环境卫生研究可能影响人类疾病的环境因素及环境相关疾病的检测、预防和控制方法进展，职业卫生关注职业环境下职员身心健康的提高和维护。放射卫生探究有关放射对环境影响以及对公众和个人健康影响，食品卫生指有关食品生产、加工、运输、储藏、销售、供给过程中影响人体健康的各种不利因素及其防控措施与方法的研究与应用。生殖卫生研究有关人类生殖的形态学、生理学、生物化学和病理学，以及生育和泌乳的生理学、医学问题。卫生事业管理指借助管理科学的理论知识和方法，对卫生事业进行研究和分析，揭示卫生管理活动的规律。医疗援助服务是指国家和社会针对那些因贫困而无经济能力进行治病的公民，所实施的专门帮助和支持。

第三部分的一级类目为健康，二级类目包括生理健康、心理健康、人际和谐、环境友好、老年健康、儿童健康、妇幼健康和健康产业。生理健康为人体生理功能上健康状态的总合，心理健康指心理的各个方面及活动过程处于一种良好或正常的状态，人际和谐是人与人在相互交往过程中所形成良好的情感联系，而环境友好侧重人与自然和谐共生的状态。此外，老年健康、儿童健康、妇幼健康分别关注老年人、儿童、母亲和婴儿的身体和心理的健康状况。健康产业是一类新兴产业，包括医疗产品、保健用品、营养食品、医疗器械、保健器具、休闲健身、健康管理、健康咨询等多个与人类健康紧密相关的生产和服务领域。知识体系中与健康相关的内容，将持续更新完善。

4. 建设路径

依托全国科学技术名词审定委员会规划组建的医学卫生健康科学名词审定工作委员会[18]，及其下设的临床医学、基础医学、公共卫生与预防医学、特种医学、护理学、药学等全国名词审定委员会，按中国医学卫生健康自主知识体系框架，同步开展相应专科医学名词术语的加工、审定等建设工作。建设过程中，参考《现代术语学引论》[19]，制定三轮审议原则。其中，一审会收集候选词并归纳，形成征求意见稿；二审会，协调统一名词概念；三审会，确定最终概念以及概念之间的关系。建成后，由全国科学技术名词审定委员会批准、医学卫生健康科学名词审定工作委员会发布工作成果。

此外，参考国家标准《GB/T 10112-2019:术语工作 原则与方法》等相关文献[20]，制定了中国医学卫生健康自主知识体系的选词和审定的基本原则。首先，名词用字应该遵守国家对于语言文字的有关规定。其次，贯彻“一个名词一个含义”的原则，具体指一个概念应确定一个与之相应的中

文标准名词；一个概念在不同学科中如果已分别有统一的名词，确实不宜强行统一为一个名词时，极个别的情况允许分别定名，在注释中以“又称”列出；为便于了解详细，建议提供参考注释，与已公布的名词概念相同的异名(包括港台名)，选择某些常用的在注释栏中分别冠以“又称”、“简称”、“曾用名”、“港台名”列出。最后，整个知识体系坚持“协调一致”的原则，努力做到名词的协调和统一。

中国医学科学院医学信息研究所在四十余年中文医学术语体系建设过程中，已自主建成医学术语加工平台（软著登字第 2486356 号）。在线平台可为知识体系建设提供系统信息技术支持，支持多人协同加工审定。此外，利用计算机技术与相关学科查重，在科技文献、电子病历、临床指南、网络信息等多源数据中进行语义计算、消歧、归一、和新词推荐等工作，为专家提供多源证据支持，提升自主知识体系建设效率和内容质量。

5. 结语

中国医学卫生健康自主知识体系建设意义重大、任重道远，亟需提出一条适合中国国情、科学可行的建设之路。中国医学科学院在借鉴国内外成果和经验的基础上，融合、创新、发展，初步提出中国医学卫生健康自主知识体系框架和建设路径。后续建设全过程，将集思广益、凝聚多方智慧展开论证，携手共同完成这一项基础性工作，推动我国医疗信息系统互联互通、健康大数据的共建共享与“健康中国”战略实施。

OR-013

基于博弈论组合赋权的医院科研能力云模型评价

王琰 杨爽 王雅晴 刘曙霞 安新颖
北京协和医学院•中国医学科学院医学信息研究所

0 引言

为我国医学领域科技评价等工作有机结合提供依据，实现医疗机构自身对科技研发全方位保障的强化[1]，使用科研项目和项目产出结果是医院学科建设和学术水平的量化体现，同时也是连接基础研究与临床转化的桥梁，是科技创新和人才培养的重要载体。对于医疗机构，尤其是医院科技评价有助于医院争取科研项目经费、完善研究基地与平台建设、促进科技创新人才培养、提升医院科研创新能力和综合实力。我国在医院科研能力评价主要评价方法主要是定性或半定性半定量方法，如王婷婷等人采用专家评价法、优序图法评价专科医院科研能力[6]，崔昌进采用访谈法对三甲医院进行评价[7]。专家评价法和优序图法虽在一定程度上解决了医院科研评价过程中的信息缺乏和定性指标模糊化问题，但未考虑评价过程中的随机性和不确定性。现阶段迫切需要建立适应新时期发展需要的，科学、合理、可行的医学领域科技评价机制和科技评价体系，让医学科技学术和科研工作评价回归科研本质，使科研管理符合科技发展规律和行业领域特点[8]。

为解决上述问题，本研究拟在使用现有医疗机构科技评价体系，对国家级科研院所、三级甲等医疗机构外等医疗机构，基于博弈论组合赋权和云综合评价模型建立不同层次的学术评价和科研评价机制[11]，已解决现有评价过程中语言中表达的不确定性、随机性与模糊性之间的关联关系的描述，完成定量指标与定性指标转化。

1 研究方法

1.1 基于博弈论综合赋权的医院科研能力评价指标权重确定

博弈论是运筹学的一种重要方法，广泛应用于多主体、多目标决策问题。基于博弈论的组合优化权重模型，在主观、客观权重中间找到均衡点，实现权重的优化组合，使权重的设定更加科学合理[1]。为了提高医院科研能力评价体系的准确性，本研究基于相关文献，进行方法优化并采用博弈论组合赋权法来确定各评价指标的综合权重。综合应用了层次分析法（AHP）、熵权 TOPSIS 法、CRITIC 法和前景理论模型（TODIM）来确定各评价指标的主观权重 $\omega 1$和客观权重$\omega 2$。假设通过 AHP 方法得到的主观权重记为$\varphi 1 = (\varphi 11, \varphi 12, \ldots, \varphi 1n)$，通过改进的 CRITIC 法、熵

权 TOPSIS 法、TODIM 得到的客观权重记为$\sigma 1 = (\sigma 11, \sigma 12, \ldots, \sigma 1n)$，确定各自的权重系数，由 φ1 和$\sigma 1$的线性组合计算出最终的综合权重（表 1）。

$$\omega = \alpha 1 \varphi T1 + \alpha 2 \sigma 2 \quad (1)$$

式中：$\alpha 1$为主观权重的系数；$\alpha 2$为客观权重的系数；$\varphi 1$为主观权重值；$\sigma 1$为客观权重值；ω为指标的权重。

表 1. 评价指标体系构建及权重确定

Table 1. Construction of Evaluation Index System and Determination of Weight

一级指标	二级指标	三级指标
科技产出	论文指标	SCIE 论文
		中文核心期刊论文
	专利指标	授权发明专利
		发明专利转化
		标准和指南
	引用	SCIE 收录论文引用
		中文核心期刊论文引用
		国际权威指南论文引用
学术影响	奖项	国家科技奖
		中华医学科技奖
		省市级奖
	人才团队	杰出人才
	任职及交流	重要学会任职
		重要期刊任职
		学术交流
科技条件	项目	科研项目
		临床试验项目
	平台	科学与工程研究类平台
		技术创新与成果转化类平台

1.2 基于云模型的评价模型

1.2.1 云模型理论

云模型的定义基于一个定量论域U和论域U上的定性概念C，若数值$x \in U$且x是C的一次随机实现，则x对 C 的隶属度$\mu(x) \in [0,1]$是具有稳定倾向的随机数。

$$\mu(x): U \to [0,1], \forall x \in U, x \to \mu(x) \quad (2)$$

则x在U上的分布即称为云模型，x即称为云滴。

云滴可以被视为定性概念在概率意义下的随机实现。不同的云滴分布会导致不同的云模型。云模型通常用三个数字特征来描述期望Ex、熵En和超熵He 3 个数字特征来反映。云模型可以实现定性概念和定量表达的内外延双向转化，云模型的实现算法主要包括正向云生成和逆向云生成两类算法。

1.2.2 建立评价指标云

对综合赋权之后的 19 个二级指标进行打分，运用 MATLAB 软件将打分转化为定量数据，通过云模型的逆向云发生器算法对二级指标的样本数据得分值进行运算，计算二级 19 个评价指标的云模型数字特征值，指标等级标准的界限是具有双边的定量指标变量，如 $x \in (a, b)$，则其云模型参数期望Ex、熵En和超熵He的计算公式为：

$$\begin{cases} Ex = (a + b)/2 \\ En = (b - a)/6 \\ He = k \end{cases} \quad (3)$$

式中a、b分别为某评价指标等级标准对应的最小和最大边界值；k 为常数，可依据评语集的模糊熵进行调整。

1.2.3 建立综合评价云

根据逆向云发生器算法计算出二级指标的云数字特征(Ex, En, He),再代入浮动云算法计算出相应一级指标层(即准则层)的各维度指标和综合绩效值(即目标层)的云模型数字特征，通过正向云发生器计算，若x满

足$x \sim N(Ex, En')$，其中$En' \sim N(Ex, He^2)$，则x对C的确定度满足

$$\mu(x) = e^{\frac{-(x-Ex)^2}{2En^2}} \quad (4)$$

2 医院科研能力评价云模型实证研究

2.1 权重确定

以某地区 64 家三甲医院科研能力评价数据作为实证研究内容，其中评价指标体系中的定量指标已通过历史数据资料获取，定性值通过专家打分获取，指标皆为正向指标，算得出客观权重值。根据博弈论纳什均衡原理，综合主客观权重二级指标的组合权重值，如表 2 所示，其中$\omega 1 = 0.405$，$\omega 2 = 0.278$。通过逆向云发生器算法，通过云模型的逆向云发生器算法对综合赋权后的二级指标的样本数据得分值进行运算，计算二级 19 个评价指标的云模型数字特征值，形成组合最终权重，指标等级标准的界限是具有双边的定量指标变量云（表 2）。

表 2. 指标计算结果

Table 2. Indicator calculation results

一级指标	组合最终权重	三级指标	层次分析法	熵权法	CRITIC	综合赋权
科技产出	0.65	SCIE 论文	0.99%	0.027	1.72%	0.024
		中文核心期刊论文	0.16%	0.015	6.76%	0.014
		授权发明专利	9.95%	0.061	0.10%	0.054
		发明专利转化	65.75%	0.083	0.02%	0.105
		标准和指南	2.88%	0.08	0.63%	0.074
		SCIE 收录论文引用	0.07%	0.033	34.34%	0.030
		中文核心期刊论文引用	0.02%	0.016	47.51%	0.015
		国际权威指南论文引用	94.45%	0.054	0.03%	0.051
学术影响	0.28	国家科技奖	38.44%	0.092	0.07%	0.087
		中华医学科技奖	30.04%	0.059	0.10%	0.050
		省市级奖	4.25%	0.049	0.19%	0.040
		杰出人才	1.19%	0.03	2.29%	0.025
		重要学会任职	0.53%	0.081	2.19%	0.050
		重要期刊任职	5.51%	0.034	0.25%	0.054
		学术交流	0.35%	0.03	1.33%	0.027
科技条件	0.07	科研项目	0.94%	0.065	1.69%	0.043
		临床试验项目	3.70%	0.083	0.19%	0.063
		科学与工程研究类平台	8.15%	0.079	0.17%	0.069
		技术创新与成果转化类平台	3.49%	0.027	0.42%	0.043

2.2 计算综合评价云

通过一级指标权重值和一级指标的组合最终权重，可得出医院科研能力评价综合值的云数字特征为（1.1910，0.5642，1.6731），以此生成综合评价云图（图 1）。通过某地区三级医院科研能力综合值的云数字特征的结果可知，Ex=1.1910 该值代表着地区三级医院科研能力的中心值。一个期望值为 1.1910 意味着整体上，这个地区的三级医院的科研能力处于中等和较高之间，在云模型中，期望值越高，表明评价的程度越高。En表示波动或分散程度，En=0.5642 可以得出 64 家医院之间的科研能力波动较大，即有些医院的科研能力较高，而有些较低。He =1.6731，$He/En < 1$，比值小于 1 说明评价结果的不确定性程度较低，评价的真实性较高，即该模型给出的评价结果较为稳定和可信。综合评价云位于中等云和较高云之间，表明该地区的三级医院整体的科研能力介于中等与较高之间（图 1）。

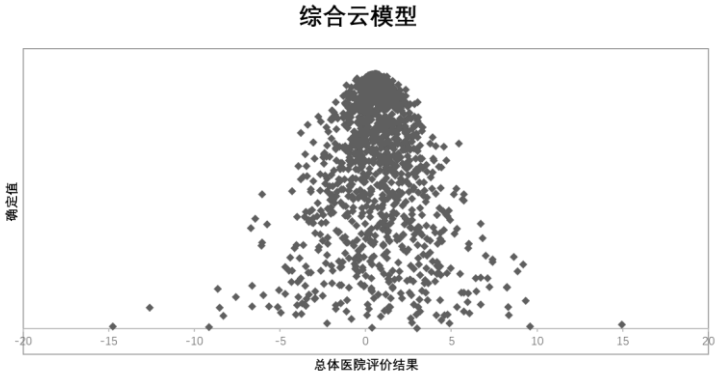


图 1. 综合云模型结果

Figure 1. Comprehensive Cloud Model Results

2.3 计算云相似度距离

选择四家医院作为案例，四家医院为已有评价结果中排名前四的医院[12]，第 7 医院Ex最小，表明其的科研能力相对较低，第 2 家医院Ex最大，科研能力水平相对较高，第 2 医院和第 4 医院En分布跨度较大，集中在-200 与 200 之间，这表明这两家医院的科研能力的不确定性较高，可能有很多因素导致它们的科研能力发生波动。He第 5 医院和第 7 医院云滴凝聚抱合程度最高，表明评价结果的不确定性较小，即它们的科研能力评价结果较为稳定。通过计算期望值发现第 2 医院为排序结果好（图 2）。这个地区的三级医院整体的科研能力介于中等和较高之间，但在具体医院中存在一些差异。

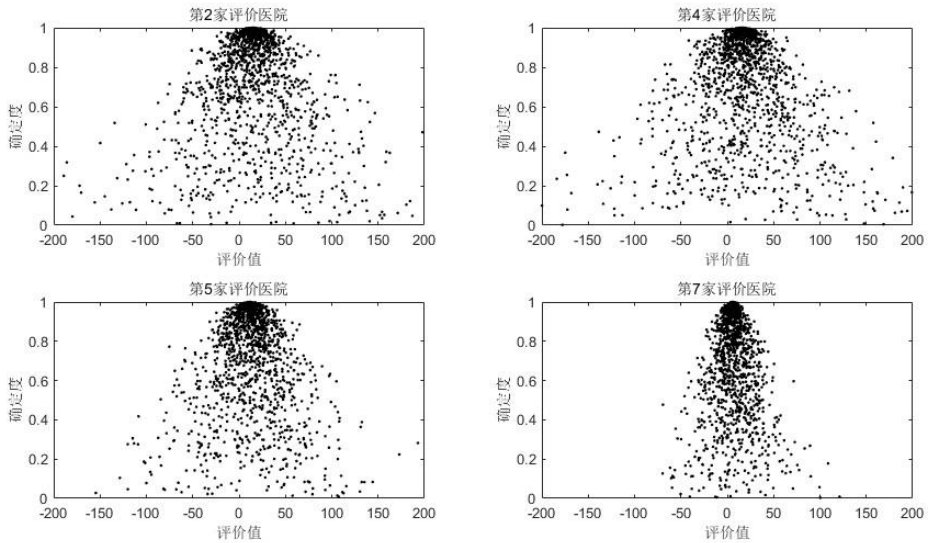


图 2. 四家医院云模型结果

Figure 2. Cloud Model Results for Four Hospitals

3 结果

为提升中国医院的科研能力评价的准确性和科学性，本研究结合了现有医疗机构的评价的不确定性、随机性与模糊性，使用博弈论结合云模型评价的方式开展了更为细致和深入的评估。在 64 家三甲医院的真实数据测试中，得出的结论与已有结果相符，证明了本方法的有效性和可靠性。整体医院科研能力评价通过云模型得到该地区三级医院的科研能力整体处于中等至较高之间，这 64 家医院的科研能力具有显著的波动性，表明存在一定的差异。使用该模型评价的结果具有较高的真实性和稳定性。这为地方卫生部门和医院管理部门提供了一个科研能力的参考，并可以进一步进行决策和规划。

相较于传统的评价方式，本研究结合了多种先进的评价方法，如 CRITIC 法、AHP、熵权 TOPSIS 法和 TODIM 法，通过博弈论组合赋权模型确定了指标权重，引入了基于云模型的评价方式，成功解决了权重赋权中的不确定性、随机性和模糊性问题。这种综合应用多种方法学和模型的方式，确保了评价结果的科学性和公正性。本研究使用的云评价模型的真实性和稳定性较高，为该地区的医院管理和决策提供了有力的参考。选取的四家医院案例进一步加深了对该地区医院科研能力的了解，为医院管理者和相关政策制定者提供了明确和具体的参考信息。针对那些科研能力较低或波动较大的医院，有必要进行进一步的调查和分析，以确定原因并制定相应的策略和措施。

本研究不仅仅是一个简单的评价过程，更是对当前医疗机构科研能力评价体系的一个重要补充。通过引入和整合各种方法学，提供了一个更加科学和精确的评价框架。对于未来的医疗评价研究，本研究提供了一个全新的、更为先进的思路和工具，有助于进一步完善和推进我国医疗评价体系的建设。

OR-014

基于 Lasso-SFO-BP 的医药专利价值预测研究

张云秋 殷策
吉林大学

0 引言

党的二十大报告在总结新时代十年的伟大变革时指出，我国在科技创新领域取得重大成果，已进入创新型国家行列。高价值的专利有助于增强国家的科技实力，使其在国际竞争态势中占据有利地位[1]。

对具有高价值的专利进行预测，可以帮助研究人员、投资者和专利权人将研发资源和投资集中在优质知识产权的研发和收购上[2]，优化其知识产权资产配置。同时，预测高价值的专利可以帮助研究人员、投资者和公司识别可能对特定行业或技术部门产生重大影响的潜在突破性技术[3]。因此，对具有高价值的专利进行预测可以从优化研发资源配置、发现突破性技术等方面帮助研究者和相关人员，从而促成高质量的科技创新发展。

1 相关研究

专利价值的确定一般从筛选影响因素开始，每项专利都同时受到多种因素的影响，这些因素既有一定区别又同时互相联系，影响因素的筛选对专利价值指标体系的构建尤其重要[4]。2017 年我国国家知识产权局发布的《专利资产评估指导意见》明确指出：影响专利价值的因素可划分为技术、法律和市场三个主要维度，并提出了测度专利价值的指标-专利价值度（Patent Value Degree, PVD）。

有学者运用生命周期理论、价值捕获理论期权定价法等理论和方法对医药专利的资产属性评估体系进行了改进，使其产生的评估结果更接近于市场价值[3-7]。还有学者以国家知识产权局公布的评价维度为研究基础，或整合更多的影响因素，或筛选更具影响力的影响因素，并通过专利内容分析、问卷调查以及专家咨询等方法对该专利价值评价体系进行更新，使评估结果更符合专利价值评估行业的特点[4,8-10]。

专利指标数据集往往具有指标维度较大、类别不平衡的特点[11]，钟炜楠[12]对比不同的机器学习算法，认为 BP 神经网络对专利价值的评估具有最好的评估性能，但其对高、中价值专利的预测准确率相对较差、且运算时间较长。张佳敏[13]将粗糙集理论与 BP 神经网络结合，对冠状病毒疫苗专利价值进行预测。

现有的将机器学习方法引入专利价值的研究还存在两个不足：（1）已有研究在专利价值评价维度基本确定的前提下，对影响因素的筛选具有一定的主观性，缺少量化的筛选途径。（2）BP 神经网络作为最适用于专利价值预测的机器学习方法，其本身具有收敛速度慢、取得误差极小值的速度慢的特点[14-15]。

2 模型构建

2.1 Lasso 变量选择

Lasso 特征筛选方法是在 1996 年由 Tibshirani 首次提出，该方法是基于惩罚方法对样本数据进行特征选择的，通过不断改变 Lambda 值，将原有的特征系数进行压缩，直至部分特征系数被压缩至 0，再将这部分特征视为非显著特征并进行舍弃，从而达到特征筛选的目的[16-17]。Lasso 模型的原理为：设x为自变量，y为因变量，n为取样次数；经n次取样后，得到的观测数据标准化值为(x,y)，其中x为维度$n\times p$的矩阵$(n>p)$，y为$n\times 1$矩阵，x的第i个观测值为$x_i=(x_{i1},x_{i2},\dots,x_{ip})^T, i\in[1,2,\dots,n]$且各观测值是相互独立的；$y=(y_1,y_2,\dots,y_n)^T$。$y$对$x$的数理回归模型一般表示为：

$$y_i=\hat{a}+\sum\beta_j x_{ij}+\varepsilon_i \qquad 公式（1）$$

公式（1）中的$\varepsilon_i\sim N(0,\sigma^2)$，定义中的$\hat{a}=\bar{y}$，标准化的数据中$\bar{y}=0$，所以，将公式（1）进行整理后可得：

$$y=\beta x+\varepsilon \qquad 公式（2）$$

其中$\varepsilon\sim N(0,\sigma^2)$，$\beta$为$n$维的参数向量，$\varepsilon$为随机扰动项，研究者如果要筛选出影响较为显著的变量，则需要给公式（2）添加一个约束表达：

$$arg\frac{min}{\{\beta_1,\beta_2,\dots,\beta_n\}}\|y-\beta x\|^2 \qquad s.\lambda\sum_j\frac{|\beta|}{\sum\beta_j^0}\le s \qquad 公式（3）$$

公式（3）中，$s=\frac{t}{\sum\beta_j^0}$，其中$s$的取值范围是$[0,1]$，$\lambda$是调节参数，$\lambda\ge 0$。

2.2 旗鱼算法

旗鱼算法也称为旗鱼优化器（Sailfish Optimizer），是由 S.Shadravan[18]提出的一种元启发式算法（Meta-heuristic Algorithm）。在旗鱼算法中，假设旗鱼为候选解，旗鱼在整个搜索空间中的位置为问题变量，旗鱼可以利用它们的可变位置向量在多维空间中进行搜索。在d维的搜索空间中，第k个搜索回合的第i个成员具有的位置为$SF_{i,k}\in R(i=1,2,\dots,m)$。使用矩阵$SF_{position}$保留所有旗鱼的位置。

$$SF_{position}=\begin{bmatrix} SF_{1.1} & SF_{1,2} & \cdots & SF_{1,d} \\ SF_{2,1} & SF_{2,2} & \cdots & SF_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ SF_{m,1} & SF_{m,2} & \dots & SF_{m,d} \end{bmatrix}$$

其中，m表示旗鱼种群中旗鱼的数量，d表示变量的数量，$SF_{i,j}$表示第i条旗鱼在第j个维度上的值。计算每条旗鱼的适应度的函数为：

$$fitness=f(SF_1,SF_2,\dots,SF_m) \qquad 公式（4）$$

沙丁鱼群的位置矩阵用$SD_{position}$来表示：

$$SD_{position} = \begin{bmatrix} SD_{1.1} & SD_{1,2} & \cdots & SD_{1,d} \\ SD_{2,1} & SD_{2,2} & \cdots & SD_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ SF_{n,1} & SD_{n,2} & \ldots & SD_{n,d} \end{bmatrix}$$

计算每条沙丁鱼的适应度的函数为：

$$fitness = f(SD_1, SD_2, \ldots, SD_n) \qquad 公式（5）$$

旗鱼算法能够保留每次迭代的最优解，算法会设定旗鱼群体中适应度的值最高的个体，作为每次迭代攻击的“精英旗鱼”$X^i_{elite\_SF}$;设定沙丁鱼群体中，适应度的值最高的个体，作为沙丁鱼群体收到每次迭代攻击后，最易被捕获的沙丁鱼$X^i_{injure_SD}$。在旗鱼群体不断迭代攻击沙丁鱼群体的过程中，旗鱼的位置更新公式为：

$$X^{i+1}_{new\_SF} = X^i_{elite\_SF} - \mu_i \times (rand \times \frac{X^i_{elite_{SF}} + X^i_{injure_{SD}}}{2} - X^i_{old\_SF} \qquad 公式（6）$$

其中，$rand$为随机扰动项，其取值为(0,1)。$X^i_{old_SF}$表示旗鱼先前所在的位置，μ_i为第i次迭代时产生的系数，其生成如下：

$$\mu_i = 2 \times rand \times PD - PD \qquad 公式（7）$$

其中PD沙丁鱼群的密度，表示每次迭代的沙丁鱼数量。由于旗鱼群体狩猎时，沙丁鱼的数量会减少，因此PD参数是更新旗鱼在沙丁鱼群周围位置的重要参数。该参数的自适应公式为：

$$PD = 1 - (\frac{N_{SF}}{N_{SF} + N_{SD}}) \qquad 公式（8）$$

其中N_{SF}表示旗鱼种群的数量，N_{SD}表示沙丁鱼种群的数量。沙丁鱼的位置更新公式为：

$$X^{i+1}_{new\_SD} = rand \times (X^i_{elite_{SF}} - X^i_{old_{SD}} + AtkP) \qquad 公式（9）$$

$$AtkP = A \times (1 - (2 \times Itr \times e)) \qquad 公式（10）$$

其中$X^i_{old_{SD}}$表示沙丁鱼先前所在位置，$AtkP$为攻击力度参数，该参数可用于调节沙丁鱼鱼群随旗鱼群体的迭代攻击而产生的位移量。A和e为攻击力度参数的系数，起调节作用。Itr为迭代次数。

2.3 BP 神经网络

BP 神经网络（Back-propagation neutral network）又称反向传播神经网络，是 1986 年由 Rumelhart[19]等提出的一种最为传统的神经网络算法。该算法将样本从输入层传入，经由隐层处理后再向输出层进一步传输。随后计算输出值与目标值之间的误差，并将误差值以某种形式反向传播到隐层和输入层的神经元上，各层的神经元通过误差来修改各自的权值并逐层调整各网络的权值和阈值，进而取得较小的误差。

2.4 Lasso-SFO-BP 模型构建

本研究构建的 Lasso-SFO-BP 模型，首先利用 Lasso 回归筛选出对专利价值影响较大的特征指标，并作为 BP 神经网络的输入；而后使用 SFO 优化 BP 神经网络的权更新方式，以提高其收敛速度；最后将 SFO 迭代寻优得到的权值阈值传递到 BP 神经网络中进行训练。本研究的模型见图 1。

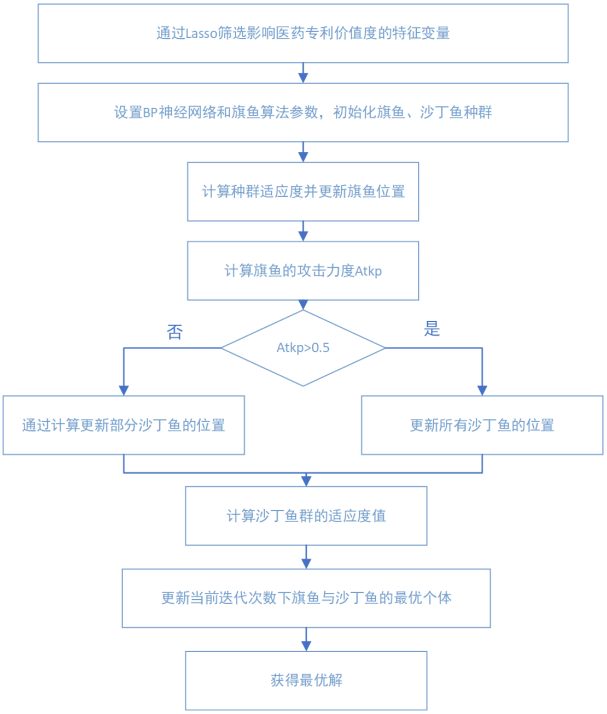


图 1：Lasso-SFO-BP 模型图

3 实证

3.1 数据来源

本研究以“血管舒张药”为例，使用国际专利分类号（IPC 分类号）在德温特数据库（DII 数据库）中进行检索，得到 3539 条专利数据，清洗后保留 3503 条专利数据纳入数据集，随机抽取 70%的数据作为训练集，其余 30%的数据作为测试集。测度专利价值的量化结果采用 incoPat 数据库使用的“合享价值度”来表示。

3.2 Lasso 特征筛选

2012 年《专利价值分析指标体系操作手册》提出了评价专利价值的三个维度：技术维度、法律维度、市场维度，又基于这三个维度提出了 18 个特征指标。本研究以上述三个维度 18 个特征指标作为医药专利价值评价的影响因素，进行 Lasso 特征筛选。

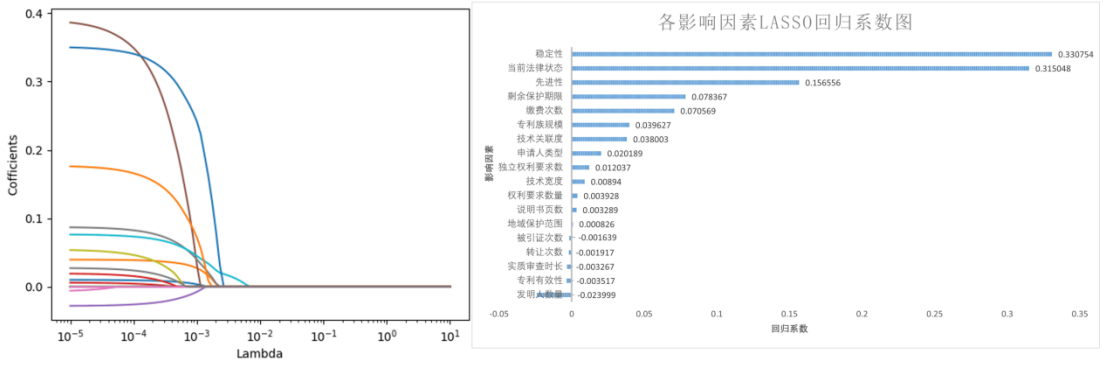


图 2：各影响因素 Lasso 回归系数图

Lasso 结果见图 2，最终确定的影响因素分别为：稳定性、当前法律状态、先进性、剩余保护期限、缴费次数、专利族规模、技术关联度、申请人类型、独立权利要求数、技术宽度、权利要求数量、说明书页数。

3.3 SFO-BP 模型参数设定

3.3.1 BP 神经网络参数设定

本研究使用的 BP 神经网络采用 3 层结构，由 Lasso 结果可知，输入层节点数量为 12，输出

层节点数量为 1，设定为专利价值度。采用 tanh 作为第一层的激活函数、sigmoid 作为第二层的激活函数。经调试，隐层设置 10 个节点时，模型的误差最小。将模型的训练次数设定为 1000 次，学习率设定为 0.01。

3.3.2 SFO 算法参数设定

本研究设定沙丁鱼种群数量为 100，定义最大迭代次数为 1000，定义旗鱼的数量为沙丁鱼的 1/3。A系数设定为 4，e系数设定为 0.001。

3.4 SFO-BP 医药专利价值预测

将 3503 条专利数据分为两组，随机选取 2452 条专利作为训练样本集，其余 1051 条作为测试集。根据公式（6）更新旗鱼种群的位置，根据公式（9）和公式（10）更新沙丁鱼种群的位置，每次更新使迭代次数增加 1，直到达到预先设定的最大迭代次数。将此时产生的最优权值和阈值输入 BP 神经网络中，训练网络并进行预测。使用 SFO-BP 得出的预测值和单纯使用 BP 神经网络得出的预测值见图 3。

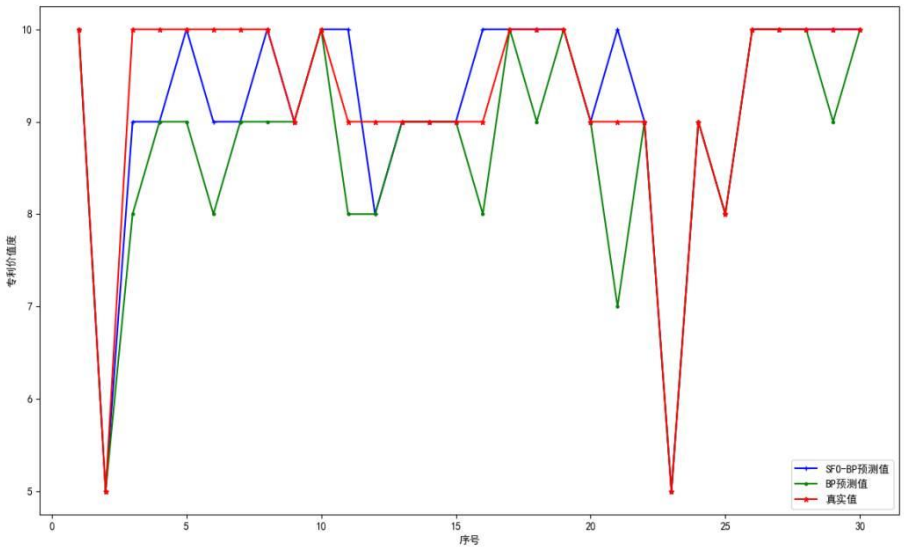


图 3：两种算法预测值（部分）

由于本研究的测试集数目较大，图中仅展示 30 例预测值输出结果。经实际测试，SF0-BP 在本研究数据集上可获得 78.32%的预测准确率，高于单纯 BP 神经网络的 70.03%。

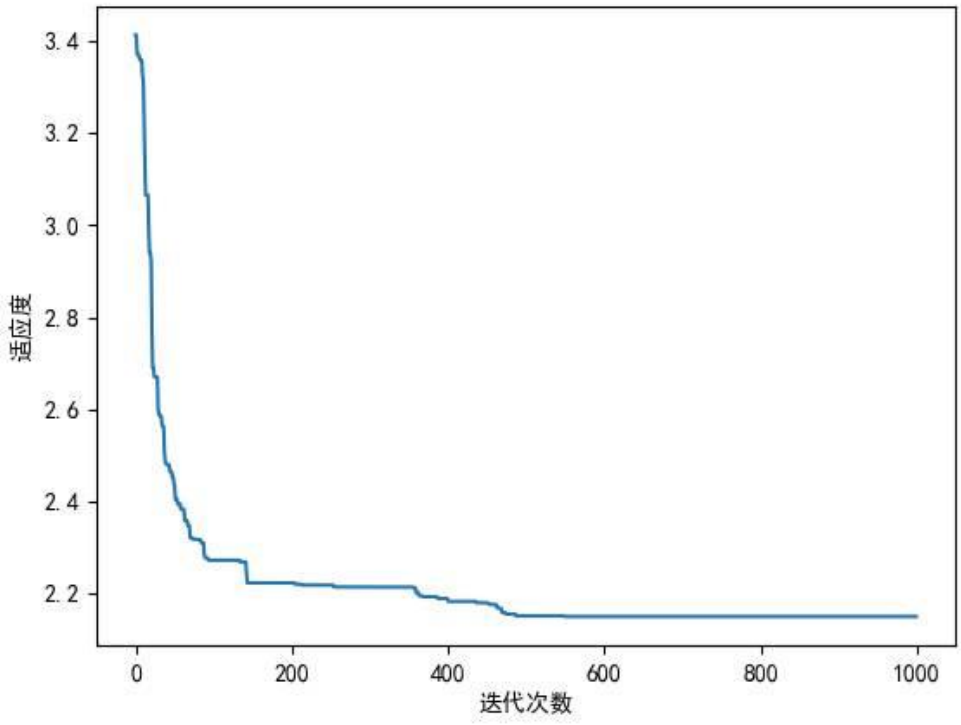


图 4：SFO-BP 算法收敛过程图

SFO-BP 算法的收敛过程见图 4，如图可知，在第 551 次迭代时，算法已经完成了收敛过程，此时的适应度为 2.14999，说明已经找到了最优权值和阈值。SFO-BP 算法与 BP 神经网络的收敛过程对比图见图 5，如图可知，在第 844 次迭代时，BP 神经网络完成收敛过程，此时的适应度为 2.18283，较 SFO-BP 的收敛值更低，收敛速度更慢。

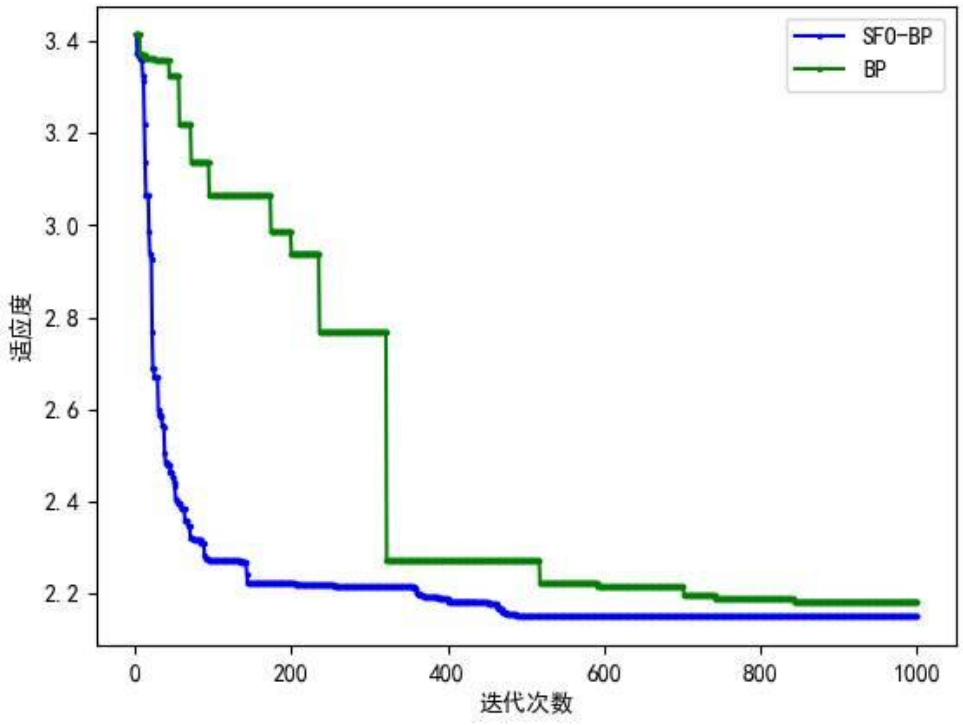


图 5：SFO-BP 与 BP 神经网络收敛过程对比图

3.5 Lasso-SFO-BP 模型性能分析

本研究选取均方误差（MSE）、均方根误差（RMSE）、平均绝对误差（MAE）、平均绝对百分比误差（MAPE）四个指标评价模型的预测准确性，这四个指标的计算公式如下：

$$MSE = \frac{1}{n}\sqrt{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad 公式（11）$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad 公式（12）$$

$$MAE(X,h) = \frac{1}{n}\sum_{t=1}^{n}|h(x_t) - y_t| \qquad 公式（13）$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \qquad 公式（14）$$

其中，$\hat{y}_t$表示 SFO-BP 算法或 BP 神经网络算法的预测值，y_t表示医药专利价值的真实值，n表示数据集的数据个数。上述 4 个指标的值越小，则表示预测的越准确。本研究同时选取收敛时间（s）和收敛值（W）来评价模型的收敛速度，其中收敛时间表示为完成收敛所需的迭代次数；收敛值表示为适应度下降值。两种算法的统计学指标值详情见表 1 所示。

表 1：SFO-BP 与 BP 神经网络模型指标值

	MSE	RMSE	MAE	MAPE	s	W
SFO-BP	0.26641	0.51615	0.23025	0.02871	551	1.26354
BP 神经网络	0.54044	0.73514	0.36727	0.04358	844	1.23071
差值	0.27403	0.21899	0.13702	0.01487	293	0.03283

从以上结果可以看出，与 BP 神经网络模型相比，SFO-BP 模型对医药专利价值的预测更加准确，预测值与真实值的差距更小，拟合效果更好，收敛速度更快。这说明 SFO 算法有效地降低了 BP 神经网络的预测误差值，证明了用 SFO 算法更新优化 BP 神经网络的权值和阈值的有效性。

4 结论

本研究将 BP 神经网络应用于医药专利价值的预测，并采用 Lasso 回归对影响专利价值测度的特征变量进行筛选，引入 SFO 算法对 BP 神经网络的权值和阈值更新方式进行优化，改进了 BP 神经网络的预测精确度和收敛速度。

OR-015

基于深度学习模型的医学文献主题识别与演化路径分析研究

王德超 李爱花 李永洁 陈逸菲 于诗睿 李晓瑛 杨雪梅 唐小利
中国医学科学院/北京协和医学院医学信息研究所/图书馆

1 引言

随着科技的快速发展，科技文献数量呈指数级增长，而科技的发展是一个循序渐进相互联系的过程。如何从文献中快速准确地获取所需主题并找出主题的发展规律，是研究者们需要面对的重要问题。主题识别可以快速地获取文献的核心信息，主题演化可以揭示科技研究的演变特征，为科技的发展提供指导。

宫颈癌是女性常见恶性肿瘤之一，研究宫颈癌文献对于促进公共卫生和保护女性健康具有重要意义。本文以 2012-2021 年宫颈癌领域文献研究为例，基于深度学习主题模型识别文献文本的主题分布与演化特征,意在拓宽深度学习主题模型的适用领域，探索在医学文献研究领域涵义丰富且有效可行的主题识别方法与主题演化分析路径。

2 国内外研究现状

2.1 主题识别相关研究

主题识别是指从文本数据中探测和挖掘隐含的主题信息 [1]，传统主题识别方法的研究主要包括基于聚类算法的研究方法 [2,3,4,5] 以及基于线性代数降维 [6,7,8] 的研究方法，目前较为主流的主题识别方法为基于概率的主题模型及其改进算法。

Blei 等人提出经典概率主题模型 Latent dirichlet allocation(LDA)模型 [9]，LDA 模型被广泛应用于主题识别研究 [10,11]，国内外学者在此基础上提出各类改进后的主题模型 [12,13]。Supervised LDA 可以潜在地学习与单个类标签内联的主题并生成相应的概率分布 [14]；Labeled LDA(L-LDA)是用于多标签文档分类的监督 LDA 模型 [15]，L-LDA 由于其效率和简洁性得到应用。

基于神经网络的深度学习主题模型利用神经网络重构文本生成过程。Cao 等人提出基于前馈神经网络的 NTM 模型，能够获得更好的主题表示和分类准确率[16]；基于变分自编码器的主题模型包括 NVDM[17] 与 ProdLDA[18] 等模型，ProdLDA 模型具有较高的主题词一致性水平，可产生更多的可解释性主题。

基于词向量模型 Word2Vec 的主题-词向量模型如 TopicVec 模型[19]、STE 模型[20] 以及 LDA2Vec 模型[21,22,23] 可表达不同主题的词向量与文档-主题分布；基于文本生成模型的主题模型如 TDLM 模型可在提升主题词抽取和文本分类能力的同时生成自然文本，提高识别主题的语义化表达能力[24]。

Wang 等人提出基于生成对抗网络的 ATM 模型，ATM 模型使用狄利克雷先验对主题建模，使用生成器网络学习潜在主题之间的语义模式，在多个数据集上表现优于 NVDM 模型与 ProdLDA 模型[25]。

2.2 文献主题演化相关研究

主题演化是指主题在时间维度上的发展、变化趋势以及不同主题之间的交互作用[26]。传统文献计量学方法进行主题演化的主要分析指标为文献的关键词与引文[27,28]，文本主题挖掘的方法包括主题时序分析、主题强度分析以及主题内容分析[29]。

主题时序分析包括嵌入式时序分析方法如 TOT 模型[30]、先离散式时序分析方法以及后离散式时序分析方法[27]。主题强度分析中主题强度表示文本主题受关注的程度，主题的演化过程可通过观测文本主题随时间变化的趋势来衡量[31,32]。主题内容分析方法通过计算相邻阶段主题之间的余弦相似性，基于主题的时间前后相似性研究主题演化模式[33,34]，除此，主题内容分析还可进行主题共现分析，如 Katsurai 等通过动态共词网络研究主题突变以反映主题演化趋势[35]。

3 研究设计

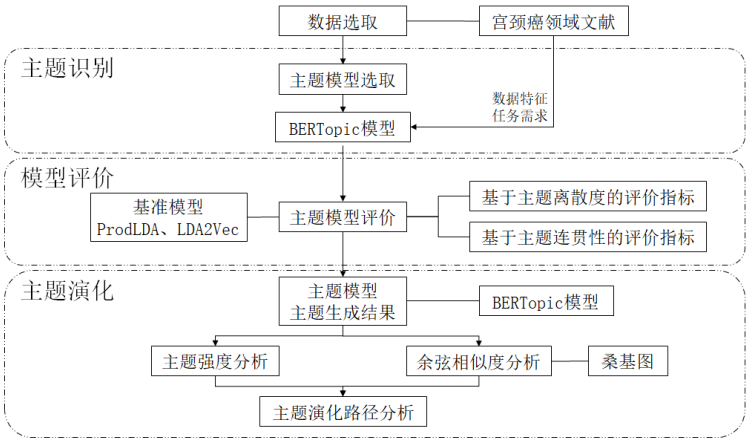


图 1 主题识别演化研究路径

3.1 数据选取

宫颈癌研究有深厚的研究基础与丰富的文献数量，该领域的研究涉及多个学科领域，包括生物学、医学、流行病学等。宫颈癌领域的文献文本能提供足够的样本数据支持，也因其具有多学科交叉性能够提供特征丰富的文本数据。

数据集选取 2012-2021 年 Web of Science(WOS)数据库中宫颈癌相关的文献摘要，共获取文献 47272 篇。

宫颈癌领域文献涉及较多医学术语和专业名词，稀疏数据较多。同时文献涉及多个领域，文献类型多种多样，具有不同的文本结构和语言风格，对模型的语义理解能力要求较高。

3.2 主题识别

Maarten 于 2022 年提出 BERTopic 模型，该模型应用主题建模的密度和层次聚类方法，使用基于类的 TF-IDF(c-TF-IDF)过程进行神经主题建模，使用最大边际相关性算法调整主题的连贯性[36]。

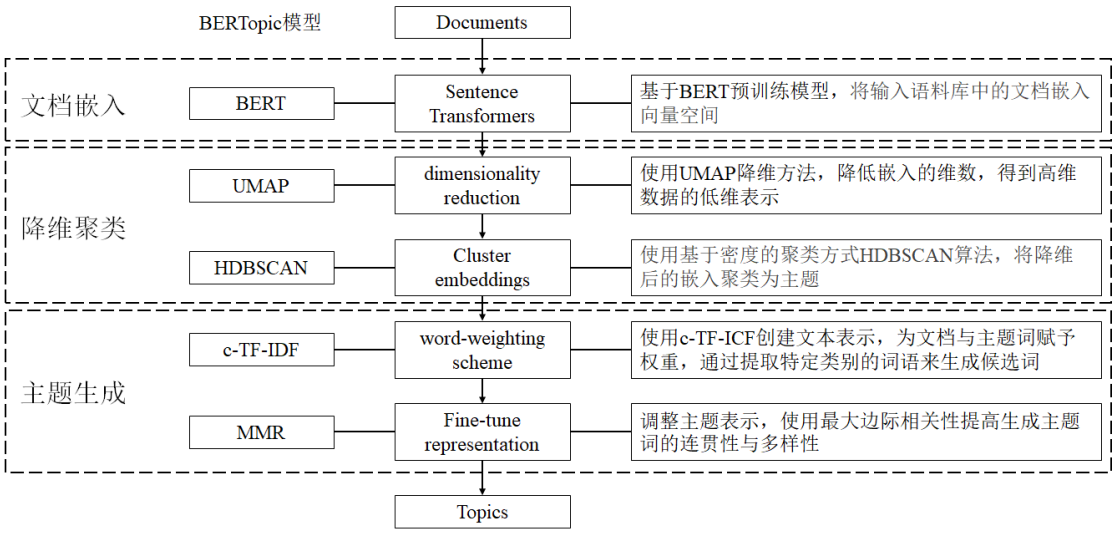


图 2 BERTopic 模型结构示意

主题识别模型的选取应根据具体任务需求和数据特征决定。

BERTopic 模型是基于 BERT 模型的无监督聚类算法，BERT 作为预训练的语言模型，能够更好地处理同义词、歧义词等问题，能够捕捉到文本中更加复杂的语义关系。

宫颈癌领域文本中的专业术语使词袋模型较复杂，文本向量的文本维度较高，BERTopic 模型使用 UMAP 方法执行降维，使用 HDBSCAN 算法进行聚类，两种方法使模型在高维数据的处理上效果较好，高维文本向量数据在低维空间能较好地保持结构。BERTopic 不需要迭代计算，在处理大规模数据时可以较好地满足处理效率和精度的需求。

BERTopic 可以自动生成 Batch 以及主题数，最终生成的主题结果如下表所示，模型生成主题数量为 237，在此选取出现频次排名前 25 的主题识别结果。

表 1 BERTopic 模型生成主题

主题编号	主题名称	主题词分布
topic_1	甲状腺癌治疗	thyroid, ptc, papillary, thyroidectomy, lymph, node, metastasis, lateral, metastases, patients
topic_2	癌症筛查	screening, health, women, care, breast, pap, participation, colorectal, cancer, years
topic_3	近距离放疗	brachytherapy, dose, gy, plans, planning, volume, bladder, rectum, applicator, bt
topic_4	宫颈癌细胞构成	compounds, synthesized, against, derivatives, activity, compound, mu, lines, ic50, hela
topic_5	HPV 疫苗宣传	vaccination, vaccine, hpv, knowledge, students, parents, about, girls, awareness, mothers
topic_6	宫颈癌病毒检测	cytology, hpv, hc2, cin2, test, hrhpv, genotyping, testing, women, cin3
topic_7	基因风险	polymorphism, polymorphisms, allele, susceptibility, risk, genotype, association, 95, controls, snps
topic_8	宫颈癌检测指标	e6, e7, viral, p53, hpv16, replication, proteins, protein, cellular, oncoproteins
topic_9	RNA 因素	lncrna, lncrnas, noncoding, cc, rna, long, expression, rnas, hotair, proliferation
topic_10	宫颈癌细胞特征	complexes, complex, ligand, activity, ligands, binding, characterized, synthesized, against, hela
topic_11	抗肿瘤效应物	extracts, extract, activity, compounds, antioxidant, gml, plant, mu, cytotoxic, against
topic_12	宫颈癌病理类型	endocervical, adenocarcinoma, uterine, mucinous, case, cases, cervix, rare, carcinoma, glandular
topic_13	宫颈癌增殖转移	proliferation, expression, migration, tissues, cc, invasion, overexpression, cell, target, cells
topic_14	头颈部癌	oropharyngeal, oral, opscc, hpv, hnscc, p16, head, hpvpositive, neck, opc
topic_15	子宫切除手术	hysterectomy, radical, laparoscopic, lrh, robotic, surgery, open, minimally, surgical, complications
topic_16	宫颈癌认知	knowledge, screening, health, pap, smear, women, about, awareness, participants, cancer
topic_17	基因甲基化	methylation, promoter, genes, dna, cpg, hypermethylation, gene, methylated, pax1, cin3
topic_18	宫颈癌药物治疗	drug, release, nanoparticles, delivery, micelles, nps, dox, drugs, ptx, loaded
topic_19	肿瘤免疫治疗	vaccines, responses, e7, vaccine, therapeutic, immune, cd8, mice, peptide, epitopes
topic_20	宫颈癌基因序列	variants, integration, hpv16, genome, e6, viral, lineage, sequences, lineages, nucleotide
topic_21	前哨淋巴结活检	sln, sentinel, mapping, slns, icg, lymph, node, detection, blue, injection
topic_22	细胞凋亡检测	apoptosis, hela, cells, mitochondrial, cell, induced, apoptotic, ros, caspase3, effect
topic_23	HPV 疫苗接种	vaccine, vaccines, vaccination, hpv, efficacy, quadrivalent, against, 9vhpv, girls, doses

主题编号	主题名称	主题词分布
topic_24	宫颈癌成像	raman, optical, spectroscopy, spectra, tissue, spectral, images, imaging, oct, mueller
topic_25	淋巴瘤诊断	lymphoma, case, rare, diagnosis, tuberculosis, mass, lymphadenopathy, bcell, lymph, tonsillar

3.3 主题评价

考虑到主题模型评价的定性方式受主观影响较大，采用定量方式衡量主题生成效果可信度更高，本研究主要选取基于主题离散度(C_v)与主题连贯性(C_pmi、C_umass、C_uci、C_npmi 及 LCP)的评价指标。

C_v(Coefficient of variance)基于主题的内部一致性来评价主题的质量[37]。C_pmi 基于点态互信息计算两个词语之间的关联程度[38]。C_umass[39]将 C_pmi 相关性得分求平均作为主题的一致性得分。C_uci 使用不同的权重函数对 C_pmi 相关性得分进行归一化处理。C_npmi 在 C_pmi 与 C_umass 的基础上为计算值提供[-1,1]的标准化范围[40]。LCP 使用对数条件概率作为一致性计算 C_pmi 部分的替代，仍然使用参考语料库和滑动窗口取样字数[41]。

本研究选择 ProdLDA 模型与 LDA2Vec 模型作为基准模型。

ProdLDA 模型：基于变分自编码器，通过正则化避免过拟合，取消了 LDA 模型中的归一化过程并重新定义了条件分布，生成的主题词具有较高的一致性与可解释性。

LDA2Vec 模型：基于 Word2Vec 词向量模型，将 LDA 模型的文档向量和 Word2Vec 的词向量联合建模，生成上下文向量并预测主题，处理稀疏数据与捕捉语义信息的能力强。

对 ProdLDA、LDA2Vec 以及 BERTopic 模型的主题生成结果分别计算主题评价指标得分，计算结果如表 2 所示。

表 2 主题模型评价得分

	ProdLDA	**LDA2Vec**	**BERTopic**
C_v	0.350	0.302	**0.564**
C_pmi	0.534	**0.766**	0.584
C_umass	-16.064	-11.973	**-3.292**
C_uci	-8.626	-6.327	**0.078**
C_npmi	0.133	0.156	**0.159**
LCP	-2.812	-2.670	**-2.518**

表中所用各指标均为数值越大，主题模型的效果越好。BERTopic 模型除 C_pmi 指标得分低于 LDA2Vec 模型，其他指标均优于 ProdLDA 模型与 LDA2Vec 模型。可以看出，BERTopic 模型生成的主题分布连贯性较好，对主题的表达能力更强。

3.4 主题演化

主题作为关键词的集合，在时间区间中的强度为该主题关联词的强度之和，主题关联词的强度取决于主题关联词在时间区间内出现的频次与主题关联词在主题中的权重。分别计算 25 个高权重主题在 2012-2021 年的主题强度，各主题的主题强度变化如图 3 所示。

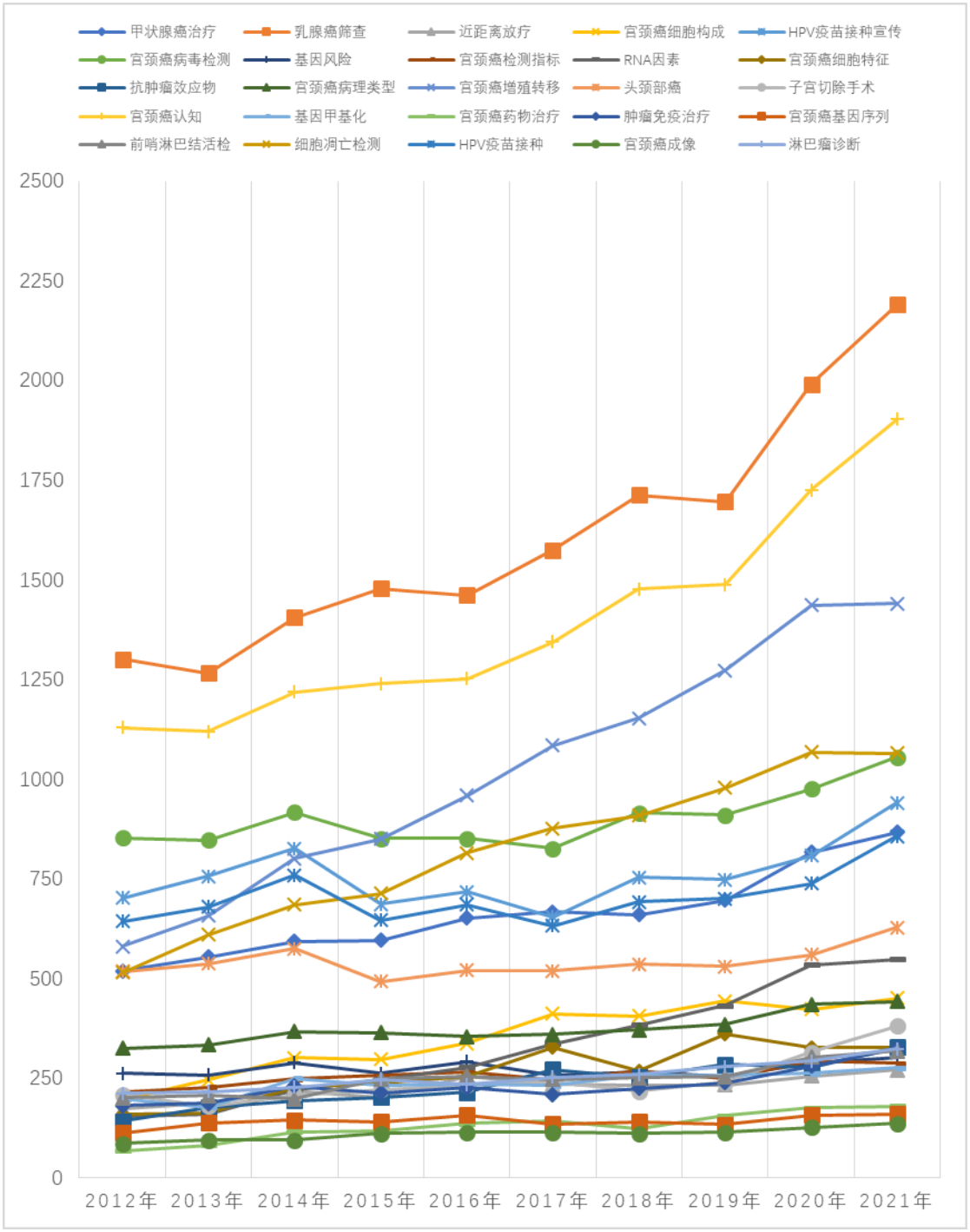


图 3 BERTopic 模型主题强度演变

以图 4 所示主题为例，五类主题总体呈增长趋势，较其他主题增速更快，可以看出宫颈癌研究领域的研究热点不断向细胞、基因等微观研究领域转移。图 5 所示三类主题的主题强度增减态势基本一致，可以推测宫颈癌的宣传、接种以及病毒检测研究之间存在一定相关性。

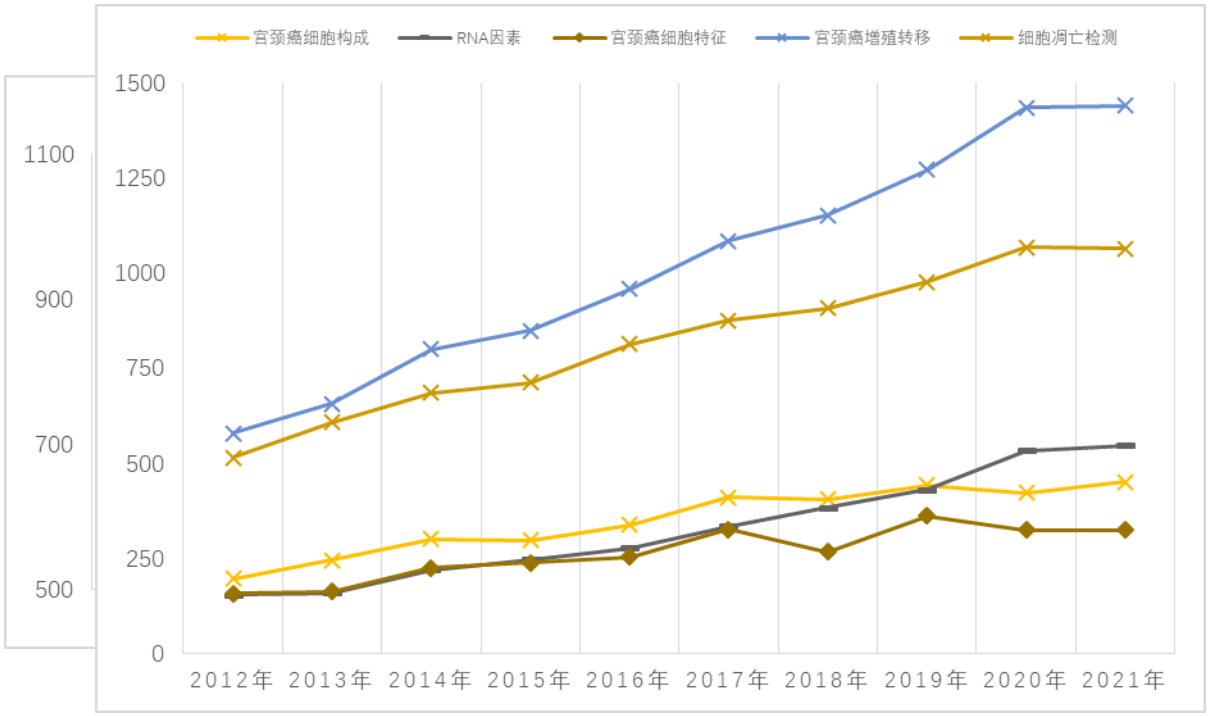


图 4 BERTopic 模型主题结果主题强度增长趋势

图 5 BERTopic 模型主题结果主题强度相关性

余弦相似度用于计算主题间的相似度，将文本特征向量化后计算相邻主题的相似程度。采用夹角余弦相似度算法构建不同年度主题间的演化关系，设置主题演化的阈值为 0.2，主题间相似度大于阈值说明具有一定的演化关系。

筛选余弦相似度计算结果，删除主题间余弦相似度低于阈值的主体间关系，得到 2012-2021 年相邻年份间存在相关关系的主题，并以此为基础对各年目标主题命名并开展主题演变可视化分析。反映主题关系的余弦相似度计算结果如表 3 所示。

表 3 主题间余弦相似度示例

主题 A-Source	主题 B-Target	A-B 余弦相似度
2019_topic1_癌细胞表达侵入因子的靶向测量	2020_topic7_癌细胞凋亡影响因子	0.656087481
2018_topic5_癌细胞表达、入侵	2019_topic1_癌细胞表达侵入因子的靶向测量	0.650812625
2015_topic1_miRNA 抑制作用	2016_topic3_miRNA 调节作用	0.649197128
2016_topic3_miRNA 调节作用	2017_topic1_癌细胞表达、侵袭、迁移机制	0.640963539

主题 A-Source	主题 B-Target	A-B 余弦相似度
2012_topic2_HPV 疫苗接种	2013_topic2_HPV 疫苗宣传	0.578002118
2014_topic10_植物提取物的细胞毒性检测	2015_topic10_抗癌药物生物活性	0.569977585

基于余弦相似度得分，本研究采用桑基图呈现 2012-2021 年宫颈癌领域科研主题的演化路径。根据 2012-2021 年主题间余弦相似度关系，绘制主题演化图以反映主题间的分裂合并关系，如图 6 所示。

图 6 2012-2021 年宫颈癌领域科研文献主题演化桑基图

主题演化图中，以从抗癌药物生物活性起始的通路为例，通过分化合并关系可以看出抗癌药物的细胞毒性活性研究分为三个方向：植物提取物的细胞毒性活性检测、抗癌化合物的生物活性检测以及复合物的生物毒性活性检测。可以推断从不同抗癌物质药物的毒性检测中，研究人员致力于寻找抗癌效果良好且副作用较弱的有效药物。

4 结语

科研文献的文本内容中隐含着科研领域的研究重点和研究热点，主题模型为科研领域显性及

隐性研究方向的挖掘提供了可能性。面对大量且非结构化的文本内容，本研究评估了 BERTopic 模型的主题识别效果，选取了主题连贯性等指标作为主题模型性能的评价标准。在主题识别的基础上，本研究计算了 BERTopic 生成主题的主题强度与余弦相似度，综合主题内容分析、主题强度分析以及主题时序分析的方法揭示宫颈癌领域文献研究热点及发展趋势，并得出了“宫颈癌研究的热点不断向微观研究领域转移”、“HPV 疫苗接种宣传与宫颈癌病毒检测研究间存在相关性”以及“抗癌药物细胞毒性活性的研究随时间逐渐细化与深入”等结论。

本研究也存在一些局限性，模型参数选取需要平衡运行效率与拟合效果，可能存在未发现的参数设置及更优的主题生成结果。此外，受到计算资源的限制，本研究在现有条件下可处理的文本数据量有限，文献数据集的代表性有限，未来需进行更为广泛的数据验证。

OR-016

基于真实世界数据的心脏疾病知识库的构建及应用探讨

印玥 胡广伟（通讯作者）
南京大学信息管理学院

1 引言

随着网络技术的飞速发展，人类获得了空前便利的信息检索和交流方式，信息资源正式进入数智化管理阶段[错误!未找到引用源。]。然而，信息资源浩如烟海，对信息的检索无异于“大海捞针”。此时，构建一个完整有序的、知识化的资源保存体系，为用户提供存档资源的有效获取与利用至关重要。而在生物医学领域，更易产生专业术语间的关系错综复杂、概念的标准化难以界定[错误!未找到引用源。]、知识的共享缺少平台[错误!未找到引用源。]等问题。在人类疾病系统中，心脏疾病复杂且需要得到及时的诊疗[错误!未找到引用源。]，具有构建知识库的必要性。因此，本文以苏南地区某三家医院的心胸外科疾病数据为样本，通过信息源的识别、确定、抽取和表达设立临床路径，对心胸外科疾病领域的知识规则进行形式化处理，最终构建出一个具有疾病诊断、治疗建议等功能的心脏疾病知识库。

知识库作为一种知识组织与服务模式，能够对特定领域多源分布、孤立分散的信息进行全面、有序的集成，经过知识获取、知识表示、知识推理等转化处理，能为用户决策问题求解提供所需的知识服务[错误!未找到引用源。]。本文试图基于本体论的思想，对心脏疾病的相关知识进行系统化的梳理和解析，明确心脏疾病的核心概念及其属性、语义关系及领域知识公理，探索心脏疾病知识库的构造方法，在此基础上，构建心脏疾病的概念框架体系，为医学领域知识资源的深度开发与再利用提供一种可行的思路和方案。

2 知识库的构建路径

2.1 信息源

信息源（Information Source）一般指信息的来源地。联合国教科文组织出版的《文献术语》将其定义为个人为满足其信息需要而获得信息的来源[错误!未找到引用源。]。一切产生、生产、贮存、加工、传播信息的源泉都可以看作是信息源。信息源内涵丰富，不仅包括各种信息载体，也包括各种信

息机构；不仅包括传统印刷型文献资料，也包括现代电子图书报刊；不仅包括各种信息储存和信息传递机构，也包括各种信息生产机构。

医学领域构建知识库的信息源主要有两种类型：一是以一体化医学语言系统（UMLS,Unified Medical Language System）[错误!未找到引用源。]、医学主题词表（MeSH,Medical Subject Headings）[错误!未找到引用源。]、系统化医学命名（SNOMED,Systematized Nomenclature of Medicine）[错误!未找到引用源。]等受控词表为主的结构化信息源，这类信息源语义结构清晰，概念与概念关系成熟、准确，可以直接作为本体构建的语义基础；二是以电子病历、医学文献、在线健康信息为主的非结构化信息源，这类信息源数量庞大，包含更加丰富和多样的术语、概念与概念关系，能够体现用户健康信息需求[错误!未找到引用源。]。目前，后者是现今构建信息源的研究重点。

2.2 信息源的识别

识别合适的信息源是构建一个权威稳定的医学知识库的基础。对于信息源的识别，首先应当根据知识库的目标和可能涉及的领域进行需求分析，明确知识库具备的功能以及需要利用该知识库解决哪些临床问题，从而确定知识库的主题和范畴；其次，应当根据需求分析结果进行知识发现，获得知识库所需的信息源；最后，从信息源抽取领域知识的过程模型及概念模型，构建半结构化的临床决策知识元本体。

2.3 信息源的确定

信息源的确定是知识抽取的基础，对构建知识库至关重要。本文基于五种类型的知识来源分析临床路径知识：（1）医疗机构正在使用的临床路径，包括国家卫生健康委办公厅在中国政府网上公布的心血管临床路径表单、胸外科临床路径表单等；（2）医疗机构正在使用的临床指南，包括中华医学会胸心血管外科学分会制订的中国胸心血管病一级预防指南；（3）领域专家的临床经验，这部分知识隐晦难以具体化，但可以在模型评估阶段给予指导意见；（4）临床路径相关文献。在这个阶段中，与临床专家共同合作，确定临床路径的知识来源，分析临床路径的诊疗流程、适用病症、入径条件、转归标准等[错误!未找到引用源。]；（5）医学术语：本论文选择了医学主题词表术语集及其语义网络作为知识库术语标准化及扩展的依据。

2.5 信息源的抽取

结构化信息源的概念关系已发展成熟准确，且经过统一标准的规范化处理，具有较为固定的存储结构，通常可以通过一些医药数据库直接批量获取。然而，随着科学技术的飞速发展和信息资源的日益增迭，知识库中的信息源需要不断更新以适应发展要求。因此，海量的知识都以科技文献、临床指南、电子病历等非结构化的形式存在，目前很多信息源的抽取也着眼于非结构化的信息源。

非结构化信息源的抽取一般分为命名实体识别（NER,Named Entity Recognition）和关系抽取（RE,Realtion Extraction）两个步骤。命名实体识别是指识别文本中具有特定意义的实体，在文本中定位预定义实体类型的过程，常用的方法包括基于词典的方法、基于规则的方法、基于机器

学习的方法以及混合的方法[错误!未找到引用源。]；关系抽取是指从文本中识别实体并抽取实体之间的语义关系进而获取三元组信息，即<主体，谓语，客体>，常用的方法包括基于共现的方法、基于规则的方法、模式识别和机器学习。

2.6 信息源的表达

本体论是一种既能从语义层面，又能从知识层面对信息系统进行描述和建模的理论，本文试图构建一种基于本体的人类心脏疾病知识库。为了使计算机理解人为添加的本体论知识并具备自动推理的功能，本体提供了对特定领域中概念及其相互之间关系的形式化表示方法——本体描述语言。常见的本体描述语言包括资源描述框架（RDF）、可扩展标识语言（XML）、网络本体语言（OWL）等。OWL 是 W3C 于 2004 年推出的用于描述网络本体的语言标准。根据 W3C 官方文档，与 XML、RDF 相比，OWL 通过扩展词汇表提供了更好的机器互操作性，因此，本文通过 OWL 进行知识库术语描述。

3 人类心脏疾病知识库的构建

心脏疾病是全球范围内导致死亡和残疾的主要原因之一。随着现代生活方式的变化和人口老龄化趋势的加剧，心脏疾病的发病率和负担一直在不断增加。为了应对这一挑战，医学界一直致力于研究心脏疾病的预防、早期诊断和治疗方法。心脏疾病发作的患者通常需要得到及时有效的诊疗，这依赖于医生丰富的经验。随着信息技术的迅速发展，知识库成为了一个重要的工具，用于整合、存储和共享关于特定领域的知识。知识库的建立为医学研究和临床实践提供了宝贵的资源和参考。针对心脏疾病的知识库可以集成和组织大量的相关信息，包括病因、病理生理学、临床表现、治疗选项等方面的知识，为医生、研究人员和患者提供准确、及时和全面的信息支持。将心脏疾病的诊疗看作一个决策过程，心脏疾病知识库则是作为决策的依据。图 1 描述了知识库支撑的心脏疾病临床路径。

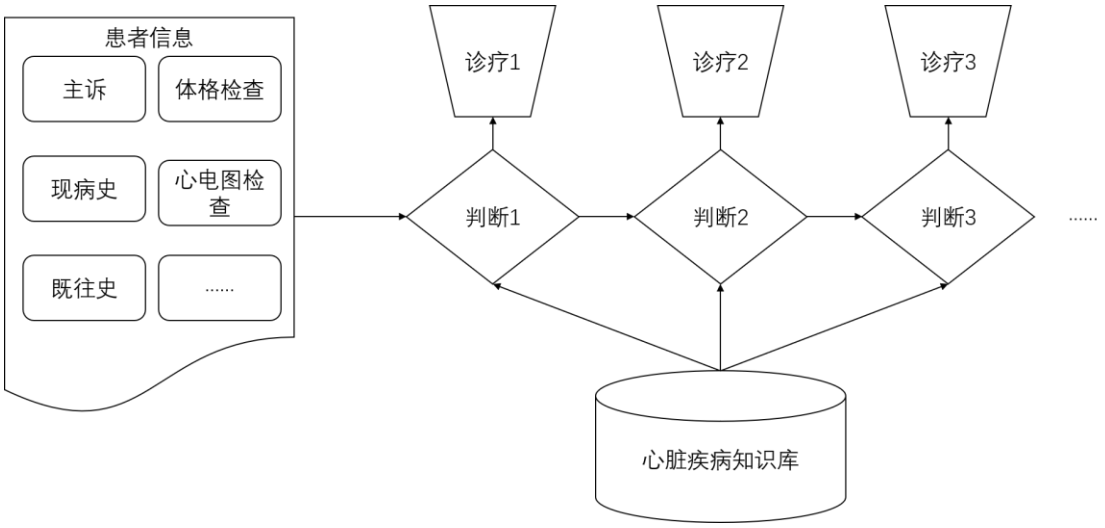


图 1 基于知识库的心脏疾病临床路径示意图

知识库是一种抽象的概念，并没有实际规定以何种形式存储知识。图结构以其可以清晰表示

实体之间的关系和连接的能力，在知识发现、推理和整合等领域得到了广泛应用，这种图结构表示的知识库即知识图谱[错误!未找到引用源。]。

构建心脏疾病知识图谱的流程如图 2 所示。首先需要定义本体及其关系，获取知识，最终构建心脏疾病知识图谱。下面进行说明。

3.1 本体与关系定义

本体的定义主要从两个方面考虑，医学知识和患者信息。医学知识主要包括疾病、症状、病因、药物、手术等，是为后面的医学决策做支撑。患者信息主要包括主诉、症状、检查结果等，是为了将患者嵌入到临床路径中。

医学知识的本体定义需要结合相关疾病，下面表 1 给出心脏疾病的部分本体定义：

表 1 医学知识本体示意表

本体	示例	属性
疾病	冠状动脉粥样硬化性心脏病、急性前壁心肌梗死	描述；分级；
症状	胸痛、胸闷、头晕、乏力	出现频率；出现时长；
部位	心脏、心血管、	
检查项目	身高、体重、脉搏、红细胞、心电图	数值；阳性与否；诊断；
治疗方法	药物、手术	描述；
药品	藻酸双酯钠片、醋酸泼尼松片	服用剂量；服用频率；
.....	.....	.....

由于构建患者信息本体的目的是使其嵌入医学知识本体组成的网络中，因此患者信息的本体应当尽可能与医学知识本体保持一致。在这里，我们只需要创建“患者标识符”一个本体，用以表示不同患者即可。

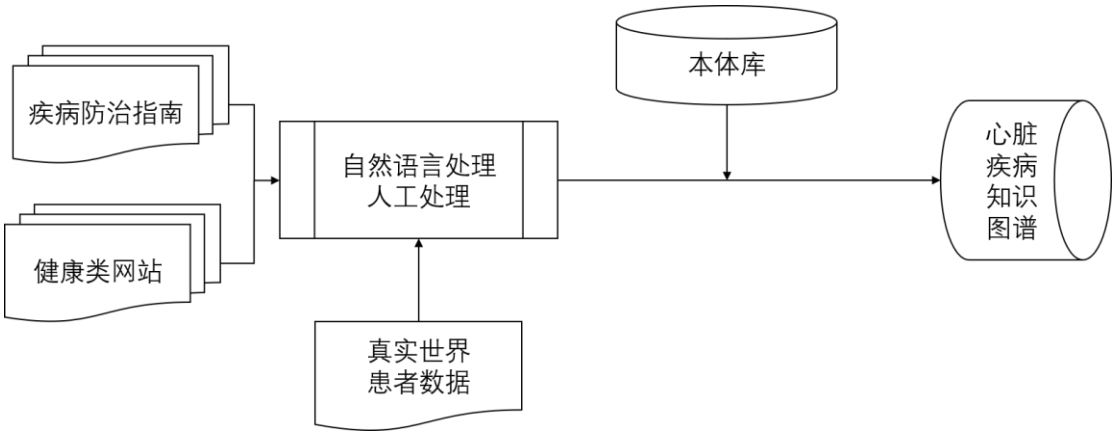


图 2 心脏疾病知识图谱的构建流程图

定义完本体后，需要定义连接这些本体的关系，所有这些定义的本体都至少拥有一条关系链

与其它本体相连接。表 2 列举了一些关系：

表 2 心脏疾病知识库本体关系示意表

本体 A	关系	本体 B
疾病	疾病-症状	症状
疾病	疾病-部位	部位
疾病	疾病-检查项目	检查项目
患者标识符	患者-疾病	疾病
……	……	……

3.2 数据来源与预处理

构建医学知识图谱，需要明确知识来源。由于医学知识对于精确性的要求，因此必须对知识来源进行严格审查。对于心脏疾病，首先参考相应的防治指南，其次可以参考健康类网站。现实临床环境是极其复杂的，因此真实世界患者数据也可以作为构建心脏疾病知识库的信息来源。

对于疾病防治指南，通过在各大学术网站如知网、Pubmed、Google Scholar 以心脏病、指南、heart disease、guideline 等关键词交叉搜索，从中选择获取相应的指南文件。

对于健康网站数据，预先分析网站的 html 代码，找到规则。使用 Python 自带的 requests 包模拟用户浏览器访问行为，爬取 html 代码，根据先前制定的规则自动提取关于心脏疾病的知识。

对于真实世界患者数据，本文通过苏南地区某三甲医院心脏的部分电子病历作为信息源，经过清洗、校对、整理，剔除无效数据后，共得到入院记录及出院记录各 1493 条，共 2986 条。其中，入院记录包括入院时间、患者主诉、病史等 65 种要素，出院记录包括出院时间、手术时间、手术名称等 42 种要素，且所有包含患者个人信息的内容均已隐去，部分记录如图 3、图 4 所示。

患者序号	入院时间	主诉	现病史	既往史
1	2022-10-27 18:33	胸闷1天	患者1天前无明显诱因出现胸闷，伴头晕，无胸痛，无发热，无晕厥，无	平素身体健康。无肝炎、结核或其
2	2022-10-27 10:38	胸闷半天	患者半天前无明显诱因下出现胸闷，伴气急，持续半小时，自诉保心丸	平素身体健康。有肝炎、结核或其
3	2022-10-27 15:12:48	胸痛一月余	患者一月余前无明显诱因下出现胸骨后压榨性疼痛，伴气喘，持续2小时	平素身体健康。无肝炎、结核或其
4	2022-10-31 10:16:06	胸闷胸痛三天，加重半天余	患者三天前无明显诱因下出现胸闷胸痛，位于心前区，呈压榨性疼痛，持	平素身体健康。无肝炎、结核或其
5	2022-10-30 17:14:29	胸闷心慌2年余	患者近两年来每次做完家务活后感胸闷，略有心慌，休息10-20分钟后可	平素身体健康。无肝炎、结核或其
6	2022-10-28 14:32:17	胸闷心悸1周余	患者1周前无明显诱因下出现胸闷心悸，无胸痛，无恶心呕吐，至海门区	平素身体健康。无肝炎、结核或其
7	2022-10-29 19:07:02	胸痛10天	患者10天前无明显诱因下出现胸痛，疼痛位于心前区，向后背部放射，持	平素身体健康。无肝炎、结核或其
8	2022-10-31 23:14:58	胸闷2月余	患者2月前无明显诱因下出现胸闷不适，表现为阵发性胸闷，持续十余分	平素身体健康。无肝炎、结核或其
9	2022-10-28 15:57:48	胸闷14年	患者14年前无明显诱因下出现胸闷气急，于上海长海医院就诊，予支架	平素身体健康。无肝炎、结核或其
10	2022-10-29 15:51:49	反复胸痛3年，加重3小时	患者于3年前始出现反复胸痛。3小时前患者自觉症状加重，胸前区为主，	平素身体健康。无肝炎、结核或其
11	2022-10-26 12:50:17	反复胸痛、胸闷20余年，加重2月	患者20前无明显诱因下反复出现胸痛、胸闷，休息后可以缓解，平时规	既往有高血压病史20余年，收缩压
12	2022-10-24 19:19:25	胸闷伴气急加重一天	患者一天前无明显诱因下出现胸闷的症状，伴呼吸困难，偶有胸痛，可	平素身体一般。无肝炎、结核或其
13	2022-10-31 15:59:17	胸闷3天	患者自诉无明显诱因下出现右侧胸痛一周，位于胸骨后，呈阵发性隐匿	平素身体健康。无肝炎、结核或其
14	2022-10-28 19:43:29	胸闷6小时	患者6小时前无明显诱因下突发胸闷，主要位于胸骨后，持续20分钟后缓	平素身体一般。无肝炎、结核或其
15	2022-10-29 19:24:05	胸痛9天	患者9天前活动后出现胸痛，为胸骨后压榨性疼痛，向后背部放射，持续	平素身体健康。无肝炎、结核或其
16	2022-10-27 06:45:37	胸痛3.5小时	患者3.5小时前无明显诱因下反复发作胸痛，位于心前区，伴肩部牵涉痛	平素身体健康。无肝炎、结核或其
17	2022-10-31 04:26:55	胸痛1小时	患者1小时前突发胸痛胸闷，位于胸骨后方，呈压榨样，持续性，有后背	平素身体健康。无肝炎、结核或其
18	2022-10-27 18:48:56	反复胸闷心悸1年余，加重1周	患者1年前无明显诱因下出现胸闷心悸，伴头晕不适，无头痛，无胸痛，	平素身体一般。无结核或其他传染
19	2022-10-14 12:37:07	突发胸闷气急4天，加重2天	患者4天前活动后出现胸闷气急，伴剑突下疼痛，无头晕黑矇，无晕厥，	平素身体一般。既往40年前有肺结
20	2022-10-28 15:25:05	胸痛8小时余	患者昨夜12点突感胸痛，无放射痛，呈针扎样持续性，发作时有呕吐1次	平素身体健康。无肝炎、结核等，

图 3 入院记录部分截图

患者序号	入院时间	出院时间	疾病编号	疾病名称	是否手术	手术时间	手术编号	手术名称
1	2022-10-27 18:33:39	2022-11-01 08:00:00	I25.103	冠状动脉粥样硬化性心脏病	0		-	-
2	2022-10-27 10:38:25	2022-11-01 09:37:50	I25.103	冠状动脉粥样硬化性心脏病	0			-
3	2022-10-27 15:12:48	2022-11-01 09:54:44	I25.103	冠状动脉粥样硬化性心脏病	1	2022-10-28 00:00:00	36.0700	药物洗脱冠状动脉支架置入
4	2022-10-31 10:16:06	2022-11-01 14:22:45	I25.103	冠状动脉粥样硬化性心脏病	0		-	-
5	2022-10-30 17:14:29	2022-11-02 08:25:00	I21.300x004	急性ST段抬高型心肌梗死	1	2022-11-01 00:00:00	88.5600	用两根导管的冠状动脉造影术
6	2022-10-28 14:32:17	2022-11-02 08:54:53	I21.401	急性非ST段抬高型心肌梗死	1	2022-11-01 00:00:00	88.5600	用两根导管的冠状动脉造影术
7	2022-10-29 19:07:02	2022-11-02 08:55:00	I21.401	急性非ST段抬高型心肌梗死	1	2022-11-01 00:00:00	88.5600	用两根导管的冠状动脉造影术
8	2022-10-31 23:14:58	2022-11-02 14:30:00	I25.103	冠状动脉粥样硬化性心脏病	0		-	-
9	2022-10-28 15:57:48	2022-11-03 07:50:00	I25.103	冠状动脉粥样硬化性心脏病	0		-	-
10	2022-10-29 15:51:49	2022-11-03 08:40:00	I21.004	急性广泛前壁心肌梗死	1	2022-10-29 00:00:00	36.0601	冠状动脉药物涂层支架置入术
11	2022-10-26 12:50:17	2022-11-03 09:02:21	I25.103	冠状动脉粥样硬化性心脏病	1	2022-10-29 00:00:00	88.5500	单根导管的冠状动脉造影术
12	2022-10-24 19:19:25	2022-11-03 09:46:46	I21.401	急性非ST段抬高型心肌梗死	1	2022-10-28 00:00:00	88.5600	用两根导管的冠状动脉造影术
13	2022-10-31 15:59:17	2022-11-03 14:00:17	I25.103	冠状动脉粥样硬化性心脏病	1	2022-11-01 00:00:00	00.6600x008	经皮冠状动脉药物球囊扩张成形术
14	2022-10-28 19:43:29	2022-11-03 15:30:00	I21.401	急性非ST段抬高型心肌梗死	1	2022-11-01 00:00:00	88.5701	多根导管冠状动脉造影
15	2022-10-29 19:24:05	2022-11-04 08:47:29	I21.401	急性非ST段抬高型心肌梗死	1	2022-11-01 00:00:00	36.0601	冠状动脉药物涂层支架置入术
16	2022-10-27 06:45:37	2022-11-04 09:40:03	I25.103	冠状动脉粥样硬化性心脏病	1	2022-10-27 00:00:00	36.0601	冠状动脉药物涂层支架置入术
17	2022-10-31 04:26:55	2022-11-04 09:57:49	I21.001	急性前壁心肌梗死	1	2022-10-31 00:00:00	36.0700	药物洗脱冠状动脉支架置入
18	2022-10-27 18:48:56	2022-11-04 10:01:22	I20.000	不稳定型心绞痛	1	2022-11-04 00:00:00	88.5701	多根导管冠状动脉造影
19	2022-10-14 12:37:07	2022-11-04 11:22:23	I21.004	急性广泛前壁心肌梗死	1	2022-10-25 00:00:00	00.6600x008	经皮冠状动脉药物球囊扩张成形术
20	2022-10-28 15:25:05	2022-11-04 14:46:24	I21.401	急性非ST段抬高型心肌梗死	1	2022-11-01 00:00:00	88.5600	用两根导管的冠状动脉造影术

图 4 出院记录部分截图

获取到这些数据后，需要将这些数据都转变为结构化数据，为后面的构建知识图谱做准备。其中，健康网站信息基本上都是结构化数据，因此可以不需要额外处理。疾病防治指南属于非结构化数据，需要重点处理。真实世界患者数据属于半结构化数据，其中一些信息如检查、疾病名称等为结构化信息，主诉、现病史等属于非结构化数据，需要额外处理。

将非结构化数据转换为结构化数据，主要就是从文本中提取出 3.1 节所定义的本体、属性、本体关系。目前有人工抽取与自动化抽取两种方式。人工抽取就是通过人工阅读相关资料，从中提取本体等信息，但是人工抽取不适合数据量多的情况，这种情况下需要用到自动化抽取。自动化抽取就是利用自然语言处理技术，自动地从文本中抽取相对应的信息。自动化抽取技术最初需要分别训练实体、关系和实体属性的模型，然后将它们的结果进行整合[13]。然而，随着预训练语言模型的兴起，通过在这些模型上进行微调，可以一次性从数据中提取三元组，减少了对大量训练数据的需求[14]。特别是自 2022 年 10 月底发布的 OpenAI 的 ChatGPT 这一大型语言模型，以其卓越的自然语言理解能力备受研究人员青睐。通过将数据抽取转化为问答任务并将相关信息输入给大语言模型，它能够学习少量样本并输出预期的三元组[15]。

3.3 心脏疾病知识图谱的构建

构建心脏疾病知识图谱就是将前面提取出的本体，以(本体 A，关系，本体 B)的三元组形式，存储到图结构数据库中。目前主流的图结构数据库有 neo4j 等。Neo4j 属于一类高性能的图数据库，它为用户提供了强大的图形储存和查询功能。通过它以图形的形式来组织数据，其中节点代表实体，关系则代表实体之间的连接。使用 Neo4j，用户可以轻松存储、检索和分析复杂的关联数据，从而发现数据中隐藏的模式和见解。Neo4j 的灵活性和可扩展性使它成为构建智能应用和推理系统的理想选择，能为用户提供更深入的数据洞察力和决策支持。

本文的心脏疾病知识图谱的构建不仅使用了防治指南和健康网站的数据，还利用了真实世界患者数据。在真实世界患者数据中，一些信息可能未被完全包含在防治指南和健康网站的数据中，例如疾病相对应的症状、治疗方法等，在指南或网站数据中有缺失情况。

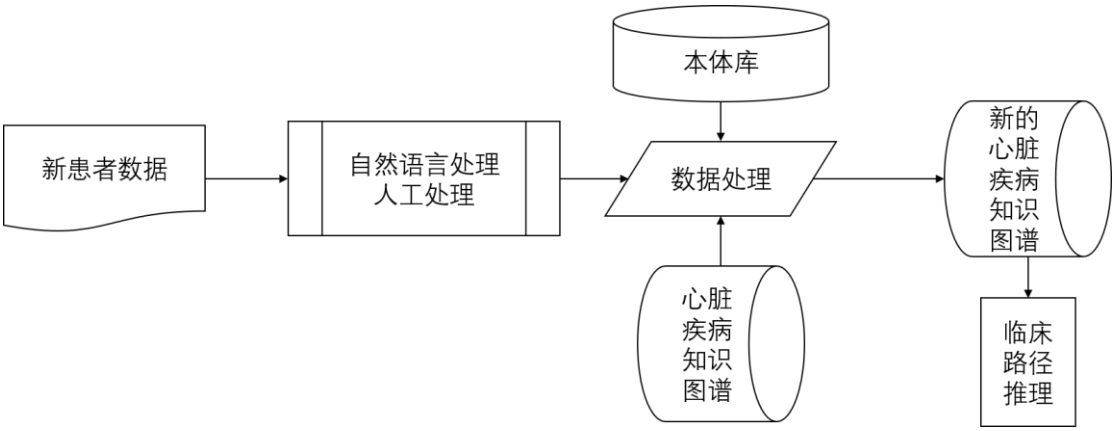
4 人类心脏疾病知识库的应用

构建好心脏疾病知识库后，可以通过编写图结构的相关规则来进行推理。当新入院一位心脏疾病患者后，通过第 3 节的方法，将新患者信息以图结构表示出来，嵌入构建好的心脏疾病知识图谱中。利用这个新的知识图谱，就可以为这位新患者进行临床路径的推理。图 5 展示了这个过程。

1) 疾病诊断。心脏疾病通常需要及时的诊疗，但这需要医生丰富的经验才可以快速准确的判断患者情况，做出合适的治疗。这时候可以利用先前采集的真实世界患者信息来训练一个疾病诊断的模型。通过对症状、检查结果等变量进行重采样，以模拟出真实世界中信息不足的情况，再使用图神经网络训练参数[16]。将新患者的各项数据转换为图结构，输入训练好的图神经网络模型中，则可以得到预测的疾病，辅助医生诊断。

2) 治疗建议。当一位新患者被确诊为某疾病，我们可以从心脏疾病知识图谱中找到与该疾病存在关系的节点，得到相对应的治疗建议。同时，也可以通过训练图神经网络的方式得到更加个人化的治疗建议。

图 5 心脏疾病知识图谱的应用流程图



5 总结

本文探讨了心脏疾病知识库的构建与应用，提出使用知识图谱存储心脏疾病的知识，并利用规则或图神经网络来应用心脏疾病知识。特别地，本文结合了真实世界患者数据描述了这个过程。但是本文没有从个性化治疗的角度思考问题，也没有能将最新期刊上的治疗方法纳入这个知识图谱中，需要更进一步地研究。

OR-017

新媒体环境下我国居民健康素养的提升研究与探讨

文怡力[1] 任慧玲[1] 黄友良[2]
1.中国医学科学院医学信息研究所
2.北京中医药大学管理学院

健康素养是指个人获取和理解基本健康信息和服务，并运用这些信息和服务做出正确决策，以维护和促进自身健康的能力[1]。健康素养是国民健康重要的评价指标，而国民健康是国家可持续发展的重要标志。党的十八大以来，以习近平同志为核心的党中央强调把人民健康放在优先发展的战略位置，把提升健康素养作为增进全民健康的前提，为提高全民健康水平做出了制度性安排[1]。2016 年，《“健康中国 2030”规划纲要》[2]明确提出“到 2020 年，居民健康素养水平达到 20%”的目标；2019 年,《健康中国行动（2019-2030 年）》[3]明确指出“提升公众健康素养是健康中国行动的重要工作内容和主要考核指标之一”。提高居民健康素养是实施健康中国战略、推进健康中国行动、抗击新冠肺炎疫情等工作的重要内容和工作目标。

本文主要研究新媒体环境下我国居民健康素养水平发展现状及特征，深入分析新媒体环境下我国居民健康素养研究现状及面临的问题，并提出有益于我国居民健康素养水平提升的发展对策与建议，为进一步研究提供借鉴和参考。

1. 新媒体环境下我国居民健康素养的内涵及提升意义

健康素养划分为基本健康知识和理念素养、健康生活方式与行为素养和基本技能素养三个方面[4]。以公共卫生问题为导向，健康素养亦可划分为六类健康问题素养，分别为科学健康观素养、传染病防治素养、慢性病防治素养、安全与急救素养、基本医疗素养和健康信息素养。健康素养能在一定程度上预测健康状况，健康素养水平高的个体，获取和辨别健康信息的能力更强，自我保健的能力更强，健康状况更为良好。

相比于传统媒体，新媒体的特点主要体现在：一、新媒体具备数字化和多媒体化的特点。数字化主要表现在信息的采集、加工、存储、管理与传递过程中的数字化转换。多媒体化主要表现在借助文本、图片、声音和视频等多种媒体形态，可以增强信息的感染力，在传播过程中给予用户全感官的体验，有助于加深该信息给予用户的整体形象。二、新媒体在信息传播上具有互动性。新媒体的互动性在于新媒体用户可拥有信息生产者、发布者和接受者等多重身份。三、新媒体在传播方式上具有碎片化特点。碎片化可表现在时间的碎片化，用户可随时随地获取或发布其想要的信息。也可体现在传播内容上，在微博、短视频等呈现方式下，新媒体传播的信息内容更加零碎[5]。

通过健康素养的监测结果可预测群体健康状况，提升健康素养水平可改善群体健康状况、增加人均预期寿命。提升健康素养水平也可有效降低社会卫生资源不平衡的现象，从而显著减少社

会经济成本。新媒体拥有的特点使信息得到更为快速、更为有效的传播。利用新媒体的巨大优势，通过视频、图像等传播手段科普健康知识、促进健康信息交流，可以有效的提升健康意识，改变其健康行为，提高健康水平和生活质量，最终促进国民健康发展[5]。

2. 我国居民健康素养的研究现状

2.1 我国居民健康素养水平的现状

居民的健康素养水平具有可监测性，可依据稳定的监测系统加以确定评价[1]。自 2012 年起，我国国家卫生健康委员会（原卫生计生委）于每年的 8 月-11 月在大陆（不包括港、澳、台地区）31 个省（自治区、直辖市）的 336 个监测点进行全国城乡居民健康素养监测工作。居民健康素养监测对居民健康素养水平的评价包含三个部分：一是评价我国城乡居民健康素养的总体水平；二是基于“知识-行为-技能”理论，从基本健康知识和理念素养、健康生活方式与行为素养、基本技能素养三个方面评价居民健康素养水平；三是从公共卫生问题的视角，从科学健康观素养、传染病防治素养、慢性病防治素养、安全与急救素养、基本医疗素养和健康信息素养六类健康问题评价居民健康素养水平[6]。本研究收集并分析 2012-2020 年共 9 年居民健康素养监测结果数据，如图 1 所示。

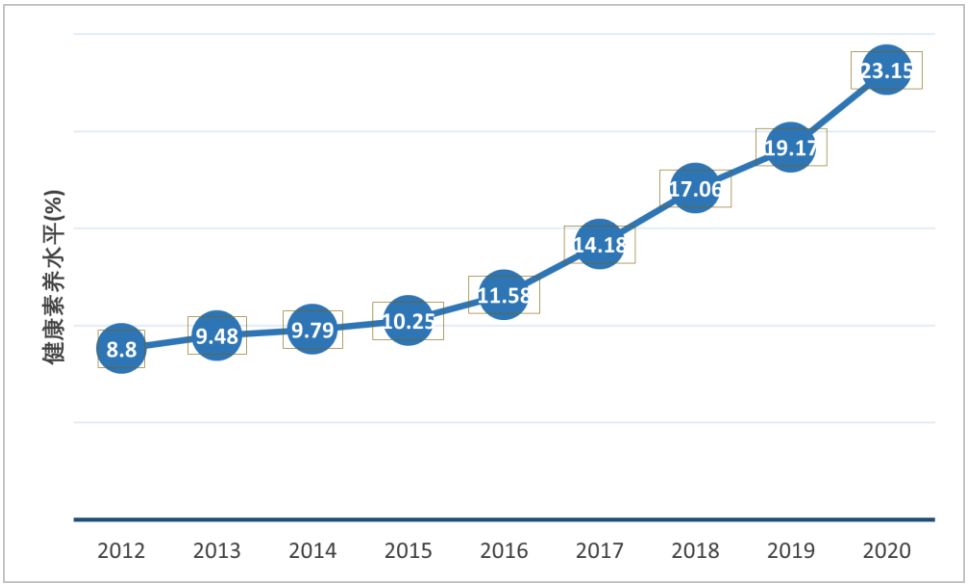


图 1 2016-2020 年我国居民健康素养总体水平

从我国城乡居民健康素养总体监测数据可得，我国城乡居民健康素养的总体水平稳步提高。从 2012 年的 8.80%到 2020 年的 23.15%，提升 14.35 个百分点。实现了《全民健康素养促进行动规划（2014-2020 年）》中提出的“到 2015 年，全国居民健康素养水平提高到 10%”的目标[7]，并在 2020 年实现了《“健康中国 2030”规划纲要》提出的“到 2020 年，居民健康素养水平达到 20%”的目标[3]。通过对比可得，2019-2020 年增长速度为历年最高点，增长率高达 163.07%，此结

果一方面反映出健康素养促进行动等专项健康素养促进项目的干预措施有了成效，另一方面与疫情的发展使居民健康意识提高有关[8]。

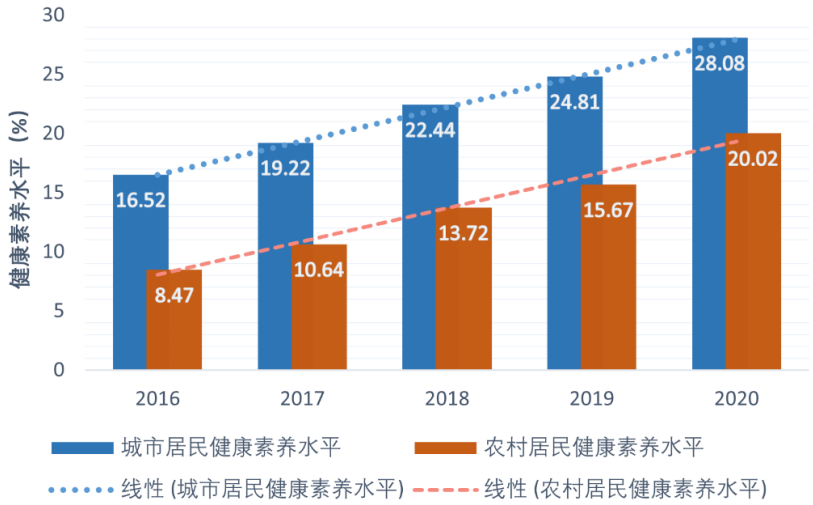


图 2 2016-2020 年我国城乡居民健康素养水平

生产力分布特点和市场机制导致了我国城乡之间的生产力发展不均衡，也拉开了城乡之间教育、医疗和收入之间的差距[9]。如图 2 所示，从城乡方面来看我国城乡居民健康素养水平都稳定增长，但城乡居民健康水平差异巨大。2016-2019 年，我国城乡居民健康素养水平差异呈逐年扩大趋势，从 2016 年健康素养水平相差 8.05 个百分点到 2019 相差 9.14 个百分点，差异增长了 1.09 个百分点。然而到 2020 年，城乡居民健康素养水平差异缩小到 8.05 个百分点，又呈缩小趋势。这与 2020 年受突如其来的新冠疫情影响有关，政府相关部门和媒体等加大宣传疫情防控，使防范疫情的意识深入人心[8]。

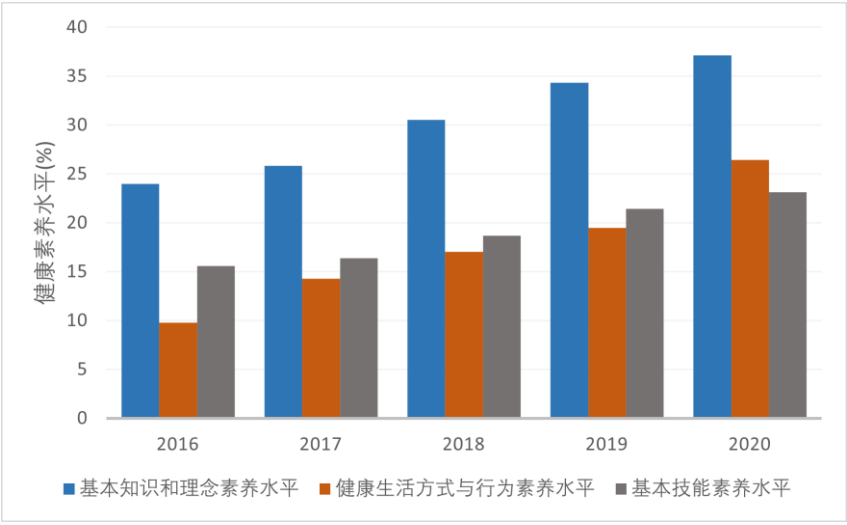


图 3 2016-2020 年我国居民健康素养水平的“三个方面”

如图 3 所示，尽管我国居民健康素养水平的三个方面均在逐年提高，然而基本知识和理念素养水平都远高于其他二者的水平。健康生活方式与行为素养水平提升速率最高，5 年间增长率为

124.00%，其次是基本知识和理念素养水平和健康生活方式与行为水平，5 年间增长率分别为47.21%和 43.29%。这反映出健康知识与健康行为的脱离，但随时间的发展这一现象逐渐有了好转。

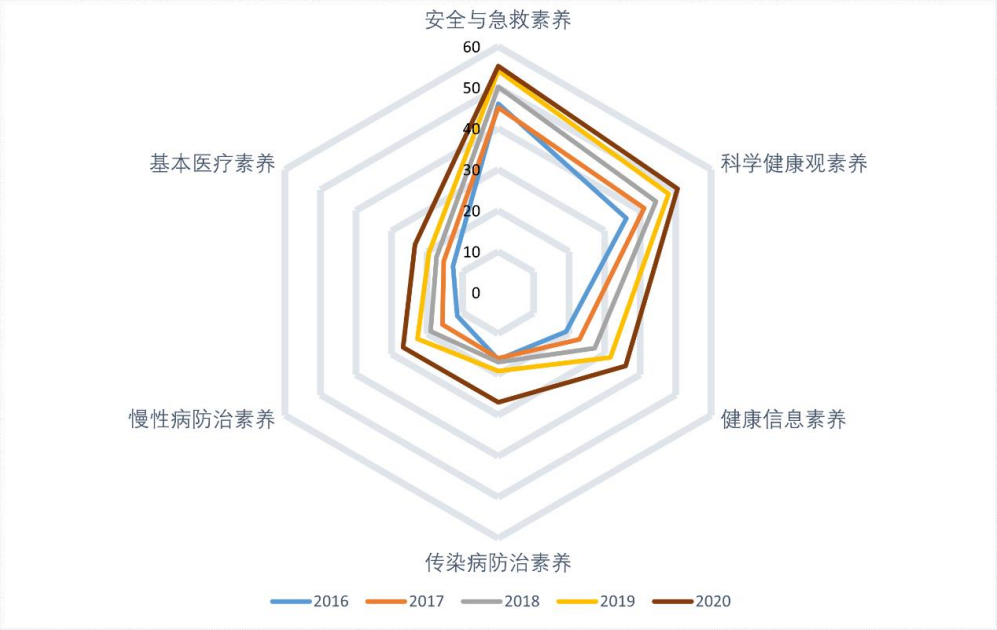


图 4 2016-2020 年我国居民健康素养水平的“六类健康问题”

通过六类健康问题的维度分析可得我国居民六类健康问题发展不均匀。如图 4 所示，主要表现为，安全与急救素养和科学健康观素养较强，而其他四类较弱，其中基本医疗素养最弱。值得一提的是，2020 年因新型冠状病毒疫情的影响，居民健康素养水平在传染病防治素养方面有了较大的提升。这提示了我国在健康教育方面、健康传播推广方面可以有针对性的选择意识薄弱的问题进行教育推广，最终达到六个方面全面发展。

表 3 2016-2020 年我国居民健康素养水平的“六类健康问题”

健康素养问题	年份				
	2016	2017	2018	2019	2020
安全与急救素养	46.00%	45.09%	50.18	54.11%	55.23%
科学健康观素养	36.18%	41.12%	44.48	48.07%	50.48%
健康信息素养	19.13%	22.95%	27.18	31.66%	35.93%
传染病防治素养	16.38%	16.06%	17.05	19.21%	26.77%
慢性病防治素养	11.48%	15.71%	18.96	22.73%	26.73%
基本医疗素养	12.76%	15.34%	17.38	19.43%	23.44%

2.2 我国居民健康素养的特征

通过研究对比发现，从时间尺度来看，我国居民健康素养在各方面都呈上升趋势。从空间尺度来看，我国城乡居民健康素养水平差异较大，且我国东部地区健康素养水平最高，其次为中部

地区，最后为西部地区。从结构上，居民健康素养的发展还在表现基本知识和理念素养远高于健康生活方式与行为素养水平和基本技能素养水平。从人口学特征来看，性别与健康素养之间的联系不大，老年人群健康素养水平明显落后于全国同期平均水平，而文化程度和经济状况与健康素养水平呈正比[10]。居民健康素养水平也与婚姻状况直接相关，一般未婚与已婚的人群健康素养水平较高。与发达国家相比，我国居民健康素养整体水平偏低[10]，因此提升我国居民健康素养已迫在眉睫。

3. 新媒体环境下我国居民健康素养提升面临的困境

3.1 发展不平衡导致健康传播存在“知沟现象”

美国第奇纳为主的传播学者提出“知沟现象”，即对于社会经济地位越高通常获得信息速度越快，随着媒介不断传送信息，不同社会经济地位之间的知识鸿沟不断扩大。而在新媒体时代，健康传播领域此现象也不断被印证[11]。我国当前社会发展不平衡，这可以体现在城乡差异巨大、区域发展不均等情况上。居民健康素养受多方面因素影响，经相关研究表明，其影响因素有文化水平程度、经济收入、性别和年龄等，其中文化程度和经济收入与健康素养水平呈正相关，而年龄与健康素养水平呈正相关[11]。

3.2 健康传播缺乏有效性导致健康知识与健康行为脱节

据相关调查显示，有相当部分的人群了解健康知识，而在其日常生活行为中却往往不采取健康行为，这印证出其健康知识与健康行为脱节[11]。健康知识与健康行为的脱节亦可体现于我国居民健康素养监测结果中。历年以来，我国居民基本知识和理念素养水平均远高于健康生活方式与行为素养水平以及基本技能素养水平，这体现出“知”与“行”的不一致。究其原因，与健康传播方式不合适而导致其缺乏有效性有关。大学生的健康素养水平总体超过居民平均水平，但大学生健康素养的三个方面掌握情况并不均衡，基本知识和理念具备率较高，而健康生活方式与行为和基本技能具备率较低，存在一定的知行分离，表明大学生将较高健康知识理念转化为相应水平的健康技能与行为方面还有滞后。此类现象，在我国居民健康素养监测中较为常见，评价和干预健康行为的研究还有待加强。

3.3 健康信息评估体系的缺少导致信息质量良莠不济

新媒体传播具有即时性和无障碍性，其参与模式较为自由，对各类消息发布与传播的限制较低，无法杜绝错误的信息掺杂其中。由于通信技术的发展,利用移动新媒体，用户可以随时发布、传递和获取信息，信息更迭更为迅速，致使信息爆炸式增长，而在缺乏信息评估体系的情况下，各类信息泥沙俱下，信息质量良莠不齐。新媒体平台上往往言论自由约束较少，而在健康信息、健康科普等领域又缺乏专业人士，这使得新媒体平台成了滋生大量健康科普的谣言。这类低质量的健康信息又往往以博人眼球的方式出现，激发人焦虑恐惧等负面情绪，使得其比一般健康科普消息更易传播，给大众造成危害。

3.4 政府监督缺失导致市场失序

相对于传统媒体信息从选题策划到刊发都有着严格的审核发布流程，新媒体平台发布信息的审核发布流程，则大部分都为机器审核发布，部分为人工发布。当其涉及到专业领域，尤其是卫生健康领域里，相关审核流程便具有局限性[13]。对于卫生健康领域的信息，若其信息不够规范、健康科普不够专业，则会对民众产生误导。尤其是中医药科普领域，一些人打着宣传中医药的名义，传播一些封建迷信的偏方，看似宣扬卫生健康知识，实则污名化中医。同样，在新冠肺炎疫情发展初期，不少严重缺乏科学依据预防新冠的知识，经过相关平台广泛传播后，不断误导民众，造成不良影响，为相关部门疫情防控工作增加负担。

4. 新媒体时代我国居民健康素养提升路径

4.1 利用新媒体技术对不同人群实施健康信息精准传播

新媒体时代信息传播更具有个性化，用户可根据个人偏好选择不同的信息内容，平台还可根据用户个性化的信息,进行精确传送。新媒体技术与大数据等新科技的融合运用,为健康资讯的精准传递提供了强大的技术支持。根据传播学理论，越精准的传播越有效[14]，新媒体平台利用算法，收集用户生活习惯及个人偏好，刻画用户像画，针对其关注的健康知识，实施精准的健康传播，以满足各类人群的需要。未来还可根据疾病的时空特征，分析用户的集体群像，预测当前地区所在人群易患疾病，并实现健康信息精准传播，从而减少其患疾病的概率，并提升其健康素养水平。

4.2 采用有效的健康传播策略，优化健康行为

在新媒体时代，信息的表达方式逐渐多元化,信息的表达模式也更加形象化和立体化,在向用户传递信息的同时给予用户以全感官式的感受,有助于使用户对信息的印象更加深刻。因而新媒体时代健康信息的呈现方式远比传统媒体的表述更易让人接受。公众会出于健康信息混杂难以分辨、避免增加心里压力和认为自己无法采取行动预防等原因，对一些健康信息产生回避行为[11]。为了避免公众的回避行为，提高健康传播的有效性，在新媒体时代，针对不同人群特点，可结合新颖的传播方式进传播。在健康促进与健康教育中，可发挥新媒体的优势，将健康知识与大众心理等内容融合，形成对健康信息的通俗化和趣味性解读。使健康传播做到对高层次用户与低层次用户的兼顾，从而增强健康信息传播的有效性，进而提升用户的健康素养水平。另外，在传播策略上也可采用公开承诺法、奖励关键行为策略、同伴教育策略风险沟通策略、揭露阴谋策略、简单量化策略等，进而优化公众健康行为。

4.3 建立健康信息评估的标准及体系

由于新媒体时代信息具有动态性、开放性、自由性及其多样性等特征，健康信息的评估难以形成标准体系。目前对于新媒体健康信息主要采用机器评估，部分采用人工评估。而新媒体时代

信息更新速度极快且又以文字、声音和视频等多种传播形式出现，使得评价工作更为繁重，因此需要建立一套适合我国的健康的信息评估标准及体系。对于健康信息评估指标，应结合定量与定性指标，以减少评价的不稳定性和主观性，使评价指标更加规范科学[15]。对于健康信息评价主体，应选择医学权威机构、政府机关、医学专业人士及用户。同时，对于健康信息评估结果应及时公布，宣传错误信息的运营主体应给予处罚，这样既能使用户懂得辨别错误信息，又能督促运营主体发布高质量健康信息，还可使信息评估更加透明化，有益于健康信息质量的提升。

4.4 加强新媒体健康传播监管

目前，我国对于新媒体健康传播监督机制不够健全，致使一些医学类谣言大行其道。要建立引导民众健康素养提升的专业机构，建立健全的卫生健康类知识信息的发布制度，并完善信息发布平台。政府部门要依托专业机构加强对新媒体健康传播活动进行监测和评估，对于宣传错误虚假的卫生健康信息的平台应给予处罚，并实时监控民众舆情，有效制止谣言的散布，以权威的态度及时辟谣。而对于健康类科普信息缺乏的问题，政府要扶持健康领域的专业团队。要出台更多的倾斜政策,引导更多专业人士投身宣传卫生健康信息的热情当中,使更多专业规范的卫生健康信息出现在人民群众视线当中，助力全民健康素养水平的提升。

5. 结论

提升居民健康素养是提高全民健康的有力保证，国民健康是国家可持续发展的重要标志。新媒体环境下，应充分利用新媒体技术对不同人群实施健康信息精准传播，采用有效的健康传播策略，优化健康行为，建立健康信息评估的标准及体系，加强新媒体健康传播监管，发挥新媒体在健康促进、健康教育等方面的巨大作用，有效提升我国居民健康素养水平。充分利用新媒体的优势，引导公众树立健康正确的价值观，养成健康的生活习惯，积极营造全社会关注学习健康知识的氛围。

OR-018

基于 CT 生成的 MRI 与常规 MRI 之间的一致性定量评估

李兆同，吴辉群
南通大学医学院

1 引言

通过 CT 数据生成的伪 MRI（synthetic MRI, sMRI）图像可以提供比常规 CT 图像更详细的病理结构信息。在一些只能进行 CT 图像扫描的特定医疗场景中，sMRI 的生成受到越来越多的关注，常规的 MRI 和生成的 sMRI 的一致性评价也越来越受到重视。普通的图像质量评价函数，如 RMSE、SSIM 和 PSNR 等，量化了真实图像与生成图像的相似性，但却无法说明生成图像在临床实践中的可行性。而通过双盲、随机抽样和一致性检验分析，可以有效地证明生成图像的临床有效性和适用性。近年来，在由标准 MRI 生成的伪 CT（synthetic CT, sCT）图像的工作中[1-12]，通过一致性统计检验的方法考察了基于 MRI 的生成 sCT 与常规 CT 在定量评价方面的等效性[13, 14]。实验结果表明，使用深度学习网络可以生成可靠的 sCT 图像，基于 MRI 的生成 sCT 与常规 CT 在临床评估中拥有一致性。

近年来，研究人员试图使用神经网络从 CT 图像生成 sMRI 图像。sMRI 可以在无法进行 MRI

检查的情况下，为单次 CT 扫描提供了更丰富的影像信息，满足了临床治疗中辅助医生 MRI 检查的需求。虽然已经有一些从 CT 图像生成 sMRI 研究的发表[8, 12, 15, 16]，但评价结果仅仅局限于基于参考图像的图像质量评价函数来评估 sMRI 的生成质量，无法说明生成图像可以满足临床应用中的需求，常规脑 MRI 和生成脑 sMRI 之间形态结构的等效性和一致性是目前急需探索的问题。针对以上问题，本研究在这方面做了初步的尝试，并通过此研究进一步评估 sMRI 图像生成质量，为 sMRI 在脑外科手术规划和临床管理中的应用提供了可能。

2 材料与方法

2.1 被试者筛选

在本研究中，为了定量评估 sMRI 与常规 MRI 在脑形态方面的一致性，我们共筛选了 2017 年 1 月 1 日至 2019 年 12 月 30 日来深圳市第二人民医院就诊的头痛患者 50 例。纳入和排除标准如下：

纳入标准：（1）年龄在 18 - 60 岁之间；（2）无高血压、糖尿病；（3）头部无外伤；（4）无大脑外科手术史。

排除标准：（1）怀孕；（2）图像质量较差；（3）有脑肿瘤、出血等改变脑实质形状的疾病；（4）MRI 禁忌症；（5）金属植入物及影像学检查不全。

2.2 MRI 和 CT 采集协议

采用 3D MPRAGE 序列从横截面方向获取 T1WI 图像，TR=1,600 ms（这里的 TR 指两个非选层 180° 反转脉冲之间的 MPRAGE$_{TR}$ 时间间隔[17]），TE =3.37 ms，体素大小=1×1×1 mm^3，图像大小=256×256，Acquisition Range=160 mm。由 GRAPPA 在 R-factor=2 的 1.5T Avanto 扫描仪(Siemens)上获得。

CT 图像的具体采集参数为：120 kV，330 mA，曝光时间=500 ms，像素大小=0.5×0.5×1 mm^3，图像大小=512×512，采用 Siemens 公司的 SOMATOM Definition Flash 获取。

2.3 sMRI 生成

利用已经发表 DRU-Net-101 网络[16]，从 CT 数据中生成 sMRI 图像。该网络是由 ResNet 和 U-Net 组成，如图 1 所示。其中 U-Net 负责对 CT 数据进行特征稀疏和 sMRI 生成估计，并随着 U-Net 层数的增加，ResNet 保证了训练过程中的梯度有效更新，防止梯度爆炸。同时，Contextual Loss 函数的引入保留了 MRI 中颅内组织的高频信息，去除了低频背景噪声。

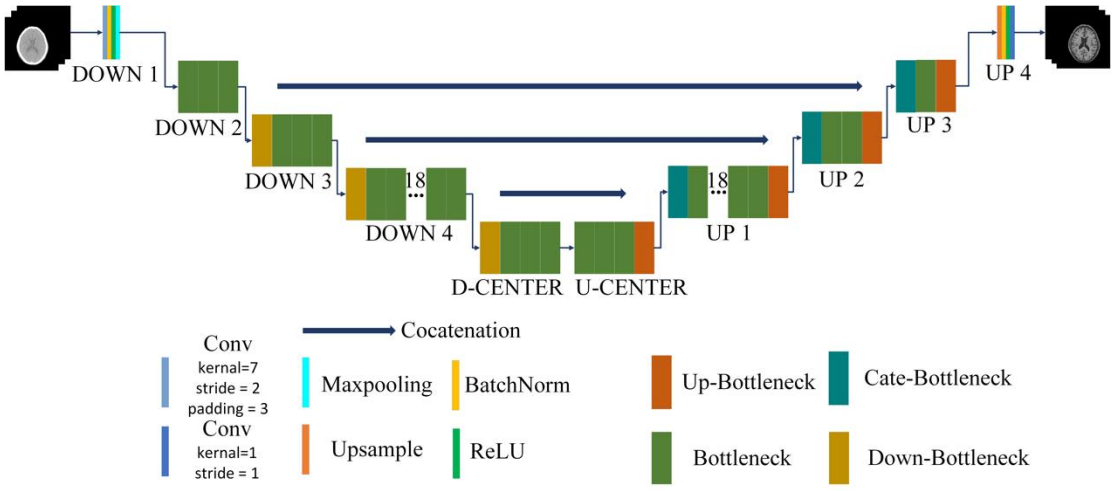


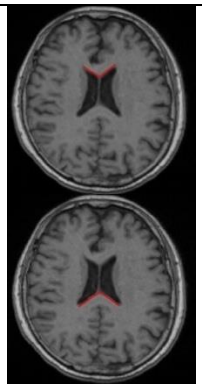
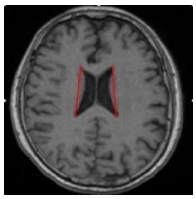
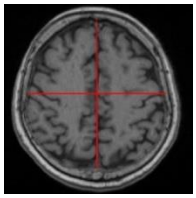
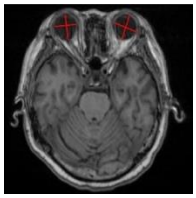
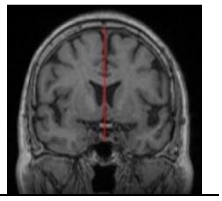
图 1 sMRI 生成网络结构图

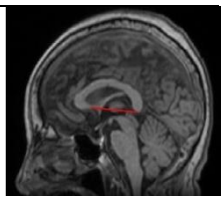
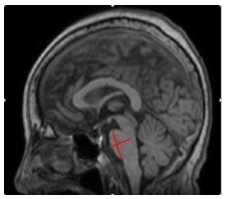
2.4 放射测量参数

在本研究中，我们聘请了具有多年经验的放射科医师，利用 ITK-SNAP 工具（3.8 版）测量 MRI 和 sMRI 图像中不同组织之间的角度和距离。我们分别在矢状位、横断位和冠状位图像中，共

确定了 14 个关于角度和距离的脑组织影像学参数。详细的测量说明如表 1 所示。

表 1 图像测量参数说明，包括测量方法、平面、切片和图示

编号	名称	方法	平面	切片	图示
1	前额角角度（Included angle of anterior ventricular angle，IAAVA）	两侧脑室前角形成的夹角	横断位	通过最大的一层透明隔断	
2	后额角角度（Included angle of posterior ventricular angle，IAPVA）	双侧脑室后角形成的夹角	横断位	通过最大的一层透明隔断	
3	右侧脑室前后角距离（Right lateral ventricle anteroposterior angle distance，RLVAAD）	右侧脑室前后角最大距离	横断位	通过最大的一层透明隔断	
4	左侧脑室前后角距离（Left lateral ventricle anteroposterior angle distance，LLVAAD）	左侧脑室前后角最大距离	横断位	通过最大的一层透明隔断	
5	半卵圆中心层面长（Intracranial length，IL）	颅内最长距离	横断位	穿过半圆形的中心层	
6	半卵圆中心层面宽（Intracranial width，LW）	颅内最宽距离	横断位	穿过半圆形的中心层	
7	右侧眼球宽（Right eyeball width，REW）	右眼最大宽度距离	横断位	通过眼球最大横截面	
8	右侧眼球长（Right eyeball length，REL）	右眼最大长度距离	横断位	通过眼球最大横截面	
9	左侧眼球宽（Left eyeball width，LEW）	左眼最大宽度距离	横断位	通过眼球最大横截面	
10	左侧眼球长（Left eyeball length，LEL）	左眼最大长度距离	横断位	通过眼球最大横截面	
11	颅内高度（Intracranial height，LH）	脑垂体窝最低点到颅骨内侧板的距离	冠状位	通过最大限度显示垂体窝水平	

12	胼胝体膝部与压部距离（Distance between knee of corpus callosum and pressing part，DKCCPP）	胼胝体膝部至按压部位最低点的连接距离	矢状位	中矢状切片	
13	桥脑长（Pontine length，PL）	从桥的最高点到最低点的距离	矢状位	中矢状切片	
14	桥脑宽（Pontine width，PW）	桥的最前边到最后边的距离，并与桥的长度垂直	矢状位	中矢状切片	

2.5 统计分析

通过测量表 1 中所示的 14 个 sMRI 和 MRI 影像学测量参数，并统计分析了方法间（Inter-Method）的可靠性和变异性，对 sMRI 和 MRI 进行等效评估。

影像学测量数据的统计信息包括均值、标准差、差值均值、差值标准差、95%的一致性界限、组内相关系数（Intraclass Correlation Coefficient，ICC）和测量标准误（standard error of measurement，SEM）[18]。95%一致性界限度表示两个测量值之间的差值均值±1.96 标准偏差。de Vet 等人提出通过计算 ICC 和 SEM 来估计测量值之间的一致性和可靠性[19]。采用双向随机效应/绝对一致性 ICC（A,1）计算观察者内、观察者间和方法间的信度，其中，ICC 值大于 0.75 表示一致性好，0.40 ~ 0.75 为一致性较好，小于 0.4 为一致性差。所有统计分析是在 SPSS Statistics Version 27 和 R Version 4.2.1 的环境下进行的。

使用 Bland-Altman 方法分析 sMRI 和 MRI 两种图像测量值的差值与平均值之间的关系，达到评估两种测量结果是否存在一致性的目的。

3 实验结果

3.1 方法间的可靠性和变异性

sMRI 和 MRI 在测量分析中表现出相似的可靠性和变异性，测量统计结果如表 2 所示。对于平均值较大的影像学测量参数，如 RLVAAD、LLVAAD，95%一致性界限较窄，一致性界限小于两次测量结果均值最小值的 1/2，一致性良好[20]。此外，所有测量的 ICC 的值大于 0.4，表明 sMRI 和 MRI 之间测量的一致性较好，但个别较高的 SEM 值表明可靠性较低。sMRI 与 MRI 之间具体的等效性检验将在下一小节通过 Bland-Altman 图进行说明。

表 2 方法间的可靠性和变异性

测量参数	均值		差值		ICC	SEM
	sMRI	MRI	均值	95% LoA		
IAAVA (°)	102.157 ±13.848	100.871 ±12.793	1.286 ±3.857	(-6.273, 8.845)	0.954	2.727
IAPVA (°)	89.343 ±13.297	89.386 ±13.219	-0.043 ±3.025	(-6.325, 6.239)	0.971	2.266
RLVAAD (mm)	62.886 ±11.237	62.886 ±11.237	-1.671 ±12.270	(-25.719, 22.377)	0.466	8.676
LLVAAD (mm)	64.029 ±12.199	63.557 ±11.692	0.471 ±2.351	(-4.137, 5.080)	0.980	1.663
IL (mm)	140.757 ±5.412	140.614 ±5.443	0.143 ±2.267	(-4.300, 4.586)	0.914	1.603
IW	120.743	120.743	0.000	(-4.648,	0.941	1.677

(mm)	±6.487	±7.162	±2.371	4.648)		
IH (mm)	94.143 ±5.822	93.314 ±6.126	0.829 ±2.036	(-3.161, 4.818)	0.934	1.440
REW (mm)	22.700 ±1.108	22.943 ±1.512	-0.243 ±1.245	(-2.682, 2.196)	0.553	0.880
REL (mm)	22.757 ±1.209	23.229 ±1.515	-0.471 1.139	(-2.703, 1.760)	0.621	0.805
LEW (mm)	22.543 ±1.059	22.643 ±1.425	-0.100 ±1.065	(-2.188, 1.988)	0.641	0.753
LEL (mm)	22.529 ±1.442	23.071 ±1.333	-0.543 ±1.099	(-2.697, 1.612)	0.641	0.777
DKCCPP (mm)	42.771 ±5.220	44.786 ±5.102	-2.014 ±2.743	(-7.390, 3.361)	0.800	1.939
PL (mm)	27.486 ±2.631	27.643 ±1.964	-0.157 ±1.807	(-3.699, 3.384)	0.698	1.278

3.2 等价评估

在图 2 的 Bland-Altman 分析中，通过分析两种图像之间基于差值和平均值的各点在一致性界限内的分布可知，绝大部分测量点都在 95%的置信区间内，这表明 sMRI 和 MRI 之间具有良好的一致性，且偏差与测量参数的多少无关。

4 结论

在本次研究中，我们比较了基于 CT 生成的 sMRI 图像和常规 MRI 图像的在临床定量评估中的等效性。通过采集临床 CT/MRI 数据对，并利用 DRU-Net 模型生成 sMRI 图像。随后确定 14 个 MRI 影像学测量参数，统计方法间的可靠性和变异性，对 sMRI 和 MRI 进行等效评估。影像学测量数据的统计信息包括均值、标准差、差值均值、差值标准差、95%的一致性界限、ICC、SEM、Bland-Altman 图和等价图等。通过以上分析，我们证明了基于 CT 生成的 sMRI 和常规 MRI 在定量评估大脑形态方面的等效性，同时也为仅接受了 CT 检测的患者提供了更多的病变信息，简化了临床工作流程。常规 MRI 和合成 sMRI 的在临床工作中的等效性是值得关注的，特别是在评估术前脑肿瘤或病变组织时，大脑形态测量的定量评估在临床实践中显得至关重要。对于医生来说，这种基于 CT 的生成大脑 sMRI 图像不仅为单个 CT 扫描提供了更多的影像信息；对于患者来说，这不仅节省了患者的检查成本，而且避免了长时间的 MRI 检查等待。

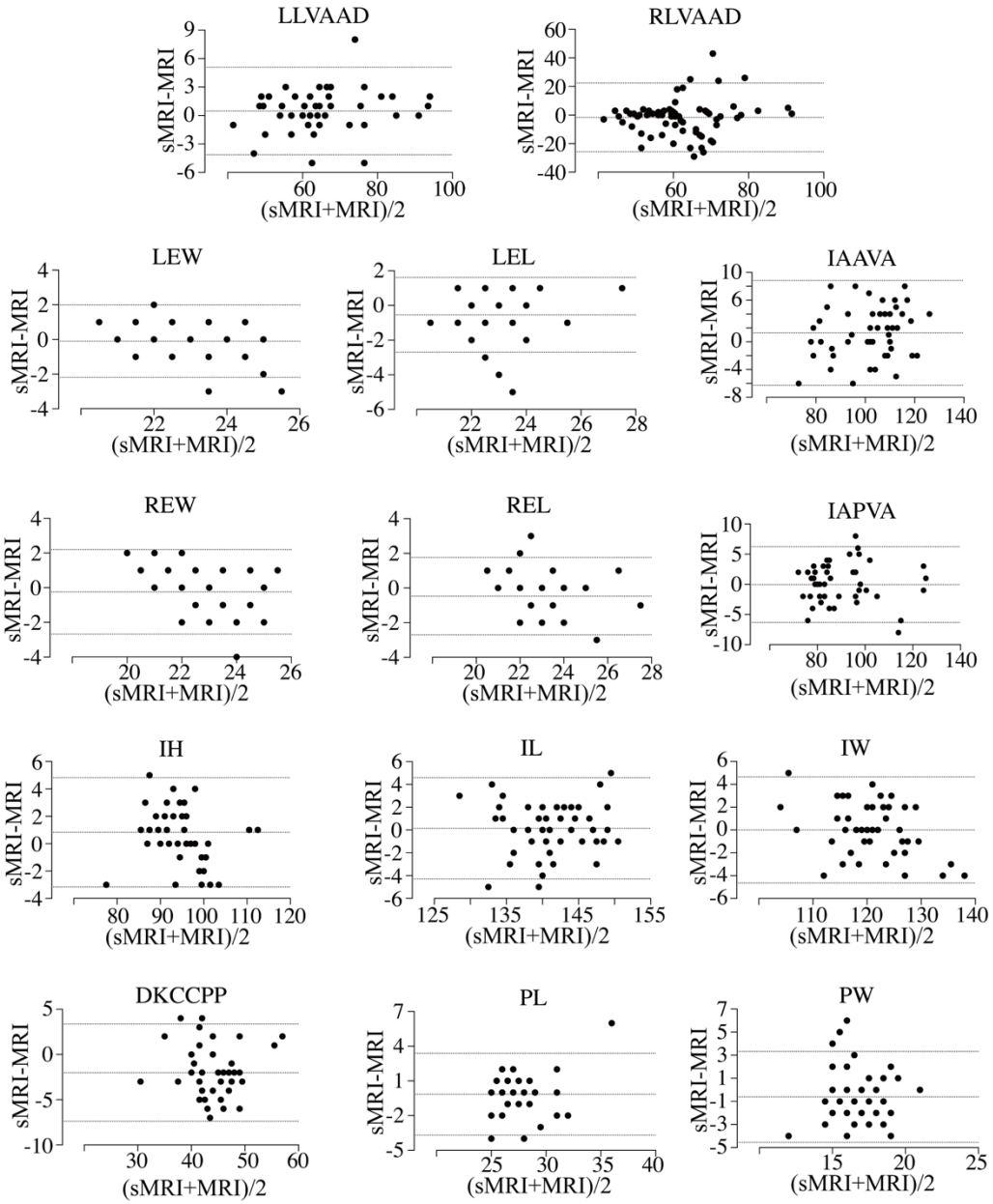


图 2 Bland－Altman 图

OR-019

ChatGPT 在公共卫生和健康领域的应用思考：以基层医疗卫生服务能力提升为例

曾华堂[1] 梁万年[2] 柯夏童[1] 许平[2] 黄鹏[3] 胡健[2] 唐尧[2] 伍丽群[1] 黄存瑞[2]
1.深圳市卫生健康发展研究和数据管理中心
2.清华大学万科公共卫生与健康学院
3.清华大学自动化系

基层医疗机构是我国公共卫生和健康领域的基底，承担着基本医疗保健、疾病预防、卫生监测等任务，对保障人民健康至关重要。然而，当前我国基层医疗机构存在服务能力不足、人员流动性大、设备不足、医疗资源配置不均等问题，严重影响了基层医疗机构的服务质量和效率[1]。近年来，人工智能（AI）和深度学习取得了巨大的进步，这些技术已经应用于包括公共卫生与健康领域在内的多个行业。在 2022 年 11 月，一家位于美国旧金山的非营利性人工智能公司 OpenAI Inc. 推出了一款新的聊天机器人（ChatGPT-4），并在短时间内引起各行业的强烈兴趣和争论[2]。ChatGPT 是一种大规模语言模型，能够处理和生成大量类人文本，提供对话、文书撰写、医学资料翻译、智能检索、报告生成和润色校验等功能，有可能在提升基层医疗机构服务能力方面发挥重要作用，例如助力患者教育、健康咨询、辅助诊断等。

然而，当前许多学术界学者和医疗行业管理者普遍认为，在将 ChatGPT 等人工智能技术应用于包含基层医疗卫生服务能力提升在内的公共卫生和健康领域时，需要充分考虑其自身功能的限制，以及在医学伦理、数据解释、问责制和其他隐私方面所面临的挑战[3]。因此，本文的研究目标和内容包括：（1）简要介绍 ChatGPT 在公共卫生和健康领域的潜在用途，以基层医疗卫生服务能力提升为例；（2）分析 ChatGPT 在公共卫生和健康领域应用过程中存在的局限性和挑战；（3）提出针对这些局限性和挑战的对策建议。

一、ChatGPT 发展现状

2018 年，OpenAI 提出了第一代 GPT（Generative Pretrained Transformer），将自然语言处理带入“预训练”时代。随着深度学习和自然语言处理技术的迅速发展，第一代 GPT 逐渐演化成为当前最先进的语言模型 ChatGPT-4，可以理解自然语言的输入并生成流畅的自然语言输出。ChatGPT 采用的是基于 Transformer 架构的深度学习算法，这种算法以自注意力机制、残差连接和层归一化等技术为基础，能够自动学习文本中的语义和语法规则，其工作原理是，先在大规模文本数据集上进行预训练，学习自然语言的统计规律和语义信息，然后在具体任务中微调模型，以适应任务的特定要求。ChatGPT 表现出来的智能化程度相比现有的搜索引擎和智能对话 Chatbot 等产品在多轮对话、上下文语义理解及推理能力、内容生成及总结等方面的表现令人震惊。全球也有不少知名企业加入人工智能军备竞赛，如微软的 Office365 Copilot，谷歌 BARD，百度文心一言等，ChatGPT 的出现开启人工智能的新纪元。

随着 ChatGPT 火爆全球，许多医疗行业相关从业者也很关注 ChatGPT 在公共卫生和健康领域的落地应用。最近，有报道称 ChatGPT 可以通过美国医学执照考试，并且在医学知识的主要能力评估中表现出了一个大学三年级医学生应有的水平[4]。一项关于测试 ChatGPT 回答患者有关乳腺癌治疗问题能力的研究发现，大约 88%的回答很恰当，但是在一个案例中它给到的治疗建议在一年前已经过时了[5]。另一项关于 ChatGPT 能否对患者问题提供与医生相当质量回答的研究结果表明，ChatGPT 对患者问题的回答质量优于医生[6]。然而，也有学者强调要注意 ChatGPT 在提供医疗协助方面的局限性，例如一项发表在 *The Lancet Infectious Disease* 的研究表明，在临床实践中实施 ChatGPT 的障碍包括：ChatGPT 对当前疾病情况的理解能力、从症状推断可能诊断的逻辑推理能力和回答相同问题保持一致性的能力不足，这可能会危及病人的安全[7]，以及 ChatGPT 有时会提供捏造或有偏见的信息[8]。因此，ChatGPT 在医疗卫生行业的应用要持有审慎的态度。

二、ChatGPT 在公共卫生和健康领域的潜在应用前景分析：以基层医疗卫生服务能力提升为例

由于我国医疗资源配置呈现“倒三角”的格局，且存在患者就医“趋高”的现象，导致基层医疗卫生机构的发展空间受到挤压，医疗服务能力提升缓慢[9]。此外，随着国家对分级诊疗制度的推广落实，强调居民要基层首诊和签约家庭医生，基层医疗卫生机构医生的工作量不断加大，然而基层医疗卫生机构医生存在配比严重不足的问题，进一步导致了在服务能力上无法很好满足居民的健康诉求[10]。因此，加快推进以 ChatGPT 为例的人工智能技术在基层医疗卫生机构的应用，提升医疗服务能力，满足居民健康需求显得十分迫切。

ChatGPT 是一个很有价值的工具，以人工智能技术为基础，自动分析患者的文本、影像、对话等多模式医疗健康数据，为基层医疗卫生机构开展卫生服务工作提供支持[11]。目前，ChatGPT 在基层医疗卫生服务能力提升的应用前景主要包括全天候和全过程的临床辅助诊疗、书写电子病历、远程管理慢性病患者、患者教育等 4 个方面。

（一）临床辅助诊疗：一是 ChatGPT 可以从患者以往的医疗记录中收集检查结果、家族史、症状等信息进行分类，帮助医生快速的评估患者的病情和治疗需求；二是 ChatGPT 帮助医生识别患者症状，并为医生提供一份可能的诊断清单，以及下一步的治疗建议；三是根据患者的病史和疾病治疗指南，ChatGPT 可能会提供对该患者治疗效果更好的药物或治疗方案的建议，这可能会降低不良反应的发生，保证患者获得最好的治疗。此外，需要注意的是，ChatGPT 仅作为辅助工具，医生仍需运用自己的临床经验和专业知识进行综合判断，并继续与患者进行有效沟通和互动。

（二）书写电子病历：ChatGPT 立足于服务医生，提供了帮助医生撰写医疗记录、电子病历等医疗文书的功能。无论是医生沟通、医疗检查说明还是医嘱等，都可以几秒之内根据需要生成自动化的文案，这将很大程度上减轻医生的工作负担，减少医生的职业倦怠，使临床医生能够花更多的时间照顾患者，提高诊疗效率和患者满意度。

（三）远程管理慢性病患者：一是基层医疗卫生机构的医生可以使用 ChatGPT 远程随访高血压、糖尿病等慢性病患者，以持续更新患者健康状况；二是 ChatGPT 帮助医生快速回复患者的询问，并根据患者的健康状况和需求创建个性化的生活方式指导；三是医生可以使用 ChatGPT 提示患者服药时间，以及有关副作用和可能的药物相互作用的信息，鼓励患者根据医嘱服药，提高依从性。

（四）患者教育：一方面，ChatGPT 可以学习简化医疗专业术语、处方，帮助患者更好地理解疾病情况和诊疗方案，从而改善对患者的教育。此外，医生工作节奏快，患者与他们直接接触并得到建议较为困难，ChatGPT 可以针对患者频繁咨询的问题，给出专业解答。例如，目前医院科室都有较多的医患沟通群，医务人员忙不过来的时候，一位经过训练的 ChatGPT 智能客服可以及时做好支撑，以更高的效率、更好的态度，为用户解答更多的问题。另一方面，ChatGPT 可以作为患者的虚拟医疗助手，根据患者的病史、当前的症状和其他相关信息提供个性化的建议和咨询。例如 ChatGPT 虚拟医疗助手可以为感冒患者提供家庭治疗建议（见图 1）。由于居民不需要去到医疗卫生机构就能获得专业和个性化的治疗建议，这既有利于降低基层医疗卫生机构医生的工作量也给居民提供了便利。

图 1 ChatGPT 对感冒患者提供治疗建议

三、ChatGPT 在公共卫生和健康领域应用存在的问题和挑战

（一）ChatGPT 自身存在的局限性：

ChatGPT 在公共卫生和健康领域的应用潜力巨大，例如可以通过辅助诊断、远程管理患者、检测诊疗错误和降低医生文书工作负担等来提升医疗服务能力。然而，ChatGPT 作为一种人工智能工具，其自身存在明显的局限性。

1. ChatGPT 的可靠性和准确性无法保证：虽然 ChatGPT 可以提供广泛的信息和诊疗建议，但是已有研究表明，由于 ChatGPT 无法分辨信息是否可信，存在被用来传播虚假医疗信息的风险，这可能会促进不准确信息的传播，对患者的健康造成危害，从而对社会产生有害影响[12]。此外，ChatGPT 甚至会编造虚假的期刊文章或捏造有偏见的信息来支持它的观点[13]，也就是所谓的“一本正经的胡说八道”。因此，它并不能代替医疗专业人员的判断能力、专业知识、培训和临床经验。总的来说，ChatGPT 可以通过向患者和医生提供快速、全面的信息来辅助基层医疗机构医生的工作，然而它远远不能替代基层医疗卫生机构医生在提供专业护理、做出治疗决定和确保患者获得最佳治疗等方面的重要作用。

2. ChatGPT 的使用缺乏医患之间的互动：尽管 ChatGPT 可以提供对问题的快速响应，但它不能取代医生与患者之间的互动在医疗服务中的价值。与书面的诊断结果和治疗方案相比，患者可能更需要医生的情感支持或细致入微的解释，而 ChatGPT 无法提供，因为它必须通过大量数据和运算资源的加持，才能够进行分析并生成语言，而且它缺乏人类的情感、价值观、道德判断力和创造力。

3. ChatGPT 的可访问性不高：ChatGPT 的访问需要账号，并非所有的患者或基层医疗机构医生都可以访问，因为 ChatGPT 只在部分国家可以注册，我国的 IP 地址无法注册。

4. ChatGPT 的语言障碍：一方面，ChatGPT 目前支持的语言数量有限，这可能会对使用其他语言的医生和患者构成挑战。另一方面，ChatGPT 可能无法理解中英文或东西方文化的细微差异，从而导致误解或误导，对医学科普产生负面影响。

（二）ChatGPT 在公共卫生和健康领域应用存在数据安全和患者隐私泄露的担忧：

医患关系信任的基础是对患者信息的保密，但是根据 OpenAI 官网公布的隐私政策，并未提及类似欧盟《通用数据保护条例（General Data Protection Regulations）》（“GDPR”）等数据保护法则，在“使用数据”条款里，0penAI 承认会收集用户使用服务时输入的数据，但未对数据的用途作进一步说明。根据 GDPR 法则，个人有权要求删除其个人数据，即“被遗忘权”或“删除权”，但像 ChatGPT 等大语言模型很难“忘记”用于训练自己的数据，也就是说，这类模型很难删除个人信息的所有痕迹，导致患者个人隐私存在风险。此外，ChatGPT 是一个开放的平台，需要将相关医疗健康数据传给它进行训练及交互，然而其还未提供私有化部署的版本，因此需要考虑医疗健康数据传给 ChatGPT 可能的数据安全风险。

（三）ChatGPT 在公共卫生和健康领域应用存在伦理道德和法律的挑战：

ChatGPT 的算法训练需要大规模高质量的数据，然而当前的算法通常是在有偏见的数据集上进行训练的，主要原因是现阶段 ChatGPT 的数据主要来源于互联网，而互联网既是一个信息宝库，但同时也是一个信息“垃圾场”，以互联网为信息源的算法训练自然会受到有偏见信息的误导[14]。这些潜在的偏见在医疗服务提供领域尤其令人担忧，因为某些疾病人群可能会被忽视和边缘化，进一步加剧健康不平等问题。此外，部分伦理道德问题还与法律相关。例如，当 ChatGPT 做出的诊疗建议对患者造成伤害时，谁应该受到指责还有待确定。当前大多数国家或地区的法律对 ChatGPT 的问责并没有明确界定，对传统医疗产品的问责制度是否可应用于 ChatGPT 也没有明确规定。尽管欧洲和美国法院将医疗软件视为医疗设备，然而，开发商可能通过将 ChatGPT 定义为服务而非产品来逃避法律责任，导致患者难以从伤害中获得应有的赔偿[15]。

四、ChatGPT 在公共卫生和健康领域应用的对策建议

ChatGPT 的强大潜力在医疗领域和其他领域中已经展现出来，对于其自身存在的局限性、数据安全、法律风险和伦理道德等挑战，我们只能期望它通过更高质量的数据来训练和优化算法，同时不断完善相关法规和问责制度。然而，禁止使用或忽视它的存在是没有意义的，因为它可以通过自动化生成大量电子文书、提供实时的医疗咨询等功能，降低医务人员的工作负担，对医疗实践带来积极影响。当然，在医疗领域中使用 ChatGPT 时确实需要谨慎，为了使 ChatGPT 充分发挥其在医疗领域中的潜力，我们不应仓促地应用它，而应该对潜在的风险进行评估，制定相应的政策和法规，确保数据的合法和安全使用，并保护患者的隐私权。其次，对于 ChatGPT 在公共卫生和健康领域的应用，需要全社会的广泛讨论，确保以审慎和负责任的态度引入和应用，达到真正改善医疗实践和提升医疗卫生服务能力的目的。基于此，本文提出的具体对策建议如下：

（一）推动高质量数据平台的搭建和使用：为了提升 ChatGPT 在公共卫生和健康领域的应用能力，需要收集和使用高质量的医疗数据来训练和优化算法。学术界学者和医疗行业管理者可以积极推动医疗数据的共享和合作，以确保 ChatGPT 能够基于准确和可靠的信息做出决策和提供建议。

（二）加快建立相关法规和问责制度：为了解决数据安全、法律风险和伦理道德等挑战，需要建立健全法规和问责制度。相关机构和管理者应制定明确的政策和法规，确保 ChatGPT 的应用符合法律法规，并建立机制对其进行监管和评估，以保护患者的权益和隐私。

（三）进行风险评估和伦理审查：在使用 ChatGPT 时，应对潜在的风险进行全面的评估和分析。医疗行业管理者可以组织专家团队进行风险评估，包括对误诊和不准确信息的可能影响进行评估，并对 ChatGPT 的伦理道德问题进行审查，确保其应用符合伦理道德。

（四）促进公众参与和知情同意：ChatGPT 在公共卫生和健康领域的应用需要广泛的社会讨论和公众参与。医疗行业管理者可以组织公众研讨会、听取意见和建议，确保公众对 ChatGPT 的应

用有充分的了解，并充分考虑公众的意见和关切。

（五）加强合作伙伴关系：医疗行业的管理者可以与人工智能技术领域的专家和研究机构建立合作伙伴关系，共同研究和探索 ChatGPT 在医疗领域的应用前景。通过合作，可以共同解决技术、伦理和法律等方面的问题，推动 ChatGPT 的可持续发展和应用。

总体而言，医疗行业管理者应以审慎和负责任的态度对待 ChatGPT 在公共卫生和健康领域的应用，如通过采取适当的措施确保其应用符合伦理原则和法律规定，以充分发挥 ChatGPT 的应用潜力，提升医疗卫生服务能力，为公众提供安全可靠的高质量医疗卫生服务。

OR-020

基于大数据的静脉血栓栓塞症主动预警预测系统与全程智能管理

曾国军
四川大学华西医院

目的 基于大数据与人工智能技术，实现静脉血栓栓塞症（VTE）主动预测、预警和全程智能管理。
方法 通过风险智能评估、智能预警、辅助决策、智能管理四大功能实现。风险智能评估：支持自动检测在院病人的医嘱，一旦下达相关医嘱即自动调用评估量表进行 VTE 风险评估，评估结果分等级按不同方式返回系统和提醒。智能预警：根据不同的分析评估模型得分进入不同的预警路径和处理途径。可支持客户端消息、短信、微信、OA 系统提醒等方式。同时对即将超时未下达 VTE 医嘱的病人列表向医生进行提醒。辅助决策：对病人的治疗方案进行智能推荐，结合病人住院数据进行预防方案推荐、治疗方案推荐、出血风险评估，智能提醒用药周期及注意事项，为临床决策提供支持。
结果 已经实现完善系统研发，并开始在医疗机构应用推广。
结论 基于大数据与人工智能技术，可以实现静脉血栓栓塞症（VTE）主动预测、预警和全程智能管理，有效提升 VTE 规范预防，降低 VTE 发生率，死亡率，复发率等。

OR-021

基于摄像头的深度学习两阶段训练法测量心率和心率变异性

兰蓝[1]，殷晋[2*]，邱甲军[2]，王伊龙[3]，秦睿[4]，唐诗[4]
1 信息管理与数据中心，首都医科大学附属北京天坛医院
2 华西生物医学大数据中心，四川大学华西医院
3 神经病学中心，首都医科大学附属北京天坛医院
4 真术相成（成都）科技有限公司

1 引言

心率变异性(HRV)是心跳周期的不同变化，一般通过分析心电图中的 R 波间隔得到。HRV 反映了心脏本身窦性心律失常的程度以及神经体液因子与窦房结相互作用的平衡。已有研究表明，HRV 可以预测心源性猝死、冠心病、高血压、慢性心力衰竭等心血管疾病以及慢性阻塞性肺疾病、糖尿病等疾病的预后[1]，并能反映睡眠、精神应激状态等多种信息[2, 3]。接触式心跳监测设备，如心电图和指尖脉搏血氧仪，被广泛使用，特别是在医院，这得益于生理健康指标检测的快速和准确。但是，接触设备对人体的长期接触会带来许多不良影响，如设备的清洁和维护成本，人体体液对设备的腐蚀和老化等[4]。毛细血管血流与心跳直接相关，毛细血管中的血红蛋白吸收一定频段的光[5]。摄像头可以作为非接触式设备的测量载体，以人脸信息为测量对象。心跳信号可以通过记录人脸的彩色图像来提取。然而，非接触式方法在采集生理信息时比接触式方法更容易受到外界环境的干扰，引入的噪声更大，丢失的重要信息也更多，因此准确性较低。因此，提高准确性是非接触设备长期的研究方向和重要需求。相关工作研究的重点是数据增强和信号滤波。信

号滤波的质量直接影响到最终结果的准确性。为了提高心率(HR)和 HRV 估计的准确性，本文总结并优化了前人的工作，设计了一个标准的基于摄像头视频 HR 和 HRV 的测量流程。

2 方法

本文算法通过信号采集、信号滤波和指标计算三个主要步骤来完成 HR 和 HRV 的估计目标。采集步骤利用深度神经网络对人脸进行检测，定位人脸关键区域，提取人脸血谱信息，同时平衡光照均匀性。然后通过分离红、绿、蓝(RGB)通道提取绿色通道，使用简单卷积滤波器(汉明窗)去除趋势，然后使用深度神经网络进行强滤波。与其他类似研究中使用的滤波器相比[6, 7]，深度神经网络滤波能更好地保留原始波形[8]。利用原始波形，可以计算出 HR 和 HRV(图 1)。

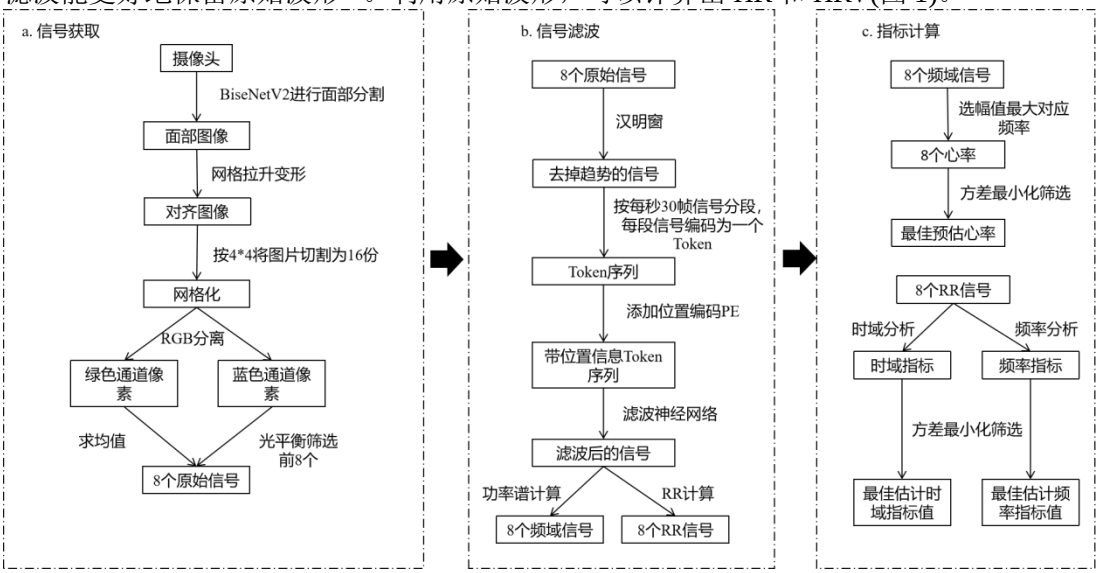


图 1 心率和心率变异性估算流程

2.1 信号采集

通过摄像机录像以每秒 30 帧的速度收集人体数据。通过 BiseNetV2 网络[9]检测面部皮肤区域。按照 BiseNetV2 网络的默认参数对原始视频进行人脸分割。然后，根据横坐标最大值和最小值、纵坐标最大值和最小值提取皮肤区域，并使用透视变换将图像转换成 512*512 的图像，用 4*4 的网格将图像划分为 16 个区域。对图像的 RGB 通道进行分离，保留绿色和蓝色通道。由于血红蛋白容易吸收绿色频率的光谱，而不容易吸收蓝色频率的光谱，因此使用绿色通道数据来计算 HR，使用蓝色通道数据来确定光线是否均匀地照射到皮肤区域。在每个网格的绿色通道的皮肤区域中对像素进行平均。由于图像每秒 30 帧，采集 20 秒，产生 16 个序列信号，因此序列信号的形状为 16*20*30。在每个网格中蓝色通道的蒙皮区域计算像素的标准差。每个网格形成 20*30 个标准差。每个网格的标准差的平均值就是该网格的光平衡得分。标准差越低，光越均匀。最后，选取得分最低的 8 个区域的信号进行后续分析。本文招募了 28 名志愿者(年龄在 18-68 岁之间)，共收集了 100 个视频和光学体积数据对，其中男性 58 人，女性 42 人。通过以上步骤，从 100 个视频中得到 8 个时间序列信号。

2.2 信号滤波

2.2.1 过滤步骤

(1)对 8 个时间序列信号进行预处理。首先，使用汉明窗去除一些趋势信号，这些信号通常是由信号收集过程中的不稳定因素引起的，例如被收集者无法完全保持静止。然后，将汉明窗的值作为卷积核，与原始信号进行卷积运算，得到滤波后的信号。

(2)将信号以每秒 30 帧的速度切片，并将每个片段定义为 1 个令牌(Token)。这 8 个信号持续 20 秒的视频总共有 8*20 个 Token。

(3)将 Token 进行位置编码(PE)叠加，并使用传统的相对 PE 计算方法。

(4)数据形状为 8*20*30(8 格，20 秒采集，每秒 30 帧)，送入滤波网络，得到滤波后的波形。

(5)利用滤波后的波形计算功率谱，用于计算 HR。

(6)提取滤波波形的峰值，计算 R-R 区间，用于后续的 HRV 计算。

2.2.2 过滤网络结构

滤波器网络采用编解码结构，编码器将输入信号编码为特征信号，解码器将特征信号解码为滤波后的信号。编码器和解码器采用变压器(Transformer)结构[10]，分别由 8 层和 4 层 Transformer 组成。在这个 Transformer 子结构中，设置了四个注意头，使用 LayerNormal 进行归一化，并使用 Eule 作为激活函数。对多层感知器(MLP)层做了一些修改，使用了两个完全连接的层，没有非线性操作。参数结构分别为 30*256 和 256*30。其主要目的是在不引入更多非线性的情况下放大参数量。

时域和频域信号可以完全对齐。对准方法采用 FFT 变换，由于对称的原因，将频率图和相位图的一半截断，然后将它们连接起来，以保证与原始信号长度完全一致。因此，频域编码器的结构设计与时域编码器的结构设计是完全一致的。

2.2.3 滤波的训练和测试

整个滤波网络的训练分为预训练和任务训练两个阶段。

(1)预训练

预训练数据集采用信号叠加生成的假数据。从理论上讲，任何周期信号都可以由不同频率的信号叠加而成。人们很自然地认为，通过产生不同频率的信号，然后将它们叠加，就可以形成一个复杂的信号。然而，来自现实世界的信号是复杂的。现实世界的信号有周期和非周期成分，而不是其他信号，除非你提取它们。对这种非周期信号的处理通常包括两种方法:将其作为无限周期信号的一个片段处理和将其作为某种噪声处理。本文处理它的方式是把它当作噪声。

(2)任务训练

对 PLE 数据集和自制的数据集进行了迁移学习训练。PLE 是 VIPL-HR 公共数据集[11]的一个子集，由 200 个经过滤波的原始视频和对光学体波形数据组成。从 100 个视频输入模型中随机选择 40 个视频与 PLE 数据集一起进行训练。这是矫正训练，需要选择一个温和的优化器和较低的学习率。因此，本文选择学习率为 0.01 的随机梯度下降(SGD)优化器，没有设置其他参数，以免破坏网络在预训练阶段学到的知识。

剩下的 60 个志愿者的视频，包括 36 个男性视频和 24 个女性视频，用于测试，以获得最终结果。

本文的方法与两种先进的模型(EfficientPhys[12]和 MTTS-CAN[13])进行了比较。

2.3 指标计算

2.3.1HR

本文首先自相关信号，然后使用 FFT。从频率信号的自相关可以得到 HR 的功率谱。最大功率对应的频率为当前 HR。由于正确的信号比较集中，通过减小方差进行选择，最终可以得到准确的信号。每个志愿者测试 3 次。每次测试的时间间隔为 2 小时，每次测试的时间为 15 秒。志愿者在面对摄像头记录面部信息时，佩戴了指尖脉搏血氧仪。在指尖脉搏血氧仪读数相对稳定后，记录 15 秒的数据和面部视频。本文设置并比较了三次指尖脉搏血氧仪显示的 HR。第一种指标判断法表明，正负误差在 1 范围内，则估计是正确的。第二个误差范围是 2。第三个误差范围是 4。也就是说，估计的 HR 被设置为正确或不正确的分类。然后，计算了每种方法的估计精度。

2.3.2HRV

利用波形的 R-R 区间，计算 HRV 相关指标，包括时域指标和频域指标。本文计算了 RMSSD、SDNN、SD1、SD2、SDSD、SD1/SD2 以及 LF、HF、LF/HF。在收集志愿者视频的同时，还使用 HRV 分析仪(型号:HW6C)收集了 HRV 和心理压力指数，其中 HRV 值作为基础真值。RMSSD 和 SDNN(两种常用的 HRV 指数)被用来比较本文的过滤网络和另外两种先进的模型的过滤效果。比较效果采用均方根误差(RMSE)来衡量，结果越小越好。对 6 个时域指标与心理应激指标进行皮尔森相关分析，以验证本文方法估计的 HRV 的有效性。

由于各 HRV 分析仪生产厂家采用的滤波算法不同，导致 HRV 频域指标的绝对值存在显著差异，不适合采用 RMSE 进行评价。本文的评价方法是将数据随机排列，然后进行前后比较。如果后一个数据比前一个数据大，记为 1；否则，记录为 0。然后，计算测试数据和目标数据之间的汉

明距离(具有相同位置的字符数)。汉明距离越大，趋势变化越一致。进行多次无序评价，最后计算平均值作为评价指标得分的趋势。分数越高越好，说明 HRV 估计越准确。

所有分析过程均使用 PyTorch 1.10 进行。

3 结果

3.1HR 测试

在进行人脸分割时，对 16 个网格进行光平衡评分。选取光平衡得分最低和最高的 8 个栅格进行后续分析。从表 1 可以看出，光平衡分数低的 HR 准确率高于光平衡分数高的 HR 准确率。

表 1 从不同光平衡得分测试的 HR

光平衡得分	范围 1	范围 2	范围 4
低	0.867	0.950	0.983
高	0.667	0.717	0.750

所有方法都在本文自制的数据集上进行了测试。从表 2 可以看出，使用神经网络滤波估计 HR 的准确率远远高于不使用神经网络滤波的准确率。通过本文的过滤网络计算的 HR 精度高于 EfficientPhys 和 MTTS-CAN。误差范围越大，估计的 HR 精度越高。

表 2 不同方法测试的 HR

方法	范围 1	范围 2	范围 4
未使用神经网络进行滤波			
Gudi et al [14]	0.167	0.200	0.217
小波滤波算法	0.300	0.367	0.417
使用神经网络进行滤波			
本文方法	0.867	0.950	0.983
EfficientPhys	0.833	0.917	0.917
MTTS-CAN	0.783	0.800	0.833

3.2 HRV 检测

3.2.1 时域

从 HRV 的时域分析可以看出，本文的方法估计的 HRV 的 RMSE 远小于 EfficientPhys 和 MTTS-CAN 算法(表 3)，这表明本文的方法在 HRV 时域分析中表现最好。

表 3 时域 HRV 测试结果

方法	RMSE
RMSSD	
本文方法	8.2
EfficientPhys	14.2
MTTS-CAN	19.3
SDNN	
本文方法	17.5
EfficientPhys	21.9
MTTS-CAN	26.5

3.2.2 频域

从 HRV 的频域分析可以看出，本文的方法估计的相关指标得分远高于 EfficientPhys 和 MTTS-CAN 算法(表 4)，这表明本文的方法在 HRV 频域分析中表现最好。

表 4 频域 HRV 测试结果

方法	相关指数评分
LF	
本文方法	0.81
EfficientPhys	0.51
MTTS-CAN	0.42
HF	
本文方法	0.76
EfficientPhys	0.46
MTTS-CAN	0.48
LF/HF	
本文方法	0.78
EfficientPhys	0.47
MTTS-CAN	0.43

3.2.3 与生理应激的相关性

本文对 6 个时域指标与心理压力指数进行了相关性分析。从图 2 可以看出，SDNN、RMSSD、SD1、SD2、SDSD 明显具有二阶函数规律。这些指标与身体压力之间的显著相关性是非常明显的。如果本文的数据无效，它们在散点图上就会显得杂乱。

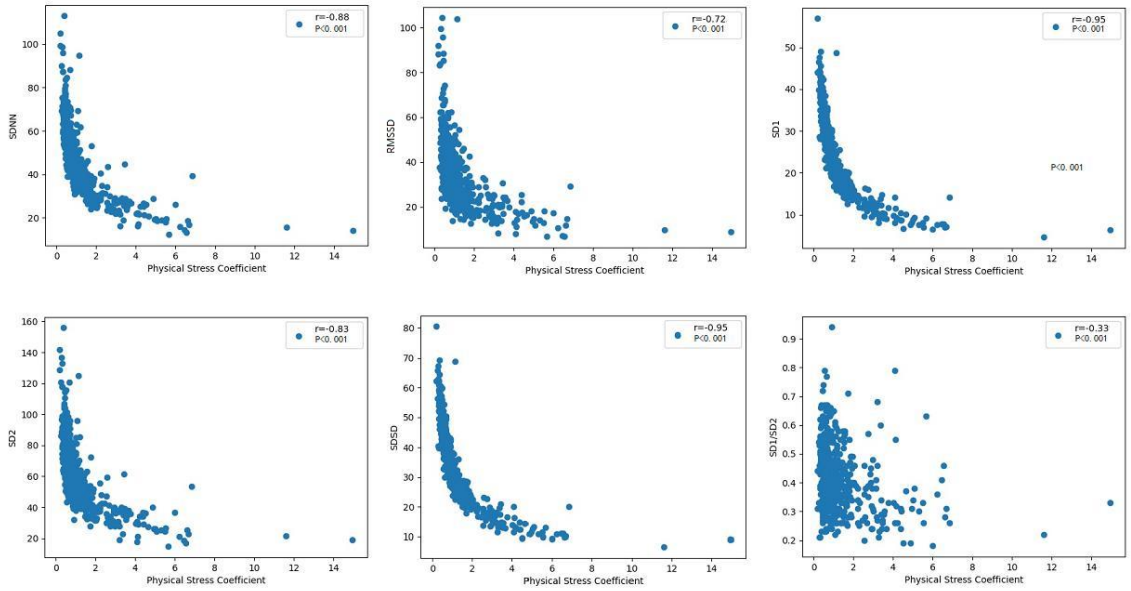


图 2 HRV 指数与心理压力指数的关系

4 讨论和结论

为了提高 HR 和 HRV 估计的准确性，本文采用了三种新的 rPPG 测量技术。第一种技术是使用光平衡分数来选择图像进行后续分析。第二种技术是使用大量的合成数据对模型进行预训练。第三种技术是首次应用 Transformer 来估计 HR 和 HRV。

以往的研究使用部分面部图像进行 HR 估计，收集到的信号不一定能反映出人体的真实 HR，从而导致结果或低或高。本文的研究将整个脸部分割成 16 个网格。计算每个网格的光平衡分数。光平衡分数越低，光线越均匀。从表 1 可以看出，光平衡分数低的 8 个网格的 HR 估计精度高于光平衡分数高的 8 个网格的 HR 估计精度，验证了更均匀的光线可以提高 HR 估计的精度。

由于视频收集成本等因素，以前的研究样本量有限。本文使用合成数据集生成 100 万个样本用于预训练本文的模型。此外，以往的研究使用的是公共数据集，如 VIPL-HR 数据集，数据质量较低。在本文中，筛选了 VIPL-HR 数据集，并选择了 200 个高质量的视频进行第二阶段的训练。最后，该 Transformer 在人工智能领域显示出良好的预测性能。本文首次应用 Transformer 估算心率和 HRV。Transformer 的主要优点是它突破了递归神经网络(RNN)模型不能并行计算的限制。与卷积神经网络(CNN)相比，计算两个位置之间关联所需的操作次数不会随着距离的增加而增加。每个注意力头都可以学习执行不同的任务。从表 2、3、4 可以看出，本文的方法在 HR 和 HRV 的估计上优于传统的信号处理方法和先进的方法，这说明大样本训练、高质量的视频和 Transformer 的优点可以提高 HR 和 HRV 的准确性。

然而，这项研究也有局限性。首先，虽然本文在合成信号时试图模拟真实世界的信号，但本文无法评估模拟信号与真实信号之间的差异。其次，当视频中的人体剧烈抖动时，即使人脸分割准确，HR 估计也不准确。在未来，HR 和 HRV 的精度还可以进一步提高，比如更真实地模拟合成数据。另一方面，可以从视频信号中分析更多的信息，比如呼吸。

OR-022

结合可解释机器学习方法的新型高性能列线图：基于 MIMI-IV 预测入住 ICU 30 天内中风患者的死亡风险

陈浩然 李姣 王敏 兰雨姗 王梓阳
北京协和医学院 医学信息研究所

1、引言

2016 年全球疾病负担研究显示，近四分之一的人一生中都有罹患卒中的风险，2019 年全球调查的数据显示，卒中仍是全球第二大死因。此外，越来越多的卒中患者被送入重症监护室（ICU），但死亡率很高，其他功能预后也很差。多项研究表明，生命体征和实验室指标的变化总是先于患者病情的迅速恶化。同样，入住 ICU 的卒中患者的预后也深受各种临床特征监测的影响。同时，重症监护室卒中患者的短期死亡风险具有更重要的意义。荷兰国家重症监护数据库的一项临床队列研究发现，在 7046 名卒中患者中，入住 ICU 的卒中患者短期（30 天内）死亡率较高，但在 ICU30 天后死亡率趋于稳定。

列线图是复杂数学公式的直观表示，因其简单明了而越来越多地应用于临床决策支持和个性化医疗中。以往的研究通常通过逻辑回归和 Cox 回归模型来建立列线图，但这些模型都是基于线性假设，无法处理临床实践中的非线性关系。相比之下，机器学习模型可以处理现实世界中的非线性关系，并表现出更高的准确性：一项研究发现，机器学习模型在预测宫颈癌死亡风险方面优于线性回归模型（logistic 和 Cox 模型）。但与列线图相比，机器学习模型的 "黑盒"特性限制了其临床适用性。因此，我们希望结合列线图和机器学习的优势，开发出性能更高、更易于使用的临床预测列线图。

2、方法

2.1 研究设计

本研究的设计包括四个阶段（图 1）：(1) 访问 MIMIC-IV 数据库，选择合适的卒中患者和相关临床特征；(2) 开发可解释的机器学习；(3) 评估所选变量和截断点；(4) 构建和评估列线图。

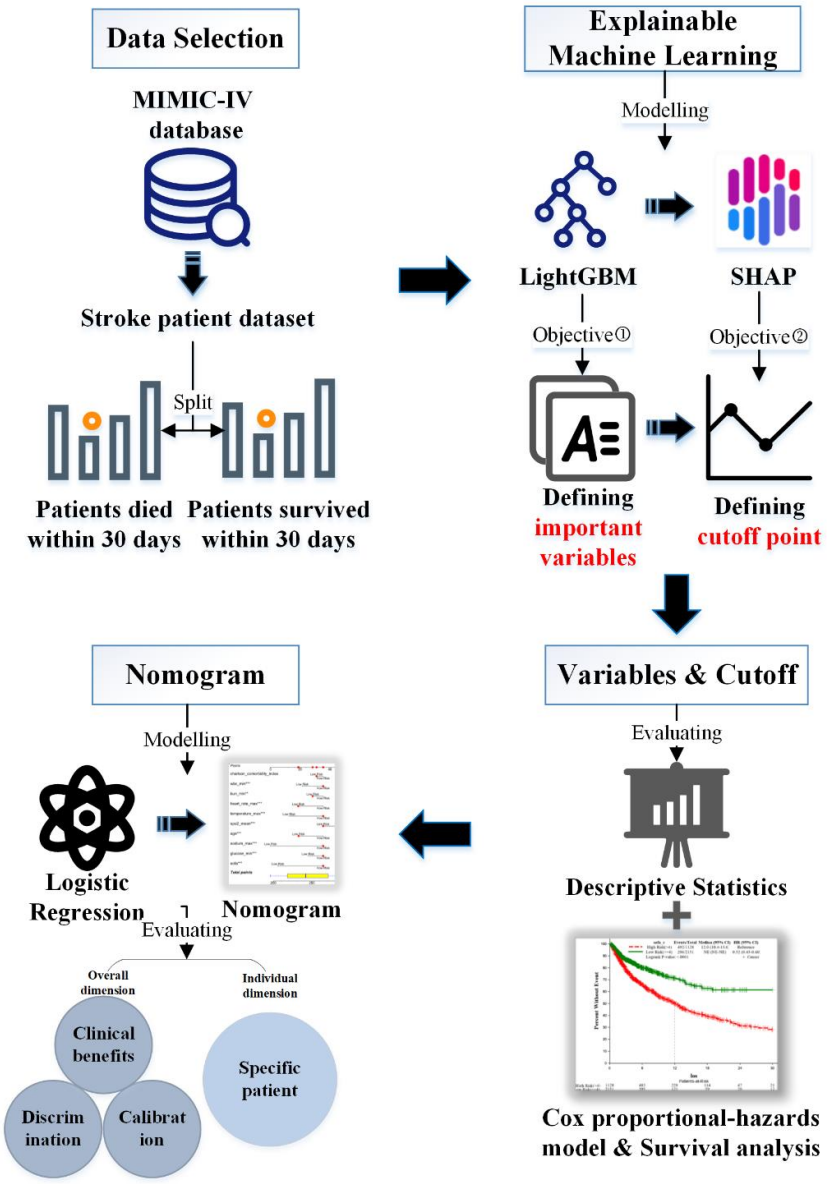


图 1. 研究设计框架

2.2 数据选择

所有数据均来自 MIMIC-IV（1.0 版）数据库。使用第九版和第十版国际疾病分类代码来识别脑卒中患者。进一步根据以下纳入标准筛选合适的脑卒中患者：(1)年龄在 18-89 岁之间；(2)仅纳入有一次住院经历的患者；(3)入住 ICU 的时间少于 30 天。卒中患者选择的详细过程见图 2。

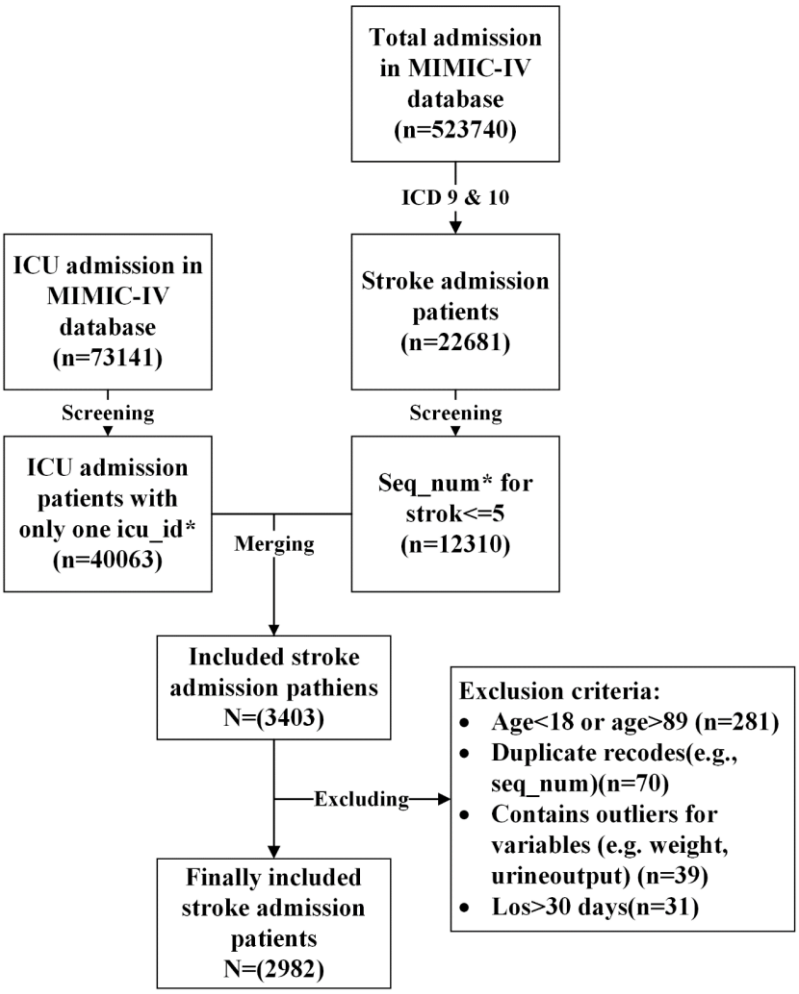


图 2 卒中患者数据选择流程图

使用 SQL、PostgreSQL 和 Navicat Premium 提取中风患者的数据，包括临床特征、实验室测量、合并症、生命体征和疾病严重程度评估等。脑卒中诊断类型（包括缺血性脑卒中、短暂性脑缺血发作（TIA）、蛛网膜下腔出血和脑内出血）也被纳入。

2.3 开发可解释机器学习模型

本研究将 LightGBM 作为机器学习算法，用于预测 ICU 卒中患者 30 天内的死亡风险。将数据集随机分为训练数据集（80%）和测试数据集（20%）。采用贝叶斯优化法寻找 LightGBM 的最佳参数组合，目标是使测试数据集的 AUC 最大化。优化后模型的性能进一步采用 5 折交叉验证方法进行评估。此外，我们还应用了 SHAP 方法来解释 LightGBM 的输出。其中，SHAP 摘要图用于确定特征的重要性，从而选择合适的变量；SHAP 部分依赖图用于找到所选变量的截断点。

2.4 评估所选临床特征和截断点

为便于开发列线图和降低临床应用的复杂性，我们选择了 SHAP 摘要图中重要性处于前 10 位的变量。采用 Kruskal-Walli 秩和检验来比较每个所选变量在生存或死亡两组之间的差异。

根据在 SHAP 部分依赖图中找到的每个变量的截断点，所有 10 个原本连续的变量都被二分为分类变量，因此脑卒中患者被分为两个亚组（高风险和低风险）。我们采用以下三种方法评估了这些截断点的性能：(1) 卡方检验比较这些分类变量在死亡组和生存组之间的统计分布；(2) 以高风险亚组为参照，使用 Cox 回归模型确定各选定变量与脑卒中患者 30 天死亡风险之间的关系；(3) 使用 Kaplan-Meier 生存曲线和对数秩检验比较基于各选定变量的两个亚组之间 30 天内生存风险的差异。

2.5 构建和评估列线图

以脑卒中患者入住 ICU 后 30 天内生存结局为因变量，对原有的 10 个连续变量和分类变量（即二分连续变量）分别构建了基于逻辑回归的列线图。在整体维度上使用了判别力（AUC 和净重分类指数）、校准力（校准曲线和 Brier 评分）和临床适用性（决策曲线分析），在个体维度上以一名特定患者（ID：2846）为例，比较两种列线图的性能。此外，通过基于 10000 次 bootstrapped 重采样的 DeLong 检验比较两个列线图的 AUC。

2.6 统计分析

由于偏度和峰度检验确定所有连续变量都呈现非正态性分布，因此用中位数和范围表示。Spearman 等级相关法用于评估变量之间的相关性，方差膨胀因子（VIF）用于评估 10 个选定变量可能存在的多重共线性。此外，分类变量以频数和百分比表示。在我们的研究中，基于 Python 开发和评估机器学习模型；使用 R 开发和评估列线图，进行数据分析和可视化。*P*<0.05（双侧）被认为具有统计学意义。

3.结果

3.1 临床特征的统计描述

本研究共纳入 2982 例 ICU 脑卒中患者，患者平均年龄为 65.3 岁，48.2% 为女性，多达 704 例（23.6%）患者在 30 天内死亡。除“性别”、“mbp_max”、“ptt_min”、“ptt_max”、“potassium_min”、“chlor_min”和“sodium_min”外，所有变量在存活和死亡的中风患者两组间均有显著差异。图 3 显示，所选的 10 个变量之间没有很强的相关性（spearman 相关系数均小于 0.5），且 10 个变量的 VIF 均小于 4（结果未列出），因此我们认为它们之间不存在多重共线性。

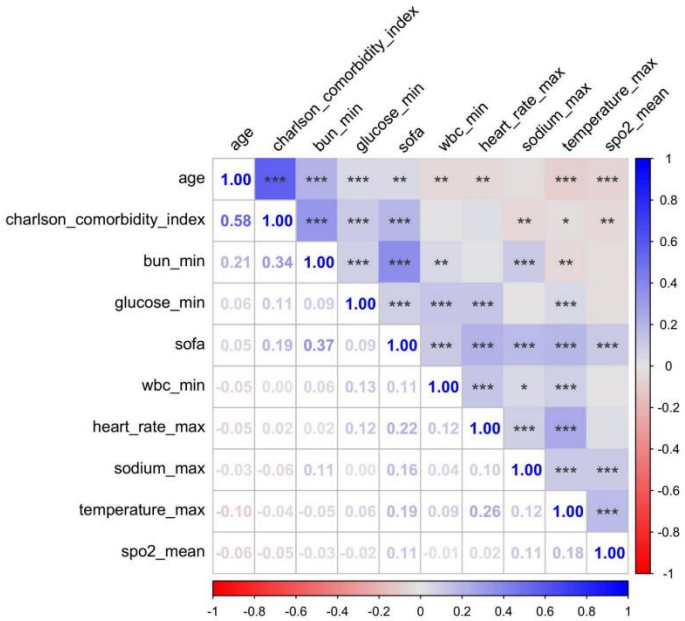


图 3 10 个选定变量之间的 spearman 相关矩阵

3.2 可解释机器学习

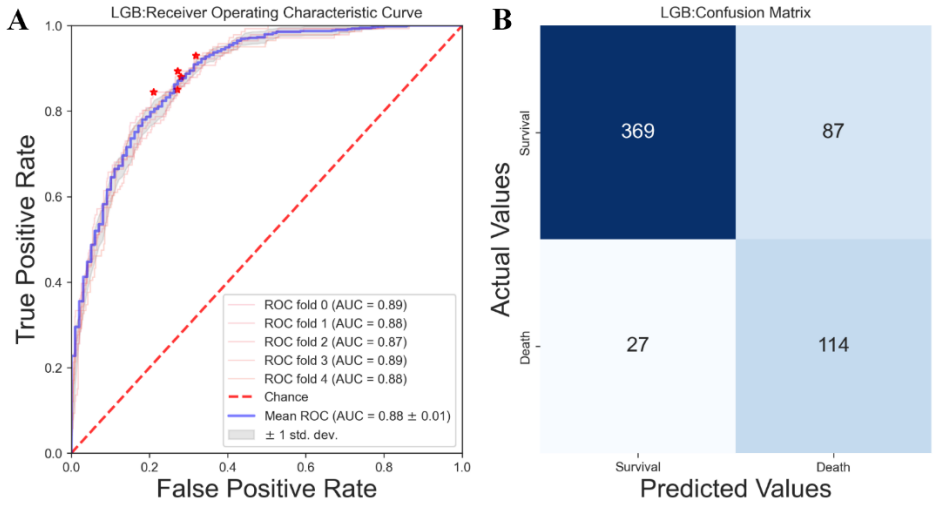


图 4. LightGBM 模型的性能。

图 4.A 和 B 显示，LightGBM 在预测脑卒中患者死亡风险方面表现良好（AUC：0.88±0.01，灵敏度：0.809，特异度：0.809）。我们观察到，"sofa"是对 LightGBM 预测能力最重要的变量（图 5.A）。其他被选中的前九个变量分别是 "glucose_min"、"sodium_max"、"age"、"spo2_mean" 、 "temperature_max" 、 "heart_rate_max" 、 "bun_min" 、 "wbc_min" 和 "charlson_comorbidity_index"。SHAP 依赖图（图 5.B）显示，在不同的特征取值下，所选的每个特征会对脑卒中患者的死亡风险产生不同的影响，而且 SHAP 值（Y 轴）在“截断点”两侧会发生变化。以"sofa"为例，当"sofa "值大于 4 时，SHAP 值一般大于 0，表明"sofa "值高对脑卒中患者的生存有显著的负面影响。其他 9 个变量的截断点为" glucose_min "，100；" sodium_max "，145；" 年龄 "，70；"spo2_mean "，93；"temperature_max "，36.5 和 37.8；" heart_rate_max"，100；"bun_min"，18；" wbc_min "，11；" charlson_comorbidity_index "，6。

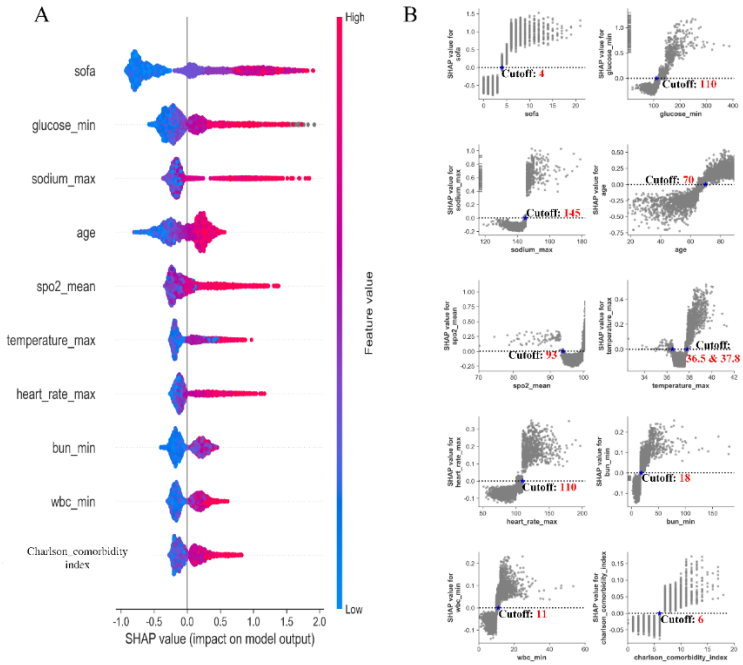


图 5.使用 SHAP 得出的可解释 LightGBM 结果。

3.3 评估选择的变量和截断值

表 1（结果未列出）显示了所选 10 个变量的统计描述结果：存活和死亡的脑卒中患者在所有 10 个变量上均存在显著差异（*P* <0.001），从而表明所选的 10 个变量能很好地区分存活和死亡的脑卒中患者。

以下三种评估方法的结果表明，所选变量的截断点在区分卒中患者存活和死亡方面具有足够的鉴别力：

(1) 卡方检验表明，所有分类变量在两组脑卒中存活者和死亡者之间均存在统计学差异；

(2) 如 Kaplan-Meier 生存图（图 6，结果未全部列出）所示，高风险亚组脑卒中患者的 30 天总生存率显著低于低风险亚组（对数秩 *P* < 0.001）

(3) 对于所有选定变量，Cox 比例风险危险模型显示，与高风险亚组相比，低风险亚组的 30 天死亡率较低（如“sofa”：HR=0.50（0.43,0.59），图 6.A）。

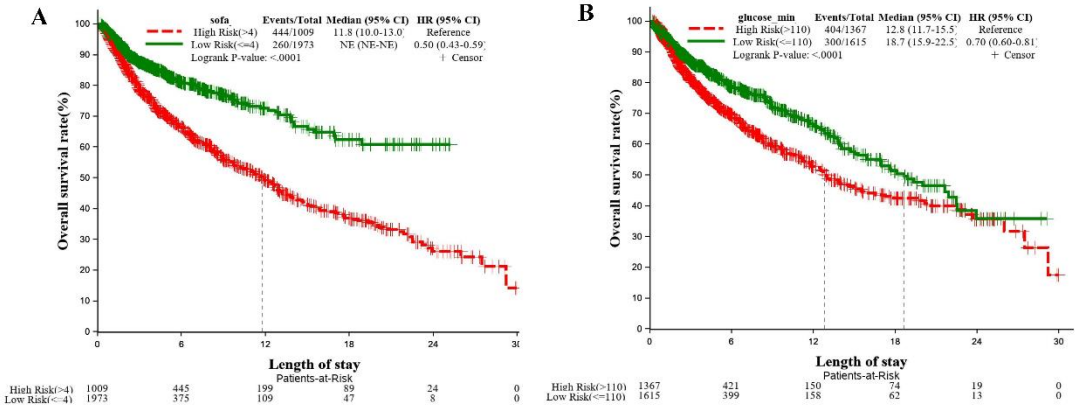


图 6. 以高风险亚组为参照，通过 Cox 比例危险度回归模型和 Kaplan-Meier (KM)生存曲线分析“sofa”和“glucose_min”与脑卒中患者死亡风险之间的关系。

3.4 列线图的评估

图 7 包括两个列线图：一个列线图（图 7.A）使用原始连续变量，另一个列线图（图 7.B）使用分类变量。我们分别通过“总体评价指标”和“特定患者的实例”对两个列线图的性能进行了比较

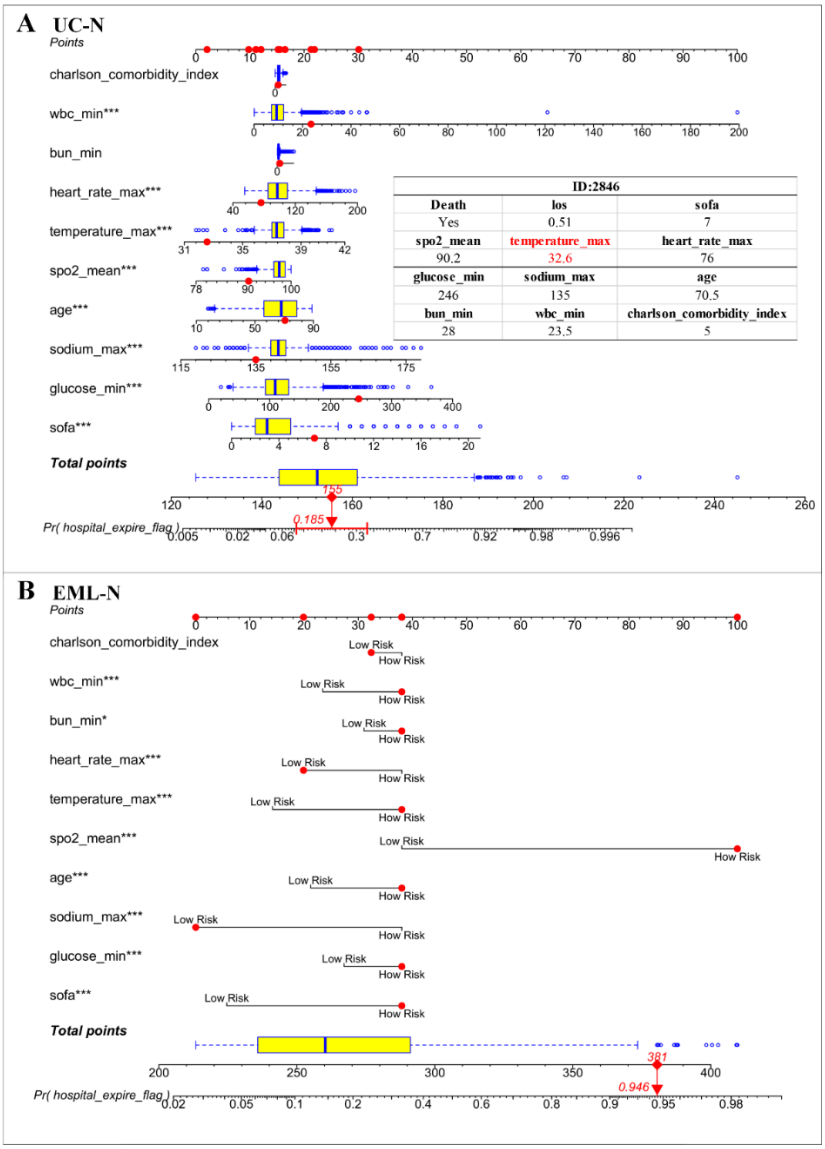


图 7. 预测脑卒中患者 30 天死亡风险的列线图。

3.4.1 整体维度

图 8 比较了 "EML-N "和 "UC-N "在以下 3 个方面的总体表现：

(1）鉴别力："EML-N "和"UC-N "的 AUC 分别为 0.837（0.800，0.873）和 0.838（0.800，0.877），DeLong 检验发现两种列线图之间无显著差异（P=0.970）；"EML-N "的 NRI 与 UC-N 相比，在预测 30 天死亡率方面有统计学意义上的正向改善（6.37% (2.11%,10.7%)）（*P*< 0.05）。

(2）校准能力：两个列线图的校准曲线均显示卒中患者的实际 30 天死亡风险与预测的 30 天死亡风险一致，但“UC-N”的“Brier”评分高于“EML-N”。

(3）临床适用性：在较高的阈值概率（0.4-0.8）下，“EML-N”的临床净获益高于“UC-N”

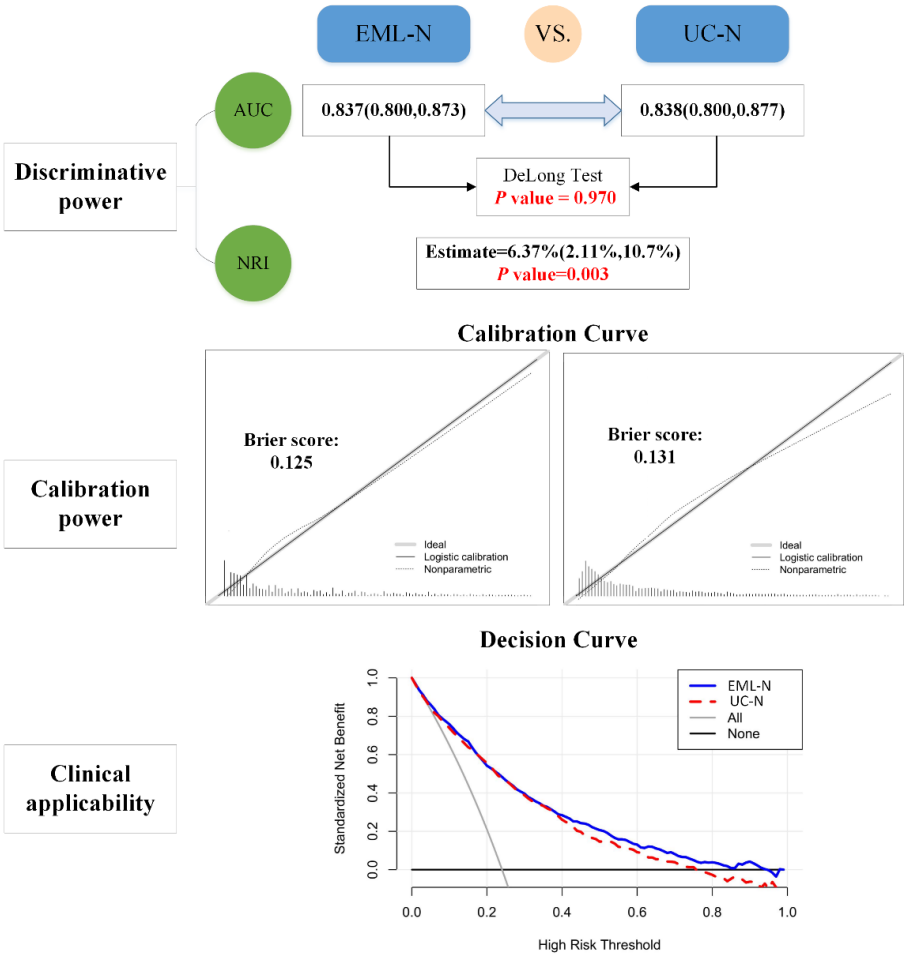


图 8. 预测中风患者 30 天死亡风险的列线图。

3.4.2 个人维度

图 7.A 显示了一位编号为“2486”的患者的详细信息（如他/她进入 ICU 的第一天，"sofa "为 7）。根据列线图（图 7）"point"轴中每个特征对应的得分:该患者在“UC-N”中的总得分为 155 分，对应的 30 天内死亡风险为 18.5%（图 7.A）；但在“EML-N”中（图 7.B），该患者的总得分为 381 分，其 30 天内在 ICU 死亡的风险为 94.6%。毫无疑问，“EML-N”对该患者的预测结果是正确的。我们热瓦内“UC-N”和“EML-N”在预测该中风患者死亡风险方面的巨大差异是由于"temperature_max"特征的评分不一致造成的。“UC-N”是根据" temperature_max "与脑卒中死亡率之间的线性相关的逻辑回归得出的，它忽略了脑卒中患者因 "temperature_max"较低而死亡的风险。此外，与“UC-N”相比，我们的“EML-N”更易于定义单个特征的分数。

4. 讨论

在本研究中，我们开发了基于“EML-N”的列线图来预测 ICU 中风患者 30 天内的死亡风险，与“UC-N”相比性能更高。此外，我们的研究结果表明，可解释机器学习能够识别重要变量并探索复杂的非线性关联，可以改善传统线性模型（如逻辑回归）的不足。因此，我们的提名图可以让临床医生在脑卒中患者入住 ICU 的第一天就轻松准确地评估其短期死亡风险，从而改善患者的治疗和护理。

鉴于脑卒中已经造成了巨大的疾病负担：仅脑卒中一项就导致全球 660 万人死亡，因此，脑卒中预后相关预测模型准确性的微小提高都会带来巨大的益处。与“UC-N”相比，我们的“EML-N”在整体维度和个体维度上都有显著改进。我们认为，正是我们对传统列线图方法上进行了以下两大改进，才使得我们的 “EML-N” 具有更高的性能。首先，线性模型（包括 logistic 回归和 cox 回归）是开发列线图最常用的方法。然而，当预测因子与结果之间存在非线性关联时，这些线性模型就不合适了。Daan 等人报告称，在评估连续预测因子与结果之间的非线性关系时，限制性立方样条回归（一种非线性建模方法，RCS）优于带线性项的逻辑回归。虽然有些研究通过在逻辑回归中使用变量的三次样条来拟合预测因子与结果之间的非线性关系，但选择的位置和个数受先验经验的影响很大。以本研究中的 "temperature_max "变量为例，我们发现使用 3 个节点（第 10、50 和 90 百分位数）和 5 个节点（第 5、27.5、50、72.5 和 95 百分位数）的三次样条回归（RCS），在 "temperature_max "超过 38 后的曲线趋势有显著差异（结果未报告）。此外，当 RCS 中包含大量变量时，选择所有变量的最佳拟合形式的工作量很大，很容易导致结果偏差。相反，本研究中的 SHAP 部分依赖图则大大降低了 RCS 中选择结点的难度，并允许同时对多个变量进行非线性拟合。其次，在我们的研究中，通过 SHAP 摘要图很容易筛选出与脑卒中患者 ICU 死亡风险相关的重要变量。在疾病研究中，选择重要变量的常用方法是最小绝对收缩和选择算子算法（lasso）。然而，lasso 是一种线性模型，它只能选择与结果线性相关的变量，而且只能从一组高度相关的变量中选择一个变量。当然，有些关键变量可能无法被选中。在我们的研究中，UN-N 和 EML-N 都是根据从 SHAP 汇总图中选择的变量建立的，它们的 AUC 值都很高。

5. 结论

我们分析了脑卒中患者在重症监护室第一天的临床特征，并开发出了一种性能更高、更易于使用的提名图。我们的研究表明，可解释机器学习处理变量间非线性关系的能力可应用于线性模型。

OR-023

基于隐私计算的健康医疗科研数据协作网络

王昌然[1]，张莺耀[2]，薛乐乐[2]，罗震[2]，刘雷[1]
复旦大学
翼健(上海)信息科技有限公司

引言

2016 年 6 月 8 日，国务院常务会议讨论通过了《关于促进和规范健康医疗大数据应用发展的指导意见》[1]（国办发〔2016〕47 号），明确提出了四大重点任务和十四项工程，对健康医疗大数据发展的指导思想、基本原则和发展目标提出了具体指导性意见。2018 年 7 月，为加强健康医疗大数据服务管理，促进“互联网+医疗健康”发展，充分发挥健康医疗大数据作为国家重要基础性战略资源的作用，根据相关法律法规，国家卫健委发布了《国家健康医疗大数据标准、安全和服务管理办法（试行）》[2]。2020 年 4 月 9 日，中共中央、国务院发布《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》[3]，将“数据”与土地、劳动力、资本、技术并列，作为新的生产要素，并提出“加快培育数据要素市场”，数据要素市场化配置上升为国家战略。隐私计算，又称为隐私保护计算，入选 Gartner 2021 年九大重要科技趋势[4]，该技术核心解决数据保护和数据共享这两个相互之间存在矛盾的需求问题。2021 年，《中华人民共和国数据安全法》、《中华人民共和国个人信息保护法》正式施行，为数据有序流通提供了法律基础。2022 年，《要素市场化配置综合改革试点总体方案》[5]、《中共中央、国务院关于构建数据基础制度更好发挥数据要素作用的意见》（简称数据二十条）[6]探索建立了数据要素流通规则：即建立数据资产持有权、数据加工使用权、数据产品经营权等分置的产权运行机制、建立高效合规的数据要素流通和交易制度等举措。为隐私计算提供了诸多明确的落地指引。健康医疗大数据作为国家基础战略性资源和国家核心资产，其开放共享可为政府实现精准决策、智慧医院发展、医药科研提供数据

支撑，提升医疗数据的再利用价值，促进医疗健康服务产业的创新，激发全社会创造力和市场活力。推进健康医疗数据共享，释放数据资源价值成为大数据时代背景下推进健康医疗数据最大化利用的关键路径。

尤其在医学科研领域，国家卫生健康委和国家中医药管理局联合印发《公立医院高质量发展促进行动(2021-2025 年)》，指出实施临床科研提升行动。通过高校医学院、附属医院及非附属的教研合作医院之间搭建数据跨域，数用分离的临床研究协作平台，通过数据不出域、可用不可见等方式，有效整合广域的医疗数据形成大样本数据源并供各级科研工作者按需使用，是提高科研成果产出转化的重要基础设施建设。

本文主要介绍在临床科研协作，特别是多中心研究中的数据协作共享方式，既支持传统的数据物理集中式计算分析，也能支持非物理集中式的分布式数据协作方式。如可以通过分布式探查，安全沙箱容器内计算等，在新的数据共享机制下，保障科研工作者在平台进行数据检索、处理、分析等应用。

目的

针对多中心临床科研场景，通过隐私计算技术，设计一套科学数据协作共享网络机制，并建立数据协作的生态，在原始数据不出机构的控制域的情况下，实现数据使用权的共享，保障各参与机构如医院对各自的数据有控制权，确保数据的所有权不会被转移。通过核心技术、产品、管理、流程相结合，实现从数据接入，数据探查、数据授权、数据使用、数据销毁的单体节点建设，在多个机构部署单体节点并互联形成健康医疗科研数据的分布式协作网络。

方法

隐私计算指在不泄漏参与方数据的前提下，由两个或多个参与方通过联合计算实现数据价值挖掘的技术体系。[7]隐私计算实现的技术路径一般包括安全多方计算[8]、同态加密[9]、联邦学习[10]、可信执行环境（Trusted Execution Environment，TEE）[11]等。

整合上述核心关键技术构建单体隐私计算节点，形成相应的数据开放应用生态。同时，在节点内必须提供数据存储和计算的资源以及资源调度的方式，可以动态弹性地利用节点的存储和计算资源。

隐私计算节点的设计首先要保证数据生态的安全性，确保数据安全和个人隐私保护。从节点的信任假设出发，建立安全威胁模型，并提出相应的技术手段保证节点的网络安全、存储安全、计算安全。

在该项目中，大部分临床数据均可在安全沙箱中开展计算和分析工作。在单体节点支持用户自助创建安全沙箱：安全沙箱实例交付形态可以是一组虚拟机或容器实例。在计算实现上，平台将所有第三方应用程序和服务容器化，使用 Kubernetes 作为容器编排基础架构，支持沙箱容器计算方式、TEE、联邦学习等多种安全计算方式。在保证第三方应用计算安全的同时，提供高效灵活地应用部署方式和计算服务。通过加强应用和 host 操作系统的安全隔离，减少 host 操作系统被攻击的可能。这样即使是恶意的应用，也无法通过系统调用的方式来控制 host 的操作系统。

用户启动程序后，平台创建一个类似于“安全屋”的安全沙箱运行环境，加载用户镜像本身，并将用户所加载的数据集作为一个外部存储加载（mount）到该运行环境中。这样用户就好像按照自己的软硬件要求配置了一台电脑，启动了电脑，并且将所需要的数据通过一个外部存储连接到电脑上来。

提供沙箱管理功能：包括使用管理（沙箱的开启、销毁）、资源管理（CPU 和内存的资源分配）等。平台通过沙箱容器计算等技术手段，保证第三方应用（哪怕是恶意的应用）在使用时不会对于平台的安全性产生任何影响，保证用户在使用第三方应用（哪怕是恶意的应用）时遵守平台用户的数据使用权限。

对不再使用的沙箱，可销毁沙箱内的所有数据，并回收沙箱资源。当应用程序运行过程中读取输入数据集时，系统会利用数据集授权时提供给用户的密钥按需实时地进行解密，进一步保证

用户无法绕过应用界面直接使用未授权数据。应用在运行过程中产生的数据会临时存储在实际所用的计算节点的本地存储中。应用执行完毕后，除了指定输出的结果数据外，其他中间数据都会被销毁掉，最大程度上避免数据泄露风险。

结果

基于单体节点的建设，在复旦大学医学院及附属医院的不同机构主体间，通过单体节点间的互联，以分布式等数据探查、数据授权使用等的建设从而形成健康医疗科研数据协作网络或联盟[7]，见图 1。实现节点网络内数据使用授权开放、用户开放并支持第三方应用开放。通过建立疾病专科间多中心科研协作联盟网络，如多中心肝癌专科数据协作网络、儿科罕见病专科数据协作网络、脑胶质瘤专科数据协作网络等等。从而解决跨数据控制域间开展数据安全有序协作流通的数据所有权/持有权保护、使用权共享的问题，促进健康医疗数据安全有序流通。

节点内所构建的数据生态内有不同类型的参与方。有数据提供方，例如复旦大学实验室提供生命科学公共数据、医疗机构提供临床数据等。有数据的服务方，在平台内将获得数据提供方授权使用的数据利用平台内部署的第三方软件或者自己开发的程序进行加工处理，满足不同的需求。例如一家 AI 公司利用诊疗数据训练疾病预测模型；或是一个医药研究所在平台内利用药品在区域内使用的使用进行真实世界研究。还有数据的需求方，他们利用平台构建的数据产品，满足业务的需求。例如，一家医药公司利用平台提供的数据产品满足药物研发的需求。最后，平台由平台运营方管理，运营方负责维护数据生态和管理工作。例如为新的用户建立账号，确保平台的存储和计算资源负载在系统允许的范围内，处理软件、硬件故障等等。

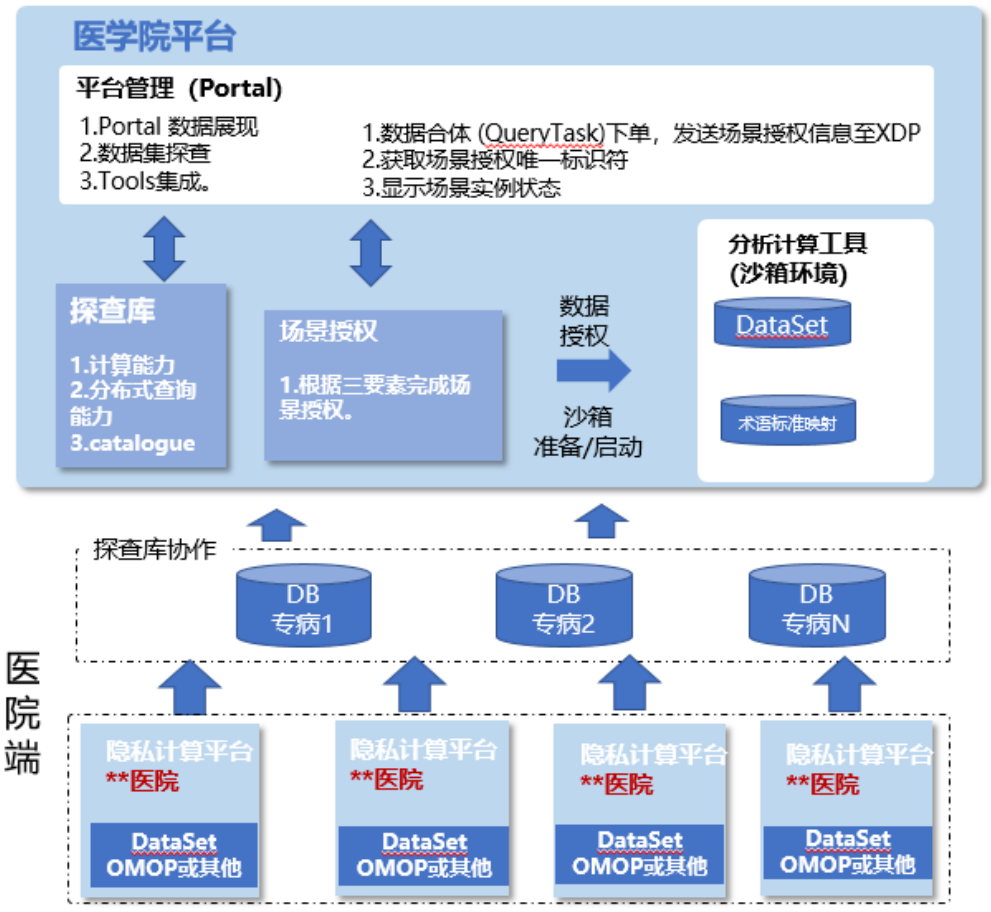


图 1

结论

在大数据、人工智能、隐私计算技术催化下，现代医学研究范式正在悄然转变。多中心健康医疗科研协作已成为现代医学研究必不可少的环节，在数据安全法、个人信息保护法等法规出台后，受限于医疗数据的敏感性和安全性，跨机构共享医疗数据发展缓慢，基于隐私计算构建跨控制域的健康医疗科研数据协作网络，在充分保护医生和患者隐私的前提下，实现各机构数据所有者便捷管理数据的控制权，共享数据的使用权，以一种全新的方式实现健康医疗数据的安全、有序的开放共享。通过这种方式，可加速多模态医疗健康数据要素安全有序的流通、共享、协作，促进多元化的数据驱动型医学研究，提升医学教育和诊疗服务创新，引领新医科发展方式。

OR-024

中国居民昼夜节律时空分布及节律扰乱与压力的关系——基于微博活跃度数据

李顺利*，向菲
华中科技大学医药卫生管理学院

1 引言

昼夜节律指每日睡眠和清醒的规律，是个人生活的最主要特征之一，也是人体内在生物钟的重要外在表现[1]。昼夜节律在调节睡眠习惯、体温、激素分泌、食物摄入和认知等[2]多种生物功能方面起着重要作用。扰乱这种节律将导致多种生理[3, 4]或心理[5]健康问题。

在现代社会中，昼夜节律的扰乱十分常见。一方面，现代生活方式扰乱了自然节律。不仅调节昼夜节律的光信号[6]受到电灯和电子产品等产生的光线干扰[7, 8]，轮换工作制[9]和高工作量[10]等工作需求也会迫使人们用闹钟提前于自然生物钟起床，并在自然就寝时间之内保持清醒[6]。另一方面，外部压力也导致了昼夜节律失调的发生[11, 12]。简单来说，暴露于压力事件将影响促肾上腺皮质激素释放激素的水平，通过下丘脑-垂体-肾上腺轴调节睡眠模式[13]。在现代生活中，压力主要来源于经济、人口和环境三个方面[14]。研究表明，过高的人口规模和生活成本将显著影响居民幸福感和心理健康[15, 16]；空气污染将导致抑郁等心理问题[17]；而收入越高的人群心理压力越小[18]。2006 年，Wittmann 等人提出了“社会时差”的概念以定量地描述这种节律被扰乱的情况，“社会时差”被定义为休息日和工作日的睡眠中间点之间的绝对差值[19]。研究发现，在工业化国家，已经有超过 70%的人口具有至少 1 小时的社会时差[6]，而在后疫情时代，这种情况只会更加严重[20]，这俨然已经成为广泛存在的公共卫生问题。

然而，尽管昼夜节律如此重要，目前的研究仍主要集中在个体层面，比如分析理想的光照-黑暗交替环境下昼夜节律的机制和功能[21]，或探究个体的生活习惯、工作习惯和昼夜节律之间的关系[22]。在群体层面上，尽管有不少研究调查了人群的睡眠持续时间和质量，但鲜有学者关注人群的昼夜节律变化，其扰动与压力的关系也尚未被充分认识。本研究通过使用微博时间数据，对以下问题进行分析：（1）微博数据能否在一定程度上反映昼夜节律？（2）昼夜节律的时空分布如何？是否存在规律性？（3）昼夜节律的扰动情况与外部压力的关系如何？

2 数据与方法

2.1 采集微博数据

本文选择微博作为数据来源以方便、快速地获得能够覆盖大范围地区的海量数据。为了对用户进行随机抽样，使用在各种语境中频繁出现的中文停用词在微博上进行检索[23]，并收集微博用户 id。随后，剔除了非中国大陆用户、无效用户（已被封号用户和不存在的用户）和可能造成误差的用户（不能查看所有微博的用户）。收集余下用户在 2023 年 8 月 15 日及以前发送微博时记录的时间。这些微博被按照不同的省、自治区和直辖市进行分类，并剔除了数据量过少（少于 10000 条）的地区（西藏）以减少数据不足造成的计算误差，最后共获得 58346 位用户发送的 4307858 条微博对应的时间（见图 1A）。另外，为了更加准确地识别用户发送微博时所属地区，对于有 IP 地址记录的微博采用 IP 对应的地区，否则才使用用户自填的账号所属地区。使用《中国统计年鉴——2022》[24]中的常住人口数据以验证采集数据的地区代表性，发现各地微博用户数量与各地人口之间呈显著正相关关系（见图 1B），因此认为采集到的用户数据具有一定的地区代表性。

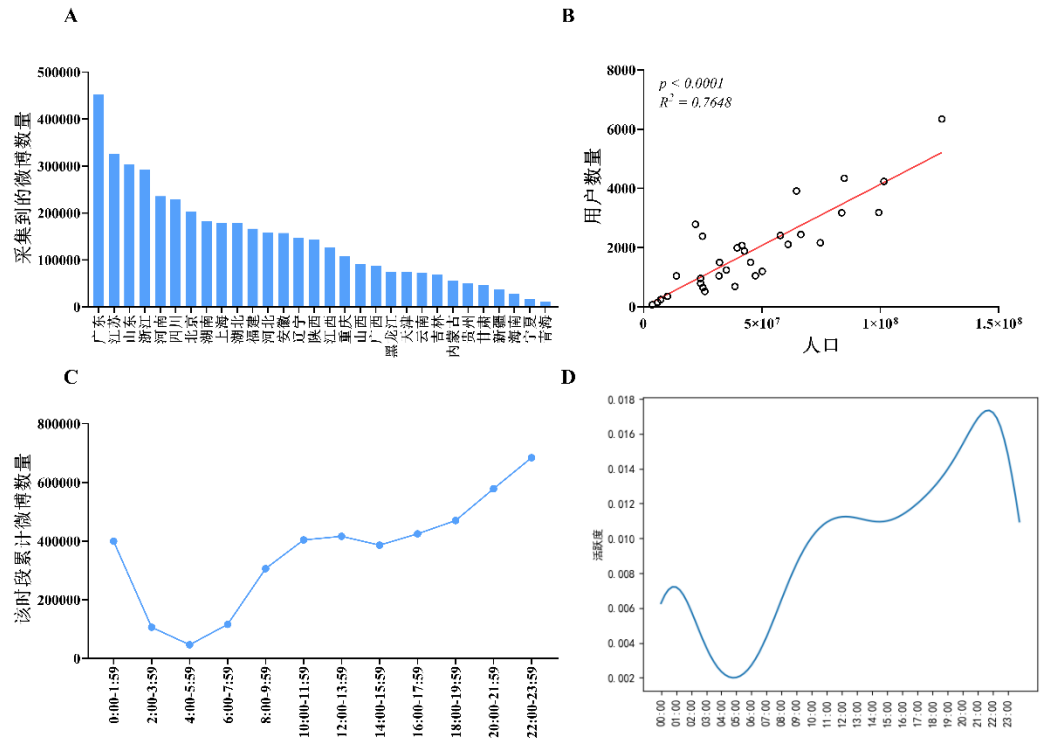


图 1 采集到的微博数据总体情况

A：采集的不同地区的微博累计数量；B：不同省份常住人口数量与采集的不同省份用户数量之间的关系；C：采集的不同时段的微博累计数量；D：经过平滑和标准化后的微博活跃度。

2.2 对微博用户昼夜节律、节律扰动和外部压力的表征与量化

为了描述微博用户的昼夜节律时空分布，需要统计不同时间节点微博用户的“活跃度”；分析节律的扰动情况则需要量化工作日和休息日之间昼夜节律的差异；而为了分析节律与压力的关系则需要一种“压力指数”描述不同地域人群面临的压力情况。本部分用于描述这些所需指标的计算方式。

2.2.1 微博活跃度的计算

为了计算昼夜节律，本文以 15 分钟为单位将一天分为 96 个区间，并统计落入每个区间的微博相对累计数量以表征该时间段内人群的活动状态[25]。具体来说，若用户 A 在某天内共发送了 k_n 条微博，在某时段发送的微博数为 k_m，这天共有 K 个用户发送过微博，那么这天某时段的微博活跃度即为$\sum_{k=1}^{K} \frac{km}{kn*K}$，将需要计算的时间段数据进行累加得到累计活跃度，为了避免不同时长中数

据累加引起的差异，每个时间点数据需要根据累加天数进行平均。本文将最终计算得到的数据称为微博的“活跃度”。此外，为了方便后续的分析，对数据进行了宽度为 5.5 的高斯核平滑，并对平滑后的数据以 1.5 分钟为单位进行了线性插值。采集到的不同时段微博数量如图 1C 所示，经过计算、平滑和标准化后的活跃度如图 1D 所示。

2.2.2 识别微博用户的昼夜分布

用户几乎不会在睡眠时发送微博，而深夜是大部分用户睡眠的时间，因此微博活跃度最低的时段很有可能反映了用户的睡眠状态。本文参考前人的研究[25]，将已经过平滑和插值后的数据中位数的 25%作为阈值，计算阈值对应的两处时间点，以这两处时间点作为微博低活跃的开始时间和结束时间，并以它们差的绝对值作为低活跃时长。为了判断这种方法计算出的数据是否能够确实反映用户的睡眠模式，本文参考《2023 中国健康睡眠白皮书》[26]以及《2022 年中国睡眠指数报告》[27]中中国各地居民睡眠质量和失眠情况数据，与计算得到的数据进行了关联性分析。

2.2.3 昼夜节律扰乱程度计算

本文通过使用社会时差定量地表示微博用户昼夜节律被扰乱的程度。社会时差被定义为工作日和休息日失眠时间中点的差值，最初通过调查问卷询问调查对象的睡眠和起床时刻进行计算[19]，本文使用微博低活跃时段的时间中点代替计算。另外，不考虑调休，将所有周末和法定节假日作为休息日，其他日期均作为工作日，计算工作日和休息日的微博低活跃时间中点的差值作为社会时差。

另外，由于中国人具有午睡的习惯，这也是生活节律的一部分，但打乱午睡节律的情况及其产生的后果尚鲜有研究涉及，因此本文对此也进行了探究。午睡时间没有明显的 U 型曲线，较难取得中点，而 Leypunskiy 等人[25]的研究显示，通过计算不同向量之间的皮尔森系数，能够达到计算中点的效果。基于此，本文计算了休息日 12 时至 18 时活跃度构成的数据序列与工作日 8 时至 22 时的连续六个小时活跃度构成的序列之间的皮尔森系数，将系数最大的两个序列之间时间点的平移量作为“午睡时差”。另外，本文参考了《2020 中国卫生统计年鉴》[28]中的门急诊数据以探究打乱午睡节律是否可能和某些疾病具有关联。

2.2.4 不同地区的压力表征

人群面临的外部压力主要包括环境压力、人口压力和经济压力。为了能够在宏观层面上表征不同地区的压力，通过熵权法将各地统计年鉴中的数据合并为“压力指数”。具体来说，在环境压力方面选择工业废水排放量和工业二氧化硫排放量，人口压力则使用人口密度表征，在经济压力方面选择城镇登记失业率、人均可支配收入和人均消费支出表示。由于 2020 年采集到的数据相对较多且统计年鉴数据较为完整，本文仅使用 2020 年的数据进行分析。

另外，需要一种已被证实能够表征心理压力的指标以验证这种计算方法的合理性。研究表明，“百度搜索指数”中“压力”、“抑郁”和“焦虑”相对搜索量的和能够表征人群的心理压力水平[14]，因此本文将该搜索指数和计算得到的压力指数进行了关联性分析。

3 结果

3.1 低活跃时间与睡眠模式相关

为了探究微博活跃度的节律是否与居民的睡眠模式相关，本文将不同地区的低活跃时长与睡眠数据进行了比较。根据《2023 中国健康睡眠白皮书》[26]和《2022 年中国睡眠指数报告》[27]，发现低活跃时长与对应地区的失眠人数比率显著负相关(图 2A)，且与睡眠指数、睡眠质量和睡眠环境均显著正相关（图 2B-D），但与睡眠信念无关。其中，睡眠质量由匹兹堡睡眠质量和睡眠剥夺构成；睡眠环境包括社会环境、家庭环境和居住环境；睡眠信念反映了受试者对睡眠的非理性认知和睡眠拖延行为。睡眠指数则是由睡眠质量、信念和环境综合计算得出。这些结果表明，微博时间数据与居民睡眠模式具有关联，能够在一定程度上代表该地区居民的睡眠时长和睡眠质量，并反映昼夜节律。

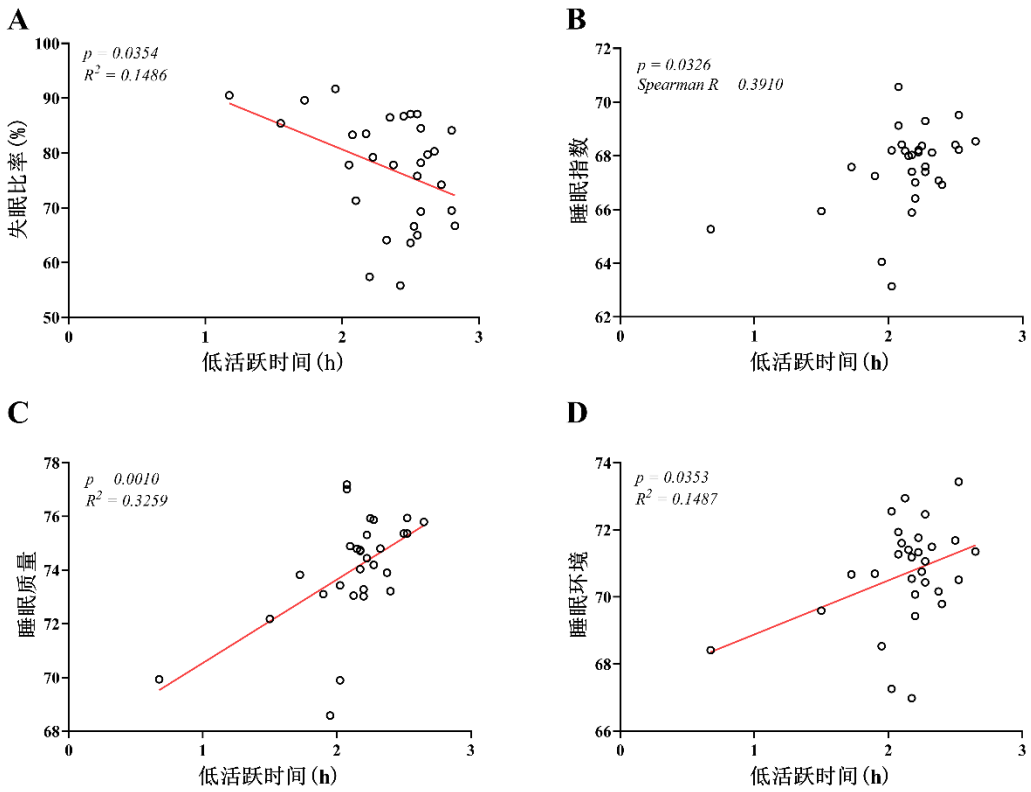


图 2 不同地区的微博低活跃时长和当地的睡眠情况之间的关系

A-D：不同地区低活跃时间分别与当地失眠比率、睡眠指数、睡眠质量和睡眠环境之间的关系

3.2 微博活跃度的空间分布

如图 3 所示，在不同地区上微博节律显示出显著的差异性。总体来说，南方沿海地区和北方地区低活跃时长较短，中部地区低活跃时长较长（图 3A，B）。分别选择南部、北部和中部的代表性地区，能够观察到不同地区的活跃度曲线形状的显著区别（图 3C）。这展现出不同地区居民的昼夜节律有所差异。

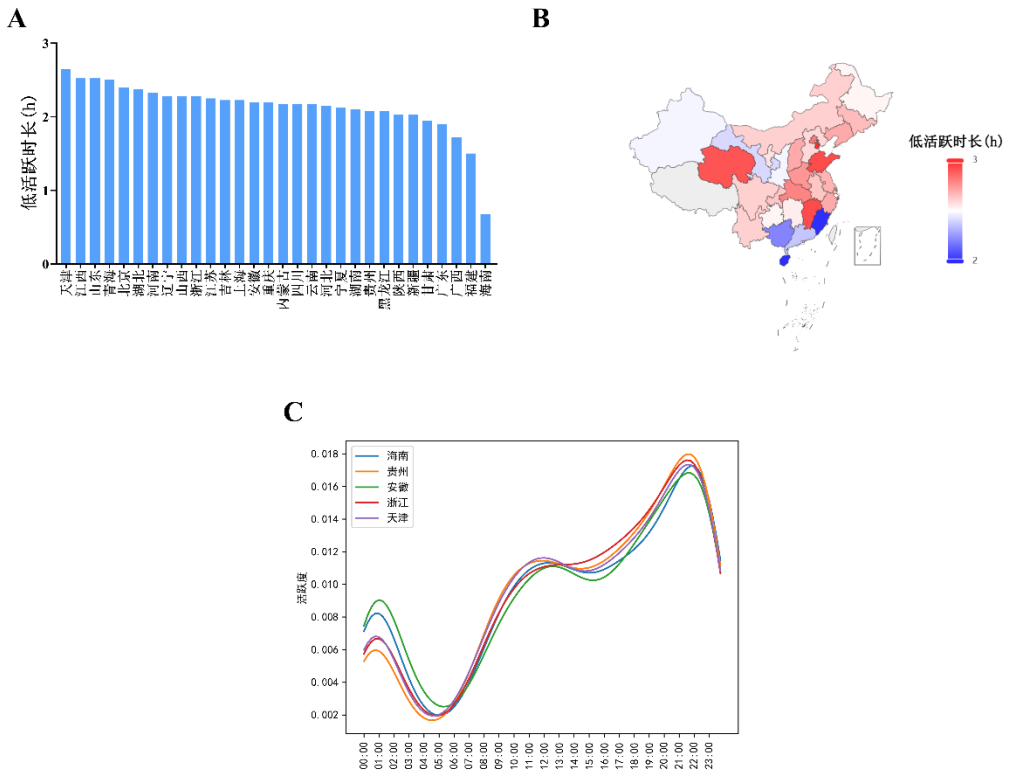


图 3 微博活跃度的空间分布

A，B：不同地区的低活跃时长；C：代表性地区的微博活跃度曲线。

3.3 微博活跃度的时间分布

微博活跃度表现出了明显的时间规律性和差异性。在不同时间尺度上，微博在 24 小时内的活跃度曲线形状均十分相似。以不同年份的曲线为例，如图 4A 所示，约 2:00 至 8:00 是微博活跃的低谷，这对应了夜间睡眠时间，此后活跃度上升，至约 12:00 达到第一个波峰，随后开始下降，约 16:00 开始再次上升，这大致对应了午睡时间。至约 22:00 达到最高峰，随后下降，而 0:00 至 1:00 又上升至一个小波峰，再降至低谷。以年为尺度时，不同时间的微博活跃曲线差异不明显(图 4A)，但仍能观察到不同年份低活跃时长之间的差异(图 4B)，即不同年份微博用户的睡眠模式具有差别。在月的尺度上，能够观察到每年 7-9 月低活跃时长较短（图 4C，Mann-Whitney U = 132.5，p = 0.0017）。值得注意的是，在 2021 年 5 月前后低活跃时间突然缩短，而在 9 月后则逐步恢复正常（图 4D）。

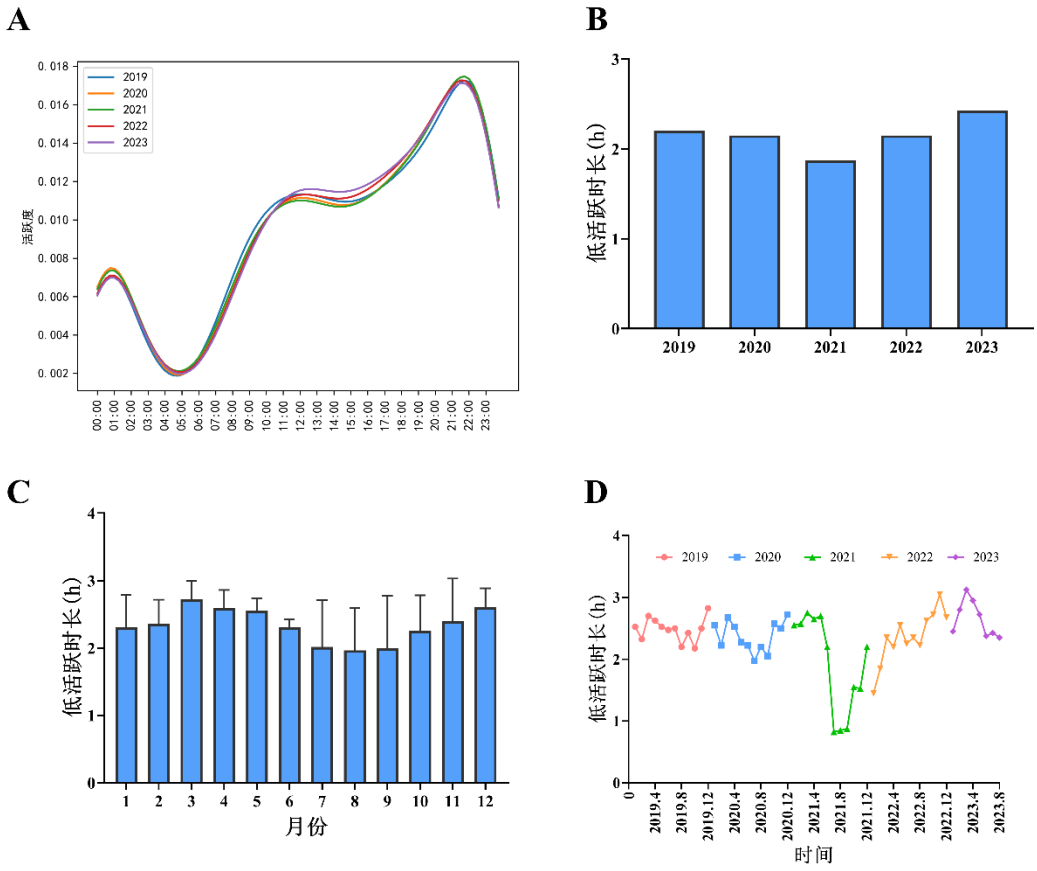


图 4 微博活跃度的时间分布

A：不同年份的微博活跃度曲线；B：不同年份的微博低活跃时长；C：不同月份的微博低活跃平均时长；D：不同月份的微博低活跃时长。

3.4 休息日低活跃时长更短且更晚出现

有趣的是，相比于工作日，休息日的低活跃时长更短，且往往比工作日更晚出现，几乎没有违反这两点规律的地区（图 5A，B）。另外，以月为尺度，这种规律同样适用，且随着时间的推移，休息日与工作日之间的低活跃时长之差、低活跃时间段出现时刻之差正在缓慢扩大（图 5C-F）。

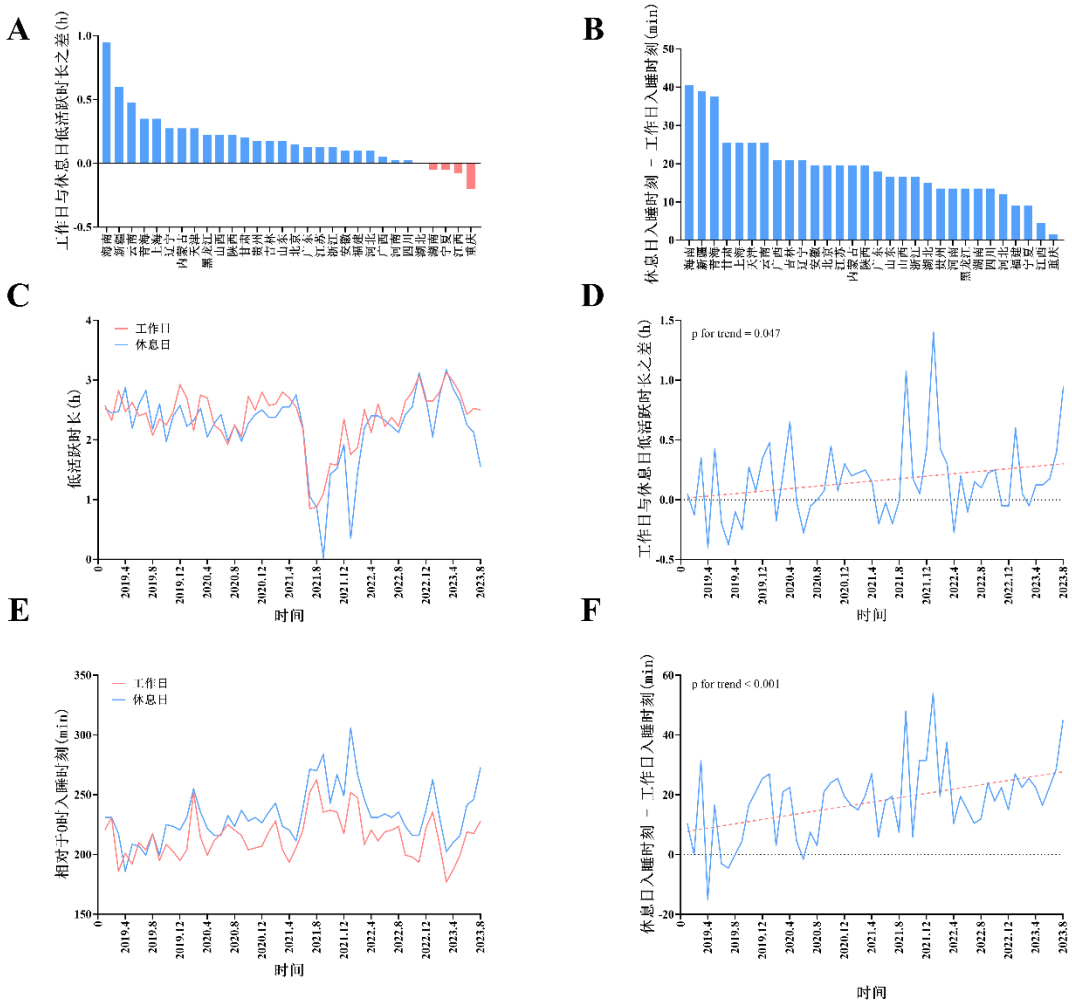


图 5 不同地区和时间的低活跃时长及其出现时刻

A，B：不同地区的低活跃时长和时刻差值；C，D：不同时间工作日与休息日低活跃时长及它们的差值；E，F：不同时间工作日与休息日低活跃时段出现时刻及它们的差值。

3.5 昼夜节律扰乱与压力的关系

本文进一步探究了昼夜节律的扰乱程度与外部压力之间的关系，首先分析了昼夜节律扰乱情况的时空分布。结果发现，社会时差在北方和南方部分地区较小，西北地区和中部地区则较大，并随着时间正缓慢增加(图 6A，C)。而午睡时差则在北方和中部地区较大，南方沿海地区较小，并随着时间缓慢降低(图 6B，D)。随后，对社会时差和午睡时差之间的相关性了探究。如表 1 所示，压力指数与“压力”、“抑郁”和“焦虑”组成的压力搜索指数显著正相关，这表明压力指数能够确实反

映某地区人群的压力状况。另外，压力指数与社会时差显著负相关，但其二级指标“环境压力”、“人口压力”和“经济压力”均与社会时差无显著性关系。这表示昼夜节律的改变是各方面压力综合作用的结果。此外，压力指数与午睡时差之间也未发现显著性关系。节律的改变对健康的影响已有不少研究进行了证实。然而，午休习惯的改变与精神科门急诊量显著正相关（见表 2），即休息日和工作日午休习惯的差异可能会导致罹患精神类疾病。

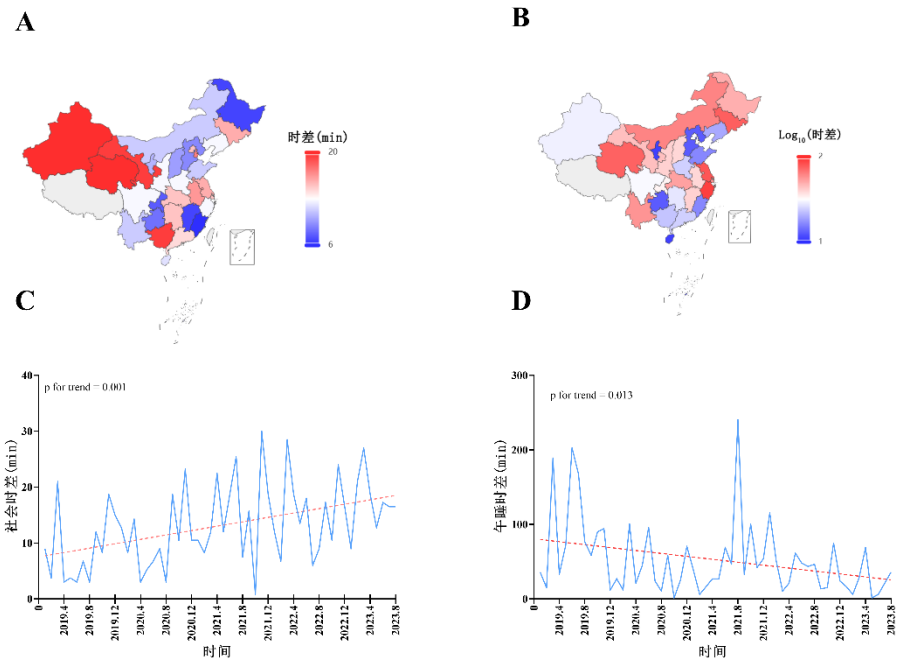


图 6 社会时差和午睡时差的时空分布

A，B：社会时差和午睡时差的空间分布；C，D：社会时差和午睡时差的时间分布。

表 1 各变量之间的相关系数

	压力指数	环境压力	人口压力	经济压力	搜索指数	社会时差	午睡时差
压力指数	1	.642**	.686**	.445*	.578**	-.414*	-0.077
环境压力	.642**	1.000	.413*	-0.070	.702**	-0.305	-0.201
人口压力	.686**	.413*	1.000	.101	.682**	-0.168	0.111

经济压力	.445*	-0.070	0.101	1	-0.083	-0.185	-0.034
搜索指数	.578**	.702**	.682**	-0.083	1	-0.358	0.167
社会时差	-.414*	-0.305	-0.168	-0.185	-0.358	1	-0.092
午睡时差	-0.077	-0.201	0.111	-0.034	0.167	-0.092	1.000

* 代表 $p < 0.05$，** 代表 $p < 0.01$

表 2 午睡时差和各科室门急诊量/常住人口的相关性分析

科室名	合计	预防保健	全科	内科	外科	儿科	妇产科	眼科
相关系数	0.149	0.047	0.111	0.113	0.145	0.231	0.088	0.143

科室名	耳鼻咽喉科	口腔科	皮肤科	医疗美容科	精神科	传染科	结核病科	肿瘤科
相关系数	0.155	0.198	0.219	0.287	0.481*	-0.022	0.138	0.239

科室名	急诊医学科	康复医学科	职业病科	中医科	民族医学科	中西医结合科	重症医学科
相关系数	0.372	0.226	-0.106	0.170	0.308	-0.214	0.158

* 代表 $p < 0.05$

4 讨论

4.1 微博适用于监测人群昼夜节律

广泛使用的社交媒体为了解人群健康状态提供了新的省时省力的方式[29]，但目前鲜有研究尝试据此对人群的睡眠状态和生活节律进行探究。本文发现，一方面，微博低活跃时长与该地区人群睡眠质量、失眠状况显著相关，即能够通过微博反映不同地区和时间人群的睡眠模式（图 2A-D），另一方面，在计算出的微博活跃度曲线中，约 2:00 至 8:00 是活跃度最低的时间段，随后缓慢上升至 12:00，再于 16:00 之前下降形成一个小波谷，这分别与夜间睡眠时间、午睡时间相对应。这些时间数据契合了用户睡眠和清醒的规律，表明可以通过微博了解人群的昼夜节律。

4.2 昼夜节律具有自身的时空分布规律

昼夜节律具有明显的时空分布规律。在空间上，不同地区的微博低活跃时长不同，这代表不同地区的居民睡眠模式具有差别（图 3A-C）。在时间上，7-9 月份的低活跃时长显著低于其他月份，这反映了人体睡眠节律随着季节、温度和光照自然变化的规律[30-32]（图 4C）。另外，在 2021 年 6 至 7 月份低活跃时长迎来了一次十分不寻常的断崖式减少，随后逐渐恢复，直到次年 4 月左右才基本回到原本水平（图 4D）。这一时期不仅与暑期重叠，且与艺人失德事件、双减政策逐渐落地等广受关注的社会事件发生时间相符合，这表明人群的睡眠状态可能受到社会因素的影响。

有趣的是，微博低活跃区间倾向于在工作日更早出现，且常在工作日持续更久（图 5C-F），这表明微博用户在工作日具有更好的睡眠质量和更长的睡眠时间，且用户在休息日可能有“报复性熬夜”的行为。另外，这种差距正在随着时间缓慢增长（图 5D，F），这可能会在某种程度上引发健康问题。

社会时差和午睡时差也具有自身的时空分布（图 6C，D）。其中，社会时差正随时间缓慢增加，而午睡时差正缓慢降低。社会时差的增加反映了微博用户的昼夜节律扰乱程度正在增加。鉴于昼夜节律与多种健康问题相关[33]，研究微博时间数据的时空分布规律将有助于为制定健康干预对策提供参考，但微博数据与具体病变之间的关系还需要更多数据支持。

4.3 昼夜节律的扰乱与压力的关系

通过使用环境压力、人口压力和经济压力合成的压力指数表征某地区人群面临的压力，发现该指标与社会时差显著负相关，即当外部压力增加时，会导致人群在工作日和休息日的昼夜节律趋于一致。由于本文的压力指数中包含了经济因素，而生存压力是加班的动机之一[34]，因此外部的压力增加可能会促进加班行为，或不得不在休息日处理工作日余留的任务。

4.4 紊乱的午睡节律和精神类疾病相关

尽管没有发现压力指数与午睡时差之间的显著性关系，但午睡时差同样具有其独特的时空分

布。同时，通过使用各地常住人口校正数据后，发现午睡时差与各地精神科门急诊量显著正相关，即紊乱的午睡规律可能与潜在的罹患精神类疾病风险有关联。

5 结语

本研究收集了微博数据，描述了用户的昼夜节律并分析了其扰乱程度与外部压力的关系，发现人群昼夜节律及节律的扰动表现出了明显的地域特性和时间规律。在自然规律上，表现出 7-9 月份睡眠时间和质量更低，在外部因素上，昼夜节律及节律扰动程度受到人口、经济、环境和社会等因素的影响。另外，昼夜节律被扰乱的程度正在随时间缓慢增加，而午睡节律的扰乱可能与精神类疾病相关，但这还需要更多的临床证据等支持。总之，微博数据适用于在大范围内通过昼夜节律角度进一步了解人群健康状态，这能为调查和解决广泛的昼夜节律失调等问题提供支持。然而，我们的数据规模相对于全国人口来说仍显得不足，且获得的用户样本以女性为主体，这可能会导致研究结果具有性别上的偏差；在未来的工作中，需要采集更多数据，补充男性用户。另外，微博数据包含传统睡眠数据中缺少的文本数据，通过深入分析文本，可能为人群昼夜节律分布的形成原因和扰乱情况的影响因素提供进一步的解释。

OR-025

A Random Forest Algorithm-Based Prediction Model for Moderate to Severe Acute Postoperative Pain After Orthopedic Surgery Under General Anesthesia

Gaoxiang Shi[1,2,3,4], Geliang Liu[2,3,5], Qichao Gao[1,2,3], Shengxiao Zhang[3,6,7], Qi Wang[1,2,3], Li Wu[1,2,3], Peifeng He*[2,3,7], Qi Yu*[2,3,6,7]

1. School of Basic Medical Sciences, Shanxi Medical University, Taiyuan, China.
2. Institute of Medical Data Science, Shanxi Medical University, Taiyuan, China.
3. Shanxi Key Laboratory of Big Data for Clinical Decision, Shanxi Medical University, Taiyuan, China.
4. Department of Anesthesiology, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Third Hospital of Shanxi Medical University, Taiyuan, China.
5. School of Management, Shanxi Medical University, Taiyuan, China.
6. Department of Rheumatology, Second Hospital of Shanxi Medical University, Taiyuan, China.
7. Key Laboratory of Cellular Physiology, Ministry of Education, Shanxi Medical University, Taiyuan, China.

Background

Postoperative pain is one of the most common complications after surgery, the incidence of moderate to severe postoperative pain varies from 25% to 66% according to the previous reports[1, 2]. The consequences of suboptimal postoperative pain control include negative effects on postoperative recovery, increased incidence of respiratory and circulatory complications, increased length of hospital stay and healthcare costs, as well as an increased risk of transition to chronic pain or neuropathic pain[3, 4]. With the advances in modern medicine, postoperative pain remains a challenge, hence improving pain control is an international initiative promoted by

multiple health organizations including WHO[5].

Orthopedic surgeries are considered to be some of the most painful procedures that have a variety of options for postoperative analgesia ranging from surgeon provided (e.g., local anesthesia) to more intensive techniques (e.g., nerve blockade or PCEA) requiring care from an acute pain service[6]. According to the author's clinical experience and some research reports, compared with patients receiving spinal anesthesia or regional anesthesia, acute postoperative pain is more severe in patients undergoing orthopedic surgery under general anesthesia[7, 8].

The ability to identify and focus care on patients at higher risk of moderate to severe postoperative pain would improve analgesia and patient satisfaction. The construction of a reliable postoperative pain prediction model based on risk factors can be applied in the early identification of orthopedic patients with a high risk of moderate to severe postoperative pain, which is vital in taking timely interventions to prevent pain from worsening.

Random forest algorithms can build a machine learning model based on sample data and be used to make predictions, and its performance advantages are mainly due to ensemble learning[9]. The previous studies demonstrated that the logistic regression model had limited performance in predicting acute postoperative pain[8, 10], while there have been no reports of using random forests to predict postoperative pain to the best of our knowledge.

Accordingly, we constructed machine learning models to predict moderate to severe acute postoperative pain of orthopedic patients under general anesthesia by identifying the risk factors. In addition, we evaluated the efficiency of the random forest algorithm-based prediction model by comparing it with the multivariate logistic regression-based model.

Materials and methods

This retrospective observational cohort study was conducted following the Declaration of Helsinki (as revised in October 2013). The study was approved and monitored by the Ethics Committee of Shanxi Bethune Hospital (Third Hospital of Shanxi Medical University). Because of the retrospective nature of the study and the patient's identity information has been concealed, the requirement for informed consent was waived by the Ethics Committee of Shanxi Bethune

Hospital. We present the following article in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting checklist[11]. The procedure of establishing moderate to severe acute postoperative pain prediction models in this study is shown in **Figure 1**.

Sample selection

Patients who underwent orthopedic surgery under general anesthesia in Shanxi Bethune Hospital from January 2020 to June 2020, were included in the study. The demographic and perioperative characteristics were extracted from the Electronic Medical Record (EMR) database.

Inclusion criteria

The inclusion criteria for this study were as follows: (1) Patients between the age of 18 and 100. (2) Patients underwent orthopedic surgery under general anesthesia.

Exclusion criteria

The exclusion criteria for this study were as follows: (1) Patients with chronic pain. (2) Patients with malignant tumors. (3) Patients retain an endotracheal tube after surgery. (4) Patients with cognitive dysfunction or who cannot communicate normally (5) Outpatient surgeries. (6) Incomplete clinical data.

Pain scoring methods and diagnosis of moderate to severe postoperative pain

The primary outcome was pain scores at rest on postoperative day one (POD1) using a visual analogue scale (VAS), with 0 representing no pain and 10 representing the most intense pain. Moderate to severe pain was defined as a VAS score of 4 or greater, which has been previously identified as a value at which patients request additional analgesias, become unsatisfied with pain control, and have interference with functional activity[12].

Variables

Demographic variables were defined and analyzed as follows: sex, age, and body mass index (BMI). Perioperative variables were physical status score based on the American Society of Anesthesiologists physical status classification (ASA score), operation score, timing of surgery (emergency surgery *vs* elective surgery), type of surgery (open surgery *vs* endoscopic surgery),

history of surgery or anesthesia, indwelling urinary catheters, indwelling drains, patient-controlled intravenous analgesia (PCIA) pumps, nerve blockade, arteriovenous catheterization, secondary surgery in a short period (within a month), duration of surgery, blood loss during surgery, time from withdrawal of medicine to awake, preemptive analgesia, immobilization, intraoperative blood transfusion, sufentanil consumption and tourniquet during surgery.

Feature selection

The pre-processed data were randomly split into training and testing sets. In the training set, demographic and perioperative characteristics above were selected as candidate risk factors because of previous reports and clinical experiences. After univariate analysis, the Lasso regression model was applied to screen the optimized variables by running cyclic coordinate descent. Age, duration of surgery, blood loss during surgery, time from withdrawal of medicine to awake, and sufentanil consumption were entered into the Lasso regression procedure as continuous variables. ASA score, timing of surgery, type of surgery, indwelling urinary catheters, arteriovenous catheterization, secondary surgery in a short period, immobilization, intraoperative blood transfusion, and tourniquet during surgery were entered as dichotomous variables. Lasso regression was generated using the glmnet package in R, the optimal lambda value was determined by 10-fold cross-validation, and the retained variables were selected as the input variables of the random forest models and multivariate logistic regression models.

Random forest modeling

The mlr3 package based on R was applied for random forest model construction and hyperparameter tuning. The data was resampled by using the bootstrapping/bagging method. AutoTuner functions were used for the automatic tuning of hyperparameters, the cross-validation technique was used to tune the number of estimators in the classifier, and all training was conducted with 10-fold cross-validation to prevent overfitting. All the indicators included in the risk prediction model were analyzed based on the mean decrease in accuracy and the mean decrease in the Gini coefficient.

Logistic regression modeling

The mlr3 package was applied for logistic regression model construction and hyperparameter tuning. The training set was conducted with 10-fold cross-validation to improve predictive performance and prevent overfitting. Independent risk factors were identified using a multivariate logistic regression model that entered variables selected in Lasso analysis, and odds ratio (OR) along with 95% confidence interval (CI) were calculated. The nomograms were applied to visualize the prediction model, the calibration curves were applied to visualize Hosmer-Lemeshow goodness-of-fit test, and the decision curves were used to determine clinical benefit.

Evaluation of machine learning models

The confusion matrixes, the accuracy of the classifiers ("classif.acc"), the classification error rate ("classif.ce") and the area under the receiver operating characteristic curve (AUC) were analyzed to evaluate the performance and clinical usefulness of the random forest classifier and the logistic regression classifier by comparing the predicted results with the true results.

Statistical analysis

Statistical analyses were performed using the RStudio software (version 2022.12.0-353), which runs R software (version 4.1.3; http://www.Rproject.org). Descriptive statistics were computed for all variables. These included means and standard deviations (SD) for continuous factors, and frequencies for categorical factors. Comparisons of the distribution of demographic variables and clinical characteristics were performed using the two-tailed t-test (or the Mann-Whitney test as appropriate) for continuous variables and the chi-square test (or the Fisher exact test as appropriate) for categorical variables. P values of 0.05 or lower were considered statistically significant.

Results

Patient Characteristics

A total of 327 patients were enrolled in this study, The rate of moderate to severe acute postoperative pain among all enrolled patients was 41.3%. After univariate analysis, fourteen characteristics were retained for subsequent Lasso analysis. The demographic and perioperative characteristics of all enrolled patients are shown in **Table 1**.

Filtered features for machine learning model establishing

Using the Lasso regression model, fourteen characteristics were tested for their ability to predict the clinical outcomes and to avoid overfitting. The Lasso coefficient profiles of features and the optimal penalization coefficient lambda (λ) are shown in **Figure 2**. The feature selection results revealed that nine variables, including age, indwelling urinary catheters, arteriovenous catheterization, secondary surgery in a short period of time, duration of surgery, intraoperative blood transfusion, immobilization, time from withdrawal of medicine to awake and tourniquet during surgery, could be used to predict moderate to severe acute postoperative pain (**Table 2**).

Random forest algorithm-based prediction model

A risk prediction model was constructed based on confirmed characteristics selected by the Lasso algorithm. The number of decision trees was set at 500, the “mtry” parameter was set at 3, the “min.node.size” parameter was set at 5, and the “max.depth” parameter was set at 6 according to the cross-validation algorithm and AutoTuner function. As shown in **Figure 3**, the mean decrease in accuracy and mean decrease in Gini for all indicators entered in the random forest model were analyzed. The mean decrease in accuracy showed that immobilization was the highest, followed by duration of surgery, blood loss during surgery, tourniquet during surgery, indwelling urinary catheters, etc. It refers to the degree of decrease in accuracy without the presence of this risk factor in the random forest model, which is equivalent to the classification contribution.

Logistic regression algorithm-based prediction model

To verify the efficacy of the random forest model, we constructed a multivariate logistic regression model to predict moderate to severe acute postoperative pain, the model is visualized in **Figure 4A**. Based on the multivariate analysis, three characteristics, namely shorter time from withdrawal of medicine to awake [OR 1.19, 95% CI (1.08, 1.31)], immobilization [OR 2.36, 95% CI (1.15, 4.85)], and indwelling urinary catheters [OR 2.39, 95% CI (1.09, 5.27)] were identified as independent risk factors. As shown in **Figure 4 (B and C)**, the calibration plots showed favorable consistency between the prediction of the logistic model and actual observations in

both the training and testing sets. Furthermore, As shown in **Figure 4 (D and E)**, the DCA plots showed that the logistic model was clinically useful and had good predictive ability in the training set.

Evaluation of predictor performance

The ROC curves of prediction models constructed by random forest and traditional logistic regression in the training and testing sets are shown in **Figure 5**. The AUC of the random forest algorithm-based prediction model in the training and testing sets were 0.972 and 0.810, respectively, which confirmed the good discrimination performance of the prediction model. Additionally, the AUC of the risk prediction model constructed by multivariate logistic regression in the training and testing sets were 0.781 and 0.764, respectively.

The accuracy and the error rate were applied in testing the reliability of prediction models in our study. The accuracy values of the random forest algorithm-based prediction model and multivariate logistic regression-based prediction model in the training set were 0.882 and 0.724, respectively. The accuracy values of the random forest algorithm-based prediction model and logistic regression-based prediction model in the testing set were 0.747 and 0.687, respectively.

Discussion`

Despite extraordinary advances in anesthesia and analgesia, a significant proportion of patients still suffer from moderate to severe pain after surgery, yet treatments and interventions for these patients are lacking[13, 14]. According to previous reports, the incidence of moderate to severe postoperative pain can be up to 66% in the United States[1]. In particular, orthopedic patients have a higher incidence of postoperative pain. In a study of 10008 patients in Canada who underwent surgery, the incidence of acute postoperative pain was highest in orthopedic patients[15]. In this study, 41.3% of orthopedic surgery patients under general anesthesia experienced moderate to severe postoperative pain.

Early identification of patients who underwent orthopedic surgery under general anesthesia with a high risk of moderate to severe acute postoperative pain is helpful for early intervention and improving analgesic effect. There are currently no models for predicting postoperative pain in the

patient of orthopedic surgery under general anesthesia, while generic postoperative pain prediction models do not accurately predict the degree of acute pain after orthopedic surgery. In addition, risk factors reported in some studies varied widely[8, 10, 16-20] (**Table 3**). In this study, we constructed a reliable risk prediction model with high discriminatory ability, which is helpful in building personalized treatment plans for patients with an increased risk of acute postoperative pain.

Most characteristics of orthopedic surgery patients between group P and group N were significantly different, so it is possible to use them to predict the clinical outcomes. Several risk prediction models were constructed to predict acute postoperative pain by typically performing univariate regression followed by multivariate logistic regression, resulting in reduced prediction accuracy. As an ensemble learning algorithm for classification, random forest is performed by constructing numerous decision trees at training time and outputting the class that is the mode of the classification of the individual trees. Compared with multivariate logistic regression, the random forest algorithm has higher accuracy in classification or prediction tasks and does not require strict assumptions about raw data[21, 22]. We applied the mlr3 package in R to establish and validate a random forest-based prediction model, which has a high ability to handle a multitude of input variables and evaluate the missing data to maintain the prediction accuracy[23]. In this study, the results of ROC analysis showed that the random forest algorithm-based prediction model had higher predictive accuracy than the logistic regression-based model in both the training and the testing sets. To our knowledge, this study is the first attempt to use random forests to predict acute postoperative pain severity in patients undergoing orthopedic surgery under general anesthesia. Our findings demonstrate the potential of random forest algorithms in predicting acute postoperative pain.

Currently, many previous studies reported that some demographic characteristics were associated with moderate to severe acute postoperative pain in patients undergoing orthopedic surgery, such as sex, age, and BMI[20, 24]. In this study, after the univariate screening, age was entered into the multivariate logistic regression model and random forest model, but it was not an

independent risk factor. Similarly, studies have found that factors such as age are associated with only statistically significant but not clinically significant associations with postoperative pain[25].

In this study, the results demonstrated that the duration of surgery, and blood loss during surgery were significantly associated with acute postoperative pain, which may be related to surgical complexity or surgical trauma size. Abrecht et al.[18] used temporal summation of pain (TSP) to predict postoperative pain accurately. Duration of surgery and blood loss during surgery may be reflections of TSP[26]. Some studies suggest that acute postoperative pain is mainly related to patients rather than surgical factors[16]. In contrast, our study found that postoperative pain was associated with surgical and anesthesia factors. In addition, the use of tourniquets during surgery, indwelling urinary catheters, and arteriovenous catheterization reflects pain from multiple causes other than surgical procedures[27, 28], all of these factors have the potential to predict postoperative pain severity.

In previous studies, it has been reported that preoperative pain can increase the incidence of acute postoperative pain[29]. In this study, two factors namely immobilization before surgery and secondary surgery in a short period of time caught the attention. These factors are related to the preoperative pain experience, immobilization is generally used in patients with fractures, and secondary surgery in a short period of time may indicate recently experienced pain. The ability of these two factors to predict postoperative pain has not been reported and can be further investigated in the future.

Time from withdrawal of medicine to awake defines the period of time from the cessation of the general anesthetic infusion to the time when the patient becomes conscious. This characteristic was extracted from the patient’s electronic anesthesia records based on our clinical experience. As far as we know, it has not been used in other studies so far. In this study, the characteristic was found to be an important risk factor or predictor of moderate to severe acute postoperative pain. In general, insufficient intraoperative analgesia leads to earlier awakening[30], so we speculate that this characteristic may reflect the adequacy of intraoperative analgesia and may

be a potential predictor of acute postoperative pain.

For the risk factors identified in this study, orthopedic surgeons, anesthesiologists, and nurses need to focus on these factors in their daily work and effectively intervene to reduce acute postoperative pain. For the model constructed in this study, we should investigate whether it can be applied to other populations of patients for predicting the degree of postoperative pain.

Some limitations of this study are worth mentioning. First, our study was retrospective. In our study, we included as many variables as possible, however, there were still a few characteristics that were not included, such as preoperative pain scores, smoking and drinking habits, psychosocial factors, etc.. Therefore, some valuable factors may not be included. Further studies are needed to investigate whether adding these variables could improve the accuracy of the prediction model. Second, the datasets in our study were collected from a single center and were not large enough. Further studies with large multi-center samples are needed. Last, as a real-world clinical study, the postoperative management of these patients employed different methods for pain management, which varied depending on the patient's condition, likely contributing somewhat to the variability in pain scores between individuals. Therefore, the risk factors screened out in this study need to be verified by rigorous RCT studies in the future.

Conclusions

This random forest model is an excellent predictor that can be used to predict the severity of postoperative pain in orthopedic surgery patients under general anesthesia, which is of potential clinical application value.

List of abbreviations

ASA: the American Society of Anesthesiologists; AUC: area under the receiver operating characteristic curve; BMI: body mass index; CI: confdence interval; DCA: decision curve analysis; EMR: electronic medical record; Lasso: the least absolute shrinkage and selection operator; OR: odds ratio; PCEA: patient-controlled epidural analgesia; PCIA: patient-controlled intravenous analgesia; RCT: randomized controlled trial; ROC: receiver operating characteristic; SD: standard deviations; TRIPOD: the transparent reporting of a multivariable prediction model for individual

prognosis or diagnosis; TSP: temporal summation of pain; VAS: visual analogue scale; WHO: World Health Organization.

Declarations

Ethics statement

1. Ethical approval

The study was approved and monitored by the Ethics Committee of Shanxi Bethune Hospital (Third Hospital of Shanxi Medical University).

2. Statement of human rights

All procedures in this study were conducted in accordance with the Ethics Committee of Shanxi Bethune Hospital (Third Hospital of Shanxi Medical University) approved protocols.

3. Statement of informed consent

Informed consent for patient information to be published in this article was not obtained because of the retrospective nature of the study and the patient's identity information has been concealed. The requirement for informed consent was waived by the Ethics Committee of Shanxi Bethune Hospital.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analyzed during this current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Department of Education of Shanxi Province [2021Y364], and the Key R&D Projects of Shanxi Province [202102130501003] & [201903D311011].

Authors' contributions

Guarantor of integrity of the entire study: QY and PH; Study design: GS and GL; Data analysis:

GS, QG, and QW; Manuscript preparation: GS; Manuscript editing: LW; Manuscript review: QY and SZ. All authors read and approved the final manuscript.

Acknowledgements

We thank the staff of the Anesthesiology Department of Shanxi Bethune Hospital for their help in clinical data collection.

Table 1 The demographic and perioperative characteristics of all enrolled patients

Variables	Total patients（*n*=327）	Group P（*n*=135）	Group N（*n*=192）	*t* / χ^2 / *Z*	*P*-value
Demographics					
Sex				0.3378	0.5611
Male	200	86	115		
Female	127	49	77		
Age, years	51.63±14.38	48.47±14.14	53.84±14.16	3.3781	0.0008
BMI, kg/m^2	24.66±3.92	24.27±4.12	24.93±3.76	1.4772	0.1408
Perioperatives					
ASA score				8.8704	0.0029
I or II	218	77	141		
III or IV	109	58	51		
Operation score				0.55496	0.4563
I or II	35	17	18		
III or IV	292	118	174		
Timing of surgery				11.634	0.0006
emergency surgery	62	38	24		
elective surgery	265	97	168		
Type of surgery				4.7037	0.0301
open surgery	289	126	163		
endoscopic surgery	38	9	29		
History of surgery or anesthesia				0.58757	0.4434
yes	86	32	54		
no	241	103	138		
Indwelling urinary catheters				14.806	0.0001
yes	227	110	117		
no	100	25	75		
Indwelling drains				2.2149	0.1367
yes	300	128	172		
no	27	7	20		
PCIA				2.444	0.118
yes	272	118	154		
no	55	17	38		

Nerve blockade				0.7377	0.3904
yes	49	17	32		
no	278	118	160		
Arteriovenous catheterization				14.299	0.0002
yes	80	48	32		
no	247	87	160		
Secondary surgery in a short period				4.9972	0.0253
yes	19	13	6		
no	308	122	186		
Duration of surgery, minutes	124 (170－80)	135 (185－104.5)	110 (159.25－72)	11.0483	7.286e-05
Blood loss during surgery, ml	150 (400－100)	200 (475－100)	150 (300－57.5)	11.1878	9.011e-05
Time from withdrawal of medicine to awake, minutes	15.77 ± 3.76	15.10 ± 3.81	16.24 ± 3.66	2.7464	0.006
Preemptive analgesia				0.21917	0.6397
yes	243	98	145		
no	84	37	47		
Immobilization				29.615	5.27e-08
yes	104	66	38		
no	223	69	154		
Intraoperative blood transfusion				10.794	0.0010
yes	95	53	42		
no	232	82	150		
Sufentanil consumption, μg	45 (50－40)	50 (60－40)	40 (50－35)	11.8357	0.0003
Tourniquet during surgery				52.177	5.071e-13
yes	65	53	12		
no	262	82	180		

Table 2 The characteristics of all enrolled patients（Training set *vs* Testing set）

Variables	Total patients（*n*=327）	Training set（*n*=228）	Testing set（*n*=99）	*t* / χ^2 / *Z*	*P*-value
Demographics					
Age, years	51.63 ± 14.38	52.32 ± 13.74	50.03 ± 15.71	-1.2564	0.2107
Perioperatives					
Indwelling urinary catheters				0.9726	0.324
yes	227	154	73		
no	100	74	26		
Arteriovenous catheterization				0.84331	0.3585
yes	80	52	28		
no	247	176	71		
Secondary surgery in a short period				0.016849	0.8967
yes	19	14	5		
no	308	214	94		
Duration of surgery, minutes	124 (170－80)	124 (166.25－81.5)	113 (170－79)	2.1046	0.673
Blood loss during surgery, ml	150 (400－100)	150 (400－80)	200 (400－100)	1.3283	0.572
Time from withdrawal of medicine to awake, minutes	15.77 ± 3.76	15.82 ± 3.67	15.66 ± 3.97	-0.34078	0.7337
Immobilization				0.60667	0.436
yes	104	69	35		
no	223	159	64		
Tourniquet during surgery				4.2322	0.03966
yes	65	38	27		
no	262	190	72		
Moderate to severe postoperative pain				0.11243	0.7374
yes	135	96	39		
no	192	132	60		

Table 3 The risk factors of moderate to severe acute postoperative pain in previous studies.

Study (year)	Country	Number of patients	Type of surgery	Risk factors / protective factors
Vasilopoulos et al. (2021)	USA	360	mixed surgery	younger age, female gender, higher anxiety, and more pain behaviors.
Sun et al. (2020)	China	1164	thoracic surgery	younger age, high BMI, preoperative pain, smoking history, and number of chest tubes.
Abrecht et al. (2019)	USA	126	orthopaedic surgery	temporal summation of pain, high BMI, number of previous knee surgeries, and female gender.
Zaslansky et al. (2018)	International	14334	orthopaedic surgery	female gender, younger age, high BMI, chronic pain, and opioid use before surgery.
Hartwig et al (2017)	International	192	gastric surgery	younger age and preoperative pain.
Borges et al. (2016)	Brazil	1062	cesarean section	preoperative anxiety, intrathecal morphine with fentanyl#.
Liu et al. (2012)	USA	897	orthopedic surgery	female gender, younger age, high BMI, preoperative pain, preoperative use of opioids, general anesthesia, preoperative use of anti-convulsants and anti-depressants, and prior surgery at the surgical site.

represents the protective factors

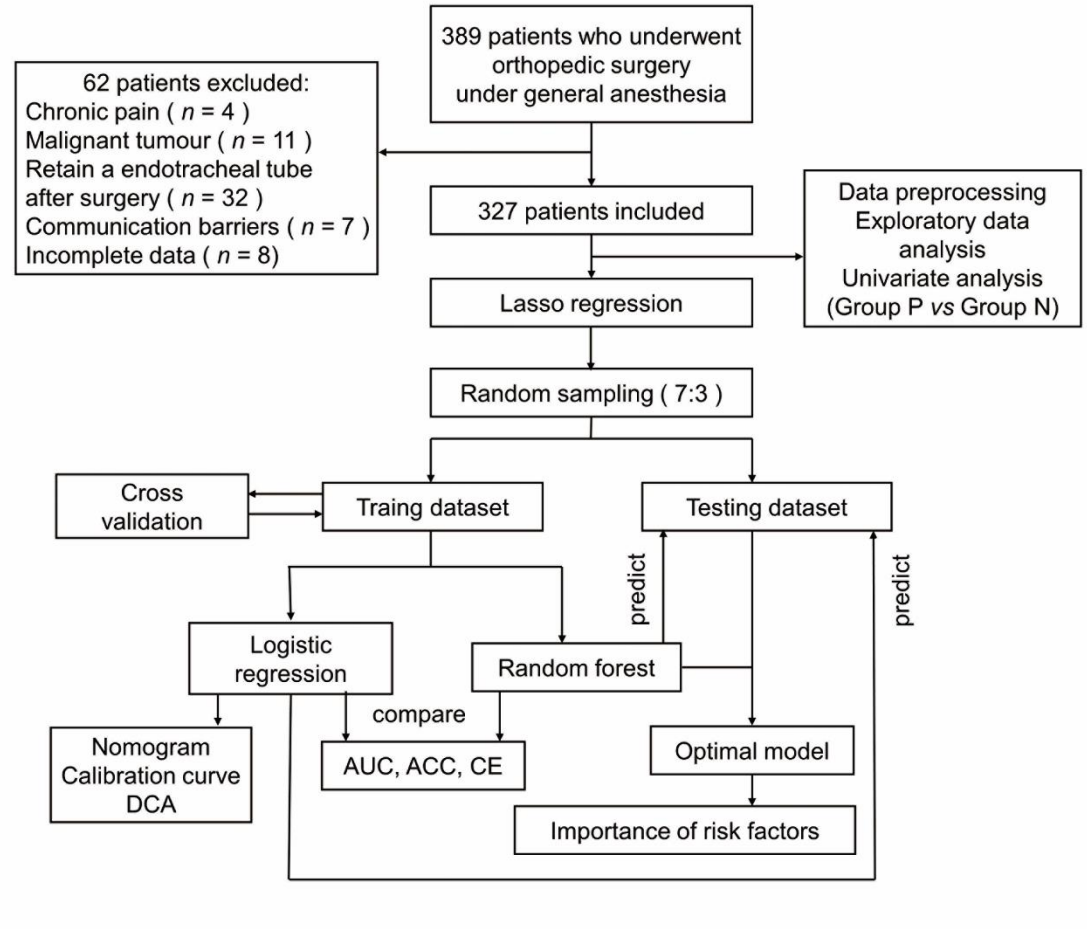


Figure 1 The procedure of establishing moderate to severe acute postoperative pain prediction models in this study.

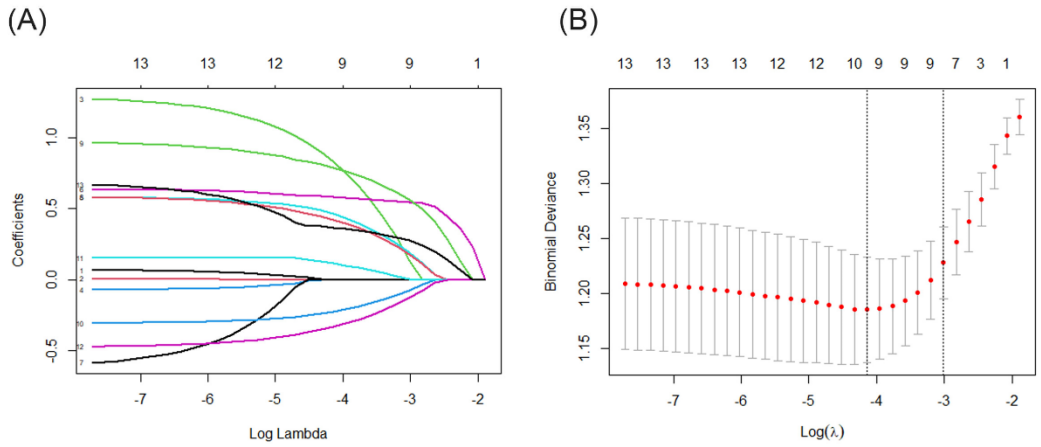


Figure 2 (A) Lasso coefficient profiles of all candidate features. **(B)** The tuning parameter λ (lambda) selection in the Lasso models used 10-fold cross-validation by minimum criteria.

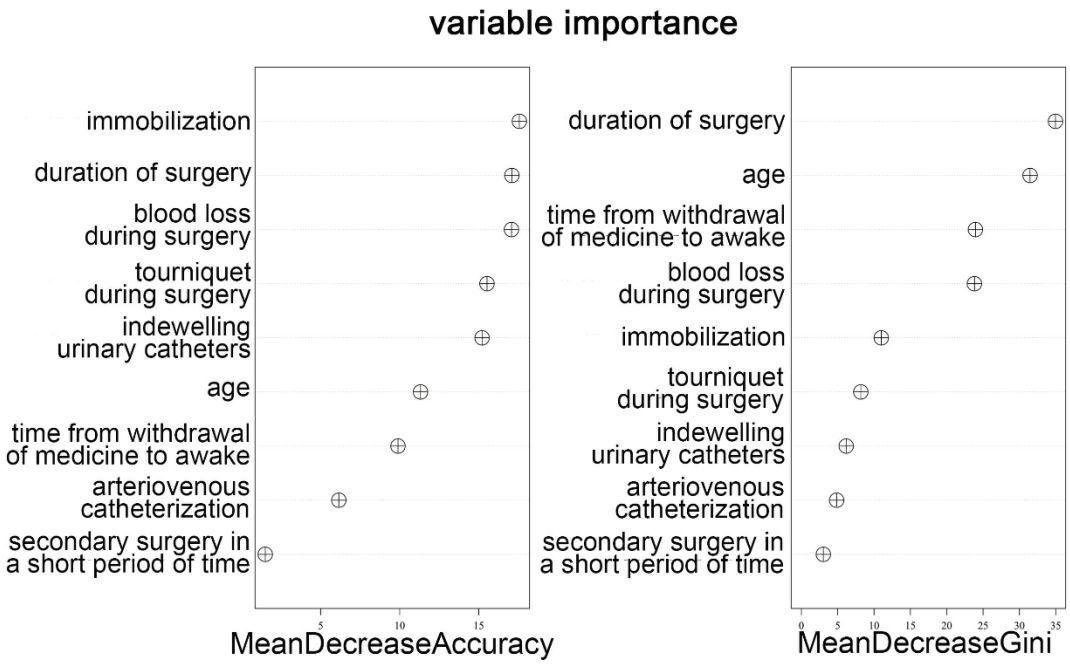


Figure 3 Importance of risk factors in the prediction model constructed by random forest.

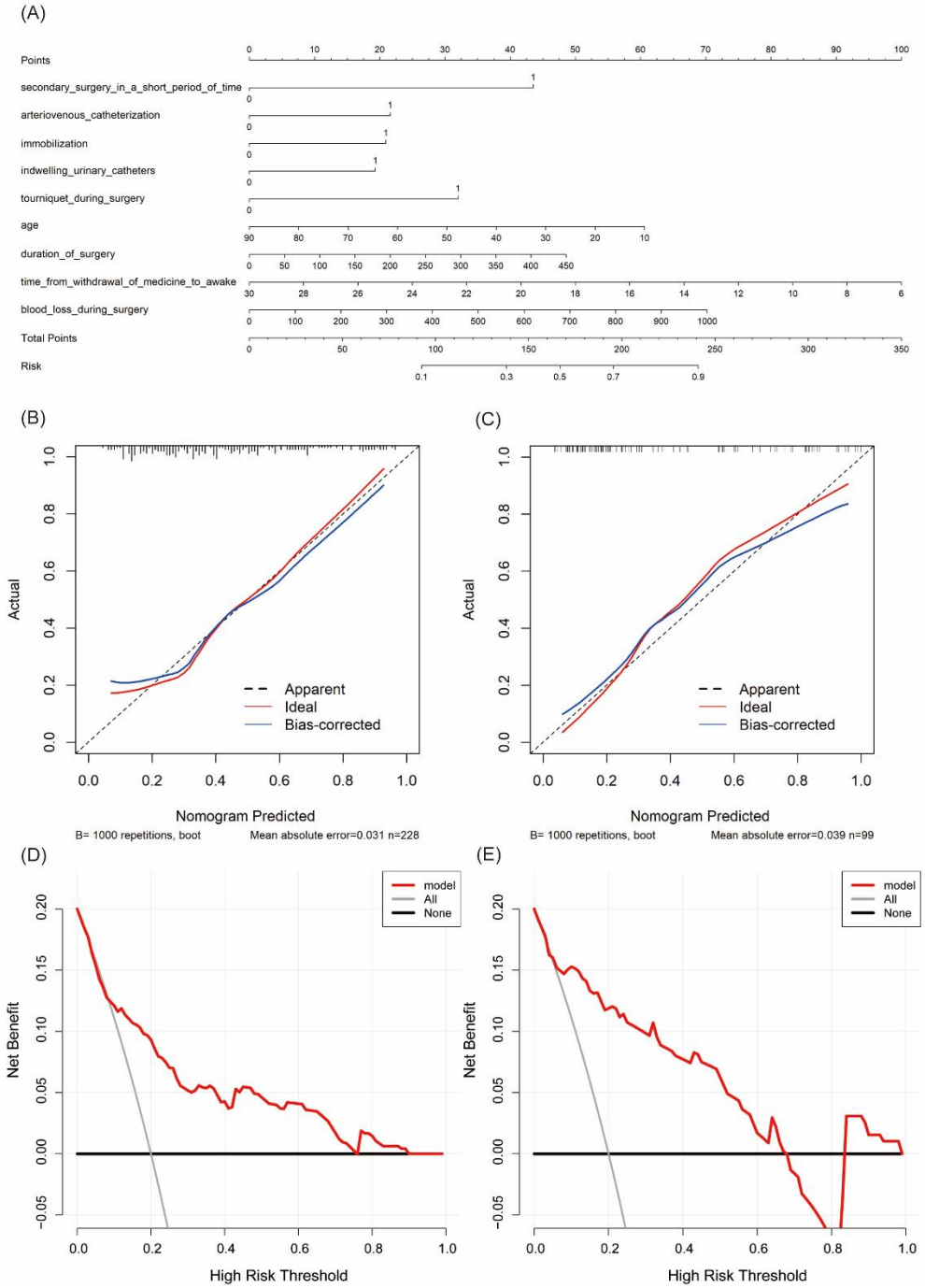


Figure 4 Visualization and performance evaluation of the predictive model based on multivariate logistic regression. **(A)** The nomogram. **(B)** The calibration curve in the training set. **(C)** The calibration curve in the testing set. **(D)** The decision curve in the training set. **(E)** The decision curve in the testing set.

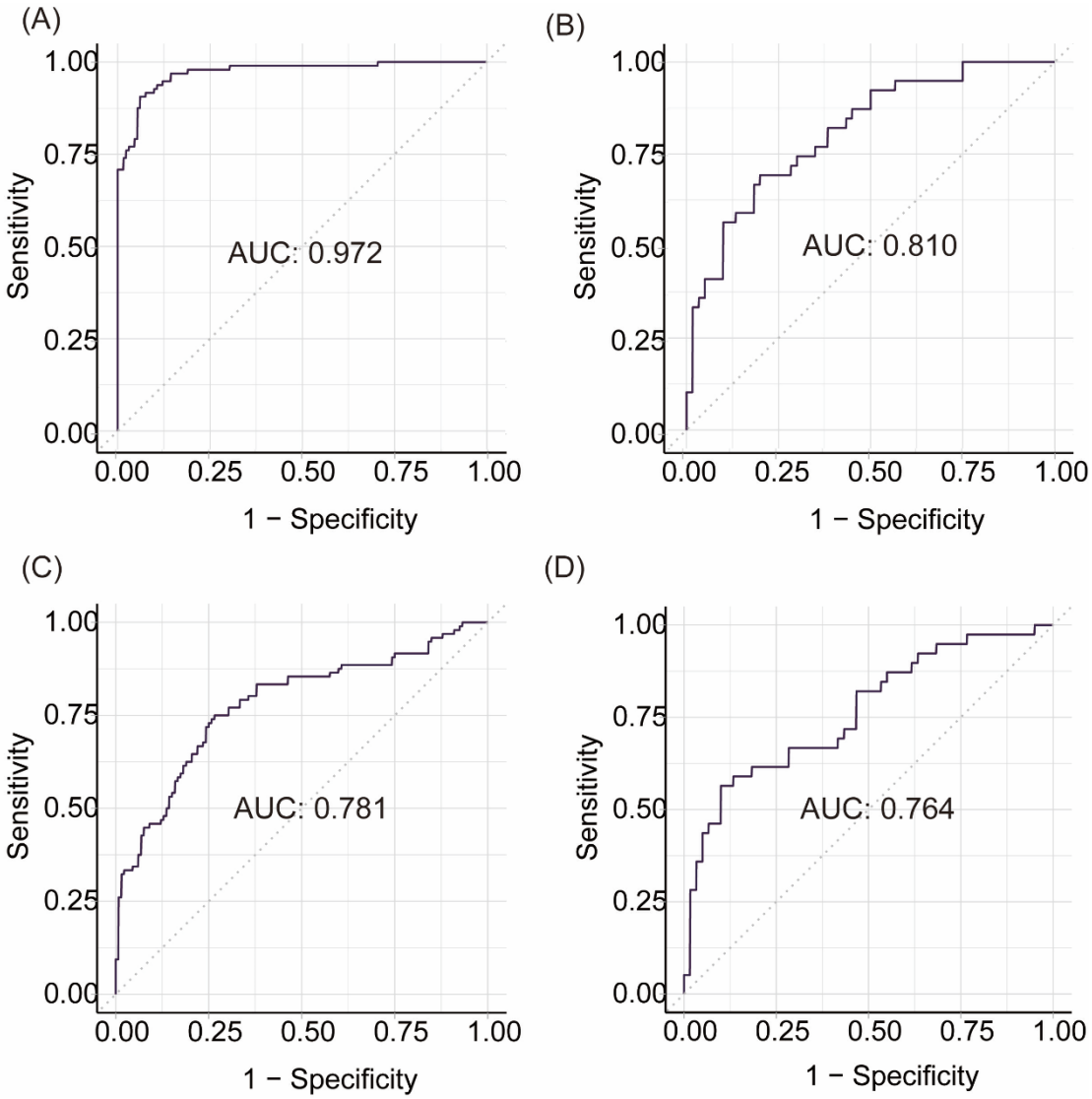


Figure 5 The ROC curves of the models in this study. **(A)** The ROC curve of the prediction model constructed by random forest in the training set. **(B)** The ROC curve of the prediction model constructed by random forest in the testing set. **(C)** The ROC curve of the prediction model constructed by multivariate logistic regression in the training set. **(D)** The ROC curve of the prediction model constructed by multivariate logistic regression in the testing set.

OR-026

基于因果表征学习的多中心结直肠癌预后预测模型

秦园炳 田雨 周天舒 池胜强 李军 丁克峰 李劲松*
生物医学工程与仪器科学学院，电子病历与智能专家系统教育部工程研究中心，生物医学工程教育部重点实验室，浙江大学

1 引言

结直肠癌作为最常见的消化道恶性肿瘤之一，我国结直肠癌发病率居恶性肿瘤第二位，死亡率居第四位，且发病率和死亡率呈现逐年上升趋势[1]。每年有超过 120w 的患者被确诊为结直肠癌，而有超过 60w 名患者直接或间接死于结直肠癌[2]。因此，寻找有效的预后预测策略对提高 CRC 的生存率至关重要。

传统上，TNM 分期[3]和 Cox 回归模型等方法常用来预测患者预后。其中基于 TNM 分期的非参数统计方法，基于肿瘤的大小（T）、淋巴结转移情况（N）和远处转移（M）等因素进行分类缺少对于其他因素的考虑，对于同一分期的患者之间可能存在显著预后差异。基于COX回归的半参数化方法，通过估计风险比评估不同因素对生存时间发生的影响，对数据中的随机性和噪声非常敏感，会对结果造成较大误差。

基于数据驱动的机器学习（ML）方法允许更有效的利用多维数据来准确预测生存率。然而，具有经验风险最小化的机器学习算法通常具有较差的泛化性能，当训练数据和测试数据属于不同分布时可能会导致模型的显著失败，例如用于预测危重患者死亡率的概率模型在首次开发时具有良好的判别性能，但在大规模、多中心临床应用评价中其预测性能仅为中等[4]。对于临床决策而言，预测结果的稳定性更为重要。因此，预测模型的泛化性对于确保其在实际临床环境中提供可靠的决策支持具有重大意义。

近年来，因果学习受到广泛关注。通过去除不相关特征和目标变量之间的关联，这些特征往往会随着环境的变化而变化，仅基于与真实标签有因果关系的特征进行预测，以在不同的数据分布下获得稳定的性能。受[5]-[8]的启发，我们引入了因果学习策略构建了一种从结直肠癌患者 EHR 数据出发，在不同医疗机构保持性能良好且稳定的结直肠癌预后预测的方法并提出如下结构：

1、我们引入了结构因果模型（SCM[9]）来描述 EHR 患者数据特征之间的关系。

2、我们使用 Lasso 和临床知识提取了与目标变量有因果关联的种子变量，基于 Pearson 相关系数和 Fisher's z 变换对所有 EHR 特征进行了条件独立性测试，将变量划分为因果特征集和非因果特征集。

3、对于因果特征集，我们构建了一个多中心预后预测模型，对于非因果特征集，我们构建了一个医疗中心识别网络。在训练阶段医疗中心识别网络生成样本权重辅助预后预测模型的更新，在预测阶段，医疗中心识别网络聚合多中心预后预测模型的结果。

2 材料和方法

2.1 数据集

我们使用了来自美国和中国的 CRC 数据。美国的数据来自美国国家癌症研究所(NCI)的监测、流行病学和最终结果(SEER)项目，包含 16 个医疗机构，中国的数据来自浙江大学第二附属医院(SAHZU)。SEER 和 SAHZU 分别包含 124636 条和 959 条记录，我们使用的数据详细描述如表 1 所示。

表 1 结直肠癌预后特征描述

名称	等级/IQR	数量
年龄	<55	25979 (20.68%)
	55-64	27868 (22.19%)
	65-74	31087 (24.75%)
	>=75	40661 (32.37%)
性别	男	62292 (49.60%)
	女	63303 (50.40%)
等级	良好区分	11272 (8.97%)
	中等区分	90972 (72.43%)
	较差区分	23351 (18.59%)
组织学	腺癌	93114 (74.14%)
	粘液或印戒细胞	12626 (10.05%)
	其他	19855 (15.81%)
位置	结肠	97372 (77.53%)
	直肠	28223 (22.47%)
T 分期	T1	13071 (10.41%)
	T2	22784 (18.14%)
	T3	74345 (59.19%)
	T4a	8998 (7.16%)
	T4b	6397 (5.09%)
N 分期	N0	76930 (61.25%)
	N1a	15119 (12.04%)
	N1b/N1c	15899 (12.66%)
	N2a	9730 (7.75%)
	N2b	7917 (6.30%)

名称	等级/IQR	数量
肿瘤大小	IQR(mm)	45.52±30.14
阳性淋巴结个数	IQR(int)	1.48±3.19
生存月数	IQR(month)	45.32±32.50
5年生存	存活	33888(26.98%)
	死亡	91707(73.02%)

2.2 方法

(1) 种子变量筛选模块

SCM 是一种用于描述和分析因果关系的统计模型，用于表示变量之间的因果关系和条件独立性。

在有向无环图中，如果路径 π 满足以下条件，则称路径 π 被节点集合 Z 而 d 分离：（1）π 包含一个链 $V_i \rightarrow V_k \rightarrow V_j$ 或 $V_i \leftarrow V_k \rightarrow V_j$，其中中间节点 V_k 位于 Z 中；（2）π 包含一个碰撞器 $V_i \rightarrow V_k \leftarrow V_j$，其中中间节点 V_k 不在 Z 中，并且 V_k 的后代节点都不在 Z 中。如果在有向无环图中，变量 V_i 和 V_j 由 Z 节点 d 分离，则 V_i 在给定 Z 的条件下独立于 V_j。

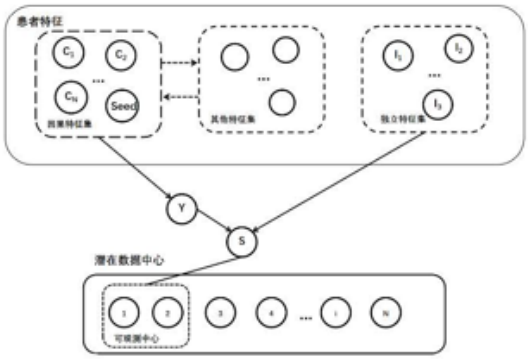


图 1 多中心预后预测模型的 SCM 模型

如图 1 所示，所有的输入变量即 EHR 数据的特征被分为了因果特征集 C、独立特征集 I 以及其他特征集。Y 表示感兴趣变量即预后结果，S 代表数据所属数据分布，C 表示导致 Y 的因果变量，Seed 表示因果特征集中可以观测并确定的其中一个因果变量，I 表示和 Y 没有直接关联的变量，但是由于选择偏差 S 的存在，即所观测的数据中心是真实世界所有数据中心的一部分，从而导致 I 和 Y 之间产生了相关性，也正是由于独立特征集 I 在不同医疗机构下差异较大从而导致模型拟合过程中如果将这一部分特征纳入会导致模型在不同医疗机构下预测结果的准确性发生显著失败。

为了最大范围利用 EHR 数据并且减少独立特征 I 对于模型不稳定性的影响。我们提出通过种子变量将 EHR 特征划分为因果特征集 C 和非因果特征集 I。基于 d 分离的定义，因果特征$C_i \in C$和独立特征$I_k \in I$基于 Y 满足 d 分离的条件，因此对于所有$I_k \in I$和$C_i \in C$满足$I_k \perp\!\!\!\perp C_i|Y$，对于$C_j \in C$和$C_i \in C$，$C_j$和$C_i$在变量 Y 条件下都不满足条件独立性。

T 分期、N 分期、M 分期被普遍用来作为 CRC 预后预测的预后因子[3]。在此基础上，采用 Lasso 回归对变量进行进一步筛选。如图 2 所示，采用 EHR 数据的所有特征作为 Lasso 回归模型的输入，其中 T 分期的系数显著大于 N 分期，仅次于年龄分组，因此我们将 T 分期作为 Seed 变量。

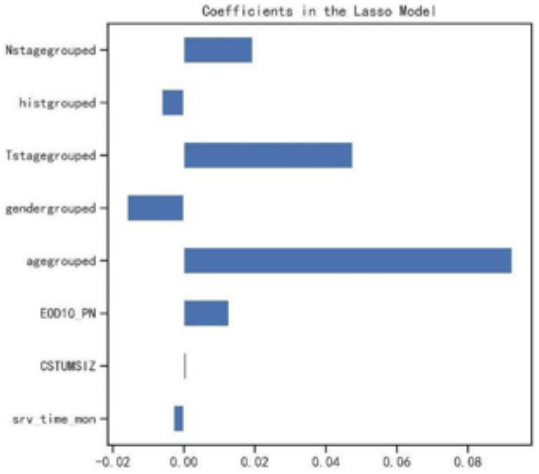


图 2 Lasso 回归模型变量系数

(2) 条件独立性测试模块

皮尔逊相关系数是衡量两个变量之间线性相关程度的统计量，如果条件独立性测试中的 $r(X,S|Z)$接近于 0，则可以认为在给定 Z 的条件下，X 和 Y 是近似独立的。

$$r_{X,S|Y} = \frac{r_{X,S} - r_{X,S} * r_{Y,S}}{\sqrt{(1-r_{X,S}^2)*(1-r_{Y,S}^2)}} \quad (1)$$

公式中$r_{X,S|Y}$表示在 Y 条件下，X 和 S 变量的相关系数。

Fisher's Z 变换是一种统计技术，用于将相关系

数转换为符合正态分布的变量。转换后的值 Z 近似地符合正态分布，可以用于各种假设正态性的统计分析。

$$z = 0.5\ln\left(\frac{1+r_{X,S|Y}}{1-r_{X,S|Y}}\right) \quad (2)$$

公式中 z 表示在将$r_{X,S|Y}$转换为正态分布后的转换值。当满足$z\sqrt{n-|Y|-3} < \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$时，即可判定变量 X 和 S 基于变量 Y 条件独立。其中 n 表示样本个数，Y 表示条件变量集长度，Φ表示标准正态分布函数，α表示置信度。

基于筛选的种子变量，本模块测试了所有 EHR 特征在感兴趣变量 Y 条件下与种子变量 T 分期的条件独立性，测试结果如图 3 所示，纵轴表示不同医疗中心，横轴表示 EHR 数据的不同特征，每一个数值表示所计算的$z\sqrt{n-|Y|-3}$，基于条件独立性测试结果，将生存时间、年龄分组、性别分组、肿瘤位置划分为独立特征集，将肿瘤大小、阳性淋巴结个数、肿瘤分期、T 分期、N 分期划分为因果特征集。

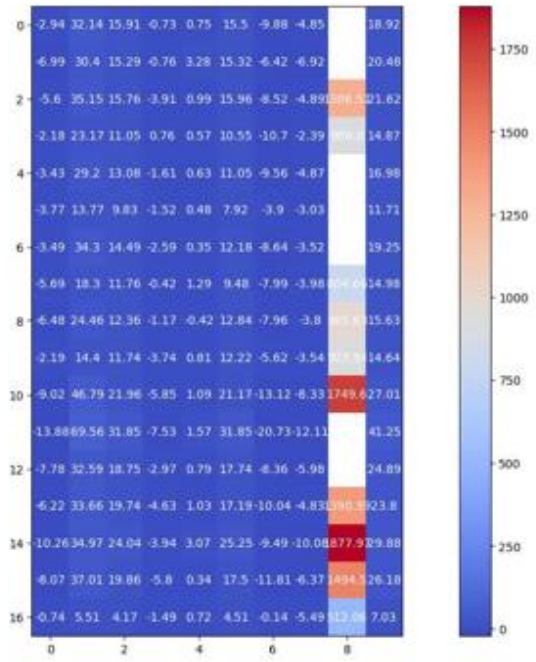


图 3 条件独立性测试结果

(3) 医疗中心识别模块

如图 4 所示，医疗中心识别模块通过将经过提取的因果特征和提取的非因果特征进行特征拼接后送入医疗机构识别网络并生成输入患者数据预测所属医疗中心的概率列表，在训练阶段，该列表作为输入患者数据的样本权重更新患者数据所属的医疗中心的分类器，在预测阶段，作为多个子分类器的模型聚合权重。

为了平衡不同医疗机构下的患者数据在表示空间中的分布，我们引入了一个额外的 loss 来衡量不同医疗机构下数据分布的距离。

$$MMD(P,Q) = \left\| \frac{1}{n^2}\sum_{i=1}^{n}\sum_{i'=1}^{n} k(x_i, x_i') - \right.$$

$$\left. \frac{2}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} k(x_i, y_j) + \frac{1}{m^2}\sum_{j=1}^{m}\sum_{j'}^{m} k(y_j, y_j') \right\| \quad (3)$$

$$loss_{imb} = \alpha \sum_{1\le i<j\le T_N} MMD(\Phi_i, \Phi_j) \quad (4)$$

其中 P 和 Q 表示两个来自不同医疗机构的数据集，k 表示 MMD 的高斯核矩阵，n 和 m 表示数据集的样本大小，x_i, x_i', y_j, y_j'表示两个数据集内的样本，T_N表示医疗机构的个数，Φ_i表示第 i 个医疗机构的数据分布。

(4) 多中心预后预测模块

如图 4 所示，多中心预后预测模块根据输入的因果特征构建多中心的预后预测模型，因果特征的表示在不同医疗机构之间是共享的，每个医疗中心下训练各自的分类网络仅使用各自的患者数据。在训练过程中，通过将医疗中心识别模块生成的样本权重赋予每一个样本来计算对应的 loss。在预测阶段，输入患者所属的医疗机构未知，通过医疗中心识别网络生成一组权重表示输入患者所属医疗机构与训练数据中的所有医疗机构的相似度，根据相似度聚合所有医疗机构下的预后预测网络的结果作为患者最终预测结果。Loss:

$$\frac{1}{N}\sum_{i=1}^{N} w_i * MSE\left(Y_i, h(\Phi(x_i), T_i)\right) + \lambda R(h) \quad (5)$$

其中w_i表示患者数据对应的样本权重，T_i表示患者数据所属医疗机构，$h(\Phi(x_i), T_i)$表示患者数据的预测概率值，$R(h)$表示模型参数的 L2 正则化，λ 是正则项的超参。

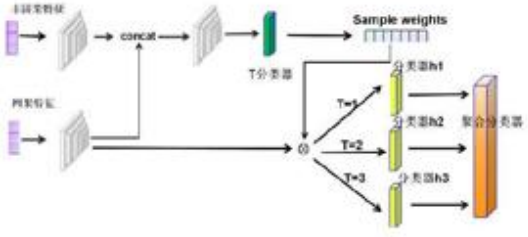


图 4 模型结构

3. 结果

3.1 评价指标

采用模型预测的准确率（ACC）、受试者工作特性曲线下面积（AUC）和估计校准指数（ECI）对模型的区分和校准能力进行评价。

3.2 结果

（1）不同医疗机构的数据集异质性评估结果

使用 wasserstein[10]距离评估不同医疗机构间数据集的异质性，评估结果如图 5 所示。其中来自中国的数据集 SAHZU 与美国其他数据集之间表现出较大的异质性，来自美国的数据集 SEER 中不同机构的数据集之间表现出不同程度的异质性。

图 5 不同数据集异质性评估结果

（2）模型评估结果

根据异质性评估结果，选择其中三个数据中心分别表示低异质性、中等异质性、高异质性数据集，评价结果如表 2 所示。

表 2 模型评价结果

内部验证集	Accuracy	Auc area	Eci
MLP	0.726	0.697	17.52
Our model	0.683	0.740	13.92
低异质性	Accuracy	Auc area	Eci
MLP	0.686	0.759	18.32
Our model	0.672	0.733	13.84
中异质性	Accuracy	Auc area	Eci
MLP	0.706	0.768	17.11
Our model	0.705	0.765	15.24
高异质性	Accuracy	Auc area	Eci
MLP	0.652	0.741	26.92
Our model	0.696	0.770	15.08

对于准确率，在内部验证集中，我们的模型准确率为 0.683 低于 MLP 模型的 0.726，随着数据异质性的增加，我们的模型准确率保持相对稳定，而 MLP 模型的准确率开始下降，在高异质性组中，我们的模型准确率为 0.696 而 MLP 的准确率为 0.652。对于 AUC 面积，在不同异质性组间，我们的模型 AUC 波动范围为 [0.733, 0.770]，MLP 的波动范围为 [0.697, 0.768]。对于 ECI 校准指数，我们的模型校准指数都显著低于 MLP，说明我们的模型具有更好的校准能力。

4. 讨论

模型的预后预测基于因果特征，非因果特征不直接参与模型的预后预测，而是通过医疗中心识别模块生成样本权重的方式调整每一个患者数据样本对于模型的贡献度，从而避免独立特征和感兴趣变量的直接关联，防止由于独立特征在不同医疗中心下分布差异变化较大而导致的模型性能大幅降低。

在预测环节，输入患者数据输入到所有子中心的预测网络，医疗中心识别模块根据患者数据与训练集中所有医疗中心的相似度生成一组权重，并对所有预测结果加权得到最终预测结果。在不同异质性的数据集中，模型的评价结果为准确率[0.672,0.705]，AUC 面积[0.733,0.770]，ECI 校准指数[13.84,15.24]，性能稳定。一方面，非因果特征不直接用于模型的预后预测避免了非因果特征与感兴趣变量的直接关联，所以模型学习到的关联更加可靠，另一方面，通过比较输入患者数据与医疗机构数据分布的相似性得到的权重保证与输入患者数据相似的子模型能得到更大的权重，使得模型保持良好性能。然而，由于模型倾向于学习更加通用的因果特征，所以当训练集和验证集的数据分布异质性较低时，基于因果特征预测的模型效果可能会比使用所有特征的模型略差。

5. 结论

本研究开发了一种基于因果表征学习的多中心结直肠癌预后预测模型，包括四个模块：种子变量筛选模块、条件独立性测试模块、医疗中心识别

模块、多中心预后预测模块。

模型在来自美国国家癌症研究所的 SEER 数据集和来自中国浙江大学第二附属医院 SAHZU 数据集上进行了验证，在不同异质性的医疗机构上测试结果分别为准确率[0.672, 0.705]，AUC 面积[0.733, 0.770]，ECI 校准指数[13.84, 15.24]结果稳定，可以有效辅助不同医疗机构的临床医生进行结直肠癌的预后预测。

参考文献

1 Brenner H, Hoffmeister M, Jansen L. Comparisons of colorectal cancer mortality between screening participants and the general population are strongly biased unless an incidence-based mortality approach is used[J]. Journal of clinical epidemiology, 2014, 67(2): 184-189.

2 Sinicrope F A. Increasing incidence of early-onset colorectal cancer[J]. New England Journal of Medicine, 2022, 386(16): 1547-1558.

3 Li J, Guo B C, Sun L R, et al. TNM staging of colorectal cancer should be reconsidered by T stage weighting[J]. World Journal of Gastroenterology: WJG, 2014, 20(17): 5104.

4 Tian Y, Chen W, Zhou T, et al. Establishment and evaluation of a multicenter collaborative prediction model construction framework supporting model generalization and continuous improvement: A pilot study[J]. International Journal of Medical Informatics, 2020, 141: 104173.

5 Shalit U, Johansson F D, Sontag D. Estimating individual treatment effect: generalization bounds and algorithms[C]//International conference on machine learning. PMLR, 2017: 3076-3085.

6 Johansson F D, Kallus N, Shalit U, et al. Learning weighted representations for generalization across designs[J]. arXiv preprint arXiv:1802.08598, 2018.

7 Johansson F D, Kallus N, Shalit U, et al. Learning weighted representations for generalization across designs[J]. arXiv preprint arXiv:1802.08598, 2018.

8 Johansson F D, Shalit U, Kallus N, et al. Generalization bounds and representation learning for estimation of potential outcomes and causal effects[J]. The Journal of Machine Learning Research, 2022, 23(1): 7489-7538.

9 Larson P D, Halldorsson A. What is SCM? And, where is it?[J]. Journal of Supply Chain Management, 2002, 38(3): 36-44.

10 Shen J, Qu Y, Zhang W, et al. Wasserstein distance guided representation learning for domain adaptation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).

作者信息

姓名：秦园炳
性 别：男
学历：硕士研究生
出生年月：1999 年 08 月 29 日
专业：生物医学工程
工作单位：浙江大学生物医学工程与仪器科学学院，电子病历与智能专家系统教育部工程研究中心，生物医学工程教育部重点实验室
邮编：310012
电子邮件地址：qyb@zju.edu.cn

OR-027

湖南省三级医院智慧管理的现状与对策

莫芳惠[1,3,4,5]，欧阳威[2,3,4,5]，王安莉[2,3,4]，戴萍萍[2,3,5]，罗爱静[1,3,4,5]*
1 中南大学湘雅二医院
2 中南大学湘雅三医院
3 中南大学生命科学学院
4 医学信息研究湖南省普通高等学校重点实验室（中南大学）
5 湖南省心血管智能医疗临床医学研究中心

1 引言

医院智慧管理是“三位一体”智慧医院建设的重要组成部分，在提高医院管理水平、优化医疗资源配置、提高医疗工作效率等多方面起重要作用。2021 年国家卫生健康委正式发布《医院智慧管理分级评估标准体系（试行）》，旨在指导各地、各医院加强智慧医院建设的顶层设计，充分利用智慧管理工具，提升医院管理精细化、智能化水平[1]。本研究依托《湖南省三级医疗机构智慧管理建设现状调查问卷》，了解和分析医院智慧管理发展现状，为我省三级医院提高智慧管理水平和加快医院信息化发展提供数据支撑和改进思路。

2 资料与方法

本次调查采用发放调查问卷的方式，采用线上发布自愿填写的方式，由各医院信息技术部负责人或具备相应水平的医院信息技术人员进行填报。通过对调查问卷的回收、统计，从中发现问题、分析问题并提出解决方法。调查问卷内容主要包括四大部分：单位基本情况、信息化组织投入与发展情况、智慧管理建设情况、基础设施及安全体系建设情况。

3 结果

3.1 基本情况

调研面向湖南省的三级医疗机构，共收回有效调查问卷 103 份，其中综合医院 64 家，专科医院 22 家，中医院 13 家，中医专科医院 2 家，中西医结合医院和妇幼保健院都仅有 1 家。医院等级以被调查医院自行填报为准，自高向低分为三级（ 甲等、乙等、未定等），其中三级甲等医院 47 家，三级乙等医院 4 家，三级未定等医院 52 家。在 103 家三级医院中，79.61%的医院核定床位数超过 500 张，88.35%的医院年门诊人次超过 10 万，86.41%的医院年出院人次超过 1 万。

3.2 医院信息化情况

在参与调研的三级医院中，共 100 家医院参加了 2021 年度电子病历系统功能应用水平分级评价，参评率为 97.09%，平均分级为 3.48，最高级别为 5 级。61 家医院参加了 2020 年度智慧服务分级评估，参评率 59.22%，平均分级为 1.10，最高级别为 2 级。

3.3 医院收入与信息化投入情况

3.3.1 医院平均每年收入 22 家医院的收入大于 100000 万，占参与调查医院的 21.36%；50001-100000 万的有 19 家，占 18.45%；从总体来看医院近五年的平均每年收入主要集中在 50000 万及以下，占所有参与医院的 60.19%。

3.3.2 医院近五年信息化平均投入占比 仅 9 家医院的投入占比大于 5.00%，占参与调查医院的 8.74%；比重在 1.00%～5.00%的医院有 31 家，占 30.09%。从总体来看信息化投入比重集中在 1.00%及以下，占所有参与医院的 61.17%。

3.3.3 医院信息化累计投入情况 23 家医院累计投入达 5000 万以上，占参与调查医院的 22.33%；2001-5000 万的有 35 家，占 33.98%；参与医院的信息化累计投入从总体上看主要集中在 2000 万及以下，占所有参与医院的 43.69%。

3.4 医院信息化人才队伍建设情况

3.4.1 信息化部门自有人数 仅 3 家医院有大于 20 人以上的信息化部门，占参与调查医院的 2.91%；25 家医院是 11～20 人，占比 24.27%；67 家医院是 4～10 人，占比为 65.05%；8 家医院是 3 人以下，占比 7.77%。从总体来看集中在有 4～10 人的信息化部门为主。其中，信息化部门最高

人数是 31 人，最低人数是 1 人，平均人数是 8 人。

3.4.2 信息化部门自有人数占医院职工数的比重 16 家医院的比重大于 1.00%，占参与调查医院的 15.53%；14 家医院的比重小于 0.50%及以下，占 13.59%。从总体来看信息化部门自有人数占医院职工数的比重集中在 0.50%～1.00%，占所有参与医院的 70.88%。

3.4.3 厂商派驻人数 仅 6 家医院有大于 10 人以上的信息化部门厂商派驻人员，占参与调查医院的 5.83%；15 家医院是 4～10 人，占 14.56%；39 家医院是 1～3 人，占 37.86%；43 家医院没有派驻人员，占 41.75%。从总体来看集中在没有派驻人员为主。其中，最高派驻人数是 75 人，最低派驻人数是没有派驻人员，平均人数是 3 人。

表 1 医院各方面综合情况

基本情况		三级医院（n=103）	
		数量	百分比（%）
医院类别	综合	64	62.14
	专科	22	21.36
	中医	13	12.62
	中医专科	2	1.94
	中西医结合	1	0.97
	妇幼保健院	1	0.97
医院等级	甲等	47	45.63
	乙等	4	3.88
	未定等	52	50.49
核定床位数(张)	≤500	21	20.39
	>500	82	79.61
年门诊人次(万)	≤10	12	11.65
	>10	91	88.35
年出院人次(万)	≤1	14	13.59
	>1	89	86.41
电子病历应用水平分级评价	参评	100	97.09
	未参评	3	2.91
智慧服务分级评估	参评	61	59.22
	未参评	42	40.78
平均每年收入(万)	≤5	62	60.19
	5~10	19	18.45
	>10	22	21.36
近五年信息化平均投入占比(%)	≤1.0	63	61.17
	1.0~5.0	31	30.09
	>5.0	9	8.74
信息化累计投入资金(万)	≤2000	45	43.69
	2000~5000	35	33.98
	>5000	23	22.33
信息化部门自有人数(人)	<3	8	7.77
	4~10	67	65.05
	11~20	25	24.27
	>20	3	2.91
信息化部门自有人数占医院职工	≤0.5	14	13.59

数的比重(%)	0.5~1.0	73	70.88
	>1.0	16	15.53
厂商派驻人数(人)	0	43	41.75
	1~3	39	37.86
	4~10	15	14.56
	>10	6	5.83

3.5 医院智慧管理各模块的建设情况

智慧管理是现代医院管理中的一种重要方式，通过利用信息技术和智能化设备，对医院的医疗、护理、管理、运营、保障等方面进行全面优化和提升，以提高医院管理效率和医疗服务质量。

在智慧管理各角色功能相关业务系统中，医疗护理管理、财务资产管理、药品耗材管理的业务系统平均建设率分别为 55.66%、70.15%、52.67%，全距分别为 48.54%、52.44%、70.87%，从整体来看三者处于中上等水平，但同一模块不同业务系统之间存在建设良莠不齐的问题；智慧管理系统建设处于中下等水平的是人力资源管理和办公管理模块，平均建设率分别为 42.07%、43.21%，全距分别为 4.85%、39.81%，前者的系统建设较为均衡，后者差距较大；设备设施管理、运营管理、运行保障管理、教学科研管理的系统建设处于较低水平，平均建设率皆低于 26%，四大模块的全距在 5%～49%零散分布。

综上所述，医院的智慧管理各模块建设存在不平衡、不充分的情况，设备设施和教学科研管理等建设相对落后，相关管理能力有限，医院的智慧管理建设需要全面推进，见表 2 和图 1。

表 2 医院智慧管理各模块的建设情况

项目	数量	百分比(%)	平均建设率(%)	全距(%)
医疗护理管理				
医疗质控管理	52	50.48	55.66	48.54
护理质控管理	32	31.07		
医疗准入管理	75	72.82		
医院感染管理	82	79.61		
不良事件管理	66	64.07		
和谐医患关系	37	35.92		
人力资源管理				
人力资源规划	41	39.81	42.07	4.85
人事管理	43	41.75		
人员考核与薪酬管理	46	44.66		
财务资产管理				
医疗收入管理	92	89.33	70.15	52.44
财务会计管理	83	80.58		
预算管理	38	36.89		
资产账务管理	76	73.78		
设备设施管理				
购置管理	31	30.10	20.63	19.42
使用运维管理	29	28.16		
质量管理	11	10.68		
效益分析	14	13.59		
药品耗材管理				
药品耗材遴选与购置	48	46.60	52.67	70.87

库存管理	94	91.26		
消毒与循环物品管理	54	52.43		
监测与使用评价	21	20.39		
运营管理				
成本控制	18	17.48		
绩效核算管理	46	44.66	24.27	33.98
医疗服务分析评价	11	10.68		
运行保障管理				
后勤服务管理	12	11.65		
安全保卫管理	61	59.22		
医疗废弃物管理	17	16.50	25.24	48.54
楼宇管控	11	10.68		
信息系统保障管理	29	28.15		
教学科研管理				
教学管理	19	18.45	15.54	5.83
科研管理	13	12.62		
办公管理				
协同办公管理	65	63.11	43.21	39.81
档案管理	24	23.30		

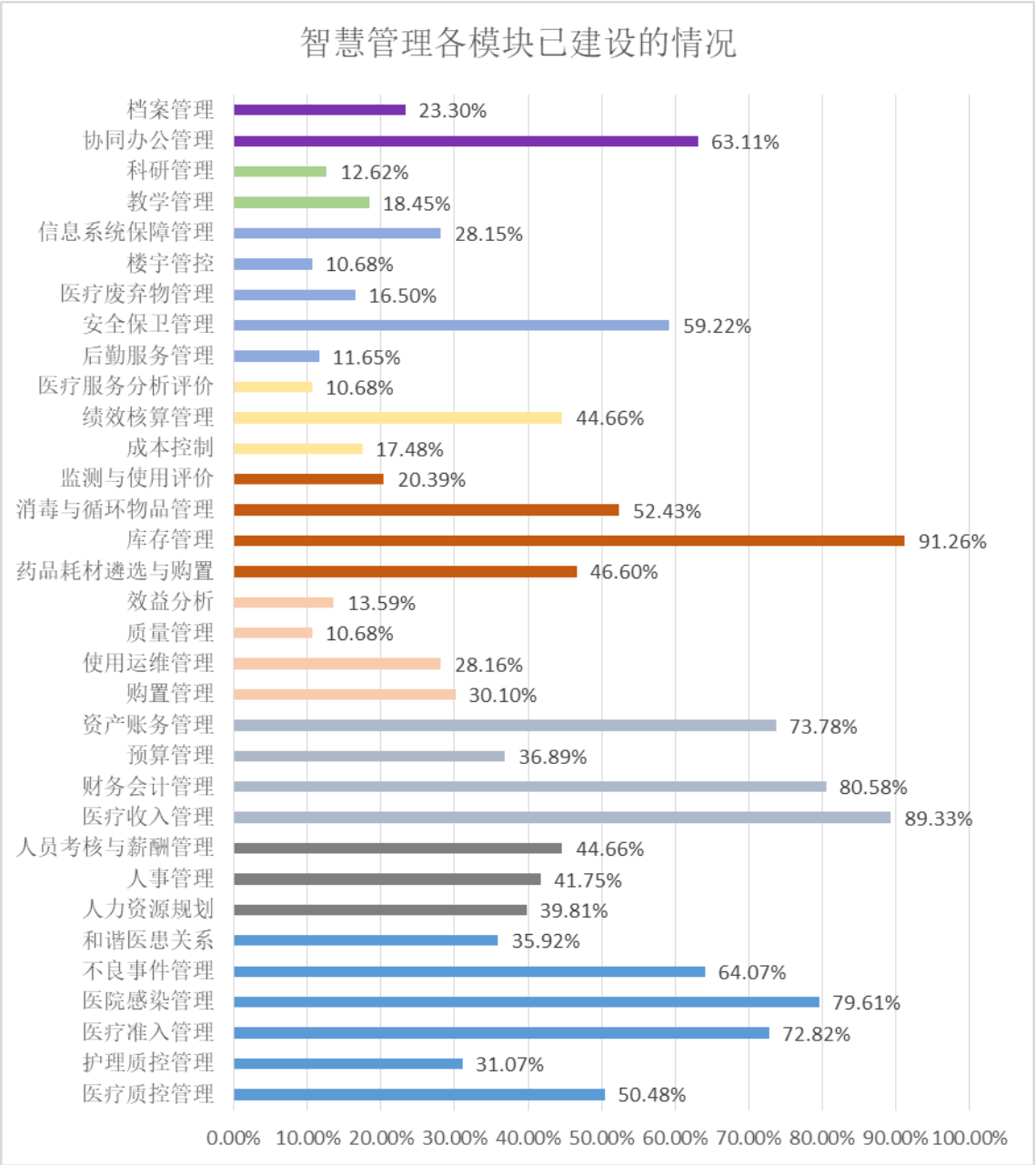


图 1 医院智慧管理各模块的建设情况

3.6 智慧管理各角色功能相关业务系统等级保护定级备案情况

信息安全等级保护制度是我国信息安全保障的基本制度、基本方法和基本策略[2]。二级等保测评和三级等保测评是信息系统等级保护定级备案的不同阶段。通过等保测评，可以帮助医院确定信息系统的安全等级，增强信息系统的抗攻击能力和安全性能，保障医院的正常运营和业务发展。同时，等保测评也是医院信息化建设过程中必不可少的环节，指导医院完善信息系统的安全建设和防护工作，促进医院信息化建设的规范化和标准化。

从调查结果来看，医院智慧管理各角色功能相关业务系统的等级保护定级备案情况不理想，除了医疗质量管理和药品耗材管理系统的等保测评通过率尚佳，通过二级和三级等保测评的医院占比达 40%以上，其他的智慧管理业务系统特别是人力资源管理、运行保障管理和教学科研管理的二级和三级等保测评通过率不到 3%，存在较大的安全漏洞和风险，系统的安全性缺乏保障，见表 3 和图 2。

表 3 医院智慧管理各模块系统等级保护定级备案情况

项目	二级等保测评		三级等保测评	
	数量	百分比（%）	数量	百分比（%）
医疗质量管理	9	8.74	35	33.98
人力资源管理	0	0.00	2	1.94
财务资产管理	9	8.74	6	5.82
设备设施管理	4	3.88	7	6.80
药品耗材管理	11	10.68	36	34.95
医院运营管理	6	5.83	12	11.65
运行保障管理	1	0.97	2	1.94
教学科研管理	0	0.00	1	0.97
办公管理	4	3.88	10	9.71

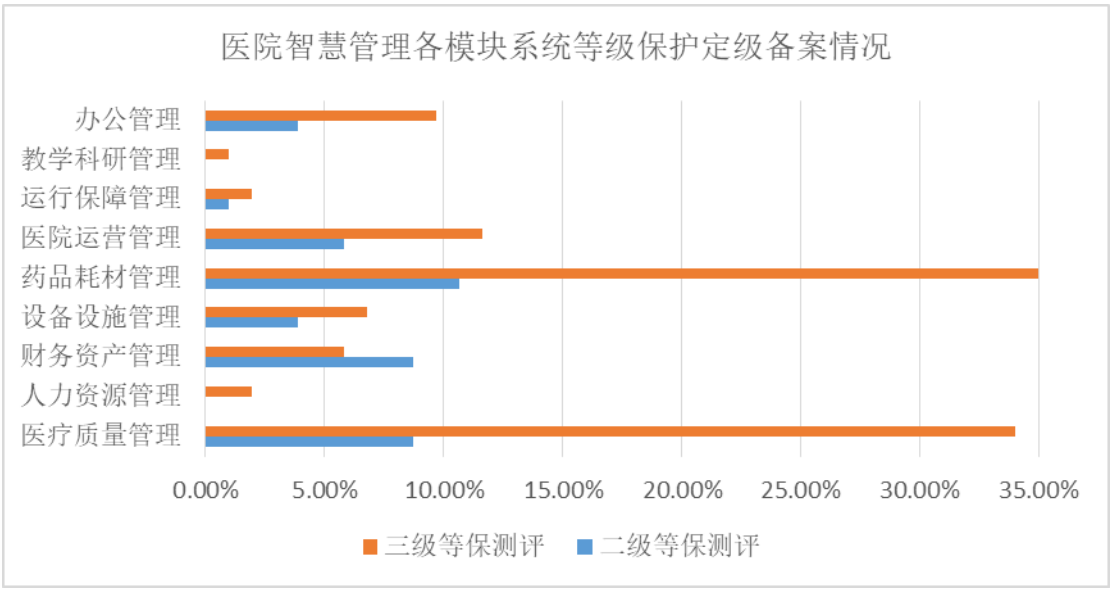


图 2 医院智慧管理各模块系统等级保护定级备案情况

3.7 问题分析

根据上述对调查问卷情况的分析结果，总结湖南省三级医疗机构的信息化建设存在以下几个突出的问题：

（1）医院信息化建设的投入严重不足

医院信息化建设的投入是影响医院智慧管理实现的重要因素之一。目前，国内三级医院在信息化建设方面的投入严重不足。据统计，国内大部分三级医院平均上年度信息化投入预算占年度总预算中的比例仅有 1.01%[3]，本研究调查结果显示我省大部分三级医院的信息化建设投入占医院总收入的比例不足 1%，与发达国家相比，这一比例明显偏低。此外，医院信息化建设的投入不均衡。很多三级医院在信息化建设方面的投入主要集中在医院信息系统和医疗管理系统 等核心系统，而对于设备设施管理系统、教学科研管理系统等智慧管理所需的系统建设投入较少。医院信息化建设投入的不足，直接导致了医院智慧管理的实施效果受到影响，这意味着医院在信息化建设方面的投入还有很大提升空间。

（2）医院信息化人才储备不足

医院信息化人才是推动医院智慧管理实现的关键因素之一。然而，目前的现状却是，国内大部分三级医院缺乏信息化人才，尤其是缺乏具备综合能力和实践经验的信息化人才。据《2019—2020 年中国医院信息化状况调查》[4]显示我国医院的信息部门职工平均为 10 人，低于国外水平

[5]。本研究结果显示我省三级医院信息部门职工平均为 8 人，无法达到国内平均水平，这意味着三级医院在信息化人才方面的缺口还很大。究其原因，主要还是由于医院信息化建设投入的不足，导致医院缺乏对信息化人才的培养和引进。另一方面，由于医院传统的管理理念的束缚，导致医院对信息化人才的重视程度不够，缺乏对信息化人才的培训和培养。以上因素共同导致了医院信息化人才储备的不足，对医院的智慧管理能力也造成了很大的限制。

（3）医院智慧管理的业务系统建设刚刚起步

医院智慧管理的业务系统建设是推动医院智慧管理实现的重要组成部分。然而，由于医院智慧管理的业务系统建设缺乏前瞻性，目前我省绝大部分三级医院的业务系统建设刚刚起步。很多系统建设是基于当时的技术环境和业务需求进行的，而随着技术环境的不断发展和变化，需要建设更多满足医院新业务需求的系统。另一方面，从本研究调查的医院智慧管理各角色功能相关业务系统的等级保护定级备案情况可知，系统安全性不够高也是当下面临的重要问题之一。随着医院信息化建设的不断推进，医院所收集和存储的数据越来越庞大和重要，因此保障数据的安全性和隐私性尤为重要。但是，医院智慧管理业务系统建设尚处于起步阶段，相关安全措施和管理体系尚未得到有效建立和完善，因此存在着数据泄露、攻击等风险。

4 建议与对策

随着信息技术的不断发展，智慧管理已经成为现代医院管理的不可逆的发展趋势[6]。三级医院作为医疗机构的高层次水平，其智慧管理的实现对于整个医疗行业的发展具有重要意义。各医院如果要实现智慧管理快速稳定的发展，必须做好以下几个方面的工作：

4.1 加大医院信息化建设的投入力度，加强信息技术人才的培养工作

医院信息化建设是实现智慧管理的重要基础，信息技术人才是推进医院智慧管理的关键。为了推进医院智慧管理的建设，医院应该加大信息化建设的投入力度。首先，医院必须重视和加强对信息技术人才的培养和选拔，将信息化建设作为人才引进和培养的重点，为医院信息化建设提供强有力的人才支持。其次，医院应该加大对信息化建设的投入，主要用于设备更新、网络建设、软件维护等方面，确保医院各项业务开展顺利进行。此外，医院应该建立完善的信息化建设管理制度，加强对信息化建设的管理和监督，确保信息化建设的高质量和高效率发展。

4.2 高度重视医院智慧管理建设，加强医院信息化平台的建设和运营

医院智慧管理是一项复杂的建设工程，需要医院领导层的高度重视和全面规划。医院相关管理人员应正确认识智慧管理对于医院发展的重要性，并将其纳入医院发展规划中，作为医院长期发展的一个方向，再切实依据医院自身发展状况制定智慧管理的具体实施方案，建立完善的智慧管理信息化平台架构及运营机制，确保平台的安全性、可用性和可扩展性。而且，医院还应积极参加智慧医院各项分级评估和等保测评工作，严格对照评价结果和国家要求，以评促建、以评促改、以评促用、评建结合[7]，逐步建成医疗、服务、管理一体化的智慧医院系统。

4.3 加强医院智慧管理的宣传和推广，推进医院智慧管理的实际应用

医院智慧管理的宣传推广和实际应用是实现智慧管理的关键及重要保障。医院应该加强对科室信息化的建设，将智慧管理相关技术应用于临床、科研、管理等多个领域，并开展智慧管理的评估和监测，及时发现和解决问题，确保智慧管理的质量和效果。此外，积极开展智慧管理相关的培训和研讨会，提高医务人员对智慧管理的认识和应用能力，不断加强与各界的合作和交流，通过与科技公司、医疗机构等合作，共同推进智慧管理的发展和应用。

5 结语

大数据时代，未来医院智慧管理建设正在进行时，实现医院智慧管理需要从多个方面入手，各方努力，加强协作，才能加快实现医院信息化建设和智慧管理的目标和愿景，推进智慧医院建设步伐。

OR-028

智慧医院多院区发展的挑战与对策

左秀然 瞿朗 杨国良*
武汉市中心医院

1. 引言

公立医院的多院区发展已成为公立医院发展的新趋势。在《医疗机构设置规划指导原则（2021-2025 年）》中首次规范公立医院分院区设置，分院区是指公立医院在原有院区以外的其他地址，以新设或者并购等方式设立的，具有一定床位规模的院区[1]。2022 年以来，国家卫健委对公立医院多院区发展进行了总体部署，引导高水平公立医院在控制单体规模基础上，适度发展多院区，促进优质医疗资源扩容和区域均衡布局，同时提升重大疫情救治能力。但同时提出了规划布局、设立条件、功能定位、平急结合、一体化管理机制五个方面的公立医院有序发展多院区的要求[2]。

在公立医院多院区发展的形势下，随之带来了智慧医院建设模式的转变需求，由单体医院纵深精细化发展，向纵深精细化+横向协同化发展的需求日益突出，给智慧医院的建设带来新的挑战。

2. 多院区架构下智慧医院挑战

2.1 对信息化顶层设计规划方式的挑战

多院区架构下的智慧医院建设并非简单的信息系统扩展院区的实施，而是以提高医疗质量、改善病人服务质量、促进医院管理水平为目标的系统设计及资源整合[3]。单体医院纵向精细化深化和横向多院区扩展，带来的业务量激增，以及多院区间业务协同需求的增强，使得传统以单院区纵深精细化为主的信息化顶层设计已无法满足集团化发展需要，对传统信息化架构提出极大挑战。

2.2 对多院区资源调度方式的挑战

医院一体化管理的重点在于各种资源的整合，而院区的分散增加了资源整合的难度。文献 4 中总结了资源分配不科学、转诊不通畅、管理不规范等问题对智慧医院的资源调度提出的挑战。多院区的运行往往带来总体运营成本的增加，运行效益边际递减，需要在院区间统筹人、设备、床位等资源，优化整体资源配置，这成为一个新的挑战。

2.3 对多院区数据资源及标准化管理的挑战

由于各院区的医疗设备、信息系统、医疗数据等存在差异，如何统一数据标准和规范，实现信息互通、数据共享，确保数据的一致性，是多院区数据管理的重要问题。业务数据未统一治理、数据字典未标准化等会阻碍多院区信息系统的互联互通，从而对业务的连续性带来挑战。

2.4 对数据中心资源配置及管理的挑战

随着业务规模的快速增长，数据中心需要不断提升算力以支撑高效率的业务运转，但传统的单中心灾备机制无法满足分散部署的数据中心间的备份和灾备恢复需求。此外，院区间的组网需要考虑安全性和经济性问题，网络出口的规划也成为新的挑战，需要优化网络路由和降低网络延迟，确保网络资源最高效的调度。同时，随着互联网需求的不断增长也带来网络信息安全的进一步挑战[5]。

2.5 对新老院区基础设施水平差异的挑战

医院规模扩大后，院区间如不能同质化服务体验，将会导致患者满意度下降，应对的关键措施是将医疗服务和医疗技术同质化。但老院区基础设施的配置短板，所带来的整体体验无法快速提升，将明显地带来患者和医务人员满意度下降，同时也带来 IT 设施管理难度的增加。

3. 多院区智慧医院建设路径

3.1 需求导向适度超前的顶层设计

建设分院区需要坚持需求导向，符合区域医疗卫生规划，契合区域内居民健康需求。采用自顶向下的视角，使智慧医院达到惠民、惠医、惠管和惠政的目的。智慧医院的多院区发展需要合理规划医院规模、制定明确的战略目标和区域定位，同时从需求、架构和技术应用等方面进行前瞻性规划和设计。需求前瞻需要对未来医疗服务的需求趋势进行深入分析和研究；架构前瞻则需要考虑如何合理构建医院信息化架构；技术前瞻则是探索新兴技术在医疗领域的应用。

3.2 跨院区资源的一体化管理统一调度

优化跨院区的资源一体化调度方式，坚持以人为本，从患者、医务人员、管理决策者三类角色的需求出发。患者角度，满足同质化就医体验及随需随时跨院区就诊的需要，在统一专家专科门诊排班、统一大型医疗设备检查预约、统一床位预约等方面，实现资源统一管理及调度。医务人员角度，满足跨院区远程互动，开展院区间远程诊断，多学科远程 MDT 联动及同质化业务信息系统操作体验的需求。管理人员角度，满足实时动态掌握多院区运营数据，运营风险综合智能研判、主动预警、一键应急指挥的需要，对多院区的医疗、教学、科研等各类资源进行整合并统一管理，如多院区号源资源池统一管理。智能规则、集中研判、风险预警和闭环处置等策略相互配合，能够有效提高智慧医院多院区的资源统一调度能力。

3.3 多院区共性支撑与个性发展的支撑能力

多院区信息化架构中，采用统一数据治理与数据资产管理，依托大数据中心构建支撑多院区诊疗及管理能力提升的共性支撑，如以患者主索引为基础关联历次诊疗数据，实现跨院区信息共享；提供临床辅助决策支持，应用于多院区快速提升医生科学高效临床管理的需要；利用数据驱动多学科联动 MDT 模式进行干预和效果评价等。信息系统的发布机制还需满足不同院区、不同用户个性化发布部署需要，采用蓝绿分布、滚动分布和灰度分布等升级发布机制[6]，帮助系统逐步更

新和扩展，保证业务的平稳运行。

3.4 线上线下一体化的应用服务体系

充分利用互联网医疗优化医疗资源的集约化配置，打通多院区线上医疗服务资源配置，统一提供线上问诊、线上随访、处方延方、物流配送等服务。基于医疗健康大数据中心，建立覆盖院前、院中和院后全流程的移动互联网医院应用，保障线上线下的闭环衔接。同时支撑分时段预约诊疗、医保在线支付、满意度评价等闭环服务应用。

3.5 云网一体稳定可靠的基础底座

云化部署弹性计算数据中心，满足多院区业务量扩展对资源弹性配置的要求。通常包含本地硬件资源云化、第三方运营商云机房构建私有云、私有+公有混合云三种做法。以一院三区为例，采用两地三中心架构，通过应用层及数据库双中心分布负载部署的方式，支持分布式数据库、容器化、微服务业务改造。核心业务采用“双活+连续数据保护”方式，一般业务采用“容灾+连续数据保护”方式，同时建立异地数据备份机制。组网架构采用双环可靠性组网，具体包括主备中心双环和院区双环，通过 OTN、裸光纤等技术建立可靠高速专线链路，如图 1 所示。A 中心作为医疗云平台业务的承载主体，B 中心作为灾备机房中心，通过两条冗余链路连接，三个院区各部署两条云专线到 A、B 数据中心，另部署 1 条云专线，用于实现到 C 中心的数据灾备。

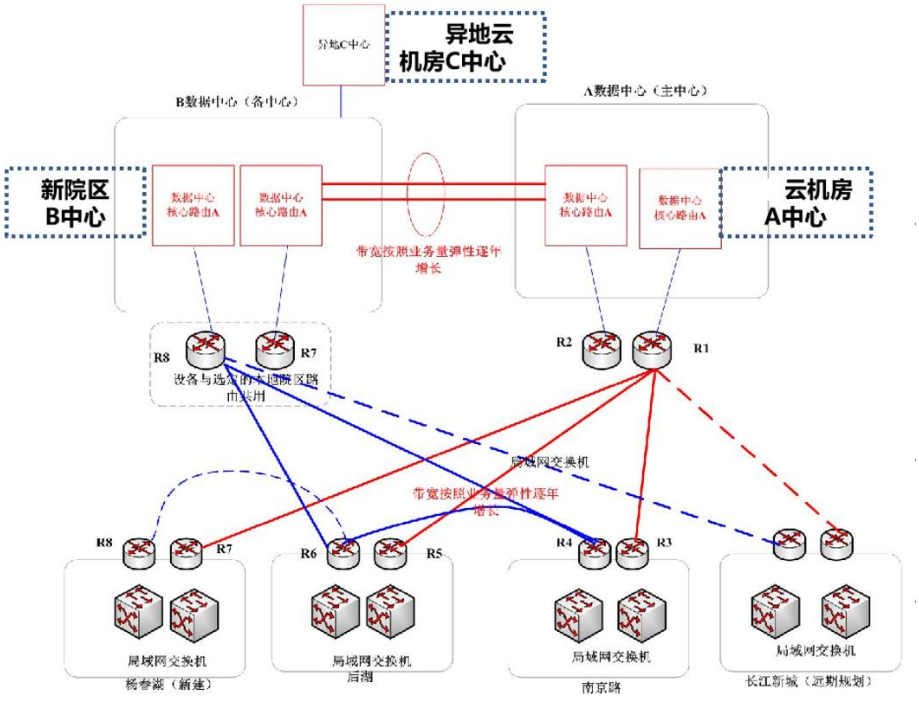


图 1 双环可靠性组网

为应对互联网业务对网络安全带来的挑战，网络架构采用互联网应用专区管理、内网数据中心加固安全防护、多院区统一互联网出口等策略，以满足网络安全要求。

3.6 主动预防一站式统一运维管理

构建完整的预防式维护体系是现代医院信息化建设中非常重要的一环，可以有效地提升设施设备、楼宇运行、业务系统的可靠性和安全性监管，保证医疗信息系统的正常运行，提高医疗服务的质量和效率。同时为了集约化配置运维团队，实现统一运维调度，多院区建立相对集中的一站式运维服务中心，并依托数据驱动的医院运行风险预警，集中式调度指挥应急事件，以提升应急事件的快速响应。

4. 展望

集团化架构下的智慧医院建设是新时期新课题，同时是智慧医院高质量发展的新契机。新院区建设需抓好新基建契机，打造“三位一体”的智慧医院，从基础建设、智能化信息化、综合管理全方位提升管理水平，再快速推广协同其他院区。在智慧管理新技术上，采取数字孪生和综合运营指挥策略，实现多院区智能运营和管理，通过一键转换、一网监测、一体指挥等手段，实现各院区的智能化运营和管理，是后续进一步深化的方向。

OR-029

数智化综合运营管理模型构建在医院智慧管理体系建设中的实践与思考

王韫秀[1] 赵玉虹[2*]
中国医科大学附属盛京医院 运营与绩效管理工作部
2. 中国医科大学附属盛京医院 党委办公室

1. 引言

随着 2009 年中国医药卫生体制深化改革以来，药耗改革、医保付费改革等系列政策举措实施，公立医院面临在 DRG/DIP 支付方式改革、医疗价格深化改革、药品耗材带量采购、集中谈判各类举措下的高质量发展道路探索。医院精益运营管理的重要性逐渐凸显[1-5]。在 2020 年，国家先后发布《关于开展“公立医疗机构经济管理年”活动的通知》（国卫财务发〔2020〕262 号）和《关于加强公立医院运营管理的指导意见》（国卫财务发〔2020〕27 号）政策文件，强调公立医院经济管理核心重要性，并明确指出了公立医院运营管理的范畴和意义。与此同时，《国务院办公厅关于推动公立医院高质量发展的意见》（国办发〔2021〕18 号）和《国务院办公厅关于加强三级公立医院绩效考核工作的意见》（国办发〔2019〕4 号），更是通过量化的指标体系引领公立医院高质量发展的明确方向和努力目标。然而，信息化建设是高质量发展和精益管理的基础，国家随后发布的《智慧医院建设指南》（DB34/T 4011-2021）和《公立医院运营管理信息化功能指引》（国卫办财务函〔2022〕126 号），为医院落实精益运营管理提供了重要的建设依据。各级公立医院面临信息化建设水平参差不齐、系统间数据缺乏互联互通和高效应用，未能充分发挥管理效能的问题比较集中。因此，在医院高质量发展重要阶段，亟待统筹考虑既有建设基础，梳

理问题和短板，加强顶层设计和整盘规划建设，以达成以大数据为基础的医院智慧管理体系，尤其是数智化综合运营管理体系构建的新局面。

2. 综合运营管理在医院智慧管理体系建设中作用

2.1 综合运营管理的组织结构

公立医院进入智慧管理重要阶段，综合运营管理在医院智慧管理中发挥重要角色，很多医院先后成立独立的运营管理部门，但在管理实践过程中暂时未达成清晰的管理权责界限，以及在和原有职能管理部门间的关系和角色上仍有困惑，主要体现在管理界限感，权责的重叠和交错部分的处理，职能管理范畴的困惑，一定程度上制约了医院智慧管理信息化建设步伐，扭扯羁绊各管理部门的工作进步和高质量发展步伐。

基于上述管理现状，一方面需要医院整体运营管理决策层的权责范围顶层设计，一方面需要建立部门间协作共赢的管理共识，形成精密仪器一样的精益运营管理组织结构和管理体系，盛京医院在探索综合运营管理在医院智慧管理实践中，摸索组织结构以及独立的运营管理部门发挥的穿针引线的作用，结构如图 1。

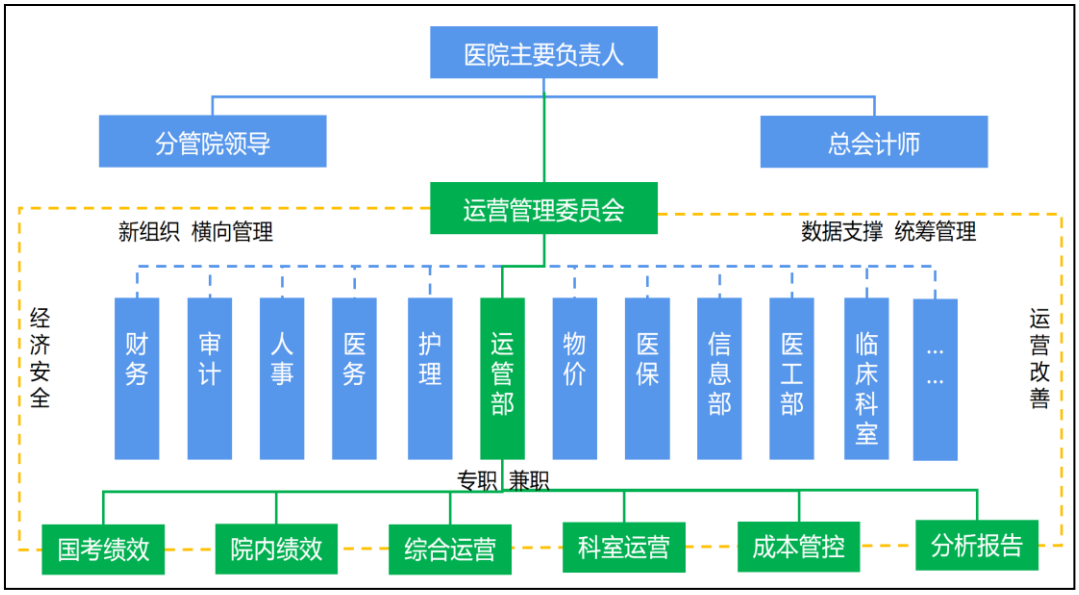


图 1 公立医院综合运营管理组织结构图

2.2 数智化综合运管信息需求

公立医院高质量发展阶段，智慧医院建设已经成为医院“新基建”，其中智慧管理部分的信息化建设一直是医院信息化发展过程中被忽视的盲区。现阶段，越来越多的管理目标集中指向了“智慧管理”在智慧医院建设过程中的核心作用[6-8]，管理者们对大数据资产建设的管理认识，在综合运营管理的“大运营”概念中得到充足的诠释，即通过整合各业务系统间的数据资产，实现各部分职能管理大数据的互联互通，构建综合运营管理的立体化全局观，填充传统管理的“缝隙和断层”，整合联动数据资产，建立“总-分-总”综合管理模式，及通过综合运营管理大数据监测和分析发现精益运营

管理的集中问题、薄弱点和管控重点，进而将集中的管理问题分配到具体管理责任部门加强专项管理，再通过大运营全局观监测管理成效，并作出进一步管理建议，形成良好的管理闭环。

3. 数智化综合运营管理模型构建

综合运营管理模型的构建一方面依靠管理思维的顶层设计，即不仅能充分了解各部分职能管理权责界限，更需要充分统筹职能管理间的相互配合和协同机制。另一方面，依靠强大的大数据信息化建设基础，为管理辅助赋能、增添羽翼。

4.1 建立运营管理数据底座，镜像汇集数据源

建立业务数据镜像汇集基础，包括但不限于如下基础数据源的整合：（1）医疗业务端数据源，如电子病历 EMR、医疗 HIS、影像 PACS、实验室 LIS 的数据汇集；（2）医保业务端数据源，如门诊和住院患者的各级各类医疗保障类型、医保支付类型、诊疗费用、统筹报销项目和自费项目费用明细、医保经办机构和医疗机构间的结算数据等；（3）人力资源数据源，如人员的类型、学科及亚专业分布、学历年资、职称职级，以及招聘和用人规划等等基础数据；（4）教学及科研类数据源，如各轨道（本科、研究生、规培生、等）教学任务、课程、获奖、平台和称号等；科研数据包括横向和纵向课题、临床试验 GCP 等；（5）资产类数据源，包括无形资产和固定资产，如设备购置、折旧、使用、报废，尤其是使用过程中的归属、调拨、使用时间、频次、耗能等各类物联网数据等；（6）全链条物资供应业务端数据源，包括各类别物资的招标、采购、入库、出库、使用、维修保养等全业务链条数据等；（7）后勤类数据源，如各类能耗、洗涤消毒、维修管理数据等；（8）基建类数据源，如大型基础建设的招标、修缮的预算和支出数据等。

4.2 达成系统间数据互联互通，标化运管数据

将上述数据源整合应用，建立数据间互联互通的结合应用基础，通过以下几种常见的连接轴线达成各类系统数据源的“手拉手”连接，为发挥整合数据资产的管理价值打下扎实的基础。

（1）以独立核算科室单元为轴线实现数据互联互通。医院各类业务系统的数据因管理角度不同，在归集颗粒度和应用名称代码方面互不相同，如第一呼吸内科，在人力资源管理系统中按照科室成立文件，为“第一呼吸重症监护病房”；在 HIS 业务端中因医疗管理需求区分为“第一呼吸门诊”、“第一呼吸病房”、“第一呼吸重症监护室”，以及“内科门诊（护理单元）”、“第一呼吸内科、重症监护室（护理单元）”更细致的颗粒度；在财务核算业务系统中，依照财务归集习惯，称为“第一呼吸”；在物资管理系统中，依照物资管理习惯，称为“呼吸内科（南湖院区）、呼吸内科（滑翔院区）、呼吸内科（沈北院区）”等。基于上述情况，以“独立核算科室单元”为轴线，统一上述不同层次不同颗粒度的数据归集名称及系统间的对照关系，达成系统间数据的连接，“标化后”的数据，不仅达成业务数据的串联，更为综合运营管理过程中评估科室单元间的收入、全成本核算及逐层归集、人力资源情况以及物资使用情况建立扎实基础。

（2）以人员基础信息为轴线实现数据互联互通。公立医院尤其是大学附属公立医院，医务人

员通常兼顾医疗、教学、科研及管理各项工作范畴，在人力资源管理上具有多岗位层次、多岗位职责的实际特点。在“大运营”管理中对人力资源应用效能的评估对各类基础数据源的互联互通提出明确要求。以人员基础信息为轴线，连接各类人员的基础信息、医疗业务数据、医保业务数据、科研业绩数据和教学任务数据，薪酬支出数据等，为规范诊疗管理、评估教学科研能力，全方位精细化人员绩效考核提供坚实基础，同时也为医院优化人员配置、动态人力资源管理提供夯实有力的大数据基础。

（3）以全生命周期设备/物资码为轴线实现数据互联互通。以管理者视角看待医院设备/物资管理，更期待能够全生命周期地了解医院采购的设备/物资的流向、评估投入和产出效能，以及在医疗过程中发挥作用和使用的合理性等各方面的信息。基于上述管理需求，以设备/物资码为轴线的各业务系统间的大数据互联互通主要体现在与设备维修数据、医疗业务数据、医保数据等方面的互联互通。通过评估设备的单机效能、物资的诊疗合理性等分析运用，为精细化物资管理、高效能设备管理提供高效的数据整合基础。

基于上述举例，建立业务系统间的大数据互联互通，形成高质量标化的大数据基础，为多维度的综合大运营管理联动评估、绩效考核，以及高效能资源配置提供了有力基础。整合数据源的“大运营”管理体系构建和工作思路是高质量发展的必备要素，系统间多维度数据源连线模型，详见图 2。

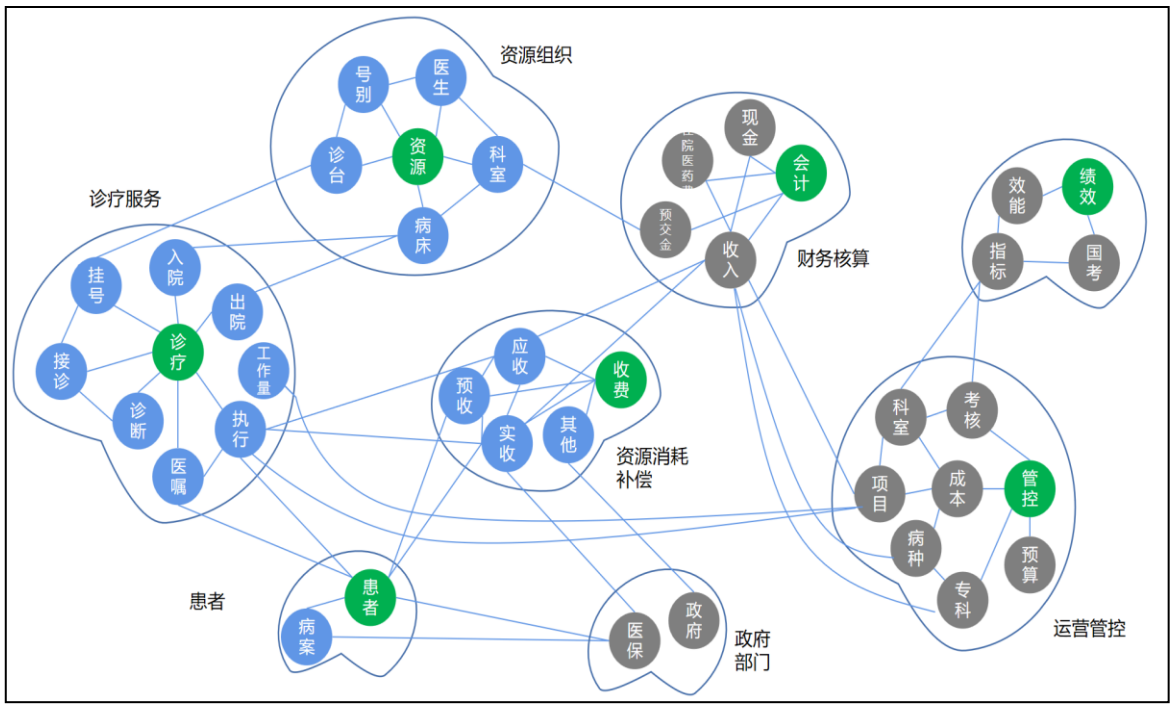


图 2 数据源的镜像汇集及系统间大数据互联互通的大运营管理模型构建

4.3 积累运管数据知识库，建立智能决策模型

4.3.1 形成运管数据分析指标库：

结合公立医院绩效考核指标体系（文件号）、高质量发展指标体系（文件号）、三级医院等级评审指标体系（文件号）、医疗质量指标体系（文件号）等政策性指导标准，建立综合运营管理指标库，统一指标计算口径，从“视图数据采集”，到“指标质控应用”，再到“与业务部门数据见面”，通过“治理数据”，达成“数据治理”，所谓“数据越用越准”，把更多的数据用起来，发挥管理效能。

监测指标和考核指标间的动态转换。在监测指标中，定向结合各方面的管理重点，做指标分配，依照管理权责分工，把管理数据权限分配给各职能管理部门。对于符合医院阶段性管理重点的关键指标，作为“考核指标”纳入绩效考核，与科室或部门 KPI 挂钩，体现管理效能的最大发挥。

4.3.2 形成临床业务管理知识库

依照管理指标库的监测和分析情况，挖掘管理的要点和阶段性重点，形成综合运营管理决策、定向职能管理、临床业务行为三者间的管理共识。在逐层协同管理的闭环管理体系中，建立业务数据知识图谱，为临床医生诊疗过程提供知识库的辅助诊疗决策支持，并且在应用过程中达成知识图谱的不断学习和优化，进而更好的提供辅助诊疗的支撑作用，数智化诊疗知识图谱模型构建详见图 3。

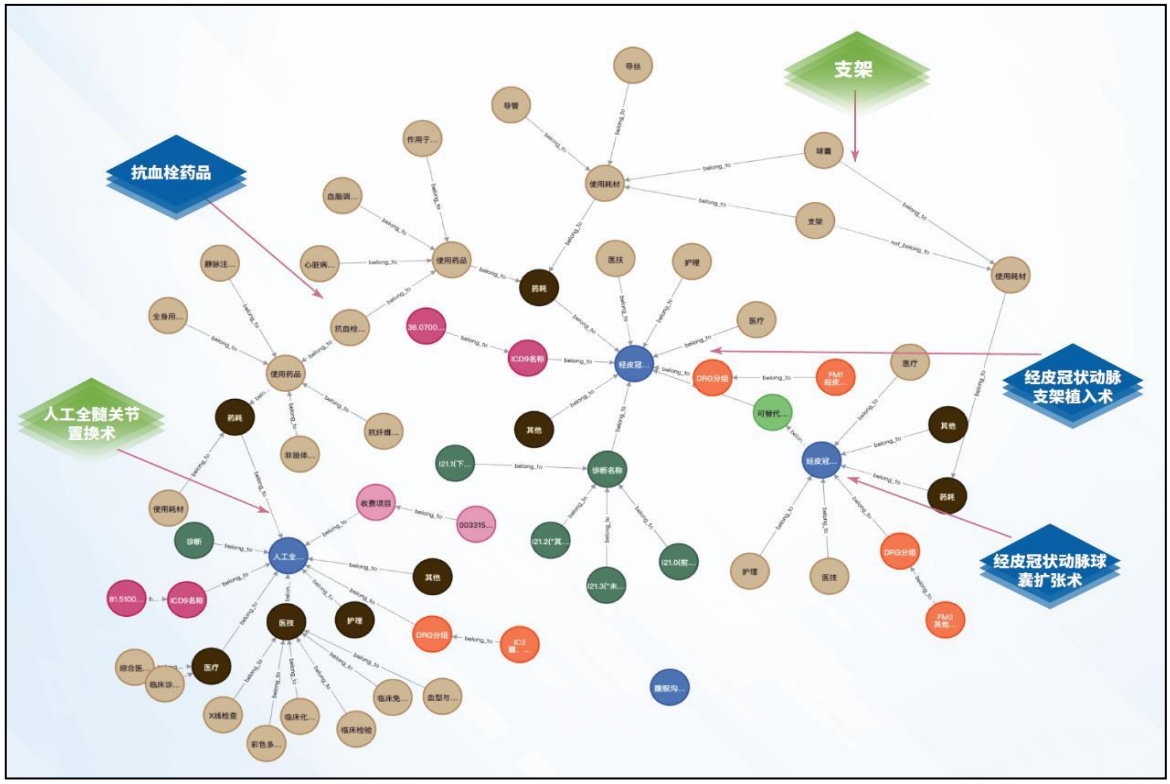


图 3 数智化诊疗辅助知识图谱的构建模型

4.4 数智化运管模型辅助高质量决策智慧管理应用

基于上述大运营管理理念，构建数智化运营管理模型，辅助医院高质量决策智慧管理应用，

主要体现在以下方面：（1）充分发挥大运营综合管理作用，横向串联管理核心要点，保障医院良性运营基石，充分达成资源优化配置和最大效能发挥；（2）针对医院资源调整规划，建立数据模型，为资源配置做好计算和预测，由拍脑门管理决策向以数据为依据的精细化科学化管理转变；（3）公立医院绩效考核统筹管理，恒量公立医院高质量发展的重要标尺是国考，基于大运营管理数据模型，构建院端国考管理数据依据，为深化国考导向实践应用；（4）专项物资精细化管理，实现高值耗材、不收费材料的精细化管理，与 DRG/DIP 资源消耗联动，真正实现降本增效；（5）专项设备精细化管理，通过单机效能评估管理，管控跑冒滴漏、资源闲置问题，真正构建“花钱问效、无效问责”的精益管理新局面；（6）内部绩效杠杆考核管理，建立各环节数据互联互通，为科室层级、个人层级的绩效考核打下坚实基础，形成以数据为基础的管理依据，科学管理、科学考核。

基于上述管理实践，搭建以“多系统集成的大运营管理数据源”为基础的数智化综合运营管理平台，信息化建设系统搭建模型图详见图 4。

图 4 以“多系统集成的大运营管理数据源”为基础的数智化综合运营管理系统模型

5.结语

公立医院高质量发展离不开医院智慧管理，大运营综合管理理念通过“治理数据”达成“数据治理”，是智慧医院管理的新基建，是承载综合运营管理思维和思路的重要载体。数智化综合运营管理模型构建的管理思路，和系统平台工具建设实践，为现阶段各医院高质量发展提供管理思路，管理思路详见图 5。

图 5 数据驱动大运营综合管理模型构建的管理思路图

壁报交流

PO-001

基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统的设计

唐跃涛 秦哲
永州市中心医院

1 引言

门诊部、病友服务中心是医院面向社会的服务“窗口”，是病友了解医院、认知医院的重要场所。病友服务中心每天需接待大量的患者及家属，是人群相对集中而流动性较大的场所，以往的门诊就诊流程复杂，普遍存在着看病难、流量多、时间慢等问题。患者就医需要排长队在窗口挂号，即使目前医院开通了电话预约，自助预约，手机端预约挂号功能，一定程度上为解决患者就医带来了一些便捷，但是号源分配的合理性与可控性存在管理上的不便，医生资源的合理分配上仍然存在较大不足，病友服务中心的医护人员需要在处理患者门诊就医问题上浪费大量时间，并且经常与患者发生争执，容易导致护患纠纷；电子健康卡是随着移动通讯技术的发展，在手机终端虚拟实际健康卡的一种形式。电子健康卡是由国家卫生健康部门统一规划，统一标准，面向全国居民发放，基于目前新冠肺炎疫情的防控需要，患者来院就医时，必须要出示电子健康卡，电子健康卡已经成为患者来院就医不可或缺的工具。因此将电子健康卡作为就诊介质应用于智能弹性门诊及住院部复诊预约管理系统不仅符合国家推行的全民健康信息化的要求，也是便民利民的重大举措，同时也是建设智慧医院的重要组成部分，而且能帮助我院门诊部优化就医流程，优化号源管理，充分调动医生的积极性，提高医院门诊量，还能解决患者排长队挂不到号，好不容易等到了一个号源，确不是心仪医生的号源的问题，从而改善病人来院就医环境，给患者带来更多福音。

2 需求分析

2.1 电子健康卡：一码在手，就医无忧

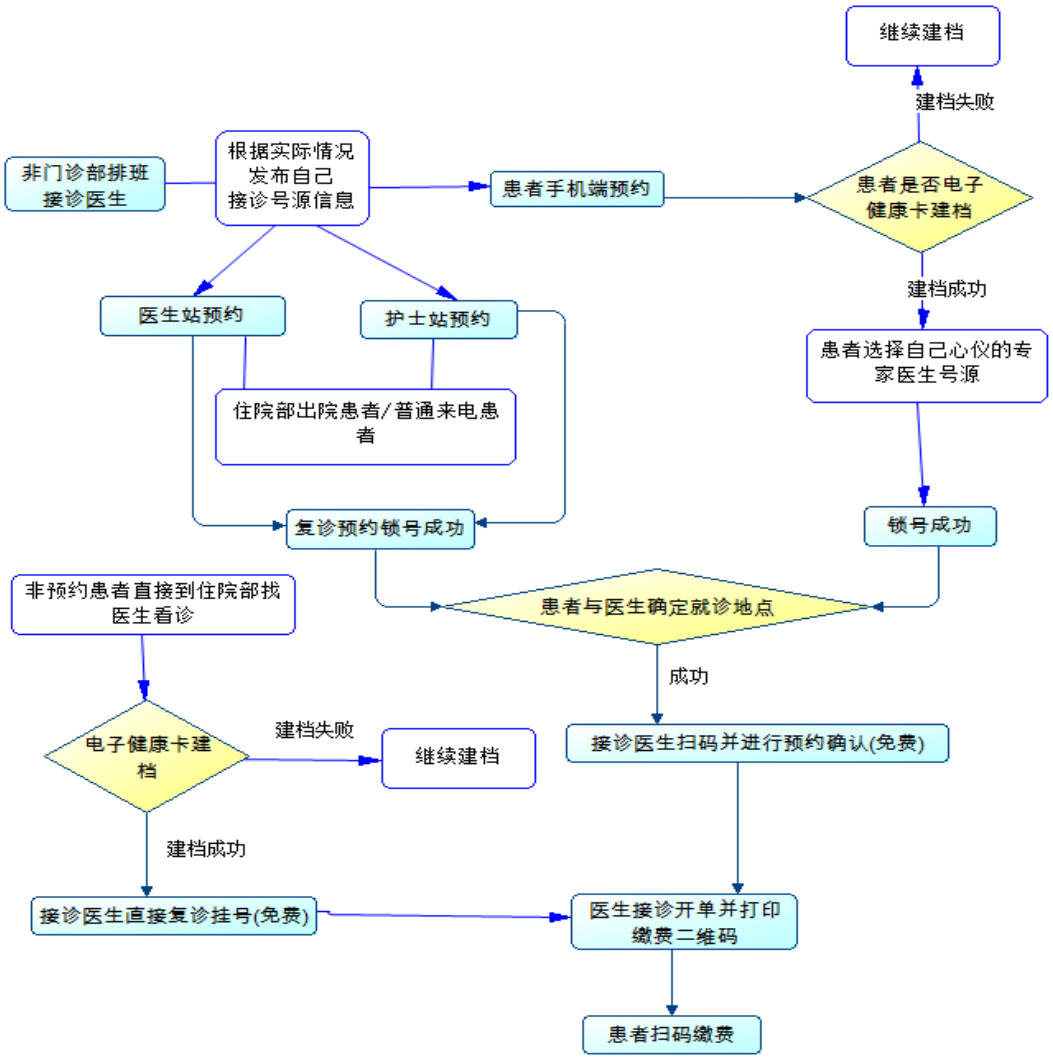
随着区域卫生信息化建设的加强，湖南省全民健康信息平台的建设，电子健康档案全程记录了居民从出生到死亡的相关医疗保健和医疗机构就医记录，随着居民电子健康档案调阅功能的开放，患者来院就医时，医生通过调阅健康档案，可实时了解患者的历史就医情况，包括历史检查检验，用药情况，为接诊医生精准诊疗保驾护航；同时基于目前新冠肺炎疫情的常态化防控要求，患者来院就医时，必须要出示电子健康卡，让医护人员实时了解患者的感染风险情况，通信大数据行程记录情况；结合目前永州市中心医院医院信息化建设实际情况与医院大力实施便民举措，我院自主研发基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统是积极响应国家卫生部门推广应用电子健康卡造福国人的一个缩影，同时也是医院积极建设智慧医院的重要举措。

2.2 优化就医流程，改善就医环境

随着医院规模日益发展壮大，来我院就诊的病人多而杂，尽管医院门诊部已经开通了电话预约，自助预约，网站预约，手机预约等渠道，但是高峰期时间段仍然存在排长队和因插队造成纠纷的现象，仍然存在患者因挂不到号，挂不到心仪医生的号而与医院医护人员发生口角的现象；研发基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统是服务全民健康的要求，是解决患者挂不到号，挂不到心仪医生号源的绿色通道，更是医院优化医疗资源，为医生充分发挥自己主观能动性开辟的一条绿色通道；传统的门诊号源由门诊部统一管控，班次的异动上不够灵活，且医生的接诊地固定在门诊大厅诊室，门诊医生坐诊方式都是实行轮班制，不便于患者复诊就医，门诊就医等候时间长，普遍存在有钱看病难的问题；因此本系统在号源设计上采取医生自己管控号源，并且可以根据自身的实际工作时间自由管控号源。无论医生身处办公楼、手术室、住院部办公室，只要身旁有一台安装有系统的终端，就可以释放自己的号源并接诊病人，极大的简化了病人的就医流程，同时复诊病友可直接与首诊医生约定复诊时间，复诊地点进行看诊，避免在门诊大厅排队等候就诊，极大地方便了患者复诊就医，系统通过与院内已建的医院信息系统实现数据共享，全程跟踪患者的就医状态，病友的医疗数据在各部门之间高速传输，使得病友就诊高效便捷，为病友就医提供更加优越的环境。

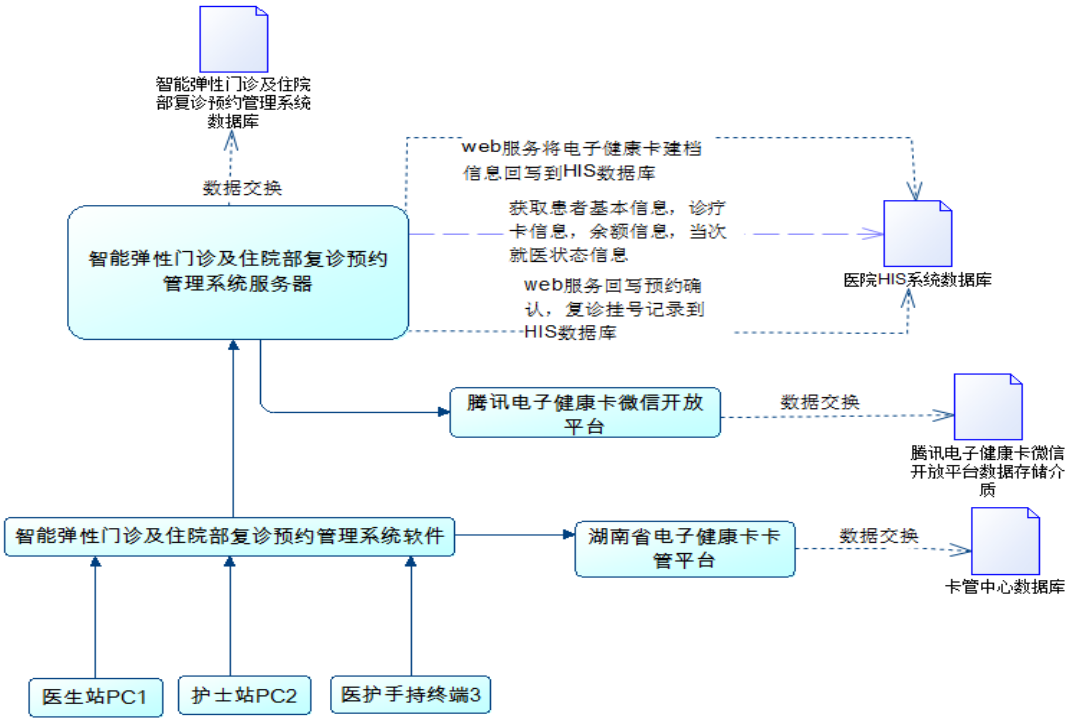
3 流程设计

3.1 智能弹性门诊及住院部复诊预约管理系统业务流程



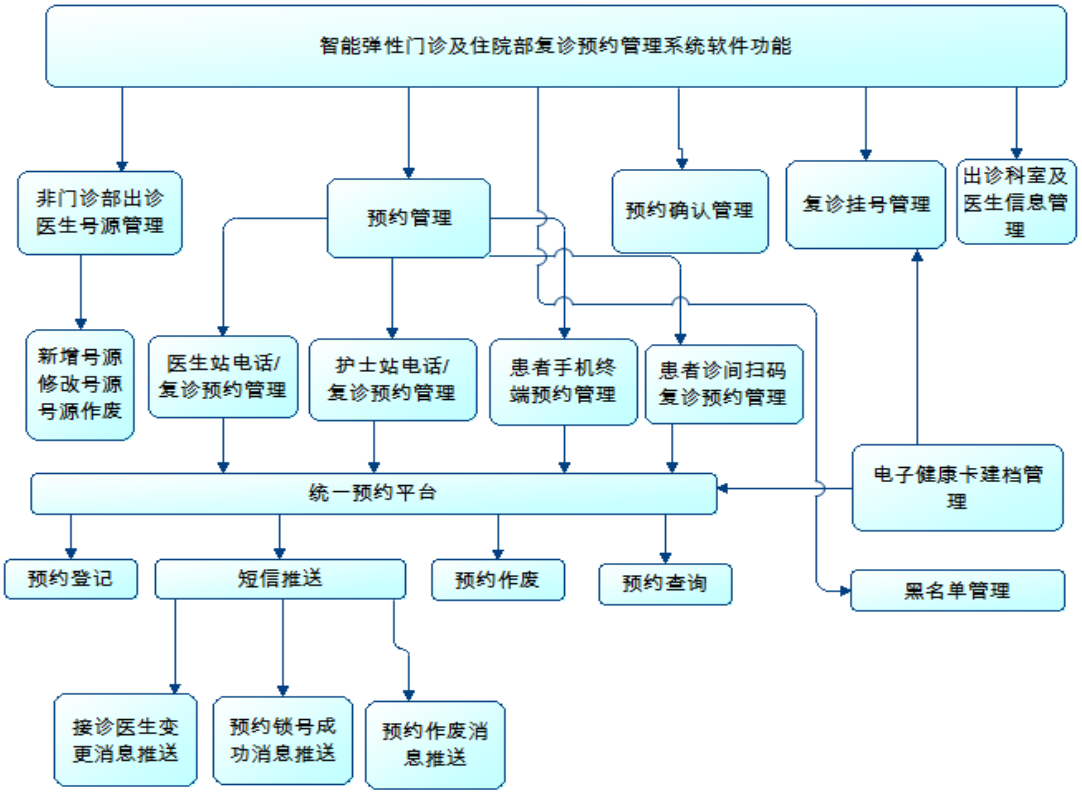
智能弹性门诊及住院部复诊预约管理流程图

3.2 智能弹性门诊及住院部复诊预约管理系统软件结构



智能弹性门诊及住院部复诊预约管理系统软件结构图

3.3 智能弹性门诊及住院部复诊预约管理系统软件业务应用功能



智能弹性门诊及住院部复诊预约管理系统业务应用功能图

3.3.1 电子健康卡建档管理

以电子健康卡作为个人身份识别读介质，持有智能手机的居民可以通过在医疗机构部署的自助

机读取身份证方式获得电子健康卡，也可以通过手机微信方式注册申领并激活。

3.3.2 非门诊部出诊医生号源管理

打开 PC 端智能弹性门诊及住院部复诊预约管理系统，医生选择自己想出诊时间段并生产号源，如果由于出诊时间段异动可进行号源修改，号源作废。对于已经预约的病友通过短信平台推送医生变更消息告知患者。

3.3.3 医护工作站预约管理

医护人员在复诊预约管理登记界面通过扫瞄电子健康卡获取患者的基本信息，选择患者来院复诊日期和对应的时间段进行预约登记，登记完成通过短信平台推送预约成功消息，消息包含就诊时间，就诊科室，就诊医生，科室联系方式，医生联系方式，就诊地址信息，就诊注意事项等，预约医生在出诊的前一天再次对所有预约患者进行短信推送提醒。

3.3.4 诊间扫码复诊预约管理

诊间接诊医生给病人看诊后，患者有复诊需求，患者可直接扫码医生二维码，进入弹性门诊统一预约平台选择该医生的出诊日期，就诊人选择建过档电子健康卡并完成预约，预约成功后通过短信平台推送预约消息，消息包含就诊时间，就诊科室，就诊医生，科室联系方式，医生联系方式，就诊地址信息，就诊注意事项等。

3.3.5 患者手机终端预约管理

患者关注永州市中心医院微信公众号后，选择弹性门诊及住院部复诊预约进入统一预约平台，患者选择就诊科室，就诊医生，就诊时间段，就诊人选择建过档电子健康卡并完成预约，预约成功后通过短信平台推送预约消息，消息包含就诊时间，就诊科室，就诊医生，科室联系方式，医生联系方式，就诊地址信息，就诊注意事项等。

3.3.6 预约确认管理

患者来院到达接诊医生约定地点后，接诊医生扫描患者出示的电子健康卡获取患者的预约记录后，医生在门诊医生工作站进行预约确认，确认完成后，医生可根据患者的病情开具处方并打印处方单，患者可使用手机扫描处方单上的二维码完成缴费。

3.3.7 出诊科室及医生信息管理

医院网络信息部维护弹性门诊及住院部复诊预约管理系统的就诊科室与医生基本信息，接诊医生可动态维护自己的个人简介及擅长信息，个人照片信息，个人联系方式等。

3.3.8 黑名单管理

系统设计自动加入黑名单功能及人工加入黑名单功能，对于 3 次无故爽约患者，自动加入黑名单，禁止通过平台预约 3 个月；5 次无故爽约，禁止通过平台预约 1 年；连续 8 次无故爽约，禁止该患者通过平台预约。对于多次无故爽约的患者，医护人员可将该患者加入黑名单，禁止该患者通过平台预约。

3.3.9 预约查询

根据系统的预约情况，形成分析报表，可以综合评估医院一段时间内的病人的弹性门诊量及复诊预约量情况，还可以对病人的诊疗情况进行规范管理。

4 关键技术及模块实现设计

4.1 接口

4.1.1 电子健康卡二维码识读终端接口程序

在智能弹性门诊及住院部复诊预约管理系统软件终端上需要安装电子健康卡二维码识读终端，系统通过与湖南省电子健康卡卡管平台接口程序改造识别二维码或者身份证号获取该病人相关信息。界面设计上，当医护理人员选择电子健康卡，系统弹出识读提示等待框，当患者将手机上激活成功的电子健康卡二维码放在终端上，扫码成功后，提示框消失，并将患者基本信息显示在预约系统界面；对于没有激活电子健康卡的患者，系统支持通过填写患者姓名，身份证号，性别，出生日期，家庭住址，联系电话等信息调用湖南省电子健康卡卡管平台进行电子健康卡注册。

电子健康卡识读提示框图

患者扫码示意图

4.1.2 预约确认、医生站直接复诊挂号与医院 HIS 系统接口程序

医护人员在门诊医生工作站界面，通过扫描电子健康卡二维码获取患者预约登记信息后，执行预约确认时，系统需要按照医院 HIS 系统第三方挂号接口协议，通过 webservice 服务将该患者基本信息，预约信息发送到 HIS 系统，HIS 系统处理完成后将处理结果反馈给复诊预约系统，系统接受到成功标识后，更新患者的预约状态，并提示医生挂号成功可以接诊患者开立处方了，同时调用短信平台接口通知患者预约记录已经进行了预约确认；当医生接诊的是非预约患者，医生通过扫描电子健康卡二维码获取患者基本信息后，系统需要按照医院 HIS 系统第三方挂号接口协议，通过 webservice 服务将该患者基本信息，挂号信息发送到 HIS 系统，HIS 系统处理完成后将处理结果反馈给复诊挂号系统，系统接受到成功标识后，将患者的复诊挂号记录进行存储并提示医生挂号成功可以接诊患者开立处方了。

4.1.3 患者手机端复诊预约与腾讯微信电子健康卡开放平台接口程序

患者通过微信公众号进行复诊预约时，对于未进行电子健康卡建档患者，系统支持患者填写姓名，身份证号，出生日期，家庭住址，联系电话等信息调用电子健康卡开发平台申领电子健康卡，为简化申领流程，对于已有电子健康卡的患者，系统支持一键注册申领，申领成功后，将建档信息进行存储，并将建档信息通过调用 HIS 提供的 webservice 服务同步到医院 HIS 系统。

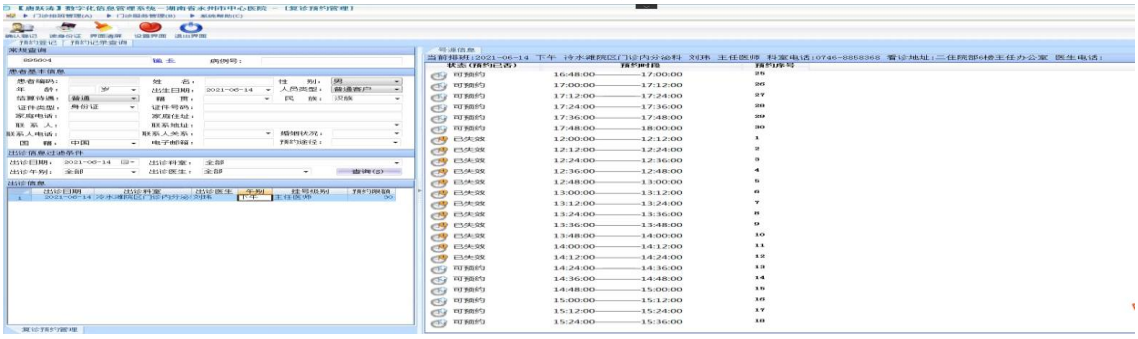
4.1.4 智能弹性门诊及住院部复诊预约与短信平台接口程序

为优化患者的就医体验，在接诊医生号源停诊，预约登记成功，预约确认成功，预约作废，诊前来院就医温馨提醒时，需要按照短信平台接口协议，将相关提醒消息推送至患者。

4.2 智能弹性门诊及住院部复诊预约管理系统界面设计

4.2.1 医护工作站复诊预约界面

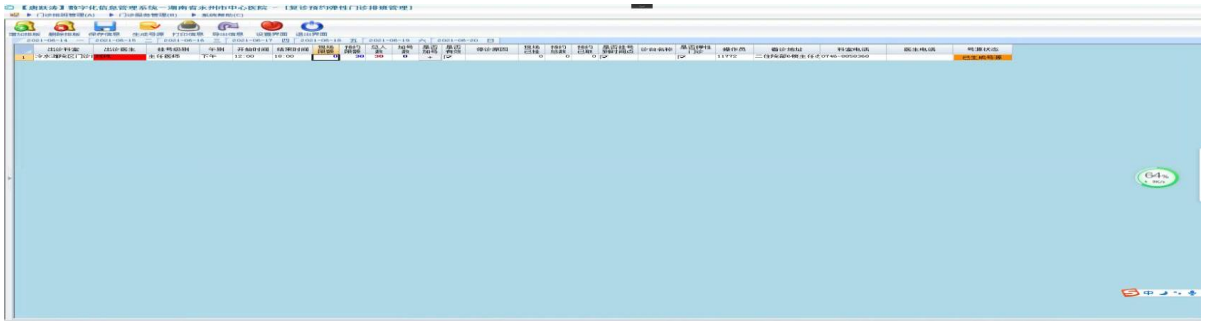
系统集成了通过扫描患者电子健康卡或身份证号获取基本信息进行预约登记，打印患者预约凭证、已登记信息的查询、预约登记作废和患者预约凭证的补打功能，



医护工作站预约登记主界面图

4.2.2 接诊医生出诊号源管理界面

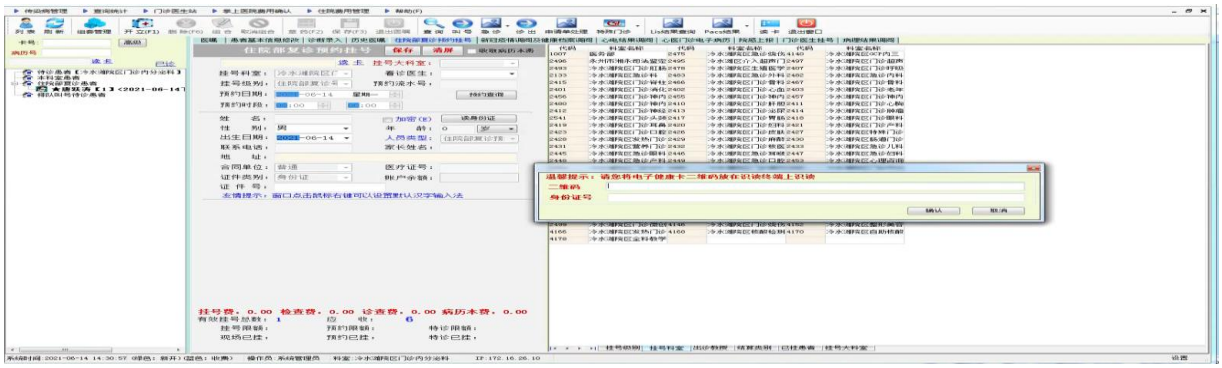
系统集成了接诊医生管理自身号源功能，医生可以自由设置自己出诊的时段，释放号源数，接诊地点，联系方式等，当医生出诊时间发生异动时，可灵活停诊，并短信通知患者。



接诊医生号源管理主界面图

4.2.3 医生工作站预约确认及复诊挂号界面

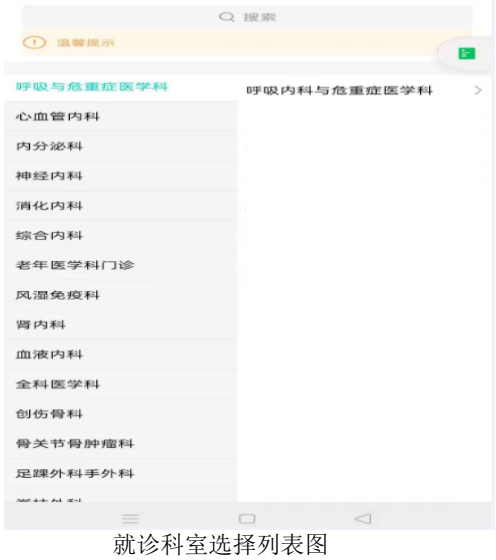
将智能弹性门诊预约确认与复诊挂号功能嵌入到门诊医生工作站，简化临床医生接诊患者操作流程，提高工作效率。



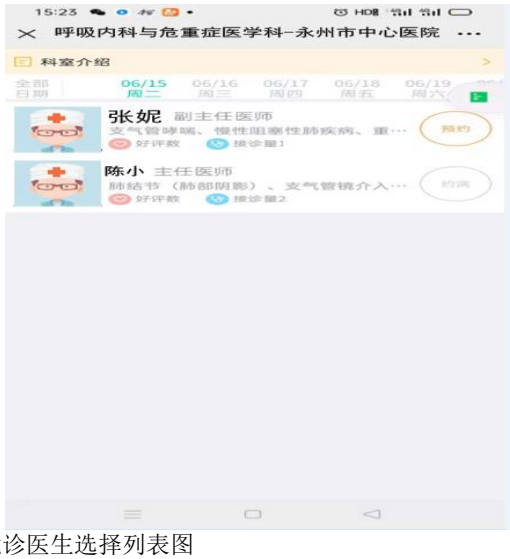
门诊医生工作站预约确认及复诊挂号主界面图

4.2.4 患者手机端弹性门诊及复诊预约界面

系统提供了院区选择功能，患者可以选择零陵院区，冷水滩院区；进入选择院区后，系统设计了二级菜单，选择相应就诊科室，再选择就诊医生，点击预约按钮后，进入医生简介，擅长详细信息界面，患者选择相关就诊号源信息，就诊人信息选择对应患者的电子健康卡后，点击提交订单。



就诊科室选择列表图



就诊医生选择列表图

5 小结

研发基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统结合了当前互联网技术并对医院普通门诊患者及复诊患者的就医流程进行了优化改进，有效弥补了传统固定门诊诊室排队就医模式的不足，与传统医院门诊就诊固定医生接诊号源，固定医生接诊时间，固定医生门诊诊室接诊地点流程相比，本系统具有强大的功能和可操作性，当门诊部诊室号源不足时，等候就诊排长队时，患者

检查检验结果出来后，当天接诊医生已不在门诊诊室接诊时，外地患者远道而来一号难求时，患者可以通过手机终端进入弹性门诊及复诊预约管理系统选择心仪的专家进行看诊。也可直接拨打科室电话与科室医生约定看诊医生与看诊地点进行就诊，从根本上解决了患者看诊排长队等候，来院就诊无医生接诊空跑腿的问题。通过本系统建设优化了医院的医疗资源，充分调动了医生的主动能动性，让医院的专家接诊患者不受时间控制，不受地域控制；医院管理者借助信息化手段将患者就医进行分流，既保证了医院门诊就诊工作的有序进行，又保证了医疗服务的高质量，同时也为患者提供了安静、温馨的就医环境。医院管理者可以借助本系统提供的综合决策报表，制定合理的绩效考核方案，不仅可以提高医院医生的工作积极性，提高他们的福利待遇，而且可以提高来院就医的门诊量，同时也为医院带来巨大的经济收益。研发基于电子健康卡的智能弹性门诊及住院部复诊预约管理系统为医院的贡献不能简单用数字来衡量和定义，本系统与院内外其他系统进行无缝对接，实现医疗数据的实时共享，它是在无形中为医院创造了价值，直接提高了医院门诊管理效率和社会声誉，进而提高了医院的经济效益。依靠信息化的手段来优化门诊就诊流程，更加关注病人就医的满意度与体验感，使科室管理变得更加规范；展望未来，随着科学技术的进步，希望能有越来也多的人工智能、大数据挖掘应用于临床服务,，为临床医护人员的提供更多的辅助支持，为老百姓就医带来更多福音。

PO-002

上海市某三级甲等专科医院提高住院病案首页数据质量实践

马傲，马晴，刘枭雄，肖进，高静，葛小玲*
复旦大学附属儿科医院

1.引言

住院病案首页作为是病案中最重要内容的浓缩，是疾病诊断相关分组(DRG)付费、公立医院绩效考核、公立医院高质量发展评价的主要数据来源[1]。2023 年版《国家三级公立医院绩效考核操作手册》[2]56 项指标中，来源为病案首页指标项 7 项；《公立医院高质量发展评价指标(试行)》[3]18 项指标中，来源为病案首页指标项 7 项；而基于疾病诊断相关分组(DRG)的付费方案则以住院病案首页为所有数据来源。同时，住院病案首页因其结构化程度高、数据内容丰富、完整性准确性高等特点，在医疗机构内部的临床、科研、管理等需求中也具有极高的数据应用价值。因此，无论在医疗机构外部评价还是在内部数据应用需求中，住院病案首页的数据质量都显得尤为重要。如何提升病案首页数据质量，成为医疗机构质量管理的重要议题。

2. 现状分析

2.1 住院病案首页信息项

一份完整的病案首页大致可以分为四个部分。第一部分为住院患者姓名、性别、年龄、住址、联系人等个人基本信息项；第二部分为出院疾病诊断信息项；第三部分为手术操作、麻醉信息项；第四部分为住院费用信息项。根据最新版《全国三级公立医院绩效考核与医疗质量管理住院病案首页采集系统对接接口标准》，病案首页 811 个信息项中必填项 71 项，选择必填项 24 项；值域校验项（含 ICD 疾病诊断及手术操作编码字典）450 项；格式校验项 125 项；补充逻辑校验项 18 项。

2.2 住院病案首页全流程调研

以上海市某三级甲等专科医院为例，运用数据全生命周期治理的理念[4]，从住院病案首页前端业务流、后端数据流两条路径开展住院病案首页全流程调研。根据调研结果，前端业务流将病案首页全流程分为源数据录入、住院诊疗记录、编码统计和上报应用四大阶段共十三个步骤，见图 1。后端数据流将病案首页全流程分为基本信息、诊疗费用信息、编码统计信息、上报信息等四大环节共 14 个系统数据的流转整合，见图 2。

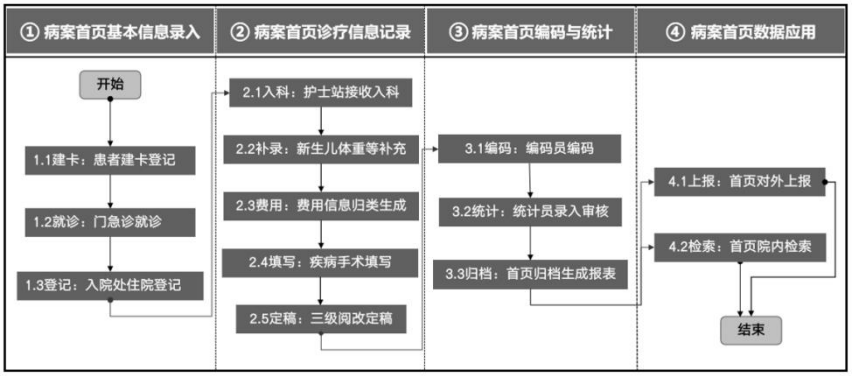


图 1 住院病案首页业务全流程

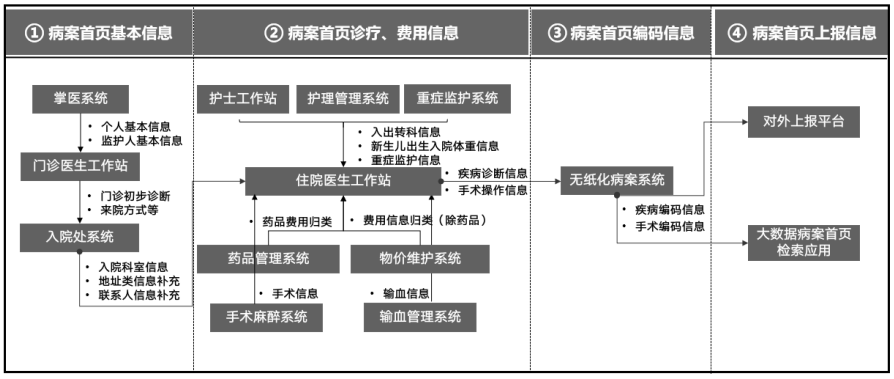


图 2 住院病案首页数据全流程

2.3 住院病案首页质量问题归类

根据调研结果，进一步分析归类全流程出现的住院病案首页质量问题。使用鱼骨图分析法[5]开展住院病案首页质量问题析因，结合“二八法则”，讨论总结住院病案首页主要问题主要有：①缺乏完整的病案首页全流程质控规则②系统功能不完善及系统 BUG③未建立统一的质控平台④人员层面诊断手术填写不规范。

3. 住院病案首页数据质量提升数据质量提升路径与措施

3.1 建立住院病案首页质控规则知识库

虽然国家、地区均发布了各类用于上报的接口标准。以上海地区为例，国家卫健委《国家公立医院绩效考核病案首页接口规范》、上海市卫健委《上海市医疗机构住院病案首页接口标准》、上海市申康发展中心《医联病案信息采集接口规范》、上海市病历质控中心《上海地区病历质控标准》等，均包含较多上报病案首页校验规则，但仍然存在质控规则不统一、缺乏针对病案首页内涵质量的控制等问题。医疗机构应从内涵质控与统一上报角度，建立满足病案内涵质控和各类上报规则映射的完整质控规则库。上海市某三级甲等专科医院利用现有的病案首页数据质量管理规范、国家及地区标准、国家疾病分类等外部数据为基础[6]，结合 DRG 疾病诊断相关分组规则[7]，梳理形成医院住院病案首页质控规则知识库。规则库包含非空检验规则 105 项，值域校验规则 92 项，格式检验规则 60 项，逻辑校验规则 122 项。其中，逻辑校验规则重点补充了病案首页诊断与编码内涵控制。如：“当性别为男性时，出院主要诊断编码不能使用以下编码：A34；B37.3；C51-C58；……”、“当性别为女性时，出院主要诊断编码不能使用以下编码：B26.0；C60-C63；D07.4-D07.6；……”、“出院主要诊断编码码范围应为：A～U 开头和 Z 开头的编码；不包括字母 V、W、X、Y 开头的编码”[8]。

3.2 完善住院病案首页系统整改流程

在医疗机构病案首页数据实际生成与流转各过程中，各类系统缺陷导致的病案数据错误无法彻底避免，如在数据传输过程中的字段丢失，字段提取错误等。建立及时、规范的系统整改流程有助于病案首页数据问题及时发现与修正。全流程的病案首页数据调研结果同时表明，如果不在系统前端进行及时的病案首页数据质量源头控制，后续进行数据清洗与修正将会产生较大的数据治理负

担。因此，将病案首页数据治理规则适当地嵌入系统，充分完善系统功能，则既不会增加系统使用者的门槛进而影响业务，又会获得质量较高的病案首页数据。上海市某三级甲等专科医院设计并上线医院系统整改流程，将数据应用中发现的系统问题通过多部门联合讨论、明确系统整改方案、规范的流程审批、执行整改与确认等步骤，实现病案首页数据应用推动系统整改，系统整改提升数据质量继而进一步推动病案首页数据应用的持续改进闭环。见图 3。

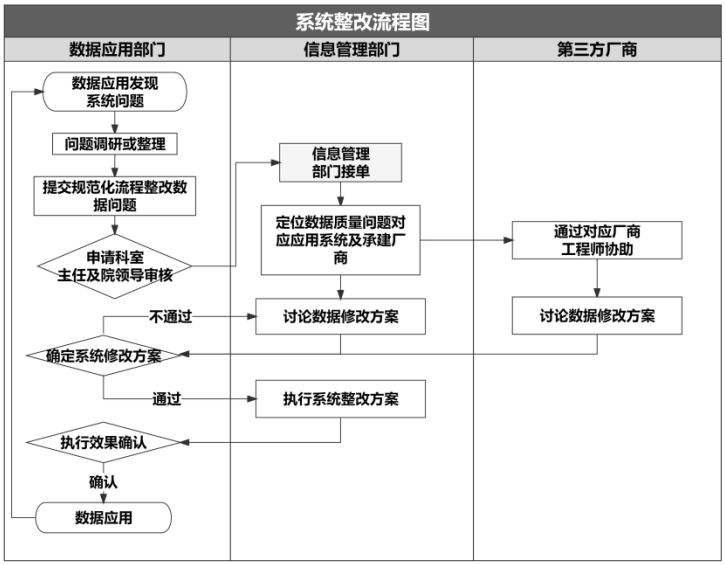


图 3 住院病案首页系统整改全流程

3.3 应用大数据技术开展病案首页质控分析

健康医疗大数据因其大体量(Volume)、多样性(Variety)、时效性(Velocity)、准确性(Veracity)、大价值(Value)，即“5V”特性飞速发展[9]，大数据技术在医院绩效管理、精准医疗、科研管理以及健康管理等方面均具有广泛应用[10-12]。医疗机构通过将大数据融合到医院内部质量控制工作中，是“将信息化全面融入到公立医院高质量发展”的重要措施。上海市某三级甲等专科医院通过建设医院大数据平台，开发大数据质控分析平台，实现病案首页数据的全流程质量控制。基于 hadoop 分布式计算框架的大数据平台[13]在数据采集层融合汇聚病案首页源系统数据表，在数据质控层将质控标准与规则与现实数据对比分析形成质控报告；针对质控报告，有的放矢地开展清洗、映射、转换、校验等数据治理工作，形成数据治理层；将治理后数据用于院内外各类病案首页数据应用，在应用中持续关注病案首页数据质量。见图 4。在大数据平台病案首页数据全流程质控中，实现规则、流程与平台结合，任一环节发现数据问题，即可回溯源头问题发起系统整改。

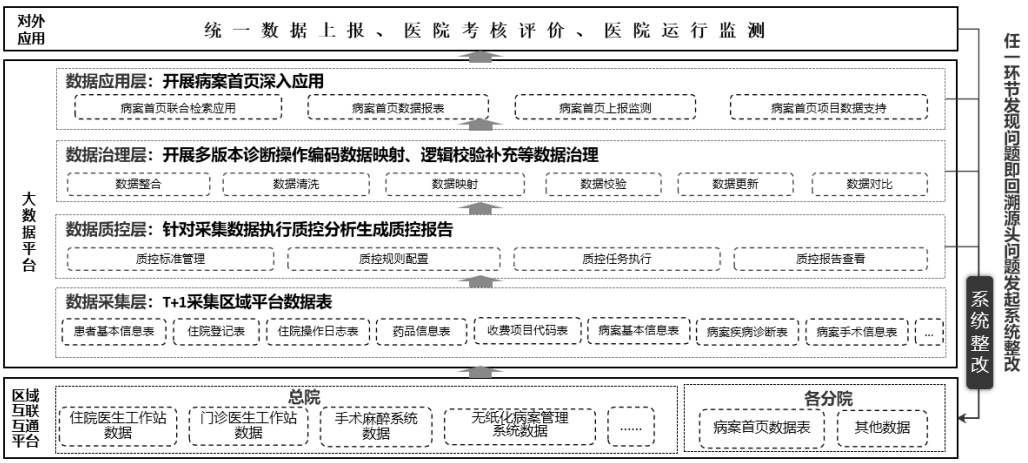


图 4 大数据平台病案首页数据质控架构

3.4 开展住院病案首页临床入科培训

针对人员层面疾病诊断与手术操作填写不规范等问题，开展临床入科培训十分必要。以病案首页出院诊断填写与编码为例。疾病诊断是为疾病起一个特定的名称，使之区别于其他疾病。而疾病诊断 ICD 则是 WHO 制定的国际统一的疾病分类方法。疾病的分类并不等于疾病。实现由疾病诊断到疾病分类，需要临床医生与编码员的共同努力。对于临床医务人员，正确、全面、准确地填写疾病诊断名称十分重要，是疾病编码的前提；对于编码员而言，只有充分、详细地阅读临床医务人员对疾病的判断与描述，才能实现对疾病的准确编码，是疾病诊断的总结与归类。开展临床入科培训，一方面培训医务人员在病案首页书写中应遵循的原则与规范，另一方面，加深临床医务人员与编码员的深度沟通，打破临床医生与编码员故有的思维壁垒，有利于编码员进一步理解临床医生基于临床视角的诊断原因，也有利于临床医生对编码员疾病编码的认可与促进，从而充分提升病案首页数据的内涵质量。同时也应注意到，由于各专科的独特性，临床培训除了做到院级层面的统一集中培训外，也应开展针对各专科的精细化培训，甚至是深入到对临床医生的重点培训。临床入科培训也不能是一蹴而就的，而应形成常态化培训与专题培训相结合的长期培训计划，将阶段性的主要问题以理论结合具体案例的形式展开培训与交流，做到活学活用，常学常新。

4. 结语

随着公立医院高质量发展的不断深化，医疗机构逐渐由粗犷的规模扩张型向精细化的质量管理转变。住院病案首页数据作为重要的质量考核依据，其数据质量提升仍是一项长期持续的工作。在病案首页数据质量提升的过程中，应充分利用建立规则、完善流程、开发平台和培训人员等多种方式，充分调动与整合各类资源，实现病案首页数据质量的持续提升的目的。同时还应认识到，病案首页数据质量提升的方法与路径也应不断与新技术新方法结合，如利用人工智能技术建立病案首页系统辅助填写 AI 模型[14]、基于新型 CDSS 辅助病案首页填报质控[15]等，不断探索更有效更智能的数据质量提升新方案。

PO-003

数字疗法在心理健康领域的发展态势及分析建议

李晋芳[1,2] 谢文照[1,2*]
1.中南大学湘雅三医院医学信息研究重点实验室
2.中南大学生命科学院

引言

当前，数字经济蓬勃发展，数字技术正在深刻改变着人们的生活。在新冠疫情的大背景下，因为患者无法亲自到医院就诊，加之部分心理疾病患者不愿意承认病情或面对医生，国内外数字化的心理健康创新产品随之出现。数字疗法(digital therapeutics，DTx)作为一种新兴的治疗方法，致力于实现智能化的软件驱动干预、跨时空服务患者，突破传统药物或者医疗器械治疗的局限性，有效弥补了传统治疗中缺失的依从性和可及性。

在世界各地，心理健康一直是一个被忽视的健康挑战。目前全球约 10 亿人正面临着精神问题，平均每 40 秒就存在 1 人自杀，然而仅仅少数人能享受到高品质的心理健康服务。最新发布的《2022 国民抑郁症蓝皮书》也显示中国当前患有抑郁症的人数约为 9500 万。在每年自杀的人中约 40%均患有抑郁症。心理健康问题亟待国家、社会及大众关注，精神卫生服务的获取问题亟待解决。而心理健康实践逐渐转向数字疗法干预，可以改善心理健康服务的获取和提供。数字医疗的兴

起对包括互联网医疗和 DTx 在内的数字医疗创造了巨大需求，推动了对全球 DTx 行业的需求发展。

1 数字疗法在心理健康领域应用的定义、特点

1.1 数字疗法在心理健康领域应用的定义

2017 年国际数字疗法联盟（Digital Therapeutics Alliance,DTA）对数字疗法的定义如下：患者利用循证治疗干预来预防、管理或治疗病情的一款医疗产品，不仅能够单独使用，还能与不同药物、设备、其他物理因子和疗法等配合使用。近年来，心理健康问题逐步成为世界各国普遍面临的社会性难题。我国同样也不例外，这些心理健康问题不仅存在于成人、职业群体，甚至扩展延伸至大中小学生群体，心理障碍患者呈现出“低龄化”的发展趋势。而心理健康领域的疾病仅依赖药物或者医疗设备很难完全根治，“低龄化”患者很大程度不愿意去医院面对面就诊，或者说因为是非观价值观未发展成熟而难以及时发现自身心理健康存在的问题。

1.2 数字疗法在心理健康领域应用的特点

数字疗法应用于心理健康领域主要通过智能手机应用程序、互联网网站、可穿戴设备、虚拟现实或视频游戏等数字化技术提供精神卫生健康信息和支持的干预方式，包括自助式、医生支持或线上线下相结合等几种模式。循证医学的证据大都表明药物不能够改善核心症状，只能控制症状，行为干预才是最主要的。应用行为分析(ABA)、融合教育、早期介入丹佛模式(ESDM)，后期在这些基础上加入改善认知、社会交往、教育的干预行为,心理学方面的 CBT 认知行为疗法等,都能够改善心理健康状况。

数字疗法 DTx 产品的主要特征是：为患者服务；以循证医学为基础，需通过临床的持续性、有效性验证；以 APP 驱动为核心。针对与心理健康领域的 DTx 产品的特征包括周期管理长、诊后干预措施较多；可低成本跨时空地远程为患者提供精准的干预方法，这在一定程度上避免了大幅度占用医疗资源，也相对减轻了社会的医疗负担。患者的心理健康问题也可以及时得到反馈和疏导；而且 DTx 属于医疗器械，使用之前需经过监管机构的审查、批准、认证。

2 数字疗法在心理健康领域应用的发展历程

2.1 国外数字疗法在心理健康领域应用的发展现状

世界卫生组织在《数字健康全球战略（2020－2025）》中，明确数字健康战略在世界各国的医疗卫生行业发展中的优先地位。目前国外有越来越多的数字疗法应用于治疗心理健康疾病，极大地减轻了对传统医疗资源的依赖。数字疗法有望在治疗抑郁症和焦虑症方向取得更好的效果。数字疗法在国外经历了多年的迭代创新。例如，可穿戴设备、VR 眼镜等的早期使用。自 2018 年以来国外一直在为在线心理治疗、聊天机器人研发或心理健康评估提供心理治疗和咨询领域的研发资

金。因为部分孤独症儿童的感觉通路是不同寻常的,他们的触觉极其灵敏,所以穿戴设备不易推广,反而是 App 大力推广,刚好大大降低了数字疗法的成本。

国际数字疗法联盟（Digital Therapeutics Alliance，DTA）是全球 DTx 具有影响力的推动者，其发布的《2018 年数字治疗行业报告》中要求在目标人群中进行一项或多项足够稳健、易于把控的且有目的的临床试验，包括通过随机对照试验来证明 DTx 的有效性和安全性；控制 DTx 产品正确处理输入数据，创建准确、可靠、准确的输出数据的能力；监测其实现临床治疗目标群体的预期目标的能力。

2013 年，美国发布了移动医疗应用指导最终版，为 DTx 的审批规范奠定了基础。2017 年，美国推出了数字健康预认证计划，为需要快速迭代的 DTx 产品提供了快速审批的渠道。FDA 批准的 DTx 在美国分类多为 II 类，FDA 于 2018 年 8 月批准的 Freespira 用于治疗恐慌症和创伤后应激障碍(post-traumatic stress disorder，PTSD)。再加上特殊的传感器、生理反馈显示器和指令，患者可以在 28 天内接受训练，使呼吸障碍正常化，这是焦虑发作和创伤后应激障碍症状（碳超敏反应）的主要生理驱动因素，89%的患者 PTSD 症状显著减轻，50%的患者处于缓解期和 77%的患者处于治疗依从性。FDA 于 2020 年 7 月批准的 Deprexis 作为 18 岁或以上正在经历抑郁发作的患者的主要或次要诊断辅助治疗慢性或复发性抑郁症。将抽取的患者随机分配到干预组和初始护理组，并在基线时通过电话对所有参与者进行诊断检查访谈，证实了 Deprexis 在干预后有效减少抑郁症。

德国在 2019 年宣布了数字医疗法案来定义数字疗法。德国联邦药品和医疗器械管理局要求 DTx 申请必须明确定义适用人群，从医疗保健中受益的数据必须来自至少一个定义的适用患者群体。在德国，数字疗法被描述为一种低风险医疗设备，用于诊断、监测、治疗和减轻疾病或支持这些功能。在实施前任何数字疗法 APP 必须首先列在德国的数字健康 APP。

2.2 国内数字疗法在心理健康领域应用的发展现状

近些年我国精神行为问题发病率为 16.6％，但是精神医护人员数量缺乏，全国仅有约 5 万名精神科医生，严重精神疾病患者有效管理率低于 68％。我国精神障碍的就诊率和患者认知远低于发达国家，大约有 92%的精神患者从未接受过治疗。“病因不明,无药可医”是业界对于这种病症的统一认识。在此背景下，我国 DTx 行业得到高度关注，并凭借极大的想象空间得以加速发展，DTx 行业具有较大的发展空间。目前，我国有 56 家初具规模的心理健康领域科技产品企业。2020-2022 年间，有 7 种创新技术产品用于治疗心理和心理健康，占 31.82%。2015 年我国相继颁布了修订的《中华人民共和国精神卫生法》和《全国精神卫生工作规划（2015—2020 年）》，心理健康服务行业迈入高速发展阶段。我国作为人口大国，心理健康服务需求大、市场成长空间大，是科技成果转化的关键领域，要争取实现尽可能多的可复制、经验可积累、低成本的数字疗法在心理健康领域的应用方案。

3 发展趋势

3.1 数字疗法在心理健康领域的整体研究态势

心理健康领域的数字治疗具有临床和商业价值，是数字治疗领域技术成果转化最有回报的投资领域。由于精神障碍的复杂性，数字治疗可以提供系统的解决方案来提高治疗效率。此外，精神数字疗法没有副作用，可以随时随地自我治疗。数字疗法的第一目标从来不是替代人工，而是结合人工，取得更高效、理想的康复效果。

对于康复机构，它可以自动监控督导师提高质量，减少培训难度和成本；并实现标准化管理从而降低运营成本。对于有心理疾病的患者来说，它则具有康复效果量化、进步可视化，用户黏性高、安全感高的优点。数字疗法的介入使长期看护、身心俱疲的看护人获得喘息和休息的时间。未来可以拓展研究数字疗法在对“冷门”的心理健康问题如人格障碍、睡眠障碍、躯体形式障碍和精神病性障碍方面的应用。此外，专业的疗法设计还能明确居家干预目标，避免居家干预不专业的问题。电子工程师协会（Institute of Electrical and Electronics Engineers，IEEE）数字疗法工作组将在 DTx 领域继续展开研究，目的是在可用性研究的基础上，进一步探索 DTx 产品用于认知治疗的安全性和其他基本性能。

3.2 数字疗法在心理健康领域面临的挑战

随着 Chat GPT 等先进人工智能（AI）崛起并被加入这些 APP，数字疗法在心理健康领域面临一系列的挑战，比如数字鸿沟带来的获取问题、技术更新过快、法律监管缺失、隐私安全问题、患者流失及成本效益问题。一个国家与数字实践实施相关的障碍取决于不同的因素，包括基础设施、法律和社会问题以及经济方面。最重要的障碍与成本和报销有关；法律责任、道德问题，如机密性、过时的设备、患者年龄和教育水平；计算机素养、带宽范围和互联网速度。数字疗法目前尚未形成清晰的商业模式，也极少被纳入保险的报销范围。目前，DTx 仍属于新兴事物，从患者到临床医师对于其认识相对有限，存在“玩游戏治病”等朴素的认知，甚至简单将其认为是心里安慰等误解，这极大地阻碍其推广。研究人员应该考虑从技术和支持的角度提高公众和医学对这一新事业的接受度。监管机构还应制定统一的标准来规范该行业，并加强医生和患者的信任。DTx 研究人员应该考虑患者的吸引力和治疗的有效性。增加患者参与度和自我导向行为的产品可以更受用户欢迎，并且可以通过增加娱乐性和丰富产品动作来实现产品改进。迄今为止，临床上对于的 DTx 相关研究受限于成本和时间的原因，其样本量和随访周期一般较为有限，而一些较为成熟的诊疗方法往往已有许多基于数万例病例长达多年甚至十余年的临床研究证据。数字疗法的核心是软件程序，监管机构传统的审批流程和周期并不适用于数字疗法产品，对于 DTx 产品的监管更需重视全生命周期的临床评价，以便更加科学、全面的评估 DTx 产品的安全有效性。

DTx 在中国的临床研究和应用仍然不足，很少有推荐的治疗方法进入现行的诊疗标准、临床指南和专家共识。医学界和工业界必须通过更多的临床研究和应用来促进科学认知。此外，针对不同

适应症的 DTx，如何在临床评价中充分识别产品风险受益亦是需要深入研究的问题，如某种特定疾病可能存在不同的分型分期(抑郁症、失眠等存在轻中重度等不同疾病程度等)，在同一适应症下，精细人群的同质性是否能够支持使用同一组临床试验，是否有必要对其亚组的数据结果进行研究和评估，以及其他方面需要针对特定疾病进行分析，并逐步建立相关的临床评估原则进行调查。

4 总结

数字疗法在心理健康领域的应用显得尤为重要，不仅可以做到未雨绸缪，及早发现隐藏的心理问题，还可以跨时空地、长期地加强患者的自我管理以及对患者进行干预治疗。全球 DTx 和医疗保健市场预计将从 2020 年的 58 亿美元增长到 2025 年的 563 亿美元，CAGR 约为 57.4%；2025 年使用人数将增长到 14 亿。正如大家都认可的药物治疗疾病，未来数字疗法治疗干预心理健康疾病也将被逐渐认可。DTx 的快速迭代性质可能会带来监管挑战，需要科学、行业和监管机构共同努力探索更广泛的监管模式。因此，政府机构、医疗机构、研究机构和相关公司应特别关注数字疗法在心理健康领域的新技术，抓住当前数字疗法在心理健康领域的发展机遇，努力推动中国数字疗法在心理健康行业的进一步发展。

PO-004

数字医疗在肿瘤管理中的应用与展望

陈珉惺 朱青 陈桑桑
上海市卫生和健康发展研究中心

恶性肿瘤成为威胁中国人群健康的重大公共卫生问题并上升为国家战略，虽然早期筛查和创新治疗手段均延长了肿瘤患者生存期望，但是肿瘤患者的生存状况与发达国家差距明显。自“健康中国 2030”发布以来，我国始终把保障人民健康放在优先发展的战略位置，要求建立以人民健康为核心，全方位、全周期保障人民生活健康的大健康管理体系，服务也从疾病治疗延伸至疾病关怀。特别是肿瘤管理生态链中的院外环节，由于多数患者完成治疗后便返回社区与家庭，但尚未支持体系和随访体系，医生难以开展患者的长期随访和管理，患者也很难与原来的医院、主治医生保持长期联系。究其原因在于，随着疾病谱及治疗模式的变化，由医疗机构主导的传统院外管理方式呈现出多方面的局限性，缺乏精准化、个体化的跟踪，导致后续随访效果未如期所愿。在肿瘤慢病化趋势下现有卫生资源配置和管理模式已不适应当前患者需求，肿瘤患者的管理模式亟需转变。

智能管理的核心技术得益于“数字疗法”的兴起与发展，该概念于 2015 年第一次提出。数字医疗(Digital Medicine)是一种基于软件程序的治疗和管理手段，为患者提供循证治疗干预以预防、管理或治疗疾病。数字疗法作为一种新的医疗方法或数字健康解决方案，可贯穿于肿瘤患者全周期康复的各个阶段，集数据收集、记录、诊疗、监测、指导、干预为一体的智能化产品，可为患者提供循证治疗和疾病管理服务。作为传统治疗手段的补充和优化，数字化赋能对患者、医疗服务提供方和医疗行业均具有重大价值。该手段可院外治疗和康复管理作为肿瘤管理体系中必要的延续和补充，而如何进行数字化赋能日益受到行业和社会的关注。目前数字医疗如何应用于肿瘤筛查、预防及诊疗中，未来的发展趋势如何，目前尚无明确答案。本文对数字医疗的政策要求、用于临床实践时的优势及在肿瘤不同临床场景中的用途进行探讨，并对其未来的发展进行展望。

1 数字赋能政策战略引导与要求

2021 年 3 月《国民经济和社会发展第十四个五年规划和 2035 远景目标纲要》提出，迎接数字时代激活数据要素潜能，以数字化转型整体驱动生产方式、生活方式和治理方式变革、聚焦教育、医疗、养老等重点领域。在国家战略规划下，各地积极响应并向下纵向推出指导意见供执行。如上海在 2021 年 1 月发布《关于全面推进上海城市数字化转型的意见》，要求推动生活数字化转型，提高城市生活品质。针对卫生健康领域提出“坚持人本导向，强化应用牵引、数据赋能、以数字蝶

变引发城市蝶变，形成《上海市“便捷就医服务”数字化转型工作方案》，纳入了智能诊后管理应用场景，探索慢性病患者健康服务延伸到家。浙江在 2021 年省卫生健康委召开卫生健康数字化改革工作专班会，推进“1314”卫生健康数字化改革。2021 年 6 月发布《浙江省数字经济发展“十四五”规划》支持应用数字技术提升医疗质量，支持开展“互联网+”服务，从预约诊疗制度、智慧结算、检查检验服务、临床路径管理等方面优化就医流程。深圳 2022 年的《深圳市卫生健康事业发展“十四五”规划》提出加强卫生健康信息化统筹管理，全面实现卫生健康信息系统互联互通，强化卫生健康行业基础管理平台建设等。针对平台建设，该规划要求全面完成 12361 工程，探索居民健康管理基础平积极探索将院内服务延申至院外。

可见数字化已成为经济社会发展的核心驱动力，数字化转型在卫生健康领域得到了政策的积极鼓励，推动了医疗信息的共享和互通、远程医疗和健康监测、健康数据隐私保护以及创新支持和资金投入等方面的发展。这些政策措施有助于实现医疗服务的智能化、高效化和人性化，提升医疗服务的效率和质量、改善患者体验，并促进医疗健康产业的发展。而肿瘤病人作为高风险的特殊群体，其诊后管理的重要性和严重性不言而喻。

2 全生命周期的肿瘤数字医疗管理

肿瘤的全生命周期管理链包括“预防-筛查-诊断-治疗-管理-临终”，数字医疗理念的革新，将数字赋能于传统肿瘤管理，打破时间、空间、人力等多维资源制约，提高肿瘤管理的可及性、精准性和扩展性，帮助构建患者管理闭环。利用现代互联网信息技术建立起联结基层和大医院的远程医疗协作网。这种“信息化+肿瘤”的发展新模式，将充分发挥各级医院的不同作用，通过需求分级搭建肿瘤患者精准服务，智能化帮助提供精准医疗服务。此外数字肿瘤管理所收集和产生的大数据还可提高科学研究与优化临床质控。

表 1 肿瘤全生命周期的数字医疗管理场景分布

肿瘤生命周期	管理应用场景	具体应用场景
肿瘤预防	健康理念、行为与生活方式促进与改善	生命体征监测、肿瘤高危风险因素识别软件等
肿瘤筛查	肿瘤标志物检查、普查与个性化筛查等	肿瘤风险分级筛查系统等
肿瘤诊断	远程诊断、多学科联合诊断（MDT）、AI 诊断等	互联网医疗、远程诊断系统、人工智能病理读片系统、MDT 系统等
肿瘤治疗	手术、化放疗、靶向药、临床试验管理等	手术导航系统、化放疗数字疗法、合理用药管理系统、临床试验管理系统（精准招募与试验登记管理）等
肿瘤管理	患者管理、健康跟踪随访、健康教育、大数据记录等	患者自我管理软件、随访提醒软件、用药记录提醒软件、疼痛评估与干预软件、可穿戴设备记录、复诊检查检验系统、转诊平台等
临终关怀	患者及家庭支持、健康教育	患者支持系统等

2.1 国外发展现状

国际上已经形成了成熟的肿瘤患者照护标准和居家管理模式，并且全生命周期地展开智能管理。肿瘤患者居家管理模式注重整合多学科团队和医疗资源，通过远程监控、远程护理和远程咨询等手段，实现对患者的全程关怀和管理。随着智能技术的发展，越来越多的国际机构开始布局智能管理，结合大数据、人工智能和物联网等技术，提供更加精准、个性化的患者管理方案，并通过智能化平台和应用程序，实现患者信息的集中管理、实时监测和智能分析，为肿瘤患者居家管理带来更高效和便捷的解决方案。

在筛查方面，德国医疗应用程序开发公司 Ada Health GmbH 开发的智能健康跟踪应用 Ada Health，可以帮助用户快速评估肿瘤风险，并提供有关癌症预防、早期诊断和治疗方案的信息；

SkinVision 是一款为皮肤癌筛查而专门开发的应用程序，旨在通过 AI 技术来识别皮肤上的异常病变，用户可以通过拍摄照片上传到应用程序中，该应用程序可快速识别和评估皮肤病变是否需要更进一步的诊断和治疗。在检查检验方面，IBM 公司的 Watson 超级计算机就开发了肿瘤智能诊断系统，通过对大量肿瘤患者的临床数据和影像数据进行分析，帮助医生快速准确地诊断肿瘤类型和病情，可以为医生提供更准确的诊断建议；Paige 公司的 Paige Prostate 软件可以识别前列腺病理切片图像上的癌细胞，协助医生读片，提高检验的速度和准确率；基于人工智能技术的临床决策支持系统（CDSS）通过整合不同的病历、文献和临床研究数据，对药物疗效、产品可及性、不良反应、患者财务状况和医疗保险类型进行评估，然后提供个性化建议，帮助临床医生优化治疗方案；美国医疗保健公司 Epic Systems 开发的应用 MyChart，可以帮助患者查看自己的检查和检验报告，以及管理自己的医学信息和预约等，患者还可以与医生进行在线咨询和交流。在康复管理方面，美国全国癌症学会开发的应用 OncoLife Survivorship Care Plan，旨在帮助癌症康复者管理自己的健康状况，它提供了定制化的康复计划、长期监测、复诊提醒和生活方式建议等，以帮助患者过上更健康和活跃的生活；Belong Life 是一款面向癌症康复者的社交支持应用，用户可以加入不同类型的癌症社区，与其他患者分享经验、获得支持和获取专业知识。该应用还提供了康复指南、症状跟踪工具和活动提醒等功能。国际上的肿瘤数字化产品众多，定位各有不同，但都有一个共同的目标，提高医疗服务的效率和质量，从而提高癌症患者生活质量。

国外许多肿瘤患者信息平台和软件应用已经被开发出来，以提供全面的患者管理和监护支持，这些信息平台和软件应用通过促进医生和患者之间的信息共享、提供实时监测和远程护理等功能，增强了医疗团队的协作能力，并帮助患者更好地管理自己的健康。但在在肿瘤数字健康领域中，由于早期诊断和个性化治疗是最具价值的环节，所以目前国外软件开发的热点和成熟度仍集中在肿瘤生命链的前中端，中后端的患者管理也在探索之中。

2.2 国内发展现状

我国的数字化转型起步晚于国外，在对国家药品监督管理局中医疗器械获批现状的查询发现，数字化产品尚不丰富。以“数字医疗”为关键词（数字健康、数字疗法均未查询到）查询自 2014 年有产品获批至今有 37 款产品，主要未手术导航系统、远程会诊系统、具体疾病的临床管理软件等。以“远程医疗”为关键词查询，自 2015 年有 17 款产品获批，主要为会诊软件。以“辅助管理”为关键词查询自 2018 年仅 3 款产品，未疾病后期患者管理，初步有医患交互与随访干预的功能。以“康复管理”为关键词查询自 2017 年有 5 款产品，聚焦术后康复训练软件。聚焦于患者管理、医患交互、肿瘤管理领域均尚未有获批的成熟产品。

表 2 数字化产品获批情况

查询词	获批数量	年度分布获批数量	代表产品
康复管理	5	2017（1）、2022（2）、2023（2）	术后康复管理软件（湘械注准20222211179）、肺功能训练及康复管理软件（苏械注准20232211057）
数字医疗	37	2014（1）、2015（3）、2016（1）、2017（2）、2018（3）、2019（5）、2020（5）、2021（2）、2022（7）、2023（8）	失眠认知行为疗法临床管理软件（湘械注准 20232210408）、神经外科手术导航系统（国械注准20153010492）、个性化三维打印人体颅内肿瘤手术规划模型（粤械注准 20232020635）、远程会诊系统（辽械注准20192210032）、
辅助管理	3	2018（1）、2019（1）、2022（1）	远程血糖辅助管理软件（湘械注准 20222211152）、乙肝母婴阻断辅助管理软件（浙械注准20192210633）

远程医疗	17	2015（1）、2016（2）、2018（1）、2019（2）、2020（4）、2021（2）、2022（4）、2023（1）	远程医疗会诊系统软件（京械注准 20162210782）、远程医疗诊断及信息管理系统软件（辽械注准 20202210005）

数据来源：2023 年 9 月查询国家药品监督管理局(nmpa.gov.cn)-境内医疗器械（注册）数据查询整理

虽然在国家大健康战略的引导下，国内针对癌症诊疗也提出了五个阶段的全生命周期管理方法，包括预防、筛选检查、诊断、治疗和康复。目前医疗机构更关注诊断和治疗，疾控部门关注于预防和筛选检查，患者出院后的康复期管理几乎成为真空地带。

目前传统的肿瘤患者出院后管理形式通常以电话随访和家庭访视为主，目前主要由各级疾控指导由社区家庭医生执行，但目前普通存在“只问生死没有服务”的缺陷。而数字赋能的肿瘤智慧管理，通过数字疗法（数字化手段+远程康复）指导对患者直接进行随访、管理和干预可提高患者的管理依从性和生存质量。现有的数字化 肿瘤康复是指远距离提供康复服务，包括远程康复评估、远程康复指导、远程康复监控等。在服务内容方面可包括患者精神、营养、生活质量、躯体功能、疼痛、用药、后续治疗等情况评估，同时通过数字疗法及其设备，给予患者从“医院-社区-家庭”的管理和干预方案。

当前国内肿瘤领域的智能应用仅局限于信息查询，对于肿瘤筛查和辅助诊断，尚未实现智能化。如癌症全球快讯，可帮助每日肿瘤领域全球前沿资讯速读，了解癌症的最新研究进展；就诊助理，帮助搜索癌症疾病，根据相关科室的诊疗内容，对比找到更适合的医院、科室，并可以直接查看医院提供的肿瘤诊疗服务；灯塔肿瘤医生患者端小程序，可以免费解读检查报告，免费咨询医生，建立病历档案整理报告，找临床项目，准备化疗、放疗期间要用的营养补充剂、缓解副作用的药物，咨询营养师；用药助手 APP，可以查药物说明书、不同剂量分型的用法。然而目前 APP 主要聚焦医患交流，帮助患者提供病情咨询、求医问诊，而站在患者角度的全程系统管理尚处空白。可见，肿瘤患者出院后管理的研究有限，目前仅停留在单个医院或科室层面，不具有推广应用的价值。由于我国医疗资源的有限性以及分配不均，造成医院与社区的患者管理割据,尚未形成良性的互动。目前国内已有的“互联网+护理服务”仅停留在利用互联网实现线上预约和咨询，并未将互联网技术融入患者护理的全流程。不论现有研究还是实践中，均缺乏大健康全系统的视角，基于患者需求的智慧管理模式。

3 政策建议

3.1 顶层规划加速肿瘤管理数字化转型发展

抓好国家智慧医疗成果转化应用，建立覆盖肿瘤患者就医全流程管理信息系统，着力实现医疗业务与健康大数据畅通共享，利用智慧医院的信息平台，建立患者诊疗信息、检验信息和影像学信息的一体化数据库，通过数据标准化和数据融合加速数据挖掘和成果产出。建立院间合作机制，实现患者数据的共享和整合。在国家数字化利好政策机遇下，加大激励、科研人员积极性大幅上升的背景下，积极搭建医学科技成果转化平台，加强与高校、研究院和企业协作交流，实现临床资源优势、学术研究优势与市场信息优势的有效融合，提高智能管理应用所需的技术。同时健全医疗保障制度，在延续护理中特别是患者出院后，从医护人员角度看如果未建立“互联网＋”医疗服务补偿制度，其劳动价值就无法得到回报与承认，会极大影响其积极性；患者方面，因医疗保障制度不完善，特别是当前网络服务费用无法计入医保，极易降低患者选择互联网健康服务的意愿。由此，相关政府部门应加强监督管理，健全网络安全相关法律法规，并可降低网络信息安全风险，同时还应建立健全医疗保障制度，制定统一合理的收费标准，最大限度地保障医患双方的利益。

3.2 多部门合作提高肿瘤院外智慧管理服务水平

数字医疗技术和管理在肿瘤领域，特别是院外环节的发展还处于起步阶段，技术的普及需要政府、医院、企业、社会等多方通力合作，共同助力数字化变革的向前发展。政府端应构建全肿瘤生

命周期的服务链，建立肿瘤院外数字康复医疗平台和管理体系，明确各级各类部门的职能与分工形成闭关管理链，在此基础上进行数字赋能加大对数字医疗的支持力度、制定数字医疗的相关行业标准、健全隐私保护相关法律法规。医院端可应用及推广数字医疗技术与康复知识普及，增加患者接受度，加强与疾控中心、社区卫生服务中心的衔接合作，对相关工作人员和患者及家属的数字化培训，提升其服务意识和服务水平。社会端营造数字健康的良好氛围，通过肿瘤院外数字管理平台实现与企业、医院、 医保、社区等相关部门的联动，加强社区、家庭医生、志愿者、患者组织等资源链接者角色的功能，分析肿瘤患者的院外需求提升服务水平。

3.3 肿瘤院外智慧产品研发提高数据管理与优化

由于肿瘤领域的复杂性，特别是院外患者管理亟需要极高的专业性也需要多医疗机构间的协作，导致数字化产品的研发滞后。对于这一特殊领域，在产品开发尤其需要五大方面的关注。其一，技术的精确性。智能管理应用所依赖的技术，如传感器、人工智能算法等需要具备较高的可靠性和准确性，避免数据收集不准确、算法误判等问题，这可能影响对患者病情的评估和管理，特别是肿瘤患者用药安全、生命体征监测等专业度与阈值判断的精准性。其二，应用的便捷性。虽然智能设备和应用在逐渐普及，但仍有一部分患者缺乏使用这些技术的能力或意愿，特别是肿瘤患者以老年人居多，导致了数字鸿沟的存在，阻碍了居家智慧管理的推广与实施。因此数字医疗产品应以患者为中心，基于受众实际情况进行开发，通过多渠道了解患者出院后的真实需求，特别是基于患者的风险评估（肿瘤凶险程度、复发转移情况、治疗手段等）平衡患者共性及不同癌种的个性化管理要求，开发不同交互形式适用不同患者管理的个性化居家智慧管理方案，包括个性化的健康指导、药物管理和营养建议等。其三，信息的互通性。目前的智能管理应用往往由不同的医疗机构、企业或开发者独立开发，导致应用之间存在互通性问题，以及患者数据的孤立和割裂。为了提高智能管理的效果，需要建立统一的数据标准和互通平台，实现患者数据的共享和整合。肿瘤患者智能管理不同于一般疾病或慢性病，更依赖于多医疗机构间协作，特别是患者发生用药副反应、出现紧急情况、肿瘤出现复发、共病或转移等复杂情况时，需三级专科医院的专业评估和转诊。其四，数据管理与算法优化。尽管人工智能已在肿瘤检测和诊断方面有了显著进步，但如何向患者提供个性化数字服务仍面临着诸多挑战，仍有很多数字医疗产品并未显示出显著优势。部分原因是对肿瘤的认识不够深入，另一重要原因是当前算法尚无法很好地适应真实世界中生成的多模态、海量维度的数字医疗数据，导致 AI 性能常常达不到预期，也即所谓的“维度灾难”。其五，数据安全与隐私保护。居家智慧管理需要收集和处理大量的个人健康数据，包括病史、生理指标、药物记录等。因此，数据安全和隐私保护问题变得尤为关键。确保数据的安全传输、存储和使用，并采取措施防止数据泄露和滥用。很多数字产品没有纳入基于证据的功能，亟需建立数据安全、隐私和数字道德方面的强有力的法规，并在通过建立基于云端的信息系统身份验证及隐私保护协议等手段，保障患者个人健康信息隐私安全。

PO-005

医院网络安全现状及建设内容

吴天赐
新乡医学院第一附属医院

医疗行业网络安全是我国网络安全的重要组成部分，受到国家高度重视。随着医院业务规模的扩大，数字化转型已经成为推动医院减本增效、提高效益的有效措施，云计算、大数据、物联网等新兴的技术的应用，在提供便利的同时，也引入新的安全风险点。网络安全风险的来源更加多样化，传统的以满足等保为主的合规型、被动型安全防护措施已无法满足当前业务环境下的信息安全保障要求。必须采用与医院业务发展相适应的新的安全保障技术体系，以提高重要业务系统安全生产和重要业务数据安全保护水平为目标，不断提高医院关键信息基础设施的网络安全保障能力，夯实网络安全保障的基础。

一、安全现状

医疗行业网络安全建设存在的缺陷和不足，可归类为技术、管理和人员能力三个层面的安全风险。

1、管理现状分析

多数网络安全事件已经暴露出医院应急响应体系部分缺失的状况。虽然部分业务系统已建立应急预案，但仍然存在其它关键预案缺失的现象，同时没有完全落实执行应急预案及应急演练计划。

IT 运维管理制度和安全管理制度已建立，也建立了专人专岗责任制，但对管理的执行落实不完善，尤其是在运维管理流程层面存在不同程度的缺失，无法对实时现状全程掌握。

2、网络威胁实时监测缺失

目前多数医院已根据等级保护要求部署的诸多网络安全设备，但仍然缺乏对网络内攻击、病毒、木马和蠕虫等安全风险的实时监测能力。单靠杀毒软件进行防护仍然存在潜在安全风险，如 PC 终端存在可被远程攻击、植入后门木马、或以此终端为跳板对内部网络实施攻击等多个潜在安全风险。

3、人员能力现状分析

信息科运维人员普遍存在网络安全知识不足、缺少安全技能、对安全设备、平台中展示的安全事件不理解，导致对事件可造成的后果敏感性不高的情况，需要定期对运维人员进行专业、定向的安全技能培训。

二、建设内容

针对上述网络安全现状的梳理，信息科应按照整体规划、分步实施的建设步骤，完善安全管理制度建设、部署先进网络威胁监测平台、内网终端安全建设、服务器区安全建设、数据安全建设等方面对全院网络安全进行提升、加固。同时根据医院现状，通过与专业安全服务厂商合作，建立及时有效的应急响应机制，保障医院业务系统的稳定持续运行。

1、管理层面建设

根据医院实际情况制定并落实执行应急响应体系的完善建设和应急演练预案的制定执行，使信息安全应急响应体系达到可适用、可落地、可执行的状态，在安全事件发生后通过对应的应急响应流程，完成对安全事件的及时有效处置。在应急预案、操作管理、变更管理和定期演练等方面建立起完整的运维管理流程，并贯彻执行，方能达到高效运维，确保医院业务的连续性和稳定性。

2、技术层面建设

通过部署终端安全检测软件，满足防病毒功能、入侵防御功能、防火墙隔离功能、数据信息采集上报、安全事件的一键处置等功能需求，利用威胁情报识别病毒信息，缩短威胁从发现到处置的时间，有效降低业务损失的影响，提升终端的安全防护能力。

部署服务器安全响应系统，对服务器进行东西向的网络隔离，并与网络威胁监测平台进行联动，实现对病毒和攻击进行自动化处理查杀。部署机房环境监测系统，实时监测机房内的温湿度变化、电流电压和 UPS 电池等的运行参数。

为提升业务性能和确保数据的安全性，部署容灾备份系统，对全业务、全系统备份和容灾，避免因物理故障或病毒、勒索攻击造成的数据丢失。部署对网络设备和服务器、存储等硬件的监测系统，能够及时发现故障和损坏的部件。

在网络关键节点部分部署网络流量探针，第一时间识别安全风险，并上报到网络威胁监测平台，对网络流量进行监测及清洗。运维人员可通过网络威胁监测平台实时监测到全网中发生的攻击事件、病毒、木马和蠕虫类的安全事件等，可做到“及时发现、及时上报、及时处置、及时改进”的完整闭环管理，降低甚至完全避免安全事件造成的损失。

3、人员层面建设

随着互联网的普及，国内外网络环境的复杂性，习总书记指出没有网络安全就没有国家安全，并多次在不同场所就网络安全问题给出重要指示。目前各医院几乎没有专业的网络安全人员，急需引进技术人才以确保医院的网络安全。对信息科运维人员进行至少一年两次的安全技能培训，对全院人员进行至少一年一次的安全意识培训计划。培训内容应包括但不限于以下内容：

1）普及开展信息安全意识教育，及时发布和宣讲信息网络突发事件应急预案；

2）普及信息安全应急处置的基本知识，提高全体人员防范意识和应对能力；

3）建立应急预案定期演练制度，不断完善应急预案，提高应急处置能力；

4）制定模拟演练和实战演练计划，提高人员参与度和应对能力。

4、安全服务外包

为了提升网络安全保障能力，除了部署成熟的安全产品，仍需要专业的安全服务的支持。以各医院目前的技术团队力量几乎无法独立完成安全服务建设，因此仍需要与第三方安全服务厂商合作达到构建整体网络安全保障体系的目的。通过专业、成熟的安全服务团队提供定制化的优质安全服务，不仅可以为医院建立完整的标准化安全管理体系和安全保障体系，还可以提升安全防护能力和应对处置效率，为核心业务安全运营的可持续发展提供强大助力。

总之，网络安全规划建设要坚持高标准要求、高目标定位、高水准建设理念，以全网“标准化”安全管控思想为指引，落实安全管控、明确安全责任、实现自主可控、加强安全防护、实现闭环安全管理，提高信息安全综合防御能力。构建以态势感知平台为综合安全管理平台的可知可见、可管可控、闭环运营的纵深安全防御体系，为医院提供全面的信息安全管理支撑、技术支撑、流程支撑和运维支撑，使信息安全保障体系具备广泛互联、全面感知、威胁可见、智能协同的网络安全保障能力。

PO-006

浅谈“互联网+医疗”模式下医学信息安全防护与隐私保护

郑健 贾晋
哈尔滨医科大学附属第一医院

1 引言

随着互联网的跃进式发展，各种先进医疗信息技术应运而生。“互联网+医疗”是病症通过数据融合后的高级医疗服务形式，是深入融合互联网资源和医疗行业的超级医疗行为，代表了医疗行业新发展方向。“互联网+医疗”以互联网为载体，以信息技术为手段，以医疗数据为基础，为人们提供远程会诊、线上诊疗、电子处方和健康宣教等多种形式的线上医疗服务。

“互联网+医疗”在医疗行业的广泛应用中通过各渠道获取患者个人基本信息和病历信息等敏感数据，这些医学信息常伴随着业务交互在各医疗系统中流转，可被用于预测疾病、健康监测和遗传学研究等，具有很高的科研价值和实用价值，因此必须确保其安全防护与隐私保护能够得到切实保障[1]。本文将重点讨论“互联网+医疗”模式下医学信息安全防护、隐私保护以及该项工作的未来发展方向。

2 医学信息安全防护与隐私保护

2.1 安全形势和防护挑战

随着“互联网+医疗”信息技术的飞速发展，医院各类信息系统应用得到不断扩充。医院信息化建设的过程是一个收益与风险并存的系统工程，一方面为患者就医诊疗和医院业务管理提供了极大方便，但另一方面数字化的医院建设对互联网的依赖性也逐渐增大，相应地各医疗信息系统所存储的医学信息也越发重要，具有高价值的医学信息成为医疗信息时代发展的重要标志。目前我国面临的信息安全与隐私保护形势十分严峻，安全漏洞、网络病毒、恶意软件和信息泄露等成急速增长趋势[2]。此外，黑客攻击、社交工程和安全意识欠缺给信息安全防护与隐私保护工作带来了严酷挑战。

2.2 等保 2.0 标准下的医学信息安全防护与隐私保护措施

等保测评 2.0 标准是我国国家信息安全等级保护制度中的一项安全评估标准。其内容包括：网络与终端安全、服务器与存储安全、数据安全、应用系统安全、秘钥管理安全、安全事件与应急管理、运维管理安全和人员管理安全等[3]。

以哈尔滨医科大学附属第一医院的医学信息安全防护与隐私保护措施为例展开探讨，我院每年定期对重要信息系统进行三级等保测评，通过测评了解自身信息系统的安全情况，对于测评中发现的问题进行有针对性地修复和改进，不断提升医院信息系统的可靠性和安全性。同时，对于医院其他信息系统的安全防护，我院也采用同样的安全标准和防护措施进行安全可靠的安全防护，措施内容包括制度建设、意识培训、风险评估、安全策略、技术措施和整体规划。

（1）制度建设

为实现医院信息系统安全，保证医学信息的机密性、完整性和可用性，建立和完善信息安全管理制度以及相应的安全体系显得尤为重要。完备的制度建设能够加强医院医学信息安全管理，能够全面提升全院职工信息安全意识和素养，能够进一步规范院内软件工程师日常运维行为，能够明确信息安全防护各环节的职责和权限，能够提高信息系统的安全防护效率，能够最大程度的规避安全风险。

我院以国家各项网络安全法律法规为标准，以医院网络安全管理要求为准则，制定了一系列规章制度，共计 35 项。内容涵盖权限分配管理、密码口令管理、数据调取管理、医疗设备管理、信息安全管理、安全工作协调、安全检查考核、安全应急响应及恢复、安全责任追究、数据安全保护等多方面内容。

（2）意识培训

定期的、高质量的信息安全意识培训，可以帮助医院建立良好的信息安全生态和安全管理体系。使医院员工深刻认识到信息安全防护和隐私保护的重要性，能够认知可能存在的信息安全问题，恪守正确的行为方式，最大程度的提升员工的医学信息安全保护意识并熟练掌握信息安全防护技能以便应对可能发生的安全事件。

我院定期邀请知名信息安全专家和高级信息安全讲师到我院进行信息安全知识讲座，并于每年不定期开展网络安全宣传周活动并逐科发放网络安全宣传手册。授课与手册紧跟国内外最新的安全技术发展与典型的安全热点事件，通过适当的简化与还原，使医院信息安全意识培训工作卓有成效。

（3）风险评估

信息系统风险评估是指通过技术手段对目标系统进行深入安全分析的过程，是针对威胁、脆弱点以及它可能导致的风险大小而评估的，旨在对其中存在的安全威胁进行识别、分析、评估和处理，判断可能遭受攻击的潜在影响程度，采取相应措施和资源应对这些安全威胁[4]。

我院不定期进行系统安全风险评估，具体措施包括：漏洞扫描、渗透测试[5]、代码审计和网络安全设备核查等。通过这些技术手段来挖掘系统的脆弱性和风险点，并及时采取应对措施进行积极整改，整改完毕后继续对目标系统进行复测，如此反复，直至安全威胁消除为止。

（4）安全策略

信息防护安全策略是指在某个安全区域内，采用的全方位安全防御规则，用于构建安全防御体系，以便实施并实现对信息系统的精准管控和安全防御。安全策略主要分为访问控制策略和访问支持策略，是信息安全的基本标准，是对系统安全需求的形式化或非形式化描述，是所有安全防护措施的指导方针。

我院对防火墙系统策略、僵尸网络木马和蠕虫监测与处置系统策略、终端防御威胁系统策略、探海威胁检测系统策略和 Waf 防火墙系统策略进行精细的、严格的配置管理和下发管理，符合安全策略的合法数据流可以通过。

（5）技术措施

信息安全防护的核心在于技术措施的实施，完整的信息安全防护技术措施包括：入侵检测技术、访问控制技术、漏洞扫描技术、身份认证技术、防病毒技术、日志收集与分析技术和系统容灾技术等。

我院定期对各类网络安全设备进行升级，以保证各库始终保持在最新状态。日常工作中采用入侵监测技术对网络安全状况进行监测，发现攻击威胁及时通过防火墙进行阻断；采用访问控制技术对入网行为进行授权管控，未经授权禁止上网；采用漏洞扫描技术对计算机和服务器主机进行检查，发现高危漏洞及时修复；采用身份认证技术对信息系统使用者进行确权，发现密码多次错误及时锁定秘钥；采用防病毒技术对全院杀毒软件终端下发病毒、木马查杀任务，发现病毒、木马及时通过杀毒软件进行隔离并删除；采用日志收集与分析技术，采集、分析和存储各类日志，并对目标日志进行分析研判；采用系统容灾技术对信息系统及数据进行容灾备份，切实保障一院两区模式下的医院信息系统平稳运行。

（6）整体规划

信息安全整体规划是医院信息安全建设的远景、目标、框架、任务和行动路线的建设过程。我院始终坚持安全和发展并重，牢固树立科学网络安全观，切实守住网络安全底线，以安全保发展、以发展促安全，推动信息安全与信息化发展协调一致，统筹提升信息化发展水平和信息安全保障能力，持续加强信息安全防护体系和能力建设。从加强组织领导、健全制度体系、强化人才培养、加强安全宣教、实施安全评估和配备安全设备等规划，确保目标任务落到实处。

3 未来发展方向与建议

随着“三位一体”智慧医院的深入建设，“互联网+医疗”模式被广泛应用，联通医院内外的信息系统逐渐增多，医学信息呈爆炸式增长，使得智慧医院在建设过程中将更加广泛的应用医学信息和信息技术。信息技术创新发展伴随的安全威胁与传统安全问题相互交织，使医学信息安全问题日益复杂隐蔽，面临的网络安全风险不断加大，医学信息安全防护与隐私保护压力陡增[6]。

未来的医学信息安全防护与隐私保护体系应该建立在各种网络安全设备协同联动的基础之上，从全局视角加强安全态势分析和安全设备统一管理。医学信息安全防护与隐私保护未来发展方向与建议：健全数据安全治理体系；加强关键信息设施保护；积极试行隐私计算技术；全面应用国产密码技术：推进网络安全云化服务；探索人工智能安全防御；医学信息分类分级管理。

4 结语

在医学信息安全防护与隐私保护工作中，要细化任务举措，着力推动落实，与时俱进，科学发展。优化医学信息安全防护技术手段，加强组织管理和意识培养。坚守安全“红线”，筑牢安全“底线”，用实际行动将医学信息安全防护和隐私保护工作落到实处。

PO-007

院科两级自助统计方法研究与系统实现

桑磊，林书艺，刘晓君，郑轲心，陈晗*
中国人民解放军总医院海南医院信息科

0、前言

随着国家“新医改”的不断深入，“以药养医”、规模经营等传统发展模式已经很难适用新时期医院发展要求，医院管理必须从“粗放型”向“精细化”转变，医院决策必须从“经验型”向“科学化”发展。以信息系统为手段，根据医院管理需求对医疗效率，医疗质量、医疗效益等指标进行分析，合理配置资源、适时调整政策导向，不断优化流程，科学制定发展战略是实现现代医院精细化管理，科学化决策的有效途径[1]。

目前我院医疗数据统计的途径繁多，有基于 PowerBuilder（PB）开发的统计软件、部分临床系统自带的统计模块、更多的是信息技术人员直接编写 SQL 语句从 HIS 数据库进行查询统计。以上因素容易造成数据统计口径不一致，修改统计算法难、统计效率低，统计难度大，数据混乱等问题。因此，建立一个适合医院各部门通过授权自助查询统计数据的综合性数据统计系统成为必要。

1、方法与目的

目前我院的数据统计问题主要表现在以下几点：

(1) 统计数据格式和统计口径缺乏统一的标准，规范性差；

(2) 统计工具比较分散，各个科室各自保管各科数据；

(3) 现有的统计系统改造难度大，无法适应新的管理需求和统计方式；

(4) 现有的统计系统没有权限设置，影响数据的安全；

(5) 数据报表格式单一，不能适应多维度的数据统计和展现；

(6) 有时需要通过 SQL 从 HIS 数据库查询数据，难度大、专业性强、对数据库不安全。

基于以上问题，将我院 HIS 系统中现有的数据统计功能和长期收集的数据统计需求进行汇总归纳，根据院科两个层级划分，设计出符合各部门统计需求的数据报表。根据用户权限不同，医院管理层可及时掌握医院运营情况，临床科室方便查询本科室医疗数据，药剂、经管、病案、质控等辅诊科室可随时查看科室相关报表。

为方便用户进行数据统计，院科两级自助统计系统应采用 B/S 架构，无需部署客户端，只要通过浏览器登陆系统即可进行数据统计。为保护数据安全，系统应有严格的权限管理功能和审计功能，用户授权后才可以查看相应的报表，数据统计行为以日志的形式保存下来，便于进行数据审计。为适应用户的不断变化的统计需求，系统应该具备快速生成报表的能力。

2、系统的构建与实现

2.1 技术路线

院科两级自助统计查询系统是基于 C#开发语言的开源框架和微软的 RDLC 报表工具的一款 B/S 架构的系统。开源软件具有代码开放、免费获得、开发模式的参与性强等优点，使用开源软件做项目，可以节约成本、缩短软件开发周期、提高自研项目的质量[2]。该开源框架具有完善的用户管理、权限管理、菜单管理、数据备份和审计日志功能，可为不同用户根据角色需求分配不同的报表查看权限。

RDLC 报表是微软开发工具 Visual Studio 中推出的一款报表插件，具有控件丰富、灵活定制、高度可编程性、数据处理能力强、与 office 兼容性好等优点[3]。RDLC 报表文件设计开发完成后需要一个 Web 页面来装载，在开源框架中以新增菜单项的形式将 Web 页面的链接地址保存后，经过授权的用户就可以通过访问 web 页面来进行查询数据的操作。

由于 RDLC 报表开发涉及到 C#编程中数据源和 Web 页面的编码，大多时间都花费在重复性的编码工作上。为使院科两级自助统计系统的报表开发更加简单快捷，本系统开发了与 RDLC 报表相结合的代码生成工具，通过简单配置后即可完成报表开发工作，大大提高了开发效率。

基于 SqlServer 数据库的开源框架提供用户管理、权限管理等基础功能，系统开发只需关注业务层面的开发即可。将 Sql 语句导入代码生成工具生成连接 His、Pacs 等数据库的动态报表数据源，通过报表设计器创建与数据源绑定的报表文件。代码生成工具根据数据源和报表文件自动生成带查询条件的 Web 页面。在系统中创建的用户授权后访问 Web 页面，输入查询条件即可得到需要的数据。用户的访问记录会保存到系统日志，便于后期的行为追溯，同时系统支持应用数据库备份，防止数据丢失。(如图 1)

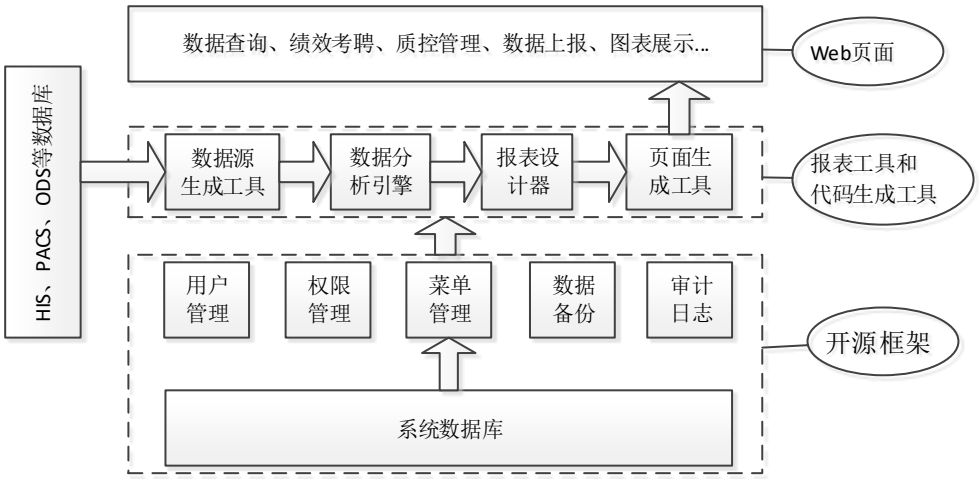


图 1 院科两级自助统计系统架构图

2.2 系统功能分类

院科两级自助统计系统基于我院 HIS 系统已经存在的多个数据统计功能和信息科保留的多年统计需求，设计了全院医疗数据统计、临床科室医疗数据统计、医技辅诊科室数据统计三大类十一个模块共一百三十余个报表（如图 2）。对医院运行过程中的医疗质量、医疗效率、医疗安全、医疗服务、合理用药、医疗经济指标实现了多维度、多粒度、图形化的统计分析，为院级管理部门、临床科室和医技辅诊科室实时掌握医院、科室的运行情况提供了数据支持。

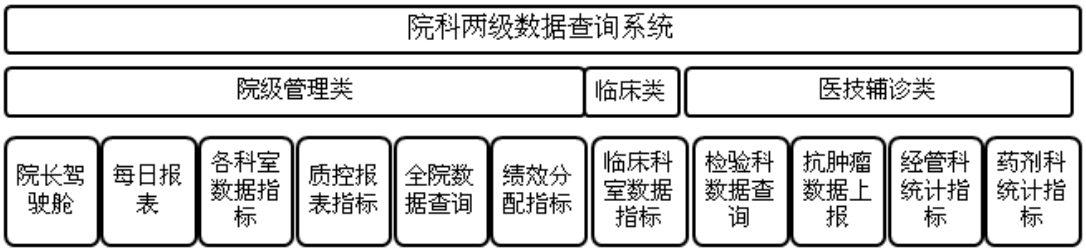
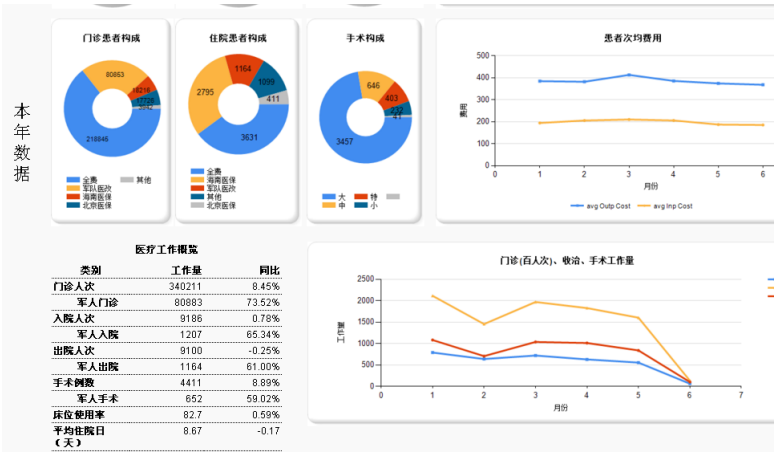


图 2 院科两级自助统计系统功能分类

2.2.1 院长驾驶舱模块

院长驾驶舱模块通过运用 RDLC 报表的图表控件，以图表结合的方式，为院领导提供全局、宏观的指标分析窗口（如图 3）。同时支持大屏显示，便于其在业务收入、工作量、工作效率、合理用药等角度，通过仪表盘，趋势图、表格等形式监控年度、当月、当日的医院运营状况和发展趋势[4]。



类别	工作量	同比
门诊人次	340211	8.45%
军人门诊	80883	73.52%
入院人次	9186	0.78%
军人入院	1207	65.34%
出院人次	9100	-0.25%
军人出院	1164	61.00%
手术例数	4411	8.89%
军人手术	652	59.02%
床位使用率	82.7	0.59%
平均住院日（天）	8.67	-0.17

图 3 院长驾驶舱年度数据

2.2.2 每日报表模块

每日报表模块包含了日报、周报、手机报等常规院长报表和住院管理科每日向上级医院上报的科室收容量、床位使用情况报表。此模块的开发优化了医院数据上报流程，由之前信息技术人员手工统计填报上报改为院领导直接在电脑终端查询，减少了手工统计的工作量，避免了数据上的误差。

2.2.3 各科室医疗数据指标模块

院级管理部门可通过此模块选择科室后查询指定科室的门急诊、收治、手术、会诊、检查、检验等相关工作量，医生出诊接诊情况，科室床位运转效率、平均住院日等绩效考核指标。(查询科室工作量概览报表如图 4)

图 4 查询科室工作量概览报表

2.2.4 质控报表指标模块

质控报表指标模块包含死亡患者统计、非计划重返住院、超长住院、重返 ICU、非计划手术、会诊不及时等涉及医疗质量方面的数据统计。医院以往因缺乏信息系统支撑和科学评审标准，经验性、主观性强，存在病历质量查不全、查不准，老问题难解决，新问题频繁出现等现象[5]。该模块基于特定的质控规则对患者诊疗全过程进行病历检索，为质控部门及时筛选出目标病历，便于下一步的质控干预。

2.2.5 全院医疗数据统计模块

全院医疗数据统计模块包括全院及各科室门急诊、收治、手术、在院患者、门诊排班、平均住院日、床位使用率等相关工作量和工作效率指标。输入统计周期后可统计到全院各科室相关指标，方便了院级管理部门对全院科室的工作效率和工作质量的宏观把控。

2.2.6 绩效分配指标模块

绩效分配指标模块包含经管科每月进行科室绩效分配计算时用到的平均住院日、床位使用率、医技检查工作量、会诊及时率、种植牙数据、院前急救指标、远程会诊、工作日及节假日各科室手术量等相关指标报表,为科室绩效分配提供准确的数据支撑。

2.2.7 临床科室医疗数据指标模块

由于临床科室的大部分医疗数据指标具有一致性，临床科室医疗数据指标模块运用了通用报表模板技术，减少了大量的重复开发工作。在系统创建用户时，用户与其所在科室进行绑定，用户登录系统后采用 Session 机制保存用户科室信息。当用户查询相关报表时，将所在科室代码作为参数传入数据源，从而得到本科室的相关数据。这样既保证了用户只能看到自己科室的数据，又减少了报表开发工作。

临床科室医疗数据指标模块包含科室工作量概览、科室节假日手术、科室工作日手术、医生接诊人次统计、医生出诊次数统计、医生手术量统计、科室检查量统计、科室检验量统计等相关工作量和工作效率指标。科室主任能对本科室及所管理的病区,每周、每月门诊工作量情况和对出院病人的治疗质量及医疗费用、病床的利用率、病区主管医师的工作质量和效率等进行查询,为科室决

策提供数据依据[5]。

2.2.8 检验科数据统计模块

检验科数据统计模块包含了多条件联合查询检验数据的报表，TAT 指标分析报表、尿复检报表、标本统计报表。此模块弥补了检验系统中缺失的相关统计查询功能。

2.2.9 抗肿瘤数据上报

抗肿瘤数据上报模块是根据《开展全国抗肿瘤药物临床应用监测工作》的通知，为药剂科开发的涉及肿瘤患者数据的 18 个数据上报报表。由于有些报表涉及的数据量达百万条，RDLC 自带的导出功能已经无法满足导出需求，专门开发了下载生成 CSV 数据文本的模块，解决了大数据量无法导出的问题，大大减轻了药剂科的数据上报工作量。

2.2.10 经管科统计指标

经管科统计指标模块包含了出院结算报表、结算未清报表、军地患者住院费用统计、收费员工作日志统计、微信挂号流水查询等，弥补了收费软件缺失的相关查询统计功能。

2.2.11 药剂科指标

医疗费的药费比例是卫生行政部门和医保管理部门考核医院的一个重要指标，控制药费比例是促进医院合理用药的重要举措[6]。合理用药是以当代系统化、综合性医药学和管理学知识来指导用药，使药物治疗达到安全、有效、经济的基本要求[7]。药剂科指标模块包含了药费比指标、用药强度指标、抗菌药使用指标、用药监测指标、药品库存查询等相关 42 个报表。此模块的应用为药剂科在全院的合理用药监测、控制药费比、药品效期管理等方面提供了及时的数据支撑。

3 应用效果

院科两级自助统计系统于 2017 年上线，在我院推广使用中根据医院管理、科室统计新需求不断修改完善。通过对医疗质量、医疗效率、医疗安全、医疗服务、合理用药、医疗经济指标等医疗数质量指标的统计分析，使院科两级管理者和医技辅诊科室能够及时发现并纠正医院运行过程中出现的问题，减轻科室手工统计数据的工作量，合理配置资源，适时调整政策导向，不断优化就诊流程，通过信息化手段向医院管理精细化、决策科学化这一目标逐步迈进。

4 结语

院科两级自助系统可以快捷的生成数据统计报表，但也不能随意在系统增加报表。在数据统计工作中把握“常规性统计纳入系统，临时性统计写 Sql”的原则，这样既保证了统计工作的效率和规范性，又不会使系统的报表混乱臃肿。方便用户的时候同时减轻了信息部门的日常统计工作。目前院科两级自助统计系统仅接入了 HIS、PASC、ODS 数据库。下一步将实现与人力资源管理系统、财务系统、ERP 系统等系统的深度融合，将统计粒度逐步推进至每个科室。同时不断优化查询统计方式，确保数据统计的准确性，提高查询效率，为医院管理决策提供有效的数据支撑。

PO-008

泛癌智能应用平台设计

刘彦霞 渠鸿竹 方向东
中国科学院北京基因组研究所(国家生物信息中心)

1. 引言

2016年，我国癌症新发病例约406.4万例，新发癌症死亡病例约241.35万人，预计未来十年将继续增加，仍然是我国的重大公共卫生问题[1]。我国的癌症负担很重[2]，国务院出台了一系列卫生政策优先推进癌症防治。为全面贯彻落实党中央对健康中国行动的决策部署，国家卫生健康委员会等10部门联合制定《健康中国行动——癌症防治实施方案（2019—2022年）》，提出要建设权威科普平台，改善健康教育；健全肿瘤登记制度，促进信息资源共享利用；打破基础研究、临床医学和公共卫生之间屏障，加快基础前沿研究成果在临床中的具体应用[3]。本文结合泛癌研究现状、问题及需求，构建与泛癌研究发展相适应的智能应用平台设计，促进癌症防治方案有效实施。

2. 泛癌研究现状及问题

到目前为止，多个国际联盟，包括癌症基因组图谱（TCGA）和MSK-IMPACT临床测序队列（MSKCC），对多种肿瘤进行了全面的基因组分析,构建的泛癌图谱提供了导致人类致癌过程的全景图[4-6]。基因组分析和临床病理注释分析为临床提供建议，改进了多种恶性肿瘤的治疗管理[7-9]。跨癌种分析使某一癌症类型的有效疗法可以扩展到具有相似基因组特征的其他癌症类型中[10]。精准肿瘤学数据库（OncoKB）包含肿瘤相关基因突变及对应的96种药物注释信息，每条信息经过临床基因组学注释委员会(Clinical Genomics Annotation Committee, CGAC)的定期审阅与修订，指导靶向治疗和研究性治疗方案。近几年，随着单细胞RNA测序技术（scRNA-seq）的发展，不同癌症的细胞异质性、免疫微环境和发病机制被相继揭示，相应的癌症单细胞表达图谱数据库（CancerSCEM）也被开发出来[11]。同时，蛋白质组学、代谢组学、表观遗传学的发现为癌症防治提供了更多的候选分子。生物大数据时代和人工智能时代的到来，为智能化临床决策系统的构建创造了条件，促进了信息医学的发展。但随着《“健康中国 2030”规划纲要》的发布和信息技术的不断进步，对泛癌研究有了更高要求。目前，泛癌研究平台还存在一些问题。

2.1 平台专业性强，公众不易懂

泛癌研究平台大多针对科研人员和临床医生，专业性强。如TCGA多为基因组的一级分析结果，要得到结论性的结果，需要生信分析人员对一级数据进行再分析。CancerSCEM更是需要有生物学背景，才能对数据库搜索的结果进行合理分析和解读。而随着信息化时代的到来，公众对癌症防治有需求，但绝大部分人没有生物学背景，这就需要平台具有针对公众的科普性版块。

2.1 平台专业性强，公众不易懂

泛癌研究平台大多针对科研人员和临床医生，专业性强。如 TCGA 多为基因组的一级分析结果，要得到结论性的结果，需要生信分析人员对一级数据进行再分析。CancerSCEM 更是需要有生物学背景，才能对数据库搜索的结果进行合理分析和解读。而随着信息化时代的到来，公众对癌症防治有需求，但绝大部分人没有生物学背景，这就需要平台具有针对公众的科普性版块。

2.2 平台数据单一，且功能尚不完善

平台数据单一，即只有一个维度的组学数据，这就使癌症的诊断、预后、监控精确度受限，不能更好地使患者受益。此外，平台的功能模块单薄，分析程度浅显，没有充分利用数据资源。

2.3 各种肿瘤平台独立存在，缺乏整合分析

多维组学的重要性不言而喻，已有研究将一种肿瘤的多维组学整合到一个数据库中进行展示。如 DBGC[12]，包含不同来源的人类胃癌相关转录组学项目、蛋白质组学项目、突变、生物标志物、药物敏感基因、流行病学统计数据以及胃癌病例注释的临床病理信息；SCDb[13]，包含 SNP、甲基化、RNA-seq、miRNA-seq 和临床数据；GliomaDB[14]是一个整合神经胶质瘤组学数据的数据库；MOBCdb[15]是一个整合乳腺癌多组学数据的数据库；以及某三甲医院牵头搭建的肺癌临床大数据管理平台等[16]。这些都是某一肿瘤的综合数据库，虽包含多维组学数据，但跨癌种分析无法实现。

3. 用户需求分析

3.1 一般人员

公众需要了解癌症防治核心信息和知识要点，推进健康生活方式，开展自我健康管理。这需要公众获得的信息是通过专业机构编制发布的，有科学依据的，而不是片面的、有购物诱导的虚假信息。现在公众缺少一个平台获得这些专业的科普资源。此外，公众希望由于挂号难、路途远等原因，不能线下请资深专家诊断的检测报告，可以通过线上上传报告图片的形式及时得到反馈。

3.2 科研人员

癌症相关的科研人员希望自己的研究成果能够被临床医生及时了解，并运用到临床得到反馈，如早诊、预后模型是否对患者的诊治有成效等。此外，科研人员虽然有生物学背景，但没有临床医学背景，对临床上诊治癌症患者真正的痛点不清楚，对发现的候选标志物的应用前景不了解等等。

3.3 医护人员

医生希望有早诊、预后模型，在不对患者进行病理切片的情况下，准确知道患者的病情，同时知道针对该患者基因型有效的治疗药物，对患者进行精准治疗。治疗后，对患者进行监控，根据无创监控数据及预后模型预测患者预后效果，及时发现异常并改变治疗方案，提高患者的生存率及生活质量。

4. 平台设计

4.1 架构设计

4.1.1 总体架构 泛癌智能应用平台总体架构由基础层、数据层、支撑层、应用层、展现层和用户层组成，见图 1。其中，数据层包括提供知识检索的基础资源库，用于辅助临床诊疗的业务资源库，以及智能咨询与一对一交互的服务主题库；支撑层融合基于深度学习的图文识别、文本抽提等多种人工智能技术，构建泛癌智能应用平台，实现基础研究、临床医学和公共卫生一体化；应用层为公众提供癌症防治知识、智能评估、智能反馈咨询等交互功能，同时为科研和医护人员提供模型上传、使用、使用度统计、方案评估、一对一交互等功能服务。

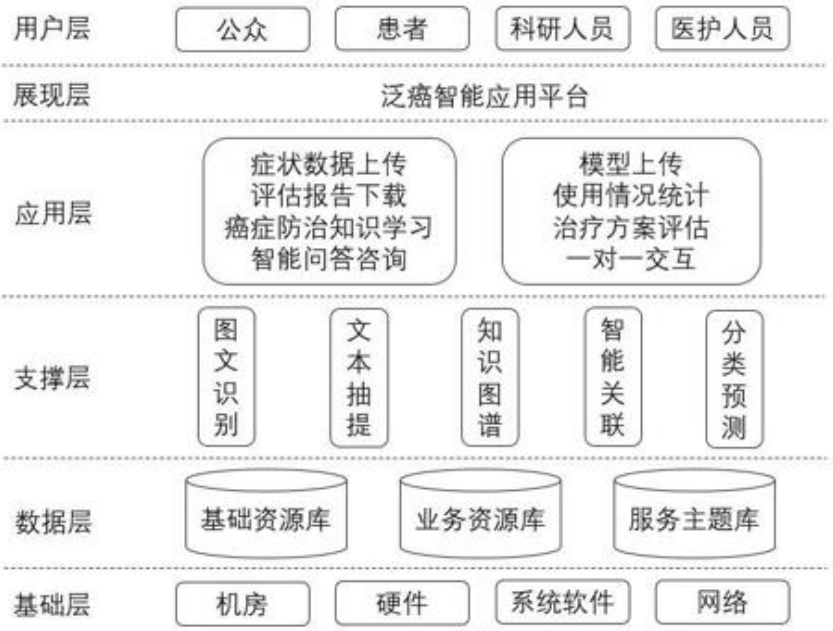


图 1 总体架构

4.1.2 数据流转过程 公众上传症状数据，如化验单和文本式主诉，系统通过深度学习图文识别和文本抽提等技术进行解析以获得结构化数据，之后基于相对应的预测模型进行预测，并反馈给公众。若需要跟医生咨询，则医护人员可从服务器获取该人员的数据，结合自己的临床经验，拟定方案，上传到服务器进行优化，最后反馈回该人员。科研人员可上传模型到服务器供医护人员使用，并可查看服务器上所有模型的使用情况及医护人员反馈意见，进行优化。并且可与医护人员一对一交流模型的使用及临床痛点等，见图 2。

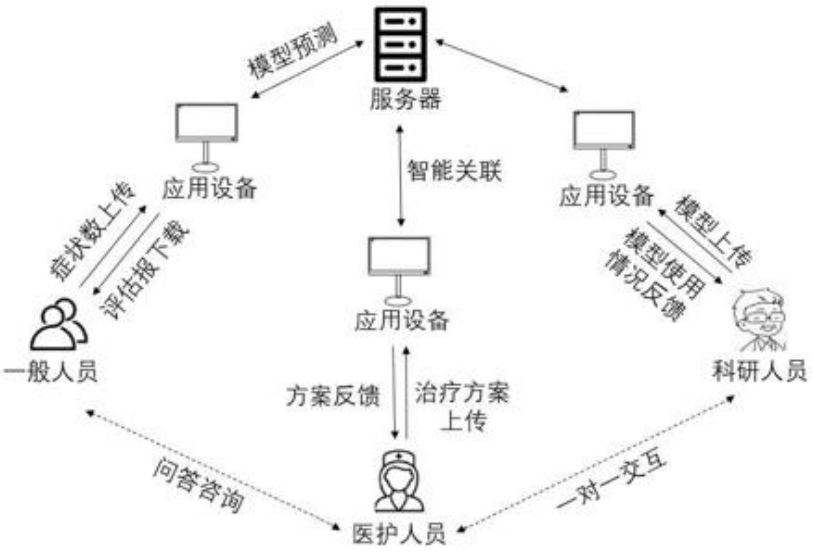


图 2 数据流转示意图

4.2 功能设计

4.2.1 一般人员 公众可通过基于知识库的检索，获得癌症防治知识，也可将症状数据上传，得到反馈的评估报告。如若需要资深专家的诊治建议，还可与医护人员进行有偿的问答咨询。此外，系统基于症状数据、历史咨询、搜索问题记录等数据为该人员个性化推荐癌症防治知识，从而提高该人员的自我护理效能与生命质量。

4.2.2 科研人员 科研人员可上传自己的科研成果-预测模型到平台后台服务器，并可随时查看平台上所有模型的使用情况及医护人员对模型使用的反馈意见。优化后的模型也可随时再次上传替代优化前模型。此外还有与医护人员一对一交互端口，为更符合临床痛点的模型的构建和优化提供了“绿色通道”。

4.2.3 医护人员 医护人员可随时免费使用平台上任一模型，了解最新科研进展，并对治疗方案进行优化。也可对患者的提问进行有偿回答，前提是医护人员需要有资格认证，对自己的回答要承担相应的责任。也可与科研人员进行交互，了解模型的使用规范等。

5. 结语

随着人民生活水平的提高，健康需求尤为突出。而癌症仍是我国的重大公共卫生问题，借助“互联网+”技术连通基础研究、临床医学和公共卫生成为必然趋势。该平台是对癌症防治模式的一种积极探索，其为公众与医护人员搭建沟通

桥梁，消除时间与空间壁垒。人工智能技术的融入极大提高医护人员工作效率，从而节约医疗资源。科研人员端口的开放使科研成果第一时间服务于临床，大大提高癌症患者生存率和生存质量。

PO-009

三级医院评审数据质量监管系统开发体会

蔡磊
平舆县人民医院

三级甲等医院(简称"三甲医院")是依照中国现行的《医院分级管理办法》等规定划分的医疗机构级别，是向所在地区以及周边辐射区域提供最高水平医疗卫生服务和执行高等教育、科研任务的区域性以上医院[1]。2018 年 11 月，国家卫健委印发了《全面提升县级医院综合能力工作方案(2018-2020 年)》明确提出，到 2020 年，500 家县医院和 500 家县中医医院分别达到“三级医院”和“三级中医医院”服务能力要求[2-4]。我院（河南省平舆县人民医院）于 2021 年 1 月 5 日，按照《河南省三级医疗机构设置规划(2020-2025 年)》完成三级综合医院执业登记。“三甲医院”也一直是我们医院坚持不懈追求的目标，为此，我院制定了创建“三甲医院”的目标和实施方案[3, 5-8]。

按照《三级医院评审标准（2020 年版）》（简称《标准》）的要求，评审分为三个部分，第一部分前置要求部分，第二部分医疗服务能力与质量安全监测数据部分，第三部分现场检查部分。其中第二部分医疗服务能力与质量安全监测数据部分在评审综合得分中的权重不低于 60%，可以看到数据指标占有很大比重，各医院尤为重视。作为医院的信息部门我们责无旁贷的要在技术上做出相应的支撑[1, 2, 4, 9-16]。

在《标准》的第二部分，共设有 74 节 240 条监测指标，内容涵盖医院资源配置、质量、安全、服务、绩效等指标监测，DRG 评价、单病种和重点医疗技术质控等日常监测数据。要求该部分的分值不低于全部评分的 60%，而且要求，如果无法提供原始数据的，即视为错误数据。同时在核查准备指引中，《标准》要求数据指标应当包含每个数据的定义、数据源、采集方式、采集时

间范畴，采集结果等因素，并且要数据有可追溯的原始数据。如果单存的靠手工采集或者数据上报的方式来实现采集这些数据，一是增加工作量、二是无法追溯原始数据，三是无法对数据进行再加工，为数据的分析提供支持。因此，需要考虑在技术上如何实现[17-23]。

关于指标体系的建设，有用 BI 或者报表来做展示的，也有把各种指标查询分析分散在各个子系统中的，当然也有手工采集最终数据上报的。无论采取哪种做法，最主要的是有数据，要能在生产的过程中产生数据，正所谓“小河有水大河满”。如果没有生产数据源，再好的软件技术也无法展示出数据。根据这个需求，我院成立了数据指标攻关小组，业务部门牵头，信息部门负责技术支撑。业务部门负责指标解读（数据的定义），分析数据指标的分子分母来源（数据源）如何从生产系统上获取，以及数据指标结果的验证（采集结果），而信息部门负责技术实现，采用定时抽取数据的方式产生每日数据指标结果，并存储到指标库，可通过指标库中数据指标追溯到生产数据源查看异常数据[24-26]。

一、需求分析

根据《标准》中第二部分医疗服务能力与质量安全监测数据中的要求分析，系统应具备以下功能：

1.1、 指标的最终呈现有两种方式：1、数字结果，2、比率结果。所以用分子，分母，结果 3 个字段来表示，既能存放数字类结果，又可以后期运算产生比率类结果。

1.2、 数据要具有可追溯，那么一定要存在可以查看异常数据（不符合分母的数据）的方式（下转数据），同时可以通过对异常数据链接到生产系统中查看原始数据内容（可以通过 CDR 视图等来实现）。

1.3、 监测指标，按照“区间赋分兼顾持续改进”原则给分，数据要能以趋势的方式来量化业务运行的基本状况。

1.4、 医院内存在异构系统的，多源数据需要能支持；

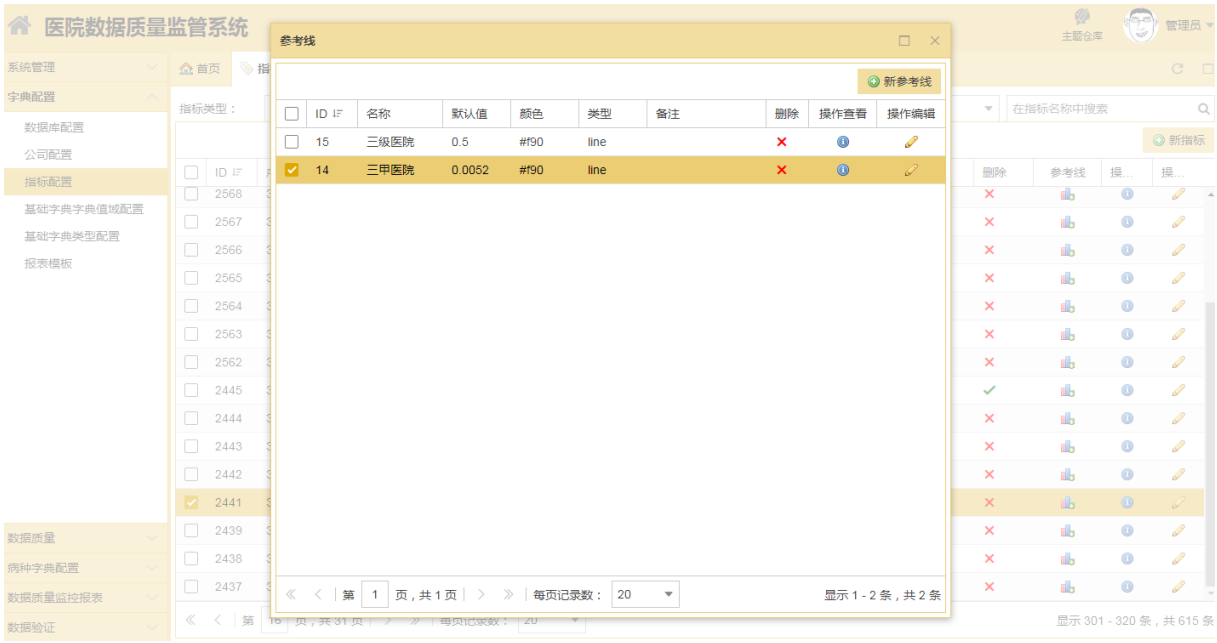
1.5、 数据能被监管、管理部门评价准确性和业务趋势发展好坏的方式；

针对《标准》中的要求，细化需求分析后，决定采用 Asp.net+Quarter 定时调度抽取数据的方式实现[15, 17, 18, 21-23]。

二、效果展示

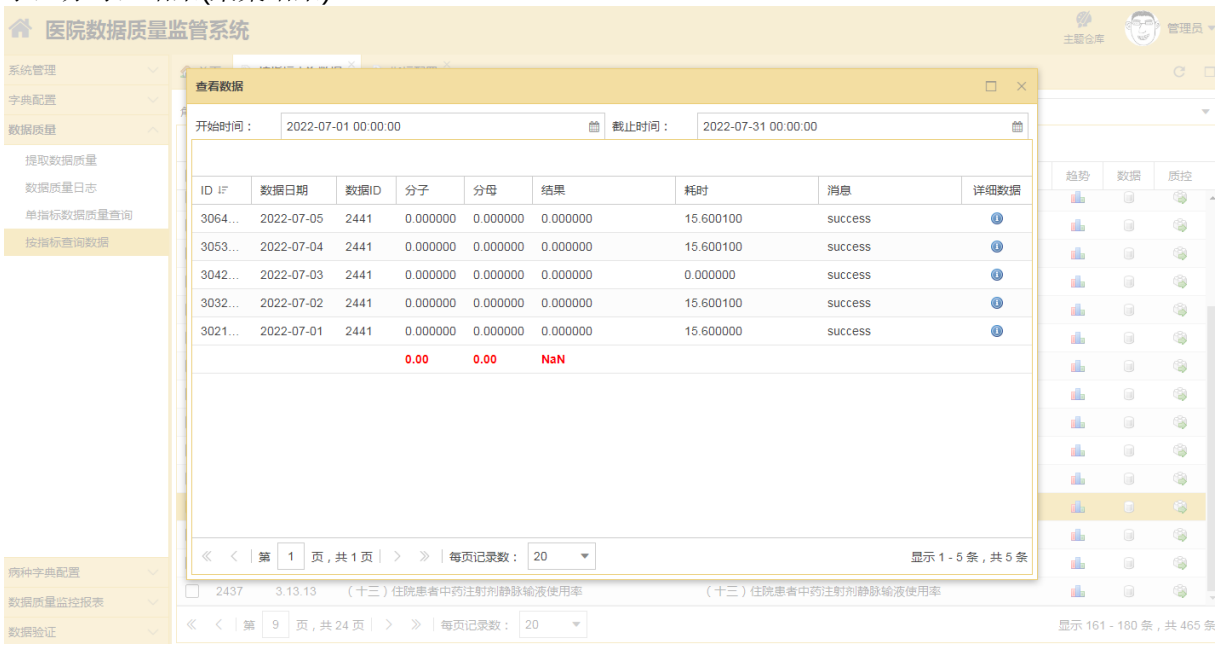
2.1、 系统需要有指标定义（包含每个数据定义、数据源、采集方式），指标名称，描述（每个数据的定义和来源），数据源，接口方式：sql,sqlcalculate（采集方式）,分子 sql，分母 sql,明细 sql 和对应数据源。

（指标配置界面）



(指标参考线（趋势参考线）界面)

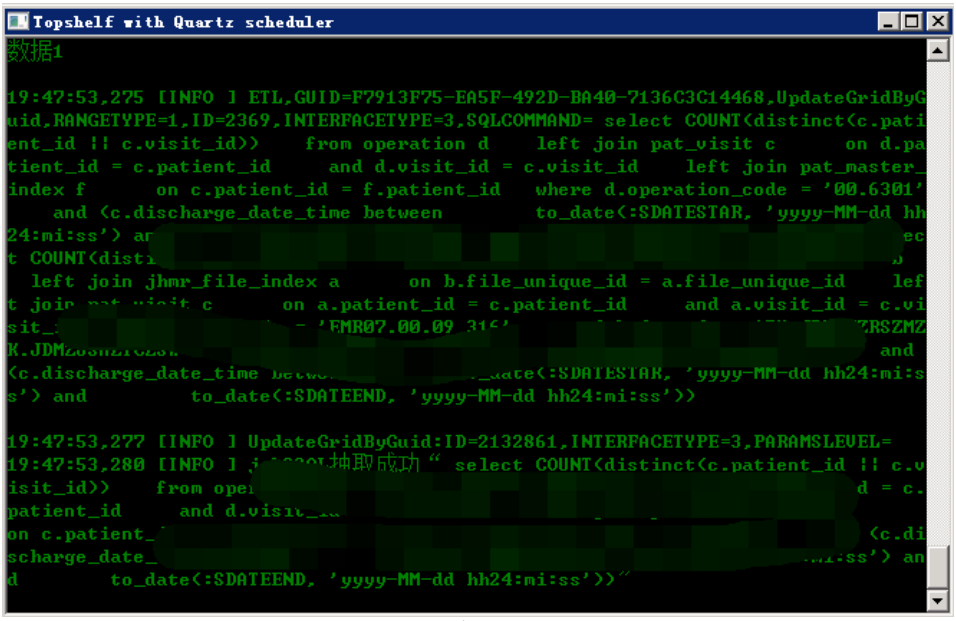
2.2、 指标采集(采集时间范畴，采集结果) 包含采集时间，数据日期（采集时间范畴 ），分子，分母，结果(采集结果)。



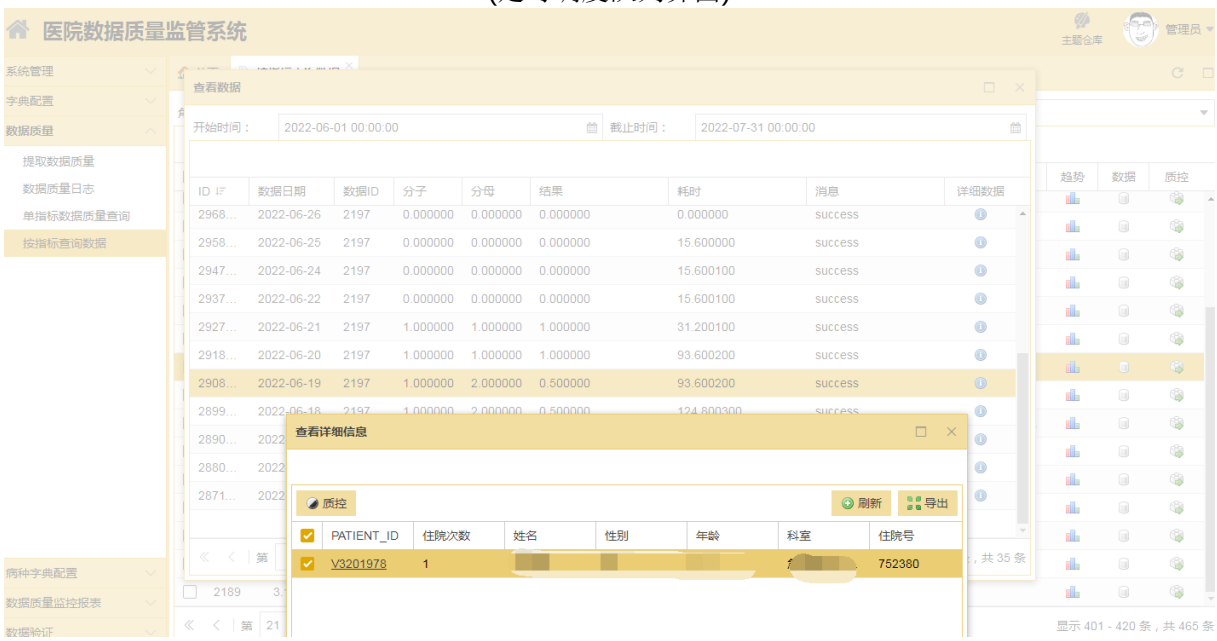
(每日数据指标界面)

(数据指标趋势界面)

2.3、 其它数据采集(Quarter 定时调度，利用夜间凌晨空闲资源)



(定时调度队列界面)



数据明细(异常数据下转界面)

三、几点体会

技术的开发不过多表述，主要是在开发完成后采集数据的过程有几点体会写出来，以期抛砖引玉。

3.1、明确数据指标的核心意义，仔细阅读指标含义，做好指标定义，不要按照自己想当然的对指标数据的理解；

3.2、正所谓“小河有水大河满”，如果没有生产数据源，再好的技术也无法展示出数据，业务部门要参与数据埋点设计（要知晓数据如何产生，在哪个系统产生，用什么样的方式产生更符合采集和管理需要）和生产系统改造；

3.3、业务管理部门要做好数据质控机制和上线验证（指标要关注的是数据，不是数字，要关注数据产生的过程，过程好了，数字结果自然也会好)；

3.4、在实现指标的过程中，要本着数据在数据发生的过程中产生，不要靠事后的补充来产生数据；

3.5、数据生产者对产生的数据负责，当数据发生时候，数据生产者要对数据负责，避免因为数据采集后又变化所带来的偏差；

3.6、“别看广告看疗效”，业务部门不要盲目信任包管一切，上了什么系统就能产生所有数据的噱头，而要真正是技术落地并生根开花结果。

以上是我们做的一点工作，目前的“三甲医院”评审细则，经过对系统的改造（主要是数据埋点），创建 480 条指标（对原指标进行了拆分），验证 165 条指标数据，该系统的应用，能够提高医院评审工作的效率和准确性，减少人工操作的错误率；同时，能够提高医院的管理水平，推动医院的数字化建设。未来，我们将进一步优化系统的性能，完善系统的功能，提高系统的易用性和安全性，以满足更多医院评审工作的需求。

PO-010

医院检验标本混采系统的设计与实现

龙凤舞[1] 胡建中[2] 冯嵩[3] 龙冠现[4]

[1]中南大学湘雅医院网络信息中心

[2]中南大学湘雅医院党委办公室

[3]中南大学湘雅医院网络信息中心

[4]湖南省郴州市妇幼保健院检验科

一．序言

在人类发展的历史上，疫情是一直存在的问题。随着全球现代工业化进程的加快和随之伴随而来的环境、气候异常等问题，各种传染性疾病和突发性公共卫生事件时有发生，快速且有效地筛查和确诊是阻断病毒传播、扩散的重要方法 。当大规模传染性疾病发生时，面对复杂多变的疫情环境，医院在承担日常基本诊疗工作以及院内疫情防控的同时，还要承担面向社会繁重的标本采集任务，基础设施、人力资源以及医学检验设施都面临空前的压力，如果能借助信息技术，在梳理、优化、重构流程的基础上，实现检验标本混采检测功能模块，支持检测工作高效安全地进行，既能确保数据的准确性和实时性，又能为数据的共享、疫情防控领导机关的决策，提供有效的数据支撑。例如在全球持续流行新型冠状病毒肺炎期间，核酸检测被作为检测新冠病毒的重要检查，常用的十合一混采检测技术就是将采集自 10 个受检者的 10 支拭子样本，集合于 1 个采集管中进行核酸检测，混检筛查中一旦发现阳性，将会立即通知相关部门对该混采管的 10 个受试者暂时单独隔离，并重新采集单管拭子进行复核，再确定这十个受检者当中到底哪一个是阳性。这种混合采样模式适用于大规模检测项目，具有筛查高效、节省资源的优势。本文在医院 LIS 系统现有架构基础上，精心设计、自主研发、实现一个高效检验标本混采系统。

二．需求分析

医院医学检验系统的基本业务流程如图 1 所示，分为门诊和住院两个业务流

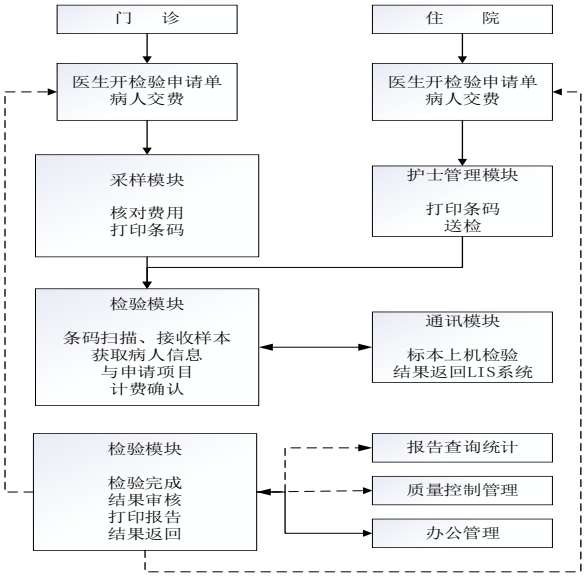


图 1 检验系统业务流程图

程，首先都由医师开具检验申请单，患者缴费以后，护士进行条码打印及样本采集，护工将标本送至检验部门后，检验部门进行接收、上机检验，待仪器将检验结果返回 LIS（Laboratory Information System）系统后，检验医师对样本进行结果审核和报告发送。

本文研究的检验标本混采系统主要应用于门诊场景，对照图1业务流程可知，传统的LIS系统仅支持一人一采，即标本单采，并且每个患者需要依照门诊流程建卡，挂号，由医师开具检验申请单，缴费后在护士站进行样本采集等，只有依照上述完整的门诊顺序流程才能完成检测。当医院承接针对不特定人群大规模检测任务时，显然目前的门诊检验流程完全无法满足实际需求，主要弊端包括以下几个方面，首先，现有门诊检验流程太过繁琐，且受检者必须完成缴费步骤，才能往下进行后续检测，而大规模的检验标本检测任务一般都是免费，这与实际情况相悖；其次，现有门诊检验流程仅支持一人对应一个采样管的标本单采，相较于多人共用一个采样管的混采模式而言，需要耗费数倍的采样管及检验试剂，而且，检验医师的工作量和检验仪器的运行时间无形中增加了数倍，总之，单采的方式显然检测效率低下且成本高企。

由于当大规模疫情发生时形势的瞬息万变，医院必须随时做好开展大规模检验标本检测的准备，同时还要重点考虑如何及时出具检验报告，及时准确地做好对国家各级卫生医疗管理机构的数据上传工作。因此，医院亟需在现有的检验系统基础上，对 LIS 系统进行流程优化和功能升级，开发一个高效可靠的检验标本混采系统，使医院现行的检验系统能实现样本混采的功能。

显然，检验标本混采模式的业务流程既有与医院现有检验流程相同的部分，又有其完全个性化的的独特要求。二者的相同之处在于，当受检者完成了采样以后，后续的标本送出，标本接收，上机检验，结果回传，检验医师审核报告，发送报告，包括受检者最终在自助设备、手机等终端获取

检验报告这些步骤，都可以完全依托医院现有的检验系统完成，甚至无需将混采样本与单采样本分开进行检验。二者的关键不同之处在于，首先，混采模式可以略过门诊业务流程中医师开单和缴费的环节，直接进行受检者基本信息登记和采样，其次，为了达到高效检测和节约人力试剂成本的目的，混采模式要求多个受检者的样本共用一个试管，最后，混采模式要求针对每个受检者出具一份单独的检验报告。

从上述分析可知，系统的流程优化和功能升级必须充分考虑混采模式与常规检验模式流程上的差异，既要满足检验标本混采的特殊需求，实现数据的互联互通和数据共享，又要充分利用现有资源并依托常规检验流程，减少检验各环节有关人员的学习成本，从而检验操作人员感受不到混采标本与其他常规检验标本的明显差异，简化系统操作，减少对现有 LIS 系统的过度介入，经过流程优化和功能升级后，系统实现无感知平滑升级，避免对现有 LIS 系统的稳定性造成影响，在有效降低风险的同时，大幅度提升用户体验。

基于上述分析，系统的流程优化和功能升级聚焦两个关键环节，即护士站样本采集环节，以及流转到检验部门之后为样本编号的环节，医院医学检验系统其他常规检验等主要功能模块保持不变。这样检验标本混采时只需针对单独的检验标本混采模块，进行条码的打印和信息采集，之后的所有业务流程跟其他的检验样本完全一致。由于没有新增其他外部系统和改变现有的检验操作流程，这种流程优化和功能升级对原有系统的稳定性影响很小，可操作性强，几乎没有增加操作员工作量，甚至不需要对操作员进行额外培训就可以直接上线，改造成本低，后期维护非常方便。

三．设计和实现

1. 整体架构。如图 2 示，在现有的系统架构上增加检验标本混采平台，受检者在混采模块完成标本采集，后续的标本送出、标本送达、检验部门接收标本、编号、上机检测、结果审核、结果发送、受检者查询打印检验报告等功能依托现有 LIS 系统完成。

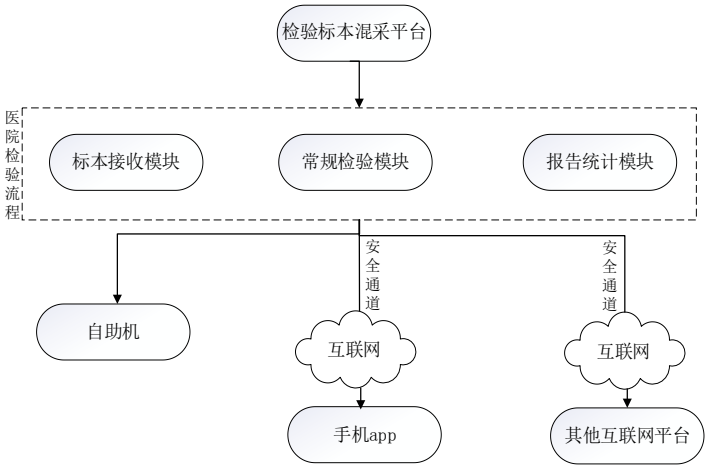


图 2 检验标本混采系统整体架构图

2. 流程优化。在医院现有医学检验系统基础上，检验标本混采流程优化之后的相关数据流如图 3 示。首先，在样本采集的环节，将一个条码对应多个受检者的申请信息，以多条记录的方式存入

新增的人员基本信息表及混采申请信息表中，其次，在 LIS 系统后续环节的数据交换过程中，该样本和其他普通样本一样，作为一条记录进行常规处理，直到检验部门编号上机，最后，再将采样时记录在混采申请信息表中的多条申请记录，同步到相关的检验报告表中，即为每一个混采受检者生成一份检验报告，满足后续数据交换过程中，检验标本混采功能模块为每一个受检者单独出具一份检验报告的需求。

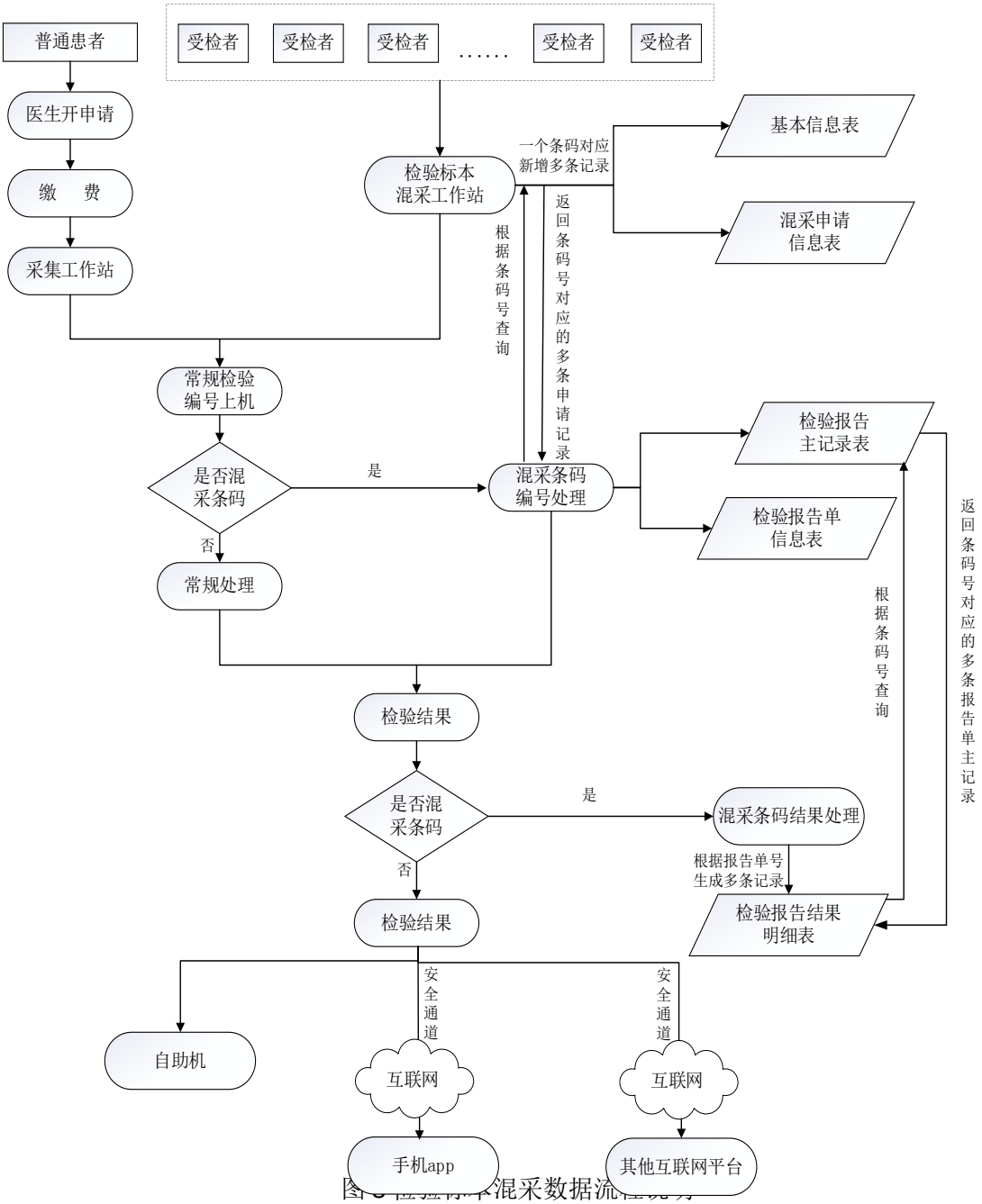


图 3 检验标本混采数据流程说明

3. 主要功能

通过对大型医院开展检验标本混采的需求分析、流程优化和功能模块设计，充分结合实际情况，最终编程实现了混合采集平台的信息登记、系统设置、混合采集、常规检验模块升级以及检验报告的审核、输出、查询等各项功能。

3.1 信息登记。受检者名单可以采用 Excel 表格的方式由申请机构提交医院，系统提前将受检者信息导入并记录在数据库中。开始检验标本采集时，只有在数据库中有记录，即通过身份认证的人员才能进入后续的检验标本混采流程。图 4 为受检者基本信息表结构。

Name	Type	Nullable	Default	Storage	Comments
OUTPATIENT_ID	VARCHAR2(50)	☐			门诊号
PATIENT_NAME	VARCHAR2(50)	☑			受检者姓名
PATIENT_SEX	VARCHAR2(20)	☑			性别
MOBILE_NO	VARCHAR2(50)	☑			手机号
ID_CARD	VARCHAR2(50)	☑			身份证号
PATIENT_DEPT	VARCHAR2(100)	☑			工作单位及部门
TYPE	VARCHAR2(50)	☑			受检者类型
SAMPLING_TIME	DATE	☑			采样时间
SAMPLING_PERSON	VARCHAR2(50)	☑			采样人

图 4 混采人员基本信息表

3.2 系统设置。LIS 系统的数据管理模块中，新增检验标本混采模块的系统设置，包括通过设置受检者数量，控制每一个混采试管包含的受检者数量，以十合一混采为例，若设置受检者为 10 人，则混采系统在检验标本采集时，当每个条码对应的受检者数量达到十人时，系统将自动完成该条码的采集，从而减少人工判断可能带来的差错，同时简化了操作流程。

3.3 混合采集。如图 5 所示，系统每生成和打印一个混采条码，对应采集多个受检者的信息，当采集的人数达到系统设置的数量时，将自动完成该条码的采集。系统在混合采集环节，通过增加检验标本混采申请信息表的方式，将混采模块采集的条码及人员信息记录至该表中，这样可以达到将检验标本混采的申请信息与门诊医师开具的检验标本单采及其他检验项目的申请信息分开处理的目的。该申请表中 10 条记录对应同一条码号，分别对应 10 个受检者的申请信息。

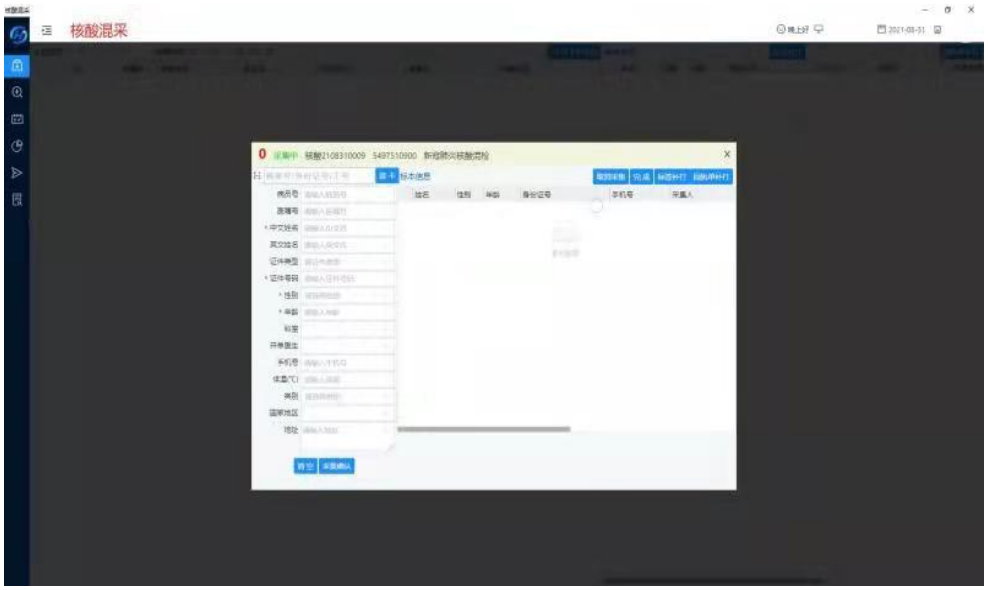


图 5 混合采集平台功能界面

3.4 常规检验模块升级。标本送达检验部门，操作员在给标本编号时，系统将自动判断条码信息，如果是混采条码，则从混采申请信息表中根据条码号提取对应的 10 个受检者申请信息，并在生成后续的检验报告相关数据流程中，生成对应的 10 条记录，这样在检验系统完成检测时，该样本检测的检验结果也将分别对应 10 个受检者生成 10 份检验报告。

3.5 检验报告。检验医师在常规检验模块完成了检验报告的审核和发送后，受检者可以通过医院的自助设备，打印纸质检验报告单，也可以通过手机 APP 或者在居民健康卡公众号查询到检测结果。通过上述功能的实现，一个完整的混采检测闭环流程就完成了。

四． 结语

医院的检验标本采样包括“单采”和“混采”两种模式，分别为一人一管和多人一管，“单采”具有较高的检验精准度，“混采”可提高采集效率，通过对检验标本混采流程的优化和医院医学检验系统功能升级。本文的研究和系统的开发，实现了上述两种采样方式的无缝衔接。医护人员在进行检验样本采集时，只需要根据实际需求灵活选择相应的检验标本采集流程，就能很好地满足疫情环境下的临床需求。同时，受检者通过自助设备打印的检测报告，或者通过移动端 APP 及其他互联网查询，均可查询检测结果，实现了信息化闭环管理，以及数据的互联互通和共享，不但达到了高效检测的目的，而且节约了人力资源和试剂成本，在实际运用中取得了很好的效果。

PO-011

中国老年人卫生健康政策主题演化研究

刘帅兵 张云秋
吉林大学公共卫生学院

1 引言

依据联合国规定的 65 岁以上的人口占总人口 7%以上的老龄化标准，2000 年我国正式进入老龄化社会。《中国发展报告 2020》预测，中国将在 2022 年左右，由老龄化社会进入老龄社会[1]。

老年人属于典型的“医学易感人群”，其卫生健康问题已成为各界关注的重点。

在 2016 年全国卫生与健康大会上，习近平总书记强调“要把人民健康放在优先发展的战略地位，为老年人提供连续的健康管理服务和医疗服务”[2]。同年国务院出台《“健康中国 2030”规划纲要》[3]，提出要实现全民健康，其中老年人更是被作为重点人群提及。2018 年我国组建国家卫生健康委员会，并设立老龄健康司，将建立完善老年健康服务体系作为国家卫生健康委员会的新增职责。2019 年 7 月国务院印发《国务院关于实施健康中国行动的意见》[4]，其中“老年健康促进”作为 15 个重大专项行动之一，将“健康老龄化”落实到行动层面。

我国关于老年人卫生健康服务政策研究尚处于起步阶段，现存的研究成果主要集中于从政策的主题、效果评价、演变历程单方面对其进行分析[5-8]，缺乏系统性及全面性。在加速老龄化背景下，现有的老年人群的卫生健康政策制度体系仍存在诸多问题。因此，本文旨在通过对我国出台的老年人卫生健康相关政策进行系统梳理，综合对政策进行发文量分析、发文部门分析及主题分析，解析政策的演化过程，探讨当前老年人卫生健康政策设计可能存在的问题或不足，为我国老年卫生健康政策体系的完善和执行落实提供参考。

2 材料与方法

2.1 政策选择

本研究选取“北大法宝”法律法规数据库及国务院办公厅官网，以“老年卫生健康”、“老年人医疗”、“老年健康服务”、“医养结合”等关键词进行系统检索，结合标题及政策全文，共收集到 1982 年 12 月——2023 年 8 月颁布的与老年卫生健康相关的中央层面的法律法规、通知公告等政策类文件 154 份，基于相关性原则将文件按与“老年人卫生健康”相关性由强到弱评为 1-3 级，数量分别为 44、49、61。将评级为 1 级的 44 份政策文件视为主要政策，提取其发文时间、发布部门、主要内容等关键信息进行后续的文本分析。

2.2 分析方法

第一步，数据预处理，首先对停用词典及过滤词典进行扩充，使用 ROSTCM 软件对政策文本进行分词及词频统计，得到符合研究主题范围的有效关键词词频表。

第二步，共词分析。将生成的有效关键词进行整理，运用 ROSTCM 软件进行主题关键词共现处理，并生成共词矩阵。其中，对角线上的数据代表该关键词在所有的政策性文本中出现的频次，非对角线上的数据代表两个相同的关键词在相同政策性文本中共同出现的联合频次。

第三步，主题分析及编码。本文使用社会网络分析软件 Gephi，将关键词共词频次作为权重值，通过社区发现进行主题聚类，相同颜色节点代表归为同一主题（如图 1 所示），并对聚类结果进行主题编码。

图 1 主题聚类结果图示

3 结果

3.1 老年人卫生健康政策文件描述性统计分析

3.1.1 政策文件出台时间分析

本研究依据 ROSTCM 词频统计结果、历年文件出台数量趋势（图 2）及我国老年人卫生健康关键政策的时间节点，将我国老年人卫生健康政策演变划分为四个主要阶段：起步酝酿阶段(1982-2000 年)，此时政策发展处于起步酝酿阶段，出台文件数量较少；初步发展阶段(2001-2014 年)，随着我国正式进入老龄化社会，老年人卫生健康政策进入初步发展阶段，各项政策陆续出台，政策内容涵盖老年人卫生健康各领域；快速发展阶段(2015-2018 年)，随着我国明确医养结合的重点任务，《“健康中国 2030”规划纲要》、《“十三五”健康老龄化规划》出台背景下，各项政策密集出台，老年人卫生健康政策进入快速发展阶段；系统整合阶段（2019-至今），国家卫健委老龄健康司正式成立，面向国内老年人的卫生健康政策发文量急剧增长，这一阶段政府正式提出建立完善老年健康服务体系，政策演变呈现体系化发展特征，进入了系统整合阶段。

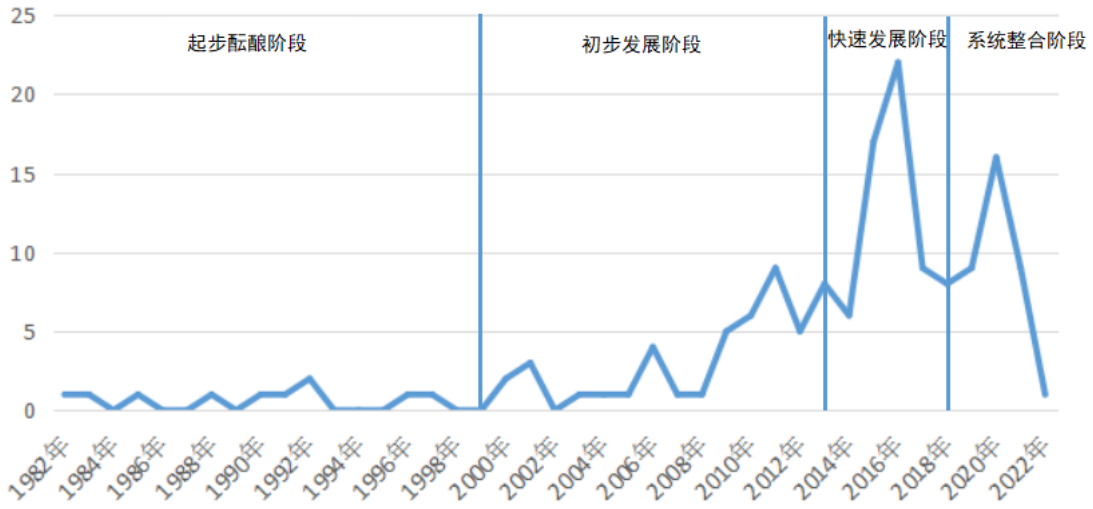


图 2 老年人卫生健康政策发文量时间趋势图

3.1.2 政策发文部门分析

从整体上看，所有 154 份老年人卫生健康政策制定共涉及 35 个部门。其中涉及部门最多的政策是 2019 年 10 月国家卫生健康委员会、民政部等 12 个部门联合发布的《国家卫生健康委、民政部、国家发展改革委等关于深入推进医养结合发展的若干意见》，这也是最早出现的多部门联合政策。

表 1 老年人卫生健康政策发文部门数分布表

发文部门数	政策文件编号	百分比
1	1，2，3，4，5，……150，151，153，154（共 110 件）	71.43%
2	10，11，22，35，39……127，142，147，152（共 21 件）	13.64%
3	18，24，29，30，56，……91，95，131，147（共 16 件）	10.39%
4	77，112（共 2 件）	1.30%
5 个及以上	124，125，133，135，138（共 5 件）	3.25%

由表 1 所示，我国老年人卫生健康政策由单个部门独立制定的数量共计 110 件，占政策总数的 71.43%，说明我国老年人卫生健康政策制定主导部门与其他部门协同较少，老年人卫生健康政策主体发文数量分布如图 3 所示。

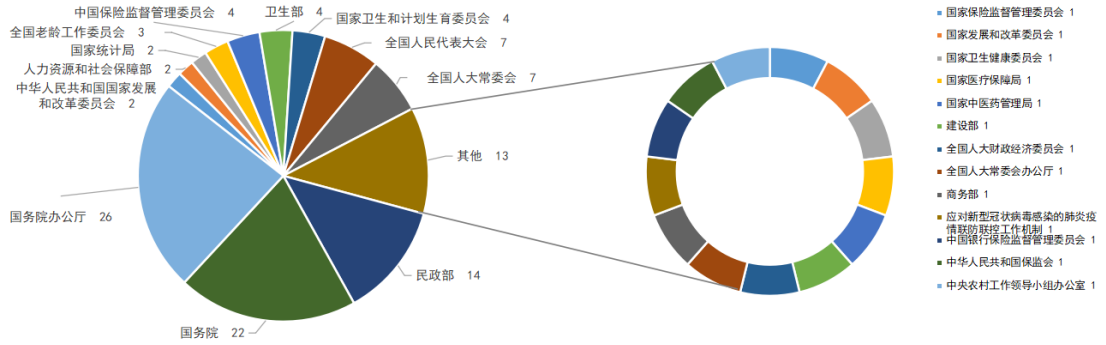


图 3 老年人卫生健康政策主体单独发文数量分布图

图 4 通过社会网络分析软件 Gephi 展示了老年人卫生健康政策联合发文部门的共现关系，从中可以看出民政部、国家发展和改革委员会、财政部等部门处于中心节点，协同程度较高。老年人卫生健康政策的发展关乎国计民生，各部门应从政策高度进一步加强对老年人卫生健康政策工作的重视，加强政策主体间的有效沟通，认真贯彻落实相关政策，从而更好满足老年人卫生健康服务需求。

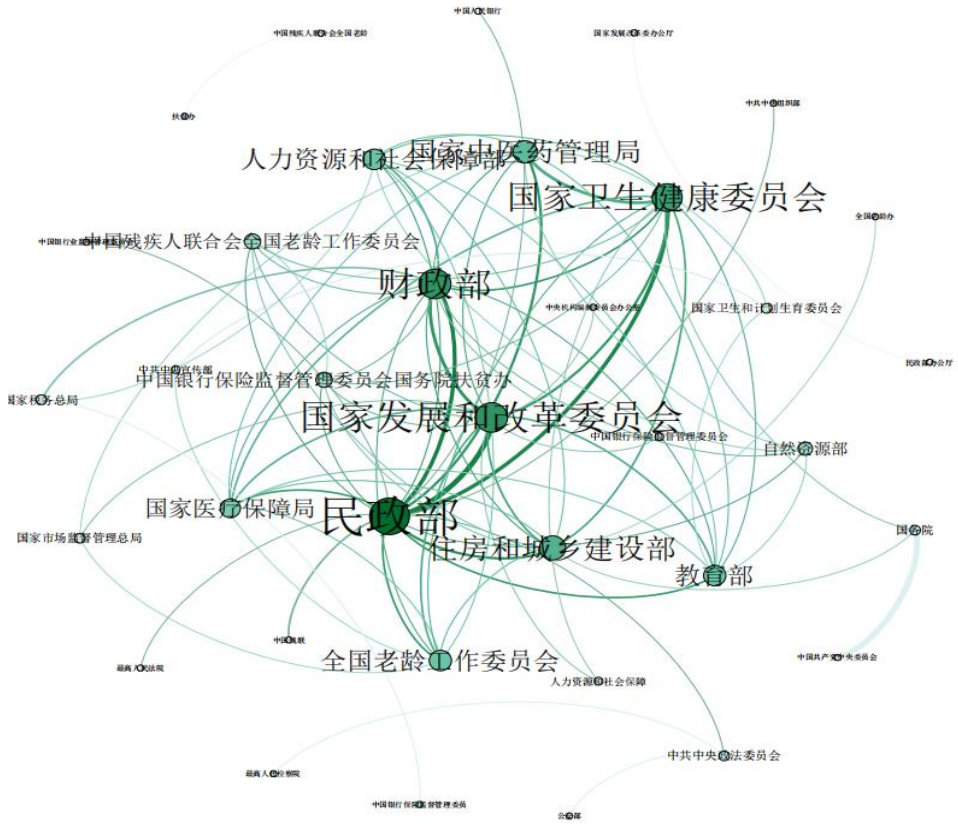


图 4 老年人卫生健康政策发文部门共现关系图

3.2 老年人卫生健康政策主题演化分析

我国老年人卫生健康政策可划分为四个主要阶段，每个阶段具有不同的政策发展重点（见图5）。

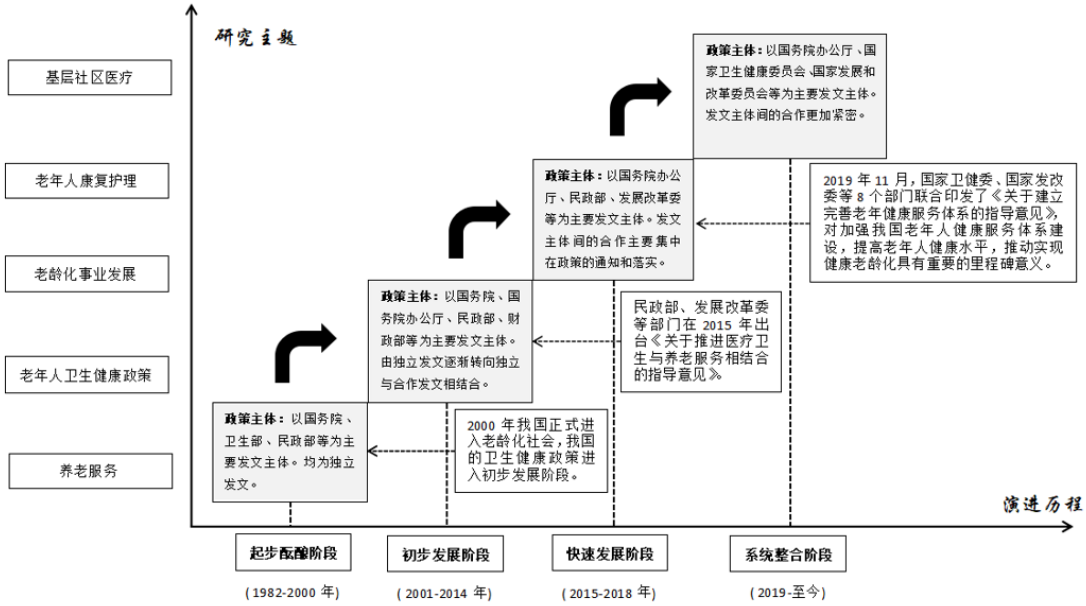


图5 我国老年人卫生健康政策演进历程

3.2.1 1982-2000年：起步酝酿阶段

1978年，党的十一届三中全会以后，老年人卫生健康问题开始得到政府的初步重视，老年人卫生健康政策进入起步酝酿阶段。1985年，《卫生部关于加强我国老年医疗卫生工作的意见》颁布，这是我国第一部专门针对老年卫生健康服务的政策文件，促进了老年医疗卫生工作的发展。1992年，民政部印发了《县级农村社会养老保险基本方案（试行）》的通知、国务院颁布了《民政部关于进一步加快发展农村社会养老保险事业的通知》，养老保险事业得到发展，老年人生活得到保障。1996年，全国人大常委会颁布《中华人民共和国老年人权益保障法》，弥补了我国老年人权益保护法律的空白，在法律层面切实保障老年人的合法权益。

在这一阶段，我国的老年人卫生健康问题开始受到关注，老年人卫生健康政策陆续出台，老龄机构初步建立，老年人卫生健康政策体系雏形开始形成，老年人卫生健康工作开始向规范化和专业化方向发展。但这一时期国家的政策发展处于起步酝酿阶段，出台文件数量较少，老年人卫生健康政策系统性的工作思路尚未形成。

3.2.2 2000-2014年：初步发展阶段

随着2000年我国正式进入老龄化社会，我国老年人卫生健康政策也进入初步发展阶段，这一时期在保障老年基本生活的同时，逐渐强调运用多元方法来满足老年人的多样化需求，体现出实用

性和大众化的特征。

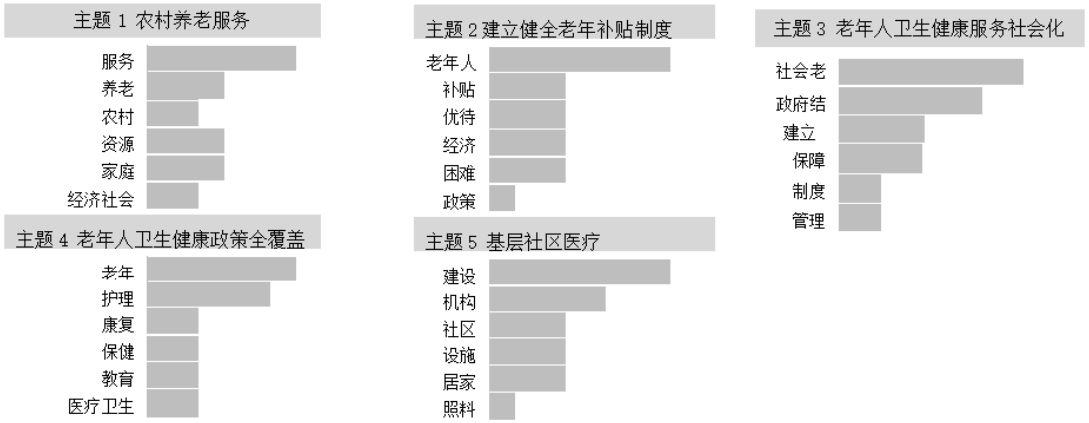


图 6 初步发展阶段政策文本主题分析结果

主题一：农村养老服务

地方各级政府重视农村老龄问题和农民的养老问题，整合资源，加大对农村养老服务的投入。充分发挥社区、基层组织和服务机构在居家养老服务中的重要作用，着力保障家庭特殊困难老年人的养老服务需求，推动农村养老服务与城乡发展相协调。2013 年国务院《关于加快发展养老服务业的若干意见》以及《关于加快发展养老服务业的若干意见》提出要切实加强农村养老服务，促进了社会和谐，推进经济社会持续健康发展。

主题二：建立健全老年补贴制度

各地陆续建立健全了经济困难的高龄、失能等老年人补贴制度，进一步推动养老服务体系建设，发挥补贴政策的带动作用。2013 年国务院出台《关于加快发展养老服务业的若干意见》，首次提出积极推动医养融合发展，建立健全老年人的补贴制度。而后又颁布《关于建立健全经济困难的高龄失能等老年人补贴制度的通知》使得实现基本养老服务均等化。[9]

主题三：老年人卫生健康服务社会化

这一阶段社会养老服务体系建设坚持政府主导，鼓励社会资本、组织、机构等积极参与，不断完善管理制度、丰富服务内容、健全服务标准，满足了人民群众日益增长的养老服务需求的持续发展，形成了多方协同的老年人服务提供格局。这一阶段国家颁布的若干政策反映我国老年卫生健康服务体系不仅仅局限于提供疾病防治，提出了康复护理、心理监测、健康教育、应急救援、日常锻炼等要点，形成了与老龄化社会需求相适应的专业化老年人康复护理队伍。

主题四：老年人卫生健康政策全覆盖

随着人口老龄化、高龄化的加剧，为实施积极应对人口老龄化国家战略，我国政府部门深入研究我国人口老龄化的趋势特征，精准识别人口老龄化带来的风险挑战，构建系统性、阶段性应对策略。这一阶段中国老年人卫生健康政策正一步步向“全面综合服务”发展，从单纯关注老年人的经济供养问题向全面保障老年医疗卫生、精神生活和老年权益等方面发展。

主题五：基层社区医疗

强调老年人卫生健康服务中基层自治的作用，老年人卫生健康资源进一步向居家社区服务倾斜，基层社区成为各项老年人卫生健康服务的主要依托平台。2012 年民政部发布《关于鼓励和引导民间资本进入养老服务领域的实施意见》对于实现养老服务投资主体多元化，加快推进以居家为基础、社区为依托、机构为支撑的社会养老服务体系建设具有重要意义。2013 年《关于促进健康服务业发展的若干意见》，以及 2014 年《关于推动养老服务产业发展的指导意见》均强调了发展老年人卫生健康服务业中社区的重要意义。

3.2.3 2015-2018 年：快速发展阶段

随着我国明确医养结合的重点任务，我国老年人卫生健康政策进入快速发展阶段。

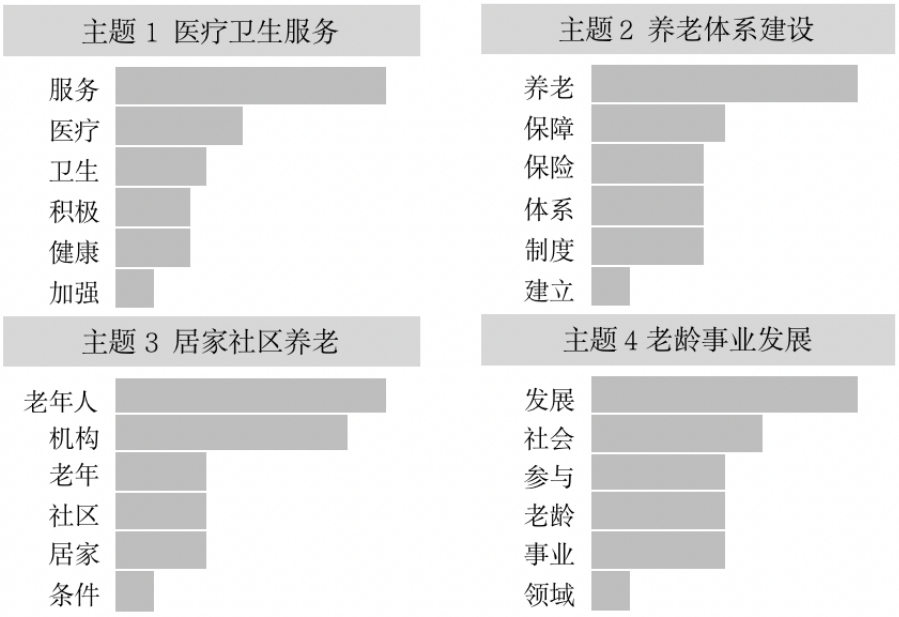


图 7 快速发展阶段政策文本主题分析结果

主题一：医疗卫生服务

这一阶段老年人的医疗卫生服务备受关注，老年人群的医疗卫生服务得到了长足的发展。这与民政部、发展改革委等部门在 2015 年出台的《关于推进医疗卫生与养老服务相结合的指导意见》这一政策相关，意见中明确了医养结合的重点任务。此后，各地陆续开展了老年人群的医疗卫生服务与养老服务相结合的试点，推动了养老机构与医疗卫生服务机构的共建。

主题二：养老体系建设

为了破解长期以来制约养老体系建设过程中的融资难问题，民政部、国家开发银行在 2015 年联合发布的《关于开发性金融支持社会养老服务体系建设的实施意见》[10]中，明确提出要推动形

成“政府引导、金融支持、社会参与、市场运作”的社会养老服务发展体制。充分发挥开发性金融资金的引领作用，吸引民间资本投入，形成以居家为基础、社区为依托、机构为支撑的社会养老服务体系建设，由此促进养老体系的建设和发展。

主题三：居家社区养老

养老资源进一步居家社区服务倾斜，为进一步提升我国老年人的获得感和幸福感，国务院办公厅在 2017 年发布的《关于制定和实施老年人照顾服务项目的意见》中，提出发展居家养老服务，推进老年宜居社区、老年友好城市建设。基层社区日益成为养老服务的主要依托平台，为居家老年人提供生活照料、医疗护理、精神慰藉等服务。全社会共同推动实现老有所养、老有所医、老有所为、老有所学、老有所乐，在此阶段的养老资源逐步向基层社区转移。

主题四：老龄事业发展

2016 年为我国“十三五”规划的开局之年，国务院在 2017 年印发的《关于“十三五”国家老龄事业发展和养老体系建设规划》，明确“十三五”时期是我国全面建成小康社会决胜阶段，也是我国老龄事业改革发展和养老体系建设的重要战略窗口期。为积极应对日益严峻的老龄化趋势，需提高社会参与老龄事业的积极性，从而推动老龄化事业全面协调可持续发展。

3.2.4 2019 年-至今：系统整合阶段

在系统整合阶段，面向国内老年人群的卫生健康政策发文量急剧增长，政策主题发生变化，政策发展进入了系统整合和规范细化阶段。

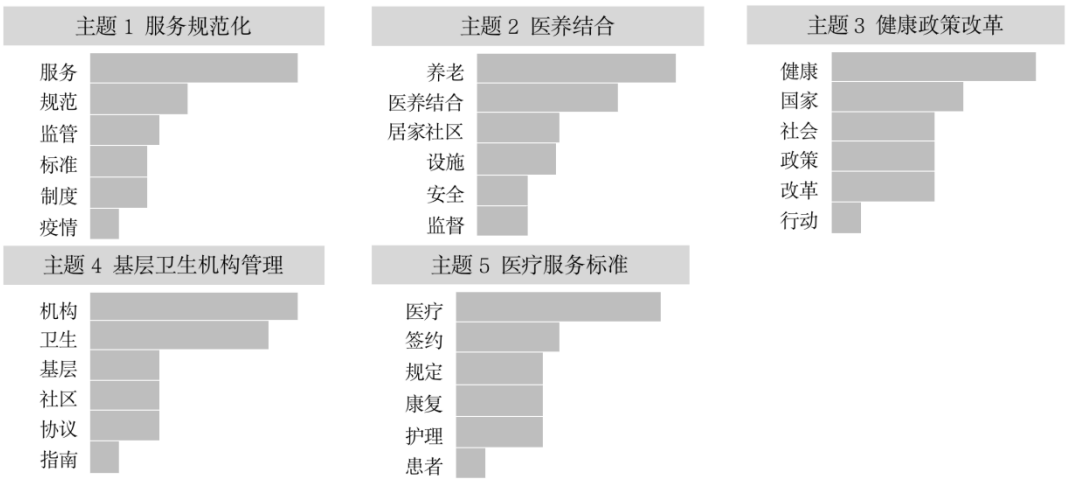


图 8 系统整合阶段政策文本主题分析结果

主题一：服务规范化

2020 年，新型冠状病毒肺炎疫情的爆发给全国卫生健康领域带来了极大的挑战，针对老年人群在疫情期间的养老服务问题，全国老龄办发布了《关于在常态化疫情防控中做好老年人照顾服务工作的通知》，切实保障疫情期间老年人的卫生健康服务需求。这一阶段我国进入“十四五”时期，

党中央把积极应对人口老龄化上升为国家战略。着力于构建和完善兜底性、普惠型、多样化的养老服务体系，不断满足老年人日益增长的多层次、高品质健康养老需求[11]。

主题二：医养结合

深入推进医养结合，将医护技术与养老服务相结合，既能看病又能养老，整合养老和医疗两方面资源，为老人提供及时、便利、精准的医疗服务。当前在中国老年群体中，居家养老仍是主要需求，但服务的创新和服务质量难达到要求[12]。2019 年 6 月国家卫生健康委制定《健康中国行动 2019-2030》发展战略，其中提到完善医养结合政策，推进医疗卫生与养老服务融合发展，推动发展中医药特色医养结合服务。

主题三：健康政策改革

深化健康政策改革，政策和产业双轮驱动，帮助社会更好地应对老龄化。2019 年 11 月，国家卫健委、国家发改委等 8 个部门联合印发了《关于建立完善老年健康服务体系的指导意见》，对加强我国老年健康服务体系建设，提高老年人健康水平，推动实现健康老龄化具有重要的里程碑意义。

主题四：基层卫生机构管理

基层医疗卫生机构既是城市社区服务体系的重要组成，也是守护人民群众健康的第一道防线。截至 2020 年底，全国有基层医疗卫生机构 97 万个，全国基层卫生人员超过 430 万人，基本实现了城乡基层社区的全面覆盖[13]。这一阶段国家致力于提升基层医疗卫生机构服务水平。鼓励有条件的医疗卫生机构，特别是基层的医疗卫生机构开展养老服务，增设养老床位，支持社会办医养结合机构，为老年人居家养老提供延伸的医养结合服务。

主题五：医疗服务标准

我国致力于完善老年人医疗服务标准，补齐康复医疗的短板，建立一个以专业康复机构、基层医疗卫生机构为主体的低成本康复医疗服务体系。2021 年国家卫健委、中国残联等八部门联合印发《关于加快推进康复医疗工作发展的意见》，力争到 2022 年，逐步建立一支数量合理、素质优良的康复医疗专业队伍[14]。同时国家探索建立长期护理保险制度，开展长期护理保险试点道路。我国社会老龄化呈现出失能人群剧增、长期护理成本居高不下、长期护理服务供需失衡等特征，长期护理保险制度一直被期待可以把医疗和养老两个问题有效衔接到一起。

4 讨论

我国老年健康服务政策逐步确立了“健康老龄化”的目标与内涵，建立起卫生健康服务体系的整体框架，各项政策的广度和深度都蓬勃发展，然而总体来看现有政策体系仍存在一些不足。

4.1 政策内容模糊

我国政府在寻求解决老年卫生健康难题方法的最优方案上出台大量政策，但对于如何真正推动

老年服务业发展、各地方政府如何因地制宜提出方案以及政策能否真正落实等问题，现有政策文件中都较少涉及。比如老年人卫生健康领域关键问题——养老问题，国家出台的养老政策覆盖面虽广，但由于各个独立政策相互间联系不足，导致难以形成合力。同时现有养老政策更多仅针对养老问题的某一方面或将养老问题作为其中一部分，并未对全国社会服务化养老发展的短期、中期、长期进行明确规划，导致社会化养老也成为一个十分模糊的概念。

4.2 养老服务城乡不平衡，不均等

我国城乡二元结构和区域发展不平衡的现状，导致社会保障与养老服务也存在城乡差异和区域差别。尽管近年来我国老年人卫生健康服务持续发展，服务参与主体多元、供给主体扩大，但主要体现在城市地区，农村老年人卫生健康服务仍存在不足，且这种情况在山区留守老人身上体现的更加明显。随着我国社会主义事业的发展进入新时代，经济发展水平显著提高，却依然面临东西部地区经济发展不均衡的难题，而经济发展状况恰恰会对老年人群医疗卫生健康服务的供给产生直接影响。老年人群的医疗卫生健康服务和养老服务的均等化程度是我国要解决的一大难题。

4.3 政策执行主体认同度低，执行力度不足

由于老年健康维护和促进的效益往往具有滞留性、潜藏性、隐蔽性，不能形成立竿见影的经济和社会效益，各级政府尚未深刻意识到健康对于个体和社会的重要意义，缺乏老年健康管理的观念与认知，政策认同度低，政策执行的力度不足。长久以来，我国医疗机构更多关注“把病治好”而非“疾病预防”，健康教育、慢性病患者管理、健康素养促进、预防与保健等健康服务尚处于起步阶段。老年卫生健康政策能否达到预期目标很大一部分取决于老年人个体的态度，但我国大部分老年人还抱有“生病后就医”的传统思想。如何提高政策执行主体认同度，是我国推进老年人卫生健康政策亟待解决的问题。

PO-012

基于三医一体化统筹的数据标准规范实践与思考

刘阳[1] 康亚西[2] 陈光焰[3] 林加论[4]

[1]中南大学湘雅医学院附属海口医院（海口市人民医院）信息管理处

[2]海南省人民医院（海南医学院附属海南医院）

[3]海南省卫生健康委员会统计信息中心

[4]海南医学院生物医学信息与工程学院

0 引言

三医联动是具有时代性、体现规律性、富有创造性的医改方略，主要涉及医疗管理体制、医保付费方式、医药监管机制的联动改革[1]。在这样的改革创新推动下，海南省通过三医一体化数据标准规范的制定与应用，解决长期以来存在于海南省医疗、医保、医药之间，医疗机构之间，卫健委与省内其他各厅局之间数据交换不规范的问题，能够有效提升数据交换的规范性，让数据的互联互通、融合共享能够有标可循、有规可依。信息互联互通、业务协同的重要基础要素就是建设三医一体化数据标准规范，同时也意味着全省医疗数据标准规范的确立。

1 医疗数据标准化建设的意义

在创新中国高质量发展和“十四五”规划新发展格局下，“海南省三医联动一张网项目”通过信息化、数字化的方式支撑并推动着医疗、医保、医药等三个领域资源的联动整合。这不仅提升了医药卫生信息化的互联互通水平、促进各级各类医疗卫生机构优化服务流程和创新服务模式，而且推动了跨机构、跨业务条线的业务协同与数据共享，助力构建新型医疗卫生健康服务体系。医疗数据标准化在三医联动项目的实施与落地，对提高海南省医疗卫生服务水平和运营效率，降低医疗机构运营成本和居民看病的负担，推动医药卫生健康行业的科学发展，都起到了极其重要的驱动作用[2]。

2 数据标准化的现状与问题

十二五以来，国内卫生信息标准化研制和实施取得了积极进展，国家层面已发布的卫生信息标准 224 项：基础类标准 60 项、医院信息化标准 88 项、区域卫生信息化标准 76 项，其中数据标准发布了 WS363-2011《卫生信息数据元目录》、WS364-2011《卫生数据元值域代码》、WS365-2011《城乡居民健康档案基本数据集》、WS373-2012《医疗服务基本数数据集》、WS445-2014《电子病历数据集》、WS599-2018《医院人财物运营管理基本数据集》等 60 多项[3]。国家医疗保障局发布的 YBJ-BZ-SJ-SJJ《医疗保障信息平台数据集规范》。国家食药监局组织发布了 NMPAB-T1008-2019《药品使用单位追溯基本数据集》、YY/T 1752-2020《医疗器械唯一标识数据库基本数据集》等 6 项数据集标准[4]。上述标准的制定与颁布，有效支撑了公共卫生、医疗服务、医疗保障、健康管理、药品管理和综合管理等医疗卫生业务应用。但是，随着三医领域对信息互联互通更高的要求，在海南省医疗、医保、医药各条线独立建设信息化系统的现状下，如何更加统一、高效的融合各种医疗数据，科学、合理的整合信息资源是长期以来困扰医疗数据互联互通的重点和难点问题。

2.1 主数据与业务术语编码不统一问题

三医领域的各系统独立建设，参照的标准不尽相同，有些系统仍使用自定义的主数据与业务术语编码，这对跨系统、跨行业的数据交换与共享利用带来了不便，需要在三医领域内构建一套统一的主数据与业务术语编码体系，通过替换或映射的方式在全省范围内统一编码[5]。

2.2 数据元名称及值域不一致问题

三医领域各系统间存在着同一个数据元，名称不统一、值域定义不一致的问题。在数据进行共享时，会存在语义与语法不通，对数据的理解利用带来了极大的挑战，因此需要在顶层设计时，规避该问题，尽量采用国家、行业已发布的数据元及值域标准。如果国家、行业标准与使用需求发生冲突时，再进行自定义和统一推广[6]。

2.3 为处理多方数据共享需求，存在重复对接问题

三医联动一张网项目建设以前，多存在医疗机构为响应不同部门的数据需求，需要多次对接传输数据，造成资源浪费。随着三医联动一张网项目规划，定位由其作为数据中台承载着三医数据汇聚、融合与共享利用的职能，负责构建“一数一源、一采多用”的数据共享格局[7]。

2.4 数据标准贯标实施不规范问题

因理解偏差，常存在对数据标准利用不规范问题，导致数据质量不高，可用性低等问题，需要加大培训力度与推动标准贯标稽核机制[8]。

2.5 数据标准维护管理不及时导致的标准落后问题

随着政策的调整，业务的不断发展，如果未建立一套良好的数据标准管理制度，则会出现标准规范与实际使用需求的偏差问题，故标准规范的管理制度建设也必不可少[9]。

3 数据标准建设方法

3.1 数据标准规范体系构建

三医一体化数据标准规范体系架构参考我国卫生信息标准体系架构，数据类标准体系分支主要

包括数据元与元数据、代码与编码、数据集、共享文档规范[10]。体系架构图如图 1 所示。

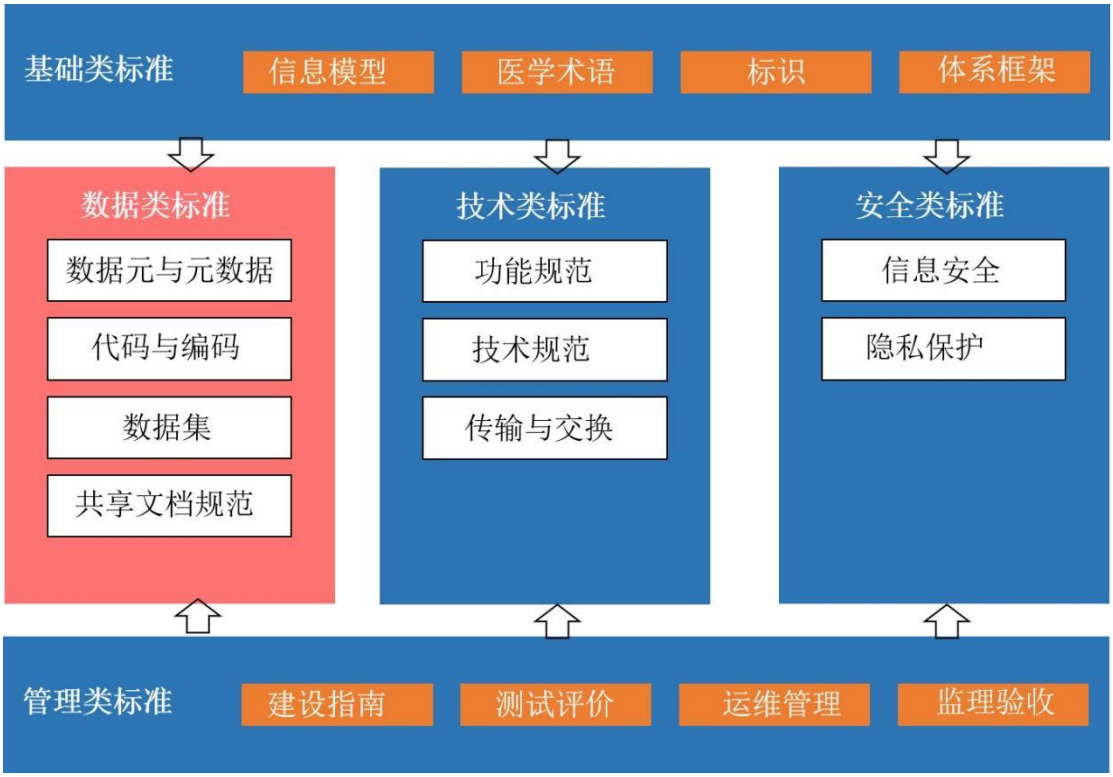


图 1 我国卫生信息标准体系架构图

依据国家卫生信息标准体系架构数据标准分支，三医一体化数据标准规范建设的重点应主要包括业务编码规范、数据集规范、数据元值域代码规范及数据共享规范，同时为了能及时、有效的支撑业务需求，还需要建立数据标准管理规范[11]。

3.2 数据标准制定过程与方法

三医一体化数据标准规范制定须充分结合三医联动一张网项目的建设内容及业务目标，梳理各种业务需求、业务流程和数据流程、以及数据流向，参照国家已发布的数据集标准规划出数据集分类框架[12]，并融合面向三医联动需求的数据元，生成最终的数据集规范、数据元值域代码规范，同时充分考虑统一业务编码，便于数据交换终端进行编码统一实施，降低数据共享成本。最终统一规划三医联动范围内的数据共享规范。

3.2.1 业务分析

三医联动一张网项目建设遵循统一设计、统一标准、统一技术架构、统一业务规范的原则，支持横向纵向业务协同、信息共享，建设智慧健康服务、药品智慧监管应用、数字疾控应用、区域血液管理等应用，并打造业务中台和数据中台，通过中台能力支撑基于三医平台的如三医联合监管、基本公卫督导、检验检查互认、药品全周期追溯监管、耗材全周期追溯监管、医保规则提醒、传染病报卡、血液直免等业务联动应用及各业务系统间的数据共享需求[2]。

3.2.2 数据标准规范制定

3.2.2.1 数据集标准规范制定

参照国家已发布的数据元、数据集标准，并结合三医联动的场景扩展产生符合三医业务需求的标准规范，主要以业务需求为目标，收集三医相关的数据集标准，例如医保的门诊/住院结算清单、卫健的门诊/住院就诊记录、处方/医嘱记录、收费记录、病案首页及药监的生产流通数据[13]，通过标准的差异分析，并参考 WS373-2012《医疗服务基本数数据集》为基础，进行标准的融合，

形成一数一源的，支撑跨部门协同场景的融合标准，以支撑卫统的综合分析、慢病诊断数据推送至基层的医防融合、药品全流程追溯等业务需求。

3.2.2.2 数据元值域代码规范制定

数据源值域代码规范制定主要分为三部分，一部分参照 GB/T 部分国家代码标准，一部分参照 WS364-2011《卫生数据元值域代码》、WS/T 789—2021《血液产品标签与标识代码标准》、WS/T《临床检验项目分类与代码》等行业代码标准，还有一部分根据需求调研过程中收集的业务表单，经多源比对，形成融合的值域代码[14]。

3.2.2.3 业务编码规范制定

业务编码规范主要是涉及三医联动的主数据及业务术语编码，主要包括：行政管理单位编码标准、医疗机构编码标准、科室编码标准、手术操作编码标准、药品编码标准、药品追溯码标准、医疗服务项目标准、医疗器械编码标准、药品企业编码标准、检验检查项目编码标准、病种编码标准、执业人员编码标准、经办机构编码标准、经办人员编码标准等。业务编码规范制定从唯一、统一、共享、便民和低成本转换等角度综合考虑形成三医统一的编码标准，如医疗机构编码参照全国组织机构统一社会信用代码制定；药品编码、手术操作编码、病种编码、执业人员编码标准、经办机构编码标准、经办人员编码标准充分考虑国家医保贯标现状，参照医保业务标准制定；药品追溯码、医疗器械编码标准参照国家食品药品监督管理局发布的编码标准制定[15]。

3.2.2.4 数据共享规范制定

三医联动一张网项目数据中心通过前置库进行三医相关数据交换，资源需求单位申请数据库和 web 服务的相关权限接口对接需求数据。为了规范数据共享行为，数据共享规范明确了数据交换与共享的流程和多种数据交换与共享的方式，其流程共经过资源需求申请、审批；共享方案制定、审批；数据交换接口开发、联调、数据 ETL 处理等 7 个流程.数据交换与共享方式主要以数据库接口及 API 服务接口为主，其它方式为辅。

3.2.2.5 数据标准管理制定

为推动数据三医一体化数据标准的制定与修订，促使标准及时与应用衔接，初步制定了数据标准管理规范，明确了数据标准规范的管理相关的参与方及职责、管理流程等。

4 数据标准构建结果

最终输出了向上依从已有国家的相关标准规范，向下可以支撑全省执行的三医一体化数据标准规范，包括 12 册数据集标准规范、1 册数据元值域标准规范、1 册业务编码标准规范、1 册数据共享标准规范、1 册数据标准管理规范。数据标准规范整体架构如图 2 所示。

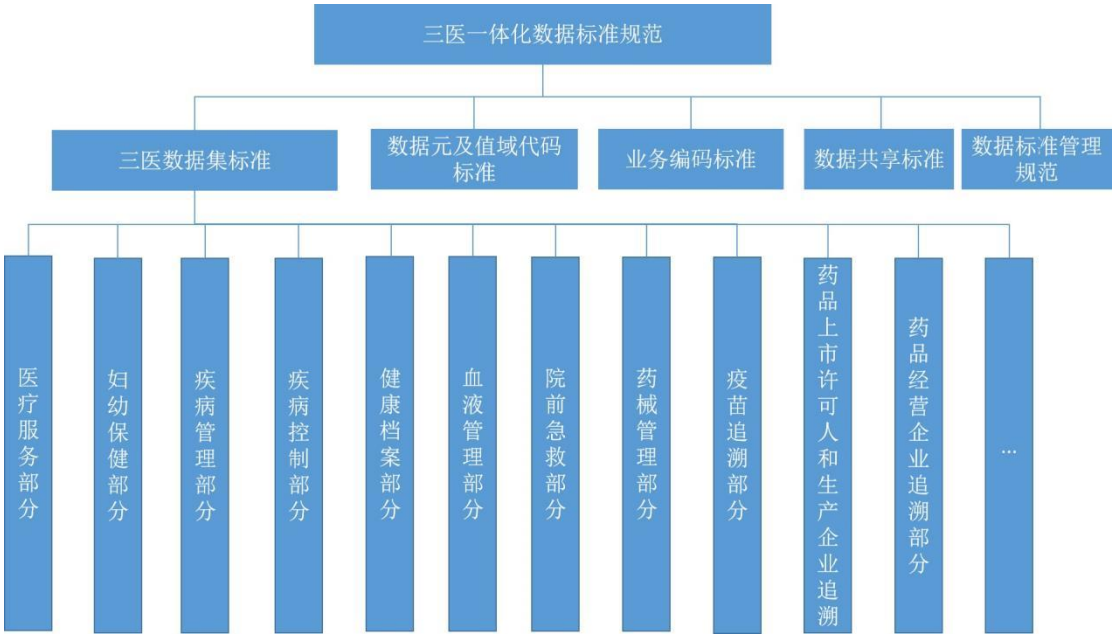


图 2 数据标准规范整体架构

5 数据标准规范的讨论与思考

5.1 数据标准规范的促进作用

通过研究国家、行业相关已发布标准，学习并吸收了标准规范编制方法与过程，结合三医联动一张网项目的业务目标进行需求调研与分析，最终编制了三医一体化数据标准规范[16]。该标准规范具有一定的先进性、创新性、扩展性与实用性，降低了数据共享的难度，破除了三医数据标准不统一、数据不共享问题，打破了三医之间的数据壁垒，实现了业务、数据的高效融合和联动，在检验检查互联互通互认、基本公卫督导、药耗材全流程追溯、三医联合监管、用血直免等业务联动场景中得到了充分应用，进一步促进了业务的数字化转型，并对同类业务具有一定的参考价值[17]。

5.2 遗留问题的思考

在标准规范的管理与应用过程中还遗留了部分问题，标准的使用单位对标准的应用与解读不熟悉、对标准的建设需求反馈意识不强，不利于标准的修订，需要定期开展标准规范的宣传、培训、知识竞赛等多种形式，营造“学习标准、掌握标准、使用标准、管理标准”的良好氛围[18]。推动三医一体化数据标准规范的持续改进。

6 结束语

标准是经济活动和社会发展的技术支撑，数据标准规范是信息化建设、信息互联互通的基础，要不断推动标准实施、强化监督、提升标准化服务能力、夯实标准化工作，同时随着业务的发展，云计算、健康医疗大数据、健康物联网、人工智能、精准医学等技术的不断升级，还要不断完善数据标准建设，加强数据标准落地应用，推动数据标准规范发展，通过标准化助力“健康中国”战略规划和深化医改工作要求贯彻落实。

PO-013

基于生物信息学的妊娠期糖尿病特征基因挖掘

喻慧心 常青 赵玉虹
中国医科大学附属盛京医院数据中心、辽宁省智能诊疗生态系统工程研究中心

1 引言

妊娠期糖尿病（Gestational Diabetes Mellitus, GDM）是指在妊娠前糖代谢正常或存在潜在糖耐量减退症状的女性于妊娠期首次出现的一种糖代谢疾病[1]。随着居民物质生活水平的日益提高以及国家全面二胎政策的放开、三胎政策的出台，我国高龄孕产妇数量激增，高危妊娠风险随之而来，为社会带来了较大的疾病负担[2]。国际糖尿病联合会 2021 年公布数据显示，全球 GDM 标准化患病率为 14%[3]。中国作为世界糖尿病第一大国，GDM 患病率为 14.8%[4]，仅一年中受到妊娠期高血糖影响的活产数高达 870 余万。目前有研究表明，尽管通过积极干预可使 GDM 孕产妇及其胎儿的症状得以改善，但在干预前 GDM 的发生已对二者健康产生不可避免的影响。因此，如何早期预测妊娠期女性 GDM 风险以及临床精准干预治疗，找寻相关生物标志物，以降低 GDM 对孕产妇及胎儿的短期、长期危害，意义重大。

因此，本研究借助生物信息学方法，基于 GEO 数据库所提供的与 GDM 相关的实验数据，采用基因差异分析及共表达网络分析两种方法，对 GDM 致病特征基因进行深入的理论挖掘，旨在找寻 GDM 相关生物特征靶点及致病机制，为 GDM 临床预测、诊断治疗提供新证据。

2 材料与方法

2.1 数据来源

为了探究 GDM 的致病基因，本研究选择纳入来自 GEO 数据库的多系列样本集进行生物信息学分析。GEO 数据库检索词为“gestational diabetes mellitus”，限定 Organism 为 Homo sapiens；Study type 为 Expression profiling by array；Attribute name 为 tissue，得到 GEO 系列数据集结果 10 例。在此基础上进行系列样本筛选，筛选标准为：（1）实验样本未受任何其他因素干扰且仅与 GDM 患病与否相关；（2）同一系列中同时包含 GDM 患病样本及健康对照样本。

通过数据筛选最终在样本差异表达基因分析中纳入其中的 4 组小样本表达谱数据，分别为：GSE19649、GSE49524、GSE65737、GSE87295（表 1）；在基因共表达网络分析中共纳入 2 组大样本表达谱数据，即：GSE70493、GSE128381（表 2）。

表 1 4 例基因表达谱系列数据基本信息

表达谱数据集 ID	平台号	GDM 样本数	健康对照样本数	来源国家
GSE19649	GPL7350	3	2	中国
GSE49524	GPL7020	3	3	意大利

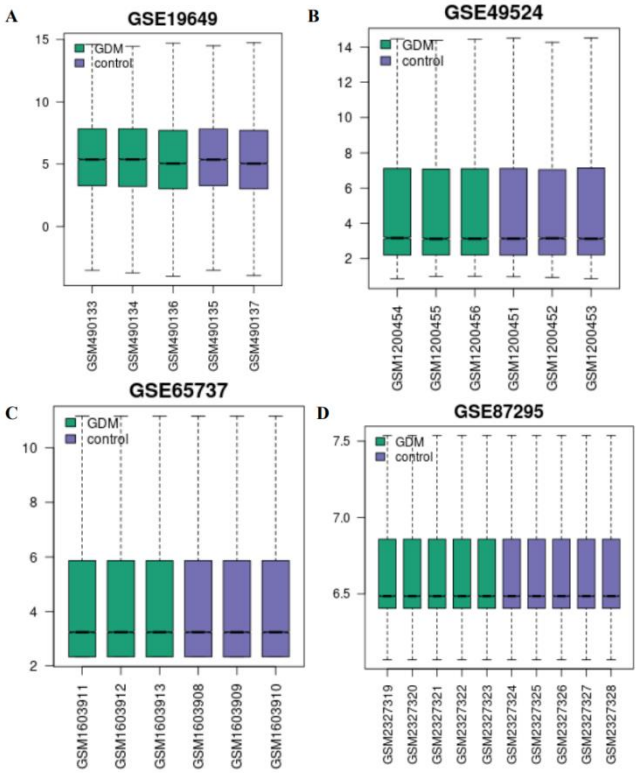
GSE65737	GPL16956	3	3	中国
GSE87295	GPL10558	5	5	新加坡

表 2 2 例基因表达谱系列数据基本信息

表达谱数据集 ID	平台号	GDM 样本数	对照样本数	来源国家
GSE70493	GPL17586	32	31	美国
GSE128381	GPL17077	6	177	比利时

2.2 基因差异表达分析

利用 GEO2R 工具，对 4 组表达谱数据分别进行数据预处理，绘制各组样本的箱线图（图



1）。

图 1 预处理后各组样本水平

以 P. Value < 0.05 为阈值初步筛选各组差异表达基因，将各组基因表达情况绘制成火山图，如图 2 所示。在此基础上，以 |log2FoldChange|＞1 为标准筛选各组差异表达基因。

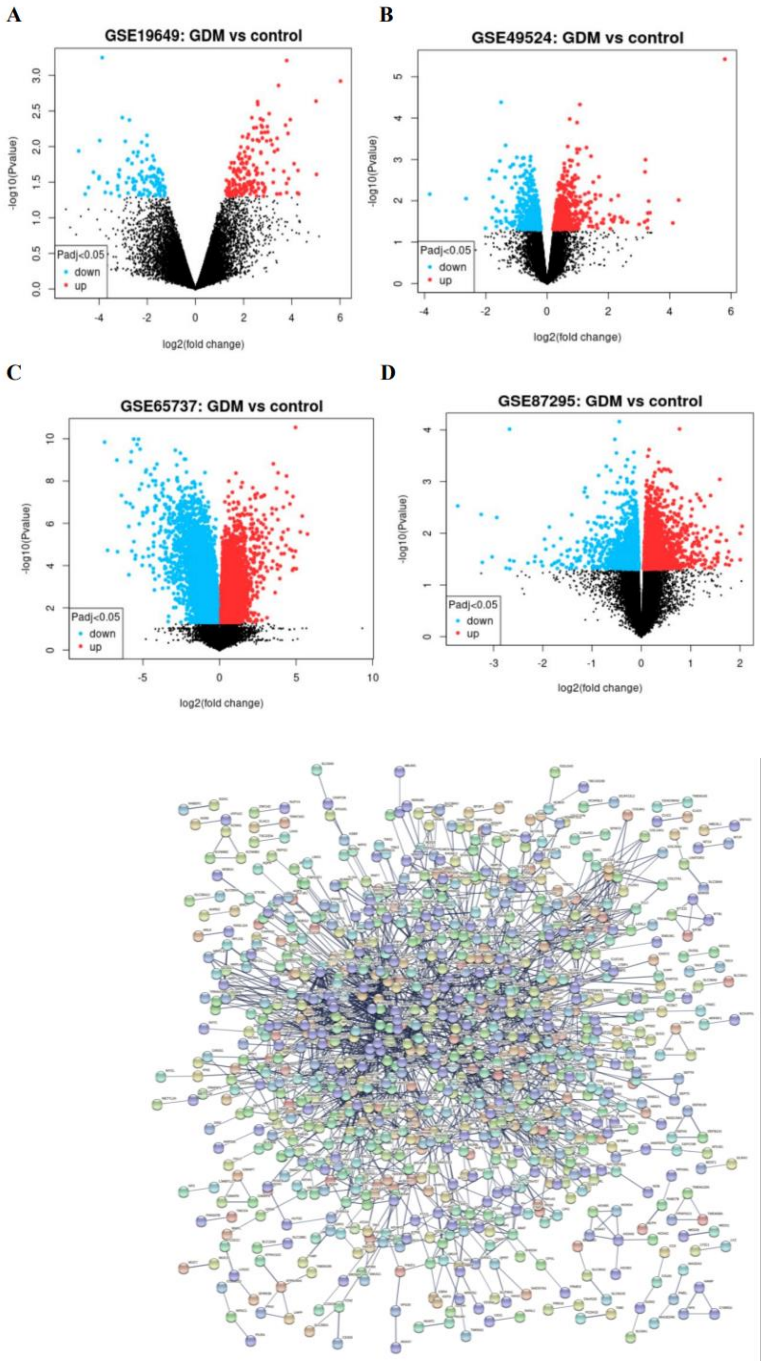
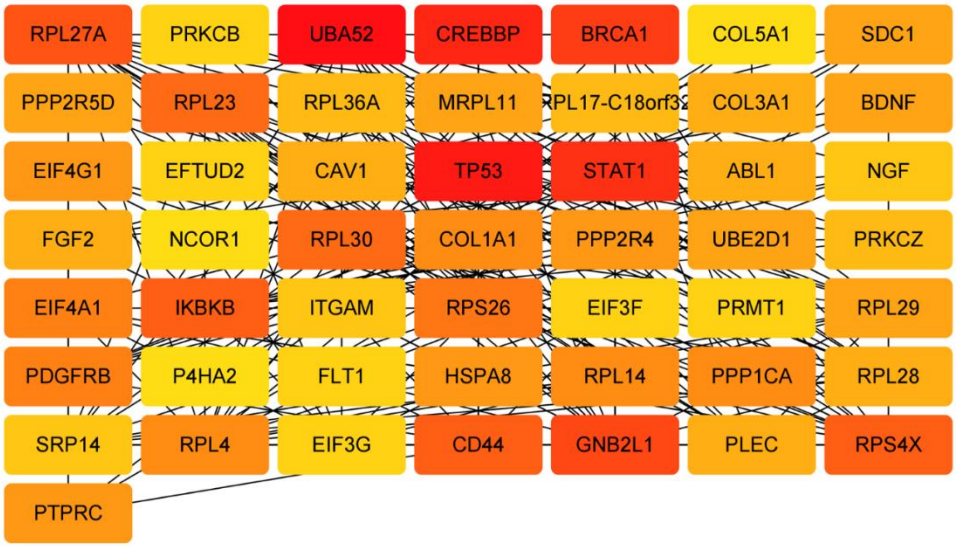


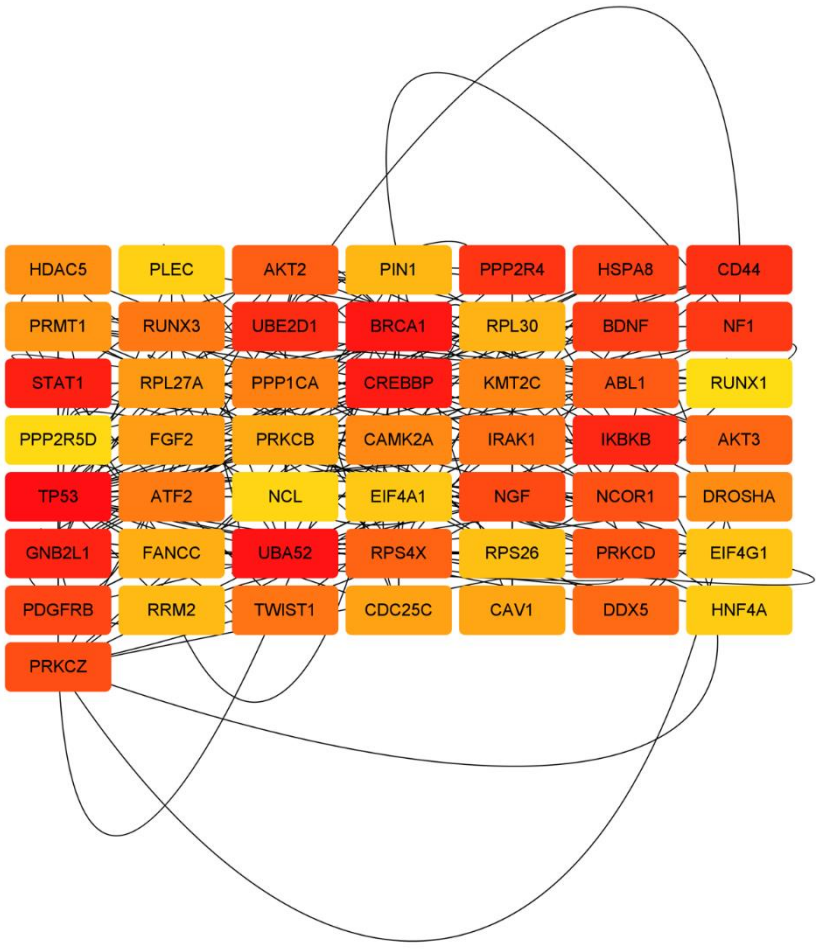
图 2 各组基因差异表达情况火山图

将四组差异表达基因情况汇总，共得到差异表达基因 2565 个。在 STRING 数据库检索 2565 个差异表达基因的具体信息，排除基因集间重复基因及差异编码基因对应非人源蛋白及未命名蛋白后，共纳入差异表达基因 1531 个进行后续分析。将 1531 个差异表达基因导入 STRING 数据库，剔除孤立蛋白节点并将网络置信度得分阈值设为 0.7 高阈值后，可得到相应的差异表达基因编码蛋白相互作用关系网络图（图 3）。

图 3 差异表达基因编码蛋白相互作用网络在对原始 PPI 网络属性进行分析的基础上，通过利用 Cytoscape 软件的 cytohubba 插件，按照网络中各基因编码蛋白节点属性值的大小排序，包括



度数中心度、接近中心度和中介中心度，筛选出各排名前 50 的中心节点（图 4-6）。图 4 度中心



度子网

图 5 接近中心度子网

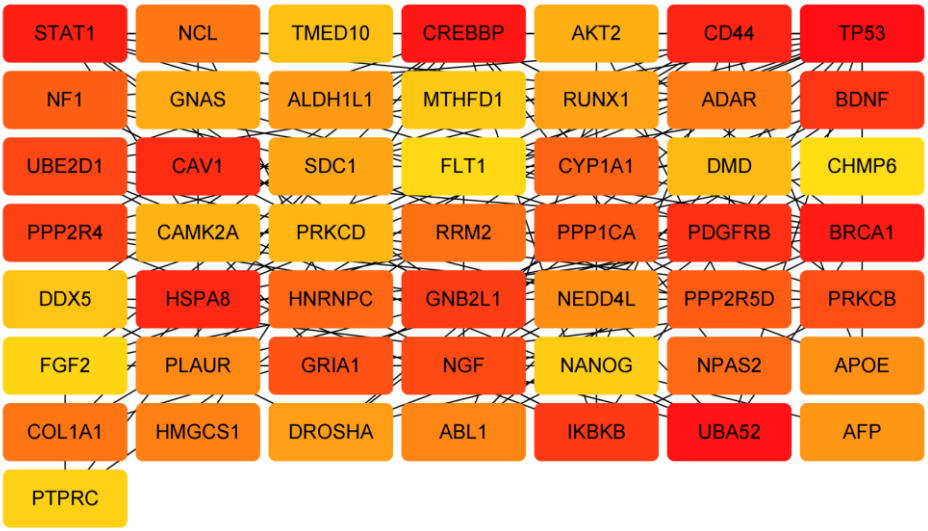


图 6 中介中心度子网

基于此 3 个网络属性值为标准所得的结果节点数量如图 7 所示。各中心性关键基因节点共有 87 个，其中中心性参数均位于总体前 50 的核心节点有 20 个。

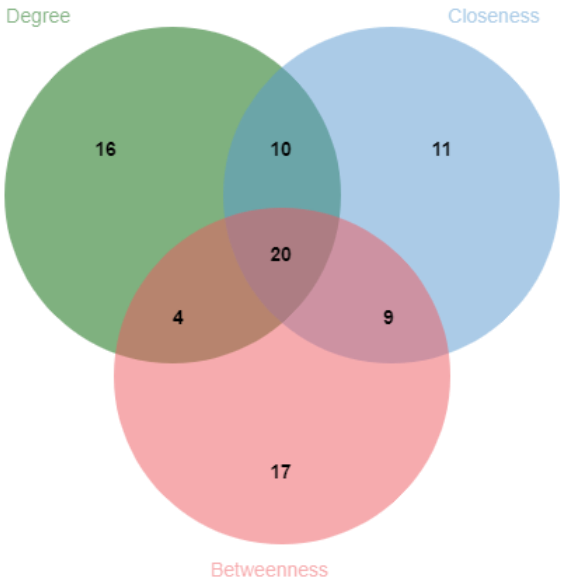


图 7 中介中心度子网

通过检索 GeneCards 数据库，检索策略为：[all]("GDM")OR[all]("gestational diabetes")OR[all]("gestational diabetes mellitus")，可得到与 GDM 相关的靶点 1017 项。限制阈值 Relevance score ＞3，最终得到靶点 286 项。在网络分析所得的 20 个核心节点中，有 3 个核心节点在靶点数据之内，分别为：PDGFRB、CAV1、FGF2。

将网络分析中各网络中心性位于前 50 的 87 个关键节点基因的 OFFICIAL_GENE_SYMBOL 导入 DAVID 数据库，选择物种为“人类”，阈值限定为 FDR＜0.01，得到 GO 富集分析及 KEGG 通路分析结果。调用基于 R 的 ggplot 包可视化图形绘制，得到相应的 GO 条形图（图 8）及 KEGG 气泡图（图 9）。

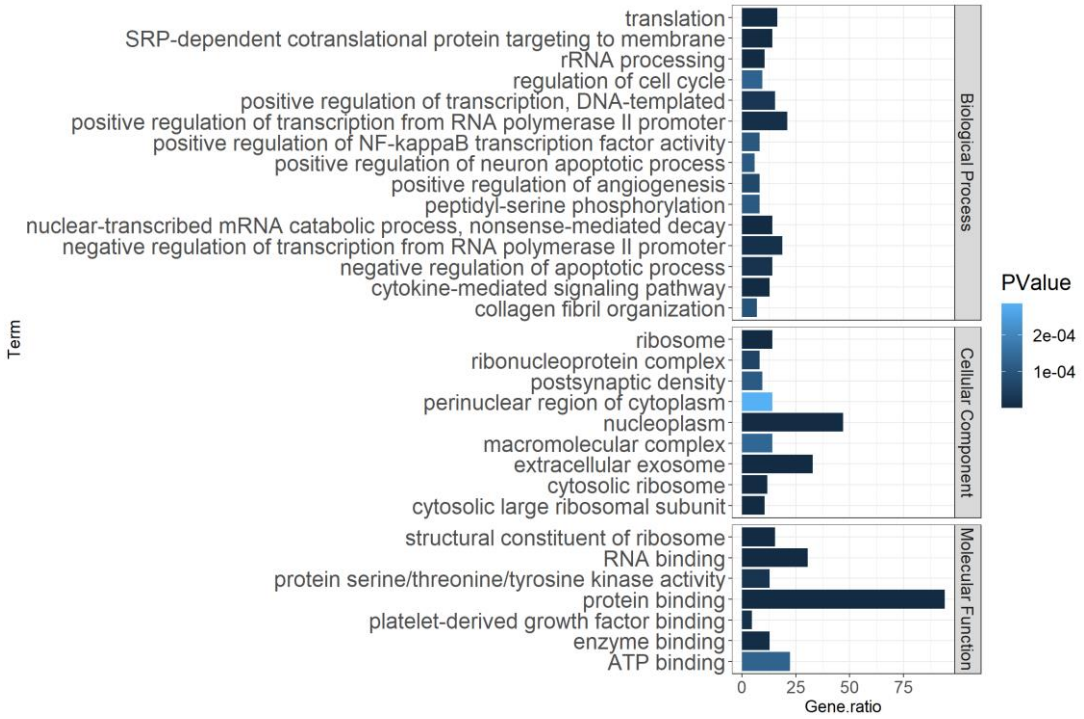


图 8 与 GDM 相关的 GO 富集结果

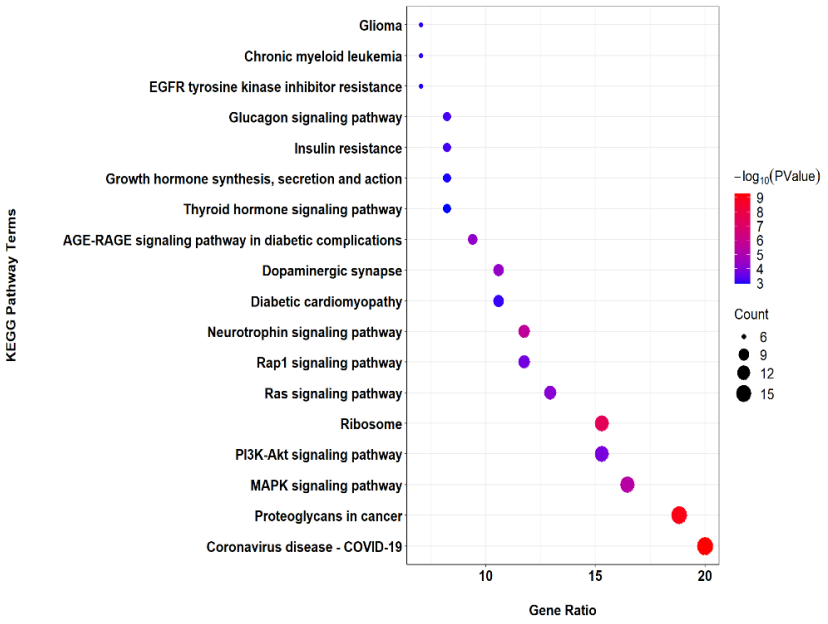


图 9 与 GDM 相关的 KEGG 富集结果

2.3 加权基因共表达网络分析

借助 R 绘制相应聚类树。将剔除异常样本后的数据集样本聚类树与性状数据匹配（图 10）。最终 GSE70493 数据集纳入 63 例样本，共 7726 个基因，486738 个表达量数据；GSE128381 数据集纳入 171 例样本，共 3457 个基因，591147 个表达量数据。

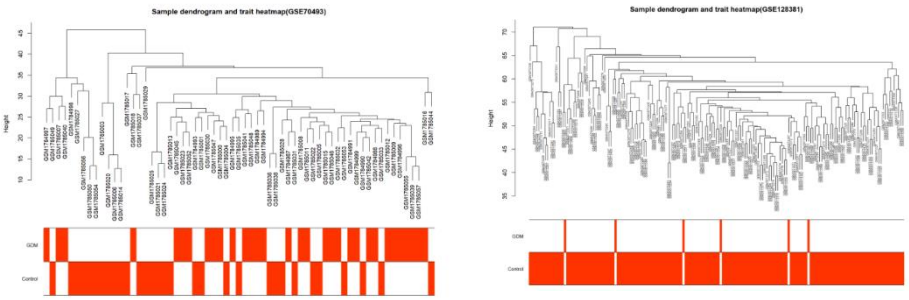


图 10 聚类树与性状热图

通过调用 R 中 WGCNA 程序包的 pickSofiThreshold 函数计算$log_{10}(p(k_i))$和$log_{10}(k_i)$相关系数 SFT.R.sq、模型相关系数 R.sq、网络连通性（mean.k、median.k、max.k）等指标。对于 GSE70493 数据集，当 β=20 时，R^2 大于 0.85 且对应网络平均连通性接近 0；对于 GSE128381 数据集，当 β=4 时，R^2 大于 0.85 且对应网络平均连通性接近 0。故本研究分别选取 β=20（GSE70493）和 β=4（GSE128381）构建后续基因共表达网络。

对于 GSE70493 数据集，采用动态剪切树的方法识别基因模块，设定模块内最小基因数为 50，可初步确定 17 个模块。对模块进行聚类分析，最终得到 5 个基因模块，如图 11 所示。

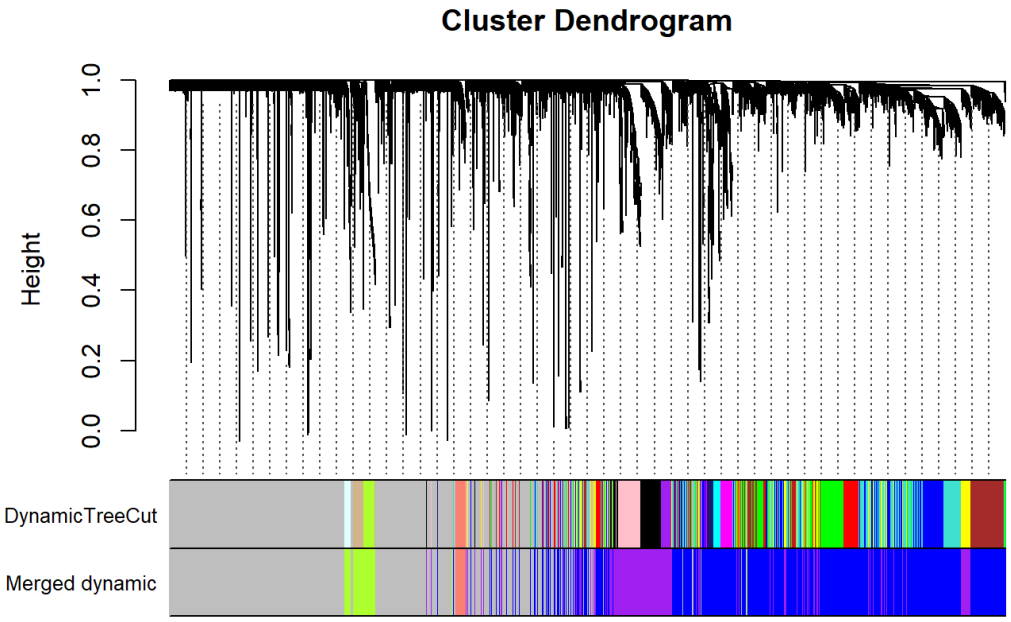


图 11 GSE70493 数据集基因聚类及模块识别

对于 GSE128381 数据集进行相同操作后，设定模块内最小基因数为 30，可初步确定 17 个模块。后通过模块合并，最终得到 13 个基因模块，如图 12 所示。

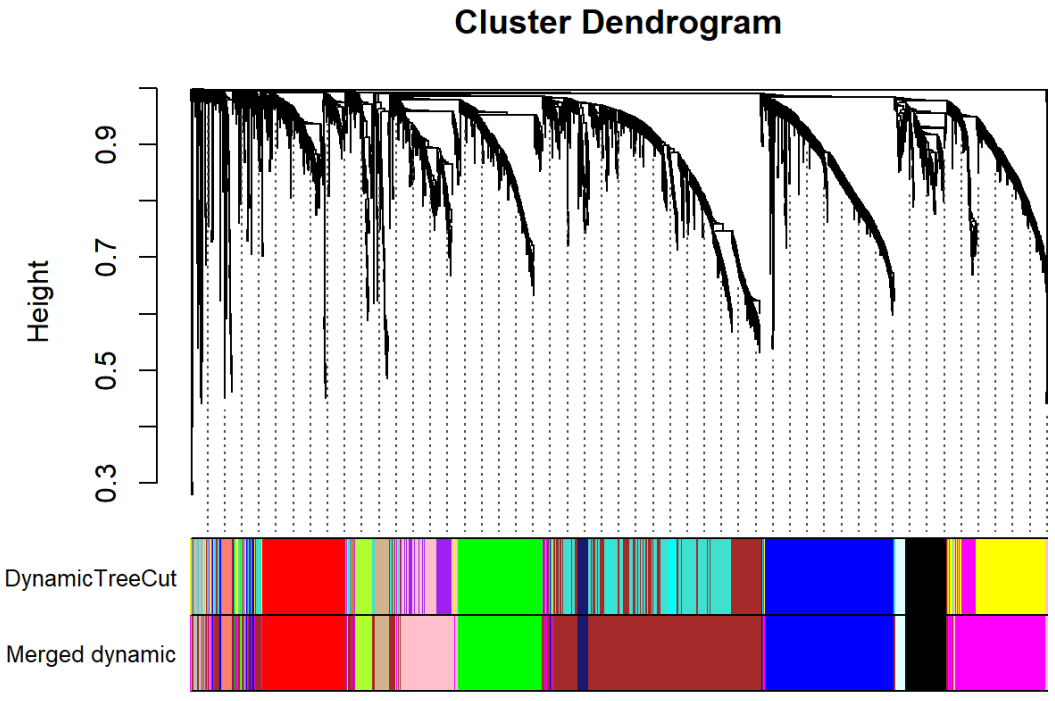


图 12 GSE128381 数据集基因聚类及模块识别

确定基因聚类模块后，将所得样本属性特征与共表达网络进行联合分析。本研究使用每个基因模块特征向量作为该模块的整体基因表达水平与已知的样本属性特征关联分析，可绘制相应热图（图 13、图 14）。

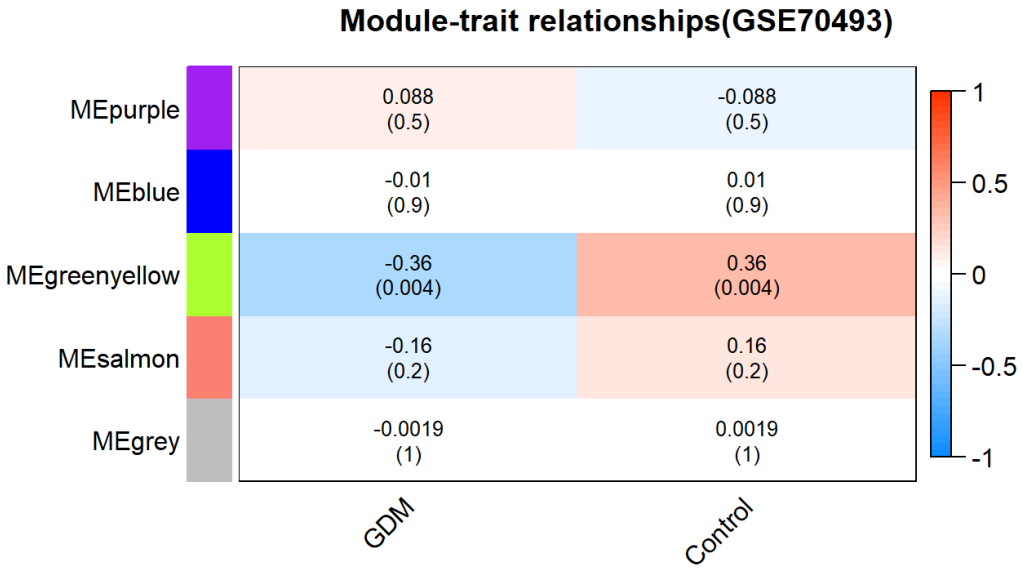


图 13 GSE70493 模块性状关系热图

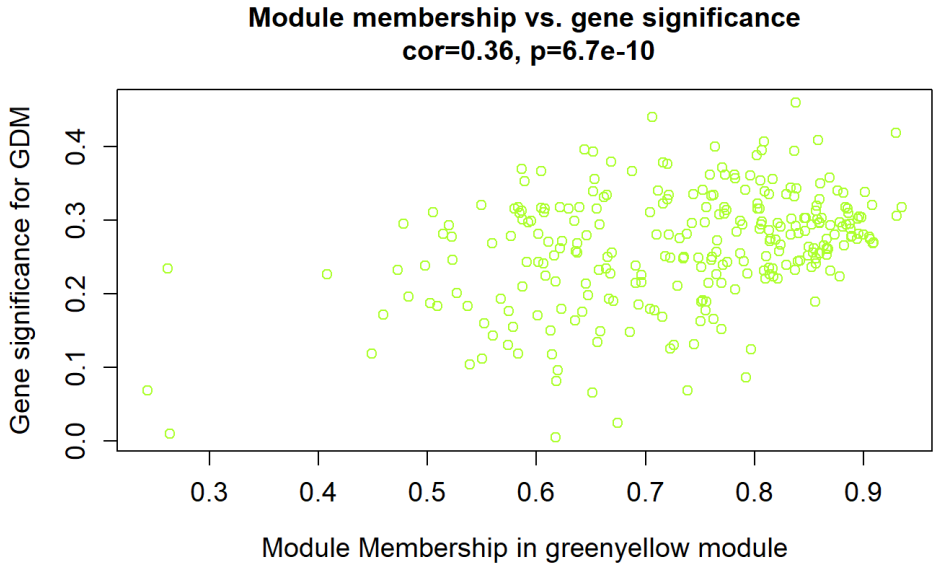
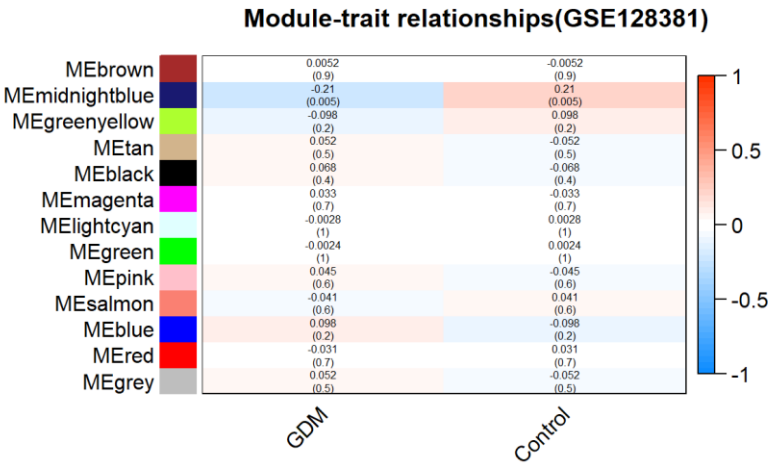


图 14 GSE70493 数据集 greenyellow 模块 GS-MM 散点图

对于 GSE128381 数据集，根据图 15 所示，midnightblue 模块相关系数最大，且模块基因与 GDM 性状呈显著负相关关系。计算此模块中 MM、GS 值并绘制散点图（图 16）。由散点图可知：该模块中 MM 和 GS 的相关系数 cor=0.059，显著性评估 P 值=0.7。

图 15 GSE128381 模块性状关系热图



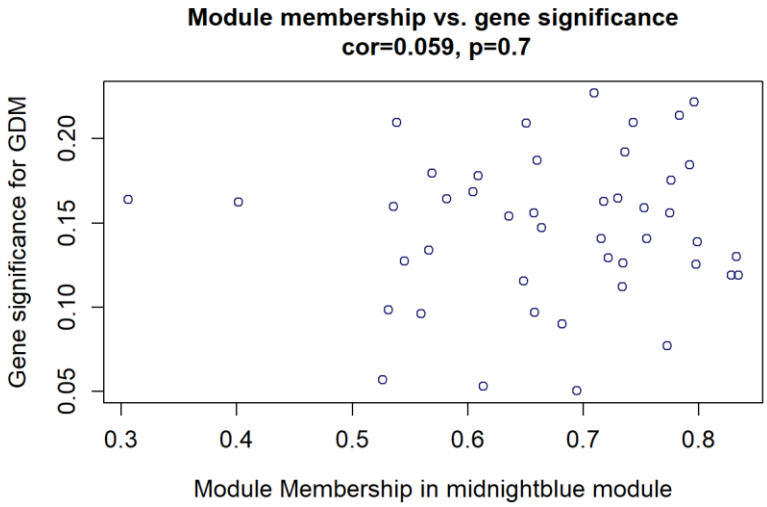


图 16 GSE128381 数据集 midnightblue 模块 GS-MM 散点图

对于 GSE70493 数据集，greenyellow 模块与 GDM 相关性最高，故对此模块中 277 个基因进行进一步分析。综合采用 GS-MM 值标准、基因共表达网络中心性标准及 PPI 网络中心性标准确定关键基因。首先，根据所得的 greenyellow 模块基因 GS、MM 值，以|GS|>0.2 且|MM|>0.8 为标准过滤模块基因，可得符合标准基因 103 个。其次，利用基因共表达网络分析所得 greenyellow 模块网络文件构建相关网络，借助 cytoscape 软件 cytohubba 插件进行网络中心性分析，汇总度数中心度、接近中心度、中介中心度前 10 的基因，可得基因 21 个。最后，将 227 个基因导入 STRING 数据库构建 PPI 网络，共纳入基因 155 个，汇总三类中心度前 10 的基因，可得基因 15 个。

将以上述三类标准所得结果取交集（图 17），得到 GSE70493 数据集与 GDM 相关的关键基因 2 个，分别是：CD74、HLA-DRA。

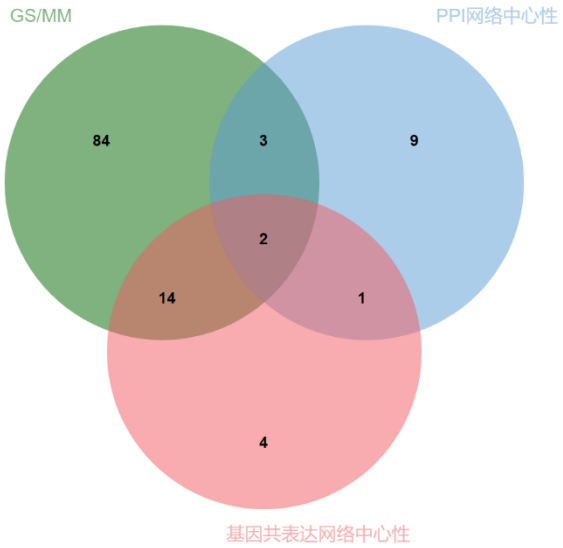


图 17 GSE70493 数据集 Hub 基因选取

借助 cytoscape 的 ClueGO 插件进行 277 个 greenyellow 模块基因的生物学功能和 KEGG 通路富集分析，结果如图 18、图 19 所示。

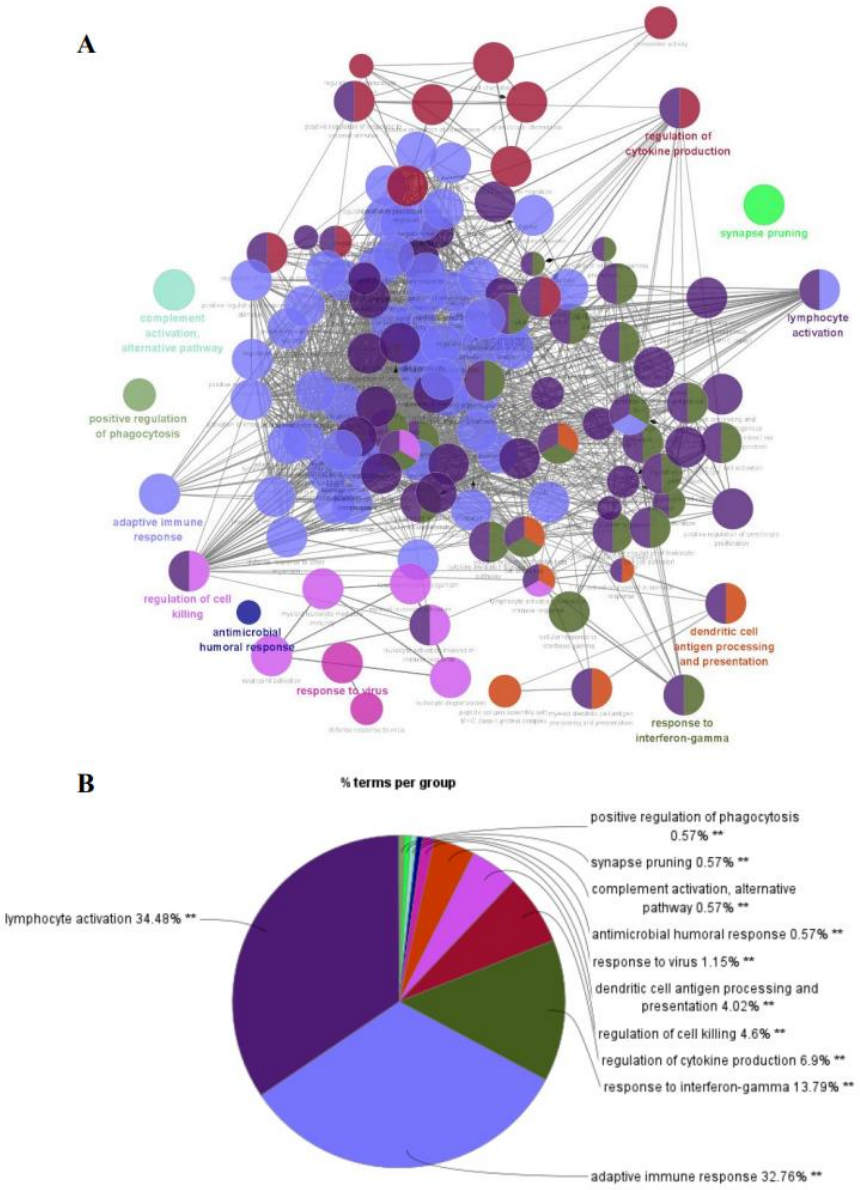


图 18 GSE70493 数据集 GO（BP）富集分析

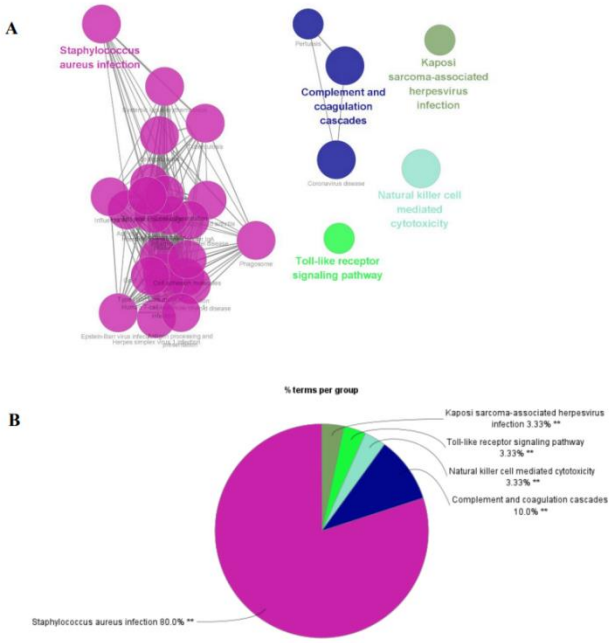


图 19 GSE70493 数据集 KEGG 通路富集分析

对于 GSE128381 数据集，采用相同的关键基因过滤方法，未发现与 GDM 相关的关键基因存在。通过模块基因分析发现：与 GDM 呈显著负相关的 midnightblue 基因模块中所含基因较少，且大部分基因为染色体上未命名的基因。

3 讨论

3.1 GDM 特征基因

PDGFRB（Platelet Derived Growth Factor Receptor Beta）基因，即血小板衍生生长因子受体 β 基因。有研究表明[5]，与健康孕产妇相比，GDM 产妇胎盘中 PDGFRB 的 mRNA 表达显著上升，且患有 GDM 个体及其后代血清中 PDGFRB 的 mRNA 表达水平也较高；同时，通过 PDGFRB 产生的高 PDGFB 水平也可能参与 GDM 女性的胎盘血管生成。由此提示，PDGFRB 基因可能与 GDM 相关，并且可通过母-胎-胎盘路径参与 GDM 相关并发症（如：巨大儿）的发生。

CAV1（Caveolin 1）基因，其编码的蛋白质是一种跨膜蛋白，是形成小窝结构和支架蛋白的主要成分[6]。有研究表明，CAV1 表达增加可导致人类肥胖及 T2DM 的发生发展，且通过研究与不同炎症标志物的关联可进一步表明 CAV1 与肥胖伴随的低度炎症有关[7]。由于 GDM 和 T2DM 具有相似的遗传学机制[8]，由此预测 CAV1 可能在 GDM 致病方面发挥作用。目前已有研究表明，CAV1 可应用于 GDM 并发症的防治。Guo 等人的研究显示，CAV1 可通过 AMPK 介导的脂肪酸转运和 GLUT1 介导的胎盘葡萄糖转运减少 GDM 诱导的巨大儿发生[9]。

FGF2（Fibroblast Growth Factor 2）基因，即成纤维细胞生长因子 2 基因。现有研究显示，FGF21 可参与多种与代谢相关的疾病进程[10,11]，例如代谢综合征、心血管疾病、T2DM 等。有研

究表明，FGF2 基因的 mRNA 及所编码蛋白质的表达水平在 GDM 患者胎盘中显著升高[12]。此外，对于患有 GDM 的孕妇，在其网膜脂肪组织中存在 FGF2 基因等血管生成蛋白和粘附分子的异常表达和分泌[13]，这可能影响母体及其后代的血管生成、炎症发生、脂质和葡萄糖代谢等，进而影响母体及后代健康水平。

CD74 基因，该基因编码的蛋白质与 II 类主要组织相容性复合体相关，是一种非多态性 II 型跨膜糖蛋白。已有研究表明，CD74 分子与糖脂代谢过程潜在相关[14]，提示 CD74 分子可能参与糖尿病相关进程。此外，靶向抑制 CD74 可延缓肥胖相关的脂肪组织炎症及胰岛素抵抗[15]，由此可见，CD74 分子与糖尿病之间存在一定负相关关系，与本研究中 WGCNA 分析结果一致。而由于炎症、肥胖及胰岛素抵抗为 GDM 主要致病因素[16]，故可预测 CD74 可应用于 GDM 疾病治疗相关进程。

HLA-DRA（Major Histocompatibility Complex, Class II, DR Alpha）基因，即主要组织相容性复合体，II 类，DR alpha 基因。一项来自山西地区的 HLA Ⅱ类基因多态性与 GDM 相关性研究[17]显示，HLA-DRA 基因为该地区 GDM 保护基因，此结果与本研究 WGCNA 分析结果一致。因此，未来 HLA-DRA 基因可能作为特征靶点应用于 GDM 诊断治疗，仍需大样本研究及临床试验辅以证实。

3.2 GDM 疾病分子遗传机制

根据 GO 富集及 KEGG 通路富集分析结果，在 GO 富集的生物学过程方面，与 GDM 相关的差异表达基因（STAT1、FGF2、CD44、TP53、CAV1、APOE 等）主要富集于细胞因子介导的信号通路、RNA 聚合酶 II 启动子转录调控、血管生成的正向调节、NF-κB 转录因子活性的正向调节、肽基丝氨酸磷酸化、细胞周期调节等过程，提示相关差异表达基因主要通过细胞因子、NF-κB 调节等过程，影响胎盘血管生成及胰岛素相关作用，进而参与 GDM 疾病进程。

本研究从基因表达量角度，借助 WGCNA 方法，对 GDM 相关基因模块所参与的生物学过程及 KEGG 通路进行深入研究。对 GSE70493 数据集进行基因共表达网络分析，结果显示：greenyellow 基因模块与 GDM 的相关性最高，呈显著负相关关系，其中包含 277 个基因，参与淋巴细胞激活及免疫反应调节、抗原加工提呈、干扰素 γ 产生调节等过程；同时与金黄色葡萄球菌感染、toll 样受体信号通路、补体和凝血级联等通路相关。

据报道，toll 样受体负责识别病原微生物的保守成分，在触发炎症性先天免疫反应中发挥重要作用[18]，进而参与 GDM 疾病进程。此外，多研究证实，胰岛素及糖脂代谢异常是导致 GDM 发生的重要原因。Monika 等人[19]的研究表明，孕产妇妊娠、分娩期间，由于自身代谢系统异常、内分泌系统紊乱，部分孕产妇会出现胰岛素分泌异常、机体糖耐量减弱现象，从而形成 GDM。而对于 GDM 孕产妇来说，糖脂代谢异常不仅影响女性孕期健康，而且对产后甚至 GDM 女性及后代远期

健康产生危害。由此可见，目前研究认为 GDM 致病主要机制为炎症及糖脂代谢，与本研究结果相符，但未来仍需大量基础实验证实这一结论。

4 结论

基于生物信息学方法探寻 GDM 相关生物标志物及分子遗传机制，本研究预测 PDGFRB、CAV1、FGF2、CD74、HLA-DRA 共 5 个基因为 GDM 特征基因。通过富集分析可得，GDM 关键靶点基因主要参与胰岛素代谢过程、糖脂代谢、炎症因子相关通路，为 GDM 分子遗传学机制研究提供新的科学依据。本研究从两个角度应用生物信息学方法进行数据处理，并对所得基因进行网络分析、富集分析，此方法可用于 GDM 特征基因的预测及疾病机制的挖掘，具有可行性及可靠性。

PO-014

人工智能在医学科技查新工作中应用模式和方法探讨

王博文
广东省医学学术交流中心（广东省医学情报研究所）

1 目的

医学查新是医学科技情报工作中的重要一环，也是各大科研项目立项与鉴定的重要参考，通过在相关数据库检索与项目相似的文献，是一项具有决策支撑作用的文献相关服务[1]。在今天飞速发展的医学科技情报的领域中，知识的不断积累产生了大量的文献。这些信息的涌入以及创新驱动发展策略的实施使得科技查新人员在高效高质量的检索、分析相关文献变得日益困难[2-3]。为了应对这一问题，多位学者提出了不同的解决方案，例如：构建科技查新区域协同创新服务联合体、基于产业链的科技查新协同、医药类科技查新站团队协作服务模式等[4-10]。而在人工智能与查新工作结合的方面，有学者进行了从数据科学和人工智能的角度来理解科技查新的业务和数据[11]、科技查新人机协作服务模式构建[12]等方面的研究，但均未探讨 AI 与医学查新工作高效有序的结合模式，在前人的基础上，本文拟探讨利用人工智能（AI）的能力增强信息检索和分析效率的方法与模式。

本研究探索了将人工智能融入医学科技信息检索领域的潜在应用模式和方法。即在医学科技领域，如何通过 AI 与现有文献数据库的协同，从而彻底改变信息检索和情报分析的方式？本研究通过利用 AI 的计算能力优化信息检索过程，使研究人员能够将更多的时间用在深入分析和解读相关文献，从而大大提高从业人员的工作效率与质量。

通过这项研究，我们深刻认识到在医学科技文献检索领域中人工智能的变革潜力。通过全面了解人工智能的应用模式和方法，一方面可以探索不同领域专家之间可能存在的合作。另一方面，在医学信息检索中采用先进的 AI 辅助技术，除了提高查新人员的工作效率外，对一般医学科研工作者在文献检索方面也提供了一定的便利性，进而推动医学科技的发展和应用。

总之，本研究的核心目的是探索如何将人工智能融入医学科技文献检索中,进而提高专业人员检索、分析和利用文献的工作效率，从而推动医学科技领域走向更加高效和创新的未来。

2 方法

2.1 文献检索

(1) 数据收集：使用 python 自动化获取已有数据库中研究成果和学术文献。

(2) 数据清洗与整合：通过工作人员指定检索用词、检索策略等信息对获取的数据进行清洗和整合，并将不同来源的数据进行统一格式转换，以便后续处理和分析。

(3) 文本挖掘与特征提取：利用自然语言处理对筛选过后的文献内容进行分析，提取关键词、主题、作者、摘要等重要信息供查新人员选择相关文献与密切相关文献。

(4) 人工智能算法应用：采用机器学习和深度学习等人工智能算法，一方面喂养大量查新报告模板，训练其使用工作人员选择的参考文献撰写相关的查新报告初稿，另一方面对相关医学文献进行自动分类、聚类和可视化分析，识别研究热点和趋势以供科研人员参考。

2.2 查新人员问卷调查

通过设计针对查新工作人员的问卷调查，收集查新工作人员对于人工智能技术在医学科技查新中的应用看法、意见和体验。问卷内容将包括关于人工智能技术应用频率、效果评价、对工作效率以及报告质量的影响等方面的问题。调查对象主要包括医学图书馆、科研机构、科研情报机构等单位从事查新工作的专业人员。通过问卷调查的数据，分析了解查新工作人员对人工智能应用的态度和观点及对其工作产生的影响，从而探讨如何进一步在查新工作中更好的应用人工智能技术提高查新人员的工作效率与查新报告的质量。

2.3 模拟应用场景

基于已有的研究成果和调查结果，模拟人工智能在医学科技查新工作中的具体应用场景。以文献查找、报告撰写和文献整合为例，探讨人工智能技术在这些方面的具体应用模式。对于文献查找，可以模拟人工智能技术自动检索、分类和过滤文献，为查新工作人员提供精准和全面的信息来源。对于报告撰写，可以考虑人工智能技术自动生成查新报告的结构、内容和格式，以提高报告的准确性和效率。对于文献整合，可以模拟人工智能技术自动整合和分析大量文献，发现研究热点和趋势，为科研人员提供重要的参考和决策支持。

3 结果

3.1 智能化查新系统构建

使用 Python 构建智能化的检索及查新系统,成功整合了多个权威医学文献数据库，其中包括了查新工作中最为常用的的 PubMed 和 SINOMED 等数据库。本系统旨在解决查新过程中信息繁琐、检索效率低下等难题，为查新工作者及医学领域的学者们提供一种高效智能的解决方案。

同时，依赖于对学术研究流程的理解。这一系统采用了自然语言处理技术，使得工作人员只需输入检索词和策略以及年限、地区范围、作者等限定词，系统便能从众多数据库中智能地筛选出与之相关的文献。这一策略不仅节省了研究人员大量的时间和精力，还极大地提升了检索结果的准确性。

在筛选出相关文献后，本系统还能够进一步提取文献的主题、关键词、作者和摘要等内容以供选择，从而使得文献的整理和归纳变得更加高效。同时，对于查新人员而言，在选定相关文献和密切相关文献后，系统能够协助撰写查新报告的初稿，为繁琐的查新报告撰写过程注入了智能的力量。

此外，研究人员通过本系统可以对特定主题、作者、单位等方面相关的文献进行计量学分析，并通过直观的可视化制图方式展现出来。这使得研究人员能够更加深入地了解研究领域的发展动态，轻松地跟踪当下的研究热点问题。

综上所述，这个基于 Python 构建的智能化检索及查新系统，不仅可以大大提高查新工作的工作效率与质量，更可以协助学者们使科研工作更加的便利，帮助他们更好地掌握当下的研究热点，推动医学研究的前进。

3.2 使用场景模拟试用及应用情况调查

共选取查新工作人员 17 人参与系统的试用及使用情况的调查。以项目委托-项目登记-报告撰写-报告审核-报告完成的流程进行智能化查新系统的试用。本系统的使用着重在报告撰写环节。包括：1，在文献查找阶段，本系统根据查新人员指定的检索词、检索策略、年限等条件自动检索出相关文献以供查新人员选择。2，在报告撰写阶段，人工智能能够自动根据查新报告的格式生成查新报告初稿，后续由查新人员审阅更改后交由审核人员审核 3，在文献整合研究方面，人工智能的技术能够通过指定的主题、作者、单位、年限等要求自动分析并整合大量文献，发现研究趋势和热点，为科研人员提供重要参考。

在试用结束后对所有参与智能化系统模拟应用的查新工作人员进行问卷调查，探讨查新人员对智能化查新系统在日常工作中应用的看法及目前存在的问题。调查结果显示：1，绝大多数工作人员（14/17）对人工智能在医学科技查新中的应用表示了积极的态度。他们普遍认为人工智能能够显著提高信息检索的效率，从而节省时间和精力。2，有一部分工作人员（6/17）认为使用智能化查新系统的学习成本较高，有待优化。3，几乎所有参与问卷调查的工作人员都认为智能化查新是今后查新工作的趋势所在。4，大部分人认为（12）本系统的使用在研究人员跟踪最新的研究热点以及进行文献计量学、统计学分析是提供了较大的帮助，相比于人工，本系统在科研选题、结果分析、可视化制图方面有着较大的优势综上，本系统在查新工作中应用显著提高了专业人员检索、分析和利用文献的工作效率。

3.3 推动医学科技发展的潜力

综合研究结果表明，人工智能在医学科技查新工作中具有巨大的应用潜力。通过智能化的信息检索系统，查新人员能够更高效地获取相关文献，提高工作的效率和质量。参与

试用调查人员的积极态度和期望进一步证实了人工智能在医学科技查新中的可行性与重要性。

4 结论

随着科技发展进步，涌现出了大量的。本研究通过查新系统构建、工作人员问卷调查和使用场景模拟等方法，全面探讨了人工智能查新系统在医学科技查新工作中的应用模式和方法，为医学查新人员提供了新的工具和方法，有望推动医学科技的发展和创新。

在科技飞速发展的今天，随着科研成果的迅猛涌现，人们发表的学术文献愈加庞大，这使得文献检索逐渐变得艰巨而繁杂。在海量文献中寻找特定目标文献的难度亦逐渐攀升。对于查新工作而言，筛选各个项目相关文献的难度也愈发凸显，这一系列问题不仅极大地降低了工作效率，还在一定程度上影响了报告的质量与准确性。针对这一问题，本系统提供了一个初步的解决方案。

本项目的核心为依赖人工智能自动化地从大量的数据库中准确快捷地检索所需信息的能力。通过输入相关的检索词和策略，在相关领域的数据库中筛选出与之相关的文献，为科研人员节省大量时间与精力，而且极大地提升了检索结果的精确度。这无疑是一个在学术研究领域的重要突破，助力科研者们更加专注于深入的研究工作，而非困于繁琐的查找环节。同时，本系统可以自动化提取文章的研究主题、关键词、作者信息以及摘要等关键内容，大大地提升了文献整理的效率，还保证了信息的准确性与全面性。在查新工作中，查新人员可以根据需要从检索到的文献中选取相关文献以及密切相关文献中，智能地选取内容，按照报告的格式协助撰写查新报告的初稿，将智能与人工结合，极大地简化了繁琐的查新流程。

因此，我们相信本系统不仅仅是一个工具，更是在科研过程中的得力助手。通过整合多个权威数据库，采用自然语言处理技术实现智能检索和信息提取，以及协助查新报告撰写、基于机器学习的计量学分析和可视化制图，该系统从根本上提升了科研人员的工作效率与质量，进而有力地推动了科技的不断发展与进步。因此，我们有理由相信，随着人工智能的不断进步和科研者的不懈努力，从庞杂的文献中获取所需信息将会变得更加便捷，科技情报工作者的工作效率和质量将得到进一步提升，从而进一步促进科技的创新与发展！

PO-015

国防科技重点领域发展态势情报分析模型构建研究

胡丽娜、郝继英、张谦
军事科学院军事科学信息研究中心

国防科技重点领域是对国家安全、军事博弈至关重要的科技领域,其发展态势情报分析研究对于了解掌握全球科技竞争焦点、支撑国家和军队科技发展战略部署具有重要意义。英国国防部在《国防技术战略》中指出：“任何时候军事优势的达成，都取决于之前 25 年的研发投资”[1]。2022 年 2 月，美国国防部发布《竞争时代国防部技术愿景》，提出了高超声速、定向能、生物技术等 14 个对美国安全至关重要的关键技术领域。2023 年 5 月，美国国防部在其发布的《国防科技战略》再次强调上述 14 个技术领域是增强美国竞争优势的关键。本文在系统梳理国内外情报分析模型研究的基础上，借鉴 Endsley[2]的态势感知理论与方法，构建了由“信息采集层”“态势觉察层”“态势理解测”“态势预测层”四个层次组成的国防科技重点领域发展态势情报分析模型，通过对领域相关的战略规划、科技项目等研究对象及其属性进行分析，进而形成对重点领域发展态势的情报认知。

1 情报分析模型研究现状

情报分析是通过系统化过程将信息转化为知识、情报和谋略的活动[3]，其中系统化过程指的就是情报分析的框架化和模型化[4]。大数据时代的情报分析范式正在发生深刻变革，正在从经验化、理论化的传统情报分析范式向多维化、模型化、智能化发展。钱学森先生[5]是基于模型进行情报分析的提倡者。情报分析模型属于概念模型[6]，即通过具象的数据对抽象事物进行系统描述，具体来讲，是根据特定目的需求，挖掘、收集、整理、分析相关信息，辅助决策者明晰问题现状及未来趋势走向的系统结构性框架。近年来，情报学领域的国内外学者大量借鉴与应用社会学、经济学、计算机科学、系统科学等领域的各类模型，从不同维度对情报分析模型的构建进行研究，总结了多种情报分析模型[6]，主要可分为面向流程的情报分析模型、面向内容的情报分析模型和综合集成型情报分析模型三种类型。

1.1 面向流程的情报分析模型

面向流程的情报分析模型是根据输入调整输出、描述动态性过程的模型，可分为周期模型、工作流程模型、思维过程模型等类型，主要包括情报需求分析、情报信息采集、情报任务处理、情报问题分析、情报任务反馈等步骤[7]。谢尔曼·肯特[8]在其情报周期论中认为情报分析是线性且有序的，包括需求、计划、搜集、处理、分析、分化等 6 个独立步骤。而 Clark[9]认为情报分析活动过程是非线性、网状的，提出了“以目标为中心”的情报分析模型，包括确定目标、明确所需信息、搜集信息源、得到新信息、用户填充新信息与需求等步骤。在国防领域，美军在其颁布的《联合作战情报支援条令》中提出“联合作战情报周期”的概念，将情报分析周期划分为“计划与指导、搜集、加工处理、分析与生产、分发与整合、评估”等六个阶段。美国陆军军事情报分析模型 ATP 2-33.4 也是基于情报周期理论建立的典型情报分析模型。麦肯锡咨询公司提出了由设定目标、定义经营单元、环境分析、战略选择、事实与监控等步骤构成的 7-S 竞争情报分析模型[10]。汤珊红[11]提出了将情报研究人员思维过程显性化的国防科技情报研究思维过程模型，由问题发现、情报立意、假设提出、素材预判、分析综合、方案形成等六个阶段组成。面向流程的情报分析模型强调情报任务解

决步骤的阐释和认知过程的反映，较少关注情报分析任务的具体内容，具有步骤界限分明的优点，但忽略了情报分析中的利益相关方网络[12]，因此，面向流程的情报分析模型也正在发展为动态的网络结构。

1.2 面向内容的情报分析模型

面向内容的情报分析模型关注情报分析任务结构及具体内容[7]，侧重于描述任务内容构成要素及其间关联关系，包括关注任务结构的情报分析模型、关注研究对象的情报分析模型与关注研究方法的情报分析模型。格伦·肯特[12]认为在情报分析实践中应从任务内容的角度，将具有共性因素的任务进行分类，明晰情报分析需解决的问题，并逐步分解为可操作的子模块，据此提出“从战略到任务”的情报分析模型。Krikelas 模型[13]认为信息需求具有特殊性，进行信息分析应将个人背景与相关文献资料相联系，才能获得符合任务需求的信息分析结果。Bystorm&Jarvelin 模型[14]认为开展信息行为前应梳理信息研究对象并判断其有效性[15]。邹益民等[16]构建了多对象在多维度上关系描述的情报关注模型，采用网状结构，将动态监测类情报任务的领域行为、机构行为等要素进行关联。张晓林等[17]将专利技术情报划分为宏观、中观和微观三个层次，采用不同的方法和计量指标对专利信息类型分群研究，构建了包含决策层、管理层与技术层的专利技术情报分析模型。在国防领域，美军《对军事行动的联合及国家情报支援条令》提出了围绕任务内容展开的情报分析模型，包括计划与指导、收集、加工与利用、分析与产出、传播与整合[18]。张静[19]借鉴目标关系模型法将任务相关的要素按照相关关系建立模型，并对南海冲突的可能性进行了战略情报分析。雷帅等[20]提出了由需求理解（R）、数据挖掘（D）、综合研判（J）和服务反馈（F）循环的国防科技战略情报研究方法，并在此基础上，借鉴“物理—事理—人理”（WSR）系统方法论，构建了基于 WSR-RDJF 的国防科技战略情报研究方法框架。面向内容的情报分析模型注重任务内容的结构化、层次化分析，将任务内容拆解为可执行的模块，指导性与实操性更强，填补了面向流程的情报分析模型的缺失。

1.3 综合集成型的情报分析模型

综合集成型的情报分析模型是指操作流程与内容要素并重的情报分析模型。在实际应用层面的情报分析模型中，大多需要情报分析工作流程与情报任务内容相整合以达到情报分析目标需求。John Salerno[21]在 JDL 数据融合模型和态势感知模型基础上建立了战略态势认知框架，并将框架应用与海湾战争的预警中。张超星[22]结合学科发展态势分析工作流程，构建了包含数据维、载体维、时间维、方法维及对象维的五维模型，总结了学科领域发展态势分析的一般性模型。汪雅君等[23] 借鉴构件技术，以霍尔三维模型与生命周期理论为基础，从情报分析内容和情报工作生命周期两个方面构建了包含构成要素、工作流程与关系构建综合型情报分析模型。同年，黄松平[24]等结合指挥控制内部态势认知流程构建了包含信息收集、态势推理、态势预判、战场影响等步骤的态势认知模型。在安全领域，徐方廷[25]从空间融合视角出发，构建由安全态势数据采集层、觉察层、理解层、投射层和预警层构成的安全态势感知模型。在国防领域，王莹莉[26]等提出了基于国防经费的投向投量数据的作战体系发展趋势预测模型，构建了包含作战概念、作战体系、作战能力、项目技术研发的知识图谱，通过聚合国防经费中的装备采购与项目数据，运用统计分析、数学模型等方法预测未来作战体系及发展方向。吕学志等[27]根据战役态势的特点对战役态势要素进行梳理，并围绕战役态势要素构建了包含态势感知、态势理解、态势预测的战役态势认知框架。综合集成型情报分析

模型集成了流程模型与内容模型的优点，结合情报工作分析流程与情报分析内容开展研究，使情报分析工作趋向系统化与工程化。较于面向流程或内容的情报分析模型更具全面性与实操性。

针对国防科技重点领域的发展态势分析大多采用定性定量方法，对领域研究对象如科学文献、专利、项目、经费预算、战略规划等进行分群分析，少有全面综合的领域发展态势情报分析模型。在大数据环境下，情报分析的属性要素由海量多源异构的数据组成，需要利用系统思维，将复杂问题具体化、结构化，全面了解研究实体内部的复杂关系及实体处在外部环境下所扮演的角色与发挥的作用，整合分析各项要素，从而实现更完整更准确的国防科技重点领域情报分析过程。

2 国防科技重点领域发展态势情报分析模型的构建

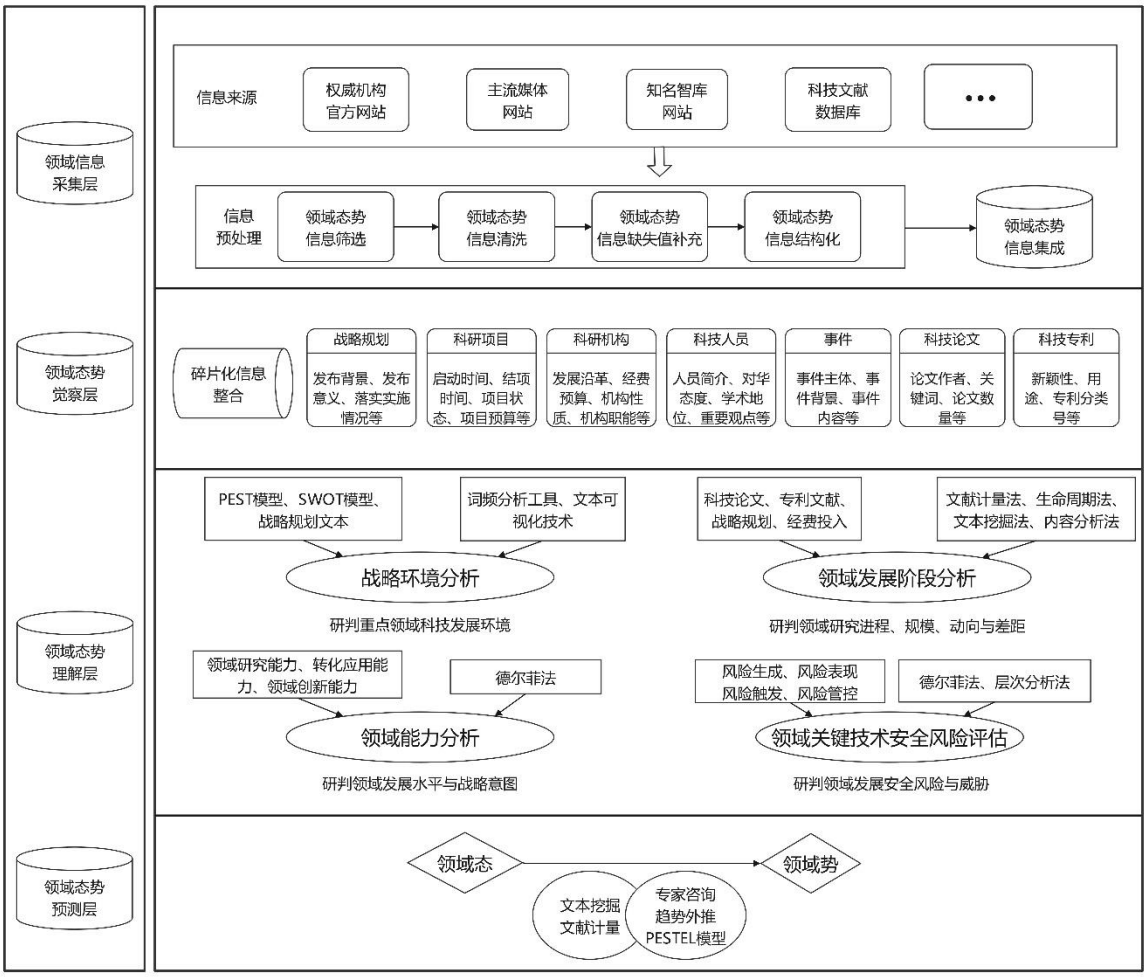
态势是指“一个人在某一时刻对一系列活动的世界观”[27]，最早用于描述战场空间中敌我及战场环境当前状态和变化趋势[28]，国防科技重点领域是指对国家安全、军事发展事关重要的前沿学科领域。本文将国防科技重点领域发展态势情报分析定义为研究与国家安全和军事发展相关学科领域的当前发展现状并预测其未来变化趋势。国防科技重点领域发展态势由政治、经济、技术、社会等多重因素相互影响共同形成。

1988 年，美国心理学家 Endsley 在分析飞行员与领航员的感知程度对其活动行为形成的影响时，提出并将态势感知定义为“在某时间空间条件下对环境中各要素的感知与理解，并进一步预知这些要素的后续变化情况”[2]。1995 年，他将态势感知抽象为三个层级：①第一级是对环境要素的感知，即感知相关要素的状态、属性和动态信息；②第二级是对现状的理解，即综合第一级要素知识及其间关联关系，形成环境的整体图景；③第三级是对未来状态的预测，即通过理解环境要素的状态和变化预测环境要中要素未来行为的能力[29]。因此，态势感知不仅仅是基于对环境中要素信息的简单感知，它包括综合理解要素信息的含义及其间关联关系，并提供对决策有价值的环境未来状态预测。这就是 Endsley[2, 29]的态势感知理论。该理论与本研究探索国防科技重点领域发展态势的研究目标相符合，能够指导国防科技重点领域发展态势情报分析模型的构建。

因此，本研究借鉴 Endsley 的态势感知理论与方法，构建了由四个层次组成的国防科技重点领域发展态势情报分析模型，首先通过开源渠道收集领域相关数据信息；其次厘清国防科技重点研究对象并获取属性信息；然后通过分析研究对象及属性信息间关联关系，获知国防科技重点领域发展“态”；最后根据国防科技重点领域发展规律合理预测国防科技重点领域发展“势”，从而揭示国防科技重点领域整体发展态势。

2.1 模型框架

本研究从国防科技重点领域发展态势情报分析的信息采集、态势觉察、态势理解和态势预测等



四项任务内容出发，构建了由信息采集层、态势觉察层、态势理解层、态势预测层四部分组成的国防科技重点领域发展态势情报分析模型。其中，信息采集层是收集领域态势信息的过程，态势觉察层表征领域发展状态的所有主观与客观信息的整合，态势理解层是基于领域整体态势进行合理评估形成认知的过程，态势预测层是根据领域研究对象及其属性的动态演变预测其发展趋势的过程。上述四层是循环往复、不断迭代的过程，可根据态势任务需求进行改变或叠加，开展综合研判。具体模型框架见下图。

图 国防科技重点领域发展态势情报分析模型框架

2.2 各层功能

面向国防科技重点领域的发展态势情报分析模型架构主要包括信息采集、态势觉察、态势理解、态势预测等四个层次，其具体功能如下：

2.2.1 信息采集层

国防科技重点领域发展态势情报分析模型的信息采集层主要从互联网、文献数据库等多种信息源，获取领域发展相关信息，并进行信息清洗和规范化处理。在信息采集过程中，需要关注与情报分析任务密切相关的信息，也需关注与情报分析任务产生背景相关的信息[30]，主要包括战略规划计

划、预算文件、智库报告、期刊论文、专利文献、动态新闻等类型的信息，其主要来源包括：①权威机构网站：政府部门网站、科研机构网站、企业网站、国家实验室网站等；②主流媒体网站；③知名智库网站；④科技文献数据库。在信息处理过程中，由于国防科技领域的信息存在拒止与迷雾现象，需要进行信息预处理，进行筛选、清洗、补充、结构化、集成等，并标准化为可抽取的统一信息格式，为挖掘信息的价值、实现多源信息的有效存储及高效利用奠定基础。

2.2.2 态势觉察层

国防科技重点领域发展态势情报分析模型的态势觉察层主要根据国防科技重点领域发展态势情报分析的要素，对采集的战略规划、项目、人员、机构、事件等领域研究对象及其属性信息进行碎片化整合，生成经情报人员初步加工的重点领域研究数据集，形成对国防科技重点领域发展的系统全面的基本认识，作为后续重点领域态势研判的重要基础。

国防科技重点领域发展态势研究对象	**属性**
科技战略规划	主要包括科技战略规划的发布背景（政治环境、军事环境、经济环境、技术环境等）、编写机构、编写人员、文件版次、版次调整情况、文件层级、文件类型、上层战略规划、发布意义、重点内容（战略布局、战略重点、战略措施）、落实情况等。
科研项目	主要包括科研项目的启动时间、结项时间、项目状态、项目预算情况（预算总额、投入情况）、主管机构、承研机构、项目管理人员、项目研究人员、所属军种、项目背景、研究内容、研究目标、主要进展、应用前景、项目主要成果。
科研机构	主要包括科技机构的发展沿革、机构性质、机构类型、机构职能、机构规模、机构年度预算、资助情况、代表性成果、荣誉奖励、社会评议及影响力、重大科研活动、科技人才队伍、主要科研设施等。
科技人员	主要包括领域相关重要人员的基本情况包括人员简介（教育经历、工作经历、获得荣誉奖励）、族裔、人员类别、技术关注重点、在职单位及职务、兴趣爱好、对战争态度、对华合作态度、人格特征、学术地位、主要学术成果、重要观点、社会评议及影响力、涉华活动轨迹、参与政策改革情况（重点内容、与现有政策契合度、产生影响）等。
事件	主要包括事件发生时间、发生地点、事件主体、事件类别、事件背景、事件内容、事件结果、产生的社会评议及影响、
科技论文	主要包括论文作者、关键词、论文数量、引文关系、合作关系、发

	文机构、发文作者、科研影响力、增长趋势等。
科技专利	主要包括专利权人、发明人、申请日期、申请号、公开日期、公开号、专利摘要、新颖性、用途、技术要点、专利分类号、法律状态、同族专利、权力要求、引用情况等。

表 2.1 国防科技重点领域研究对象的主要属性

2.2.3 态势理解层

国防科技重点领域发展态势情报分析模型的态势理解层是在领域态势觉察的基础上，利用内容分析法、文本挖掘法、德尔菲法、文献计量法、关联分析等定性定量方法，从领域发展战略环境、领域发展阶段、领域能力分析、领域关键技术安全风险评估等视角，对领域研究数据集进行分析与研判，形成对重点领域发展现状的深度认知。

（1）战略环境分析

领域发展战略环境分析主要包括分析重点领域现有资源与政策，识别重点领域发展的问题与威胁。可采用宏观环境分析 PEST 模型，通过系统梳理国防科技重点领域的政策、战略、规划、计划等研究对象，对其属性要素进行关联分析，厘清影响其发展的政治因素、经济因素、社会文化因素、技术因素等，评估影响重点领域发展的宏观环境；可采用态势分析（SWOT 模型），从优势、劣势、机遇、威胁等四个维度分析重点领域的科技发展环境、战略威胁等；可采用词频分析工具、文本可视化等技术，对重点领域战略规划及相关政策实现定量化分析研判，识别国防科技重点领域相关战略政策文本中领域的发展重点及热点方向。

（2）领域发展阶段分析

领域发展阶段分析是根据历史数据及现有数据对领域发展过程进行研究，并根据萌芽期、发展期、瓶颈期与成熟期等维度对领域发展进行阶段性划分。可将国防科技重点领域相关的科技论文和专利文献作为研究对象，采用文献计量、生命周期测算等方法，通过对国防科技重点领域文献增长趋势进行定量分析，对重点领域发展阶段进行测算和确定，从宏观角度上把握领域整体研究进程及趋势；可将国防科技重点领域相关战略规划计划、经费投入、项目部署等作为研究对象，采用文本挖掘、内容分析等方法，通过引述战略规划计划中的官方观点、分析经费投入和项目部署等情况，对领域发展进程进行阶段性划分。

（3）领域能力分析

领域能力是领域发展现状的直接反映，基于领域科技政策环境、投向投量、科技机构、科研项目、科技人才规模结构、投入应用实际等信息，采用德尔菲法等方法，可对一个国家或地区的领域研究能力（包括基础研究能力、学术成果水平、科研实验能力）、领域创新能力（技术先进水平、技术成熟度、技术体系完备水平、技术自主可控程度）、转化应用能力（成果转化程度、成果生产能力）等进行分析，以“客观信息+专家智慧”的模式完成领域能力评估。

（4）领域关键技术安全风险评估

国防科技重点领域的关键技术大部分具有典型的两用性特征，对其进行安全风险评估是防患未

然的必要措施。建立合理的领域关键技术安全风险评估监管机制是实现对领域关键技术安全风险监管的重要前提。深入分析国防科技重点领域关键技术的特征，从风险生成、风险触发、风险表现、风险管控等多个维度，对领域关键技术安全风险的影响因素进行分析，提出合理的领域关键技术安全风险评估指标体系，再通过层次分析法、熵值法等确定指标权重，最后借助德尔菲法等为领域关键技术的安全风险进行打分评估。在此基础上，提出应对风险的对策建议。

2.2.4 态势预测层

国防科技重点领域发展态势预测是在态势觉察层和态势理解层的基础上，对领域未来发展的“势”进行预测。态势觉察层与态势理解层描述的是时间t内领域的发展情况，态势预测层则是在此基础上预测领域在时间$t+n$（n代表时域长度）内的发展态势。具体地说，态势预测层就是通过文本挖掘、文献计量等技术预测方法预测潜在的新兴技术等，再结合专家咨询法、趋势外推法、PESTEL 模型等方法综合研判领域的整体发展趋势。

近年来，新兴技术将对经济和社会发展产生革命性影响已成为全球的共识。如何选择最有价值的研究领域和新兴技术进行发展部署，已成为政策制定者和科学家关注的焦点。因此，国防科技重点领域发展态势情报分析模型将态势预测作为模型的最后一层，通过挖掘新兴技术、预见未来趋势，占据科技发展的制高点。

3 总结与展望

“情报”的本质在于还事实以真相，国防科技领域发展态势情报研究的本质在于还原国防科技重点领域发展的规律与事实，为我国防科技事业提供发展依据。当前关于国防科技重点领域发展态势情报分析的研究主要集中在具体模型、方法、技术和工具等微观层面，缺乏从体系框架的宏观视角进行研究，对研究实践的指导性和牵引性也存在不足。本研究遵循情报工作“数据-信息-知识-情报”的转化链条，借鉴 Endsley[29]态势感知的理论方法，融合结构化、系统化思维，在系统总结情报分析工作中所用的流程与方法的基础上，构建了国防科技重点领域发展态势情报分析模型。该模型是一种流程与内容并重的综合集成型情报分析模型，由信息采集层、态势觉察层、态势理解层和态势预测层等四层构成，具有较广泛的适用性，为当前高度关注的人工智能、量子科学、生物技术等国防科技重点领域发展态势分析提供了一种可操作性较强的技术路线和方法。

该模型主要具有以下特点：一是情报分析要素多源化。从信息采集的角度上看，该模型从互联网、文献数据库等多种信息渠道获取战略规划计划、预算文件、智库报告等多种类型数据，能满足不同国防科技重点领域态势分析的普遍性需求。二是情报分析方法集成化。该模型结合多源数据，综合使用文献计量法、统计分析法、专家调查法、引文分析法等多种定性、定量方法进行情报问题研究，从领域发展战略环境、领域发展阶段、领域能力分析、领域关键技术安全风险评估等视角，形成对领域发展现状不同维度的认知。三是情报分析结构化。该模型将国防科技重点领域发展态势分析分解为信息采集、态势觉察、态势理解、态势预测等四个流程，并针对战略规划、项目、人员、机构等领域研究对象及其属性的不同特点，分别确定了上述研究对象的结构化分析维度。此外，本研究构建的国防科技重点领域发展态势分析模型基于开源情报信息而设计，但国防科技情报信息具有较强的隐性特征，不易获取齐全，这对情报分析结果会产生一定影响，需要在今后的研究中继续加以深入研究。

PO-016

基于 PubMedBert + TextRNN 的肿瘤文献分类模型构建研究

张颖，李晓瑛，刘懿，杨雪梅，李爱花，唐小利*
中国医学科学院北京协和医学院医学信息研究所

1 引言

随着社会发展与人民生活方式的改变，慢性非传染性疾病已逐渐成为人民所面对的主要疾病负担。以肿瘤、糖尿病、心脑血管疾病、慢性呼吸系统疾病为代表的慢性非传染性疾病给我国人民造成了 70%以上的疾病负担，已发展为阻碍人民健康发展的重要因素[错误!未找到引用源。]。恶性肿瘤在我国的发病率及死亡率呈现逐年上升的趋势。世界卫生组织国际肿瘤研究机构（IARC）提供的数据显示，2020 年全球新发肿瘤病例 1929 万例，其中中国新发肿瘤 457 万人，占全球新发肿瘤总量的 23.7%。肿瘤作为致死率最高的疾病之一，不仅严重威胁到人民生命健康，还给社会带来极大的经济压力[错误!未找到引用源。]。

据美国国家科学基金会统计，科研人员花费在查资料上的时间需占全部科研时间的 51%[错误!未找到引用源。]。肿瘤相关的研究热度和公众关注度一直处于较高水平，肿瘤领域科研人员在从事科研工作时面临着更加严峻的数据处理、信息过滤、知识捕获困难等问题[错误!未找到引用源。]。文献分类是一种有效的知识组织方法，系统精细的文本分类可以通过相应前台服务缓解科研人员的目标文献获取压力，而当前常见的期刊级别的文献分类方法，难以满足数据规模迅速增加以及“交叉学科”带来的细粒度文献分类需求，如何有效辅助肿瘤领域科研人员提高科研效率等问题亟需解决。

为此，本文构建了一种基于“PubMedBert + TextRNN”的针对篇级文献的细粒度单层文献分类模型，并选取前期系列研究构建的“Bert + TextRNN”，“BioBert + TextRNN”进行基线分析，发现“PubMedBert + TextRNN”模型在准确率、召回率与 F1 值上都表现优异。

2 相关研究

Bert(Bidirectional Encoder Representation)通过引入注意机制而成为一种新的语言表征模型，被广泛应用于机器翻译[错误!未找到引用源。]、图像描述生成[错误!未找到引用源。]、机器阅读理解[错误!未找到引用源。]和文本分类[错误!未找到引用源。]等领域。目前，Bert 在自然语言处理尤其是文本分类中有重大进展。，Shen 尝试训练基于中文语料库的 BERTbase 模型，用于中国社会科学、科技期刊的文献分类，并探索其在实际生产中的应用[错误!未找到引用源。]。此外，Lu 利用 Bert 加 CNN 的组合得到了一个用于专利分类的多层模型[错误!未找到引用源。]。然而，将通用语言模型直接应用于生物医学领域的准确率并不高，这促使相关研究从头开始训练生物医学 Bert。典型的例子是在 PubMed 引文和 PMC 全文上进行预训练的 BioBert[错误!未找到引用源。]，以及使用 PubMed 文本和临床笔记进行混合域预训练得到的 PubMedBert[错误!未找到引用源。]。

多标签分类作为学术界和工业界长期面临的挑战之一，一直是众多研究人员的工作重点，深度学习模型为多标签分类作出了重要贡献。在生物医学领域，Du 等人提出了一种针对生物医学文本的端到端深度学习模型 ML-Net[13]，而 Glinka 等人则专注于特征选择方法、过滤器和包装方法的混合[14]。此外，Hughes 等人尝试基于 CNN 对医疗文本片段进行句子级分类[15]，Yogarajan 等人使用医疗文本分类的多标签变体来增强并发医疗代码的预测[16]。Wasim 等人针对生物医学问答系统中的事实类问题和列表类问题提出了一个多标签问题分类模型[17]。同样，Baumel 等人提出了一种分层注意力双向门控循环单元，利用注意力权重更好地理解对决策影响最大的句子和单词[18]。本文提出一种基于预训练的方法，训练了一种分类模型，对巨大体量的肿瘤出版物进行分类，以便更好地进行检索、学术分析和统计评估。

3.方法

本文训练了一种基于“PubMedBert + TextRNN”的多标签分类器，可以实现对肿瘤文献的多标签篇级分类。研究的总体框架如图 1 所示。首先，本文从 DIMENSIONS 数据库中获得标题和摘要组成的语料集，经过预处理后按 7:3 的比例分为训练子集和测试子集。随后，本文利用训练集对文献进行训练。由于摘要和标题具有不同的语义特征，比如，摘要中蕴含的信息可以存在于上下文关联中，而标题中的信息更多来自于关键词，为了获取更多的文本语义特征，本文在模型中分别设置标题层和摘要层，以实现文献的标题和摘要的独立训练。最后，利用测试集将本文提出的模型与我们之前训练的基线模型进行对比，发现本文提出的模型在各项指标值上都具备显著优势。

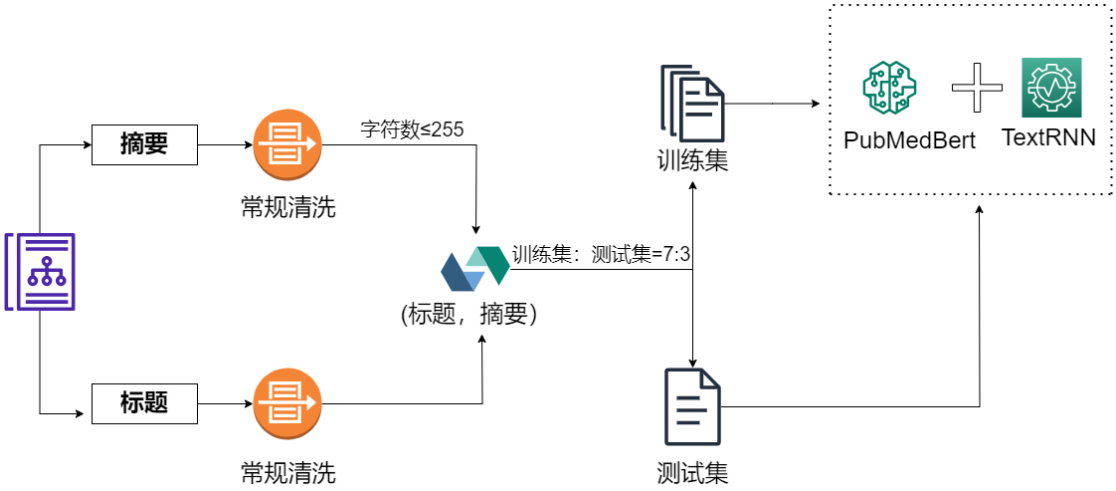


图 1 方法框架图

3.1 数据获取与预处理

本文的数据来源于 Digital Science & Research Solutions 开发的 DIMENSIONS 数据库，该数据库拥有超过 1.35 亿份出版物、650 万份基金和 1.53 亿份专利。DIMENSIONS 数据库对肿瘤文本分类采用国际肿瘤研究合作组织（International Cancer Research Partnership，ICRP）维护的分类术语体系——“Cancer Types，ICPR CT”，ICPR CT 用 62 种肿瘤类型来表征单篇肿瘤文献，并已与世界卫生组织的国际疾病分类(International Classification of Diseases ICD)关联映射[错误!未找到引用源.]，在 DIMENSIONS 数据库中用于对肿瘤文献进行细粒度标注。

3.1.1 语料库构建与预处理

以“肿瘤”为检索词，时间设定为“2003 年 1 月 1 日——2022 年 5 月 1 日”在 DIMENSIONS 数据库中进行检索，获得了 70,599 篇带有 ICRP CT 标签的文献。本文根据语料库中肿瘤文献的数量从高到低排序绘制了图 2，可以看出语料库中肿瘤文献分布极为不均匀，其中排名前 9 位的肿瘤总量（Breast cancer, Not Site-Specific Cancer, Colon and Rectal Cancer, Lung Cancer, Prostate cancer Ovarian cancer, Stomach cancer, Cervical cancer, Pancreatic cancer）占整个语料库的 76.35%（53,899/70,599），占据了语料库中 3/4。如果不对原始语料进行处理直接训练，在训练过程中容易产生过拟合现象，在应用过程中极大降低模型的准确率。

本文对原始语料进行均衡采样。采样过程如下，首先，将语料库各类别中包含文献数量按照从多到少排列。结合实际情况确定阈值（70），汇总文献数大于阈值（70）的类别计算相关指数（平均值、中位数、方差、标准差）见表 1。随后，根据抽样数据的实际分布情况可以看出，平均数、方差、标准差难以代表语料分布的真实情况，为此选择了中位数作为初始指数。考虑到文献以 7:3 分为训练集和测试集，为保证有充足的语料用于训练，本文将最终指数设定为 500。最后，对不同类别的文献数量均衡处理，即大于最终指数的文献类别仅保留 500 篇文献，小于 500 篇文献

的类别保留全部文献。

表 1 原始数据统计表

指标名称	初始指数
平均数	1,127
中位数	282
方差	2,427
标准差	5,889,927

3.1.2 二元组构建

本文所采集到语料由肿瘤文献的标题和摘要组成，二者所包含的训练重点有差异。标题是论文的标题,居于论文的首要、醒目位置,起着概括全文、揭示主旨的作用，用简洁的语言概括文章主旨。摘要是研究工作的主要对象和范围，采用的手段和方法，得出的结果和重要的结论，有时也包括具有情报价值的其它重要的信息。标题更多是对关键词的训练，而摘要涉及到上下文关联的训练，为此本文将标题和摘要分别输入模型的“标题层”和“摘要层”，以实现在训练时侧重于对应语料的特点，在后文中称之为二元组。

3.1.3 模型设计与构建

本文训练的基于“PubMedBert + TextRNN”的模型由上游模型 PubMedBert 与下游模型 TextRNN 构成。预训练一指将大量低成本收集的训练数据放在一起，学习数据的共性，然后用少量特定领域的标注数据来调整具有共性的模型。本文采用的 PubMedBert，使用 PubMed 生物医学数据库直接进行训练，利用 Masked Language Modeling（MLM）与 Next Sentence Predication（NSP）获取更为全面的语义信息[错误!未找到引用源。]。下游采用 TextRNN，该模型是 RNN 的变形，采用递归方式通过单向或双向循环神经网络将前一个时间步的输出作为当前时间步的输入，以此更好地捕获文本中的上下文信息。

该模型将设置标题层和摘要层，训练时将预处理后语料库中的标题和摘要分别输入，通过 PubMedBert 将输入语料转化为词向量并传递至 Encoder 层，随后标题层和摘要层输出的词向量将拼接在一起，然后传输到全连接层进行归一化处理，最后利用 sigmoid 生成多标签分类最终结果。

本文将语料库中 30%的语料作为测试集，测试集中的文献涵盖 ICRP CT 中所有类别，评估我们训练的分类器的性能，测试时基于准确率、召回率与 F1 值对模型性能进行验证。其中，F1 值是准确率和召回率的调和平均值，F1 值越大，分类模型性能越好。同时，本文根据前期已有研究，选择 Bert + TextRNN, BioBert + TextRNN 作为基线模型进行对比测试。

4 结果

本文把基于“PubMedBert + TextRNN”的模型与基线模型（Bert + TextRNN，BioBert + TextRNN）测试结果在同一测试集上进行对比（表 2）。测试结果表明“PubMedBert + TextRNN”在所有测试指标方面都拥有最优表现，准确率 95.20%，召回率为 93.67%，F1-score 为 93.17%。基于“PubMedBert + TextRNN”的模型参数见表 3.

表 2 模型指标对比

模型	准确率（%）	召回率（%）	F_1-score（%）
PubMedBERT + TextRNN	95.20	93.67	93.17
Bert + TextRNN	93.09	87.75	90.34

BioBert + TextRNN	88.32	87.66	86.34

表 3 模型参数

Parameters	Value
Num_ train_ epochs	100.0
Max_ seq_ length	256
leaning_ rate	0.0001
train_ batch_ size	32
Predict_ batch_ size	32
TextRNN	256*2
LTSM_UNITS	5
BERT_OUTDIM	768

5 讨论

5.1 性能优异性分析

PubMedBert + TextRNN 在肿瘤文本分类中相比我们前期训练的基线模型（Bert + TextRNN，BioBert + TextRNN）获得更优异的性能主要有以下几个原因。

（1）BioBert 首先利用 Bert 在通用语料库（Wikipedia、BooksCorpus）进行训练，而后在生物医学领域语料库（PubMed Abstract、PMC 全文）进行预训练；而 PubMedBert 模型从头开始在 PubMed 数据库中的生物医学文本上训练。在生物医学领域的自然语言应用中，PubMedBert 的效果远远领先于之前所有的语言模型。

（2）肿瘤领域生物医学文本是知识密集型、高价值文本，具有大量生物医学领域名词与表达，领域性 PubMedBert 比 Bert 更适用，采用领域性预训练模型可以生成更合适策略。

（3）BioBERT 在生物医学命名实体识别、生物医学关系提取、和生物医学问题回答三种具有代表性的生物医学文本挖掘任务上的性能更具优势，但是文献分类任务重并未具备显著优势。

（4）利用预训练与传统深度学习模型强化配合训练新模型，可以提高训练效率，从多方面多角度增强训练效果，相比单一使用 TextRNN 相比，本文提出的模型从多效果上显现出优势。

5.2 局限性分析

基于“PubMedBert + TextRNN”模型提出的分类模型存在三个问题。第一，虽然该种结合的强化模式匹配模型增强了训练效果，但是由于关键词提取造成的误差积累，此种缺陷本文提出的模型通过调整模型参数和增加自检功能实现了提高。第二，本文使用的数据较为单一，仅限于文献文本，在实际使用上非常受限，目前未使用电子病历、多媒体数据等进行训练，后续，我们将不断改进，不断拓展数据来源于数据类型，增强模型的功能。第三，以 GPT 为代表的大模型以优异的成绩为知识组织提出了新的解决路径，当前，生物医学领域的大型语言模型逐渐成熟，BioGPT、PubMedGPT、草本等大语言模型都拥有了丰富的应用实践，基于此开展分类同时研发衍生功能是辅助解决生物医学领域分类问题的新途径，本文未对此进行对比分析研究，日后将利用大模型展开肿瘤领域文本分类研究，提高研究的利用率。

6 结语

本文基于"PubMedBert + TextRNN"构建了一个多标签分类模型。该细粒度分类模型在肿瘤领域文献的分类领域是有效的。在未来，我们将尝试利用大模型扩大研究深度，同时把现有研究结果逐渐

扩展到具有高度专业化、统一实体名词、标准化长文本等特征的其他领域文献分类中（心血管疾病、糖尿病），辅助实现文献的细粒度分类，提高科学数据的使用效率。

PO-017

新医科背景下医学信息专业虚拟教研室建设的思考与探索

常青、喻慧心、鲁兴胜、孙雪超、赵玉虹*
中中国医科大学附属盛京医院临床研究中心、辽宁省智能诊疗生态系统工程研究中心

1 引言

当前，新技术、新产业、新模式不断涌现，大数据、人工智能等前沿科技在医学、教育等多领域飞速发展[1-2]。2019 年，教育部启动实施“六卓越一拔尖”计划 2.0，提出“四新建设”（新工科、新医科、新农科、新文科），其中新医科建设是健康中国战略下医学教育发展的最新要求，也是培养医学创新人才的重要组成部分[3]。新医科打破专注于治疗的传统医学服务模式，形成生命健康全周期医学服务新格局[4]；将医学专业创新人才培养从“医学模式”转变为“医学+X”交叉学科模式[5]；通过推进医工理文融通，将传统医学与大数据、人工智能等进行融合[6]，推进医学信息学专业向新的发展方向转变**错误!未找到引用源。**。在此背景下，如何创新发展医学信息教育，重构与新医科发展相适应的新型基层教学组织，促进医学人才信息能力提升，提升医学信息人才培养质量，对新医科建设具有重要意义。

2021 年，教育部出台《关于开展虚拟教研室试点建设工作的通知》，指出虚拟教研室是信息化时代新型基层教学组织的重要探索[7]，具有跨界性、多样性、灵活性、开放性等特点。在国外，麻省理工学院于 2001 年最先提出“开放课件运动”，标志着高校网络课程的兴起[8]。这一举措拓宽了传统课堂的时空界限,将大学教学从教授的个体行为转变为跨校的群体合作,创造了网络资源课程建设和运维新模式,奠定慕课教学发展的基础。在国内，随着信息化建设的蓬勃发展，我国高校教研室在教学研究领域不断创新探索，借助新型的科学技术和双线教育模式[9]，通过开展线上线下、跨越时空、虚实结合的课堂教学实践和教学研究活动，打造精品教学资源库、优质教师培训资源库等，促进多学科交叉、多技术融合，实践互联互通、共建共享的虚拟教研室建设。

中国医科大学附属盛京医院于 2021 年开始建设“智能+医学信息与大数据虚拟教研室”，在虚拟教研室模式构建、课程群建设、教学模式创新、教学资源建设、师资协作模式等方面进行实践探索。本研究重点阐述“智能+医学信息与大数据虚拟教研室”实践建设经验，并就虚拟教研室建设的重点和瓶颈问题提出思考与建议。

2 医学信息专业虚拟教研室建设实践

“智能+医学信息与大数据”虚拟教研室是以中国医科大学健康数据科学教研室、中国医科大学附属盛京医院临床流行病学教研室为依托，联合中国医学科学院医学信息研究所、中国科学院北京基因组研究所、南通大学数字医学研究所、华中科技大学、东软集团等全国领先科研院所、高校、企业多学科专家合作构建的课程群教学类虚拟教研室。

2.1 夯实实体教研室基础，支撑虚拟教研室建设

高校基层实体教研室是虚拟教研室建设的基础[8]。中国医科大学健康数据科学教研室、中国医科大学附属盛京医院临床流行病学教研室建设可追溯于 1987 年，从信息、数据、证据出发进行知识发现和教学研究工作，协同培养医学生信息素养和医学创新能力。教研室医学信息学专业课程与教材建设紧跟信息科技变革,实现从传统文献检索到生物信息挖掘、医学大数据研究与实践等方向转型。教研室主干课程《生物医学文献与网络资源》为国家级精品网络资源共享课，在此基础上，主编《医学文献检索》、《卫生信息系统与项目管理》、《卫生信息学导论》等多部教材资源。此外，依托国家重点研发计划项目,教研团队已建成生物医学大数据研究平台和精品资源课程在线平台,创新应用信息分析挖掘、数据分析软件等辅助教学工具，为开展多渠道线上线下、虚拟现实结合的交互式学习模式，为医工结合的信息化教学场景实现奠定基础。

2.2 医教研产多学科协作，构建虚拟教研联合体

教研室汇聚校内相关学院的优势资源,联合多家高校、科研院所和企业,建立医教研产一体化的医学信息与大数据教育协同发展联盟,实现医学信息学、临床医学、预防医学、基因组学、计算机科学等多学科融合，形成学科交叉、医工结合创新型虚拟教研联合体，共同开展虚拟教研室课程群建设、教学改革研究、系列教材及数字教学资源建设和师资培训。将相关企业纳入虚拟教研室联合单位，致力于将基础教学成果迅速，形成高校、科研院所与产业界的成果转化链，提升人才培养成果的临床应用性、产业化前景。

2.3 新医科课程群建设，创新教学内容

教研室根据新医科建设目标，以科学方法重新定位新时代医学人才医学信息胜任力需求，将医学人才信息胜任力模型内涵转化为可测量可评价的培养目标和评价标准（milestone)。基于该标准，针对不同类型教学对象设计教研室核心教学课程。针对医学生，根据不同培养阶段，设定课程教学目标，构建以生物医学信息与网络资源、循证医学与临床研究方法、卫生信息系统与健康医疗大数据等为主干课程的进阶式、一体化课程体系，培养医学生医学信息获取与大数据处理能力、临床思维与决策能力及医学创新能力。

此外，针对医学信息专业学生，虚拟教研室从基础理论-专业思维-创新实践出发，以健康大数据为切入点，以健康大数据产业和生物医药行业发展需求为导向，以培养具备医学信息专业素养和行业能力的交叉人才为目标建设，创新建设“数据驱动的医学创新与决策”课程群（见图 1），实现教学内容改革，提升人才培养质量。

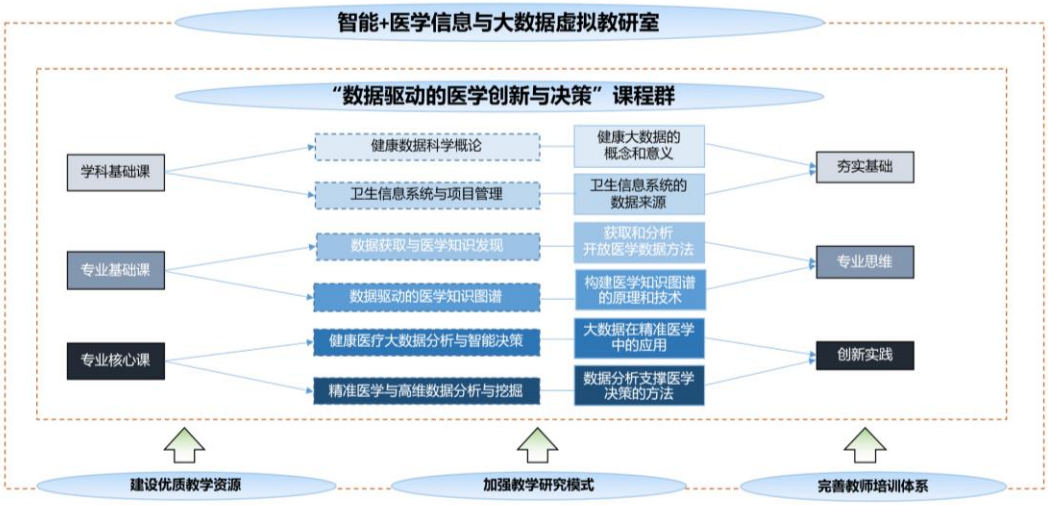


图 1 “智能+医学信息与大数据”虚拟教研室课程群

2.4 教学模式创新，完善教学资源建设

教研室充分运用团队信息技术优势，建立多渠道“线上+线下”相结合的教研模式，提供丰富的医工结合信息化场景实现。强调以学生为中心，综合使用在线课程、电子学档（e-portfolio）、网络学习社区和辅助教学工具开发相结合的方法，提高学生反思学习和自导学习能力，提升教学可及性。此外，采用研究型教学形态，以项目为牵引，引导医学生利用医学信息及数据解决实际问题，促进教学成果的实践转化。此外，教研室基于现有慕课、网络学习平台等资源，建设群智协同虚拟教研室平台，构建立体共享资源体系。制订标准化教学资源共建共享机制，形成教学大纲、教学案例、教学视频、教学课件、习题试题等一系列优质标准化、规范化资源库。

2.5 创新师资协作模式，提升教学水平

教师在虚拟教研室建设过程中起着关键的作用。教研室面向全国，建立多机构多学科融合的师资数据库，组建以国家级教学名师为核心的教学团队。教研室建立开放、共享的师资协作与培训模式，定期组织线上线下集中学习、研讨交流，提升师资同质化教学水平。通过加强校企合作,引入权威行业培训资源,拓展师资培训内容。此外，应用教师电子档案，引导教师不断反思教学及培训心得，提升教学水平。

3 虚拟教研室建设的思考与建议

随着对教学研究领域的不断创新与探索，跨学校、跨地域、跨学科的高校虚拟教研室在建设过程中必然会面临各种严峻的挑战。首先，虚拟教研室的搭建及稳定运行离不开网络技术支撑及多样化的课程资源库与虚拟社群；其次，虚拟教研室涉及多个学校、地域和学科的参与者，教研室需在“求同存异”的基础上建立制度规范，促进不同团队的多元化发展；再次，虚拟教研室的建设对教师提出了更高的要求，需要教师积极创新教研方法，开展跨学科合作；最后虚拟教研室对教研成果的创新转化落地提出了更高的要求，以上这些都需要在建设过程中的积极思考和不断完善。良好的新

型基层教学组织建设需要依托原有教学基础，进行严谨的教学、教研活动系统工程设计并付诸实践，并在实践中接受检验，不断修改、完善。

3.1 虚拟教研室需要明确的教学目标

虚拟教研室建设需有明确、具体的目标追求和价值导向，且目标需要做层次化分解。从“教”向“学”转型，向校企合作共同创新型师生学习团队的转型，不仅聚焦教学内容的创新研究，更重要的是深入开展对学生创造性学习特点和规律的研究。

3.2 智能化平台是虚拟教研室的重要技术支撑

信息化教学创新和智能化教学平台是虚拟教研室的基本生存环境和发展空间，智能化教研平台能够支撑不断创新的技术集成、线上线下结合的学习与教学创新的集成以及网络学习与教学服务创新的集成，为教研室组织、教学、评价、资源建设等多方面提供支撑。

3.3 开展深层次校企合作的资源建设与人才培养

虚拟教研室需要开设跨时空的网络课程，技术平台功能复杂，是传统教学模式没有的课程运维，教育企业往往有很强的技术研发和项目管理服务团队，进行校企深度合作，可以弥补教师技术、课程运维服务的短板。

3.4 智能化评价模式和体制创新是虚拟教研室的发展支柱和政策空间

虚拟教研室建设容易受高校现行评价模式和教学、教师管理模式的制约。创新型教学形态评价体系和课程教学管理体制机制将促进虚拟教研室的稳定和可持续发展，有效打破学校和区域的壁垒、科研和教学的壁垒，保证教学创新系统沿着正确的目标发展。

4 结语

尽管面临着困难与挑战，但高校虚拟教研室建设仍是具有全局意义的重要战略决策和系统工程。中国医科大学附属盛京医院“智能+医学信息与大数据虚拟教研室”通过建立医教研产群智协同的虚拟教研平台，致力于打破校际与区域间、科研与教学间、企业与用户间、技术工具应用与培训间的壁垒，推动以课程群为载体的协同教研，提升课程的建设质量，促进学科交叉融合，助推医学信息人才培养模式更新，切实践行新医科背景下推进高等医学教育教学改革纵深发展，有利于达成新医科多渠道、多元化育人目标。

PO-018

数智时代医学信息管理类本科人才核心素养培养框架研究——以重庆医科大学信息管理与信息系统专业为例

侯筱蓉* 赵文龙 贺向前 黄成 吕亚兰
重庆医科大学医学信息学院

1 数智时代对医学信息管理类人才培养的新要求

成果导向教育（Outcome-Based Education）观点认为，提倡以学习者最终掌握的技能、取得的成绩为根本，反向设计学习者的培养结构及评价体系，是一种基于学习成果或结果的教育理念。[1]结合数智时代对人才培养提出的新要求，医学信息管理类专业的人才培养体系培养也可考虑采用反向设计的方法，根据所在院校的学科背景和专业发展历史，从技术的发展和产业结构的升级出发，

倒推人才培养计划的制定和课程体系的设置，以促进专业人才核心竞争力的形成。

数智时代处于信息化技术发展的第三个阶段，以大数据、人工智能为主要技术特征，主要体现在技术思维、技术产品和技术应用上的全面创新[1]。其本质是将过去的小数据转变为现在的大数据，并进一步发展为面向未来的智慧数据，以推动数据治理和智慧治理的融合[2]。医学信息管理类专业的发展，离不开技术的赋能，面临着大数据、人工智能等一系列新的技术、方法和模型的全面介入，这对该类专业人才的培养提出了新的要求。

2 医学管理类专业人才培养框架搭建

本科专业教育，要体现“基础-主流-发展”的关系结构，也要注重知识更新和实践应用[3]。“中国高等院校信息系统学科课程体系”课题组先后制定了“中国高等院校信息系统学科课程体系2005（CISC2005）”和“中国高等院校信息系统学科课程体系 2011（CISC2011）”，2018 年教育部发布了《普通高等学校本科专业类教学质量国家标准》，这些文件均强调应重点培养信息管理和信息系统（以下简称“信管”）专业学生的信息技术能力、信息分析能力和管理决策能力[4]。美国计算机协会（Association for Computing Machinery, ACM）提出信管专业教学应该包括信息技术、组织与管理及系统开发理论三个部分。

从经典的学科领域特征来看，信管专业同时面向技术、管理等多个层面。当前，技术与管理并重是该领域发展的主流特点。技术学派侧重信息的处理以及系统开发的理论、方法与应用；管理及行为学派侧重信息技术/系统的使用、管理和行为[3]。同时，结合数智时代对该专业提出的新要求，传统的信管课程，逐步升级为数智赋能的信息学理论与方法、多模异构信息的融合与挖掘、智能驱动信息组织与检索、全文本科学计量分析及创新评价、数据驱动的信息行为与人机交互和情报服务模式工程化及智能化等。因此，信管人才的培养，需要对“资源大数据化”“方法智能化”“服务精细化与模式智慧化”这三方面有及时的反映[5]。

结合以上分析，重庆医科大学信息管理与信息系统专业，在最新一轮的人才培养计划修订中，遵循“管理与技术并重”的学科特征，以培养适应新医科建设需求的高级信息管理与信息系统专业人才为目标，做到：培养德、智、体、美、劳全面发展，适应我国卫生事业信息化发展需要，具有医学基础知识，熟悉医疗信息系统管理，掌握医疗健康信息的有序化和标准化组织与管理、疾病分类编码，具有医疗数据智能治理、挖掘分析、辅助决策和利用等能力，具备良好职业素养的智慧医学信息管理人才。针对人才培养目标，搭建了适宜学校学科发展与数智技术发展的专业人才培养框架，见表 1。

2.1 专业基础素养

业基础素养是智慧医学信息管理人才必备的关键能力，它为高阶核心素养的实现提供坚实的支撑。为了培养具备综合素质的医学信息管理专业人才，要求学生在医学基础素养、数学基础素养和计算机基础素养方面具备扎实的知识基础。

（1）医学基础素养

学生需要掌握人体学、临床医学等基本知识，这些知识是理解和应用医学信息管理的基础。学生需要深入了解人体的结构、组织和功能，同时掌握临床医学中的常见疾病和诊断方法，以便能够准确地处理和分析临床医学数据。

（2）数学基础素养

学生需要具备高等代数、线性代数和概率论等数学知识。这些知识是进行数据分析和模型建立的基础，学生需要掌握各种数学方法和技巧，以便能够对医学数据进行准确的统计和推理。

（3）计算机基础素养

学生需要掌握基本的系统设计思维，具备良好的逻辑思维能力和问题解决能力。他们需要学习计算机的基本原理和结构，了解常用的编程语言和工具，以便能够设计和开发医学信息管理系统。

2.2 专业核心素养

专业核心素养是在基础素养培育的基础上，重点培养具备专业核心竞争力的素养代表。这些素养包括数智驱动的管理决策素养、知识组织与数据挖掘素养以及智能技术支持与智慧系统开发素养。

（1）数智驱动的管理决策素养

该素养是指掌握支持组织管理决策的相关理论，能够对业务进行分析，以恰当应用信息技术和信息分析解决管理问题的能力。其中尤其强调数智驱动管理决策（DDD: Data & Intelligence Driven Decision-Making），具体是指依赖大数据和智能化技术做出管理决策[6]。因此除了要求学生掌握经典的管理知识框架之外，还要求学生能够将数据分析和挖掘的结果转化为实际的管理决策，并能够评估和优化这些决策的效果，以应对复杂的管理决策环境。在该素养中，对创新思维能力有较高要求，需要学生具备创新思维和问题解决能力，能够从多个角度思考和分析问题，并提出创新的解决方案。

（2）数据挖掘与知识发现素养

该素养具体是指知识组织、知识获取、数据挖掘和数据分析等一系列技能。在大数据环境下，各领域的知识通过丰富的语义关系形成多领域的数据语义网络。基于人工智能技术的知识获取具有供给特征，使人类在知识获取上表现出颗粒化、分散性、主观性和个体性等特点[7]。学生需要掌握数据特征选择和构建的原理和方法，能够从原始数据中提取有用的特征，并构建新的特征以增强数据的表达能力。同时，需要掌握常见的数据挖掘算法，能够理解算法的原理和应用场景，并能够使用相应的算法进行数据挖掘和知识发现。同时，需要掌握知识管理和知识应用的基本原理和方法，包括知识表示、知识存储和知识检索等，能够将数据挖掘和知识发现的结果转化为实际的知识，并能够有效地管理和应用这些知识。最好的结局是，学生具备数据可视化的能力，能够使用可视化工具和技术将数据挖掘的结果以直观和易于理解的方式展示出来，以支持数据分析和决策制定。

（3）智能技术支持与智慧系统开发素养

该素养是指利用计算机技术实施信息采集、信息建模和数据分析，从而实现管理创新和决策支持[8]。在信息管理类专业中，计算机技术已经成为工具，不局限于计算机系统本身，也不过分强调程序设计和软件系统的开发，与计算机专业形成较为清晰的边界。因此，学生需要具备的能力应该涉及信息系统分析与设计、项目管理、信息安全与风险管理以及医院信息化架构等方面。信息系统分析与设计方面，能够分析组织的信息需求，并设计和实施相应的信息系统，以支持组织的管理创新和决策支持。项目管理方面，能够有效地组织和管理信息系统项目，以确保项目按时、按质地完成，并能够解决项目中的问题和风险。信息安全与风险管理方面，能够评估和管理信息系统的安全风险，并采取相应的安全措施，以保护组织的信息资产和业务运作。医院信息化架构放方面，需要了解信息化架构设计、信息战略规划和信息技术治理等，能够将信息系统与医院业务需求相匹配。

3 支持人才培养框架的课程体系构建

基于 OBE 教育理念培养体系的关键特征之一是制定课程体系（curriculum framework），产生明确具体、可衡量的结果[9]。课程体系设计是高校人才培养模式变革的重要环节，研究本科课程设置，构建与之相适应的课程体系对提升本科人才培养质量有重大推动作用。[10]根据专业人才培养框架的设定，进行了详细的课程体系设计，以分别支持专业基础素养和专业核心素养的达成。具体课程体系如表 1 所示。

在新医科背景下，信息管理类人才的培养目标是培养具备医科背景的专业知识和技能的信息管理专业人才。因此，在专业核心层的课程体系中开设了“病案信息管理”“国际疾病分类”“自然语言处理与电子病例挖掘”“医院信息系统开发”等课程（表 1 中标注*的课程）。“病案信息管理”旨在培养学生掌握病案管理的基本理论和方法，涉及病案管理中的标准化、电子化、信息安全和信息分析等核心问题。“国际疾病分类”旨在培养学生掌握国际疾病分类系统的原理和应用，以支持医疗机构的统计分析、质量评估和医保结算等工作。“自然语言处理与电子病例挖掘”旨在培养学生利用自然语言处理技术和数据挖掘方法进行电子病例的信息提取和分析，以发现潜在的医学知识和支持医学研究和决策。“医院信息系统开发”旨在培养学生掌握医院信息系统开发中涉及的需求分析、系统设计和软件开发等问题，以满足医疗机构的信息化建设需求。

在双语教学方面，在“信息服务与用户研究”“信息检索”“医学信息分析”等课程中进行了双语教学。这是为了培养学生的跨文化沟通能力和国际视野，使他们能够适应全球化的信息管理环境。通过双语教学，学生不仅可以学习到专业知识和技能，还能够提升自己的语言表达能力和跨文化交流能力。

在行业和社会的关联性方面，与相关医疗机构和企业展开了深入合作，了解他们对人才的需求

和要求，并将这些需求和要求融入到课程的整体设计中。其中，“病案信息管理”与“国际疾病分类”的教学团队由我校多所附属医院的病案科骨干成员联合组成，通过与行业的紧密联系，确保了课程的实际应用性和针对性，使学生能够具备满足行业需求的能力和素养。

在学生的实践能力和创新精神培养方面，为学生提供了丰富的实践机会，包括实习、项目实践和科研实践等。通过实践，学生能够将所学知识和技能应用到实际问题中，培养解决问题和创新的能力。

表 1 基于素养的人才培养框架与课程体系映射表

素养大类	分支素养类型	课程名	课程特色
专业核心素养	数智驱动的管理决策类	管理学	
		运筹学	
		卫生组织与信息管理	
		信息服务与用户研究	开展双语教学
		病案信息管理*	
	数据挖掘与知识发现类	医学信息组织与知识发现	
		信息检索	开展双语教学
		国际疾病分类*	
		医学信息分析	开展双语教学
		自然语言处理与电子病例挖掘*	增加课程设计环节
	智能技术支持与智慧系统开发类	JAVA 程序语言与开发框架	
		医院信息系统开发*	增加课程设计环节
专业基础素养	医学基础类	人体结构与功能	与本学院医学信息工程、智能医学工程专业同步配置基础类课程
		预防医学	
		诊断学	
		疾病学	
		临床医学概论	
	数学基础类	高等数学	
		线性代数与矩阵论	
		概率统计	
	计算机基础类	计算机网络	
		数据结构	
		数据库原理与应用	

4 总结

在数智环境下，提出基于核心素养的医学信息管理本科人才培养框架，旨在解决医学信息管理类人才培养中的两个关键问题，即清晰界定专业边界和聚焦培养目标。由于信息管理专业的学科交叉性和综合性，确定专业边界成为挑战。同时，通过详细的课程体系设计，医工、医理充分融合，注重双语教学、行业联系和实践能力培养，旨在培养具备专业知识和技能、具有创新精神和实践能力的高素质人才，为他们未来职业发展奠定基础。但由于技术的不断发展，人才培养体系需要具备动态调整的机制，以适应数智新时代的技术发展要求。这意味着课程体系亦需要不断更新和调整，以保持与行业发展的同步性。

PO-019

智能医学工程专业医校企协同全过程育人体系构建与探索

姚蓓 耿兴云 吴辉群
南通大学

随着智能时代的到来，医学信息学人才的培养也逐渐转型为智能医学工程专业人才培养，同时也为医校企协同育人带来了新的机遇。然而，智能医学工程专业融合了计算机科学、信息工程、生物医学工程等多个学科的知识，具有很强的跨学科性质，在专业的人才培养过程中始终面临一些问

题[1-2]。医学和人工智能专业本身都是热门专业，都已经有了核心的课程体系。医学是一种精英教育，除了本科阶段的培养，医学生还需要在研究生阶段以及临床实践中积累经验。因此，医生的培养需要长期的投入。计算机专业毕业后就业机会较多，近年来企业对计算机人才的需求非常大。在计算机人才的培养中，许多学生往往觉得技能教育不够充分，师资力量不足，因此即使学完了本科课程也不能熟练进行编程。因此，他们需要在进入企业之前接受企业内部培训，甚至在毕业后参加一些社会组织的培训课程，以提高实际操作能力。他们可能觉得医学课程内容繁多，需要记忆的知识非常多，同时还需要具备强大的动手能力。因此，在这个背景下，如何培养相关的复合型人才，需要医、校和企之间相互协同。建立智能医学工程专业医校企协同全过程育人体系，可以提供更加全面、深入的教学内容和实践机会，为医疗人工智能行业提供了计算和信息技术的应用解决方案的创新创业人才。通过建立医校企协同全过程育人体系，可以促进医学教育与实际工作的有机结合。学生们可以通过虚拟仿真和实践项目，模拟真实的医疗场景，提高自己的技能和实践能力。同时，企业也可以参与到教育过程中，为学生提实践机会和指导，培养出更加适应行业需求的人才。

一、开发医校企共同参与的智能医学工程理论与实践课程

医学与信息技术的结合为医学教育带来了巨大的机遇[3]。智能医学工程是一个交叉学科，学生往往感到学分压力比较大，因此在本科阶段要做到专业化很困难。为了解决这个问题，需要将医院和企业的研发文档共享出来，让学生了解工程研发的实际过程，并参与需求分析和软件工程实践。此外，教材来源和符合实践需要的案例也存在一定的问题。在智能医学工程专业医校企协同全过程育人体系中，课程设置应该更加贴近实际需求。针对不同的专业和岗位，设置相应的课程内容和实践环节。同时，应该注重跨学科的融合，将医学、信息技术、管理等多个领域的知识进行整合，提高学生的综合素质。为了满足市场需求并优化学校课程，建议在现有的统一课程基础上，发挥学校特色，形成各自带有鲜明特色的方向。对于医学和信息技术，学生需要具备一定的视野和意识，能够通过项目与企业对接，利用人工智能解决医学领域的问题。通过应用人工智能和大数据技术，可以更加高效地处理这些数据，发现潜在的规律和趋势，为医学研究提供更加准确和全面的支持。在智能医学工程专业医校企协同全过程育人体系中，课程设置应更贴近实际需求，注重跨学科融合，提高学生综合素质。在本科阶段进行专业教育的机会应增加，通过项目与企业合作解决实际问题。

另外，实践环节在院校企全过程育人体系中的重要性不可忽视。实践在智能医学工程教育中扮演着不可或缺的角色，医学院校与企业之间的合作与协同更加重要。通过与企业的合作，学生可以参与实习、实训和科研等活动，进一步了解行业需求和企业文化，提高实践能力和创新能力。为了培养具备企业家精神，能够创新研发并深入改变和优化医疗就诊过程的高端人才，必须进行本科和高端研发人才的全过程培养。在这个过程中，需要联合区形成创新创业的实习实践基地，并由实践基地的管理者提供组织方面的建设，以保障人才培养的建设并为其长期稳定的发展提供支持。我校智能医学工程专业先后开展“地方综合类高校智慧医疗方向师生双创孵化器建设研究”、“基于 AI 黄斑裂孔疾病辅助诊断研究与应用“、”虚拟颈动脉斑块内视系统的研发与应用“、”智能医学工程专业数字孪生实验项目创新实践”、“面向智能医学专业的 VR 教学实验创新实践”、“人工智能+”新医科人才培养模式探索”等教育部产学协同育人项目，丰富学生的产业实践经历。

二、基于 CDIO 导师制的育人过程

导师制度是医学教育中的重要组成部分[4-9]，它不仅是医校企协同全过程育人体系的核心，也是培养医学人才的关键环节。然而，在医学+信息的背景下，导师制度需要不断进行改进和完善，以适应时代的发展需求[10]。首先，导师制度应注重导师的素质和能力。导师作为在特定领域具有丰富经验和专业知识的专家，他们的专业水平和教学能力直接关系到学生的学习效果和发展潜力。针对于医校企协同全过程育人体系，导师需要更好地了解企业需求和行业动态，以便将最新的实践经验和专业知识传授给学生。为此，医学院校应积极与企业进行合作，建立长期稳定的合作关系，通过实践项目和任务，让学生在企业实践中接受更加深入、全面的教育。在导师制度下，每位学生都将配备一名导师，导师将负责学生的学习、实践和科研等方面的指导。导师制度强调个性化、全方位的指导，旨在帮助学生掌握专业技能和实践经验。通过与导师的互动交流，学生可以更好地理解和应用所学知识，提高自己的综合素质。同时，企业作为医学教育的重要组成部分，应积

极参与导师制度的实施。企业可以为学生提供实践机会和实践环境，让他们亲身体验医学实践的各个环节。通过与企业的合作，学生可以接触到真实的医学工作场景，锻炼专业技能和解决实际问题的能力。导师制度还将促进医校企之间的合作与交流。企业作为教育的主体之一，与医学院校进行紧密合作，共同探索医学教育的发展方向和未来趋势。双方可以通过交流合作，相互借鉴经验和资源，推动医学教育与实践的融合，培养更多高素质的医学人才。

三、构建过程与结果相融合的评价机制

学校可以提供人才储备，培养学生在医学和计算机领域的基础知识和技能，为他们的未来发展打下坚实的基础。医院可以为这些人才提供实践的机会，让他们在实际工作中学习和成长。企业则可以为他们提供智能相关技术的支持，使他们能够将理论知识应用于实际工作中，提高他们的实际操作能力。通过医院、学校和企业的合作，可以建立起一个完善的培养体系，培养出具备医学和计算机双重知识背景的复合型人才。通过与企业的合作，学生可以在实践中学习到更多的知识和技能，了解行业的需求和企业的文化，提高自己的实践能力和创新能力。同时，学校、医院和企业之间的协同合作也为学生提供了更多的机会和平台，让他们够更好地实践自己的专业知识和技能，医学和计算机领域的发展做出更大的贡献。评价机制是智能医学工程专业医校企协同全过程育人体系中不可或缺的一部分。为了确保评价机制的有效性和公正性，需要关注以下几个方面。首先，评价应该具备科学性和客观性。评价结果应该基于可量化和可衡量的指标，并且评价过程应该遵循科学的方法和标准。这样可以确保评价结果的客观性，减少主观因素对学生的不公平影响。其次，评价结果应该被用于反馈和利用。评价结果应该及时反馈给学生，帮助他们了解自己的优劣势和发展方向。同时，评价结果也应该被用于改进教学方法和培养计划，从而提高教学质量和学生的综合素质。明确评价的目的是评价机制中的关键一步。在医校企协同全过程育人体系中，评价的目的是提高学生的综合素质和就业竞争力[11-13]。因此，评价机制应该侧重于评估学生的综合素质和就业情况。可以利用学生综合素质评估和毕业生就业率等指标来评价医校企协同全过程育人体系的实施效果。在确定评价指标时，需要考虑医校企协同全过程育人体系的特点。综合素质评估可以包括专业能力、创新能力、团队协作能力和社会责任感等方面。毕业生就业率评估可以包括就业岗位与专业对口程度、就业薪资水平和就业稳定性等方面。选择合适的评价指标可以准确衡量学生的综合素质和就业竞争力。建立评价体系是评价机制的核心。评价体系应该包括评价指标、评价方法和评价标准三个方面。评价指标应该根据实际情况选择合适的指标；评价方法应该基于科学、客观、可靠的方法进行评价；评价标准应该公正、客观，并且能够量化评价结果。最后，建立数据收集和分析机制是评价机制中必不可少的一环。通过建立完善的数据收集和分析机制，可以及时收集评价数据，并进行科学分析。这样可以及时发现问题并进行改进，提高评价机制的有效性和可靠性。

四、建立资源共享型创新创业实践基地

资源共享在医校企协同育人体系中扮演着重要的角色。为了促进合作和信息共享，应该加强学校与企业之间的合作关系。同时，建立多样化的交流平台可以增进学生、教师和企业之间的互动与交流，提高合作效率和质量。目前，创新创业实践基地的建设有多种形式之一是三方授牌。这种形式下，企业认可学校方提供的优质学习资源和储备人才，而创新创业实践基地为学生提供实习和产品研发所需的硬件环境，对于医学专业创新创业具有重要意义[14-20]。通过企业的实际介入，学生能够从企业分享的相关问题和经验中真正体会到实践过程中的困难和挑战。相对于传统的医学信息培养方案，这种形式更加符合企业需求的实际情况。另一种方便的方式是企业委托学校相关的智能技术经纪人来发布需求，并与当地政府科技园区对接，高效地联系学校的教授和研发团队，以横向科技为主的方式为企业提供原型甚至测试方面的帮助。

建立创新创业实践基地首先为相关企业提供所需人才的储备。这些复合型人才能够组织和带领多个成熟学科的人才，依托高校的海归或高端创新性人才。相比于临床经验的医生，医学交叉类产品研发往往缺乏领军人才。其次，在实际经济中，做好政策配套以及提供相关服务非常重要。目前，创新创业往往仅停留在授牌的表面，缺乏全面的管理。联合实验室可以作为一种有效的合作模式，政府、教授或高效资源以及企业将研发项目投入实际生产，共同建立联合中心。然而，联合创新的实际运营方式是将教授或相关人员派驻到某个飞地进行研究院的建设，这会加强研究生和博士生的培养。然而，这些研究生和博士生在非地联合实验室中可能无法与原来实验室中的优势形成良好联系。很多人反映，他们只是换了一个工作环境，但并没有很好地了解到新的工作方式。有部分

企业与高校和中小企业有很好的联系，并有政府组织的展会和博览会，以及相关政策的配套，引导他们与国内外同行之间密切互动。政策支持能够让专业人士做专业事情，不是让一个专业人士分散精力去做不同的事情。初创企业受到欢迎，政府会给他们背书，甚至有前瞻性眼光的政府会设立引导基金来促进行业发展。相关投资人也会给创业团队提供支持。2022 年南通大学被评为首批国家级创新创业教育实践基地，依托国家级南通大学科技园等科创孵化器进行产学研一体的。为进一步加强大学生创新创业教育，推动创新人才培养模式改革，2023 年南通大学成立创新创业教育学院，正处级独立建制，列入院系、科研机构序列，为创新创业型人才的教育提供了坚实保障。医学院智能医学工程教学团队分别与上海、山东等省市的智能医疗企业开展协同育人合作，拓展本科生实习实践基地。另外，还积极与东软、华为等行业进行“数字医学图像处理”等课程产教融合探索，拓展了学生的实践技能。

综上所述，本文通过构建医校企协同全过程育人体系，为智能医学工程专业的人才培养提供了一种新的模式和思路。这种模式不仅能够满足学生的学习需求，还能够促进医院、高校和企业之间的合作与交流，推动智能医学工程专业的发展。

PO-020

基于全蛋白质组关联分析探究治疗炎症性肠病的新型药物靶点

郝佳炜[2]，陈淼然[2]，宁宇鑫[2]，谢文佳[2]，杨奕璇[2]，杨彤彤[2]， 张耀虹[4]， 于琦[3]，贺培风[1*]
1 山西医科大学临床决策大数据重点实验室
2 山西医科大学细胞生理学重点实验室
3 山西医科大学管理学院
4 山西医科大学第一医院

1 引言

炎症性肠病（IBD）是一种慢性复发性胃肠道疾病。尽管 IBD 的发病机制尚不清楚，但它可能是由遗传易感性驱动的(1)。IBD 主要包括克罗恩病（CD）和溃疡性结肠炎（UC）。与溃疡性结肠炎相比，克罗恩病的临床表现更具异质性。特别是，CD 可累及胃肠道的任何部位，尤其是回肠末端或肛周。与 CD 不同，UC 仅局限于大肠，尤其是阑尾周围(2, 3)。尽管在过去几年中，IBD 的治疗取得了长足的进步，但目前可用的标准治疗方案疗效有限，存在对药物反应不充分、产生耐药性或长期失效等缺点。因此，有必要确定治疗靶点，以促进新药物的开发。

过去十年间，大规模的全基因组关联研究（GWASs）发现了各种非重叠遗传风险位点，以筛选候选靶基因用于药物开发。然而，所发现的变异只解释了少数遗传风险(4)。将这些遗传风险位点开发成药物靶基因仍是一项重大挑战。Katrina 等人通过对包括 25,305 人在内的 GWAS 数据进行荟萃分析，确定了 25 个新的基因位点，并进一步揭示了已知治疗策略的可能生物学机制(5)，然而，全球基因组研究很少发现介导性状变异效应的因果基因(6)。因此，Virginia 等人使用一种新的转录组范围关联研究（TWAS）统计方法，通过将基因表达与特征和疾病相关联，发现了 39 个影响 IBD 发病机制的新型易感基因(7)。尽管之前的研究已经揭示了与 IBD 发病机制相关的基因位点，但其结果都是基于基因组或转录组（信使 RNA [mRNA]）分析而非蛋白质组分析得出的。蛋白质是基因表达的最终产物，可作为潜在的药物靶点和生物标志物(8)。

近年来，PWAS 研究引起了人们的极大兴趣。这是一种对所有影响蛋白质编码基因的变异信号进行整合，并使用机器学习和预测模型来评估它们对蛋白质功能的整体影响的方法(9)。Zhang 等人结合基因和蛋白质表达数据，提出了首个基于血浆蛋白质组的 PWAS 框架，并确定了三种特定的基因变异，从而揭示了治疗疾病的潜在靶蛋白(10)。总之，PWAS 是一种全面、有效和可靠的分析方法。

在这项研究中，我们整合了来自血浆样本的高通量蛋白质组和基因数据，以确定与 IBD 易感性相关的蛋白质以及 IBD 的潜在治疗靶点。研究人员对 IBD 的 GWAS 和蛋白质定量性状位点（pQTL）

数据进行了 PWAS 研究。孟德尔随机化（MR）验证了所发现的蛋白质与 IBD 之间的因果关系。最后，进行了贝叶斯共定位分析，以检验 GWAS 和 pQTL 数据是否与共享的因果变异一致。据我们所知，本研究是首次报告与 IBD 易感性相关的蛋白质组学研究，为该疾病的潜在治疗方法提供了新的见解。

2 方法

2.1 pQTL 数据

根据社区动脉粥样硬化风险（ARIC）招募的 7,213 名欧洲血统的 4,657 种血浆蛋白数据，确定了顺式-pQTL(10)。该研究根据第三阶段千人基因组计划（1000Genome）中的个体计算了模型的权重，并将其基因型数据作为权重模型。利用表达残差概率估计（PEER）方法，共鉴定了 1,348 个常染色体基因编码的独特蛋白质或蛋白质复合物，并对其数据进行了基于等级的反正态变换和质量控制。

2.2 GWAS 数据

有关 IBD（包括 UC 和 CD）的公开 GWAS 数据来自 de Lange KM 最近的一项研究。该研究纳入了来自英国 IBD 遗传学联合会（UKIBDGC）和 UK10K 联合会的 25,305 名欧洲血统个体（12,160 名 IBD 患者和 13,145 名对照个体）。(5)经过质量控制后，最终纳入了来自 4,474 名克罗恩病患者、4,173 名溃疡性结肠炎患者、592 名未分类 IBD 患者和 9,500 名对照组个体的 296,203 个变体的数据进行分析。

2.3 GWAS 数据的质量控制

本研究只包括欧洲血统的个人数据。数据经过了以下严格的质量控制：(1) 将数据与 hgl9 人类参考基因组对齐；(2) 保留常染色体 SNPs 双偶合；(3) 去除重复或缺失 rs ID 的 SNPs；(4) 过滤 MAF 值小于 0.01 的 SNPs。

2.4 全蛋白质组关联研究

通过整合蛋白质组学和遗传学数据（即 pQTL 数据中血浆蛋白的表达和 GWAS 数据中的 SNPs），使用 FUSION 标准管道进行 PWAS。(11)对于具有显着遗传性（即遗传性 P 值小于 0.01）的蛋白质，采用 ENet 模型（一种更精确的预测模型），通过评估 IBD、UC 和 CD 中各基因与血浆蛋白的关系，研究 SNP 对蛋白质丰度的影响。在只考虑遗传蛋白的情况下，研究了血浆蛋白表达与蛋白顺式位置的 SNPs 之间的关联强度，以预测 SNPs 对蛋白丰度的影响。连锁不平衡（LD）是通过一个 LD 参考面板来控制的，从而减少其对估计测试统计量的影响。

假发现率（FDR）阈值为<5%，映射窗口为±500 kb，在此基础上评估 Z-分数，以衡量血浆蛋白表达与 GWAS 数据之间的遗传协方差(12)。根据重要顺式遗传血浆蛋白或蛋白复合物的估算模型总数，P 值为 3.7×10^{-5} 表示存在重要遗传位点。

2.5 MR 分析

进行 MR 分析是为了验证 IBD、UC 和 CD 的 PWAS 显著顺式调控血浆蛋白是否与 IBD 丰度相关，并确定候选因果血浆蛋白。在 MR 分析中，根据 PWAS 结果得出的与蛋白质表达密切相关的基因变异（pQTL）被视为工具变量（IV）(13, 14)。根据严格的纳入标准筛选出与 IBD（包括 UC 和 CD）相关的全基因组意义重大（$P < 1\times10^{-5}$）的 SNP，然后进行 LD 聚类（R2 < 0.01）。如果筛选后只有一个 SNP 可用，则使用 Wald 比值来估计暴露与结果之间的因果关系。Wald 比率计算的是蛋白质生物标志物中 IBD 风险每标准差变化的对数几率变化与 SNPs 的工具风险等位基因的关系(12)。当使用一个以上的 SNP 作为 IV 时，则使用逆方差加权法（IVW）和 MR-Egger 法(15)。此外，使用 MR-Egger 检验水平多向性是否为零，P 值小于 0.05 表示存在水平多向性。MR 分析使用 R 中的 "TwoSampleMR "软件包实现。

2.6 贝叶斯共定位分析

贝叶斯共定位（COLOC）分析用于评估同一因果单核苷酸变异导致 IBD（包括 UC 和 CD）风险和调节血浆蛋白水平（即 pQTL）的概率(16)。使用了 $p1=10^{-4}$、$p2=10^{-4}$ 和 $p12=10^{-5}$ 的默认 Coloc 检验。p1 是给定一个变异与 IBD（包括 UC 和 CD）相关的概率，p2 是给定变异是显著 pQTL 的概率，p12 是给定变异与 IBD（包括 UC 和 CD）相关且是 pQTL 的概率。R 中的 Coloc

软件包用于 pQTL 和 IBD GWAS 数据集之间共享 SNP 的共定位分析，并计算贝叶斯因子，以评估五个相互排斥的假设的后验概率 （PP）：（1） 没有 SNP 与 GWAS 或 pQTL 相关（H0）；（2） 只有性状 1（GWAS）有一个因果 SNP，而没有 SNP 与 pQTL 相关（H1）；（3） 只有性状 2（pQTL）有一个因果 SNP，而没有 SNP 与 GWAS 相关（H2）；（4） 两个性状（GWAS 和 pQTL）都有独立且不同的因果 SNP（H3）；（5） 两个性状都有一个共享的因果 SNP（H4）。我们主要关注最后一个假设（H4），其 PP 被称为 PPH4。PPH4≥0.75 被定义为共聚焦的有力证据。

3 结果

3.1 在 PWAS 中发现的相关血浆蛋白

我们利用 FUSION 管道将 IBD、UC 和 CD 的 GWAS 数据与 1，348 个血浆蛋白质组的数据整合在一起，进行了 PWAS 分析。结果显示，62、22 和 30 个顺式调控血浆蛋白的丰度分别与 IBD、UC 和 CD 显著相关。在这些蛋白中，有 4 种蛋白在 IBD、UC 和 CD 中很常见。此外，在 IBD 和 UC 之间、IBD 和 CD 之间以及 UC 和 CD 之间，分别有 15、17 和 6 种蛋白质是常见的。在较高的置信水平下，这些蛋白质的 FDR 均小于 0.05。与 IBD、UC 和 CD 相关的血浆蛋白的详细信息见图 1 和表 1。

3.2 通过 MR 验证血浆蛋白质与 IBD、UC 和 CD 的关系

为了验证血浆蛋白与 IBD、UC 和 CD 风险之间的关系，并阐明具体的因果关系，对血浆蛋白进行了 MR 分析。共有 32、8 和 9 种具有强烈因果效应的蛋白质分别被确定为 IBD、UC 和 CD 的生物标志物（$P < 0.05$）。与 IBD、UC 和 CD 风险相关的蛋白质之间存在部分重叠。与 IBD 风险相关的前五种血浆蛋白是 MST1 （$P = 6.14 \times 10^{-8}$， OR = 0.82， 95% CI = 0.77–0.88），PARK7 （$P = 1.76 \times 10^{-6}$， OR = 0.81， 95% CI = 0.75–0.89）， NADK （$P = 3.25 \times 10^{-5}$， OR = 0.84， 95% CI = 0.78–0.91）， RIPK2 （$P = 6.22 \times 10^{-5}$， OR = 0.62， 95% CI = 0.49–0.78）和 TALDO1 （$P = 1.14 \times 10^{-4}$， OR = 0.60， 95% CI = 0.46–0.78）。与 UC 风险相关的前五种血浆蛋白是 MST1 （$P = 6.22 \times 10^{-8}$， OR = 0.84， 95% CI = 0.79–0.90）， CADM2 （$P = 1.46 \times 10^{-4}$， OR = 0.64， 95% CI = 0.51–0.80）， VSIR （$P = 2.77 \times 10^{-4}$， OR = 0.89， 95% CI = 0.83–0.95）， PRKCB （$P = 6.20 \times 10^{-4}$， OR = 1.19， 95% CI = 1.08–1.31） and PIGR （$P = 7.14 \times 10^{-4}$， OR = 0.78， 95% CI = 0.67–0.90）。与 CD 风险相关的前五种血浆蛋白是 FLRT3 （$P = 4.99 \times 10^{-8}$， OR = 0.90， 95% CI = 0.87–0.93）， MST1 （$P = 6.08 \times 10^{-6}$， OR = 0.83， 95% CI = 0.77–0.90）， ABO （$P = 3.96 \times 10^{-5}$， OR = 1.11， 95% CI = 1.06–1.16），TNFRSF1A （$P = 5.32 \times 10^{-4}$， OR = 1.35， 95% CI = 1.14–1.60） and C7 （$P = 8.41 \times 10^{-4}$， OR = 1.14， 95% CI = 1.06–1.23）。详细信息见图 2 和图 3 以及补充表 1-3。

3.3 与疾病风险相关的血浆蛋白质共定位

为了验证基因共定位，对 pQTL 和 IBD（UC 和 CD）的 GWAS 数据中符合 MR 分析中 FDR 校正 P 值阈值的基因进行了 PP 评估，以确定它们之间共享的因果变异体。结果显示，GWAS 和 pQTL 数据有可能共享一个因果变异体（PPH4 ≥ 0.75）。根据 PPH4 值≥0.75，发现 5、3 和 2 个蛋白质在 IBD、UC 和 CD 的进展中起重要作用（图 2 和 4）。在通过共定位分析发现的蛋白中，CADM2 是 IBD 和 UC 的重要共用蛋白。

4 结语

总之，本研究揭示了可能导致 IBD（UC 和 CD）发病机制的九种蛋白质生物标志物。本研究的发现可为今后有关遗传和功能机制的研究提供理论依据，从而有助于未来开发治疗 IBD（UC 和 CD）。

PO-021

脑卒中患者的神经系统损伤对预后影响的研究

王敏，王梓阳，李姣*
中国医学科学院医学信息研究所

1 引言

脑卒中（stroke）具有高发病率、高死亡率和高致残率的特点，是全球致残和血管性死亡的主要原因。国内每年新发约 240.8 万病例，且患病率和发病率逐年上升，严重威胁着人们的健康。目前几乎没有干预措施可以改变脑卒中的预后，但更好地了解影响脑卒中预后的因素可能有助于开发更好的预后工具和新治疗策略。美国国立卫生研究院（NIH）的脑卒中量表（NIHSS）已被用于评估脑卒中的长期预后，目前还不清楚除 NIHSS 之外的其他风险因素对脑卒中死亡的影响。根据流行病学调查研究，在存活的脑卒中患者中，62%的患者遗留各种各样的神经功能损伤，导致不同程度的生活质量受损和残疾，给家庭、护理人员和社会带来了沉重的负担。因此，早期准确评估神经系统损伤程度对脑卒中患者预后影响有助于做出正确的临床决策以改善患者预后。

对于脑卒中患者而言，首次出现症状后的神经功能损伤是一种常见且令人担忧的并发症，并直接关系到患者的预后。格拉斯哥昏迷量表（Glasgow Coma Scale，GCS）提供了一种评估神经系统损伤程度的实用方法，得分值越高，提示患者的意识状态越好，意识障碍程度是脑损伤的临床标志。GCS 采用三重标准评分系统：睁眼反应（最高 4 分）、肢体运动（最高 6 分）和语言反应（最高 5 分）。通过这些反应可以比较不同患者之间的情况以及患者随时间的变化，从而为治疗提供重要指导，这三个部分可以单独计分，也可以合并计分。尽管人们越来越认识到脑卒中患者神经系统损伤的普遍性，而且已有少量研究分析了急性脑卒中患者院前和入院后，早期神经功能损伤的频率、以及影响因素。然而，关于脑卒中的 ICU 患者住院期间的神经系统以及患者的睁眼反应、肢体运动和语言反应情况（不同 GCS 评分系统）对患者预后的影响尚未得到很好的描述。

生存分析，也被称为时间-事件分析（time-to-event analysis），是一种非参数统计模型，其中因变量是 "时间间隔"，自变量是给定事件的 "发生与否"。在循证医学中，生存分析被广泛用于检查时间到事件序列，它可以说明生存/死亡事件的发生时间。Kaplan-Meier 曲线又称生存曲线，是一种生存分析的常用方法，定义是在给定时间内存活的概率，同时考虑多组之间的变化，该方法主要分析单一因素对生存期的影响，用于估计患者生存率和绘制生存曲线。为了对神经系统损伤程度和脑卒中预后之间的关系提供更加客观和定量的评估，本研究将通过生存分析法评估脑卒中患者的神经系统损伤对预后的影响。

2 方法

2.1 数据来源以数据集构建

研究的数据主要来源于公开数据集 MIMIC-IV，由麻省理工学院（MIT）计算生理学实验室和合作研究小组的研究人员提供，收集了 2008-2019 年间美国马赛诸塞州波士顿市三级学术医疗中心住院患者的相关信息，包括人口学信息、实验室测量值、给药方式、记录的生病体征等。本研究的作者通过了 NIH 网站上名为“保护人类研究参与者”的培训课程，并获得了拥有数据的许可（证书编号：44835986）。纳入标准如下：（1）首次入院时年龄小于 18 岁的患者；（2）出院诊断中前三位诊断代码中包含卒中相关 ICD 代码；（3）在 ICU 内停留时间>24 小时的患者；（4）多次入院 ICU 的 IS 患者，仅提取首次入院数据。同时排除标准如下：（1）文件或其他重要病历不完整的患者;（2）生存结局资料缺失的患者。根据上述标准提取对应的卒中患者，共提取 6021 人。

研究选取的主要终点是 30 天内的 ICU 生存状态以及生存时间（death_30icu_days、los_icu）。在文献调研的基础上，使用 PostgreSQL 工具提取了与脑卒中相关的特征，包括 9 项人口学特征；12 项首次进入 ICU 后的基础生命体征（如心率、血压等）；32 项实验室数据（如血常规、肝肾功能、电解质和凝血功能等）；14 项基础合并症（如糖尿病、冠状动脉疾病、充血性心力衰竭等）；以及相关你评分（评估神经系统损伤状态的 GCS 评分）。其中，本能研究选取年龄、性别、神经系统损伤量表 GCS 评分以及 GCS-睁眼反应、GCS-肢体运动、GCS-语言反应评分，GCS 分值越低表示神经功能损伤程度越高。我们根据 GCS 评分（总分 15 分）将患者分为 3 个类别（<9 分[重度意识障碍]、9~13 [中度意识障碍]、≥13 [轻度意识障碍]）；根据 GCS-睁眼反应评分（总分

4 分）将患者分为 2 个类别（=4 分[睁眼反应正常]、<4 分[睁眼反应异常]）；根据 GCS-肢体运动评分（总分 6 分）将患者分为 2 个类别（=6 分[肢体运动正常]、<6 分[肢体运动异常]）；根据 GCS-语言反应评分（总分 5 分）将患者分为 2 个类别（=5 分[语言反应正常]、<5 分[语言反应异常]）。

2.2 生存分析

我们采用 Kaplan-Meier 分析法进行多组生存分析并绘制生存曲线。生存分析中，事件是指研究中规定的生存研究的终点，本研究中指患者死亡。生存时间，从检测开始到事件发生所经过的时间，本研究中指 30 天内 ICU 患者生存时间（los_icu）。删失是指研究对象在观察时间内没有发生事件，本研究中指 30 天内患者未出现死亡事件(death_30icu_days)。Kaplan-Meier 曲线显示存活概率（未发生死亡事件）与时间的函数关系，时间位于 X 轴，存活概率位于 Y 轴，$p < 0.05$ 时，具有统计学意义。采用 Log rank 检验进行生存分析组间比较，统计量 χ2 的计算公式为$\chi^2=\frac{[\sum W_i(d_{gi}-\tau_{gi})]}{V_g}$，累积风险常用来估计风险概率，累积死亡率定义为$H(t) = -\log(\text{survivalfunction}) = -\log(S(t))$。

2.3 统计分析

我们用 R 软件实现所有的数据分析，按 GCS 总分、GCS-睁眼反应（gcs_eyes）评分、GCS-肢体运动（gcs_motor）评分、GCS-语言反应（gcs_verbal）评分分别计算人群特征，计量资料以均数±标准差（$x \pm s$）表示，计数资料以频数(%)表示。分类变量采用χ^2检验，连续性变量采用 t 检验，P≤0.05 被认为具有统计学意义。

3 结果

共有 6021 名脑卒中患者具有完整的 GCS 评估 、CGS-睁眼反应评估、GCS-肢体运动评估、GCS-语言反应评估以及生存时间等相关预后测量数据的参与者被最终分析，表 1 列出了按照 GCS 不同评分系统分层的参与者的基本信息。我们对 GCS-睁眼反应评估、GCS-肢体运动评估和 GCS-语言反应评估 3 个子表进行进一步分析，表 2 展示了睁眼反应评估、肢体运动评估和语言反应的不同组之间的基本信息。

图 1(A)显示脑卒中患者 GCS 总分预后情况的生存曲线，不同组的生存时间有明显差异（P <0.0001）。Kaplan-Meier 分析显示，轻度和中度意识障碍组的中位生存期长于重度意识障碍组。图 1(B)显示了随着时间变化不同程度意识障碍组患者累积死亡率的变化情况，根据累计死亡率曲线显示，轻度和中度意识障碍组差距不大，但都明显低于重度意识障碍组。

图 2 分别展示了随着时间变化不同睁眼反应组、肢体运动组、语言反应组患者的生存率的变化情况，每一个“台阶”都对应着一个发生终点（死亡）事件的时间点。Kaplan-Meier 曲线起点所有组患者的生存率都是 100%，随着时间的推移，有患者发生死亡，终点事件发生，生存率从 100%开始下降。在 12.76 天时，睁眼反应异常组的存活概率约为 50%，睁眼反应正常组的存活概率约为 75％。在 17.56 天时，肢体运动异常组的存活概率约为 50%，肢体运动正常组的存活概率约为 75％。在 15.69 天时，语言反应异常组的存活概率约为 50%，肢体运动正常组在 30 天后的存活概率才降到 50%。在随后 ICU 住院过程中，我们发现睁眼反应异常的生存率在任何一个时间均低于睁眼反应正常组，肢体运动异常的生存率在任何一个时间均低于肢体运动正常组，同时语言反应异常的生存率在任何一个时间均低于语言反应正常组（P 均<0.0001）。图 3 显示了随着时间变化不同睁眼反应组、肢体运动组、语言反应组患者累积死亡率的变化情况。综上，睁眼反应、肢体运动、语言反应正常组的患者的生存率明显高于睁眼反应、肢体运动、语言反应异常组。

表 1 基本信息表

变量	存活（x±s，n=5051）	死亡（x±s，n=967）	P
GCS	11.75±3.46	8.57±4.68	***
GCS_eyes	2.95±1.02	1.70+1.03	***
GCS_motor	5.29±1.39	3.22±1.95	***
GCS_verbal	3.35±1.72	2.62+1.85	***

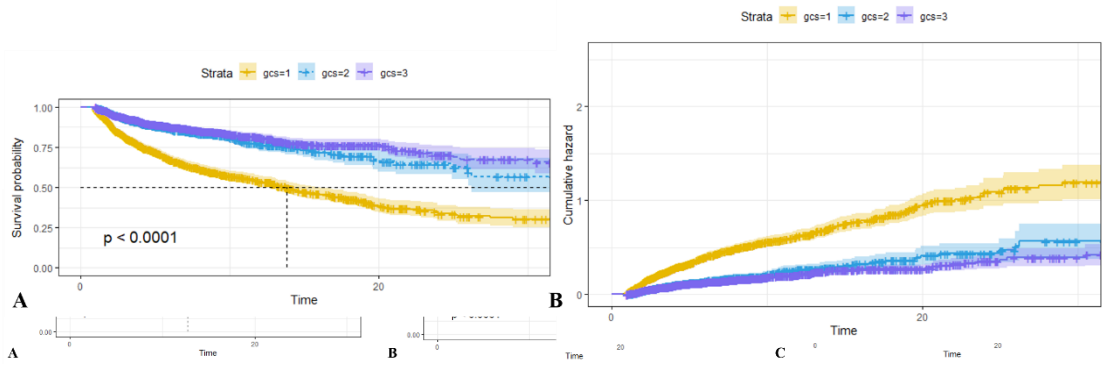
年龄	69.09±15.47	70.31±15.08	0.04
性别			
男(n=3136)	2640(84.18%)	496(15.82%)	-
女(n=2882)	2409(83.59%)	473(16.41%)	-

注：***: p<0.001

表 2 神经系统损伤不同分组人群基本信息表

变量	睁眼反应		肢体运动		语言反应	
	正常(x±s)	异常(x±s)	正常(x±s)	异常(x±s)	正常(x±s)	异常(x±s)
年龄	69.27±15.46	69.34±15.25	66.85±15.19	68.30±14.87	69.37±15.25	70.88±15.05
性别						
男	2427(77.39%)	709(22.60%)	1920(61.22%)	1216(38.78%)	1114(35.52%)	2022(64.48%)
女	2218(76.96%)	664(23.04%)	1681(58.33%)	1201(41.67%)	937(32.51%)	1945(67.49%)

图 1 GCS 总体评分生存分析

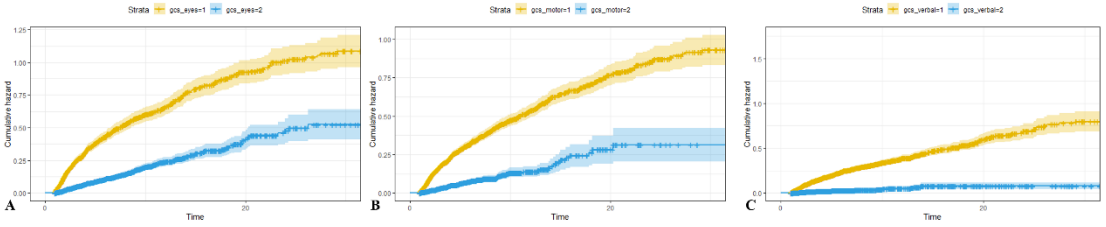


注：gcs:昏迷量表；1：重度意识障碍；2：中度意识障碍；3：轻度意识障碍。横轴（x 轴）表示以天为单位的时间，纵轴（y 轴）表示生存的可能性或生存的人口比例。线代表三组的存活曲线。曲线中的垂直下降表示事件。曲线上的十字叉表示此时患者存活（删失）。

图 2 A：睁眼反应 B：肢体运动 C：语言反应评分生存分析

注：gcs_eyes：睁眼反应评分；gcs_motor：肢体运动评分；gcs_verbal：语言反应评分；1 表示睁眼反应/肢体运动/语言反应异常组；2 表示睁眼反应/肢体运动/语言反应正常组。横轴（x 轴）表示以天为单位的时间，纵轴（y 轴）表示生存的可能性或生存的人口比例。线代表两组的存活曲线。曲线中的垂直下降表示事件。曲线上的十字叉表示此时患者存活（删失）。

图 3 A：睁眼反应 B：肢体运动 C：语言反应累计死亡率



注：gcs_eyes：睁眼反应评分；gcs_motor：肢体运动评分；gcs_verbal：语言反应评分；1 代表睁眼反应/肢体运动/语言反应异常组；2 代表睁眼反应/肢体运动/语言反应正常组。横轴（x 轴）表示以天为单位的时间，纵轴（y 轴）表示累积死亡率。

4 讨论

本研究存在以下局限性，1）这项研究仅对 GCS 评分以及不同 GCS 评分系统与脑卒中 ICU 患者预后的关系进行了分析，实际在特征提取时，我们共提取包括人口学特征、实验室数据、基础合并症在内共 55 项特征。其他特征之间的关系以及对生存率的影响还需综合分析；2）我们仅对年龄和性别 2 项人口学特征进行了混杂因素分析，可能存在其他人口学特征（如种族）的混杂偏倚；3）在这项小型初步研究中，我们通过统计学方法验证了不同程度神经系统损伤对脑卒中 ICU 患者预后的影响，我们的研究结果只是提出了假设，还需要进一步的研究来验证这一假设。

5 结论

本研究首次探讨了脑卒中患者的神经系统损伤对预后影响的研究。基于 30 天内不同神经系统损伤程度的脑卒中 ICU 患者生存分析结果，可以推测随着时间的变化，神经损伤程度越低，患者的生存率越高。对于睁眼反应、肢体运动以及语言反应情况，无反应异常的患者的生存率较高。未来，还需要进一步的严格的研究来验证这一推测。

PO-022

基于因果发现的卒中死亡早期预测模型

王梓阳，杨林，王敏，兰雨姗，李姣
中国医学科学院医学信息研究所

1 引言

卒中（stroke）是一种急性脑血管疾病，具有高发病率、高死亡率和高致残率的特点，是全球范围内重大的公共卫生问题。国内每年新发病例 240.8 万，且患病率和发病率逐年上升。同时根据流行病学调查研究，在存活的卒中患者中，62%的患者遗留各种各样的神经功能障碍，导致不同程度的生活质量受损和残疾，给家庭、护理人员和社会带来了沉重的负担[1,2]。当前临床上将卒中根据病因分为四大类，根据一项囊括了 31 省共计 480687 人的流行病学研究，我国卒中病例中，缺血性卒中（ischemic stroke，IS）占比为 77.8%、颅内出血（intracranial hemorrhage，ICH）占比为 15.8%、蛛网膜下腔出血（subarachnoid hemorrhage，SAH）占比为 4.4%以及 2.1%的未分类卒中[3]。急性缺血性卒中（AIS）的治疗是与时间赛跑的过程，卒中院前的急救管理是卒中急救生命连的关键环节之一，对于卒中患者的治疗和预后有决定的作用[4]。对卒中死亡率进行预测可以有针对性地对死亡风险较大患者及时进行风险管理，降低患者死亡风险，同时合理分配有限的医疗资源和人力资源。目前临床上常使用评分系统对卒中的严重程度进行评估，但评分本身并不能作为卒中死亡率预测的指标。

机器学习方法允许通过自动化算法处理非线性关系和变量之间的复杂相互作用，从而发现数据中的经验模式和相关关系，进而提高预测准确度[5]。而相关关系并不是稳定的，这种不稳定直接导致了模型的鲁棒性和泛化性不佳，当训练集和测试集的数据存在偏差，模型的预测效果就会变得不可信，这类问题可以被归结为分布外泛化（out of distribution，OOD）问题[6]。与相关关系相对应的是因果关系，如果能够让预测模型在训练阶段重点关注因果关系，则能够极大的减少 OOD 问题。此外，机器学习模型是黑盒模型，尽管有一些如沙普利加和（SHAP）等可解释性的方法能够揭示其预测的过程，但仍不能令人信服。尤其是在临床医学这类严谨学科中，即使模型的输出准

确，也会极大的限制其在行业的发展和应用[7]。清华大学的崔鹏教授团队提出的稳定学习方案，在当训练数据集与测试数据集中的数据差异增大时，稳定学习相比于传统训练方法有着更小的波动[8]。而其本质则是让模型在训练中以因果关系进行预测。

因果关系识别也被称为因果发现（causal discovery），通过分析观测数据中的数据分布，利用一定的规则和方法找出背后的因果关系。基于观测数据的因果发现一直以来是多学科学者所关注的焦点问题。为了提高模型对于卒中死亡预测的准确性、鲁棒性和可解释性，本研究拟采用因果学习的方法，改善当前研究在卒中的死亡预测中存在问题。

2 研究方法

本研究以 MIMIC-IV 为数据，利用因果关系识别方法研究数据集中进行特征提取，根据提取的特征进行卒中死亡的建模分析。

2.1 数据来源及数据集构建

研究的数据主要来源于公开数据集 MIMIC-IV[9]，由麻省理工学院（MIT）计算生理学实验室和合作研究小组的研究人员提供，收集了 2008-2019 年间美国马赛诸塞州波士顿市三级学术医疗中心住院患者的相关信息，包括人口学信息、实验室测量值、给药方式、记录的生病体征等。本研究的作者通过了 NIH 网站上名为“保护人类研究参与者”的培训课程，并获得了拥有数据的许可（证书编号：44835986）。

应用电子病历进行疾病谱的分析，最主要的一点是要能准确识别目标患者群体，常用的方法有国际疾病分类编码（ICD）等表型算法[10]。Tingting Zhang 等[11]人提出一套用于识别卒中患者的 ICD 代码组合，多数基于 MIMIC 数据库研究的学者也应用同一套代码来对卒中患者进行识别[12,13]。纳入标准如下：（1）首次入院时年龄小于 18 岁的患者；（2）出院诊断中前三位诊断代码中包含卒中相关 ICD 代码；（3）在 ICU 内停留时间>24 小时的患者；（4）多次入院 ICU 的 IS 患者，仅提取首次入院数据。同时排除标准如下：（1）文件或其他重要病历不完整的患者;（2）生存结局资料缺失的患者。根据上述标准提取对应的卒中患者，共提取 6021 人，其人群对应卒中分型如表 1 所示。

表 1. 卒中患者对应分型表

	ICD-9-CM	人数	ICD-10	人数
缺血性卒中（AIS）	433	632	I63	1202
	434	977		
	436	0		
短暂性脑缺血发作（TIA）	435	46	G45	62
		445	G46	1
蛛网膜下腔出血（SAH）	430	724	I60	273
颅内出血（ICH）	431	632	I61	614
合计	6021			

研究选取的主要终点是 30 天内的 ICU 死亡。在文献调研的基础上，使用 PostgreSQL 工具（版本 9.6）提取了与卒中相关的特征[14–17,13,12]。包括年龄、性别、种族等 9 项人口学特征；首次进入 ICU 后的基础生命体征的最值和平均值（如心率、血压等 12 项）；32 项实验室数据（如血常规、肝肾功能、电解质和凝血功能等）；14 项基础合并症（如糖尿病、冠状动脉疾病、充血性心

力衰竭等）。但需要说明的是，MIMIC 中并没有提供与卒中诊断相关的 NIHSS 评分，为了尽可能弥补 NIHSS 评分的缺失带来的预测问题，研究在上述四大类特征的基础上，又提取了评估合并症的 CCI、评估生理状态的 SAPS、评估中风严重程度的 OASIS、评估器官损害的 SOFA、评估验证的 SIRS、评估昏迷状态的 GCS，用以从不同维度评估患者状况。

2.2 特征筛选

特征筛选是训练模型中的重要一步，其选择将直接影响最终预测模型的效果。之前的方式更多倾向于应用统计学检验、互信息法、特征重要性、PSI 等方式进行筛选。本研究主要通过利用因果关系识别的算法来从观测性数据中识别因果图，选择直接指向结局变量的特征作为特征筛选你的方式，如何让模型能够专注于因果特征的学习，是本研究的关键部分。

2.2.1 因果关系识别方法

目前从获取因果关系最常用的方式是随机临床试验，但是其受限于成本、可行性、实验设计等问题难以普遍应用。因此利用观测性数据进行因果发现是一种非常必要的补充，其核心思想便是条件独立。目前主要有基于约束发现方法、基于函数的因果发现和基于评分的发现方法三种，其共同目的是从观测性数据中得到因果图，用有向无环图来表示（directed acyclic graphs，DAGs）。基于约束的的算法以条件独立关系作为约束，生成满足条件的因果图（通常用有向无环图表示）。从 20 世纪 90 年代开始，数据中的条件独立性已被用作约束条件来重构因果信息。这类基于约束的方法包括 PC（Peter-Clark）算法和快速因果推理（Fast Causal Inference, FCI）算法[18]。基于函数因果模型的因果发现方法主要通过恰当定义的函数来区分等价类中的不同因果结构，常用的方法有 LiNGAM 及其拓展方法[19]、ANM[20,21]、PNL[21]等。相比于基于约束的算法在进行中需要进行多次条件独立性检验，基于评分的贪婪等价类搜索（Greedy Equivalence Search, GES）[22]可直接在等价类空间进行搜索算法而无需要做独立性检验，因此已经得到很多应用。在本文中主要应用 GES 方法进行因果发现。

GES 是用于学习贝叶斯网络结构的一种经典算法。贝叶斯网络是一种图模型，用于表示变量之间的概率依赖关系。GES 算法的目标是根据观测数据找到最佳的贝叶斯网络结构，以最好地解释数据并捕获变量之间的条件独立性关系。研究选择贝叶斯信息准则（BIC）作为评分指标，进行贪婪搜索，不断通过可能边的添加、删除和翻转操作，选择使得 BIC 评分函数最小化，进而得到最后因果图。

$$BIC = -2\ln(L) + k\ln(n) \qquad (1)$$

2.2.2 互信息法

本研究从 MIMIC -IV 数据库中提取了大量备选特征后，采用互信息法对特征进行筛选。互信息（Mutual Information）是衡量随机变量之间相互依赖程度的度量。

$$H(X;Y) = H(X) - H(X|Y) \qquad (2)$$

其中$H(X)$是指X的信息熵，$H(X;Y)$是特征X和因变量Y的互信息值，表示在已知X时，对Y不确定性的减少。互信息的取值范围为 0～1，0 表示已知特征之后不能减少标签的不确定性，即该特征与标签不相关；同理 1 表示特征与标签完全相关。互信息法从信息熵的角度出发，相比较于方差分析，卡方检验，除了能够判断线性相关，还可以判断非线性相关。

2.3 实验设置

本研究进行实验的操作系统为 Ubuntu 18.04.5 LTS，算法均在 Python 3.9 和 R 3.6 的环境的基础上实现，因果学习的应用的软件包为 Causal Discover Toolbox（0.6.0）[23]，机器学习软件包为 Scikit-learn（1.3.0）[24]。

本研究的评价指标主要采用分类算法的评价指标，包括受试者特征曲线下面积（ROAUC）、F1 分数、MCC、精确度（Precision），召回率（Recall），等指标。同时，为了避免偶然误差，研究采用 10 次随机种子的 10 折交叉验证的方式计算上述指标。

卒中按照病因可以粗略的分为缺血性卒中和出血性卒中，两种卒中发病率不同，其死亡率也有一定的差异。现实中，不同医院接诊的卒中患者的危急程度也不同，为了检验因果学习的效果，研究将按照卒中类型的比例构造训练数据集，以此来模拟现实中不同医院的差异。

3 结果与讨论

从 MIMIC-IV 数据库中提取了共计 6021 名患者，死亡率 16.1%。研究人群的基本特征的描述性统计如表 1 所示，从表可知，缺血性卒中患者约为出血性卒中患者的 2 倍，同时出血性卒中患者的死亡率显著高于缺血性卒中患者（21.2% vs. 13.5，p=1.82E-14）。

表 1 特征与生存状态、卒中类型的描述性统计

特征	生存 (n=5052)	死亡 (n=969)	统计量	p 值	出血性卒中 (n=2056)	缺血性卒中 (n=3965)	统计量	p 值
体重	76.56±25.08	79.87±26.15	-3.7	***	76.56±25.08	79.87±26.15	0.4	0.716
年龄	72.93±14.06	68.6±15.56	8.6	***	72.93±14.06	68.6±15.56	15.1	***
GCS	8.59±4.69	11.75±3.46	-20.0	***	8.59±4.69	11.75±3.46	8.5	***
SIRS	2.89±0.92	2.32±0.99	17.1	***	2.89±0.92	2.32±0.99	0.8	0.399
OASIS	40.47±8.51	31.57±8.56	29.8	***	40.47±8.51	31.57±8.56	0.8	0.439
SOFA	6.83±3.92	4.19±3.04	19.8	***	6.83±3.92	4.19±3.04	8.7	***
SAPS	45.27±14.45	33.49±12.14	23.8	***	45.27±14.45	33.49±12.14	9.7	***
CCI	7.34±2.81	6.36±2.64	10.1	***	7.34±2.81	6.36±2.64	16.8	***
性别			1.3	0.251			0.5	0.490
女	2403(83.3%)	481(16.7%)			998(34.6%)	1886(65.4%)		
男	2649(84.4%)	488(15.6%)			1058(33.7%)	2079(66.3%)		
卒中类型			58.7	***			-	-
出血性	1621(78.8%)	435(21.2%)			-	-		
缺血性	3431(86.5%)	534(13.5%)			-	-		
通气类型			544.0	***			46.0	***
氧疗	3063(93.6%)	209(6.4%)			1042(31.8%)	2230(68.2%)		
高流量养料	76(82.6%)	16(17.4%)			18(19.6%)	74(80.4%)		
非侵入性同期	73(89.0%)	9(11.0%)			23(28.0%)	59(72.0%)		
有创行通气	1761(70.9%)	724(29.1%)			923(37.1%)	1562(62.9%)		
气管切开	79(87.8%)	11(12.2%)			50(55.6%)	40(44.4%)		
AKI 分级			270.9	***			23.8	***
正常	1949(91.1%)	190(8.9%)			803(37.5%)	1336(62.5%)		
阶段 1	745(84.7%)	135(15.3%)			310(35.2%)	570(64.8%)		
阶段 2	1720(83.6%)	337(16.4%)			665(32.3%)	1392(67.7%)		
阶段 3	638(67.5%)	307(32.5%)			278(29.4%)	667(70.6%)		

表中连续性变量以平均数±方差表示，采用 t 检验，分类变量以计数（组内占比）表示，采用卡方检验；

*代表 P<0.05，**表示 P<0.01，***表示 P<0.001。

研究数据集包含 125 个特征，应用 GES 和 MIC 方法作为特征筛选的方法，图 1A 为 GES 生成的因果图，从中选择结局变量的直接父节点作为因果特征，控制 MIC 选择的特征数量于因果数

量相同，预测模型以随机森林为例，在不同的卒中类型比例的随机采样数据集中，预测的 AUC 值如下图 1B 所示。从中明显看出，随着缺血性患者比例的提高，模型的预测准确度都有所下降，但 GES 选择的因果特征的效果均优于 MIC 筛选方法。同时模型的预测准确率也比当前研究的效果更好。评价指标的比较如表 2 所示。表中可知 GES 因果特征下，模型的预测效果显著高于 MIC 特征筛选方法。

表 2. 随机森林模型在不同特征下预测评价指标

评价指标	MIC 特征	GES 因果特征	统计量（t）	p 值
AUC	0.88±0.02	**0.89±0.02**	-11.77	**7.30E-30**
AUPR	0.63±0.07	**0.68±0.08**	-10.07	**1.17E-22**
F1	0.58±0.06	**0.61±0.07**	-7.84	**1.31E-14**
MCC	0.49±0.06	**0.53±0.07**	-9.46	**2.62E-20**
准确率	0.79±0.05	**0.82±0.04**	-8.33	**2.94E-16**
召回率	0.8±0.07	**0.81±0.06**	-2.16	**3.11E-02**
特异度	0.79±0.07	**0.82±0.06**	-6.69	**3.94E-11**

图 1C 中展示了两种方法的特征韦恩图，如所示，由此可知两种方法筛选出的特征差异较大。为了进一步展示因果特征筛选方法与普通的 MIC 特征筛选方式的区别，我们绘制了特征与结局事件的皮尔逊相关系数的热力图，并按照韦恩图进行分组，如图 1D 所示。其中可以明显看到，因果方法选择的特征具有极差大的特性，相关系数分布较为极端，而 MIC 方法的相关系数大多分布在 0 的附近。但目前仍有大量的研究应用 MIC 作为一种特征筛选方式[25]。

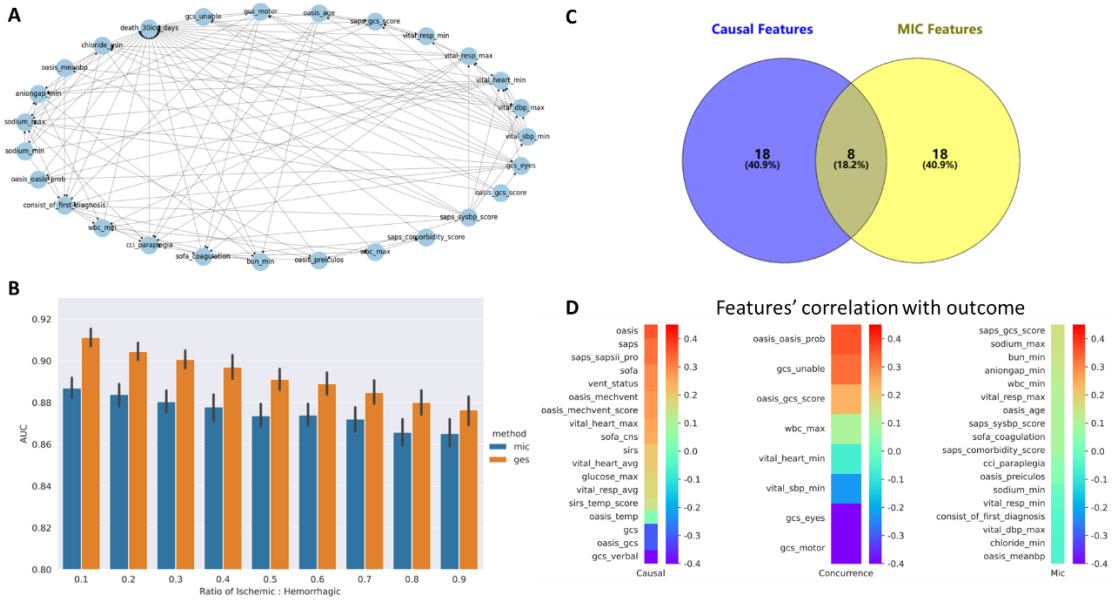


图 1 卒中 ICU 内死亡预测结果图

4 讨论

我们的研究发现应用因果发现的方法进行机器学习的特征筛选，相比于 MIC 的方法能够显著提高模型的效果，从可解释的角度来看，GES 因果特征与结局变量具有比较强的相关性，而 MIC 方法的特征则以弱相关为主。但研究仍然有许多待改进的部分如应该与更多的特征筛选方式进行横向对比，以进一步验证结论的可靠性；同时结论应该在更多的机器学习模型中进行验证。并且应该添加外部验证来进一步证明结论的准确性。

PO-023

基于区块链的医患电子健康数据交换体系的研究

胡佳[1] 杨麟[1] 李颜[1] 陶新娟[1] 曹丁文[2*]
1. 岳阳市人民医院信息科
2. 岳阳市人民医院设备科

1 引言

随着信息技术的发展，不同医疗实体之间的患者电子健康信息调阅已越来越频繁。电子健康信息主要由患者的电子健康记录（Electronic Health Record, EHR）组成的。EHR 在世界范围内得到了广泛应用，跨医疗实体的健康信息及时地交换可以显著降低医疗成本，提高医疗质量和增强疾病监测的效果。区块链是一种相对较新的技术，可用来提高物联网医疗系统领域的安全和隐私水平。由于区块链的不可变特性，EHR 数据可以存储在安全的医疗区块中，以维护数据的隐私和完整性[1,2]。由于其分散的性质，区块链可以在不依赖第三方的情况下实现某种程度的安全性[3]。然而，在医疗环境中，EHR 的数据源涉及许多结构化和非结构化数据类型——文本和图像，这些数据特别是图像影像数据占用空间可能会非常大，无法通过现有的物联网络系统进行传输。需要对现有的区块链和物联网解决方案进行系统性调整，以便能高效地、有隐私保护地交换医患电子健康数据，并能实时监控患者的状态。

一方面，区块链是一个分布式的体系，可以在多个实体之间共享数据，并能保持数据的完整性和一致性[4]；另一方面，将边缘计算分布式部署在各个终端，它能够快速、便捷地检测边缘不良事件，并在医院过度拥挤的情况下提供医疗实体的及时响应。因此在本文中，我们提出了一种新的基于区块链技术的高效、安全、智能的医患电子健康数据交换框架，该框架在医生/患者这一端使用边缘计算，并部署采用群交换技术的私有区块链，利用群交换技术将患者的 EHR 数据高效安全地传输给利益相关者（如医生）。这样，可以提高医疗物联网的安全性和隐私性。同时，医生也可以使用相同的安全基础设施向患者发送药物或诊断报告。

2 资料和方法

2.1 数据收集：

因为 EHR 文件可以是任何类型的，为了不失一般性，本文通过计算机模拟了 30 个患者 3 种类型的数据：体温、血压和脑电图（Electro Encephalon Graph, EEG)，其中体温和血压数据为 TXT 格式（模拟文本文件），EEG 数据为 PNG 格式（模拟图像文件格式）。

时间	体温（℃）	血压（mmHg）
2023-01-01 07: 00: 00	36.5	122/75
2023-01-01 11: 00: 00	36.2	125/77
2023-01-01 16: 00: 00	36.4	123/74
2023-01-01 20: 00: 00	36.3	126/76
2023-01-01 23: 00: 00	36.1	124/74

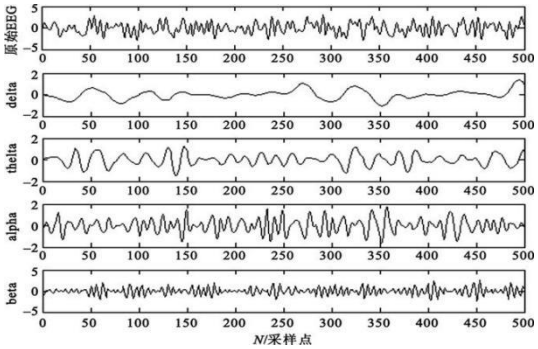


图 1 某患者的三种类型数据：体温、血压和 EEG

2.2 构建高效安全的医患数据交换体系架构

为了解决大多数方法存在的可扩展性差、计算成本高、易遭受网络攻击和响应速度慢的问题，我们提出了基于边缘计算和群交换的区块链医疗数据交换体系。该体系由多种电子健康实体组成。该体系分为两个主要网络：本地网络和区块链网络。本文所提的数据交换体系如图 2 所示。

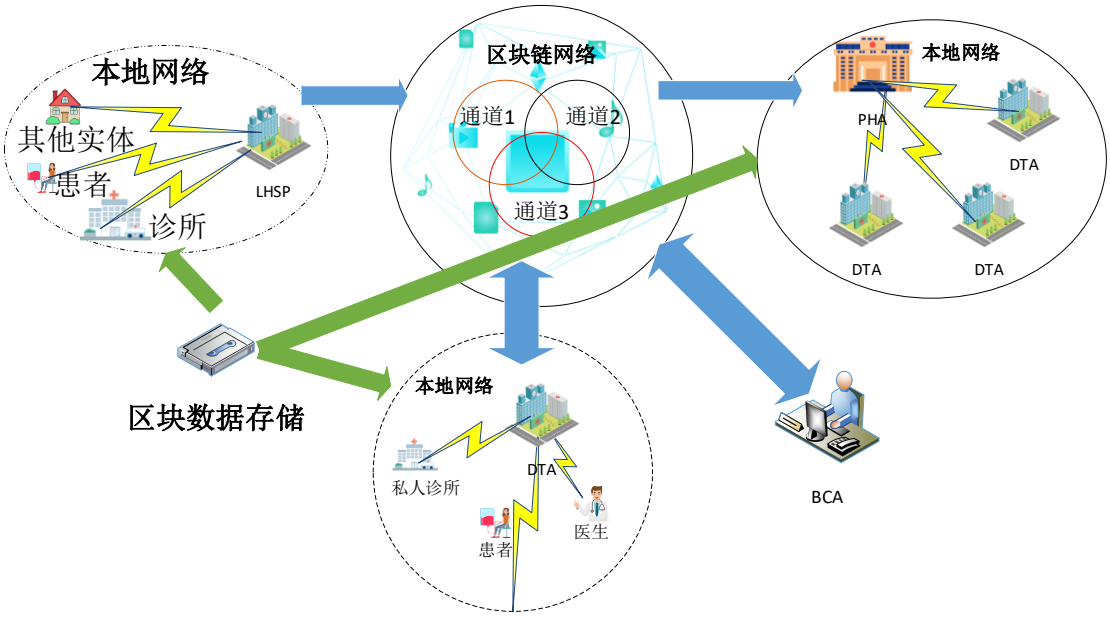


图 2 区块链框架下的高效安全医患数据交换体系示意图

本文提出的数据交换体系架构分为两个主要的网络：本地网络和区块链网络。本地网络主要是连接患者身上或患者周围的数据源到本地医疗服务机构（local healthcare service provider, LHSP），比如私人诊所、社区医院等。它包含以下几个主要的组件：

(1) IoMT: (Internet-of-Medical Things, IoMT) 连接/靠近患者的 IoT 设备组合，用于监测智能辅助环境中的健康状况和活动。包括植入式或可穿戴式的传感器、智能手机、IP 摄像机以及外部医疗和非医疗设备等。

(2) LHSP: LHSP 是一种医疗设施，包括私人诊所、社区医院等，用于监测当地患者并为其提供所需的医疗服务，记录患者的状态，并在需要时提供及时的急救服务。

LHSP 在监测患者状态方面发挥着重要作用，不仅是在医疗设施内部，而且在此类设施外部也能提供高级的护理服务，甚至可以与患者的家庭圈子连接，以跟踪患者的病情。

(3) DTA: (Diagnosis and treatment area, DTA)，诊疗区域，包括私人诊所、社区医院和大型医院等，它们是能够提供精确诊断和治疗的医疗服务机构，负责数据存储，应用复杂的数据分析技术，以及在公共卫生实体之间共享重要的健康相关的信息。利用边缘计算能力，每个实体都可以在区块链内共享医疗数据之前验证 DTA 医疗数据的真实性和完整性。

(4) PHA: (Public health administration, PHA)，公共卫生管理部门，包括医疗行业协会、卫健委和疾控中心等政府部门和协会组织，主要作用是协调不同卫生实体，监测 LHSP 和 DTA 的医疗服务进行监管、监控和评估，以确保可接受的医疗质量。

(5) 其他实体：其他实体也可以是本地网络的一部分，如医学药学科研组织、保险公司、药店等。其主要的作用是通过本地网络获取相应的权限提取感兴趣的数据以完成本实体的工作，更好地为患者服务。

本文提出的数据交换体系最终的目标是通过区块链网络实现以下主要功能来满足 LHSP、DTA、PHA 等监控和诊疗活动能够安全、高效地执行：A 数据收集、特征提取和患者状态监测，以确保关键医疗事件的高可靠性和快速响应；B 不同实体之间随时随地安全访问数据。它包含以下几个主要的组件：

(1) DS: Data Sender，数据发送方，主要包括患者、私人诊所、医院等实体。主要功能是收集不同患者的数据；从获取的数据中识别与患者状态相关的信息性特征；利用已识别的特征检测患者状态的重大变化。

(2) 区块链数据存储：该模块存在于区块链各个节点中，对网络中运行的数据以分布式存储在其中。

(3) BCA: (Block Chain Administrator, BCA)，区块链管理者，执行链码管理功能，当区块链网络收到 DS 以“事务”的形式发送的数据并上传到附近的 BCA 时，BCA 充当验证者：它将未经验证的数据块分发给自己的验证模块进行验证，触发验证者之间的共识过程，并将验证的数据块插入区块链中。

(4) Channel: 区块链中数据流的通道。在本数据交换体系中，本文定义了三个数据流通道：A 紧急的数据流通道（如影响患者生命安全的数据流向本通道）；B 非紧急的数据流通道，但需要高安全级别关注（如不影响患者生命安全但异常的数据和机密法律数据流向本通道）；C 正常数据流通道。

2.3 医患侧边缘计算的实现

在不失一般性的情况下，本文提出的数据交换体系模拟 30 个患者在三个不同疗程中服用特定药物前、中、后的监测场景。采集的数据包括 EEG 信号和常规观测的体温、血压数据。

A 特征提取

本文数据交换体系的第一步是边缘计算识别主要的统计特征，这些特征信息丰富、具有代表性，并且与患者的生命体征相关。在众多的服药数据中，医生很难快速区分和检测到显著的变化。通过分析数据，发现数据呈现不同的统计特征，因此本文使用均值、方差、均方差、峰度来提取特征的变化。

均值：$\mu_{ij} = \frac{1}{N}\sum_{k=1}^{N} y_{ij}(k)$

方差：$\sigma_{ij}^2 = \frac{1}{N}\sum_{k=1}^{N}\left|y_{ij}(k) - \mu_{ij}\right|^2$

均方根：$R_{ij} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left|y_{ij}(k)\right|^2}$

峰度：$v_{ij} = \frac{\frac{1}{N}\sum_{k=1}^{N}\left(y_{ij}(k)-\mu_{ij}\right)^4}{\left(\frac{1}{N}\sum_{k=1}^{N}\left(y_{ij}(k)-\mu_{ij}\right)^2\right)^2}$

其中，$y_{ij}(k)$是通道i和患者j的输入体温、血压或 EEG 信号值，N是样本数。因此，对于给定的患者j，将针对每个生命体征信息通道i计算上述特征，以表示患者在N个样本的时间窗口内的状态。

B 变更检测和共享

边缘计算的第二步是检测患者状态的主要变化。如检测到重大变化（即紧急情况）将通过区块链紧急通道共享紧急通知，以及可能需要进一步调查原始数据；如检测到微小/无变化，将只共享获得的特征；如只检测到一个或两个体征的变化，这意味着 IoMT 出现了一些误差，测量可能不准确，因此会通知医生/患者重复测量。

为此，本文定义了一个变化指标向量λ和患者状态η，其计算过程如下式子组，在本文中，$P = 30$为患者数量，$C = 3$为通道数量。

$$\delta_{ij} = \mu_{ij} + \sigma_{ij}^2 + R_{ij} + v_{ij}$$

$$\bar{\delta} = \frac{\sum_{j=1}^{P}\sum_{i=1}^{C}\left(\delta_{ij}^{befor} + \delta_{ij}^{during} + \delta_{ij}^{after}\right)}{3CP}$$

$$\lambda_{ij} = \left[\frac{\left| \delta_{ij}^{befor} - \delta_{ij}^{during} \right|}{\delta} + \frac{\left| \delta_{ij}^{during} - \delta_{ij}^{after} \right|}{\delta} \right]$$

$$\eta_j = \begin{cases} Major,\ if \quad \|\lambda_i - \varphi\|_0 > 2 \\ Minor,\ if \quad \|\lambda_i - \varphi\|_0 = 0 \\ Repeat,\ if \quad 0 < \|\lambda_i - \varphi\|_0 \leq 2 \end{cases}$$

其中φ为阈值，在本论文中为$\varphi = 20$，这是一个经验值，此值越大，则状态变化越不敏感。$\|\cdot\|_0$表示零范数，即该矩阵中取零的个数值。

通过上述两步，本文的数据交换体系就能获取患者的状态，从而优化通过区块链共享的数据。同时，它可以快速检测患者状态的主要变化并保持较低的复杂性，因此可以在任何移动边缘上实现。

2.4 医患电子健康数据群交换技术在私有区块链中的应用

群是一种用于 P2P(点对点)分散文件传输的范例，其传输文件的内容通过各自的散列来寻址，即内容可寻址。群的主要优势是低延迟的激励机制。群的意义在于，去中心化网络的区块链中现有节点可以从指定的源获取文件。群交换是处理去中心化网络中各节点的激励机制、通信机制和实现自稳定对等池的策略，它可以在挖掘数据时减轻可检索性的证明的计算压力。

本文提出了一种新的群交换技术模式，整个私有区块链网络中群交换节点或者处于监听状态，或者宣告在实体内上传或下载 EHR 文件。首先，患者使用群交换节点利用混合密钥加密技术，通过指定医生的公钥加密 EHR 数据。随后加密文件被上传到监听群交换模块，并宣告私有区块链中其他的群交换节点的接收情况。同时，获取 BCA 授权后使用标准 SHA-3 算法对 EHR 文件进行散列，并将其值插入私有区块链中。因此本过程会有两个输出文件：加密的 EHR 文件和加密 EHR 文件的哈希值。

指定医生需要提供三个参数（加密的 EHR 文件、私有区块链上的散列哈希值和自己的私钥）使用混合密钥解密技术进行解密。如果块的散列值与输入的 EHR 文件匹配，则该 EHR 文件可以从私有区块链中下载到医生的群交换节点上。否则，攻击者不能下载加密的 EHR 文件，即使下载来也无法打开。这就赋予了群交换节点双层安全性和隐私性。这就可以利用 EHR 文件的双层保护机制实现了一个健壮且防篡改的系统。在这个体系中，借助于群交换范例，所有的群交换节点都被更新，新加密的 EHR 文件被包含到私有区块链系统中，并且可以根据需要，EHR 文件可以被移交给特定的群交换节点。同时，EHR 哈希块在前面提到的工程中提供了双层结构，其中通过私有区块链实现不变性和透明性，其中 EHR 数据只能在预先指定的节点组内访问。

根据患者的 EHR 数据，医生诊断患者的病情并加密诊断报告。诊断报告的所有群交换和散列在系统内以无缝遍历的方式上传至私有区块链中。患者最终能下载并读取指定医生发送的诊断 EHR 数据，以便采取必要的措施治疗疾病。不管是患者/医生上传数据，还是下载数据，EHR 文件双层保护机制可以无缝运行，以增强 EHR 数据源的安全性、隐私性、不变性和透明性。本数据交换体系采用群交换范例、安全哈希块和混合加密-解密技术，使得该体系在本质上非常严格且防篡改。

3 结果

3.1 群交换范例对在实验中的运行模式

群交换范例由两部分组成：群命令接口和群守护程序，如图 3 所示监听群节点向群命令接口提供有效的 EHR 相关参数，然后群命令接口调度器将输入发送到命令模块。命令模块可以通过传统的 PreRun、Run、PostRun 函数处理命令。在脱机模式下，EHR 文件直接从调度器发送到输出端，用于联系其他群节点。在线模式下，群守护程序将激活群 HTTP 服务器，通过类似的命令模块处理 EHR 文件。群核心 API 和核心群节点通过 P2P 通信，HTTP 服务器和调度器支持群守护程序高效地执行群交换操作。这种模式极大地帮助群节点为灵活的任务建模提供了双重选择。因此，患者和医生的群集节点可以在高安全性和可用性级别上相互通信，从而实现 EHR 文件的平滑传输。

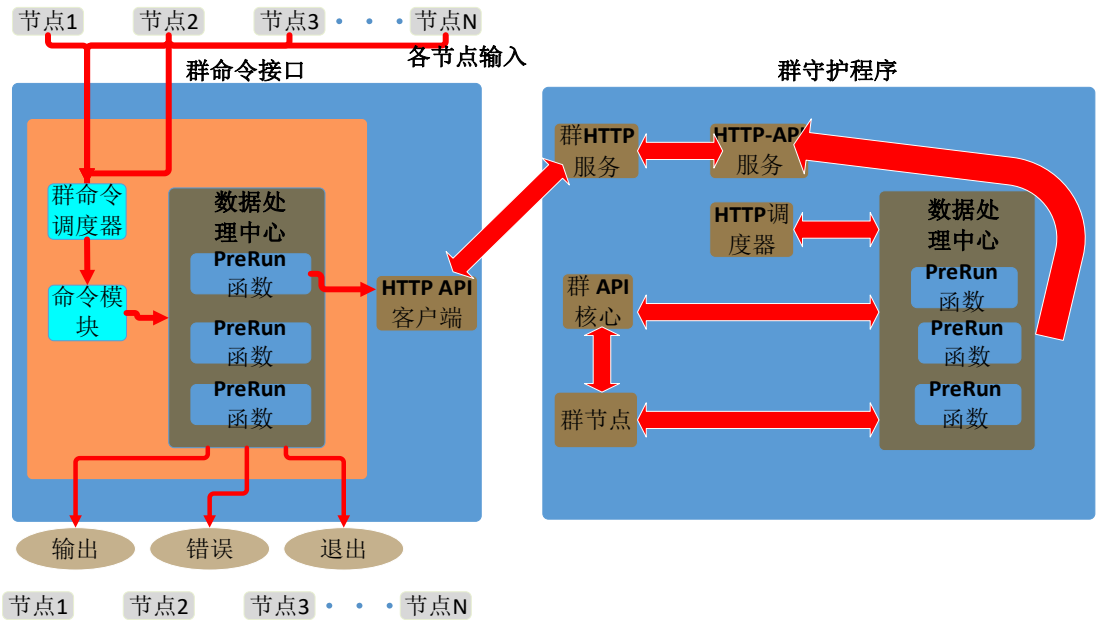
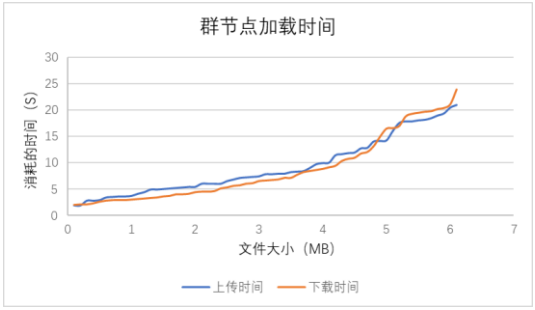


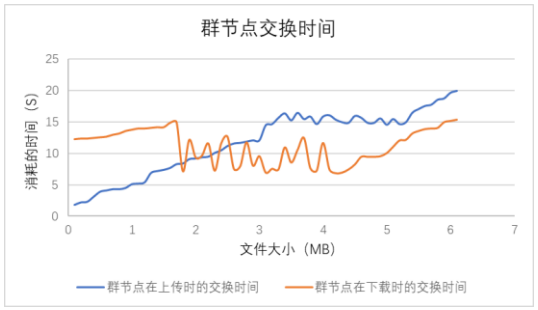
图 3 群交换的核心模型示意图：群中各节点向群命令接口提供有效的 EHR 相关参数，通过调度器和命令模块将参数发给数据处理中心，并输出处理结果

3.2 群交换技术在模拟实验中的结果

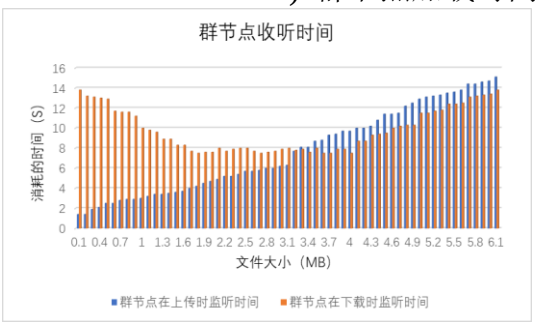
下面分别通过群节点加载时间、群交换时间、群收听时间、群公告时间、群可用时间几个方面分析本文所提的体系在医患健康数据交换中的性能。



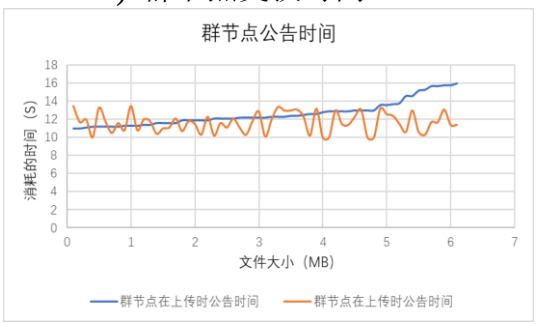
A) 群节点加载时间



b) 群节点交换时间



c) 群节点收听时间



d) 群节点公告时间

图 4 模拟现实场景的实验结果

从图 4 中可以看出，在群节点交换时间实验中，EHR 文件的平均上传和下载时间存在差异，上传时间随着文件大小的增大而变长，而下载情况不同，有个“凹”字形的形状表现。在群节点收听时间实验中，上传时间随着文件大小的增大而变长，但下载情况呈现出 U 形。值得注意的是，群节点公告时间不会随着文件大小的增大而变化，但在下载阶段的群公告并不稳定，因为群公告是在下载阶段随机执行的。实验表明，本文提出的数据交换体系在区块链物联网场景、群节点交换和 EHR 传输的各个方面都表现出良好的通过性和 EHR 的可靠性。而现有的医疗数据交换体系花费时间大都在 100S 以上，甚至出现拒绝连接行为。

4 讨论

区块链是一种开源分布式账本技术，最早应用于金融行业[5]，最流行的应用是比特币加密货币，它一级证明区块链技术是稳定、安全和健壮的[6]。在医疗保健领域已有区块链的应用，如药品供应链[7]、临床试验管理[8]和病历管理[9]。然而，医疗领域的大多数区块链应用仍处于实施的早期阶段。文献[10]提出了一个利用美国 PHIPAA 法规（个人健康信息隐私和访问法案）来解决集可用性和隐私于一体的 EHR 数据安全问题的体系，该体系也讨论了 EHR 存储网络建模、无权限可用性和运营成本最小化方面的方案。为了利用 EHR 服务的供应能力，文献[11]开发了一个分散的数学模型，为患者和

其他利益相关者提供一个防篡改的 EHR 管理系统。在文献[12]中，研究了 LoRAWAN(Long Range Wide Area Network，远程广域网)协助的概率验证实施，其中物联网数据存储在群交换服务器中，但本项工作没有使用任何私有区块链网络，因此未能准确显示系统的内部状态图，使得系统更容易受到网络攻击。文献[13]研究了基于无人机的群交技术，通过共享密钥的方式在区块链网络中访问数据。因为缺少私有区块链，使得系统更易于访问，这可能导致终端用户之间的利益冲突，这也间接验证了在私有区块链中存储数据可以提高系统的安全性。

区块链中的安全力量来自体系中实体的集体资源，实体必须使用共识算法验证每个数据块，如委托权益证明（Delegated Proof of Stake, DPOS）。任何对区块链体系的攻击必须击败整个群体的资源，才能破坏数据的完整性，这使得对区块链的攻击不切实际[4]。因此只要大多数实体是无伪的，就可以保证所获取数据的正确性。如 Wang 等人[14]利用区块链将患者、医院、卫生管理部门和不同的医疗机构联系起来，参与的实体数量很多，其中医院、卫生管理部门和众多医疗机构都可以被认为具有较高的数据安全性和无伪性，因此可以实现全面的病历安全共享和审查。Liang 等人[15]提出了一个以用户为中心的医疗数据交换方案，使用移动应用程序从可穿戴设备收集数据，然后使用私有区块链与医疗机构和保险公司共享数据。Xia 等人[16]提出了一个支持不同实体之间共享医疗数据的系统，该系统利用智能合约和访问控制方案来检测共享数据上的恶意活动，并拒绝违规实体的访问。然而，上述大多数方法都存在可扩展性差、计算成本高和响应速度慢的问题。

本研究提出了一种新的医患健康数据交换体系，该体系通过采用区块链技术作为框架，群交换技术整合区块链物联网系统，用于医疗系统中的 EHR 安全、隐私地传输。隐私是为了防止未经授权的第三方访问，其目标是更改、转换、分发和隐藏敏感信息以防止在第三方系统处理期间破坏原始数据。由于其特性，区块链技术是保护物医疗网络中电子健康信息的比较好的解决方案，物联医疗网络是包含医患健康信息的区块链。对于医患健康数据来说，EHR 的不变性数据源对于诊断患者的健康状况非常重要。必须保留 EHR 数据，以便不进行任何更改。区块链的不变性特征有助于将 EHR 保留在一个完整的区块内。

在本研究中，我们模拟基于物联网的传感器节点，以生成连接到数据交换体系中的患者的体温、血压和 EEG 图。此外，此类物联网传感器、文本和图像形式的附加患者健康数据已与 EHR 合并，以便在区块链生态系统内进行处理。在本文中，我们设想一数据聚合器模块，以 EHR 字符串的形式连接三个不同的物联网传感器数据。此类 EHR 随后会选择性地增加关于患者的其他必要细节，以便在私有链和群交换平台中进行拓展。

鉴于医院的负担越来越大，特别是在传染病传播期间（例如最近的 COVID-19 暴发），人们总是建议将病情温和的患者转移到家庭护理中。然而，对来自不同地点的大量患者进行远程监控会给无线网络带来很大的负担。根据文件大小的增加，本研究提出的系统在上传阶段的群集加载时间平均增长了 64%，图像 EHR 文件也反映了类似的结果。在群交换阶段，在上传阶段，交换时间有适度的急剧增长，但在下载阶段，增长不那么快。还观察到，小型 EHR 文件在上传过程中花费的群交换时间较少，而 EHR 文件在上传过程中花费的时间最少，从 3.3 MB 到 4.6 MB 不等。

下一代医疗保健系统正在通过整合新兴技术来形成，以在医疗保健服务方面提供根本性的改进。因此，本文提出的基于区块链技术的医疗数据交换体系，有望实现有效的大规模医疗数据交换，并提供各种医疗实体产生的大量医疗数据的安全传输。特别是，本文的边缘计算方案可以有效监控患者状态，以确保对紧急事件的早期检测和快速响应。并通过群交换技术范例使用混合加密/解密机制管理、监控与控制患者和医生之间的 EHR 流。使用区块链技术的群交换技术，在不需要第三方的情况下增强 EHR 便利化，同时将高可用性、灵活性、透明性、隐私性和安全性注入本文的体系。

由于区块链是一项相对较新的技术，没有很多成功的区块链模型可供遵循，在实际工作中会面临一些问题，比如目前只有少数基于该技术的成果倡议，这就造成了一个不确定的局面。本文的工作基于模拟的场景进行，在大规模网络环境中还有待验证。值得一提的是，区块链技术并不是医疗数据交换的唯一解决方案，本文验证了区块链技术在医疗数据交换中的独特性，为区块链技术应用于医疗系统提供了一个可行的方案。

PO-024

基于 MRI 多个解剖参数预测术后尿失禁的人工智能方法

李佳坤 沈百荣
四川大学华西医院

背景 深度学习方法在医学领域的应用越来越多；然而，它们缺乏可解释性仍然是一个挑战。Captum 是一种可以通过计算特征重要性权重来解释神经网络模型的工具。尽管 Captum 是一个可解释的模型，但它很少用于研究医学问题，而且在接受机器人辅助的前列腺癌根治术（RARP）后，关于癌症患者的 MRI 解剖测量数据也很少。因此，使用多种解剖 MRI 测量的失禁预测模型是有限的。

方法 我们通过分析 MRI 测量结果，探讨了深度学习模型预测失禁的能量效率。我们分析比较了各种统计模型，为可解释的深度学习模型的临床应用提供了参考实例。本研究纳入了 2019 年 7 月至 2020 年 12 月期间在我们机构接受 RARP 的患者。使用这些患者的一系列临床 MRI 解剖测量来发现失禁特征，并使用一系列统计方法和计算模型初步评估其对失禁的影响。

结果 年龄和其他六个解剖测量结果被所提出的 UINet7 模型确定为失禁的前七个特征，准确度为 0.97，这些特征中的前四个也是通过初步统计分析发现的。

结论 由于临床数据和适用模型的局限性，本研究填补了 RARP 后失禁特征深入研究的空白。我们提供了一个将深度学习模型应用于临床问题的开创性例子。深度学习模型的可解释性分析具有临床应用的潜力。

PO-025

基于图神经网络的癫痫脑电发作预测研究

石雨菲[1]，王瑶[1]，陈子怡[2]，周毅[1*]
1 中山大学中山医学院
2 中山大学第一附属医院

引言

癫痫（Epilepsy）是一种由大脑神经元群高度同步化异常放电引起的反复发作的神经系统疾病[1]。目前癫痫患者在全球范围内近 6500 万[2]。癫痫患者人群中，约有 30%的癫痫病人无法通过药物治疗得到有效的控制，我们称之为顽固性癫痫患者[3]。在顽固性癫痫患者中有 25%可以通过手术除去癫痫灶的方法来进行治疗[5,6]。剩下的癫痫患者的癫痫症状尚没有任何合适的治疗手段可以控制[4]。这几种类型分布经计算如图 1 所示。

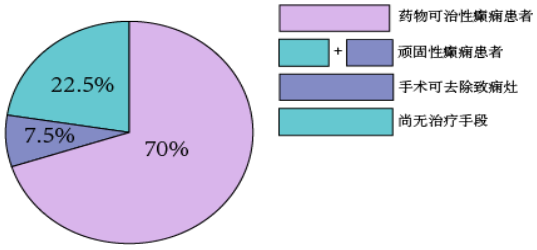


图 1 癫痫患者人群分布扇形图
Fig.1 Fan chart of population distribution of epilepsy patients

本文的研究工作正是面向图 1 中 22.5%左右尚无有效治疗手段的难治性癫痫患者。对于无法去除癫痫灶的难治性癫痫患者，临床能做的主要是缓解癫痫病人的痛苦。而缓解病人痛苦的最好办法就是对癫痫发作进行预测[7]，并给出一定的预防和缓解措施。
研究出一种能有效预测癫痫发作的模型和算法就是本文的主要研究工作所在。在以往的研究者所做的分类实验中，训练数据和测试数据常常来源于同一病例的 EEG 信号，跨患者的癫痫脑电检测研究成果有限，模型算法泛化性不足；针对于公共癫痫脑电数据集的研究成果可观,但针对于临床真实数据的研究还不够,现有算法的鲁棒性不足。本着解决现有检测模型的泛化性和鲁棒性不足问题的出发点，以及考虑到脑电信号的时序性，我们计划使用图神经网络对癫痫脑电发作前期特征进行分类，将变体门控递归单元（GRU）和图卷积神经网络（GCN）相结合，最后使用多层感知机（MLP）生成预测序列。

1 国内外研究现状及发展趋势

1.1 基于全导联的癫痫脑电的自动检测研究

头皮脑电（electroencephalogram，EEG）中包含了大量的生理和病理信息，在癫痫等脑科疾病的诊断、预测、治疗中起着非常重要的作用。其中针对于临床发作不明显的患者常常需要借助于脑电图完成癫痫的诊断。但是其数据量之大，其中还夹杂着医务者的主观因素不仅使得医务工作者的工作不堪重负，也使得诊断错误的情况时常发生。因此，癫痫脑电信号的自动化分类诊断在当前临床应用上有着巨大的潜力。

癫痫的检测更偏向于智能诊断软件以及可穿戴设备，这就要求实现的算法具有计算方便快捷的特点，癫痫发作检测的任务包括区分正常患者和癫痫患者，以及癫痫患者是否处于发作间期（癫痫患者非发作时期）或者发作期（癫痫发作的时期）。一般来说，癫痫检测过程分为两部分[8]：特征提取和分类。基于人工特征和机器学习分类器的技术研究很多。一方面采用时频分析、非线性动力

学、复杂度、同步和累积能量增量方法作为特征提取方法。另一方面，机器学习分类器包括贝叶斯网络、传统神经网络和支持向量机（SVM）等。事实上，特征分类器工程技术已经成功地用于癫痫检测任务。然而，特征提取基于有限的手工工程操作。最重要的是，考虑到不同患者的癫痫发作特征不同，并且可能随着时间而改变，有必要从 EEG 数据中自动提取和学习信息性特征。

在过去十年中，深层学习的最新进展在检测和预测数据分析方面，特别是在卫生保健和医疗实践中引起了更多的关注。它是一个强大的计算工具，能够自动地从数据中学习特征。研究已经证明，深层多层感知器神经网络比传统的 logistic 回归和支持向量机具有更好的性能。近期有研究学者在小型波恩大学公共数据中，采用 13 层深层卷积神经网络（Convolutional Neural Network,CNN）算法的准确率、特异性和灵敏度分别为 88.67%、90.00%和 95.00%[9]，递归卷积神经网络学习了癫痫发作的一般空间不变表示，大大超出了先前在跨患者分类器上获得的结果[9]。深层无监督神经网络用于癫痫发作的自动检测，由于稀疏策略，可能会遗漏重要信息。其他的技术如深度信度网络、转移学习等也被应用于癫痫检测[10]。其中卷积神经网络在图像处理领域奠定了癫痫检测研究的基础。

开发实用可靠的智能诊断算法对癫痫自动检测具有重要意义。虽然已经采取了许多努力来推动这一领域，但我们必须得出结论，癫痫检测分析尚未进入临床实践。

1.2 基于深度学习的癫痫发作预测研究

癫痫发作预测在于检测癫痫发作前的信号特征变化。大量研究发现，在发作间期和发作期之间存在一个过渡期，称为癫痫发作前期[11]。因此，癫痫发作的预测可以被视为发作前期的早期检测。图 2 给出了癫痫发作过程不同时期的过渡。

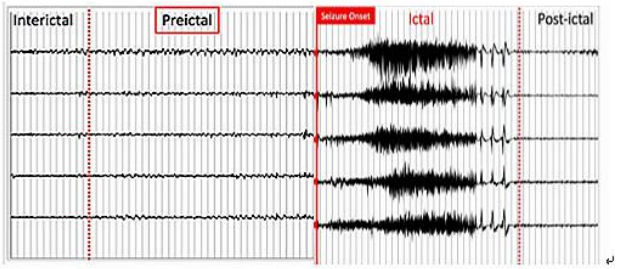


图 2. EEG 癫痫发作过程

Fig.2 EEG Epileptic seizure process

癫痫发作过程一般从发作间期过渡到发作前期再到发作期直至发作结束[12]。结合癫痫发作不同时期转变，建立相关癫痫综合分析模型来研究预测问题，是癫痫研究的一大难点[12][13]。《柳叶刀神经病学》中的开创性研究表明，连续跟进 15 位癫痫患者近三年的 EEG 数据表明预测癫痫发作具有可行性 [14]。国内一些学者及本课题，也结合临床数据和动物实验进行了癫痫预测的有关尝试[14]。但是基本上都是基于小样本的传统机器学习方法，导致癫痫预测的研究缺乏统计验证。

癫痫预测的重点在于检测发作前期，一种方法是因此根据发作间期设定正常阈值，通过发作前期是否超过阈值来判断。Aarabi ,A 和 He,B 计算发作间期的特征值确定阈值置信区间，设置三次阈值，如果发作前期特征值超过该阈值，则标记发出预警[15]. 此外对数据进行标注转化为分类问题。使用若干种分类器包括径向基网络，BP 网络，Elman 网络以及贝叶斯网络等比较预测癫痫发作的性能[15]. 然而这涉及实现非手工设计的包括特征提取和长期预测模型的长期系统。需要进一步选择以匹配每个患者。非深层学习算法在脑电分析中是一个有限的选择，并且已经广泛地与各种特征提取算法相结合。这种情况很容易导致缺少统计验证的算法，这需要大量的数据。

图神经网络(GNN)旨在建立图论下的神经网络来处理图域中的数据。图卷积神经网络(GCNN)[16]是将 CNN 与谱理论[15]相结合，对传统 CNN 方法的扩展。与经典的 CNN 方法相比，GCNN 在处理离散空间域信号的鉴别特征提取方面具有更大的优势。更重要的是，GCNN 方法提供了一种有效的方法来描述图中不同节点之间的内在关系，这将为研究脑电情绪识别过程中脑电多通道之间的关系提供一种潜在的方法。基于 GCNN 模型的成功，每个脑电信号通道对应一个顶点节点，而两个不同顶点节点之间的连接对应图的一条边。虽然 GCNN 可以根据不同节点的空间位置来描述节点之间的连接，但在应用它建立情绪识别模型之前，我们需要预先确定不同脑电通道之间的连接。另一方面，脑电通道间的空间位置连接与功能连接不同。也就是说，更紧密的空间关系

不一定能保证更紧密的函数关系，而函数关系将有助于判别性脑电特征提取在情感识别中的应用。因此，根据图节点的空间位置预先确定其连接是不合理的。

2.实验设计

2.1 实验整体设计概述

本论文采用循环神经网络(RNN)来学习数据的时间特性。循环神经网络 RNN 是用于解决序列预测和序列分类等数据具有时序特性的任务的最有效的模型之一。早期简单的 RNN 中存在梯度爆炸和梯度消失的问题因而不能学习太长的序列，研究者们为了解决这个问题提出了门控机制的循环神经网络，常见的门控机制循环神经网络有 LSTM、GRU。实际上 GRU 是 LSTM 的改进版本，它的参数比 LSTM 更少，运算速度比 LSTM 更快，但可以达到和 LSTM 相同的效果。因此在本论文中系主要使用 GRU 来学习时序特性，GRU 的更新公式如 2.1 所示。其中 是上一个时刻 GRU 的输出或隐藏状态， 和 用于控制过去记忆的遗忘和新信息的记忆。 为 t 时刻的输出或隐藏状态，通常情况下会把每个时刻的输出作为该时刻的隐藏状态。

$$
\begin{aligned}
u_t &= \sigma(W_u[f(A, X_t), h_{t-1}] + b_u) \\
r_t &= \sigma(W_r[f(A, X_t), h_{t-1}] + b_r) \\
c_t &= \tanh(W_c[f(A, X_t), (r_t * h_{t-1})] + b_c) \\
h_t &= u_t * h_{t-1} + (1 - u_t) * c_t
\end{aligned}
\tag{2.1}
$$

本论文采用图神经网络(GNN)来学习数据的空间依赖特性。图是一种在生活中十分常见的数据结构，例如城市交通道路、人际关系和文件结构等，但传统的神经网络例如 CNN、RNN、MLP 只能处理欧式数据，而图是一种非欧数据。得益于近年来图神经网络的大力发展，如今构建一个神经网络处理这种非欧数据非常容易，常见的图神经网络有图卷积神经网络 GCN、图注意力神经网络 GAT 等，GCN 和 GAT 的原理不在这里过多赘述，但有必要列出 GCN 和 GAT 的更新公式，多层 GCN 的特征矩阵更新公式如 2.2 所示，其中， 上一层的参数， 表示这一层的参数。

$$
\begin{aligned}
f(X, A) &= \sigma(\hat{A}\, RELU(\hat{A} X W_0) W_1) \\
\hat{A} &= D^{-\frac{1}{2}} \overline{A} D^{-\frac{1}{2}} \\
\overline{A} &= A + I_N \\
D &= \sum_j \overline{A_{ij}}
\end{aligned}
\tag{2.2}
$$

2.2 数据与模型结构

本研究中，我们的预实验数据来自于波恩大学公共癫痫脑电数据集，正式实验数据拟取 13 名真实临床患者；计划使用图神经网络对癫痫脑电发作前期特征进行分类，将变体门控递归单元（GRU）和图卷积神经网络（GCN）相结合，最后使用多层感知机（MLP）生成预测序列。目前搭建了 4 种图神经网络模型，分别是 GRU、GCN、GAT 和 GCN-GRU(TGCN)，其中 GCN-GRU(TGCN)是先学习空间依赖再学习时间依赖的混合模型。本文假设导联信号之间存在一定的图的关系，因此使用图神经网络对脑电信号进行学习和分类预测，但本文没有提供图的先验信息，而是使用一个随机生成的且权重不变的图或学习一个图。GCN、GRU 和 GCN-GRU(TGCN)的模型结构分别如图 3、4、5 所示。

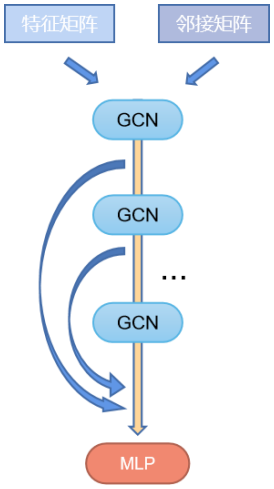


图3 癫痫分类中的GCN模型

Fig.3 GCN model in epilepsy classification

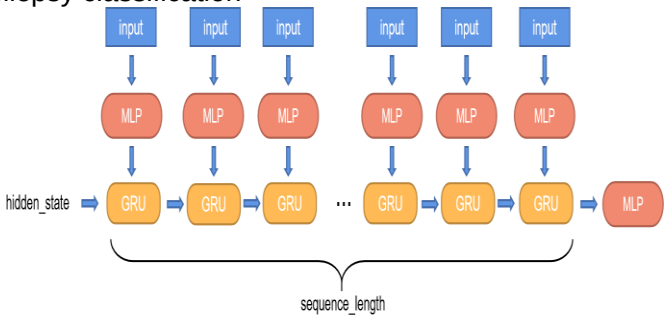


图4 癫痫分类中的GRU模型

Fig.4 GRU model in epilepsy classification

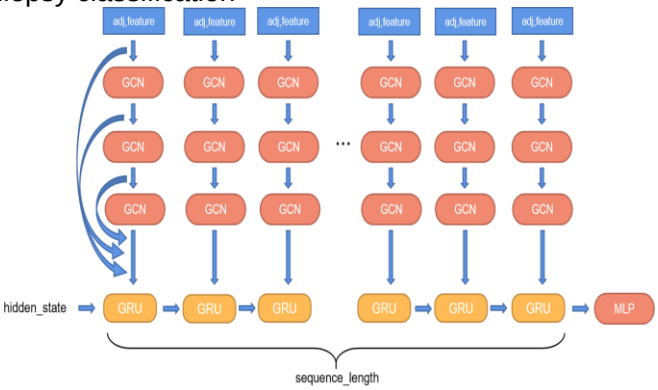


图5 癫痫分类中的GCN+GRU(TGCN)模型

Fig.5 GCN+GRU(TGCN) model in epilepsy classification

本文中拟解决的问题是一个二分类问题，即给定 n 个数据 以及对应的标签，学习一个模型对每一个数据能够正确预测其属于哪个分类。对于二分类问题最常用的损失函数是二分交叉熵，设模型对每个数据的分类预测结果为，则该分类预测结果的代价函数如下式所示：

$$Loss(Y,\tilde{Y}) = -\sum_{i=1}^{n}(y\cdot\log(\tilde{y}) + (1-y)\cdot\log(1-\tilde{y})) \tag{2.3}$$

二分交叉熵是一个凸函数，具有最优值点。二分交叉熵损失和 KL 散度 形式上非常相像，KL 散度如式 2.4 所示，因此优化二分交叉熵的过程近似等价于在优化模型输出分布和观测数据分布之间的距离。

$$KL(p \| q) = -\int p(x) \ln(\frac{q(x)}{p(x)})dx \quad (2.4)$$

2.3 评价指标与实验结果

我们常用以下三个参数来评价癫痫分类性能：敏感性(Sensitivity，Sen)，特异性(Specificity，Spe)，总识别率(Accuracy，Acc)。Sen 用来识别系统对于发作期脑电数据的敏感程度;Spe 用来识别系统对于发作前期脑电数据的识别能力[15]。Acc 表示识别系统的识别能力，具体公式表示如下：

$$Sen = \frac{TP}{TP + FN}$$

$$Spe = \frac{TN}{TN + FP} \quad (2.5)$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP}$$

其中，公式的含义如下表 1 所示。

表 1 模型评价的混淆矩阵

Tab.1 Confusion matrix for model evaluation

	真阳性	真阴性	合计
预测阳性	TP	FN	TP+FN
预测阴性	FP	TN	FP+TN
合计	TP+FP	FN+TN	TP+FN+FP+TN

图 6 至图 9 对比了几个模型在训练损失、准确率、敏感度和特异度上的差异。从结果上来看，GAT 模型的综合表现都不错，在敏感度上最高达到了 98.5%，但非常不稳定。而 TGCN 模型在准确率、特异度表现都比较稳定，其准确率达到了 87%，特异性也达到了 95.9%，并稳定在 94%以上。无论从哪一个指标来看 GCN 和 GRU 都不能成为最理想的模型。

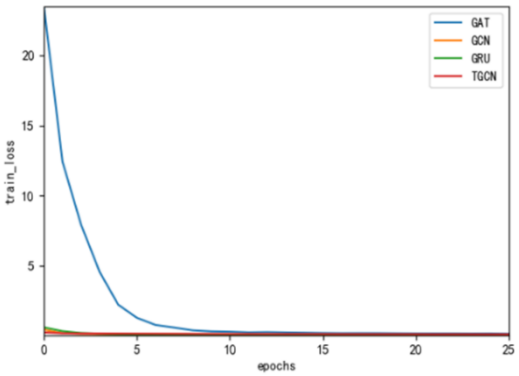


图6 训练损失函数对比

Fig 6 Training loss function comparison

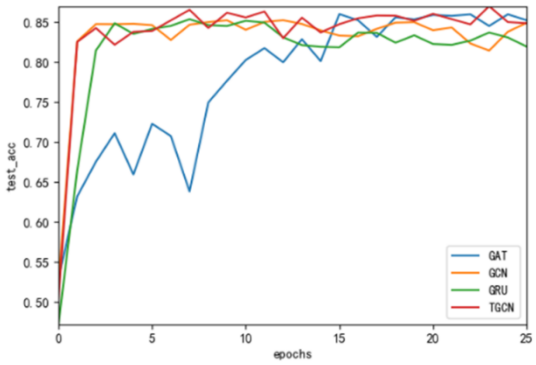


图 7 模型的准确率对比

Fig 7 Comparison of model accuracy

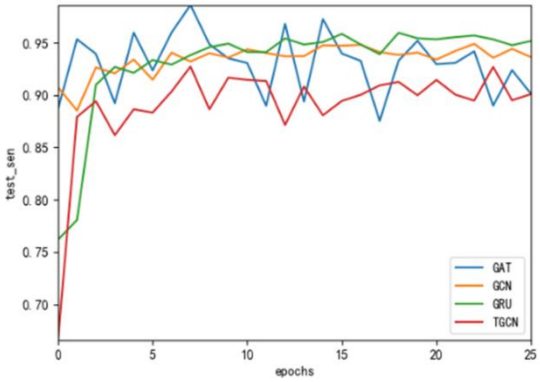


图 8 模型的灵敏度对比

Fig 8 Comparison of model sensitivity

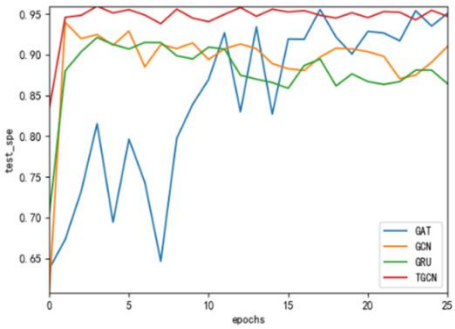


图9 模型的特异度对比

Fig 9 Comparison of specificity of models

3.结论与分析

本文在前人研究的基础上，设计了一种基于 GCN+GRU(TGCN)的分类模型。本文利用图神经网络学习一个将脑电数据映射到可分类特征向量的映射函数，并基于 Adam 梯度下降算法优化 和 Y 之间的二分交叉熵损失以达到一个对癫痫脑电数据分类的效果。其中 是 sigmoid 函数，常作为激活函数用于二分类任务中。本研究虽然使用了图神经网络，但并没有给网络提供一个邻近矩阵先

验，而是利用网络以一个非常低的学习率去学习这个邻近矩阵，背后的原因在于本文认为不同导联数据之间存在一定的关系，即可能存在一个相关系数矩阵表达各导联之间的联系强度，因此可以使用图神经网络来对癫痫脑电数据预测和分类。

从本文最后得到的结果来看，这些模型能够成功提取癫痫脑电数据的特征并用于分类，但在准确率上几个模型没有明显的差异。直觉上，本研究期望 TGCN 能够表现得足够出色，但实际结果显示 TGCN 甚至与 GCN 相比都没有明显的优势。一个可能的解释是 TGCN 中的 GRU 模块忽略了 GCN 学习的空间依赖。实际上脑电癫痫数据可以手工计算不同导联直接的相关系数，未来的工作将可能继续深入图神经网络在脑电癫痫上的应用，不使用学习邻近矩阵的方式，而是使用手工输入特征邻接矩阵。

本研究旨在提高癫痫检测模型的泛化性和鲁棒性。虽然目前在本数据集上的结果有了改进，但仍是针对小群体患者的研究。后续研究希望可以建立癫痫专病库，形成一个大型的针对各年龄层患者的癫痫数据集用以研究。

PO-026

元宇宙医学视角下自闭症儿童康复训练模型构建的探索

钟小云 王康龙 吕艳华
山西医科大学管理学院

1 引言

现如今，我们已经迈入元宇宙时代，元宇宙是人类运用数字技术构建的，由现实世界映射或超越现实世界，可与现实世界交互的虚拟世界，具备新型社会体系的数字生活空间[1]。元宇宙不是新技术，而是集成了一大批现有技术，包括 VR（虚拟现实技术）、AR（增强现实技术）、MR（混合现实技术）、数字孪生技术、区块链技术以及新一代信息技术 5G、人工智能和大数据技术[2]。随着科学技术的不断进步，元宇宙的应用范围也不断扩大，元宇宙与医学也将碰撞出新的火花，但元宇宙医学的概念尚处于探索和发展阶段，复旦大学附属中山医院呼吸科教授、元宇宙医学创立大会主席白春学在 Clinical eHealth 杂志称元宇宙医学是通过 AR/VR 实践的物联网医学，元宇宙医学是将虚拟世界拉入现实世界并多维互动的混合现实医学网络平台[3]。元宇宙医学可应用于治疗场景（精神治疗、模拟手术、远程医疗和医院就医全场景仿真等）、个体预防和疫情防护、康复和健康管理以及科研和教学中[4]。

自闭症(autism)又称自闭症谱系障碍(autism spectrum disorder, ASD)，被归类为一种由于神经系统失调导致的发育障碍，是以社会交往障碍、沟通交流障碍、兴趣受限和行为重复刻板为典型特征的精神类疾病[5][6]。2006 年自闭症被列入我国的残疾人目录，在国内 ASD 患病率的最新报告指出中国目前的患病率与西方国家接近，目前在中国有超过一千万人罹患 ASD，ASD 个体大都精神残疾且生活不能自理，该疾病导致的家庭、社会及医疗费用非常高昂[7]。自闭症多起病于婴幼儿，儿童自闭症患者居多，如果能及时对自闭症儿童进行康复训练，可以大大提高其在交往、认知、生活自理等方面的能力。

本研究在元宇宙医学视角下基于远程通信技术、VR、MR、AR 以及大数据和人工智能技术探索自闭症儿童康复训练模型的构建，以期为自闭症儿童的康复训练提供参考，助力自闭症儿童康复训练模型真正落地。

2 自闭症治疗现状

目前对自闭症的治疗方式有很多种，康复机构常规开展的康复治疗方法有行为干预治疗、感觉统合训练、言语训练和结构化教育，研究表明，常规的康复治疗也取得了一定效果，除了以上常规疗法，艺术疗法（包括美术治疗和音乐疗法）、同伴支持治疗、动物辅助疗法也逐渐兴起[8]。有临床研究及临床试验结果证明干细胞疗法在自闭症治疗中存在潜在价值[9]。在药物治疗方面，目前对

自闭症通常是对症治疗，没有特效药物，可能会使用一些抗精神病药、抗抑郁药、促进脑细胞功能药、维生素和其他补充剂等缓解症状。

3 构建自闭症儿童康复训练模型

元宇宙医学视角下自闭症儿童康复训练模型如图 1 所示。

图 1 元宇宙医学视角下自闭症儿童康复训练模型图

3.1 远程诊疗

远程诊疗（Telemedicine）是指通过远程通信技术对患者的病情进行初步诊断、制定治疗方案或者进行健康教育的一种诊疗方式。视频会议平台或其他医疗平台为专家对患儿进行远程诊疗提供了条件，专家可直接通过在线手段检测患儿的语言水平，也可对患儿家长和教师及相关治疗人员进行培训以及远程诊断实践，对自闭症儿童进行准确地诊断评估，专家通过视频观察、指导和监督相关人员，使其掌握各类干预手段的实施原理和操作方法，从而对自闭症儿童实施多样化的干预康复治疗[10]。

3.2 VR 创建沉浸式体验环境

VR(Virtual Reality)是虚拟现实技术，用户通过特殊的头戴显示器设备体验 VR 创造的虚拟世界，用户只能看到设备显示屏内的画面，而看不到真实的环境。

利用 VR 技术，可以为自闭症儿童创造一个仿真的、沉浸式的康复环境。这个环境包括各种视觉、听觉和触觉刺激，例如色彩鲜艳的场景、动态的视觉元素、音乐和声音效果等，以刺激他们的感知和认知能力。陈阳[11]等基于 VR 设计沉浸式自闭症康复系统，将自闭症儿童在社交、语言、行为及认知方面的障碍康复治疗转化为三种不同层级的能力训练，包括适应能力、交际能力和接受能力，构建社交发展训练、情绪识别训练、语言发展训练和特定恐惧治疗四大模块，由浅入深、先易后难进行训练。杨晓庆[12]基于 VR 设计出一款裸眼自闭症儿童辅助玩具“蘑菇点点”，引用高频音乐疗法，刺激听觉系统从而刺激大脑；引用全息投影拉近与他人距离感，增近感情；通过卡通图片方式教会孩子一些安全知识。安然[13]设计出一款以 VR 头显、操作手柄为硬件端，以配套 APP 为软件端的产品来帮助患儿进行辅助康复治疗。

3.3 AR 实现真实环境叠加虚拟元素

AR(Augmented Reality)是增强现实技术，可以识别真实环境，通过在真实环境中添加虚拟元素，利用设备的摄像头获得周围的环境信息，将预先制定好的虚拟画面叠加到周围环境中，来增强用户的现实体验。

利用 AR 技术，可以根据自闭症儿童的具体情况和需求，定制个性化的康复课程。这些课程包括各种虚拟活动，如社交互动、游戏、学习和运动等。通过这些活动，孩子们可以提高他们的社交技能、运动技能、认知能力等。陕玉娟[14]基于 AR 开发自闭症儿童面部表情识别能力干预系统，将现实人物面部表情转换为儿童感兴趣的卡通皮卡丘形象，设计学习模式和测评模式，利用动画和游戏帮助自闭症儿童识别基本面部表情。罗琳丹[15]提出以 APP 为载体融合 AR 手段可以辅助自闭症儿童治疗康复，通过移动端摄像头扫描、图像上传功能实现 AR 互动和虚拟成像，与绘画疗法结

合，促进情绪表达；借助 AR 形式进行通感艺术疗愈，使用设备扫描真实物品转化为针对自闭症儿童所设计的特有形象，配置声音、动画效果以及视觉语言，产生多重感官刺激，刺激感知能力；利用 AR 手段塑造可行性虚拟情景教学，创建不同虚拟角色，与其互动，改善语言理解能力。

3.4 MR 构造虚拟—现实人机交互环境

MR(Mixed Reality)是混合现实技术，MR 结合了 AR 和 VR 的特点，不仅可以在真实的环境中添加虚拟元素，还让真实环境以及用户与虚拟环境进行互动。

利用 MR 技术，可以将虚拟环境融入到真实环境中，创造特定情境，让自闭症儿童与虚拟环境进行互动，实现人机交互，迎合自闭症儿童的兴趣从而吸引他们的感知觉，提高训练效果。靳新培[16]设计出国内首个针对自闭症儿童打造的混合现实系统，它将虚拟海洋世界与真实的鱼缸结合，通过投影设备和液晶调光膜展现 3D 虚拟场景，将认知训练课程整合到混合现实鱼缸系统中，通过视觉化、具象化的效果促进自闭症儿童与训练内容、康复师之间的互动，训练其认知能力、动手能力和表达能力[17]。“迷雾之地”混合现实系统加入一系列视觉、听觉效果并在地板投影虚拟环境，自闭症儿童手持真实捕蝶网，系统对捕蝶网进行追踪，通过捕蝶网与投射在地板上的虚拟昆虫元素进行互动，既可以自己单独进行游戏，也可以邀请同伴参与，共同开展具有创造性的探索类游戏，这能让自闭症儿童表现得更加放松，为自闭症儿童提供一个舒适的社交环境，提高社交能力[17][18]。

3.5 数据分析与反馈

利用大数据和人工智能技术，对自闭症儿童在元宇宙中的行为表现进行实时监测和分析。通过对数据的挖掘、整理和分析，可以了解孩子们的表现和进步情况，评估患者的康复进展，并将数据分析的结果反馈给家长、教育工作者、医生等相关人员，获得干预支持和建议，从而不断调整和优化康复方案，制定个性化训练方案，提高康复效果。合理的数据收集、分析和反馈，可以为自闭症儿童提供更个性化、有效的支持和帮助，促进他们的发展和适应能力的提升。

4 结语

随着医疗技术不断发展，传统的康复训练模式已不能满足当下日益严峻的患病环境，而元宇宙医学为自闭症康复训练提供了全新的视角和可能性，利用 VR、AR、MR 技术辅助自闭症儿童康复治疗在后期实验中均被证明可以有效提高自闭症儿童的认知、语言和行动能力。基于远程通信技术、VR、AR、MR、大数据和人工智能技术构建自闭症儿童康复训练模型，创建高度仿真的、沉浸式的康复环境，结合个性化的康复课程和交互式训练方式，以及进行后续的数据分析和反馈，我们可以更好地满足自闭症儿童的需求，提高他们的康复效果和生活质量。

PO-027

基于多任务学习融合外部知识的临床实体识别

程铭, 谭健
郑州大学第一附属医院

1 引言

随着电子病历系统的快速发展，人们对将文本挖掘和信息提取应用于电子病历越来越感兴趣。这些技术可以为医学研究和应用带来巨大的好处。在医学文本挖掘任务中，医学实体识别是一项基础任务，它定位医学命名实体的提及，并在非结构化医疗/保健记录中对其进行分类（如症状、测试、药物、手术和疾病等）[1-4]。然而，医学领域的临床实体识别是一项具有挑战性的任务：（1）临床记录中经常出现各种非标准表达，同一实体的多种变体[5-9]，以及（2）临床记录的句子结构往往不完整，上下文较少，语法错误较多[10-12]。

本文提出了一个基于字典的多任务神经网络模型。除了主实体识别任务外，我们还使用了一个辅助但相关的辅助任务——医学实体分割。医学实体分割是一个二进制分类任务，其目标是预测字符是否是实体的一部分。我们使用这两项任务来联合训练网络。此外，为了覆盖更多的临床实体，我们设计了五个 n-gram 特征模板来构建字典特征。最后，在两个医学数据集上对所提出的方法进行了实验评估。

2 整体框架

为了缓解对大量标注数据依赖的问题，提出了一种融合外部字典的深度神经网络多任务学习模型（DicMT），整体框架如图 1 所示，其中句子采用字符分割的方式进行切分，医学实体采用“BIO”标签进行标记。从图中可以看出该模型主要包括以下模块：特征表示层，多任务神经网络，输出层。

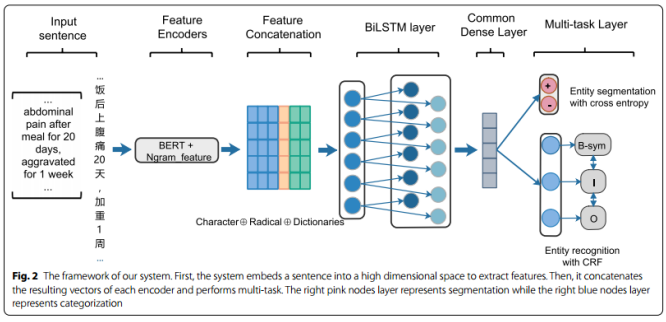


Fig. 2 The framework of our system. First, the system embeds a sentence into a high dimensional space to extract features. Then, it concatenates the resulting vectors of each encoder and performs multi-task. The right pink nodes layer represents segmentation while the right blue nodes layer represents categorization

图 1 技术路线

2.2 特征表示层

特征表示层，包括中文字符特征表示和外部字典特征表示。中文字符特征表示，采用双向 Transformer 编码器（Bidirectional Encoder Representations from Transformer，BERT）作为语言模型在大量的无标注中文电子病历医疗文本上进行预训练。获取 7.8G 的原始电子病历数据，删

除隐私信息后只保留原始医疗记录，经过处理后剩余可用数据 1.2G。这些医疗语料来自不同的临床科室，如胃肠外科、心血管内科、妇科、骨科等。对于预训练 BERT 模型，基于已有的 BERT 检查点，在特定的无标注电子病历数据集上运行额外的预训练步骤，从而优化 BERT 模型。另外，汉字通常由更小的子结构组成，称为部首。部首具有汉字的潜在特征，并存在额外的语义信息。另外，汉字的书写形式通常具有共同的子结构，其中一些子结构具有相同的语义信息。例如，字符“肝”、“腺”、“腹”共享相同的部首“月”，利用部首来增强中文字符的表示已经被广泛使用。受此启发，在模型中增加部首特征以增强中文字符的表示形式。

表 1 n-gram 特征模板

类型	模板
2-gram	$x_{i-1}x_i, x_ix_{i+1}$
3-gram	$x_{i-2}x_{i-1}x_i, x_i\ x_{i+1}\ x_{i+2}$
4-gram	$x_{i-3}x_{i-2}x_{i-1}x_i, x_i\ x_{i+1}\ x_{i+2}\ x_{i+3}$
5-gram	$x_{i-4}x_{i-3}x_{i-2}x_{i-1}x_i, x_i\ x_{i+1}\ x_{i+2}\ x_{i+3}\ x_{i+4}$
6-gram	$x_{i-5}\ x_{i-4}\ x_{i-3}x_{i-2}x_{i-1}x_i, x_i\ x_{i+1}\ x_{i+2}\ x_{i+3}\ x_{i+4}\ x_{i+5}$

外部字典特征表示，先前的研究已证明在实体识别任务中将字典信息融入到字符特征中是有效的。假定一个句子 X 和外部字典 D，基于 x_i 的上下文，采用预定义的 n-gram 特征模板构建文本片段，n-gram 特征模板如表 2-1 所示。

N-gram 特征模板共生成 10 个文本片段，设计 5 个二进制向量表示字典中不同的实体类型。CCKS2017 数据集中，疾病实体表示为（0,0,1），解剖部位（0,1,0），症状（0,1,1），检查（1,0,0），治疗（1,0,1）；CCKS2018 和 FCCd 数据集中，药物实体（0,0,1），解剖部位（0,1,0），独立症状（0,1,1），描述症状（1,0,0），手术（1,0,1）；（0,0,0）表示这个文本片段不是一个临床实体。图 2 展示了 n-gram 特征模板的例子。

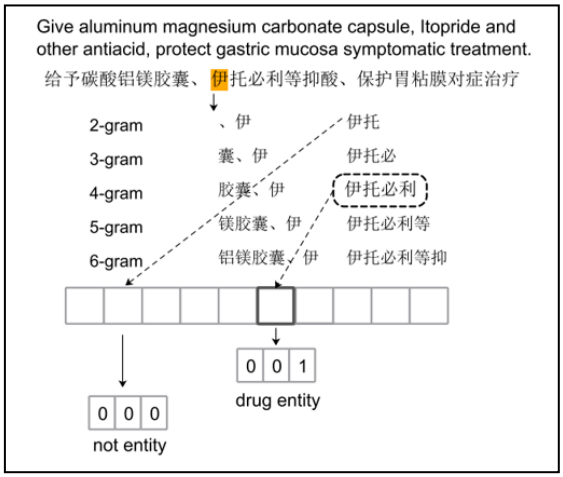


图 2 n-gram 特征模板例子

2.2 多任务神经网络

临床实体分割与识别是两个相关的任务，它们的输出可以相互受益。具体说，实体分割的输出可以减少实体识别任务的搜索空间，反之亦然。因此，提出一种多任务学习框架来联合训练医学实体分割和识别任务模型，模型共享训练参数。

每个字符特征表示由三部分组成：预训练的 BERT 字符嵌入，部首特征，字典特征。例如，给定一个中文临床句子 X= x_0 x_1… x_n，部其中 $x_t(1 \le t \le n)$是第 t 个字符，xt 表示为$[p_t \oplus r_t \oplus d_t]$，其中 p_t、r_t、dt 分别是预训练字符嵌入、字符部首嵌入以及字典特征嵌入，$\oplus$表示拼接操作。将字符特征输入到 BiLSTM 网络中，然后将 BiLSTM 的输出同时作为实体分割和识别任务的输入。对于实体分割的二分类任务，使用 sigmoid 激活函数和交叉熵损失；而对于主要的实体识别任务，采用 CRF 来预测字符可能的标签。此外，使用从共享网络层学习的权重来捕获两个任务的泛化特征，同时两个任务的总损失在训练过程中被反向传播从而优化网络参数。

训练目标，对于实体分割的二分类任务，采用交叉熵损失。实体标记为“1”和非实体标记为“0”，交叉熵损失表示为：

$$H(p,q) = -\sum_{c \in C} p(c)\log(q(c))$$

实体分割的损失函数表示为：

$$loss_1 = -\sum (p(1)\log(q(1)) + p(0)\log(q(0)))$$

对于实体识别任务，采用 CRF 预测字符的标签。CRF 层在最大条件似然估计下进行训练。对于训练集（z_i，y_i），最大似然估计损失函数如下：

$$loss_2 = -\sum_i logpp(y|z;W,b)$$

直接将两个单任务的损失结合起来作为多任务的损失。引入正则化参数α和β，多任务框架的损失函数定义为：

$$loss = \alpha\ loss_1 + \beta loss_2$$

2.3 输出层

CRF 的输出作为最终的字符标签，使用 Viterbi 解码以获得最佳的输出：

$$y^* = argmin_{y \in Y} p(y|z; W, b)$$

最终，得到结构化的输出序列 $Y=\{y_1,\ldots,y_n\}$。

3 实验结果

实验评估采用两个数据集：CCKS2017、FCCd，数据集统计如下表所示。

表 2 数据集统计

CCKS2017	**症状**	**疾病**	**检查**	**治疗**	**部位**
Total(1596)	10142	1275	12689	1513	13740
FCCd	**部位**	**手术**	**药品**	**独立症状**	**描述症状**
Total(736)	9686	1164	1105	4117	3061

① 不同字符嵌入对模型性能的影响。“random”表示随机生成的字符特征表示，“bert”表示在通用领域训练得到的预训练模型，“ft-bert”表示在特定临床文本上微调训练得到的预训练模型，“ft-bert-radical”表示微调预训练模型联合部首特征的字符表示。实验结果表明，中文字符部首特征联合特定领域的预训练模型能够得到较好识别性能，优于其他方法，如表 3 所示。

表 3 不同字符嵌入对模型的影响

CCKS2017	**precision**	**recall**	**F-measure**
random	89.56	89.29	89.42
bert	90.73	90.51	90.62
ft-bert	91.27	91.21	91.24
ft-bert-radical	91.69	91.34	**91.51**
FCCd			
random	84.23	83.32	83.77
bert	86.11	85.52	85.81
ft-bert	86.21	85.83	86.02
ft-bert-radical	86.95	86.56	**86.75**

② 字典信息对模型性能的影响。添加字典信息后，模型性能得到显著的提高，字典信息可以有效的识别稀少实体，如表 4 所示。另外，调查字典大小对模型性能的影响，通过随机选择 70%、80%、90%、100%的字典信息，构建 4 个子字典。实验结果如表 5 所示，随着字典大小的增加，模型性能也随着提高。实验结果证明，字典中包含的实体信息越多，则模型的性能越优越。

表 4 字典信息对模型的影响

CCKS2017	precision	recall	F-measure
ft-bert-radical	91.69	91.34	91.51
ft-bert-radical +dictionary	91.91	91.78	**91.84**
FCCd			
ft-bert-radical	86.95	86.56	86.75
ft-bert-radical +dictionary	87.32	86.79	**87.05**

表 5 不同字典大小对模型的影响

CCKS2017	precision	recall	F-measure
70%	91.79	91.61	91.70
80%	91.83	91.69	91.76
90%	91.87	91.73	91.80
100%	91.91	91.78	**91.84**
FCCd			
70%	87.19	86.65	86.92
80%	87.23	86.71	86.97
90%	87.28	86.75	87.01
100%	87.32	86.79	**87.05**

4 思考与展望

本文提出了一种新的用于中文临床 NER 任务的多任务神经网络模型。它将字典和中文部首信息融合到多任务神经网络中。由于字典包含罕见的实体，我们提出的方法可以更好地处理它们。在两个数据集上进行实验，我们发现，将字典信息合并到模型中可以提高性能。在未来的工作中，我们打算进一步研究如何应用基于字典的多任务学习方法来识别临床文本中的嵌套实体，以及所提出的模型在其他相关 NLP 任务中的应用。

PO-028

面向智慧医院临床数据中心的多源异构数据融合平台研究与实践

刘阳[1] 文霞[1] 康亚西[2] 陈伟聪[3] 苏子仁[4]
[1]中南大学湘雅医学院附属海口医院（海口市人民医院）信息管理处 急诊医学部
[2]海南省人民医院信息工程部
[3]海南医学院第一附属医院 信息中心
[4]海南医学院第二附属医院 信息科

0 引言

我国医疗信息化领域快速发展，医疗人工智能辅助、智慧医疗无不需要医疗大数据的支撑，因此对于医疗大数据应用，首先需要做的就是临床数据的融合，但是在基于日志变化数据捕获技术的多源异构临床数据融合共享领域，其关键技术一直被 IBM、ORACLE 等少数欧美信息技术企业所垄断，国内医院所使用数据抓取工具始终无法摆脱被国外信息技术企业的技术控制，在业务系统安全性和数据加密保护方面也存在极大地安全隐患[1]。聚焦医院的业务烟囱和数据孤岛现状，突破长期以来困扰医院信息系统性能的瓶颈，研究具有自主知识产权的数据捕获引擎，构建数据融合共享平台，成为促进智慧医院高质量发展的当务之急。

1 面临现状和主要问题

1.1 数据分散在不同的业务系统

近十年来，由于历史原因，医疗机构在不同时期根据当时的业务需求先后建设了不同类型的应用系统，截止目前，大型医院中同时运行着几十甚至上百套医疗业务信息系统已经司空见惯，主要包括：HIS、EMR、CIS、PASS 以及医技科室信息系统、HRP、OA 等[2]。这些系统采用不同数据库系统和不同设备在医务活动中采集并产生数据，不同的数据源所采用的操作系统、数据库管理系统不同，数据的存储模式和逻辑结构不同，数据的产生时间、数据用途、代码协议等也不同。不同类型的数据在产生的过程中没有统一标准，因此造成了数据“异构”的特征。

1.2 业务核心数据库不堪重负

临床信息系统围绕着患者所产生数据，不仅数量大而且增长速度快，每个临床应用系统都是数据的产生者，同时也是业务核心数据库的资源消耗者[3]。因此在传统的网状架构中，不断增加的临床业务系统需求，就导致了业务核心数据库不堪重负，医院系统核心主机面临崩溃的边缘。

1.3 数据访问跨库操作的资源消耗高

在传统的网状系统架中，不同类型数据库间的查询采用数据库网关的方式进行跨库数据访问，比如 ORACLE 到 SQLServer 数据库的查询，需要通过 ORACLE Gateway 实现，数据跨库访问的资源消耗非常高，严重影响临床应用，比如：医保飞行检查、病案报表查询、多学科会诊等需要多个异构数据库查询数据时，查询时间都需要几个小时，导致全院系统响应速度迟缓。

2 方法

2.1 数据融合平台总体架构

总体架构采用数据融合的设计思路，对多个不同业务用途的数据库按需采集数据，将数据交换到融合平台，按数据应用需求对数据进行分组、合并、加工、清洗等，将处理后的数据融合到一个或多个数据资源库，用于数据统计分析、复杂查询，消除跨数据库访问带来的性能问题。多源异构数据融合平台详见图 1，包含以下 5 个部分的内容，数据采集组件（数据捕获与发布）、数据加工组件（ETL 数据转换）、数据比对与修复组件、数据订阅与应用组件、数据融合管理平台。

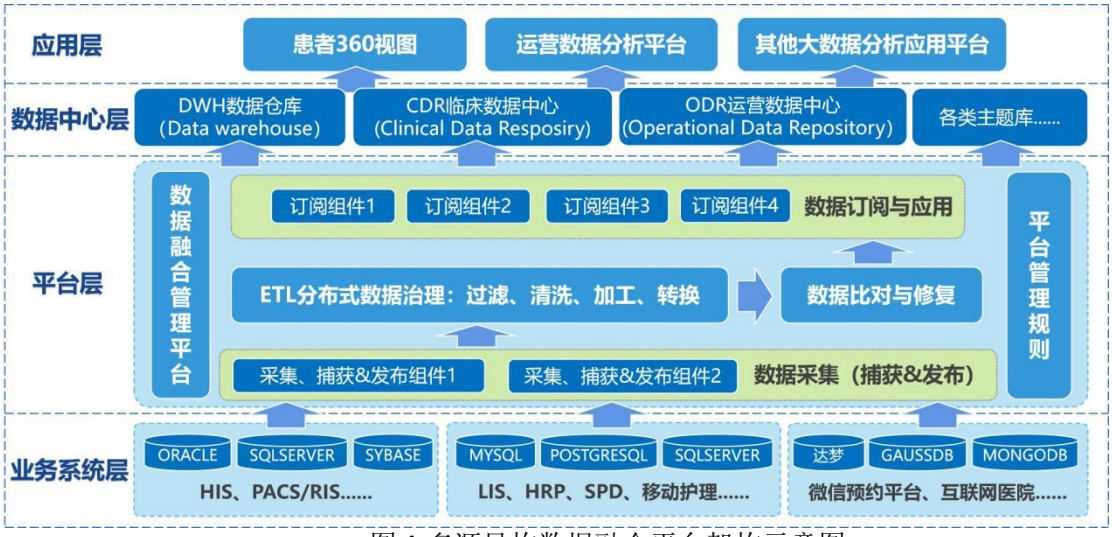


图 1 多源异构数据融合平台架构示意图

2.2 数据融合平台的开发方法和技术路线

各个系统的数据都能按照统一的数据管理机制进行信息集成，确保数据的准确性、一致性和完整性，实现数出同源、数用同标，满足跨部门、跨系统的数据应用需求[4]。具体的开发方法和技术路线，首先需要适配医院业务系统常用的 ORACLE、SQL Server、MySQL 三种数据库，以上三种数据库的业务系统分别对应 HIS、LIS、RIS，完成技术层面开发和适配后，已陆续适配主流的商用数据库和开源数据库产品，包括国产数据库，主要包括：DB2、Postgresql、Sybase、达梦、人大金仓、TiDB、OceanBase 等。

2.2.1 数据采集组件（数据捕获与发布）

按需支持 ORACLE、SQLSERVER、MYSQL 数据库全量、增量发布；

按需发布到 KAFKA 集群平台，统一管理，根据需实现其业务逻辑，上层操作仅为用户名，表名，对列名，以及数据库组件等信息

以 ORACLE 数据库为例，ORACLE 支持并包括从 9i 到 20c 之间的全部版本，源端数据捕获支持 2 种方式，采用第三方 Agent 软件或无代理模式，无代理模式根据版本不同可选择通过 LOGMNR 或 XSTREAM 实现数据的捕获：

使用 LOGMNR 模式实现增量数据捕获,支持 ORACLE 9i-ORACLE 11.2.0.4 所有版本，增量数据从数据库准实时获取出来，对生产压力较小；

表级开启附加日志

alter table 用户名.表名 add supplemental log data(all) columns;

库级开启附加日志

alter database add supplemental log data(all) columns;

alter database add supplemental log data;

2.2.2 数据加工组件（ETL 数据转换）

数据交换过程中的数据加工，实现如下加工能力：

多种数据转换方法：包括字符串转换、字段赋值、时间类格式转换、数学运算、身份证格式转换等。

交换过程中的数据质量检查，根据数据逻辑判断规则，将干净的数据装载到目标中，将判断有问题的数据路由到数据库表或者数据文件中[5]。

多种逻辑处理，包括：格式匹配检查（如日期格式、数据格式、身份证格式、自定义格式等）、字符串逻辑检查（包含、以结束、以开始、在列表中、等于等）、内容为空检查、重复记录检查、范围内检查（如在列表内、字典或代码表内、包含、等于）、表表外键关联检查、逻辑检查（=、>、>=、<、<=、为空、非空、为真、为假等）、复合逻辑检查（以上逻辑的 and、or、&&、||组合）、自定义逻辑检查等。

多种路由策略，路由条件可以是逻辑判断，也可以是等于、不等于、小于、小于等于、大于、

大于等于、规则表达、是空、非空、包含、开始为、结束为、为真、为假等方法。

当生产数据变化后，事后触发功能，参考触发器功能，实现业务动作衍生。实时 ETL 数据转换，对生产零影响。支持触发方式：SQL 语句，存储过程，BAT/SHELL 处理。

2.2.3 数据比对与修复组件

数据比对组件，支持条数级别数据比对，实现生产库和目标库记录数比对；支持数据分流模型中，分流数据比对；在使用 ETL 数据转化会后，实现数据分流，或者数据转换，则对数据分流之后的数据进行统计、比对。支持全量、实时数据比对；支持对整体数据库表中的数据进行数据比对；支持模拟生产业务变化，对插入，更新，删除等动作进行数据比对，比对分为比对主键和比对内容两种模式[6]。

支持条件、行级别数据比对同库，如 ORACLE，使用 DBLINK 等方式进行数据比对实现数据内容全量比对。

数据修复组件，支持表级全量数据修复；支持表级行级数据修复功能；支持指定时间段之内业务重新模拟变化，重新应用数据。

2.2.4 数据订阅与应用组件

支持任意一份数据，1:N 输送到多个目标端;支持对集群中任意表进行组合，提高业务灵活性;支持数据应用到 ORACLE、SQL Server、MySQL 等关系型数据库中;支持数据输送到 REDIS、ROCKSDB、LEVELDB、HBASE 等 NOSQL 数据库中;支持按列复制;支持补充生产操作的时间戳;补充生产操作的类型。在实际应用中，医技检查实时的排队情况、危机值等信息，均可及时展现在规划的位置或推送给相关干系人[7]。

2.2.5 数据融合管理平台

管理平台完成对系统内部的所有数据库重新编排，编码，为集群内部建立数据库统一编号，库与库之间的同步，管理，全部依赖此编号，管理平台以 Web 界面展现，在单一界面中完成数据融合平台的全部管理任务[8]。

在同一个平台内，可视化实现数据节点的管理、交换服务建模、数据加工建模、共享服务建模、数据质量的管理、服务目录管理、资源目录管理、运行及监控管理等[9]。通过相应服务的可视化操作界面生成相应类型的服务模型，可以根据业务需要创建文件传输、ETL、数据交换和处理、数据质量、安全共享、流程等各类服务，并可以根据需要分类、编码形成服务目录，便于重用、编辑、安全授权等需要，提供服务创建、编辑、删除、部署、查看等功能。

3 结果

海口市人民医院信息化建设经过 10 余年的发展，业务系统繁多，数据库类型多样，数据接口杂乱无章，面对这样的困境，我们研发了多源异构数据融合平台，将医院应用范围最大的核心业务系统 HIS（含 EMR）、LIS、RIS 三个系统进行了数据集成，把以往分散在以上系统中的医疗数据进行了融合[10]。其中 HIS 系统实现了医院门诊医生工作站、住院医院工作站、医院经济管理、药品管理、医保控费等多项功能，HIS 系统数据库采用 ORACLE Database，数据库部署在 2 台 IBM S824 小型机服务器上，通过 ORACLE RAC 高可用集群保障业务的连续性和可靠性；其他临床应用系统和医技应用系统通过点对点应用接口方式与 HIS 数据库建立连接；LIS 系统采用独立的服务器部署在 Windows Server 操作系统平台中，数据库采用 MSSQL Server，用于管理和存储检验报告相关数据；PACS 及 RIS 系统分别部署在独立的 Windows Server 和 Linux 操作系统中，采用 MSSQL Server 和 MySQL 数据库，用于管理和存储检查类相关医学影像和报告。研究利用数据复制与交换技术，该项技术能够基于统一的规则和元数据、统一架构、统一工具可视化配置完成源端数据采集、数据交换、数据加工、数据转换、数据按需共享等不同的数据服务，实现数据交换和数据管理需要（国内尚没有此类解决方案，目前只能依赖美国 IBM、ORACLE 的 CDC、OGG 等产品，已有国外产品存在很大的局限性，不支持异构数据库；并且国外产品不支持国内的信创要求，不能够适配国产化数据库）。将各个应用系统、不同类型操作系统、不同类型数据库中的业务数据融合到一个集中平台上，达到多源异构临床数据融合从而形成临床数据中心 CDR 的目的，向智慧医院上层服务和大数据应用提供支撑[11]。目前，在已有研究成果——数据融合平台 V1.2 的基础上，在海口市人民医院已经完成的部署和实践应用。在海南省人民医院、海南医学院第一附属医院、海南医学院第二附属医院的对接实施、安全测试和风险评估工作正在进行中。

图 2 海口市人民医院数据融合管理平台主界面截图

如图 2，海口市人民医院数据融合管理平台主界面截图，可以对通过平台进行治理的数据进行管理和监测，从首页查看数据发布统计、数据处理统计（入）、数据处理统计（出）、订阅统计，通过点击相关数据融合的工具：容器、发布、策略管理、数据处理、订阅、监控、系统、资源管理、计划任务、客户端管理、权限管理，查看数据融合平台与各业务系统数据采集、发布的配置情况和事实数据治理的监测状况。

4 讨论与思考

4.1 建设成效的讨论

采用数据融合技术，能够基于统一的规则和元数据、统一架构、统一工具可视化配置完成源端数据采集、数据交换、数据加工、数据转换、数据按需共享等不同的数据服务，实现数据交换和数据管理需要。将各个应用系统、不同类型操作系统、不同类型数据库中的业务数据融合到一个数据工厂按需进行数据的加工处理，形成各种用途的数据资源库，达到多源异构临床数据融合从而形成临床数据中心。

4.2 遗留问题的思考

4.2.1 多平台协调工作量大，业务实现复杂

医疗信息化蓬勃发展的同时也催生了众多的系统开发商，医院的业务系统往往承建厂商不同、技术标准各异，涉及多个系统的联动、协同与交互，协调工作量大，业务实现复杂。解决方法：医疗机构组建了标准委员会，首先明确技术层面的接口规则；其次，成立用户和开发商共同组成的协调组或者专班，分工协调不同类型的工作任务，并且明确项目中各角色的工作面分工，清晰边界、明确责任。

4.2.2 数据共享带来的个人隐私保护等安全问题不容忽视

个人隐私保护等安全问题不容忽视：互联网时代数据的暴露面越来越大，各业务系统之间患者数据的共享在提供了数据访问便利的同时，个人隐私的泄露风险同步增加。解决方法：融合平台与 CA 相结合，为患者病历数据的共享提供授权服务，必须使用患者授权的二维码进行解锁，方可查看患者病历数据；严格落实患者病历授权访问管理制度，有效加强对患者隐私数据的授权访问保护，使得患者的病历、检验、检查报告等信息既能够被方便共享利用，又能够通过患者授权后的开放读取方式保障数据安全合法利用。

5 结论与展望

展望未来，我们希望能够在已有研究成果——数据融合平台 V1.2 的基础上，解决 V1.2 版本存

在的数据适配性、完整性、一致性方面的问题，继续改造数据融合平台升级版本，将不同应用系统、不同类型操作系统、不同类型数据库中的业务数据融合到一个集中平台上，达到多源异构临床数据融合从而形成临床数据中心 CDR 的目的，向智慧医院上层服务和大数据应用提供支撑。

PO-029

临床专病数据库的建设现状与展望

王耀国[1] 龙思哲[2] 刘迷迷[3] 周毅[3]

1 新疆医科大学
2 中山大学附属第一医院信息数据中心
3 中山大学中山医学院

1 引言

专病数据库的发展与国家政策和临床医学需求密切相关，我国卫生健康信息化领域顶层设计体系基本形成[1]。同时疾病研究更加依赖临床数据。但是医疗信息系统前后建设差距大导致大量临床数据无法直接支持临床研究。提高临床科研数据的可用性是专病数据库产生的直接原因，同时采用专病数据库提高数据的互操作性是实现数据互联互通的一种方式。目前国内临床各分支学科希望通过建立专病数据库对目标患者群体的临床数据进行整合，并发掘有价值的信息，以提高专科疾病治疗效果[2-4]。

2 专病数据库发展现状

2.1 国外专病数据库发展现状

本文以“disease and database”为关键词，在 PubMed 上搜集了 2020～2023 年期间的外文文献共 6633 篇，大多数针对特定疾病的国际临床研究通常建立大型队列研究，收集前瞻性数据，且数量也相对较少。欧美等国家已经建立了国家级专病数据库[5-8]，美国大约在 2016 年左右完成了医疗信息大规模的互联互通工作，为构建国家级专病数据库提供了平台基础。日本医疗主体以私人为主，实行严分级诊疗制度，全民医保，这也在一定程度上导致了日本的临床医疗数据与保险公司数据有密切的关联。Delaney 等人从一些医疗器械、生物制药公司抽取相关医疗数据，构建大规模数据库[9-11]；Tsuneishi、Kyoung DS 等人，利用日本国家健康保险索赔和特定健康检查数据库、韩国国民健康保险服务和健康保险审查与评估数据库进行数据分析[12-15]。

2.2 国内专病数据库发展现状

本文以“专病数据库”为关键词，在知网等网站检索收集相关文献共 52 篇，总结发现我国专病数据库的发展程度和方向深受我国医疗体系和医疗信息化程度的共同影响，且仍处于发展的早期阶段。我国三甲医院在医疗系统中占主导地位，然而大多数三甲医院正处在以数据为导向，建立院内及区域互联互通医疗平台的探索阶段[16]，并且专病数据库建设主体也主要在三甲医院。目前我国国家级重大疾病专病平台数量较少，缺少业内公认的、标准的设计和开发方法[17]。该领域工作发展状况不均衡，上海、广州、北京等地的医院专病数据库的建设起步时间早[18-24]，上海交通大学转化医学研究院院长陈赛娟院士在 2022 年全国两会期间，提交《关于发展重大疾病专病数据库》的提案；中山大学孙逸仙纪念医院院长宋尔卫院士牵头“中国乳腺癌标准数据库”建设，取得了实质性进展；南方医院侯凡凡院士围绕肾脏疾病防治发展需求，加强肾脏病专病大数据标准化建设等。但西北等地的医院相较于内地，专病数据库的发展程度和规模较为落后。

3 专病数据库建设方案

目前，国内专病数据库的主流构建模式有两种：第一种依托医院业务信息系统构建科室专科或专病数据库；第二种基于科研数据平台构建。专病数据库通用构建流程大致分为三部分：数据整合、数据处理和数据安全。专病数据库中多源异构数据来源包括医院信息系统（Hospital Information System，HIS）、电子病历系统（Electronic Medical Record，EMR）、检验信息系统（Laboratory Information System，LIS）、医学影像系统等。数据整合一般采用数据仓库技术（extract-transform-load，ETL）完成数据的抽取、转换和装载。数据处理一般采用自然语言处理技术对数据进行二次加工。数据安全则贯穿始终，包括确保数据不泄露，做数据脱敏处理。专病数据库通用构建流程见图 1 。

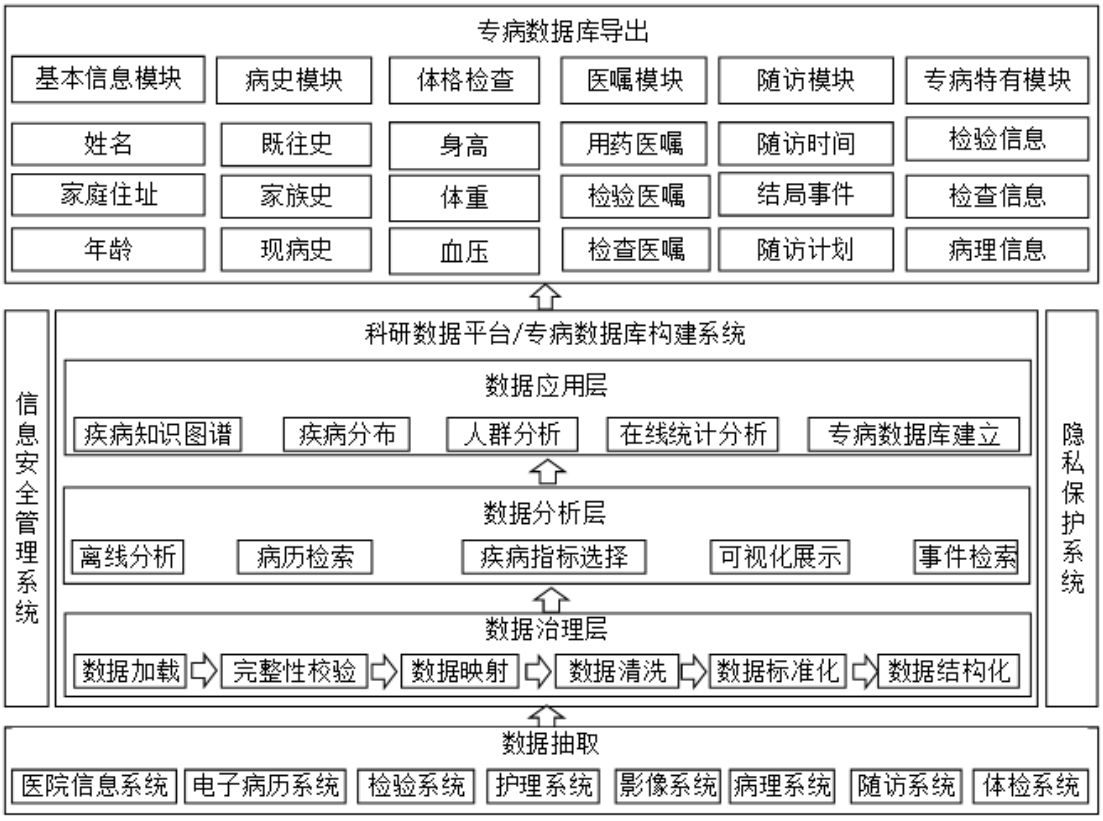


图 1 专病数据库通用体系架构

专病数据库中多源异构数据的处理方案见图 2 。

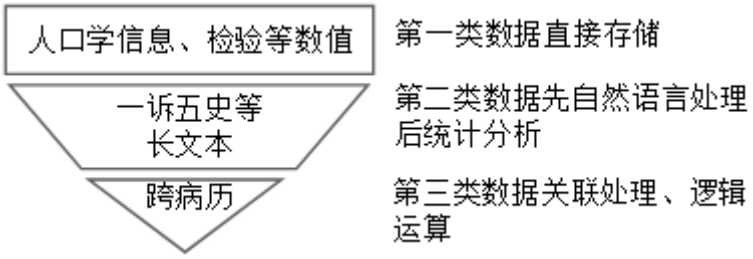


图 2 多源异构数据通用处理方案

数据处理方案分三类：第一类数据直接存储；第二类数据需辑处理，例如就诊年龄，需要根据就诊时间与出生日期进行算术运算；第三类数据一般先用自然语言技术再进行算数运算，例如病程，根据疾病史中的一段描述从中提取疾病的开始时间，再结合出院时间确定病程[26]。专病数据

库数据模块分为两类：第一类是通用信息模块，包括患者人口学信息模块、病史模块等；第二类是专病数据信息模块，包括检验模块、检查模块、等[27, 28]。

3.1 基于业务系统

该方案根据临床科室确定的目标数据集，直接从医院内各种业务系统中抽取数据[25]，主要用于构建小数据量和单病种专病数据库。搜集的临床数据主要是数值和文本，对检查影像等临床数据的处理一般仅搜集其文本描述信息如检查时间、检查所见、检查结论等，不整合影像结果。主要为临床科研提供数据支持。

3.2 基于科研数据平台

基于科研数据平台构建专病数据库是目前医院建设专病数据库的主流方案。首先，持续抽取全院数据建立面向全院数据的科研数据平台。然后，对医院所有接入系统的数据整合、管理、可视化展示等。最后，在科研数据平台的基础上抽取某种疾病的全部临床数据。

4 讨论与展望

医院信息系统数据库与专病数据库最大的区别在于医院信息系统数据库面向临床业务，而专病数据库面向临床科研，重点在于多源异构数据的整合。目前我国医疗信息化领域的痛点在于海量临床数据闲置，上述痛点的主要原因在于医院内部各种信息系统存在'孤岛效应'。建设专病数据库是为了解决多源异构数据的交互问题，目前国内构建专病数据库的方案已商业化落地，主要聚焦于临床数据的可用性。本质上仍是一套独立信息系统，处理后的数据与医院信息系统的数据交互效率较低，有的甚至不交互，没有实现真正意义上解决"数据孤岛"问题，实现数据层面的互联互通。本文提出一种专病数据库构建展望，能否将专病数据库与医院的互联互通建设结合在提高临床数据可用性的同时实现数据层面的互联互通。

PO-030
基于区块链的医疗电子病历大数据共享方案

刘曙霞[1] 杨斯博[2] 黄文波[3] 胥美美[1]

([1]中国医学科学院医学信息研究所 北京 100020 [2]天津大学管理与经济学部 天津 300072 [3]南开大学金融学院 天津 300072）

1 引言

目前，区块链已经应用于医疗信息领域，并且其核心应用技术优势正日益凸显。国内外有很多区块链医疗应用的成功案例，如 IBM Waston 与 FDA 签署的区块链在共享病人数据的项目协议，通过使用区块链技术设计和创建网络，提高医疗保健行业的透明度和互操作性；阿里健康与常州市合作"医联体+区块链"试点项目，旨在利用区块链技术保证部分医疗机构之间安全、可控的数据互联互通。

近年来，国内外也有不少区块链在医疗信息存储方面应用的研究。Ekblaw 等[1]提出去中心化电子病历管理系统 MedRec。Xiao Yue 等[2]提出使用区块链技术设计的数据共享架构--Healthcare

Data Gateways。以上研究缺少具体的实施机制和操作流程。Shrier 等[3]提出基于具体机制的医疗数据保护方案，即基于麻省理工学院研发的 OPAL/Enigma 加密平台，使用区块链技术建立一个安全的医疗环境以存储分析医疗数据，但操作起来效率极低。为了在保护数据安全的同时提升效率，徐建等[4]提出基于智能合约的区块链网络医疗记录安全储存访问方案，能够帮助快速识别不利于保护数据完整性的行为，但其仅仅针对云端数据，未能解决数据源头的安全性问题。而 BitHealth[5]基于点对点的传播文件方式，实现区块链对医疗健康数据的存储保护，解决了安全性问题，但是在资源利用方面仍存在一些缺陷。

2 医疗大数据电子病历特征分析

2.1 传统电子病历的发展

根据我国卫生部《电子病历基本架构与数据标准》，电子病历是由医疗机构以电子化方式创建、保存和使用的，重点针对门诊、住院患者（或保健对象）临床诊疗和指导干预信息的数据集成系统（图 1），是居民个人在医疗机构历次就诊过程中产生和被记录的完整、详细的临床信息资源[6]。

与纸质病历相比，电子病历可以有效提高工作效率，提升医疗质量。然而，传统的电子病历存在许多问题，比如无法对数据进行统一管理，患者在不同医院的医疗就诊数据无法共享；安全性低，数据很容易泄露或者被篡改等。

图 1 传统医疗电子病历信息

2.2 医疗行业大数据化趋势

随着医疗行业信息化的建设，大部分的病历数据和临床数据都已经信息化，数据量的大规模增加给数据的存储和管理带来很大挑战。如图 2 所示，医疗行业大数据采集有多种渠道[7]，智能可穿戴健康设备的发展也使收集公众行为健康数据变得更为简单。

医疗大数据不仅具有规模大、类型多、时效性、准确性、价值高等传统大数据的特征，还具有不完整性、冗余性、隐私性等特征[8]。大规模的数据如果能够通过集中和共享得到有效利用，那将势必给医疗行业带来前所未有的机遇和发展。

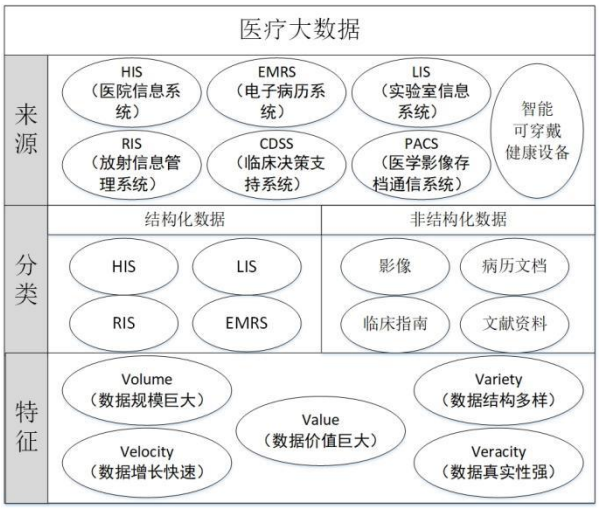


图 2 医疗大数据来源及其分类和特征

3 我国当前医疗电子病历数据共享情况调研分析

为了解医生和患者对医疗资源共享的需求，我们走访天津市 4 家三甲医院，包括天津市第一中心医院、天津市南开医院、天津医科大学总医院和天津市中医药大学附属第一医院，对医生和患者进行访谈。根据访谈得到的信息设计调查问卷，面向全国医生和患者开展调查，共回收信度效度较高的有效问卷 329 份，覆盖全国 24 个省市和地区（图 3）。329 份有效问卷中，医生回复的问卷占比为 15%。问卷结果显示，75%以上的医生（图 4）和 55%以上的患者（图 5）对现有医疗资源不够满意，经济落后的地区满意度更低，这在一定程度上显示出我国当前医疗资源匮乏、资源分配不均的问题十分严重。

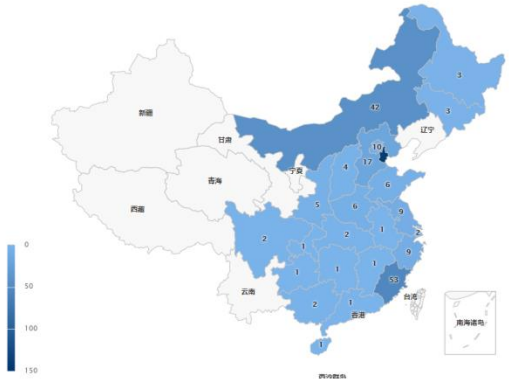


图 3 医疗资源满意度调查有效问卷的地区分布

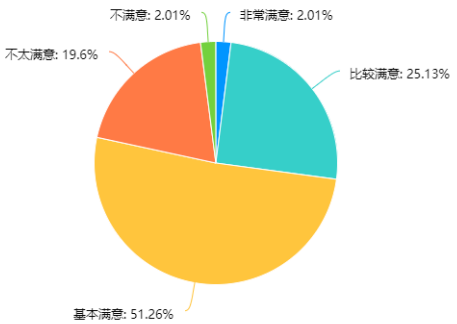


图 4 医生医疗资源满意度调查结果统计

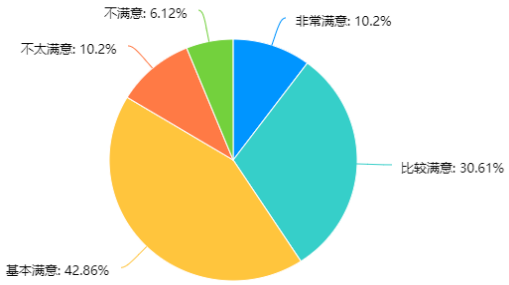


图 5 患者医疗资源满意度调查结果统计

调研中我们了解到电子病历的普及提高了医院工作效率，但是其安全性差、中心化、访问权限不明确等问题也很突出。现有的电子病历系统只能记录患者单次的就诊信息，对于需要转院的慢性病患者或者病情与病史相关的患者，不完整的就诊记录会给后续治疗带来许多不便。为解决上述问题，我们提出基于区块链的医疗电子病历大数据资源共享系统。

4 基于区块链技术的医疗电子病历大数据资源共享方案设计与应用

4.1 系统体系框架设计

区块链技术是底层技术，具有普适性。一般来说，区块链系统由数据层、网络层、共识层、激励层、合约层和应用层组成。数据层主要用来存储经过哈希加密的医疗大数据，网络层通过 P2P 协议将系统中的所有终端联系起来，共识层封装了不同节点之间的共识机制，激励层设计发行机制和分配机制，合约层封装了交易双方签署后的智能合约，应用层主要是系统使用者可直接进行操作的客户端（图 6）。

图 6 基于区块链的医疗大数据资源共享系统基础架构

4.2 系统交互机制设计

（1）医生与患者在就诊过程中系统交互机制设计

在系统中，每一个区块记载着患者自身的病史记录，其中患者在使用区块链系统的医院进行就诊的记录皆可保存在区块链中，利用密钥对其进行加密处理。患者可随时利用自己的公钥对自己以往的病例、诊断记录等相关信息进行查看。

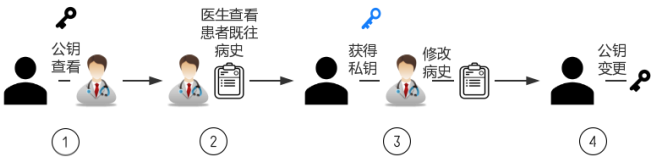


图 7 医生与患者就诊过程系统使用流程图

如图 7 所示，当患者与医生在就诊过程中需要查看或修改病历时，与系统交互机制如下：

1）医生登录医疗系统，系统验证医生的账号密码有效性并反馈登陆验证情况。

2）在需要查看既往病史时，系统跳转至患者权限索要界面，患者需向系统提供其公钥并等待系统进行身份验证，验证成功即可查看患者病历档案，在此期间仅有查看权限，无修改权限。

3）就诊期间，医生如需向节点上添加新的就诊记录，需进入相应记录创建界面，系统对患者私钥进行验证，验证成功后医生获得修改权限。

（2）医生与医生在数据共享与学术交流中系统交互机制设计

与患者同理，每位医生对应区块链中的一个区块，记录着医生从医以来的诊断记录，与患者信息相关的部分会加密处理，以保护患者个人隐私。

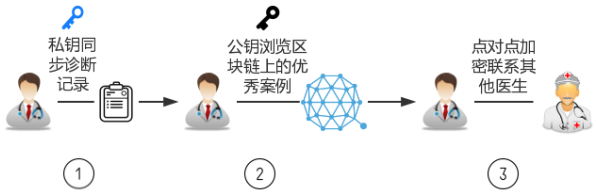


图 8 医生访问系统交互流程图

如图 8 所示，当医生在浏览过程中遇到较为感兴趣的病例想要进一步研究或遇到很难解决的医学问题想向其他医院的医生寻求帮助时，可以进行点对点的交流，向该医生发送会诊申请，经过双方同意后，双方（一方）可交换（分享）各自的医疗记录。

在点对点的信息交流过程中，两人的对话内容会被进行严格地加密处理，确保交流内容不被第三方窃取。

4.3 系统应用方案设计

在日常生活中，患者到使用本系统的医院进行就医时，其流程大体如下（图 9）：

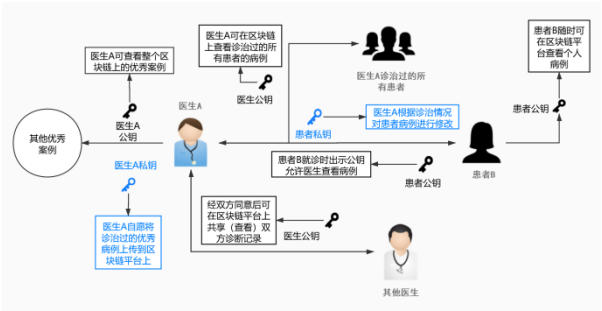


图 9 患者就医时系统使用情景模拟

1）患者 A 到使用本系统的医院就医时，向医生 B 出示公钥。

2）医生 B 通过公钥查看 A 的既往病史，了解其身体状况和体质等相关信息。

3）简单问诊后，患者 A 向医生 B 提供私钥，使得医生能够将就诊记录添加至患者所在区块。

4）就诊结束后，系统自动更新患者公钥，以保证隐私安全。

如果在此过程中，医生遇到较为棘手的疾病而缺少就诊经验时，可浏览区块链中其他加密处理后共享的案例，根据需求与相应的医生进行交流，从而弥补自身知识的局限，提高医生诊治的治愈率。

至此，在保留原有的医院医疗数据系统基础上，较合理地利用区块链技术对其进行完善，同时在数据共享方面，不仅实现了原有医疗系统无法实现的功能和远程信息交互，还使医疗大数据的安全性得到了进一步保障，为整个医疗系统的信息交流提供了安全可靠的平台。

5 结束语

本文将区块链技术与医疗信息系统相结合，提出基于区块链技术的医疗电子病历大数据共享方案。方案可以保护医生与患者双方的个人隐私，形成双方各自独立的诊断记录，并且具有开放共享、安全交互的特点，可以在一定程度上缓解我国东西部医疗资源差异过大，患者看病难、转院难、就医难的问题，有助于深入进行医疗领域改革、建设更和谐高效的医疗环境。

PO-031

基于上皮细胞-间充质转化特征基因的胃癌预后风险特征研究

王琪 张一凡 李晨龙 贺培凤

山西医科大学

1 前言

胃癌是全球第五大常见恶性肿瘤，也是癌症相关死亡的第四大原因，严重威胁着人类的健康。我国是胃癌高发国家之一，新增病例和死亡病例分别占全球总数的 42.6%和 45.0%[1]。大多数患者早期症状较轻，但一经确诊时就处于晚期，甚至有远处转移。胃癌的治疗和预后与诊断时肿瘤的病理分期、大小、部位等有关 [1]。近年来，由于手术联合免疫治疗、靶向治疗等新兴治疗方式的发展，显著改善了进展期胃癌的预后，使胃癌的发病率和死亡率持续下降。但是由于药物研发遇到的瓶颈、适用人群狭窄、易产生耐药性等因素，大多数患者的总生存期较低，预后差 [2]。现有的预后

生物标志物在预测胃癌患者预后方面的效果并不理想。因此，需要开发有效的预后风险评估模型，识别早期复发的高风险患者对于改善预后至关重要。

上皮细胞-间充质转化是指上皮细胞去分化为间充质细胞的生物学过程 [3]，它赋予了上皮细胞间充质特性，将极性上皮细胞转化为具有迁移性的间充质细胞 [4]。大量研究表明，上皮细胞-间充质转化在肿瘤中发挥着重要的作用，包括肿瘤发生、肿瘤进展、转移、迁移、血管内浸润和化疗耐药性等 [5,6]。上皮细胞-间充质转化相关基因与胃癌的发生发展密切相关，有研究证实了 LINC01235-TWIST2 反馈环通过抑制 THBS2 促进胃癌上皮-间充质转化 [7]。因此，深入了解上皮细胞-间充质转化相关基因在胃癌中的作用可以进一步加深我们对胃癌侵袭转移的认识。

胃癌作为一种高度异质性的恶性肿瘤，疾病进展和复发不仅仅与癌细胞的内在活性相关，还与肿瘤微环境中募集和激活的基质细胞和免疫细胞密不可分。肿瘤微环境的改变可能通过各种基因和信号通路的外源性激活促进胃癌的发展。因此，阐明胃癌中肿瘤微环境的特征，将有助于增强我们对胃癌病理机制的理解，并为进一步的个体化治疗提供基础。

本研究从 TCGA 数据库中下载了 407 例胃癌样本的转录组和临床数据。鉴定了上皮细胞-间充质转化相关的差异表达基因，基于生物信息学方法构建上皮细胞-间充质转化特征基因的胃癌预后风险评分模型。并在 GEO 数据集对该模型进行了验证。此外，我们研究了高低风险组的免疫微环境的状况，并探究了两组对化疗药物敏感性差异。为临床医生制定有效的治疗策略提供可借鉴的思路，从而提高患者的生活质量。

2 资料与方法

2.1 数据来源

从 TCGA 数据库获取胃癌的转录组表达谱（FPKM 值）和临床信息 [8]，包括 375 个胃癌样本和 32 个癌旁样本。TCGA 数据作为训练集。在 GEO 数据库中，以 “(Gastric cancer) AND "Homo sapiens"[porgn:__txid9606]”检索式进行筛选，将 GSE84437 作为验证集（N = 387）。训练集和验证集样本筛选标准：生存时间大于 1 个月且样本相关性大于 0.7。从 MsigDB v7.1 数据库中获取了 1384 个上皮细胞-间充质转化相关的基因。将胃癌基因表达谱（FPKM 值）转换成 log2 (TPM+ 1) 进行数据标准化。

2. 2 筛选上皮细胞-间充质转化相关的胃癌差异表达基因

利用 R 软件包 “Limma” 筛选 TCGA 数据中胃癌组织和癌旁组织之间的差异表达基因，筛选标准：|log2FC| ≥ 1, 矫正后的 P 值 < 0.05。将差异基因和 1384 个上皮细胞-间充质转化相关基因取交集，得到差异表达的上皮细胞-间充质转化相关基因。使用 Venn 进行可视化。使用 R 软件包 “pheatmap”和“ggplot2”绘制热图和火山图。

2.3 上皮细胞-间充质转化基因预后模型的构建及验证

使用单因素 COX 回归分析、LASSO 回归分析和多因素 COX 回归分析鉴定与胃癌患者总生存

期（OS）相关 (P < 0.05) 的独立预后基因并构建模型。LASSO 回归分析中，将随机数种子设置为“123”。危险比 (HR > 1) 的候选基因被认为是危险的预后基因，(HR < 1) 表示保护性预后基因。得到风险评分公式：风险得分 = ∑ (coefficient$_n$ × expression of gene$_n$)，计算每位患者的风险得分[9]。基于中位风险评分值，将患者分为高风险组和低风险组。进行主成分(PCA)分析和 t-SNE 分析，以评估预后模型区分高风险和低风险患者的性能。采用 R 软件包“survminer”绘制生存曲线，比较高风险和低风险组的生存差异；R 软件包“timeROC” 绘制 ROC 曲线，通过曲线下面积 AUC 评估模型的性能；通过单因素和多因素 COX 回归分析，确定风险模型和临床信息 (年龄、性别、分级、分期) 是否可以作为 OS 的独立预测因子。利用 R 软件包 “rms” 和 “forestplot” 构建列线图和森林图。

2.4 富集分析鉴定高低风险组间生物学过程特征

通过 R 软件包“limma” 筛选高低风险组的差异基因并利用“Clusterprofile”包进行基因本体论（gene notology，GO）[10] 和京都基因和基因组百科全书（kyoto encyclopedia of genes and genomes，KEGG）[11] 通路富集分析。

2.5 肿瘤微环境特征评估

采用 CIBERSORT 算法，以 LM22 标记为参考，推断 22 个肿瘤浸润免疫细胞的相对丰度。ESTIMATE 算法评估免疫得分（Immune Score）、基质得分（Stromal Score 和 ESTIMATE 评分 (ESTIMATE Score)。此外，通过 Pearson 相关分析计算风险评分与免疫浸润的相关性。

2.6 实时荧光定量聚合酶链式反应（RT-qPCR）验证风险基因

选取 2021 年 6 月至 2022 年 1 月山西医科大学第一附属医院胃肠外科收入院行手术治疗的 6 例胃癌患者的胃癌及癌旁样本。根据试剂说明书，使用 Trizol 试剂（Invitrogen, CA, USA）从胃癌和癌旁组织中提取总 RNA。使用 PrimeScript™ RT reagent Kit with gDNA Eraser（RR047A，Takara）从总 RNA 反转录 cDNA。使用 TB Green® Premix Ex Taq™ II (Tli RNaseH Plus)（RR820A，Takara）进行定量 PCR。GAPDH 被用作内参基因，PCR 引物购于日本 Takara 公司，具体引物序列见表 1，基因相对表达量使用 $2^{-\Delta\Delta Ct}$ 方法分析。

表 1 RT-qPCR 引物序列

Table 1 Primer sequences for RT-qPCR

引物名称	序列（5’-3’）
MATN3-F	GAATGACAGAACAGGGTCCCATC
MATN3-R	ATGTGCTGGCAACCATGAGAG
SERPINE1-F	CTGCATGACCTACCAGGACAGAAC
SERPINE1-R	CAGCCGGAAATGACACATTGA

3 结果

3.1 上皮细胞-间充质转化特征相关胃癌差异表达基因的鉴定

通过对 TCGA 数据集的 375 个胃癌样本和 32 个癌旁样本中进行差异表达分析，共获得 1704 个胃癌差异基因，包括 725 个上调基因和 979 个下调基因。差异基因的火山图和热图如图 1A-1B 所示。将胃癌差异基因和 1384 个上皮细胞-间充质转化特征基因取交集，共获得 249 个上皮细胞-间充质转化特征相关的胃癌差异表达基因（160 个上调基因和 89 个下调基因）（图 1C）。

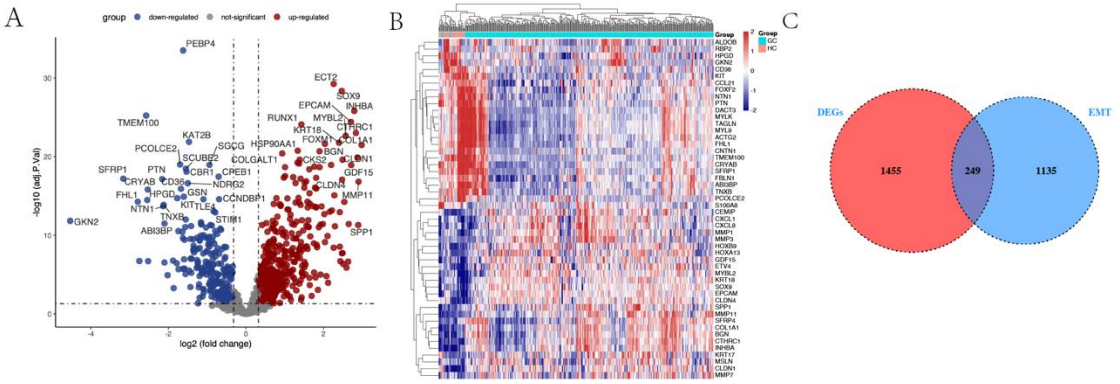


图 1. 胃癌差异表达的基因的鉴定。(A) 差异表达基因的火山图。(B) 差异表达基因的热图。(C) 胃癌组织的差异表达基因和上皮细胞-间充质转化基因相互作用的 Venn 图。

3.2 风险评分模型的建立

基于获得的 249 个差异表达的上皮细胞-间充质转化基因构建风险评分模型。通过单因素 COX 回归分析，得到了 25 个预后相关的基因，如图 2A 所示。从图中可以看出，这些基因都是高风险的预后基因（HR > 1）。进一步进行 LASSO 回归分析，得到了 3 个预后相关的基因，如图 2B-2C 所示。最后，采用多因素 COX 回归分析得到 2 个预后相关的基因并构建风险评分模型，两个预后基因分别为：MATN3（人母洗蛋白 3）和 SERPINE1（纤溶酶原激活物抑制因子 1）。风险评分公式为：风险得分（risk score）= 0.205* MATN3+ 0.159* SERPINE1。基于风险得分的中位值，将 TCGA 和 GEO 数据集中的患者分为高风险组和低风险组。PCA 和 t-SNE 分析表明，在 TCGA 和 GEO 队列中，高/低风险组被很好地分离(图 2D-2E 和图 2F-2G)。构建模型的两个基因 MATN3 和 SERPINE1 在胃癌患者中高表达，与患者的低生存有关（P <0.001）(图 2H-2I)。

3.3 风险评分模型的验证

在训练集（左）和验证集（右）中进行生存分析，可以看出低风险组患者和高风险组患者的生存存在差异（P < 0.001），且低风险患者的生存预后明显优于高风险患者（图 3A-3B）。风险曲线表明（3C-3D）表明，随着风险得分的增加，患者的死亡率升高。通过 ROC 曲线下面积对模型性能进行评估，训练集 3 年、5 年、8 年的 AUC 值分别为 0.665、0.791、0.729。验证集 3 年、5 年、8 年的 AUC 值分别为 0.647、0.30、0.637。ROC 分析结果表明，该模型具有较好的预测性能。单因素和多因素 Cox 回归分析表明（图 4A-4B），风险评分（risk score）可以独立于其他临床信息作为胃癌患者的预测因子 (P < 0.05)。校准曲线表明，预后模型预测概率与 3 年和 5 年 OS

的实际观测结果有很大的一致性，表明该模型具有良好的预测能力（图 4C）。列线图可用于多指标联合诊断或预测疾病的预后（图 4D）。

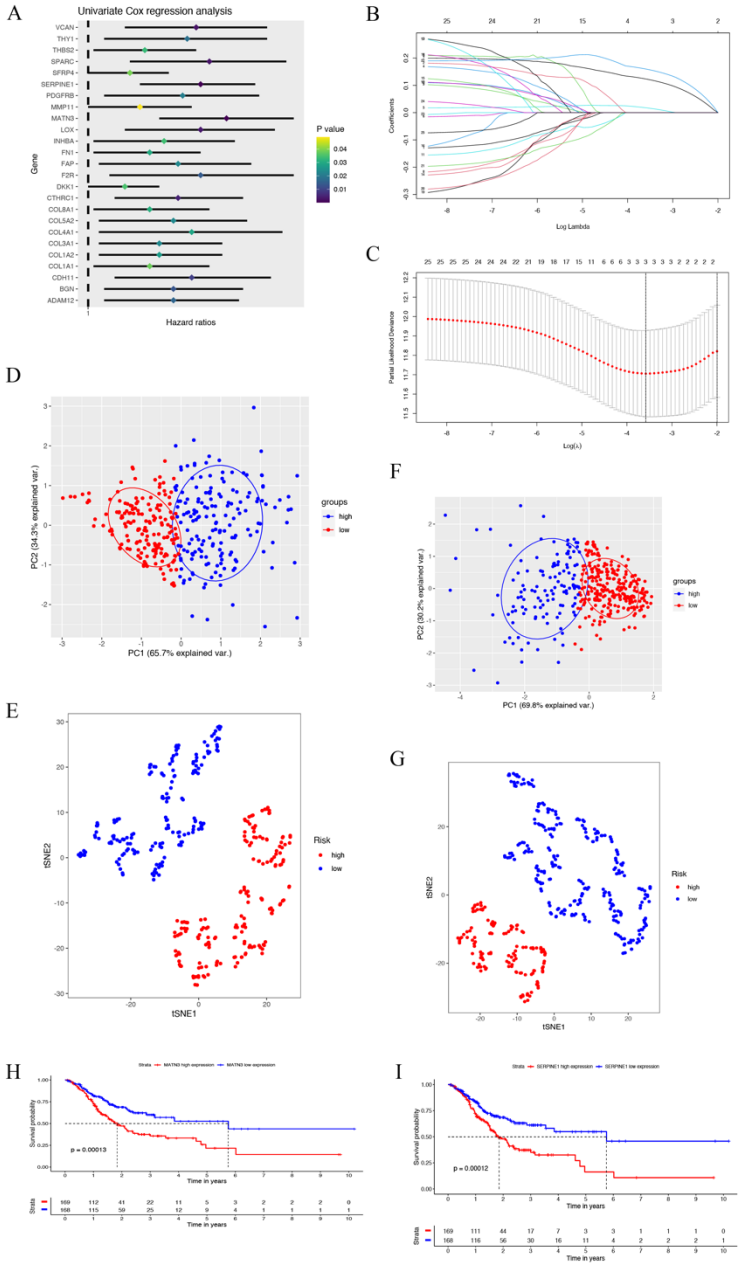


图 2: 构建预后风险模型。(A) 25 个候选基因的单变量 Cox 回归分析。(B) 3 个候选基因的 LASSO 系数谱，在 10 倍交叉验证选择的值处画一条垂直线。(C) LASSO 回归中调优参数选择的十倍交叉验证，左边的垂直线代表最终确定的 3 个基因。(D) TCGA 队列的 PCA 分析。(E) TCGA 队列的 t-SNE 分析。(F) GEO 队列的 PCA 分析。(G) GEO 队列的 t-SNE 分析。(F-G)预后基因的生存曲线。

3.4 富集分析

GO 富集分析的结果表明（图 5A），在生物过程中，基因主要参与白细胞细胞粘附、淋巴细

胞增殖、单核细胞增殖的调节。细胞组分的分析结果表明，基因主要在 MHC 蛋白复合物和内质网膜处富集。分子功能的分析表明，基因主要参与多肽、MHC II 类受体和细胞因子受体的结合和活性。KEGG 富集分析的结果发现基因主要涉及单纯疱疹病毒、沙门氏菌和人类 T 细胞白血病病毒感染和肿瘤坏死因子信号通路（图 5B）。

3.5 肿瘤微环境

利用 CIBERSORT 算法计算了 22 种免疫细胞在高风险组与低风险组的表达情况（图 6A），发现 T cell CD4 memory activated ($P < 0.01$)、Macrophages M0 ($P < 0.05$) 和 Macrophages M2 ($P < 0.01$) 在高低风险组的表达存在差异。通过免疫细胞和风险得分（risk score）的相关性分析发现，风险得分和 B cell memory ($R = -0.15$, $P = 0.0046$) 和 T cell CD4 memory activated ($R = -0.15$, $P = 0.0054$) 呈现负相关；和 Macrophages M1 ($R = 0.12$, $P = 0.034$)，Macrophages M2 ($R = 0.12$, $P = 0.049$) 呈现正相关（图 6B）。与低危组相比，高危组的免疫评分(Immune Score)、基质评分(Stromal Score)和 ESTIMATE 评分(ESTIMATE Scoe)较高，提示高危组 TME 中免疫细胞浸润和基质成分水平较高，但肿瘤纯度较低（图 6C）。

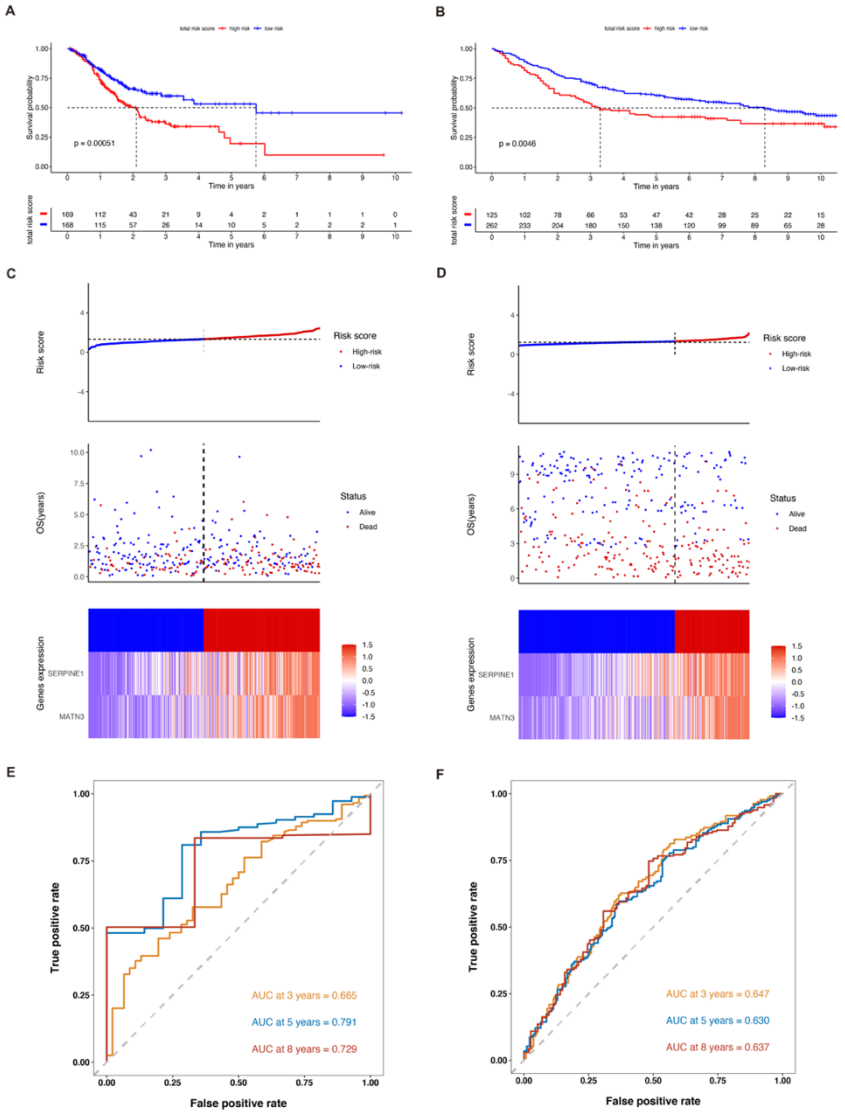


图 3: 风险模型的验证。(A-B) TCGA 和 GEO 数据集中高低风险组患者的生存分析。 (C-D) TCGA 和 GEO 数据集中高低风险组患者的风险曲线。(E-F) TCGA 和 GEO 数据集中风险评分模型的 ROC 曲线。

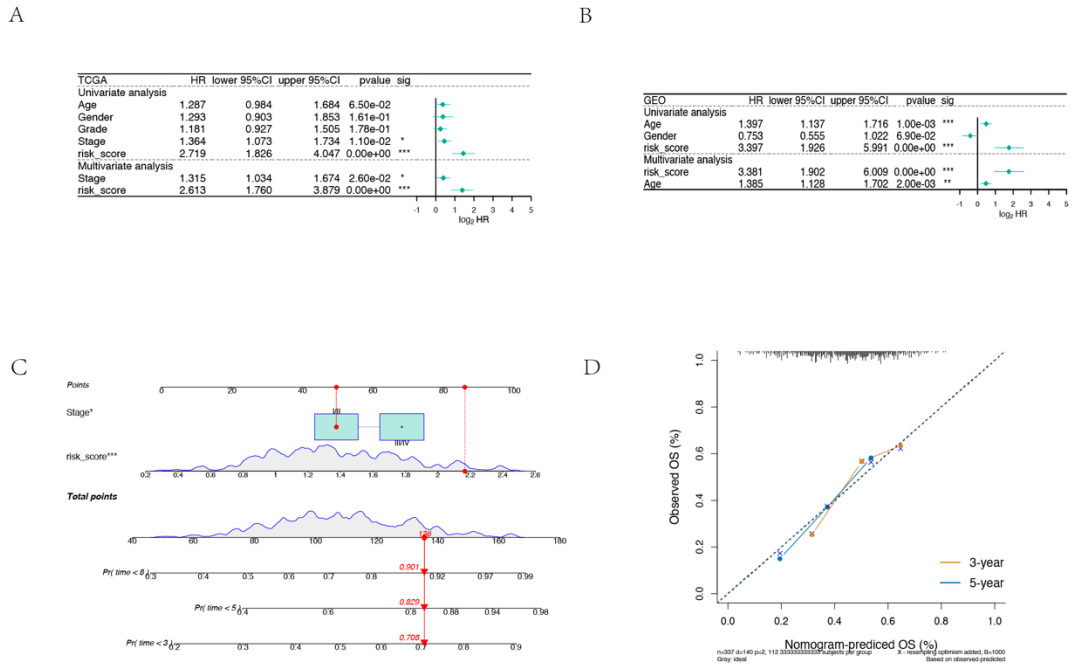


图 4. 风险模型的独立预后分析及基于多因素回归分析构建列线图。 (A-B) 训练集（A）和验证集（B）的森林图。（C）胃癌患者预后的诺莫图。（D）校准曲线用于评估诺莫图的准确性。

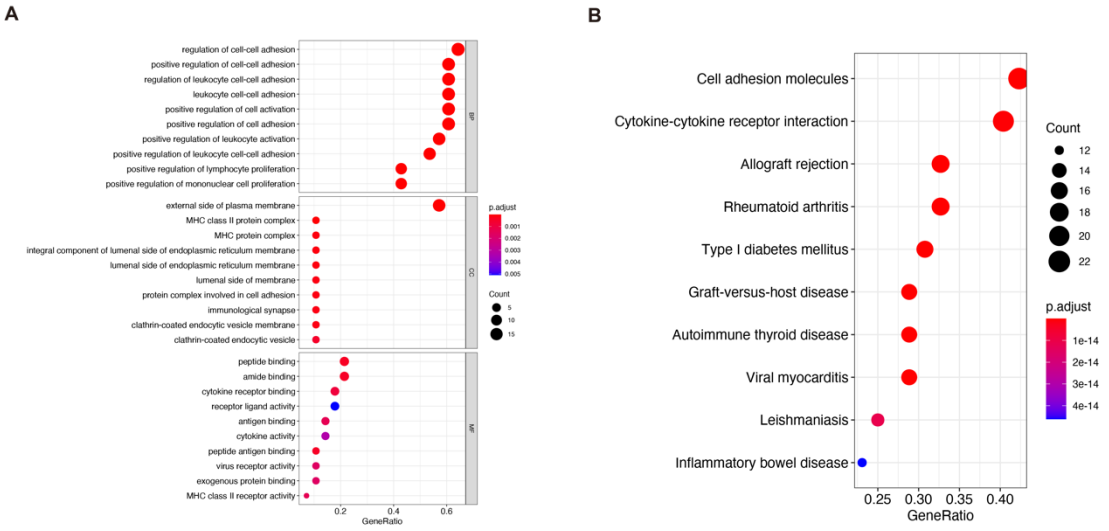


图 5. 差异表达的上皮细胞-间充质转化相关基因的富集分析。（A）GO 富集分析。（B）KEGG 富集分析。

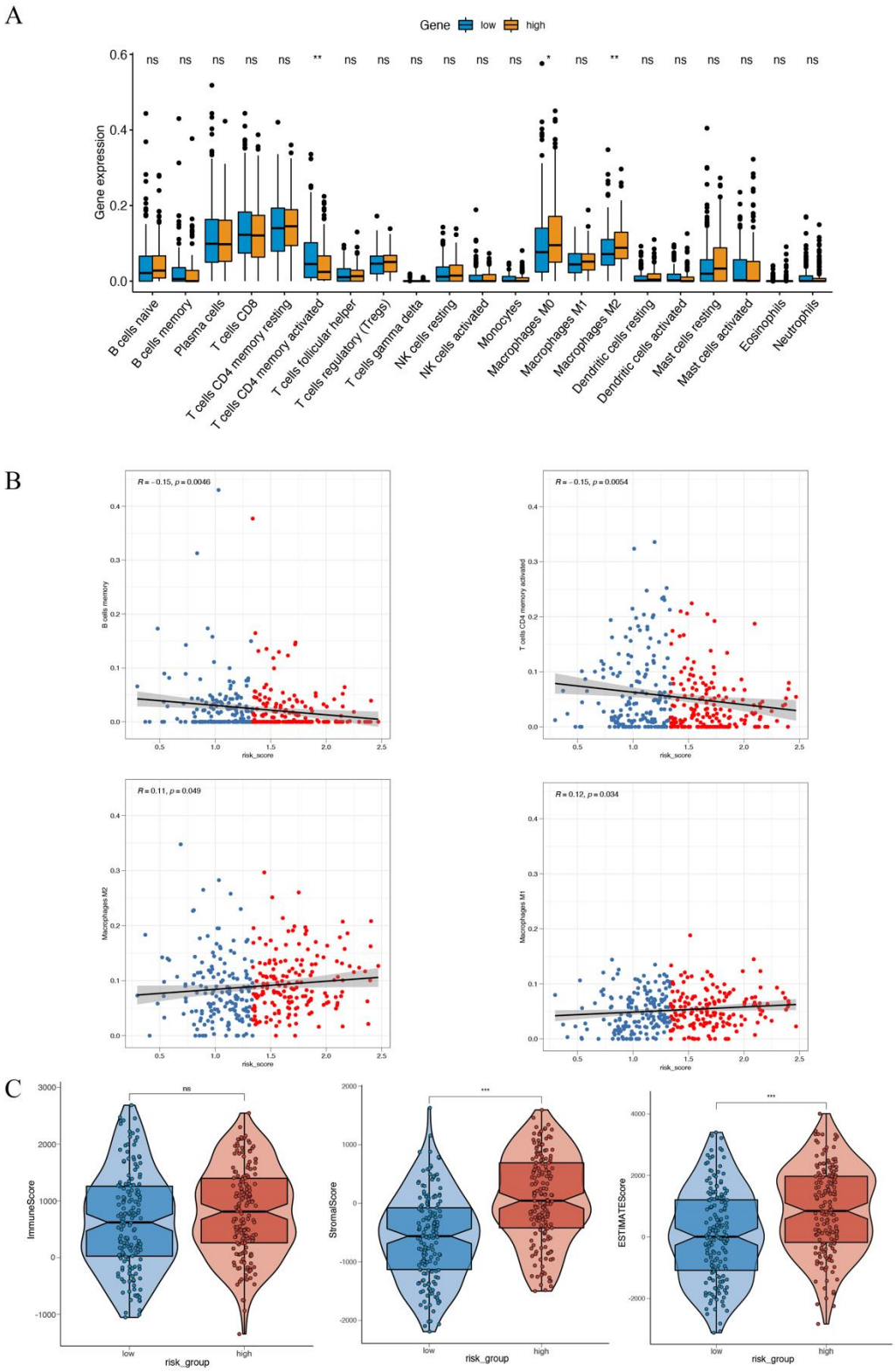


图 6. 免疫微环境分析。（A）22 种免疫浸润细胞在高风险组与低风险组的表达情况。（B）风险得分和免疫细胞的相关性分析。（C）ESTIMATE 评分在高低风险组的表达情况（免疫评分 (Immune Score)、基质评分 (Stromal Score)、ESTIMATE 评分 (ESTIMATE Score) ）。

3.6 验证风险基因的表达

通过 RT-qPCR 分析显示，MATN3 和 SERPINE1 的表达水平在胃癌组织中上调（图 7）。

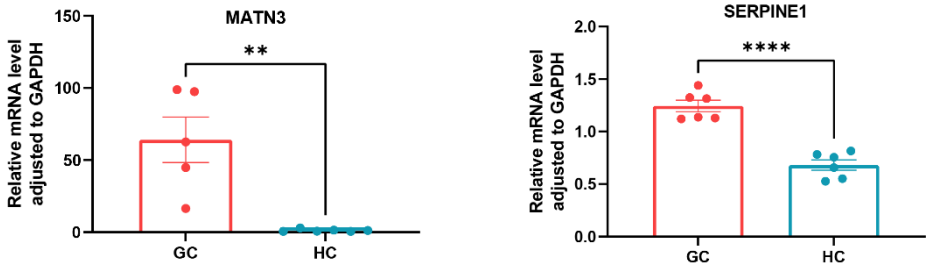


图 7. 通过 RT-qPCR 验证风险基因的表达水平。

4 讨论

本研究构建了一个包含 2 个上皮细胞-间充质转化基因的胃癌预后模型,并在 GEO 数据库中做了验证，且该模型 5 年生存率 ROC 曲线线下面积（AUC）为 0.791，表明该模型具有较好预测价值。人母系蛋白 3（MATN3）是母体蛋白家族的成员，是一种非胶原细胞外基质。相关的研究表明，MATN3 在胃癌中高表达与患者的预后不良及相关 [12]。在胃癌中，MATN3 常发生异常甲基化和功能失调，与 MATN3 可能是参与胃癌细胞生长和转移过程的关键靶基因有着密切的联系 [13]。异常 SERPINE1 DNA 甲基化参与卡铂诱导的上皮性卵巢癌上皮-间质转化 [14]。SERPINE1 它通过调节上皮-间充质转化（EMT）来促进肿瘤细胞的增殖、迁移和侵袭 [15,16]，影响胃癌患者预后 [8]。本研究和相关的研究都表明，这两个基因都与上皮细胞-间充质转化有着密切的联系，都作为高风险基因，不利于胃癌患者预后。

高低风险组的富集分析表明，GO 富集分析表明这些基因主要参与白细胞细胞粘附、淋巴细胞增殖、单核细胞增殖的调节；多肽、MHC II 类受体和细胞因子受体的结合。KEGG 分析表明这些基因主要参与细菌感染和肿瘤坏死因子信号通路，这些信号通路与 TME 的特点有关。与低风险组相比，高风险组中活化的 CD4+记忆 T 细胞浸润程度显著降低，而 M0 和 M2 巨噬细胞浸润水平增加。研究表明，M2 巨噬细胞通常受到由肿瘤微环境中的致瘤信号而由巨噬细胞 M0 分化而来。其作为免疫抑制细胞，通过促进血管生成、转移扩散和局部免疫抑制表现出促肿瘤活性，与肿瘤患者的不良预后相关 [17]。此外，高危组 TME 中免疫细胞浸润和基质成分水平较高，但肿瘤纯度较低。相关研究表明，TME 的特征可分为三大类 [18,19]: (1) 免疫炎症: 免疫细胞存在于肿瘤细胞附近; (2) 免疫排斥: 免疫细胞存在于间质周围但不穿透肿瘤; (3)免疫荒漠: 免疫细胞缺乏浸润。我们推测高风险组患者的 TME 符合免疫排除亚型。在尽管 TME 表现出大量的免疫细胞浸润，但免疫细胞无法有效穿透肿瘤实质清除肿瘤细胞。因此，高风险组患者的预后较差。

综上所述，本研究基于上皮细胞-间充质转化基因集构建胃癌患者的风险预后模型，并通过收集临床胃癌样本，进行外部实验验证，证明模型构建基因于胃癌样本及癌旁样本具有显著表达差异，证实模型应用的可靠性。有利于系统的评估上皮细胞-间充质转化基因在胃癌中的表达模式及其作用。此外，我们还评估其预后价值以及对肿瘤微环境的影响。基于风险模型，可识别出高风险的患

者，了解不同风险患者的发病机制，以改善预后风险分层并有助于医生对患者制定有效的治疗决策，进行有效的治疗，提高我国胃癌患者的生活质量。

PO-032

自然语言处理技术在在线健康社区研究中的应用现状：系统综述

王哲轩[1,2,3,4]，舒泳搏[1,2,3,4]，罗爱静[1,2,3,4*]

(1 中南大学湘雅二医院，长沙410011
2 中南大学生命科学学院

1. 引言

在线健康咨询社区是一种通过互联网平台提供健康信息和医学咨询服务的虚拟社区。这些社区汇聚了大量的医学专业人员、健康专家以及关心健康的用户。在在线健康社区中，医疗健康知识突破了空间的限制进行传递，成为了医疗信息领域中重要的一环。一方面，在线健康社区可以使患者随时随地通过在线健康咨询社区获取医疗信息，包括疾病诊断、治疗方案、用药建议等，无需前往医院，节省时间和精力[1]。另一方面在线健康社区提高了患者的健康知识和素养，为患者提供了丰富的健康教育资源，增加患者对健康问题的认知和理解[2]。目前在线健康咨询社区在在现代医疗领域的重要性越发提高，除了本来具备的咨询功能外，其在使用过程中还产生了大量的医学信息，包括健康咨询和医学问题、医学知识分享、病例讨论和经验分享、医学新闻和研究进展等，比如 Huh, Jina 等[3]对在线健康社区中临床知识的应用情况进行了研究。在线健康社区中产生的信息具有来源广泛、更新及时、表达自由等特点，因此如何更好地利用这些医疗信息也成为了当下研究的一大热点。

然而，随着用户在在线健康社区中产生了海量的信息，其中也存在着大量干扰噪音。因此需要研究者选用合适的方法去对其进行处理与分析，从而提升处理效率并得到更好的分析结果。

自然语言处理（Natural Language Processing，NLP）是一种利用计算机技术和自然语言学原理[4]，实现人与计算机之间以自然语言进行通讯的科学技术。NLP 旨在让计算机能够理解和处理人类自然语言，从而使得计算机能够更智能地服务于人类生活。自然语言处理与文本挖掘技术的高速发展很好的满足了研究人员的需求通过自然语言处理和文本挖掘技术，深入分析和理解在线健康社区中的内容。

目前大量研究都使用了在线健康社区的信息与 NLP 技术，早在 2014 年 Liu, Yang 等[5]研究者便开发了一种新的集成生物医学 NLP 技术，用来从在线健康论坛中提取一组全面的患者人口统计数据和医疗信息。随着 NLP 技术的不断发展进步，其在在线健康社区中的应用也变得更为广阔，目前尽管已经有一些研究针对特定的应用方向进行了综述，但是本综述旨在更全面地回顾和总结 NLP 技术在在线健康社区中的整体应用现状而非特定方向或者特定 NLP 技术的现状，从而分析得出：（1）当前自然语言处理与文本挖掘技术在在线健康社区的应用现状与特征，（2）未来 nlp 技术在在线健康社区信息挖掘分析的潜在价值与发展趋势。为未来深入研究和实践提供一个坚实的基础。

在接下来的章节中，我们将会详细介绍本综述的研究方法，呈现我们的综述结果，并对其进行讨论和解释。最后我们将总结本综述的主要发现，并在此基础上提出关于在线健康社区和文本分析技术的未来研究建议。

3. 研究材料与方法

本综述的结构遵循 PRISMA（系统综述和 meta 分析的首选报告项目）的原则[6]，但由于本综述研究目的是对这一研究领域的综合概述，而不是对特定研究方向的质量评估。所以，本综述并未进行质量评估与偏倚性分析。但在人工筛选过程中，我们使用了较为详细的评判的标准来增加综述可信度。

3.1 研究纳入与排除标准

结合我们的研究目标，对纳入本系统综述的研究采用了以下排除标准：

1. 研究应该明确说明其使用了至少一个在线健康社区的数据。2. 研究应该明确说明其设计使用了至少一种 NLP 技术并对其应用方法进行了详细描述。3. 文献类型应该为发表在相关期刊上的 Article。4. 文献发表年份应该限制为近五年内。

此外，如果文献有以下情况应该排除：

1. 重复出版物。2. 非 article 类型文献，比如综述、社论、编辑材料。3. 未发表在相关期刊上，比如学术论文。4. 发表年限早于 2019 年 1 月 1 日。5. 未使用 NLP 技术或未使用在线健康社区作为研究对象。

3.2 检索策略与数据来源

本文检索了 web of science（WoS）和 PubMed 两个数据库，由于文本挖掘和文本分析是 NLP 技术应用最广泛的领域之一，而在对在线健康信息的利用中，文本处理也主要集中获取在文本数据以及分析文本数据两个方面，因此我们限定 nlp 技术应用的数据来源未在线健康社区，NLP 技术的体现方式为文本分析与文本挖掘，具体两个数据库的检索策略与

结果见表 3-1：

表 3-1 检索策略与检索结果

数据库	检索策略	检索结果
Web of Science（所有数据库）	((TS=("online health communities") OR TS=("Health Social Media")OR TS=("Online health consultation"))AND (TS=(Text processing) OR TS=(text mining) OR TS=(text analysis))) AND DOP=(2019/2023) NOT review	49
PubMed	((((("online health communities")OR ("Health Social Media") or ("Online health consultation")) AND((text analysis) OR (text mining) OR (Text processing))) AND (("2019"[Date - Publication] : "2023"[Date - Publication]))) NOT (review[Publication Type])	23

3.3 审核筛选

本研究使用了 Zotero 文献管理软件对检索文献进行处理，手动对检索结果去重以及筛选了文献类型后，由两名评审员通过标题与摘要内容进行独立筛选，如有分歧则需要第三名评审员共同协商，并通过全文内容进行审查选定。整体核查流程见下图 3-1

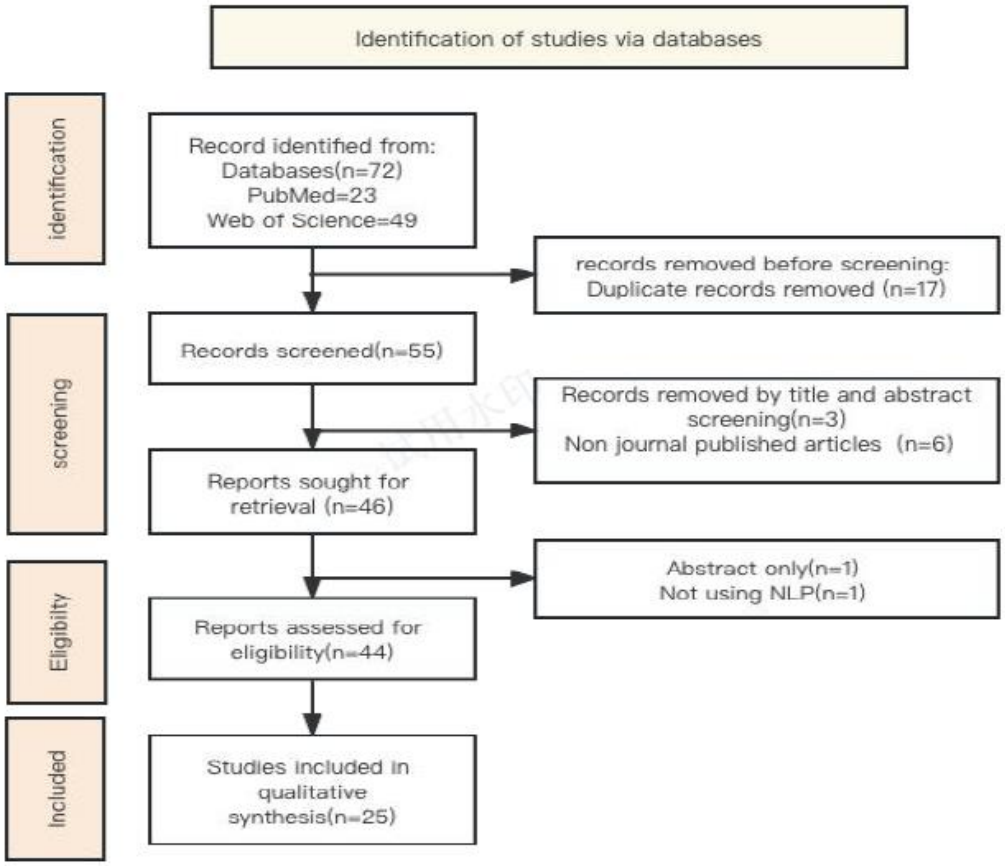


图 3-1 PRISAMA 研究筛选流程图

4. 结果

由于本综述是对 NLP 技术在在线健康社区领域内的应用现状进行研究，所以上述选取文章的研究目的、研究方法和研究对象都并无统一的研究对象、研究方法或者是评价标准，很难在这些方面将其整体情况进行分类讨论。最终我们选择把选取的文献的整体情况列为两张表概括，表 3-1 列出了 NLP 处理相关的特征：数据量大小、数据收集方法、文本分析方法；表 3-2 列出了在线健康社区内容相关的特征：在线健康社区种类、研究病种、研究目的。由于本综述主要是对各研究的研究方法与研究对象进行特征分析，所以对于研究的结论并未具体列出。

由表 3-1 可知，每一篇纳入研究的文献在数据收集方法与数据分析方法两个方面至少有一项使用了 NLP 技术，大部分文献使用了网络爬虫收集文本数据，然后对其进行文本分词筛选等预处理，最后再对其进行分析研究。

此外我们对表 3-2 中出现的所有在线健康社区以在线健康社区网站的开设国别进行分类（避免跨国机构干扰），以上文献中选择美国的在线健康社区的数量最多，且涉及的不同在线健康社区数量也最多，一共包括 20 个相关社区，其中 Reddit 是被选择研究次数最多的在线健康社区。Reddit 和 twitter、facebook 一样是美国的社交新闻和论坛网站，其与以健康相关信息分享为主的传统在线健康社区，比如 Breastcancer.org、Cancer Survivors Network 等有所不同，这些网站的社交属性更强，健康社区只是属于其中的一个子版块。而以中国的在线健康社区为研究对象的文献数量排名第二，其中使用最多的是好大夫在线这一平台。好大夫在线是中国目前最具有影响力的在线健康社区，其功能包括在线问诊、医学科普、医生推荐等方面，是以医学信息分享为主的传统在线健康社区。而在中国其他在线健康社区中，只有百度贴吧与美国的 Reddit 类似，都是在社交网站下开设的在线健康信息分享子板块。其他国家的在线健康社区研究相对较少，但其研究价值也十分重要。位于英国的 HealthUnlocked 是全球最大的在线健康社交网络平台之一，拥有超过 120 万名注册会员，该网站与美国的 Reddit 相比，其专业度更高，患者可以与医护人员取得联系获取建议与支持。同时它又与中国的好大夫在线有所不同，其社交功能的自由度更高，比如患者之间可以进行病情交流。另外 Camoni 平台是一个使用以色列语言的在线健康社区，其功能与英国的 HealthUnlocked 类似，都是信息交流自由度相对较高的专业健康信息平台。

从在线健康的用户研究角度来看，以患者与患者之间的交流文本为研究对象的文献数量最多，其次是患者与医生之间的交流文本内容，但也有文献使用了医生与医生之间的交流文本内容。从选择的研究话题来看，糖尿病是研究次数最多的文本主题，艾滋病、抑郁症、乳腺癌、新冠肺炎等主题研究次数也相对较多。

4.1 总体特征

表 3-1 NLP 技术特征表

作者	数据量情况	数据挖掘方法	文本分析方法
Xie, Jiaheng 等[7]	纳入超过 100,000 条	文本挖掘和网络爬虫	多方法框架去分析相关文本数据，包括一种新颖的医学增强深度学习（MEDEL）特征提取方法，经济计量分析，和一个随机实验。
Yao, Zhizhen 等[8]	使用了两个数据集：34,718 条帖子 和 46,136 条帖子	网络爬虫	BERT 分类器。
Haldar, Shefali 等[9]	30 名受试者的相关评论	调用平台使用日志和相关信息	IBM Watson 的自然语言理解 API（一种文本分析工具）进行了情感分析。
Fang, Jiaming 等[10]	5,452 名医生的 246,294 条文本评论数据	python 程序	一种混合情感分析服务质量的方法，结合手动选择、基于语义的无监督学习模型和深度学习。
Alzahrani, Sara 等[11]	4301 条推文	通过 Twitter API 收集	用 TF-IDF 方法将语料库转换为词向量表示，并应用了 11 种分类模型，采用交叉验证对语料库进行了验证。
Liu, Jingfang 等[12]	17,891 个文档的语料库	python 程序	TF-IDF 算法对文本进行变换，Latent Dirichlet Allocation（LDA）主题生成模型用于文本主题挖掘。
Liu, Xuan 等[13]	3,802 用户的 341,834 消息帖子	网络爬虫	手动编码训练集指南，使用这些分类消息来训练机器学习分类器，最后使用随机系数模型。
Naderi, Hamid 等[14]	3,000 条帖子	PHP 脚本	使用自然语言处理工具 MetaMap 来提取问题和答案中的医学概念,映射到 UMLS 的 Concept Unique Identifiers。基于词频-逆文档频率为概念赋予权重。
Bulcock, Alexander 等[15]	1,359 份完整的调查问卷和 20 名焦点小组参与者的讨论数据	调查问卷回收、手工记录	定性软件（NVivo）对数据进行了主题分析。
Wang, Xiangyu 等[16]	2,700,000 个帖子	Breastcancer.org 提供	人工注释训练集，机器学习分类，回归模型来调查用户的首次寻求支持体验是否可以预测他们在 OHC 中的未来发布行为。
Wang, Xiangyu 等[17]	约 50,000 名用户的 2,800,000 条帖子，	Breastcancer.org 提供	深度学习方法、并行 LSTM 神经网络模型。
Hu, Lei 等[18]	1,605 条记录	网络爬虫	使用生根理论用于指导文本分析进行编码处理。编码阶段分为以下三个主要阶段：（1）开放编码，（2）轴向编码和（3）选择性编码。
Yang, Hangzhou; Gao, Huiying[19]	53,198 个帖子、331 个用户档案和 895 对朋友关系。	未具体说明	使用 LDA 提取主题信息，然后使用 XGBoost 模型学习用户之间的影响关系，并最终使用改进的 SimRank 算法计算用户之间的行为模式相似度。
Ouyang, Peng 等[20]	20,000 多名医生的数据	网络爬虫	结合了情感标注、无监督学习技术（LDA 主题建模）、Gensim 中的语义主题多样性估计和分层聚类方法。
Lee, Young Ji 等[21]	46 名参与者发表的 885 篇帖子	网站提供	使用 LIWC 词典评估了单词使用情况，使用了 LDA 模型分析主题。
Chen, X [22]	9,500 条帖子	网络爬虫	通过情感分析自然语言处理模型对提取的数据进行系统编码、提炼和分类。
Xu, Qian 等[23]	1,889 条帖子	网络爬虫	RoBERTa-BiLSTM-CRF 模型进行实体识别。
Luo, Aijing 等[24]	1,000 条帖子	C# program	基于 Python3.6 和 R 语言实现了文本关键词共现分析和聚类分析。
Gao, Yuyang 等[25]	2,882 名用户数据	未具体说明	使用了一种名为 DynAttGraph2Seq 的框架来对这些文本数据进行分析建模，从而预测用户的健康阶段。
Foufi, Vasiliki 等[26]	17,624 个文本帖子	网络爬虫	使用 PKDE4J 文本挖掘系统实现基于词典的实体抽取和基于规则的关系抽取。
Liu, Yi-Hung [27]等	40 个主题下的 1,624 帖子， 包含了 19,743 句子	未具体说明	NLTK 对文本句子进行处理，使用 MetaMap 进行医疗实体提取。
Sampath, Pradeepa 等[28]	1,824 条帖子	web API	Python NLP NLTK 和 Scikit 学习包进行数据集预处理。卡方检验提取单词的基本特征，用 K-means++ 算法对它们进行聚类，生成概率模型（LDA）来确定每个集群中的基本主题。
Ansari, Bahareh; Du, Shichao [29]	来自 1,500 个成员的 1,1069 条留言	Python 程序	基于 LIWC 词典来请求留言中的单词特征，使用机器学习算法自动编码文本信息和情感支持的程度，最后使用线性回归来评估两者关联度。
Wang, Jie 等[30]	8,178 条帖子	未具体说明	基于 python 的数据预处理，基于“同义词词库扩展版”生成词汇意义共现矩阵，基于词汇意义共现 CL-LDA 模型得到文本主题聚类结果
Li, Mingda 等[31]	25,208 条帖子	网络爬虫	先使用关系提取模块提取文本关系。然后分别使用文本相关性度量模块、问题概率计算模块和动作概率计算模块提取相关文本特征。最后，在特征组合模块中使用深度学习模型对所有特征进行组合。
Wang, Jie 等[32]	60,478 条帖子	网络爬虫	预训练模型 SBERT 为问题文本生成句子向量，采用 k-means 聚类算法得到特定疾病主题，使用 TF-IDF 从聚类结果中提取关键字。
Chowdhuri, Sanchari 等[33]	10,584 条帖子和评论	网络爬虫	用 HPO 概念识别软件从文本中识别人类表型本体（HPO）术语。用 MMLite 从文本中识别统一医学语言系统（UMLS）元词典中的生物医学术语。
Yu, Haiyan 等[34]	9,841 名医生和 704,467 条投票	未具体说明	使用倾向得分模型和反向概率加权法估计了电子咨询服务对反应变量的平均处理效应。
Jiang, Shan 等[35]	无	未具体说明	使用 TextMind 分析中的词语使用和语言风格。
Yao, Z 等[36]	抑郁症社区中 164,100 个帖子和 809,165 条评论，糖尿病社区中 65,688 帖子和 76,218 评论	网络爬虫	使用 LDA 的主题生成模型建立文本主题模型，使用 TextMind 进行情感极性识别，利用 Ucinet 6.0、Gephi 0.9.2 进行社交网络分析。

Park, Insu 等[37]	53,831 个帖子	Python 程序	人工在短语层面进行内容分析编码，再使用 Atlas.Ti 软件，将编码后的文本数据转换为计数数据。
Lee, Young Ji 等[38]	50,626 篇帖子	未具体说明	人工标注特征，使用 Neo4j 图数据库构建知识图谱。
Bi, Q 等[39]	5,802 条帖子，45,181 评论	网络爬虫	使用了 TF-IDF、LDA 主题模型和朴素贝叶斯分类器相结合的文本主题分类与挖掘方法。
Chen, Yuangao 等[40]	63,188 记录	Python 程序	建立了一个经验模型来验证假设，使用文本信息作为模型变量，采用了逻辑回归和普氏回归作为主要的模型分析方法，同时辅以一些基础的统计分析来评价模型。
Balakrishnan, Athira 等[41]	150,000 条帖子	未具体说明	使用了一种名为 BiLSTM 的深度学习模型来分析用户在社区中的情感变化。
Bitton, Yonatan 等[42]	1,000,000 条帖子和 100,000 条搜索日志	网络爬虫	使用了基于深度学习的方法——MDTEL（医学深度音译实体链接）来分析相关文本数据。
Peng, Li 等[43]	28,546 条记录	网络爬虫	使用文本挖掘和机器学习的方法，包括词频统计、情感分析、主题建模、聚类分析和回归分析，来提取和分析咨询文本中的信息和特征。
Liu, Mengque 等[44]	6,262 条帖子	网络爬虫	基于单词共现网络分析，保留文本的语义结构。从语义结构中提取信息得到网络结构定量和定性分析的相互作用的支持。
Alnashwan, Rana 等[45]	Lyme 病论坛：1,491 篇，Lupus 病论坛数据：1,085 篇	未具体说明	对文本的词汇特征、情感特征、关键词特征进行了挖掘，然后基于机器学习算法对这些特征进行分类。
Dreyfus, Brian 等[46]	76,912 个帖子	Python 代码	使用 XGBoost， TF-IDF 和 Count 向量化的组合进行文本分类，拼写校正（CSpell）和医学实体识别技术（cTAKES）从在线卫生社区提取医学信息。
Zhang, Yunqiu 等[47]	4,916 名医生数据	Python 程序	提取词汇特征，使用词频、TF-IDF 等方法将文本转换为向量，比较经典的机器学习分类方法对其分类效果。
Wang, Xuqi 等[48]	无	网络爬虫	使用词向量、Bi-LSTM、注意力机制、量子计算等技术对文本进行表示和特征提取，并构建了量子 Bi-LSTM 模型。
Du, Yinfeng 等[49]	4,412 条关于 10 位医生的有效评价	网络爬虫	TF-IDF 技术用于提取关键词和主题，SKEP 模型用于进行情感分析，PLTSs 用于将非结构化的文本评价转化为结构化的数据，PL-PT-MULTIMOORA 方法用于综合考虑医生的基本信息和患者评价。
Xiang, Minhao 等[50]	10,728 条评论	网络爬虫	BERT-LDA 模型去分析相关文本数据。

表 3-2 在线健康社区特征研究表

作者	在线健康社区名称	研究的病种	研究目的
Xie, Jiaheng 等[7]	Caring.com, AgingCare.com, AARP.org	老年护理	医学语言的正式程度如何影响在线健康信息的有用性。
Yao, Zhizhen 等[8]	百度贴吧	糖尿病	探讨在线糖尿病社区中用户寻求和分享健康信息的模式和内容，以及他们表达的方式和主题的异同。
Haldar, Shefali 等[9]	研究者自行搭建的平台	无	医院同伴支持的性质以及这种支持对患者住院的影响。
Fang, Jiaming 等[10]	好大夫（www.haodf.com）	无	医生评论文本中的服务质量信号对患者择医的影响。
Alzahrani, Sara 等[11]	Twitter	新冠肺炎	构建一个关于沙特阿拉伯方言推文的情感分析语料库。
Liu, Jingfang 等[12]	百度贴吧	心脏病，高血压、抑郁症、强迫症	探究生理和心理疾病患者之间兴趣话题的差异，并讨论如何满足这些差异。
Liu, Xuan 等[13]	甜蜜家园（ bbs.tnbz.com）	糖尿病	影响患者获得社会支持的社会资本因素，旨在为在线寻求社会支持的患者提供指导。
Naderi, Hamid 等[14]	Yahoo Answers, WebMD, NetWellness	无	比较提交给在线健康社区的问答之间的主题间和主题内的相似性。
Bulcock, Alexander 等[15]	HealthUnlocked, Facebook, Twitter	类风湿性关节炎，纤维肌痛，呼吸系统疾病，甲状腺疾病和/或心理健康问题	本研究旨在探讨公众对使用自动化方法检测和报告社交媒体上提及的 ADRs 以加强药物警戒工作的看法。
Wang, Xiangyu 等[16]	Breastcancer.org	乳腺癌	通过分析用户在 OHC 中的首次寻求支持体验，预测他们随后在在线健康社区的参与行为。
Wang, Xiangyu 等[17]	Breastcancer.org	乳腺癌	通过学习 OHC 用户在帖子数量和内容方面的时间轨迹来预测他们未来的发布活动。
Hu, Lei 等[18]	未说明	糖尿病	为了明确糖尿病患者日常生活中自我管理的真正需求，并明确指南等医学标准与患者需求之间的矛盾和错位。
Yang, Hangzhou; Gao, Huiying[19]	甜蜜家园（ bbs.tnbz.com）	无	提供一种内容推荐方法，为在线健康社区（OHC）的用户推荐合适的内容。
Ouyang, Peng 等[20]	好大夫（www.haodf.com）	无	调查医生自我介绍的信息对患者在线健康社区决策的影响，研究医生在线声誉的调节作用。

Lee, Young Ji 等[21]	the Network	癌症	通过对帖子进行计算文本分析，为患有癌症的性少数群体女性（SMW）探索在线健康社区。
Chen, X [22]	丁香园	新冠肺炎	如何利用或增强在线健康社区功能，帮助老年人克服疫情带来的困难。
Xu, Qian 等[23]	好大夫（www.haodf.com）	糖尿病	利用 RoBERTa BiLSTM CRF 的组合模型来完成与糖尿病相关的医疗实体识别任务，从而提高实体识别性能。
Luo, Aijing 等[24]	好大夫（www.haodf.com）	高血压	如何在网络环境下充分揭示高血压患者在问题中表达的健康信息需求，以及如何利用互联网帮助高血压 患者接受个性化的健康教育。
Gao, Yuyang 等[25]	breast cancer community，American bladder cancer society	抑郁症、焦虑症	验证作者提出的一个文本处理框架在预测在线健康社区用户的健康阶段方面的有效性和可解释性。
Foufi, Vasiliki 等[26]	Reddit	无	研究在线健康社区用户生成的文本帖子中与慢性病相关的实体及其关系。
Liu, Yi-Hung [27]等	Diabetes Community，POZ Community Forums	糖尿病、艾滋病	介绍一种新颖的文本摘要方法，以准确有效地从在线患者帖子中获取信息量最大的摘要。
Sampath, Pradeepa 等[28]	Med Help	糖尿病、甲状腺疾病和结核病	从 OHC 文本提取关键字并探究三种慢性疾病之间的相互关系。
Ansari, Bahareh; Du, Shichao [29]	forum.drugs.com	无	探索在线药物康复论坛中社会支持的交流过程，并确定交流中的语言特征。
Wang, Jie 等[30]	39 健康网、好大夫在线、问答网、求医问药网、微医网和春雨医生网	新冠肺炎	中国在线健康社区（OHC）对新型冠状病毒肺炎（COVID-19）的信息需求。
Li, Mingda 等[31]	CSN breast cancer forum		确定与医疗保健决策相关的讨论线程中帖子的影响关系。
Wang, Jie 等[32]	寻医问药	无	结合在线健康信息需求的统计分析和主题分析，探讨中国青少年相关信息的分布、影响因素和关注话题。
Chowdhuri, Sanchari 等[33]	BabyCenter.com	产后抑郁	在线健康支持小组对经历产后抑郁症 （PPD） 等临床状况的人的社会支持作用。
Yu, Haiyan 等[34]	好大夫（www.haodf.com）	糖尿病、癌症、心脏病、抑郁症和艾滋病	研究电子咨询的可获得性如何影响病人对医生专业能力的评估。
Jiang, Shan 等[35]	甜蜜家园（ bbs.tnbz.com）	糖尿病	探讨在线健康社区中写作风格对社会支持的影响。
Yao, Z 等[36]	百度贴吧	抑郁症、糖尿病	比较在线心理和生理疾病社区在主题、情绪、参与和情绪传染模式方面的差异，并讨论如何满足用户不同的信息和情感需求。
Park, Insu 等[37]	tudiabetes.org	糖尿病	探讨在线医疗社区中信息支持的作用机制。
Lee, Young Ji 等[38]	Cancer Survivors Network	卵巢癌	研究通过文本挖掘在线健康社区的数据来构建症状知识图谱的可行性和价值。
Bi, Q 等[39]	百度帖吧	抑郁症	利用多种方法确定中国 OHC（百度帖吧）白化病的主题分类和情感极性
Chen, Yuangao 等[40]	好大夫（www.haodf.com）	白血病，肺癌、肝硬化、高血压，关节炎和胃炎	探讨在线健康社区（OHC）中患者在电话咨询和文本咨询之间服务选择的决定因素。
Balakrishnan, Athira 等[41]	Breastcancer.org	乳腺癌	在线癌症社区中患者在接受治疗期间的情感变化。
Bitton, Yonatan 等[42]	Camoni	糖尿病、多发性硬化和抑郁症	如何在跨语言将希伯来文中的医学术语链接到英文的 UMLS 知识库中，以减少音译带来的变异性和噪声，并提高信息检索和数据挖掘的效果。
Peng, Li 等[43]	好大夫（www.haodf.com）	糖尿病	在线健康社区中的礼物赠送对医生服务质量的影响，以及不同病种和医生类型的差异。
Liu, Mengque 等 [44]	Patientslikeme.com	肺癌、乳腺癌、肠癌、基底细胞皮肤癌、前列腺癌、卵巢癌、胰腺癌、肾细胞癌"	检查不同类型癌症患者之间社会支持沟通的差异。
Alnashwan, Rana 等[45]	未具体说明	莱姆病	研究现有的自然语言处理（NLP）工具是否能够分析面向医学的公开讨论，识别其中包含的相关准确医学信息。
Dreyfus, Brian 等[46]	MedHelp，WebMD，Drugs.com，HealthBoards.com	无	建立一个自然语言处理模型，以识别这些患者帖子中提到的药物的标签外使用情况。
Zhang, Yunqiu 等[47]	好大夫（www.haodf.com）	无	疾病严重程度对在线医生行为影响的异质性问题。
Wang, Xuqi 等[48]	Twitter	无	验证基于 QBi-LSTM 的药物不良反应检测方法的有效性。
Du, Yinfeng 等[49]	好大夫（www.haodf.com）	无	如何在在线健康社区中选择最优和次优的医生，以帮助不同特征的患者（理性或情感）做出决策。
Xiang, Minhao 等[50]	好大夫（www.haodf.com）	无	在线健康社区用户的健康信息需求和情感表达。

5. 讨论

本系统综述的客观研究问题是“如何从在线健康社区产生的文本信息提取更多研究价值？”，为了解决这一问题，我们选取并分析了 44 篇利用了 NLP 技术对在线健康社区进行研究的文献，综述结果表明，目前 NLP 技术在文本信息收集与文本信息分析两方面都有着广阔的应用方向，但为了提取更多价值，我们需要结合在线健康社区的特征去选择或者开发合适的自然语言处理技术。因此我们将从在线健康社区与 NLP 技术两个方面对综述结果进行讨论分析。

5.1 在线健康社区特征与趋势讨论

从在线健康社区角度来看，以上文献主要是从社区特征、用户特征、信息特征进行研究，这一点符合当前研究者将社区、用户、信息视为在线健康社区三大要素的观点[51]，三者的关系如图 5-1 所示。

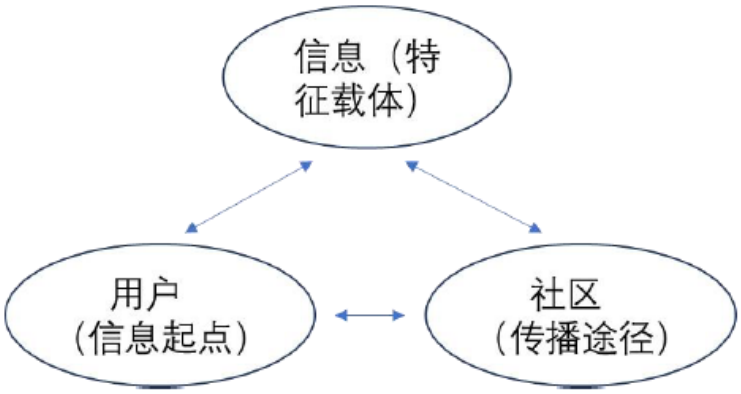


图 5-1 在线健康社区维度

（1）社区分类与特征讨论

社区不仅是信息赖以传播的路径，也是帮助用户生成信息的环境。在线健康社区的特点对其用户产生文本信息的影响是巨大的，社区往往会对用户的交流方式、文本信息格式、文本信息内容进行规范要求。根据用户交流的特点，我们的综述研究发现目前在线健康社区的主要类型有三种：

1. 医患咨询平台，这类平台往往会严格区分专业医护人员和咨询患者，患者提出信息需求，医护人员提供相关信息，这样做的好处是可以减少许多无意义的信息产生，比如过多的主观描述以及不相干的话题信息，从而提升信息质量，好大夫在线是中国最大的医患咨询平台之一，其内容专业度使其成为了许多实体医院对其患者的线上补充部分。使用这一类平台的文本信息作为研究对象可以更好地研究文本中的医学价值，比如识别疾病的症状描述、患者的求医行为研究等。

2. 社交媒体下的健康话题交流平台，在 2020 年新冠疫情之后越来越多人开始意识到在线健康信息的重要性，而他们平时使用的社交媒体平台就成了最直接的健康信息来源，比如推特、贴吧等。这类以社交为基础的交流平台的降低了健康信息产生的门槛，大量用户基础使其很容易在短时间内产生大量健康信息，这一特点有利于研究者利用这些文本信息快速对一个新兴的健康话题进行分析或者发现现有健康话题的新发展方向。但其产生的在线健康信息具有往往缺乏科学监督，容易产生噪音和虚假信息[53]，这导致了虽然通过这些网站获取的在线健康信息数量远多于医患咨询平台，但是质量却不如医患咨询平台。

3.基于特定健康信息的交流平台，这类平台结合上述两种平台的交流方式，它不会严格限制沟通的方式，但交流的信息内容存在一定的限制，比如要求产生的信息是与某一病种有关的规范表达。美国的 Cancer Survivors Network 就是一个关于癌症的开放式讨论社区，在一个讨论贴下出现的回复用户可能有同类病种患者群体、医生及其专业人员群体、关心该类病种的群体。此类社区中的文本信息的价值特点介于上述两类平台之间，是很好的研究信息补充来源。

（2）用户的交流模式与特征讨论

用户是信息的出发点，用户的需求决定了社区的发展方向。传统观点把在线健康社区的交流分为了两种：在线支持小组与有专业人员主持的在线平台[54]。而在本综述研究中发现在线健康社区的用户可以分为医生、护士等专业人员与病患、家属等非专业人员两类，根据交流方式的不同可以把健康社区信息交流的模式分为三类：

1.患者与患者之间的信息交流：对应在线支持小组；2.患者与医护人员之间的信息交流：对应有专业人员主持的在线平台；3.医护人员之间的信息交流：医护人员之间的信息交流模式目前研究较少，但其对医生行为研究具有重要的价值[55]。

不同的交流模式所产生的信息有着不同的特点，并不仅仅只有医护专业人员的信息是有价值的，也不是只有他们可以成为健康信息的产生者，有些类似疾病诊断的用户也可以分享其经验知识[56]。研究者往往会根据研究目的选择不同种类的信息，两者产生文本信息具有着不同的研究价值。在本研究选取的论文中，大多数选择的是患者与患者之间的信息，其次是医患之间的信息，最少的是医生之间的信息。具体来说，以上三种交流有以下几点差异：

① 内容差异：

医患交流的文本通常包含与医疗和健康相关的主题，如症状描述、诊断、治疗建议等[56]。患者之间的交流可能涉及到他们个人的医疗经验、治疗效果分享、情感特征，特别是社交支持这一主题在患者之间的交流中十分常见[57]。医生之间的交流通常是更专业化的主题，包括临床案例讨论、医疗研究、治疗方案等。

② 语言风格差异：

医患交流的文本通常更加正式和技术性，医生和患者可能使用专业术语和正式语言。患者之间的交流可能更加亲切和非正式，他们更有可能使用日常语言和情感性语言。医生之间的交流通常专业性极高，文本中会涉及大量专门的医学术语和病例描述。

③ 情感差异：

医患交流中，情感特征可能更加显著，患者情绪往往集中在焦虑、担忧、希望等几个特点情感。患者之间的交流可能更注重情感支持和共鸣，涉及到更多种类的情感体验，比如好奇、同情等。医生之间的交流通常更加专业和客观，情感因素可能较少。

（3）信息的病种主题与特征讨论

在线健康社区的信息内容具有相当高的研究价值，信息中暗藏着用户的需求与社区的特征，因此对于在线健康社区的信息内容进行分析与研究是一大热点，。而在以上选择的研究中，大部分研究者中往往会选择一个或多个特定的病种作为在线健康信息的主题特征进行数据采集，用于下一步分析。

图 5-2 话题词云图

由话题词云图 5-2 可知，中选择次数最多的前五种分别是：糖尿病、抑郁症、乳腺癌、心脏病、新冠肺炎。除了这些明确使用了病种信息的研究外，还有一些文献并未针对特定病种，比如 Naderi[14]等在研究在线健康社区中的问答信息之间的相似度时，随机抽取了相同病种主题和不同病种主题的问答信息。

以上文献中研究的病种以慢性病为主，慢性病的病程长、症状体征缓解与复发交替出现，一般无法痊愈，需要长期治疗和管理的疾病。基于此特点，在线健康社区对慢性病患者的帮助更加明显[58]，正如 Amann, J[59]在关于在线健康社区对慢性病患者的影响研究中提到的结论：在线健康社区是一项潜力巨大的以患者为中心的健康管理工具。因此慢性病在线健康社区的信息内容是当前研究者主要的研究对象。

但是新冠肺炎疫情的爆发给在线健康社区的研究方向带来了新的趋势变化，由于这类型的急性传染病容易出现规模性集中快速传染爆发的情况，为了更好地阻断传播途径与更合理地分配医疗资源，在线健康社区成为了有效的健康管理工具 [60]。综合学者对新冠肺炎有关的在线健康信息进行研究，我们发现在线健康社区的文本信息对研究急性传染病的预防、治疗以及预后都起到以下帮助作用：（1）急性传染病的爆发很容易引起患者恐慌，而使用文本分类模型可以将在线健康社区中关于某急性传染病的帖子或评论进行情感分析，判断用户的情感特征，如焦虑、宽慰、困惑等，有助于了解用户在传染病话题下的情感状态，从而能更好提供治疗建议。（2）对于一些新出现的急性传染病，其疾病症状、传播途径、预防措施等信息都尚不可知但又十分重要，而对患者在在线健康社区中产生的相关信息使用信息提取技术，可以快速从文本中自动抽取关键信息，这有助于将用户提供的信息转化为结构化的数据，加快对该病的后续分析。（3）通过对大量的在线健康信息进行 NLP 分析，可以发现传染病话题的趋势，如在特定时间段内关注度的变化，有助于及时识别可能的传染病爆发。

5.2 自然语言处理技术总结讨论

从上述文献中，我们总结了在线健康社区中主要应用的一些自然语言处理技术与文本挖掘技术。

1.命名实体识别（NER）：NER 技术用于从文本中识别和标记重要的实体，如疾病名称、药物医疗程序、医生姓名等。

2.文本挖掘和关系抽取：文本挖掘技术用于从文本中提取有关患者和医疗实体之间的关系，如患者-医生关系或疾病-治疗关系。这一技术通常与命名实体识别技术结合使用，有助于研究者对某特定病种或者药物进行系统性研究以及构建医学知识图谱。

3.主题建模： 利用主题建模技术，如 Latent Dirichlet Allocation（LDA），可以发现文本中的主题或话题。这有助于从在线健康社区中识别关于特定健康问题的讨论主题并观察趋势。

4.情感分析：情感分析是在线健康社区信息研究的一个热点，情感分析技术用于确定文本中的情感倾向，如正面、负面或中性情感。这可以帮助医生了解患者对特定医疗问题或治疗的情感反应以及变化，从而提供更好的医疗服务。

5.信息检索与推荐：信息检索技术使用关键词匹配和文本相似度排名可以用于构建搜索引擎，以帮助用户查找与他们的健康问题相关的信息。在此基础上结合用户特征与机器学习、深度学习模型可以实现在线健康信息的个性化推荐。

6.词嵌入和文本向量化：这是对文本的一种特征转换技术，使用词嵌入模型（如 Word2Vec、BERT 等）可以将文本数据转化为向量表示，以便深度学习模型进行分析和建模。

7.迁移学习：利用迁移学习技术，可以将在一个医疗领域上训练的文本分析模型迁移到另一个医疗领域，以减少对大规模标记文本数据的需求，降低模型训练成本。

5.3 挑战与局限性

当前利用自然语言处理与文本挖掘技术从在线健康社区中提取更多信息这一研究方向也存在许多挑战：

1.文本分析技术效果有待提高：当前文本分析技术的功能存在局限性。由于患者缺乏专业的医学知识，所以他们在表达时常常会存在语义模糊，特别是中文表达中还存在多义词、反问等特定表达，这让研究者很难准确对其进行主题、情感分类。此外还有网络特定的表达方法也给自然语言处理技术提高了处理难度，比如加入“表情包”改变句子原意，而 nlp 技术无法对其进行识别。

2.文本信息隐私问题：虽然从在线健康社区中获取的相关数据信息已经经过了相关的脱敏处理，但在实际信息挖掘时仍然存在隐私信息泄露的风险。例如，用户提到了他们所在的城市、医院或特定医生的名字，过多的信息挖掘与信息匹配可能会导致他们的身份被识别出来。

此外，本综述研究也存在以下局限性：

1.本研究采用的检索数据库较少，检索的数据样本不足，可能无法完全概括研究整体情况。

2.对选取的文献并未采用质量评估，相关研究可能会存在质量问题，影响结论的可靠性。

3.对于以上研究的结果与结论未做讨论评价。

6. 结论

本系统综述对自然语言处理技术在在线健康社区的应用情况进行了综合探讨，研究表明：

1. 目前自然语言处理与文本挖掘技术已经成为了文本信息分析的基本工具，这一工具的功能十分丰富，其在在线健康社区文本分析处理中有着广泛的应用前景，研究者可以根据他们的研究目的与研究方向选择不同的技术。

2. 在线健康社区三个维度（用户、信息、社区）都有着不同的发展特点，研究者应该根据这些特点选择合适的数据源、文本处理方法，从而挖掘出更多价值。

3. 从在线健康社区的文本信息中挖掘更多价值有两个基本趋势：①提高自然语言处理技术对健康信息的处理与分析效果；②结合社区发展特点发现更多文本信息的价值挖掘方向。

PO-033

基于生物信息学方法筛选原发性转移肾透明细胞癌关键基因及潜在治疗药物

莫状[1,3]，耿杰[1,3]，刘格良[1,2]，柴霞[2]，黄泰[1,2]，于琦[1,2*]，贺培凤[1]
1 山西省大数据重点研究实验室
2 山西医科大学管理学院
3 山西医科大学基础医学院

1.引言

肾脏肿瘤中较为多见的恶性肿瘤是原发性肾细胞癌，其直接致死率即使在所有泌尿系统恶性肿瘤中的排名也始终居于前位。据统计，在 2018 年全世界新出现了约 40 万例肾细胞癌，约占整个世界新发肿瘤患者数量总和的 2%，而且该病例数几乎每年都在增加。肾细胞癌有许多的构造亚种，每个亚种都各有其独特的分子景观。其中最为广泛常见的一个亚种即为肾透明细胞癌（clear cell renal cell carcinoma，ccRCC），约占到全部患者总数的 75%[错误!未找到引用源。]。ccRCC 来源于近端肾小管细胞，原发性 ccRCC 患者在术后有较大概率发生转移，且常规疗法对发生转移的患者治疗效果不佳，预后效果也较差，对患者的生命安全构成了极大的威胁。有文献报道，术后有20%~30%的病人出现了不同程度的肿瘤复发及远处转移[错误!未找到引用源。]，确诊患者五年生存率不足10%。ccRCC 病人在癌症初期并没有明显的病症，一些病人在 ccRCC 中后期的情况下才会被发觉，且随着病情不断发展，ccRCC 也伴随着转移，对于疾病的诊断和治疗又增加了难度。目前，ccRCC 的发病机制尚不明确，并且有研究发现 ccRCC 患者手术之后经常伴有局部复发和远端转移，造成预后不良**错误!未找到引用源。**。因此，为了能够早期识别 ccRCC 患者的转移趋势，探索

ccRCC 转移机制，进而寻找潜在的治疗药物，本文运用生物信息学的方法筛选出与转移性原发肾透明细胞癌有关的关键基因，并筛选及预测潜在的治疗药物，为原发性转移 ccRCC 的诊断和治疗提供新思路。

2.材料与方法

2.1 数据收集与处理

在基因表达综合数据库（Gene Expression Omnibus data base，GEO）输入检索词肾透明细胞癌（ renal clear cell carcinoma），并选择人类物种，逐一对检索结果进行查看，最后选择了原发性转移和未转移的肾透明细胞癌患者样本的微阵列基因表达数据集。下载原始数据并使用 R 语言的“Affy”和“Simpleaffy”包对数据集进行背景调整和归一化处理。

2.2 差异表达基因筛选及富集分析

接着使用 R 语言的 limma 包，对所获取的样本的基因芯片上的所有基因在转移性原发 ccRCC 和非转移性原发 ccRCC 之间构建矩阵和比对模型，平行比较基因在两类样本中的表达量。随后使用 lmFit 函数进行线性模型拟合，接着进行贝叶斯检验，经由 Bonferroni 校正，筛选出|log2（fold change）|＞0.5 且 adj.P.Val<0.05 的基因作为两组值的差异表达基因（differentially expressed genes，DEGs）。随后使用 R 语言 clusterprofiler 包对 DEGs 进行 Goterm 富集度统计学分析和京都基因与基因组百科全书通路富集分析（Kyoto Encyclopedia of Genes and Genomes，KEGG），其中 P<0.05 的通路被认为是显著富集的通路。

2.3 蛋白互作网络的构建与关键基因筛选

利用 STRING 在线工具对差异基因进行 PPI 网络的制作，将其中的与整个网络无关的作用节点剔除，最后生成蛋白互作网络。接着使用 cytoscape 软件的 cytoHubba 插件来查看其中得分位于前十的 Hub 基因，这些 Hub 基因就是蛋白互作网络的关键，可能与原发性肾透明细胞癌发生转移密切相关。

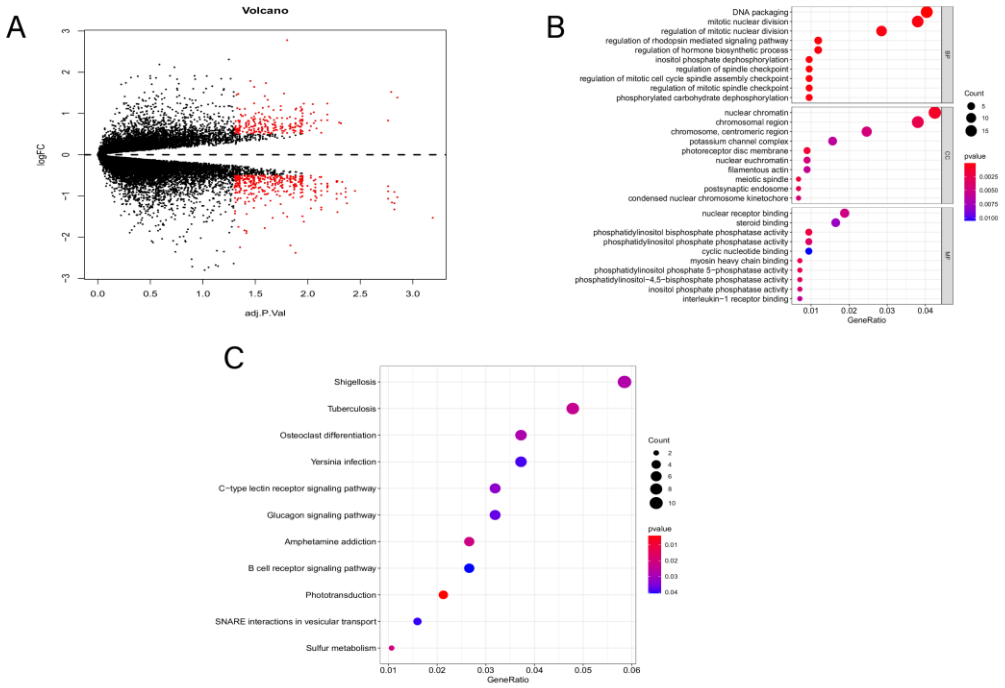
2.4 潜在治疗药物预测与筛选

基于转录组学鉴定出的 10 个关键 DGEs，其中上调基因 9 个，下调基因 1 个，使用 L1000 数据库检索并下载可能逆转基因表达方向的目标药物。使用 R 语言和 NCI-60 数据库进行药物敏感性分析，NCI-60 数据库涵盖 60 种不同种类的肿瘤细胞系及药物活性数据，进一步预测药理学层面的潜在治疗药物。

2.5 分子对接

从 Uniprot 和 RCSB PDB 数据库检索下载靶基因的蛋白质 3D 结构模型，使用 AlphaFold 数据库进行补充，选择方法为 X 晶体衍射法（X-RAY DIFFRACTION），分辨率< 2.50 A 的共晶体结构用于分子对接。从 PubChem 数据库下载 Ibrutinib 和 Dasatinib 的 3D 结构模型。

使用 AutoDockTools 软件受配体 3D 结构进行去水、加氢、计算电荷、选择扭转中心和扭转键（小分子）等预处理，并生成用于对接的 pdbqt 格式文件，随后使用 vina 程序进行分子对接计



算，导出对接结果。pymol 软件用于可视化受配体相互作用力。

3.结果

3.1 差异表达基因筛选

通过使用 R 语言的 limma 包进行 DEGs 的筛选，条件为|log2（fold change）|>0.5 且 adj.p.Value＜0.05，最终共得到 629 个 DEGs，其中上调基因 218 个，下调基因 411 个，并绘制出 DEGs 的火山图（图 1 A）。图中 logFC 为差异倍数，adj.p.Val 为校正后的 P 值，其中的每一个红色的小点就代表着一个差异基因，在 0 值以上的小红点为上调的差异基因，0 值以下的小红点为下调的差异基因。

图 1 A：原发性转移和未转移肾透明细胞癌的差异表达基因火山图。B：差异表达基因的部分 GO 富集通路图。C：差异表达基因的部分 KEGG 富集通路图。

3.2 GO 富集分析

使用 R 语言对筛选出的 DEGs 做 GO 功能富集分析，结果发现 DEGs 中有 266 个基因呈显著富集性（pValue＜0.05）。通过使用 R 语言做出部分 GO 功能富集气泡图（图 1 B）。通过气泡图中横轴为富集系数，纵轴为路径，不同颜色的 pvalue 值，由蓝向红代表 pValue 由大到小，富集程度逐渐增加。圆点的尺寸表示该路径上聚集的基因数目，圆点越大，数目越多，反之则相反。对于生物过程（BP），主要参与 DNA 甲基化、染色体凝聚和 G2 DNA 损伤检查点等生物学过程；对于细胞组分（CC），其主要表达在钾通道复合物、肌动蛋白丝、染色体区域等；对于分子功能（MF），其主要参与磷脂酰肌醇-4,5-二磷酸磷酸酶活性、肌球蛋白重链结合和白细胞介素-1 受体

结合等生物学过程。

3.3 KEGG 富集分析

用 R 语言对 DEGs 进行 KEGG 通路富集分析，结果发现 DEGs 中有 37 个基因呈显著富集性（pValue＜0.05）。通过富集气泡图（图 1 C）可以发现富集显著性由高到低依次为光转导、硫代谢、苯丙胺成瘾、结核、志贺氏菌病、破骨细胞分化、C 型凝集素受体信号通路、胰高血糖素信号通路、耶尔森菌感染、囊泡运输中的 SNARE 相互作用、B 细胞受体信号通路。

3.4 构建差异表达基因蛋白互作网络及筛选关键基因

将经过筛选得到的 629 个 DEGs 导入 STRING 在线工具制作出 DEGs 的 PPI network。从图中可以发现共有 152 个节点，171 条边，图中连线的粗细程度代表节点基因之间的联系强弱程度，在此列出 combined_score＞0.95 的节点。通过 Cytoscape 对不同蛋白的交互作用进行了研究，从中找出了主要的基因。利用 Hubba 插件，筛选出前十的 DEGs，这些基因即为 DEGs 蛋白互作网络的关键枢纽基因。

3.5 潜在治疗药物预测结果

L1000 数据库中共获取到每个基因的 100 个潜在治疗药物，其中 CDCA5 基因检索无结果；药物敏感性分析（图 2 A）结果显示，每个基因都存在正相关与负相关的调控药物。将上述结果取交集得到，Ibrutinib 可能是 MCM3 的潜在靶向药物；Dasatinib 可能是 CHEK2 的潜在靶向药物。

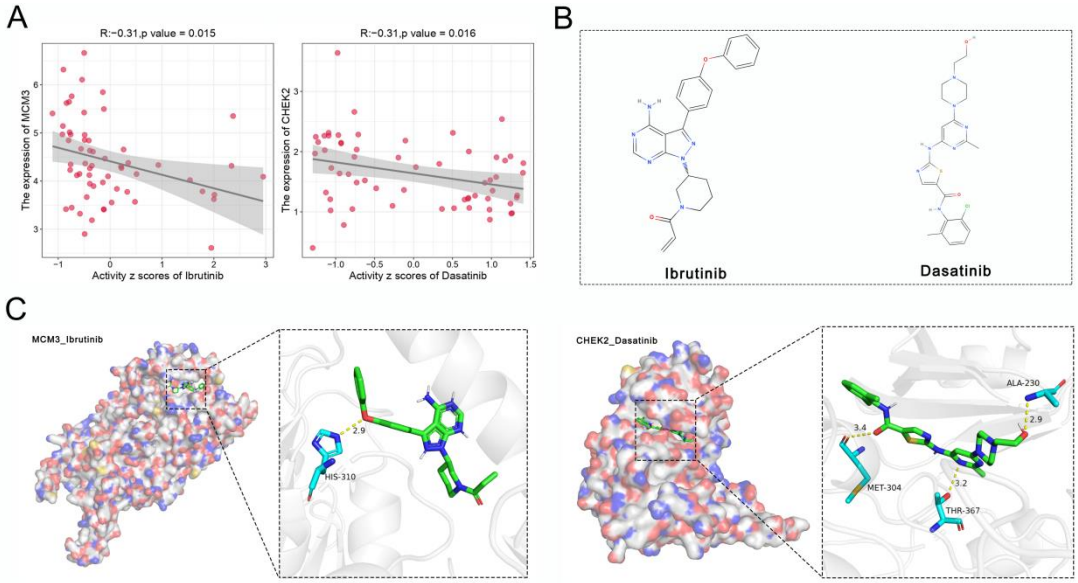


图 2 A: 药物敏感性分析：MCM3 和 CHEK2 基因表达与 Ibrutinib 和 Dasatiib 的药物活性 ZSCORES 的相关性分析。B: 小分子化物物 Ibrutinib 和 Dasatinib 的 2D 结构图。C：Pymol 可视化分子对接结果：小分子化合物 brutinib 和 Dasatinib 与 MCM3 和 CHEK2 蛋白的 3D 相互作用图

3.6 分子对接结果

目标蛋白质的 Uniprot ID 为：MCM3：P2520；CHEK2：O96017；小分子化合物 Ibrutinib 和

Dasatinib 的 Pubchem ID 分别为：24821094 和 3062316。对接结果显示（图 2 C），MCM3 蛋白与 Ibrutinib 的最低结合能为：-8.2 kcal/mol，共有 1 条氢键。CHEK2 蛋白中的与 Dasatinib 的最低结合能为：-9.7 kcal/mol，共有 3 条氢键。

4.讨论

原发性肾透明细胞癌发病相对隐蔽，大约 20~30%的病人在他们开始治疗时就已经有了扩散和转移。另外，30%的病人在接受肾脏手术后仍然有肿瘤的转移[错误!未找到引用源。]。为探究原发性 ccRCC 转移的机制，对转移性原发 ccRCC 尽早诊断并进行治疗。本文从 GEO 数据库中下载转移性和非转移性原发 ccRCC 样本数据，随后运用生物信息学方法进行分析，包括 DEGs 筛选，GO 功能富集分析和 KEGG 通路分析，蛋白互作网络的构建与分析以及药物预测与分子对接验证。最终筛选出的枢纽基因以及潜在治疗药物可能与原发性 ccRCC 转移及其治疗密切相关。

通过比较转移和未转移样本组织，结果共获得 629 个差异基因，上调有 218 个，下调有 411 个。GO 功能富集可以看出 DEGs 主要表达在钾通道复合物、肌动蛋白丝、染色体区域等，同时参与 DNA 甲基化、染色体凝聚、G2 DNA 损伤检查点、磷脂酰肌醇-4,5-二磷酸磷酸酶活性、肌球蛋白重链结合和白细胞介素-1 受体结合等生物学过程。DNA 甲基化的作用就是调控基因的表达。有学者对大量肾透明细胞癌患者样本数据进行分析，结果表明异常高甲基化与肾癌的分期和分级相关。Hildebrandt 等人的研究发现 hsa-miR-9 甲基化状态与 ccRCC 患者的癌症发展和转移复发有关**错误!未找到引用源。**。KEGG 通路富集显示这些 DEGs 富集在硫代谢、苯丙胺成瘾、结核、志贺氏菌病、破骨细胞分化、C 型凝集素受体信号通路、胰高血糖素信号通路、耶尔森菌感染、囊泡运输中的 SNARE 相互作用、B 细胞受体信号通路等。肿瘤发生的特征之一就是代谢紊乱，硫代谢过程对于 ccRCC 的发展和转移具有重要的影响作用。Tsai-Kun Wu 等人研究发现对甲酚硫酸盐（p-Cresyl sulfate ,pCS）通过涉及 miR-21/HIF-1α 信号通路的机制直接诱导 ccRCC 细胞的增殖和迁移[错误!未找到引用源。]。骨骼是晚期 ccRCC 最常见的转移扩散部位之一。已有研究证实癌症引起的溶骨性转移中的骨质流失是由于引发了癌症和骨驻留细胞之间的恶性循环，造成骨骼生成与退化的失衡。单核-巨噬细胞谱中的骨髓巨细胞，主要控制着骨骼的分解。成骨细胞负责骨形成。Chiara Spadazzi 等人研究发现在溶解性病变中，当肿瘤细胞到达骨骼时，它们会导致骨稳态失衡，导致骨吸收失控，从而导致可溶性因子过度释放，进而维持肿瘤细胞生长[错误!未找到引用源。]。接着对 DEGs 做出 PPI network，并将蛋白互作数据导入 Cytoscape 筛选出前十个关键基因。其中 Paul A. Stewart 等人的研究发现 MCM3 的表达在原发性肿瘤中显著增加，并且在骨转移中进一步上调，表明 MCM3 在前列腺癌进展和转移中起着关键作用[错误!未找到引用源。]。Chunhui Wang 等人的研究发现 CDT1 在前列腺癌细胞中的过度表达增强了细胞迁移、侵袭、肿瘤转移，并与细胞周期调节相关。CDT1 目前在肾细胞癌方面的作用仍需进一步探明，目前暂无研究结果**错误!未找到引用源。**。B Xu 等人通过对 26 对新鲜骨肉瘤组织及邻近正常组织的样本进行 RT-PCR 分析，发现 NDC80 在骨肉瘤组织中表达上调，并与肿瘤晚期和远处转移呈正相关[错误!未找到引用源。]。I-Wei Chang 等人检测 36 个上尿

路尿路上皮癌（upper urinary tract urothelial carcinoma，UTUC）和 30 个膀胱尿路上皮癌（urinary bladder urothelial carcinoma，UBUC）中的 CDCA5 转录水平，并且通过 H 分数评估的免疫组织化学来分别确定 295 个 UBUC 和 340 个 UTUC 中的 CDCA5 蛋白表达。最后得出结论 CDCA5 过表达与尿路上皮癌的晚期临床特征相关[错误!未找到引用源。]。Zhou Jiangqiao 等人通过病理学和分子分析评估 ASF1B 在细胞增殖和迁移中的功能，结果发现 ASF1B 依赖于 AKT/P70 S6K1 通路过表达促进肿瘤细胞增殖和迁移[错误!未找到引用源。]。Sijia Liu 等人通过对 DTL 在宫颈腺癌中的表达量测定，发现 DTL 的过表达在体外诱导了肿瘤细胞的迁移和侵袭，在体内促进了肺内转移。此外 DTL 的过表达通过调节 RAC1-JNK-FOXO1 轴增强了细胞运动并促进了宫颈腺癌的肿瘤转移**错误!未找到引用源。**。Narges Ansari 等人通过研究 CHEK2 在乳腺癌中的突变情况，发现 CHEK2 的突变导致细胞凋亡受损和癌细胞不受控制的增殖[错误!未找到引用源。]。Seung Ho Choi 等人分析了放射敏感性细胞和放射抗性细胞的基因表达谱，确定了一组可能参与肺癌细胞脑转移后放射抗性的基因。最后发现 TopBP1 和 Claspin 的表达对癌细胞的存活有积极的影响。Ze Gao 等人根据免疫浸润活性将 ccRCC 患者分为两组，并且使用使用单变量 Cox 回归分析，无监督聚类分析和主成分分析（PCA）等方法，最后发现 BRCA1 相关蛋白 1（BAP1）突变的患者预后较差[错误!未找到引用源。]。其中 JUN 在 ccRCC 中尚未有相关研究，其在 ccRCC 转移中的作用机制需进一步研究阐明，此外其余关键基因与癌细胞的增殖迁移密切相关。因此，本文认为上述关键基因可能与 ccRCC 患者发生转移有关，其可作为 ccRCC 发生转移时的标志物，帮助人们提早鉴定和发现 ccRCC 的转移情况。

MCM3 和 CHEK2 在原发性转移肾透明细胞癌中表达异常，在疾病的发生发展起到重要的调控作用。Ibrutinib 是一种有效的、不可逆的布鲁顿酪氨酸激酶（BTK）抑制剂，目前在临床上主要用于治疗套细胞淋巴瘤 （MCL）、慢性淋巴细胞白血病 （CLL）/小淋巴细胞淋巴瘤 （SLL）、瓦尔登斯特伦巨球蛋白血症 （WM）、边缘区淋巴瘤 （MZL） 和慢性移植物抗宿主病 （cGVHD）等疾病。有临床前研究证据表明，Ibrutinib 能够抑制 CLL 细胞存活和增殖，并影响 CLL 细胞迁移和归巢[错误!未找到引用源。]。但 Ibrutinib 在原发性转移肾透明细胞癌治疗相关方面研究较少，这也为临床治疗提供了新的思路和方向。Dasatinib 是一种酪氨酸激酶抑制剂，通过阻断许多酪氨酸激酶起作用。达沙替尼用于治疗费城染色体阳性的慢性粒细胞白血病患者和急性淋巴细胞白血病患者[错误!未找到引用源。]。有研究指出在前列腺癌的原位小鼠模型中，达沙替尼抑制 SRC 和 SFK 的活性，减少肿瘤生长和淋巴结转移的发展[错误!未找到引用源。]。同样地，达沙替尼还减少了肿瘤生长和胰腺腺癌细胞的播散[错误!未找到引用源。]。然而，与 Ibrutinib 相同，Dasatinib 在原发性转移肾透明细胞癌的相关研究较少，但这也为临床研究人员提供了一种新的方向，为原发性转移肾透明细胞癌的治疗与诊断提供了新的可能。综上所述，Ibrutinib 和 Dasatinib 具有抑制 MCM3 和 CHEK2 基因表达的潜力，可能成为治疗原发性转移肾透明细胞癌的潜在治疗药物。

PO-034

贵阳市健康医疗大数据治理平台建设实践

李 挺① 邵建华②

①贵阳市卫生健康信息中心

②贵阳市第一人民医院信息科

1.健康医疗大数据治理平台建设背景与需求

随着新医改的深入开展，医疗卫生信息化进入了快速发展期,要求夯实人口健康信息化和医疗健康大数据基础，推进互联互通信息标准落地应用，消除信息壁垒，畅通部门、区域、行业之间的数据共享通道，探索健康医疗大数据信息互联互通机制。国家卫健委在卫生信息化十三五规划中将“加强大数据质量体系建设，规范数据采集，保障数据质量，优化数据治理”。作为十三五期间各地全民健康平台的主要任务之一。

贵阳市卫生健康局在十三五期间建设了连通全市各级医疗机构的全民健康信息平台系统，实现了居民健康档案和电子病历数据的互联互通。信息平台建立之后，对于医疗大数据数据质量和更高要求应用的需求随之浮出水面。

为此贵阳市卫健局在 2019 年底启动了健康医疗大数据治理平台建设项目，目标是在现有区域卫生平台建设基础上，针对医疗大数据的标准、质量建立完整管理机制，以数据资源目录建设为核心的行业治理、居民服务、临床研究提供更加完备的健康医疗数据资源服务。

2.平台建设中的关键工程实现

医疗大数据平台提供的大数据服务为：①数据挖掘和建模。提供基于大数据架构下海量数据读取、数据处理、数据计算服务，通过可视化的数据探索工具、数据挖掘模型、简易模型训练支持数据挖掘与分析服务。②数据应用服务。支持快速数据集成、在线数据检索、多人协同等工具，提供大数据的检索、归并等应用服务。③数据治理。通过规范流程和规则库，基于流程引擎构建统一的、可配置的数据转换、清洗、比对、关联、融合等加工处理过程，对异构异源海量离散的数据资源加工生产，生成易于分析利用的、可共享的数据[2]。

贵阳市健康医疗大数据治理平台涉及的医疗机构包括贵阳市属的 27 个二、三级医疗机构和 156 家基层卫生机构。实现了对辖区内健康医疗核心数据的完整采集。经过初期摸底调研，系统需要处理的数据涵盖了部署在 35 个场地和 12 个业务系统，总数据量达到 15T，数据的时间跨度超过 20 年，总涉及患者人数超过 2 千万人。

为应对庞杂的异构数据汇集融合，大量历史数据的清洗转化工作，数据治理系统在设计初期明确了：大数据平台架构重点通过利用 Hadoop 架构建立分层数据处理机制以及利用数据库 ETL 工具建立数据融合机制的工程方案。

2.1 大数据平台的数据分层设计

医疗机构数据治理分为技术和管理两种视角：一是技术视角，将医院数据全生命周期各个环节的治理，包括数据采集、数据归集、数据共享、数据利用、数据开放、数据价值等；二是管理视角，指技术、组织、能力、法规等不同层面的治理[3]。

为实现对贵阳市跨医疗机构数据的融合清洗和质量控制，大数据平台需要为数据资源的生产建立覆盖多个环节供需的加工流程，而进入加工流程的数据需要借助大数据平台良好的可扩展性和分布式存储计算能力实现多层存储。

在贵阳市健康医疗大数据治理系统中，平台数据设计分成四个层次处理：

第一层：原始数据层：是指通过前置机提取后进入 Hadoop 存储的数据，这层数据保留了原生业务系统的完整数据内容和数据结构，作为数据治理系统的原材料层为后续加工提供数据来源。由于这层数据已经脱离原生关系型数据库存储，因此在数据的存储扩容和访问性能，以及数据备份冗余能力上得到极大增强。

第二层：ETL 层：是指通过脚本语句将原始数据层的数据抽取映射到平台标准的数据结构中，这层的映射转换屏蔽了异构系统带来的数据访问困难，为上层数据访问提供了最基础的原始数

据内容。

第三层：清洗层：将 ETL 层的原始数据内容根据一定规则完成清洗转换，尤其是对数据格式错误，基础数据字典转化的处理，经过清洗后的 ETL 数据屏蔽了各个医疗机构业务系统的数据标注差异，对上层访问来说更加友好易用，

第四层：数据资源层：将清洗后的数据按照一定业务主题重新调整数据结构，封装成为可访问可应用数据，在本项目中主要是以患者为核心维度的居民健康数据，以就诊为核心维度的诊疗监管数据等。

大数据治理平台的分层设计“以空间换时间”，支持了数据治理加工流程的过程可监控、数据可溯源、任务可调度，为信息部门全面掌控数据质量和数据资源情况提供了便捷通道。

2.2 跨医疗机构的数据汇集实现

十三五初期贵阳市在建立区域健康信息平台项目中，基于 WebService 接口开发连通了区域平台与各级医疗机构的信息系统，接口开发的数据集成方式一方面保证了数据同步的及时性，但同时也暴露出灵活性不足，开发调试成本高、数据质量难以保证等问题。

在数据治理系统建设中，面对大量医疗机构的业务系统数据融合需求，平台采用了：前置机全量数据备份采集，后台 ETL 数据抽取的数据采集机制。利用服务器的强大算力和大数据平台性能优势，减少业务系统的代码开发工作量，保证数据采集的完整性和准确性。

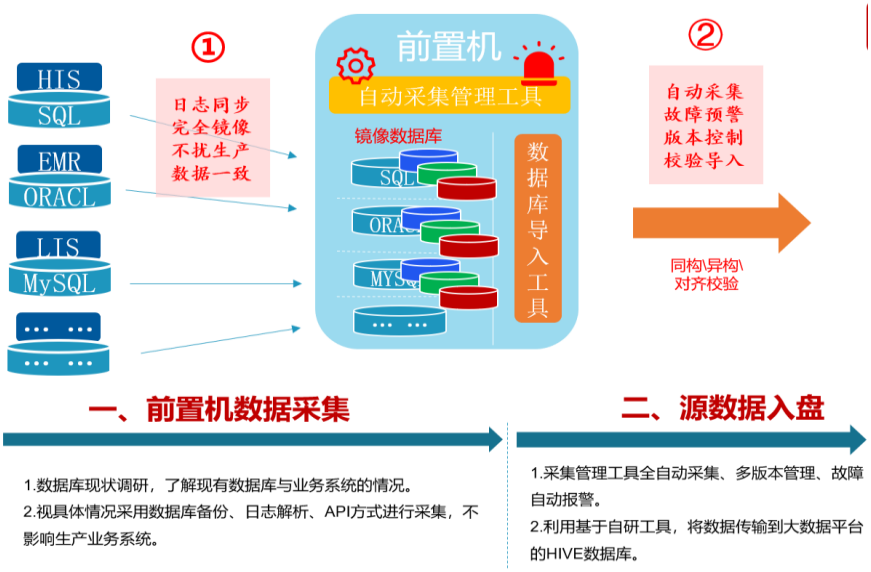


图 1 数据采集流程

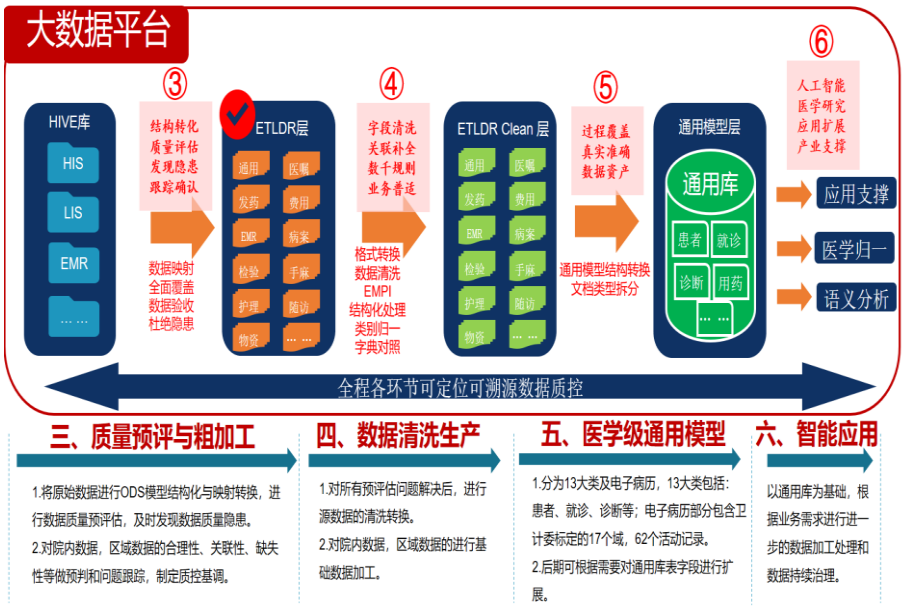


图 2 数据清洗治理流程

具体方法为：首先在各级医院和区县卫健局云 HIS 机房部署前置机，利用数据库备份同步机制在前置机上备份目标业务系统的数据库，通过前置机的数据验证的比对能力，定期抽取增量数据传输给市卫健局数据中心集群。数据中心集群采集到增量数据后自动进入 Hadoop 集群的原始库数据层，然后利用事先编制的 ETL 脚本从增量数据中抽取数据内容进入后期数据加工和质控，最终生成可用数据资源。

由于整个采集过程不需要业务系统做任何接口代码开发，克服了异构系统的数据转化困难、减少了对业务系统日常运行的扰动，在前台没有感知的情况下完成数据采集，极大地减少了项目实施风险。

2.3 数据资源与数据应用层建设

完成数据处理加工后的数据资源通过数据资源目录方式对卫健局使用部门提供数据共享服务，通常的政务数据资源目录为数据使用放提供数据导出、数据库访问和 API 接口访问等方式。贵阳市健康医疗大数据平台考虑到本地数据应用实际需求在上述访问方式基础上扩展了数据沙箱和数据指标集可视化访问方式。

图 3 数据沙箱管理

数据沙箱为有一定数据分析能力的用户提供了兼顾数据安全性和访问自由度的数据使用环境。 用户只要登录数据沙箱界面，就可以对数据资源层的总体数据分布情况做预览，并且通过高级查询条件设置圈定自己所关心的数据范围设置分析数据集，在将分析数据集倒入沙箱之后，服务器集群提供单独的数据存储和计算资源，支持用户通过高级脚本对数据做进一步分析处理。 数据沙箱功能对数据的访问、检索、倒入沙箱、分析和最终的数据销毁做全程授权管理和监控，实现了数据资源的安全访问。

针对业务管理部门的需求，数据治理平台之上搭建了针对行业监管指标的可视化分析层。通过事先对指标计算口径，取值范围，分析纬度的设置和预计算，形成可视化分析指标集。平台的指标集可视化访问功能经过数据质量处理后的数据资源以数据指标的形式呈现给管理部门，各级医疗机构的运行情况、医疗质量、处方管理等主题指标可视化展示，互动界面直接调整指标参数，支持管理部门对指标的深度观察使用。

图 4 监管指标可视化分析

3.健康医疗数据质量与数据资产

3.1 数据质量的分析和度量

参照国标《GB/T-36073-2018 数据管理能力成熟度评估模型》，在数据质量管理中设定明确的数据质量目标和数据质量评价分析体系是开展数据质量管理工作的首要步骤。

考虑到贵阳市健康医疗大数据的主要应用场景仍然是以患者为中心的疾病诊疗数据和以诊疗为中心的医疗机构运营数据，贵阳市健康医疗数据治理平台选在数据的完整性、时效性、一致性为主要质量评价分析纬度。其中，完整性规则是指基于常规医疗业务逻辑判断的完整性，比如病案首页有手术操作，则电子病历中应该有手术记录或有创诊疗操作记录等；时效性是指数据中时间逻辑的校验，比如手术开始时间应大于入院时间等；一致性是指跨系统的业务数据一致性校验，比如 HIS 住院费用明细合计金额应与病案首页的应收总金额保持一致等，其中一致性中的关联性指业务数据中患者、就诊、处方医嘱和收费记录之间的业务关联对照一致性，关联性对于通过分散记录的数据重建患者就诊，医疗机构诊疗活动路径至关重要。

经过对原始数据质控规则分析，可以发现不同机构之间的数据质量得分偏差比较大，尤其是在数据完整性方面，即便使用同一类型软件产品数据质量得分也很悬殊，说明医疗机构对业务数据记录的详尽程度与使用习惯有很大关系而与业务系统功能关系不大。

总体来说，在三个方面上完整性暴露的问题最大，其次是一致性和时效性。这为后续反馈各个医疗机构改进数据录入习惯，优化系统功能从而改善数据质量提供了管理切入点。

3.2 数据可用与数据增强

面对跨机构数据的融合问题，为了进一步提高数据可用性，数据治理平台在数据常规清洗处理之后加入了 EMPI 处理， 数据归一和数据结构化处理，形成增强层数据用于支持病历数据搜索等应用场景。

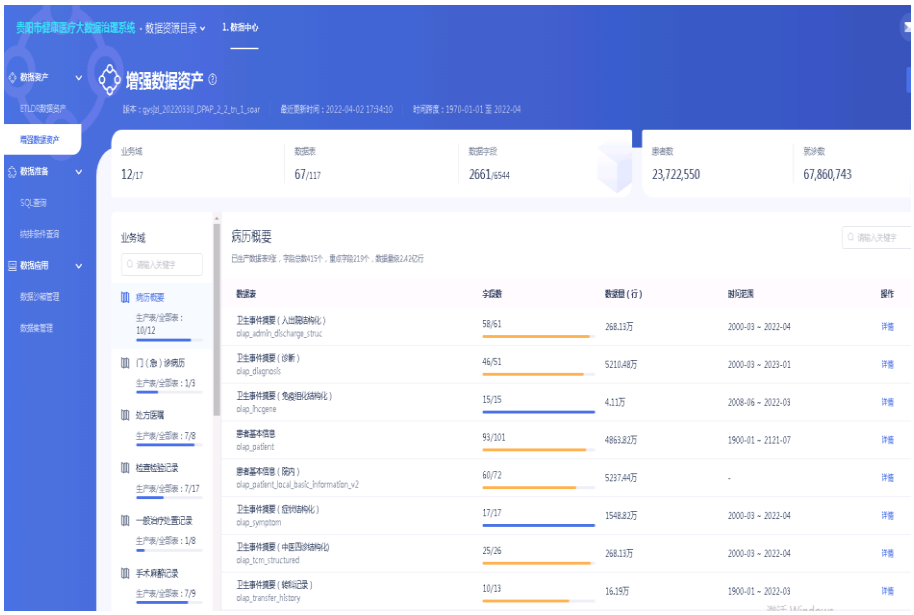


图 5 数据增强层数据管理

其中 EMPI 处理是指通过患者的姓名、出生年月、身份证号等信息，将来自不同医疗机构的患者数据做关联处理，实现患者跨机构就诊时间轴的重建。在 4409 万原始患者基本信息数据中，通过 EMPI 处理合并出 2352 万患者数据。考虑到数据中存在大量实名就诊制之前的历史数据，导致身份证字段空值率高达 68%，上述患者数据在未来一段时间内随着新就诊数据进入还有进一步去重关联处理的空间。

数据归一处理则是对诊断、药品名称、检查名称等数据做标准化处理；结构化处理是对电子病历中的文本数据做切词和关键信息提取转译处理，方便对诊疗过程做复杂搜索的时候精确定位病例数据，提高数据召回率。

3.3 数据资产管理与数据开放

面对数据治理平台建立后的多个数据层，通常数据应用是从数据资源层和数据集市层抽取数据以便上层应用充分利用数据治理成果。然而在实际业务中，数据分析人员对增强处理后的数据与原始数据的偏离缺少了解，在使用增强数据层数据的同时还会使用清洗前的数据做抽样对比，这时候完成数据映射处理解决了数据异构问题，而尚未进入数据增强处理的 ETL 层成为最理想的比对对象。因此对 ETL 数据层单独建立数据开放成为本项目的重要特点。

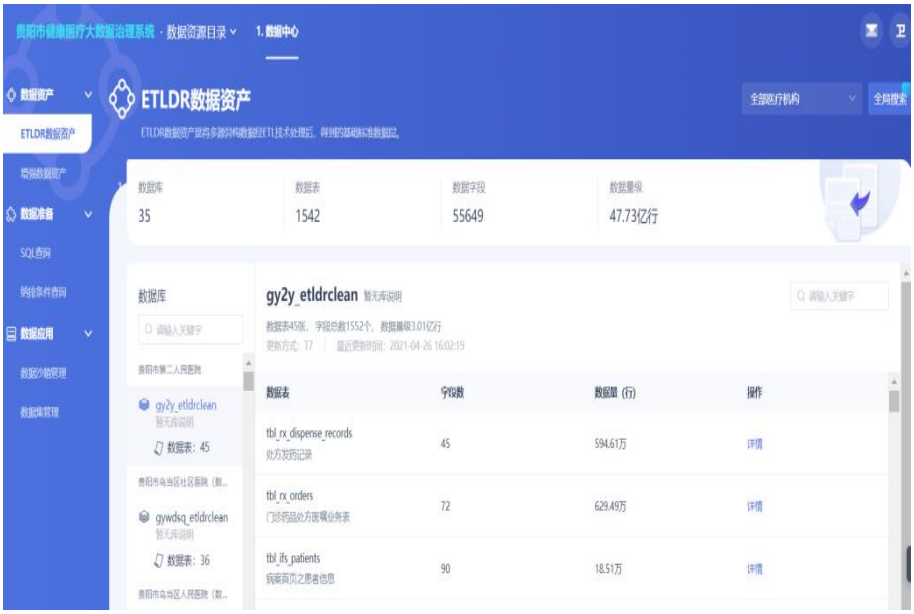


图 6 数据抽取层数据管理

数据治理平台中的数据沙箱功能提供了 SQL 脚本工具，这个模块基于 Presto 技术实现对存放在 HDFS 中的上亿条数据做秒级查询反馈，对于信息部门人员来说比较容易上手。因此在系统上线之后信息部门工程人员成为数据沙箱第一批用户，为卫健局相关部门的数据需求提供了快速响应支持。

4.思考与建议

经过两年多的系统设计、建设、运营和使用，贵阳市卫健局从十三五初期的信息化建设阶段跨入了十四五初期的数据资源建设阶段。2021 年底国家发布的《“十四五”大数据产业规划中》对健康医疗大数据的要求是“提升对医疗机构和医疗行为的监管能力，助推医疗、医保、医药联动改革。”未来几年内，包括贵阳市卫健局在内的各地卫生行政管理机构都会在健康医疗大数据的高质量数据治理方面面临更大的挑战，依托《数据管理能力评估模型》（DCMM）标准，建立强化数据分类分级管理，推动数据资源规划，打造分类科学、分级准确、管理有序的数据治理体系是必经之路。

PO-035

空间转录组技术发展及在肿瘤研究中的应用

王国梁[1,2], 渠鸿竹[1,2], 方向东[1,2]

1. 中国科学院北京基因组研究所（国家生物信息中心）

2. 中国科学院大学

背景

经典转录组学利用高通量测序技术对组织或一群细胞中所有RNA反转录而成cDNA文库进行测序，提供样品RNA含量的平均读数。该技术不能用于研究细胞异质性、疾病微环境、稀有细胞类型等[2]。单细胞RNA测序技术是在单细胞水平对转录组进行测序的一项新技术，可以研究单个细胞内的基因表达情况，同时解决用组织样本测序无法解决的细胞异质性难题，让解析单个细胞的行为、机制及其与机体的关系成为了现实[3]。目前最具有代表性的单细胞测序平台是 10X Genomics单细胞转录组测序平台，该平台利用微流控、油滴包裹和Barcode标记等技术来实现高通量的细胞捕获技术，能够一次性分离、并标记 500–10000 个单细胞，从而获得每个细胞的 3'端的转录组信息[4]。该技术主要用于细胞分型和标记因子的鉴定，从而实现对细胞群体的划分与细胞群体间基因表达差异的检测，此外该技术还可以预测细胞分化与研究发育轨迹，在当下疾病、免疫、肿瘤领域以及组织、器官、发育研究中发挥越来越重要的作用[5]。尽管单细胞转录组测序技术比经典转录组测序技术有所改善，但是进行单细胞测序之前需要将样本解离成单细胞悬液，会在很大程度上破坏样本的空间位置，同时在解离过程中会诱导应激基因的表达，引起细胞转录模式发生人为改变，造成转录偏好。这样获得的数据无法真实的反应样本的细胞转录状态，结果的可靠性大大降低[6]。空间转录组学技术通过在天然组织环境中测定细胞转录状态可以避免这些技术伪影[3]。在这里，我们总结了空间转录组的不同技术、分析方法、数据资源、生物学应用并对未来展望进行讨论。希望能对从事空间转录组研究的科研人员提供入门级的指导。

空间转录组技术

空间转录组学的关键在于检测组织样本中基因转录本数量。目前空间转录组学技术已经超过 50 种，可以分成基于二代测序（Next Generation Sequencing，NGS）技术和基于成像两类。

1. 基于NGS技术的方法

基于NGS技术，在测序前转录本被空间条形码原位捕获，测序完成后解析条码序列，映射出原始地址。2016 年，瑞典卡罗琳斯卡研究所发布了空间转录组技术（Spatial Transcriptomics，ST），该技术通过空间原位捕获mRNA，继而测序获得全转录组信息。ST检测法的每个单位点检测直径为 100 um，一共可以检测 1040 个位点[8]。2018 年底，10x Genomics公司收购该技术后进一步研发，升级后的 10x Genomics Visium的单个检测位点直径为 55 um,检测位点数量增至 4992 个。最近空间条形码的分辨率已经提高到了单细胞大小甚至是更高的水平。Slide-seq将冷冻组织切片转移到带有空间DNA条形码的磁珠表面上，随后对其进行测序并解析空间条形码序列[8]。升级版Slide-seqV2，对barcode珠的合成、阵列测序的流程以及cDNA的酶法处理三个方面进行了改进，显著提高了空间分辨率和转录本检测的灵敏度[9]。随后高分辨率空间转录组技术（HDST）也在同时期发布[10]。DBiT-seq技术是一种基于微流体条码标记的空间多组学测序技术，RNA逆转录过程是在细胞内进行的，而不是释放细胞进行，该技术的空间分辨率可以达到 10 um像素[10]。华大自主研发的空间增强分辨率组学测序技术Stereo-seq采用DNA纳米球(DNB)模式阵列，实现纳米级分辨率[11]。美国密歇根大学医学院发布的Seq-scope实现亚微米分辨率的空间转录组学，可将各种组织中细胞和亚细胞水平转录组学异质性可视化[12]。Nanostring GeoMX DSP技术采用独家专利技术在抗体或者RNA上偶联 DNA Oligo。每个 DNA Oligo对应一个靶标，当抗体或 RNA与样本切片的靶标结合后，利用激光切断两者的连接物，从而释放出DNA Oligo进行测序[13]。

2. 基于成像的方法

基于原位测序技术（In Situ Sequencing,ISS）的方法可直接读出组织内转录本的序列。具体来说，首先mRNA被逆转录为cDNA，通过循环扩增和原位测序检测。BaristaSeq是将特异性核苷酸序列作为条形码挂锁探针杂交传递到这些基因，其读取长度仅为 15 个碱基。STARmap使用条形码挂锁探针，与靶标杂交，通过添加针对挂锁探针旁边的位点的第二个引物，，不仅不需要进行逆转录，避免了转化效率问题，还降低了噪音[14]。前面所述的技术都是基于对靶标的先验知识，而荧光原位测序技术（FISSEQ）是基于非靶向原位测序，这种方法的优点是能够覆盖全转录本[15]。

基于原位杂交技术（In situ hybridization，ISH）的方法是利用荧光标记探针定位核酸分子。单分子水平FISH（smFISH）使用多条寡核苷酸探针精确定位mRNA在细胞中的空间位置。该技术具备较高的灵敏度和亚细胞的空间分辨率。但由于光谱颜色的固有限制，每次只能检测数十个RNA[16]。seqFISH在smFISH技术上进行了改进，使用寡核苷酸探针对序列进行多轮杂交，基因通量提升到了 250 左右，但是这种方法不仅昂贵而且非常耗时[17]。针对这个短板，庄小威团队开发了多重抗误差矫正荧光原位杂交技术（MERFISH），MERFISH技术使用二进制条形码识别细胞中的mRNA，以大规模复合方式标记和连续成像的方式来检测这些条形码，该技术可以鉴定单个细胞中数千种RNA的拷贝数[18, 19]。

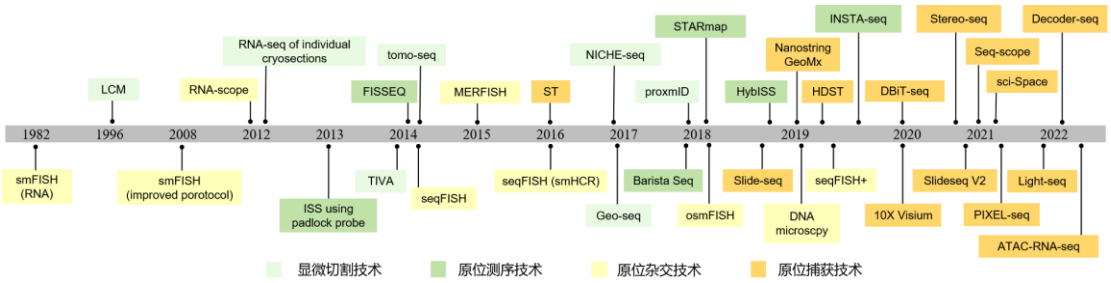


图 1：空间转录组技术发展

【图注：不同的颜色代表不同的时空转录组技术类型。sm FISH:单分子荧光原位杂交;LCM:激光捕获显微切割;ISS:原位测序技术;FISSEQ:荧光原位测序;TIVA:体内转录组分析;seqFISH:顺式荧光原位杂交;tomo-seq:RNA断层成像;MERFISH:多重荧光原位杂交技术;ST:空间转录组;Geo-seq:地理位置测序;Barista Seq:条形码原位靶向测序;STARmap:空间分辨转录扩增子读数映射;osmFISH:环状单分子荧光原位杂交;HybISS:基于杂交的原位测序;INSTA-seq:原位转录组可及性测序;DBiT-seq:基于微流体条形码标记的空间多组学测序;PIXEL-seq:聚合酶克隆索引文库测序技术】

空间转录组分析软件

随着空间转录组测序技术的发展和数据海量增长，空间转录组分析工具层出不穷，应用于不同测序方法和研究目的的分析流程也逐渐开发形成。目前标准化空间转录组数据分析主要分为预处理、空间变异基因分析、集成单细胞转录组解卷积细胞类型以及细胞互作等。预处理分析流程Space Ranger、STAR和zUMIs等可以基于原始测序数据进行上游分析产生表达矩阵；分析工具箱Scanpy、Seurat和Scater等能够从基因表达矩阵开始完成一系列复杂的下游分析任务，如聚类、差异基因分析、细胞类型注释等。针对空间转录组数据分析的不同需求，目前开发了很对针对于特定分析步骤的工具。例如ClusterMap、CoSTA、BayesSpace、FICT和SpatialCPie等工具可以根据基因表达和空间位置信息对空间聚类进行解析，帮助理解空间上的细胞群落和区域；SpatialDE、SPARK、SPARK-X和trendsceek等工具能够帮助阐明基因表达的空间模式；cell2location、RCTD、SPOTlight和Tangram等工具通过集成匹配的单细胞转录组数据进行去卷积分析以注释出单细胞精度的空间转录组数据；Cellchat、cellphoneDB、SpaTalk和iTALK等工具利用空间邻近原则挖掘细胞间的通讯信息，帮助理解细胞之间的相互作用和通讯网络。

表 1：空间转录组生物信息学分析工具

工具	语言	功能	URL	参考文献
ClusterMap	Python	空间聚类	https://github.com/wanglab-broad/ClusterMap	He et al. [20]

CoSTA	Python	空间聚类	https://github.com/rpmccordlab/CoSTA	Xu et al. [21]
BayesSpace	R	空间聚类	https://github.com/edward130603/BayesSpace	Edward et al. [22]
FICT	Python	空间聚类	https://github.com/haotianteng/FICT	Teng et al. [23]
SpatialCPie	R	空间聚类	https://github.com/jbergenstrahle/SpatialCPie	Joseph et al. [24]
SpatialDE	Python	SVG* 鉴定	https://github.com/Teichlab/SpatialDE	Svensson et al. [25]
SPARK	R	SVG 鉴定	https://zhoulab. github.io/SPARK	Sun et al. [26]
SPARK-X	R	SVG 鉴定	https://zhoulab. github.io/SPARK	Zhu et al. [27]
trendsceek	R	SVG 鉴定	https://github.com/edsgard/trendsceek	Edsgard et al. [28]
GPcounts	Python	SVG 鉴定	https://github.com/Manchester/Bioinference/Gpoounts	BinTayyash et al. [29]
SOMDE	Python	SVG 鉴定	https://github.com/Xuegonglab/somde	Hao et al. [30]
SPADE	Python/R	SVG 鉴定	https://githubcom/mexchy1000/spade	Bae et al. [31]
HMRF	Python	SVG 鉴定	https//bitbucket.org/qzhudfci/smfshhmrf-py/src/master/	Zhu et al. [32]
Meringue	R	SVG 鉴定	https://jefworks/MERINGUE/	Miller et al. [33]
stereoscope	Python	细胞注释	https://github.com/almaan/Stereoscope	Andersson et al. [34]
RCTD	R	细胞注释	https://github.com/dmcable/spacexr	Cable et al. [35]
SPOTlight	R	细胞注释	https://github.com/MarcElosua/SPOTlight	Elosua et al. [36]
Tangram	Python	细胞注释	https://github.com/broadinstitute/Tangram	Biancalani et al. [37]
DSTG	Python	细胞注释	https://github.com/Su-informatics-lab/DSTG	Song et al. [38]
Cell2location	Python	细胞注释	https://cell2location.readthedocs.io/en/latest/	Kleshche et al. [39]
spatialDWAS	R	细胞注释	https://giottosuite.com/	Dong et al. [40]
DestVI	Python	细胞注释	https://docs.scvi-tools.org/	Lopez et al. [41]
STdeconvolve	R	细胞注释	https://jef.works/STdeconvolve/	Miller et al. [42]
AdRoit	R	细胞注释	https://github.com/TaoYang-dev/AdRoit	Yang et al. [43]
Cellchat	R	细胞通讯	http://www.cellchat.org/	Jin et al. [44]
CellPhoneDB	Python	细胞通讯	https://github.com/Teichlab/cellphonedb	Efremova et al. [45]
SpaTalk	R	细胞通讯	https://github.com/ZJUFanLab/SpaTalk	Shao et al. [46]

*SVG，spatial variable gene，空间可变基因

空间转录组数据平台

空间转录组技术的快速发展，产生了海量的数据集，并在生物科学领域被广泛的利用。目前，空间转录组数据集成有 5 个代表性的数据库资源，分别是 SpatialD[47]、CROST（https://ngdc.cncb.ac.cn/crost/home）、STOmicsDB[48]、SPASCER[49]、Aquila[50]。SpatialDB是空间转录组领域的第一个数据资源库。该数据库于 2020 年发布，包含 8 种空间转录组技术生成的 24 个数据集，不包含下游分析并且持续至今无更新，数据量在同类数据库中最少[47]。CROST提供用户友好界面方便浏览、搜索、分析、可视化和下载空间转录组数据。它包含来自 8 个物种、35 种器官和 56 种疾病的 182 个空间转录组数据集，其中包括 1033 个子数据集和 48043 个肿瘤相关空间可变基因(SVG)。CROST搭建了标准化的空间转录组数据处理流程，集成单细胞RNA测序反卷积空间转录组数据，并评估相关性、共定位、细胞间通讯和生物功能注释分析。此外，CROST整合了转录组、表观基因组和全基因组来探索肿瘤相关SVG，并全面了解它们在癌症进展和预后中的作用。CROST还提供单样本基因集富集分析和SpatialAP两个在线工具，使用户不需要任何编程技能即可对空间转录组数据进行个性化分析和生物学功能注释。STOmicsDB共有五个模块，分别是资源中心模块、数据探索模块、定制数据库模块、在线分析模块和数据提交模块。资源中心模块集成了 21 种空间转录组技术生成的 141 个数据集，共涵盖 12 个物种。该模块还收集了空间转录组学相关的 5618 篇文献和 674 个工具。数据探索模块提供了空间变异基因的浏览和下载。定制数据库模块能够让其他研究人员参与到该数据库的建设之中。在线分析模块允许用户提供自己的单细胞转录组数据集结合数据库中的空间转录组数据进行整合分析。最后，数据提交模块提供了空间转录组数据的归档标准和系统，允许用户将他们自己的研究数据提交到STOmicsDB。该数据库仅提供与空间转录组学相关的文献和工具，并没有深入挖掘出版物、作者、工具、技术和数据集之间联系。同时该数据库与SpatialDB一样，提供的数据集较少且没有对这些数据集进行进一步的分析[48]。SPASCER包含来自 43 项研究的数据集，包括来自四个物种的 16 种组织类型的 1082 个样本，且包括与空间转录组数据配对的单细胞转录组数据。该数据库通过整合单细胞转录组数据，进行了丰富的下游分析，深入挖掘数据。包括scRNA-seq的细胞类型注释、空间模式基因分析、空间模式通路分析、基因调控网络分析、使用scRNA-seq进行细胞间相互作用分析、空间转录组学数据反卷积分析、空间转录组学的细胞间相互作用分析。此外，该数据库从DisGeNet数据库和DrugBank数据库中提取了与空间变异基因相关的疾病和药物信息。SPASCER的用户界面使用起来有些难度，没有提供在线分析工具，缺少与用户交互的功能，并缺少对已收集的数据集进行整合分析[49]。Aquila是一个空间多组学的数据库和分析平台，包括空间转录组学和空间蛋白质组学的各种技术产出的数据，目前集成来自 15 种技术的 107 个数据集。Aquila为用户提供多种形式的空间组学数据的可视化，例如细胞图谱，表达图谱和标记基因的共定位图谱。Aquila允许用户上传自己的数据集进行空间组学分析与可视化，同时该数据库提供了空间转录组数据的归档标准和系统，允许用户将他们自己的研究数据提交到Aquila进行存放。Aquila目前不支持组织图像[50]。

空间转录组在肿瘤领域的应用

肿瘤的空间结构与肿瘤的发生、进展和治疗反应密切相关，不同区域的组织学相似的肿瘤在分子上表现出差异，并具有不同的致瘤性，显示了肿瘤起始和空间位置的关联[51]。此外，空间相邻的细胞具有更强的相互作用，可以在功能上重塑肿瘤免疫微环境，促进肿瘤进展和进化[52]。这种异质性的细胞组成和相互作用机制，赋予了肿瘤细胞不同的侵袭、增殖和免疫抑制能力，因此，通过研究基因原位表达和免疫细胞与肿瘤细胞之间的相互作用，可以发现新的免疫治疗靶点，从而增强免疫系统对肿瘤的攻击能力[3]。目前，多项研究报道肿瘤空间异质性，如肝细胞癌、卵巢癌、结直肠癌[53-55]。Qi等人发现结直肠癌种FAP+成纤维细胞和SPP1+巨噬细胞在空间位置上紧密定位，并通过体外实验证明破坏FAP+成纤维细胞和SPP1+巨噬细胞间的相互作用能够显著提高免疫治疗效果[56]。Liu等人结合单细胞转录组和空间转录组测序技术发现了肝细胞癌中的肿瘤免疫屏障结构，并发现在小鼠模型中通过阻断SPP1 可增强抗PD-1 治疗肝细胞癌的疗效，同时减少肿瘤成纤维细胞的浸润并增加细胞毒性T细胞浸润[57]。总之，肿瘤细胞和其他细胞类型之间的空间共定位影响了肿瘤组织中潜在的细胞间相互作用和复杂的肿瘤免疫微环境。

未来展望

目前，空间转录组技术在表征空间异质性和肿瘤领域已显示出其强大功能。越来越多的研究表明空

间转录组技术在病理诊断中的能力。但是空间转录组仍然面临一些问题，比如基于测序技术空间分辨率不够、基于成像技术基因通量不够等问题正在迅速得到解决。空间转录组学的方法正在尝试对生物库中收集的石蜡包埋的组织进行兼容，希望能对这些数十年收集到的资源进行回顾性分析。通过未来的技术发展，有可能通过分析更大的组织区域来重建 3D水平的空间生物图谱。除了解决这些技术挑战外，还需开发新的算法工具以及具有创造性的分析思维来更好地挖掘空间转录组数据，提供潜在的生物学的见解。

PO-036

医院微博的对话沟通能力分析

铉婧雯 张韦
华中科技大学

1 引言

根据 WeAreSocial 维奥思社发布的《2023 年全球数字报告》显示，截至 2023 年 4 月，全球活跃的社交媒体用户数量达到了 48.0 亿[1]。中国互联网络信息中心（CNNIC）发布的第 52 次《中国互联网络发展状况统计报告》显示，截至 2023 年 6 月，我国网民规模达 10.79 亿人，较 2022 年 12 月增长 1109 万人，互联网普及率达 76.4%[2]。传统的医疗服务模式往往存在信息不对称和沟通不畅的问题，随着互联网和移动技术的发展，越来越多的医疗机构开始使用社交媒体进行交流和信息分享[3]。

研究表明，社交媒体良好的对话沟通能力可以促进医院更好地与患者进行互动和交流，塑造医院品牌形象，增加患者的信任度和满意度，有助于医院吸引访问者、品牌机构和患者等。1998 年，Kent 和 Taylor 等人[4]提出了对话沟通的 5 个原则：（1）对话循环原则（2）信息有用性原则（3）形成重复访问原则（4）用户友好界面原则（5）挽留用户原则。尽管现在很多医院都开设了社交媒体的官方账号，但是并不是所有医院的对话沟通能力都很强，实践上也存在对话沟通失败的例子。在社交媒体的对话沟通领域，一方面，现有文献主要强调医院社交媒体账号的传播效果[5]，但是传播效果跟对话沟通能力并不能完全等同，且两者的影响因素不一样。另一方面，相关研究主要集中在卫生行政机构，而对医院对话沟通能力的评估相对较少，如果忽视医院的对话沟通可能会导致信息传递障碍、患者误解和焦虑、医患关系不良、医院声誉受损等。

因此，本文以 33 家医院微博为研究对象，以对话沟通理论为基础理论框架，采用内容分析法对医院微博的对话沟通现状进行定量分析，研究提出以下问题：

问题 1：对话沟通原则在医院微博的使用情况如何？

问题 2：医院可以实施哪些实践来吸引公众参与对话沟通？

2 方法

2.1 数据收集

为了说明上述问题，本研究以新浪微博作为研究平台，时间跨度为 2023 年 3 月 27 日至 4 月 26 日，数据收集期为近 1 个月[6]，选取医院作为分析单位，借助复旦大学医院管理研究所发布的 2021 年度全国综合医院排行榜，对全国前 100 所医院微博进行筛查。通过新浪微博账户查找功能纳入医院官方认证微博，将目前不活跃的（即在爬取时间范围内无任何动态）账号筛除得到 33 家医院微博作为研究对象，抓取内容包括：主页信息内容（微博名称、关注数、粉丝数）、文本内容、发布日期、点赞数、转发数、评论数等，爬取推文共 1027 条，将收集到的数据进行编码并评估。

2.2 对话原则操作化与测量

由于微博和网站设计不同，我国微博账号的进入页面并没有很大差别，所以“用户友好界面”这一原则不纳入测量。如表 1 所示，基于 17 个编码规则，用“1”表示满足条件，用“0”表示不满足条件，医院微博对话沟通能力的计算方式为所有项目的分数总和与项目总数（17）之比，并将其乘以 100%，以百分数记值[7]。

表 1 编码规则

对话沟通能力	测量指标
对话循环	1.回复评论
	2.发起投票、问卷
	3.使用提及功能（@）
	4.未关闭评论
	5.回答问题
	6.提出问题
信息有用性	7.健康科普、疾病介绍
	8.医院新闻/活动
	9.声明医院的理念，使命和目标
	10.联系方式（电子邮件/电话/地址）
形成重复访问	11.使用主题标签（#）
	12.事件日历或指向包含此类日历的网页的链接
	13.邀请返回（关注后续报道、请求关注与转发）
	14.每日更新
挽留用户	15.链接到医院官网
	16.链接到医院其他社交媒体平台
	17.信息相关链接/图片/视频

3 研究结果

表 2 为 33 家医院的全国排名、地域分布、床位数、关注数、粉丝数、在观察期内的发帖数量的总体信息统计。医院地域分布在北京、上海、四川、重庆、湖北、湖南、江苏等地。其中，北京地区医院个数最多为 11 家，其次是四川、上海各有 4 家。在关注粉丝数方面，北京协和医院的粉丝数最多为 130.7 万，南方医科大学珠江医院微博粉丝数最少仅 205，北京大学肿瘤医院关注数最多为 1330，南方医科大学珠江医院关注数最少为 23。在观察期内中南大学湘雅三医院发文数量最多为 132 条。

表 2 医院微博总体情况

微博名称	排名	地域分布	床位数	关注数	粉丝数量	博文数量
北京协和医院	1	北京	2000 多	1012	130.7 万	42
四川大学华西医院	2	四川	4300	812	65 万	43
瑞金医院	4	上海	2139	165	14.7 万	22
复旦大学附属中山医院	5	上海	2005	741	16.7 万	65
北京大学第三医院	14	北京	2332	91	17 万	41
中南大学湘雅医院	15	湖南	3500	1217	68.2 万	22
北京大学人民医院	21	北京	NA	107	34.5 万	42
江苏省人民医院南医大一附院微博	22	江苏	4200	274	13.3 万	16
中国医学科学院肿瘤医院微博	25	北京	NA	307	22.1 万	6
我是天坛人	26	北京	1650	97	34.8 万	54
复旦大学附属肿瘤医院	28	上海	2136	111	10.7 万	24
中国医学科学院阜外医院	30	北京	1291	216	16.7 万	3
首都医科大学宣武医院	37	北京	NA	217	15.4 万	20
四川省人民医院官方	38	四川	4300	1153	94 万	4
武汉大学人民医院	40	湖北	5238	408	35.6 万	15
四川大学华西口腔医院微博	44	四川	NA	85	6.8 万	2
重庆医科大学附属第一医院	48	重庆	3200	68	4193	23
邵逸夫医院	49	浙江	3300	155	15.5 万	47
山东省立医院	52	山东	3889	66	3874	14
天津医科大学总医院	59	天津	2468	56	3966	12
重庆医科大学附属儿童医院	61	重庆	2480	87	9134	18
北京大学肿瘤医院	62	北京	801	1330	92.9 万	8
西安交通大学第一附属医院	63	陕西	NA	135	6.3 万	2
北京朝阳医院官微	68	北京	2500	117	16.4 万	32
四川大学华西第二医院	70	四川	1580	290	16.6 万	7
北京友谊医院	72	北京	2306	464	57.6 万	70
中南大学湘雅三医院	73	湖南	2200	299	7836	132
北京积水潭医院	76	北京	1800	38	21.2 万	124
上海市胸科医院	80	上海	1000 多	65	3444	12
河南省人民医院	83	河南	NA	313	13.4 万	81
复旦大学附属妇产科医院	85	上海	820	114	98832	2
南方医科大学珠江医院微博	87	广东	3500	23	205	9
浙江省妇女保健院	91	浙江	1520	689	17.1 万	13

注：1.数据来源：床位数来源于各个医院官网

2.数据统计截至 2023 年 9 月 12 日

问题 1：对话沟通原则在医院微博的使用情况如何？

如表 3 所示，在观察期内，医院微博总体的对话沟通水平得分为 49%（SD=0.37），这表明医院微博均在一定程度上采用了对话原则，其中对话循环为 39%（SD=0.33），信息有用性为 63%（SD=0.41），形成重复访问 36%（SD=0.37），挽留用户为 68%（SD=0.43）。

在四项原则中，对话循环原则的评分较低（39%），并不是所有的研究医院允许用户进行评论，只有 12%的医院能够对用户回复，至于发起投票以及问卷、回答问题、提出问题则相对较少为 3%。在分析信息有用性原则时发现，90%的医院在发布信息时都包括健康和疾病的科普，在分析微博页面，有一半以上（57%）的页面标注了联系方式。在形成重复访问原则中，医院微博的每日更新略低，得分为 3%，相比而言，邀请返回也较少为 21%。挽留用户原则中，90%的医院微博能够在帖子中加入信息相关链接/图片/视频，一半以上的医院微博有链接到医院官网以及链接到医院其他社交媒体平台。

总体来看，对话沟通原则中的“信息有用性”“挽留用户”原则在微博样本中的整体适用情况较好，但在微博内容方面，集中于科普类信息，特别是疾病和健康类信息，且多使用图文结合的形式吸引受众停留。另一方面，医院微博的对话沟通原则的使用出现传播单向度情况，“对话循环”“形成重复访问”原则的整体适用程度较差，主要问题体现在微博主体对于微博的基础功能和扩展功能的开发利用程度不够，例如在医院回应、医院发起投票问卷等具有互动性的功能虽然适合建立和公众的对话沟通关系，但是使用频次不高。

表 3 医院微博的对话沟通水平

医院微博内容	M(%)	SD
对话沟通能力	49	0.37
对话循环	39	0.33
1.回复评论	12	0.33
2.发起投票、问卷	3	0.17
3.使用提及功能（@）	63	0.48
4.未关闭评论	75	0.43
5.回答问题	3	0.17
6.提出问题	78	0.41
信息有用性	63	0.41
7.健康科普、疾病介绍	90	0.29
8.医院新闻/活动	78	0.41
9.声明医院的理念、使命和目标	30	0.46
10.联系方式（电子邮件/电话/地址）	57	0.50
形成重复访问	36	0.37
11.每日更新	3	0.17
12.使用主题标签（#）	75	0.43
13.事件日历或指向包含此类日历的网页的链接	45	0.50
14.邀请返回（关注后续报道、请求关注与转发）	21	0.41
挽留用户	68	0.43
15.链接到医院官网	57	0.50
16.链接到医院其他社交媒体平台	57	0.50
17.信息相关链接/图片/视频	90	0.29

问题 2：医院可以实施哪些实践来吸引公众参与对话沟通？

1）重视对话循环和形成重复访问

对话循环原则在对话沟通中有重要地位，不仅可以解决患者询问的问题，而且还能拉近医院与患者的距离。医院应该对公众的评论及时回复，畅通沟通渠道，如设置电话、邮箱等联系方式，并且对于公众提出的问题予以公示答复。医院还应善于利用微博的优势，开展私信互动和在线投票增加公众的参与感。除此之外，为使用户形成重复访问，医院微博应尽量频繁更新高质量推文，邀请用户关注后续报道。

2）发掘关注对象的需求，精心发布内容

赵加奎[8]研究结果显示, 经常使用智能手机的人群希望通过手机获取的健康知识的形式是图片(68.8%)、文字(66.7%)和视频(40.0%),说明受众接受多样性的健康传播形式，在发文时注重多媒体形式结合。在信息有用性方面，微博不能单一地、呆板地宣传医院的科室、技术等，而是注重对用户的有用性。在微博中运用“@”和“#”符号，使微博内容被辐射式传播，扩大作品影响力。

3）完善管理制度，提高微博使用率

在医院微博的应用中，不仅需要强化领导的使用意识和以身作则，还需要加强医院微博运营人

员的培训，提升他们的专业素养和运营能力。此外，医院还应加大对微博平台运营的投入，包括技术设备、人力资源和推广宣传方面的支持。

4 讨论

表 2 表明不是所有的医院都开通了微博或者微博状态活跃，比如排名前 10 的医院有 5 家没有开通微博如中国人民解放军总医院等医院均没有开通微博，这可能是因为不同医院分配在社交媒体平台的注意力不同，有的医院可能抖音是其主要的对外宣传平台，比如华中科技大学附属协和医院。根据表 2 所示，中南大学湘雅三医院在观察期内的发文量最多为 132，但是对话沟通水平得分为 0.52，对话沟通能力并不高。而在观察期内发文量仅为 7 篇的四川大学华西第二医院，对话水平却较高为 0.7，表明发文量与对话沟通水平并不是正向关系，还要确保发文的质量。通过对各个医院微博的对话沟通能力进行分析对比，发现北京协和医院微博的对话沟通能力水平较高为 0.82，可能的原因在于该医院地理位置优越，处于北京市东城区，人力、物力资源较为丰富。有研究[9]结果显示博物馆在“对话循环”较弱，这与本研究的结果相似，社交媒体的使用普遍存在对话循环能力较弱的问题，因此应该对该问题加以改善，提升社交媒体对话循环的水平。

纳入分析的 33 家医院微博其对话沟通能力水平的均值为 0.49，较低的原因在于对话循环能力和使用户形成重复访问能力较差，比如说医院微博几乎不回复评论以及部分医院存在不允许评论的情况。存在对话沟通水平较差的原因可能在于：地域、重视程度、开通微博的时间、粉丝数等方面。根据使用与满足理论，分析受众对媒介的使用动机和获得的需求满足，了解大众传播给公众带来的心理和行为上的效用，从而利用在医院社交媒体的使用上，提高医院的对话沟通能力。

PO-037

基于扎根理论的在线问诊服务质量研究：服务提供方的视角

刘菁菁[1]，张韦[1]

[1]华中科技大学同济医学院医药卫生管理学院

1 研究背景

近年来，我国居民对于医疗健康的需求与日俱增，在线问诊平台借助互联网平台为公众提供了一个便捷的就医渠道。2020 年，《关于在疫情防控中做好互联网诊疗咨询服务工作的通知》[1]中提出要充分发挥互联网医疗服务的优势，进一步完善“互联网＋医疗健康”的服务功能，提升在线问诊服务质量。为解决在线问诊服务质量参差不齐的问题，进一步发挥互联网医疗方便公众就医的作用，早在 2018 年，国务院办公厅研究制定《关于促进“互联网＋医疗健康”发展的意见》[2]，提出要健全“互联网＋医疗健康”的服务体系，加强行业监管与安全保障，《互联网诊疗监管细则》[3]也进一步明确了线上诊疗的业务监管、医疗机构监管、质量监管、责任监管等要求，为提升在线问诊服务质量提供政策支撑。

在线问诊服务依托互联网平台开展，远程问诊不能触诊查体、医疗硬件设施需求小、医患关系绑定性弱，显然线下医疗服务质量的评价指标不适用于在线问诊平台，应从线上角度出发，探寻在线医疗的服务质量的评价。医护人员作为在线问诊服务的提供方，直接影响患者的线上问诊服务体验。鉴于此，本文将研究视角转向服务提供方，通过半结构化访谈以及扎根编码的方法，明晰医护人员对于在线问诊服务的看法，总结出医护人员视角下在线问诊服务质量的构成及影响因素，为在线问诊服务质量提升提出策略。

2 研究方法

2.1 访谈对象

由于质性研究更注重样本的丰富性和代表性，本次访谈尽可能地考虑到医护人员在职业、年龄、科室、省份等方面的异质性，选取了湖北省、山西省、湖南省、河南省、江西省 5 个不同省份

医院的 25 名医护人员。受访者具体信息如表 1 所示。

表 1 受访者基本情况

序号	性别	所属医院等级	所属科室	职位	年龄	工作年限	是否开通	有无在线问诊经历
1	男	二甲	呼吸内科	医生	40 岁	10 年	是	有
2	男	三甲	肝胆外科	医生	36 岁	8 年	是	有
3	男	三甲	门诊科	实习医生	24 岁	1 年	是	无
4	男	三甲	党办	党办主任	未告知	30 年	是	有
5	女	三甲	儿科	护士	27 岁	6 年	是	无
6	男	三甲	肛肠科	医生	53 岁	20 余年	是	有
7	男	二甲	中西医结合	医生	27 岁	3 年	否	无
8	女	三甲	未告知	医生	38 岁	10 年	是	有
9	男	三甲	急诊科	医生	未告知	20 余年	是	有
10	男	未告知	心内科	医生	42 岁	15 年	是	有
11	男	三甲	胃肠外科	医生	43 岁	20 年	是	有
12	女	三甲	儿科	医生	未告知	6 年	是	有
13	男	三甲	门诊办公室	门办主任	未告知	未告知	是	有
14	男	三甲	呼吸内科	医生	未告知	13 年	是	有
15	女	三甲	儿科	医生	未告知	34 年	是	有
16	女	三甲	妇产科	医生	未告知	15 年	是	未告知
17	女	三甲	未告知	护士	28 岁	5 年	是	无
18	女	三甲	未告知	医生	33 岁	7 年	是	有
19	男	三甲	麻醉科	医生	47 岁	15 年	是	有
20	男	三甲	外科	医生	50 岁	25 年	是	有
21	女	三甲	心内科	医生	48 岁	25 年	是	有
22	男	三甲	未告知	护士	30 岁	7 年	是	有
23	男	三甲	未告知	医生	42 岁	17 年	是	有
24	男	三甲	未告知	医生	32 岁	5 年	是	无
25	男	三甲	小儿内分泌	实习医生	24 岁	1 年	是	无

2.2 访谈提纲

本研究运用半结构化方法，根据访谈对象、访谈目的制定访谈大纲，具体访谈问题见表 2。通过整理医护人员对于在线问诊服务的评价的原始资料，分析医护人员视角下在线问诊服务平台存在的问题，得出医护人员视角下在线问诊服务质量的构成及其影响因素。

表 2· 访谈大纲

序号	问题
1	访谈对象基本信息
2	您所在医院在线问诊服务开通情况怎么样？
3	您是否提供过在线问诊服务，对其认知有何变化？
4	作为医护人员，您对医院在线问诊服务质量的评价怎么样？
5	患者对医院在线问诊服务的评价怎么样？
6	您认为在线问诊平台存在哪些问题及改进措施？

2.3 分析方法

本文基于扎根理论，借助编码软件 Nvivo20 质性分析软件，对原始资料依次进行开放性编码、主轴编码、选择性编码。同时，为检验理论模型是否饱和，余下 5 份访谈材料进行饱和度检验。

3 研究结果

3.1 开放式编码

根据医护人员对在线问诊服务的评价，从访谈资料中提取出 36 条相关语句，并提炼出 36 条概念，并将概念范畴化，形成 16 个范畴，部分开放式编码见表 3。

根据医护人员认为在线问诊服务存在的问题以及医患双方的评价，从访谈资料中提取出 26 条相关语句，并提炼出 26 条概念，并将其范畴化，形成 15 个范畴，部分开放式编码见表 4。

表 3 在线问诊服务质量构成开放式编码（部分）

范畴	概念提炼	原始资料示例
诊疗统一性	线上线下绑定，患者开药方便	现在云医也都和医生绑定了，可以直接通过这上面在医院药房开药，然后邮寄给患者，免去了患者去药房自己开药的麻烦和风险。
诊疗连续性	基层药物不足	从县里的医院回来咨询，但是我们开了药发现县里没有这种药，还是需要从市里调药。
平台功能	不能解决复杂病情	对于比较复杂的问题，比如腹痛、头痛这些可能很多疾病可以引起的症状，或者非常紧急的情况，在线问诊就很难符合要求。
平台监管	监管难，利用流量卖药	二来，监管难也成为阻碍其发展的一大痛点，第三方互联网医院平台“电商化”特征显著，利用流量卖药变现成为常见模式，互联网医院成为互联网医药。
病史掌握程度	不能获得完整病情信息	单纯靠在线问诊，难以保证医护人员全面准确获悉患者的病情信息。
医务人员资质	具有相应的职业资质要求	应当依法取得相应执业资质，具有 3 年以上独立临床工作经验，并经其执业注册的医疗机构同意。

表 4 在线问诊服务质量影响因素开放式编码（部分）

范畴化	概念提炼	原始资料示例
医生服务态度	医生回复时效慢	患者不满意的因素主要是部分在线门诊回复慢，患者等候时间长。
医生沟通能力	医生沟通能力差	个别医护人员与患者的沟通能力较差、水平不高、服务态度生硬，解释不到位。
患者患病类型	慢病患者评价较好	有很多慢病患者和复诊的患者对在线就诊服务的评价挺高的，给他们带来了极大的便利。
患者沟通能力	患者沟通意识弱	患者没有沟通意识，一直在自顾自地描述自己患病感受，而对医生的问题置之不理。
平台技术	画面卡顿	视频问诊画面信息传输较慢，会出现卡顿。
平台局限	无法健康查体	医生看不到病人，做不了健康查体，不能全方面地了解到病人的情况，基本上只能通过文字才知道病人现在是什么状态。
行业监管	缺乏规范和监管	我觉得是没有一个专门的监管部门，医生可以随意地想开诊就开诊，想停诊就可以停诊，没有一个规范。
质量标准	缺乏统一质量标准	互联网医疗无法实现线下医疗的服务质量。医疗健康类的服务是人对人的服务，很难标准化。

3.2 主轴式编码

主轴编码是将开放式编码中的概念进行范畴化，并挖掘范畴之间的类属关系，并借鉴 Donabedian 教授的医疗质量三维度[4]理论，将在线问诊服务质量构成归纳为结构质量、过程质量以及结果质量。具体编码内容如表 5 所示。同理，将在线问诊服务质量的影响因素归纳为医生、患者、平台以及政策四个维度。具体编码内容如表 6 所示。

表 5 在线问诊服务质量构成主轴式编码

主范畴	范畴化	概念范畴
过程质量（A01）	医患沟通效率（B01）	沟通效率低；医生交互感受差，病人无法准确描述；仅通过文字描述，可能增加误诊风险
	病史掌握程度（B02）	不能获得完整病情信息；病史未互通；患病情况了解度不够
	医生接诊效率（B03）	医生接诊方式和时间灵活；节约就诊时间
	医疗费用（B04）	节约就医成本，方便老百姓，解决基础医疗。
	医疗资源（B05）	分流患者，优化医疗资源，减轻医院压力。
结构质量（A02）	基础检查能力（B06）	看不到病人；无法触诊；做不到视触叩听；不能体格检查
	平台监管能力（B07）	监管难，利用流量卖药
	平台技术水平（B08）	网络卡顿
	医务人员资质（B09）	具有相应的职业资质要求
	医生专业程度（B10）	官方公众号，医生专业；微信小程序，医生对患者熟悉；服务不满意可退款
	平台服务量（B11）	年龄大的老教授需要辅助；年龄高的患者很少使用在线医疗；年龄大，不会使用在线功能
	平台功能（B12）	不能解决复杂病情；不能完全满足患者需求；不能开处方；分流作用不明显，医保尚未打通
结果质量（A03）	诊疗统一性（B13）	方便患者看病寻药；线上线下绑定，患者开药方便；复诊开药方便
	诊疗连续性（B14）	术后恢复、慢病保健便利；线上开药、线下取药的一致
	诊断能力（B15）	仅作参考，不能完全下诊断结论
	接诊能力（B16）	及时性，跨越时间和空间；解决基础医疗，方便老百姓

表 6 在线问诊服务质量影响因素主轴式编码

主范畴	范畴化	概念提炼
医生（A04）	服务态度（B17）	接诊回应速度慢；交流回馈不及时；医务人员积极性不高。
	工作时间（B18）	医生线上问诊时间紧张；医生排班紧张，工作繁忙
	沟通能力（B19）	医生沟通能力差；医患之间缺乏真实情感交流
	专业程度（B20）	医生熟悉病情且可靠；看重医生专业水平
患者（A05）	患病类型（B21）	慢病患者评价较好；复诊续方便利；不复杂疾病的效果好
	表达能力（B22）	患者描述不清；患者表达能力弱
	沟通能力（B23）	患者沟通意识弱；患者难以提供有效信息
	就医成本（B24）	解决交通不便和路费问题；时间成本低；问诊时间灵活
平台（A06）	平台机制（B25）	无反馈机制；隐私泄露问题；在线问诊小程序功能单一；在线问诊流程复杂
	平台局限（B26）	病历查询不互通；在线问诊科室单一；无法健康查体；没有辅助技术和既往病历查询；患者没有指引，可能找错医生
	平台技术（B27）	画面卡顿；视频、图片看不清
政策（A07）	医疗环境（B28）	缺乏激励机制；医护人员权益缺乏保护；存在执业风险隐患
	市场环境（B29）	缺乏统一的质量标准；定价机制有待完善
	行业环境（B30）	缺乏规范和监管；医生坐诊时间随意

3.3 选择式编码

选择式编码旨在进一步厘清主范畴间的关联，梳理范畴之间的逻辑关系，阐述其中典型关系结构和脉络。本研究以“在线问诊服务质量”为核心范畴，4 个主范畴均影响在线问诊服务质量。医生和患者维度作为在线问诊服务的交互主体，发挥决定作用。平台维度是在线问诊服务的载体，影响医患双方的服务体验。政策维度是从外部环境下影响在线问诊服务体验。具体关系如表 7 所示。

表 7 选择式编码形成的主范畴间的关系结构

关系结构	关系类型	关系内涵
医生维度→在线问诊服务质量	决定作用	医生作为在线问诊服务的提供方，是在线问诊服务质量的决定因素，医生专业能力、沟通能力，与在线问诊服务质量正相关。
患者维度→在线问诊服务质量	决定作用	患者作为在线问诊服务的需求方，是在线问诊服务质量的决定因素，患者表达能力、沟通能力，与在线问诊服务质量正相关。
平台维度→医生维度→在线问诊服务质量	内在结构	平台是在线问诊服务的管理方。平台机制、平台功能、平台技术的越完善，方便医生利用，在线问诊服务质量的评价可能越高。
平台维度→患者维度→在线问诊服务质量	内在结构	平台是在线问诊服务的管理方。平台机制、平台功能、平台技术的越完善，方便患者使用，在线问诊服务质量的评价可能越高。
政策维度→在线问诊服务质量	外部影响	完善医疗、市场、行业环境，医生对于在线问诊的积极性可能会增加，行业规范性也会加强，相应的，在线问诊服务质量的评价可能就越好。

3.4 理论饱和度检验

本研究遵循扎根理论的分析方法，随机抽取 20 份访谈数据进行编码和分析，将得到的概念和范畴进行重复比较，挖掘各范畴之间的类属关系，并用剩下的 5 份访谈材料进行饱和度检验，发现没有新的概念及范畴出现，范畴之间也没有新的关系，说明该编码体系通过了理论饱和度检验。

4 构成阐释与模型构建

4.1 在线问诊服务质量构成阐释

借鉴 Donabedian 教授提出的医疗质量三维度理论[4]，可将在线问诊服务质量的构成分为结构质量、过程质量、结果质量。过程质量关注诊疗过程的技术能力以及医疗服务是否得到恰当的实践。结构质量关注医疗发生的环境和医疗产生的工具。结果质量是医疗服务有效性和质量的最终验证因素[5]。具体服务质量构成如图 1 所示。

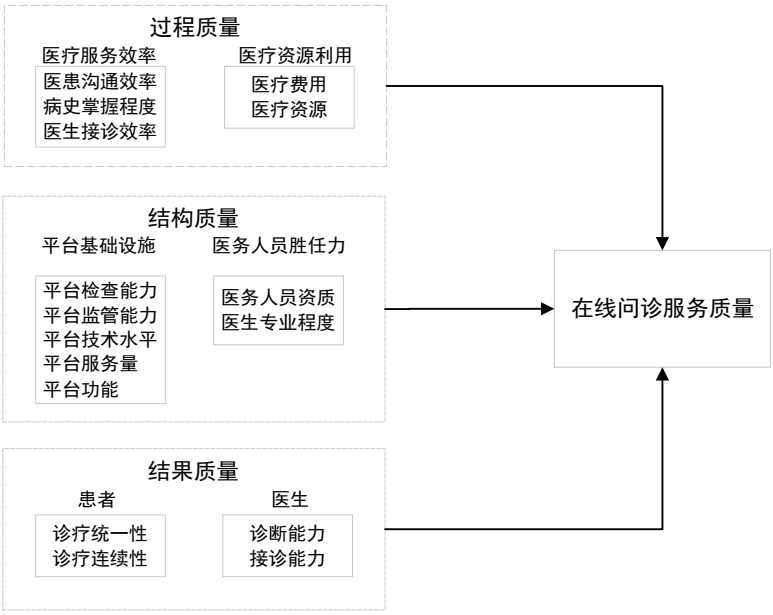


图 1 医护人员视角下在线问诊服务质量构成

4.2 在线问诊服务质量影响因素分析

医护人员视角下，在线问诊服务质量受到医生、患者、平台以及政策四个维度的共同作用，其中，医生维度和患者维度是在线问诊服务质量的决定因素，平台维度通过医生和患者两个主体，对在线问诊服务质量产生作用，政策维度通过对医生产生激励或监管，对医疗环境、市场环境、行业环境产生作用，从而影响医生提供的在线问诊服务质量。在线问诊服务质量影响因素模型如图 2 所示。

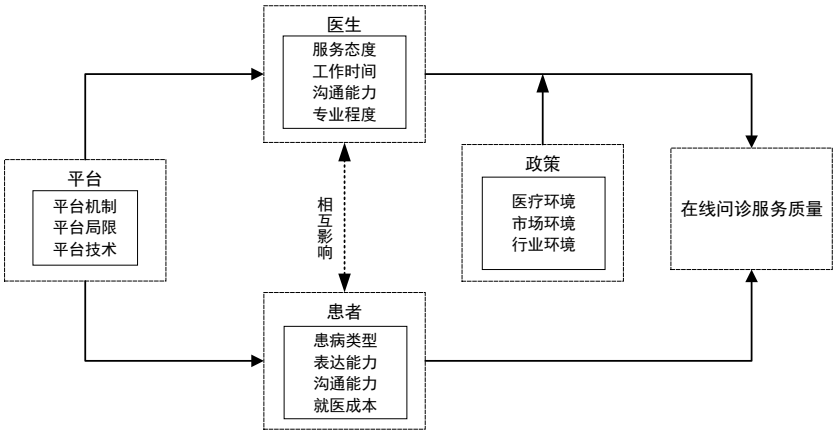


图 2 医护人员视角下在线问诊服务质量影响因素模型

（1）医生维度

医生在开展在线问诊服务时，可能存在工作时间安排紧张、医生个人综合素质有待提高的问题。在工作时间安排上，医院应统筹医生线上线下的工作安排，完善线上问诊医生的工作流程，为医生安排合理的工作时间。并适当提高线上医生的待遇，提高线上坐诊医生的积极性。在服务能力上，需要医生提高自身综合素质和专业能力，加强线上沟通能力。

（2）患者维度

患者不仅是在线问诊服务的需求方，也是在线问诊的服务的主体。一方面，由于患者缺乏专业的医学知识，对自身的症状和病情难以准确描述；另一方面，慢病患者大多为年龄较大的群体，语言表达和沟通能力可能较差，医生难以获取有效的病情信息。为解决以上问题，一是可以开展医学常识的普及工作，尽可能提高群众的医学信息素养，如现在很多医院小程序设置了健康宣教板块，为公众普及医学常识。二是在线问诊平台可推出线上导诊服务，如武汉市第一医院互联网医院推出智能自诊服务，通过描述具体症状，帮助患者选择正确的科室。武汉市第一互联网医院还设立了人工客服分诊服务，提前了解用户情况，并将信息转达给指定医生，提高医生接诊效率。

（3）平台维度

在线问诊平台是连接服务方和需求方的桥梁，影响在线问诊服务过程的用户体验，平台功能单一、就诊流程复杂、缺乏技术支持问题较为明显。作为在线问诊服务平台的管理方，平台设计上，完善平台功能，优化用户的使用体验，实现病历信息互通共享，同时需要关注到用户隐私泄露的问题，做好患者隐私保护工作。平台服务运营上，平台作为在线问诊服务开展的载体，加强相关技术支持，保证平台的稳定性和时效性。

（4）政策维度

政策是在线问诊服务的外部环境，影响医生提供在线问诊服务的体验，有研究证实政策可以吸引更多医生提供在线问诊服务，应改善线上医疗体系的政策制定，提高医护人员积极性。因此，政策制定上，可以完善激励机制和权益保护机制，为医护人员提供良好的医疗环境。通过营造统一的质量标准和定价机制，改善医疗环境和市场环境。为解决医护人员坐诊随意，监管部门需要加强线上行业的规范，营造良好的行业环境。

5 讨论

患者视角下的在线问诊服务质量研究中，质量评价中虽有差异，但都强调线上医疗服务的有形性、可靠性、响应性。本研究基于服务提供方视角，提炼出了其他维度，不仅关注诊疗服务效率、资源利用、平台基础设施、医护人员能力等软硬件设施，还强调线上线下、基层医院与三级医院诊疗连续性、统一性的重要性。一个完整的诊疗服务包括诊前、诊中、诊后三个环节，在线问诊服务

需要加强线上问诊与线下医院的一体化，完善就诊前的分诊、就诊中的问询和就诊后取药等服务的衔接，加强在线问诊医疗的一站式服务能力。

患者视角下的影响因素研究中，在线医疗服务存在医生耐心不足、接诊速度较慢等问题[6]，由于患者视角的研究评价较为简单，易于将服务质量的影响因素归因于医生维度。而医生视角下的研究更为关注医患沟通效率、问诊时间和所在医院的激励机制[7]，发掘出医院组织层面对在线问诊服务质量的外在影响。因此，本研究的不同之处在于对服务提供方的访谈材料进行扎根编码分析，将政策维度细化为医疗环境、市场环境、行业环境 3 个要素，既强调了激励机制的促进作用，又提倡加强对线上医疗的行业监管，营造和谐的医疗环境、规范的行业环境、有活力的市场环境。

6 结语

本研究运用扎根理论分析方法，以不同省市 25 位医护人员作为访谈对象，借鉴 Donabedian 教授提出的医疗质量三维度理论，将在线问诊服务质量的构成分为结构质量、过程质量、结果质量，强调了结果质量中诊疗服务连续性和统一性的重要性。将在线问诊服务质量的影响因素拓展为医生、患者、平台以及政策四个维度，提倡改善政策维度中的医疗环境、市场环境、行业环境。

本研究也存在一定局限性。采用扎根理论分析方法具有主观性，构建出的在线问诊服务质量的影响因素模型可能需进一步验证完善。后续研究可结合结构方程建模等定量研究方法，提升研究模型的普适性。

PO-038

大型综合医院入院服务中心的信息化建设实践与应用创新

陈启岳[1] 沈亮[1] 陈灿达[1] 周敏[1]
浙江大学医学院附属第一医院

1 引言

近几年医疗卫生行业主动拥抱互联网+思维、技术、服务势头迅猛，蓄力构建医疗健康服务业的新流程、新服务、新体验。浙江大学医学院附属第一医院（简称“浙大一院”）为进一步探索与践行 “最多跑一次”改革行动，设立四大服务中心深层次延伸医疗服务 O2O 触角，包括综合服务中心、入院服务中心、检查服务中心、志愿服务中心。其中融合了“互联网+”思维的入院服务中心，跨界创新与探索实践统筹全院床位资源与检查资源，重塑医院入院服务流程，为患者提供一站式、集约化、智慧化入院服务品质保障。

2 服务集约化融合

平均住院日是衡量医院服务质量和综合实力的重要标尺之一。为多方位提升患者就医体验，医院基于2018年度改革行动实战成果基础上，继续拓展面向住院服务的提升工程，打破传统入院服务“多部门、多窗口，多方式”及床位管理松散化壁垒。梳理住院服务各部门业务，床位资源集中调度，重塑业务流程服务，将床位协调、待床预约、入院手续、院前检查、入院宣教、住院咨询等多业务在入院服务中心集约化管理。入院服务所涉部门业务整合详见图1所示。

图 1 入院服务集约化整合

入院服务中心运行目标为统筹全院床位资源，加快床位周转，降低均次费用，提升服务质量与满意度。落实浙江省对日间手术患者的医保支付政策，同时面向患者建立预住院服务流程。预住院流程是以患者自费性质办理虚拟病区入院手续，待完成院前检验检查医嘱执行，待病房、手术室、麻醉科明确治疗方案后，患者再入住病区。预住院流程具体为门急诊入院申请，入院审核确认，虚拟床位预约，入院登记缴费、院前医嘱预约执行、手术排程(含日间手术)、入院知情宣教、实体病区入住。

整个预住院流程均由入院服务中心一窗式办理，入院服务中心实时跟踪待床病情动态，结合床位动态、入院诊断、检查结果、危急程度、治疗方案、手术安排等多维度权衡协调实体床位安排入住，院前检验检查具体费用出院时统一结算。预住院流程相对传统模式大幅度缩短住院中等待检查和术前待床的时间，降低平均住院日及费用支出，提升待床患者收治率及满意度。

。-保障了新流程在全院各部门高度配合下从试点到全覆盖过程中有序有质推进建设。

3 系统创新性建设

入院服务中心作为患者住院服务的调度枢纽中心，融合了多部门服务内容和业务流程。新流程贯通了门急诊电子病历，入院服务管理、医护医嘱管理、临床检验系统、医技预约登记、住院费用管理、住院电子病历及护理管理等业务系统。

1） 一体化入院管理

入院服务中心服务内容覆盖整个院前流程，具体包括入院申请单管理、院前检查医嘱管理、全院床位协调管理、患者入院预约通知，病种收治指南梳理、综合运营数据等，具体分为窗口服务和坐席服务。窗口服务主要业务包括申请单审核、床位协调预约、入院登记缴费、医嘱核对执行、检验检查预约、宣教咨询答疑等；坐席服务主要业务包括入院语音通知、病区床位协调、手术日患者安排、患者来电咨询、入院知情宣教等。通过建设入院服务中心综合管理系统来贯通各业务，即将原分散、独立的部门业务归并整合，具体业务详见图 2 所示。

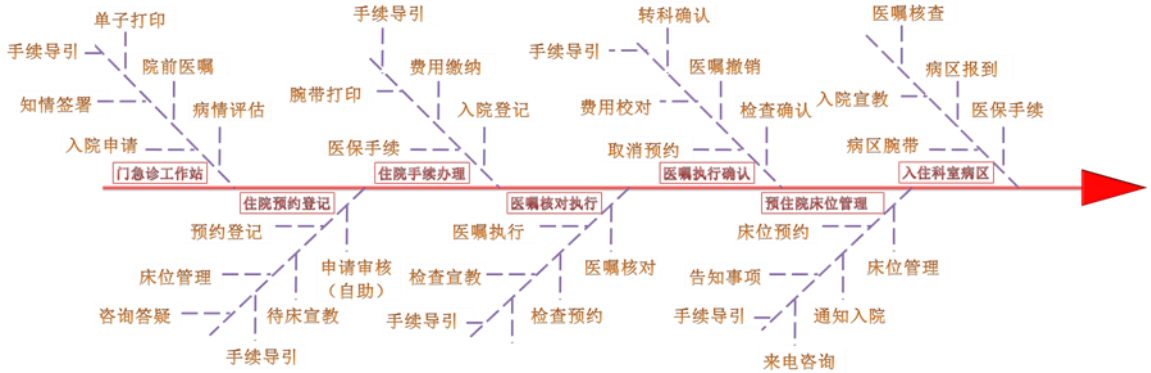


图 2 一体化入院服务流程

2） 自助式入院预约

新业务流程及管理机制下，入院申请单作为入院服务流程的起点，医院严格执行首诊负责制。门诊电子病历系统支持入院申请单及院前检查医嘱项目开具，检查项目在线评估，知情同意书电子版打印。通过优化建设实现入院信息数字化、数据可溯化、管理规范化，提升了门急诊医生工作效

率。解决传统版印刷版手写住院证、检查单、知情同意书等易遗漏遗失，难辨识弊端。

利用互联网+思维及移动用户黏性，每张入院申请单带有二维码数据属性，支持患者扫码自助确认档案信息、申请信息、待床地址、到院路程耗时，床位类型意向、紧急联系电话等关键属性信息。当日新开具入院申请的患者及家属可直接离院，手机端自助审核后进入待床状态，等待入院服务中心通知再到医院办理入院手续。二维码自助审核解决了业务窗口当日办理入院患者与当日预约入院叠加压力效应，缩短患者院内滞留时间，降低患者院内感染风险性。

3）一站式窗口服务

入院服务中心设置 8 个窗口，6 个窗口用于申请单审核、床位预约、入院登记、费用缴纳、医嘱执行、检查预约、宣教咨询等，2 个窗口用于财务费用缴纳(现金)、医保事务，窗口坚持一岗多责、一站受理制。

患者持身份证、就诊卡、申请单二维码即可通过自助取号机一键取号，同时校验入院申请单状态参数，确认入院服务当前业务队列，一次取号即可完成全部业务办理，避免多次取号，等候队列实时校验 HIS 端申请单状态不同状态在入院中心业务窗口队列不同。将窗口划分为三大业务队列，具体为：住院预约、入院登记、院前检查，三个队列有业务级联属性，可因实际业务需求办理转诊处理换队列。所有窗口支持非现金(如移动支付、银联卡支付)，无需通过转诊方式过渡财务专窗。叫号系统实时更新各业务队列等候患者数量关系，窗口服务人员可动态调整呼叫队列，优化了窗口队伍秩序及服务体验感。

4）智能化语音坐席

床位统筹协调作为入院服务中心重点工作之一，面向医院内部用于病区床位协调、院前检查等业务沟通，面向社会用于待床患者入院通知、手续流程信息、院前检查安排、待床患者咨询热线等服务沟通，依赖于信息系统与话务通讯两者相辅，共设置 9 个管床护士岗位，每个护士管理 10 来个病区床位。

结合管床护士人员配置、全院床位周转率、日均话务量、工作环境空间，融合桌面呼叫的智能话机成为首选方案，通过话机系统与 HIS 系统集成，主叫时管床护士可以直接从 HIS 端入院申请单管理患者联系电话直接发起呼出，被叫时 HIS 端根据呼入号码校验入院申请单留存号码，自动定位患者入院申请单信息，设计方案详见图 4 所示。

图 2 智能语音服务

利用现有市话通信网络及布线工程，项目成本、进度、质量一步到位，解决了传统话机诸多弊端。如主叫号码盘拨号易错率高，市外加拨“0”重拨率高，手持听筒影响系统交互；非标准环境窜音噪音严重，话务绩效监管缺失等所致通话耗时长、质量低、效率低，改善了床位协调员话务服务能力。

3） 联动式协同管理

医院的待床患者具体门急诊开单医生、申请入住科室医生、入院服务中心由三方协同管理。HIS 端系统互联互通，各方均可通过各所属业务信息系统，实时监测患者待床期间入院申请单状态、院前检查执行状态及报告、科室待床人数及天数等。若需要附加检查则由科室医疗组和入院服务中心双方，通过语音坐席系统联系患者沟通明确后，在医护医嘱系统开具附加院前医嘱，预约具体检查时间，及相关注意事项宣教。若出现危急值则启动包括医技部门在内的四方联动紧急处置机

制，保障患者第一时间得到专业治疗及处置指导，坚持“安全至上”原则。在信息流中衔接了各关联部门角色，予以保障患者待床状态高质量管理。

4 成效与思考

医院持续深化和践行“最多跑一次”改革行动，2019 年初组建入院服务中心并明确部门职责，同时一举取消加床挂床合计 1100 余张。全院两个院区保持 3300 张床位，门急诊量日均约 1.45 万人次，入院申请单日均约 1200 人次，二维码自助审核率达到 92%，入院量日均约 900 人次，预住院手续日均约 260 人次，院前检查日均约 600 余人次。平均住院日较 2018 年度下降 1.45 天/人次，术前平均待床日下降了 2.1 天/人次。随着“最多跑一次”行动的深入推进，入院服务中心将全院床位向更集约、透明、高效、科学方向转化，也助推了各临床科室服务质量及管理水平的联动提升。

随着医院多院区一体化管理模式建设推进中，入院服务中心将面临多院区患者共享、床位共享、服务共享的创新，建设标准化的入院服务中心已纳入新的日程，持续推进患者院前、院中、院后全周期高质量管理。

建设智能化语音客服，基于智能语音及网络电话技术，将入院服务、床位咨询、手续办理等常见患者咨询问题，梳理成规范性答疑知识库。通过 AI 语音实现解决患者日常入院咨询及待床进展自助解答，释放现有管床护士语音话务压力。

建设线上化移动服务，基于移动终端及互联互通技术，将整个入院手续移植到患者移动端线上自助服务，包括申请单确认审核、待床查询、床位预约、入院登记、网络支付、电子票据、检查预约等业务，打造线上自助服务有利于打破现有窗口式服务理念与业务瓶颈，逐步向无纸化医院有序过渡。

建设可视化数据中心，基于 CDR 数据及大数据技术，将多院区门诊就诊、入院申请、床位资源、检查资源、疾病诊断、手术安排、DRGS 付费等入院服务资源整合。动态监测住院临床运营数据，进行深度数据关联分析，利用数据行辅助临床及管理智慧决策。

5 结语

浙大一院作为浙江大学“双一流”建设的重要组成元素，且正在打造国际级医学中心，入院服务也将面临多院区一体化、同质化管理模式创新。随着“互联网+医疗健康”创新服务持续拓展，同时坚持业务导向、服务导向，聚焦业务流程创新再造与服务理念破垒新出，建立成熟、规范、感知的入院服务体系是新时代医院建设发展的关键落点。

PO-039

基于医院信息平台智慧病房的建设

曾丽娟[1] 段涛* 文嘉希[2] 向波[2] 刘文容[2]
川北医学院附属医院

1 引言

近年来随着互联网与物联网技术不断发展，各医院数字化建设朝着智慧医院方向逐步推进。智慧医疗建设已成为增强医院竞争力与创新力的关键，被称为医院第三次信息化革命。《新一代人工智能发展规划的通知》就明确提出鼓励运用人工智能技术发展智慧医院建设。智慧病房以辅助医疗为目标，是智慧医疗中物联网使用的一个主要应用场景，同时也是医院信息化程度提升的一个重要标志。智慧病房旨在提升患者住院体验、减轻医护人员工作量、美化病房环境。

2 医院现状

我院信息化建设在经历了一系列发展阶段后取得了飞速的进展，各种信息系统都顺利推行，也实现了医疗资源的高效利用，但智慧病房建设尚处于空白。我院肿瘤科、泌尿外科、神经内科目前核定床位 430 余张，随着住院病人的增加，医护人员均处于高负荷工作状态。传统的病房管理模式已无法满足当前的治疗需求和就医体验，主要存在以下问题：（1）护理人员：输液、用药、治疗、宣教等医疗环境的身份识别均采用口头核对方式，存在出错的风险；床头卡、护士站一览表均采用手工书写模式，增加护理人员工作量；难以第一时间有效掌握危急重症病人的体征变化；（2）医生：查房需要携带纸质病历，才能全面的了解病人病情，并且无法实时下医嘱写记录；

（3）患者：对住院期间治疗安排、病情发展、健康宣教、住院费用不清楚明白，缺乏获取相关信息的便捷渠道；无快捷有效的反馈途径，传统的问卷调查影响了反馈的主观性并且不方便后期统计管理；缺乏人文化，没有系统能获取娱乐内容及生活服务渠道，患者住院品质得不到提升。基于以上问题，利用信息化技术建设智慧病房将是很好的解决方案。

3 系统建设目标

智慧病房建设的主要目标是以患者为中心，智能交互终端为载体，以患者的健康数据管理为主线，通过医院信息平台，将患者、医护人员、医疗设备通过多网络紧密联接，低成本、高效率建设以患者为中心的信息服务平台，为患者提供主动式、多样化、智能化、精准化的医疗生活服务，提高病人就医体验，降低护理人员劳动强度，提升医护工作效率，创造舒适病区环境，降低医院运营成本[5]。

4 系统设计

4.1 系统架构设计

智慧病房依托物联网技术，以智能交互终端为一站式信息服务平台，充分利用软硬件协同开发。整个架构主要分为用户应用层、业务应用层、数据接口层、物联网络层、智能设备采集层。系统架构图图 1 所示。

图 1 架构图

4.2 系统网络结构

系统硬件终端包括护士站大屏、床旁屏、门口屏、走廊液晶屏、护士站电话机、床头呼叫器、卫生间呼叫器、台式电脑、楼层前置机、交换机、机房服务器等。床旁屏、床头呼叫器、卫生间呼叫器安装在病房内，床旁屏为患者提供相应的住院信息服务，床头呼叫器安装在设备带上，卫生间呼叫器安装在病房卫生间内，方便患者在卫生间发生意外时紧急求助；走廊显示屏安装在病区走廊区域，门口屏安装在病房门口，可同时响应患者呼叫；护士站大屏、护士站电话机、台式电脑安装在护士站区域，方便护士统筹管理整个病区的呼叫，护士通过台式电脑登录系统 web 端后，可进行后台维护。系统以标准的 TCP/IP 网络通讯协议与组网布线方案实现各终端在病区场景的连接，网络结构图如图 2。

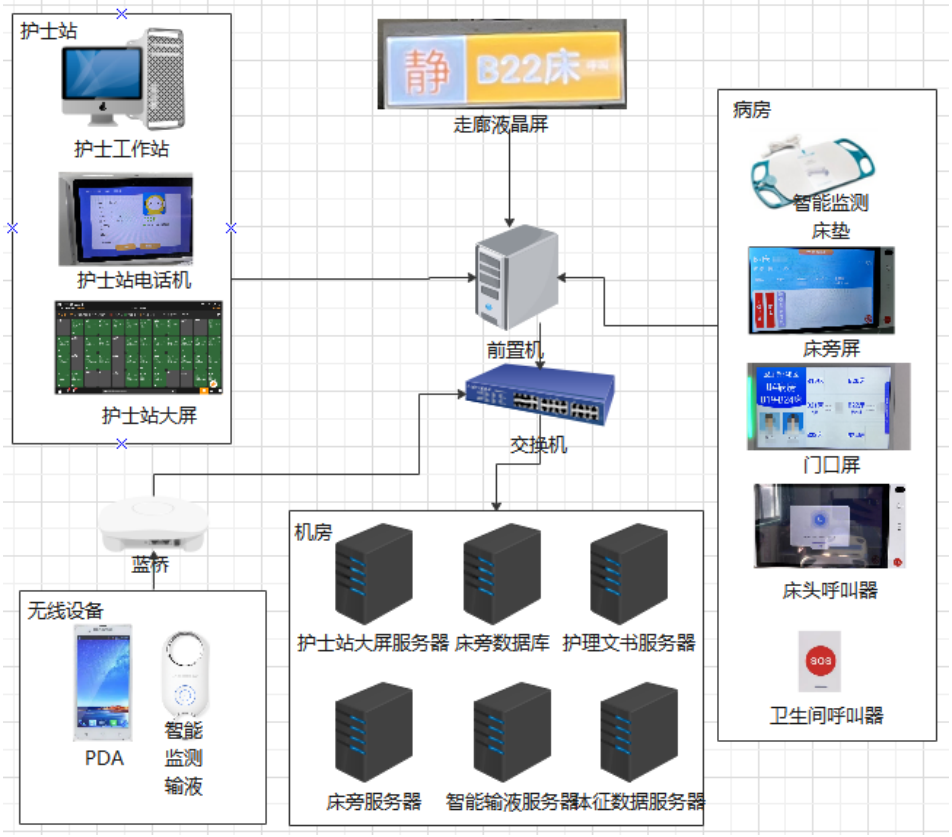


图 2 网络架构图

4.3 系统功能设计

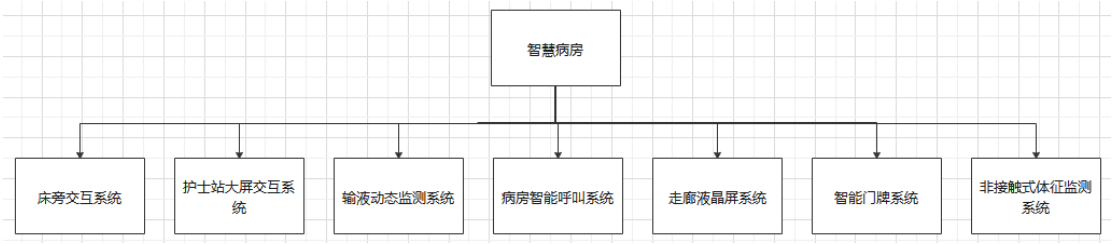


图 3 系统功能结构

本次智慧病房建设主要包括床旁交互系统、护士站大屏交互系统、输液体征动态监测交互系统、病房智能呼叫系统。

4.3.1 床旁交互系统

床旁交互系统，以床旁显示屏为信息交互终端，基于病房设备带中有线网络的连接，通过平台深度对接医院各类业务系统，借助大量模块化服务引擎，实现为患者高效精准的推送服务。主要包括病人床头卡功能、信息查询、信息推送、床旁支付、健康宣教、满意度调查、医院介绍、生活服务等模块，功能结构图如图 4 。

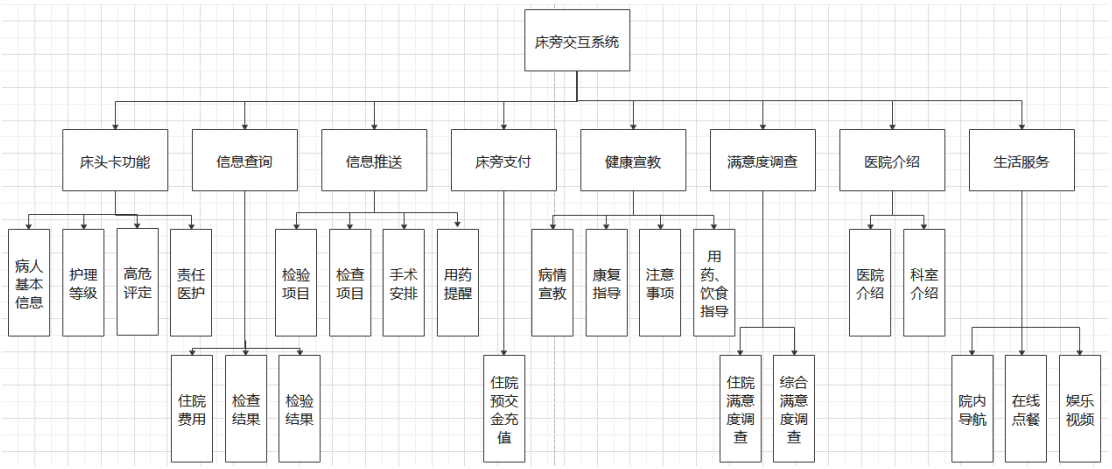


图 4 床旁交互系统功能结构

床旁交互系统通过数据中心与院内 HIS、EMR、PACS、LIS 等多个数据库对接，实现床头卡信息、检查检验报告查询、费用查询等功能；通过医嘱智能提取，将医嘱进行分类，以消息推送的形式提醒患者用药、检查、检验、手术等，从而帮助患者了解住院期间的治疗项目及注意事项。床旁交互系统能将院内的健康宣教资源进行整合，内容包括病情宣教、康复指导、注意事项、饮食指导、药事指导等，形式包括图片、文字、视频、语音等，护理人员可针对患者病情选择性的推送宣教知识，患者也可反复观看并及时将观后情况反馈给护士，这种新模式的健康宣教不仅能分担护士工作，对患者健康宣教也更为直观、易理解、更具针对性。床旁屏作为信息交互载体，集成医院满意度调查管理系统，患者填写满意度调查能避免传统调查问卷的真实客观性，并便于后期整理分析数据，为医院管理层提供优化住院医疗服务决策。为了缓解患者住院期间焦虑紧张情绪，患者可在床旁屏上选择观看娱乐视频，后台可以定期对视频进行更新，这样能给予患者更多的人文关怀。床旁交互系统与医院支付平台对接，支持床旁便捷支付，实现住院预交金的充值；后期与医保系统对接之后，护士进行出院操作、住院收费室审核后，患者即可在床旁进行出院结算。床旁屏上可以添加医院介绍、科室介绍、医生介绍及相关医疗技术，方便患者更能了解医院。系统与院内导航、食堂在线点餐系统进行集成，方便了患者就医体验，后期还可与其他第三方软件进行集成，为患者提供便利的生活及医疗服务。

4.3.2 护士站大屏交互系统

护士站大屏交互系统通过平台及时、可靠、全面的获取 HIS、护理文书、手麻、LIS 系统中临床数据，使护理人员能够更为精确地了解病区动态、护理任务。硬件采用 72 寸超清触摸大屏，方便信息交互。系统采用护眼模式，给护理人员带来贴心体验，主要包括病人一览表、病区动态、护理任务、特殊交班、功能助手等模块。

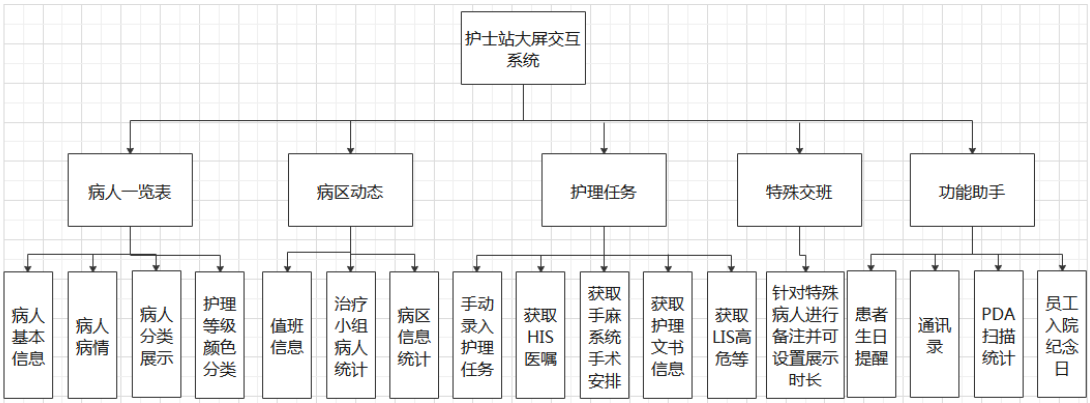


图 5 护士站大屏功能结构图

病人一览表主要展示患者基本信息、住院信息，并且根据护理等级、危重状况、高危评定、手术等进行分类，方便护理人员快速了解患者病情；病区动态主要包括病区整体情况，如值班信息、治疗小组管床统计、出入院情况、手术安排等内容；护理任务依据各个科室实际需要可定制展示模

板，系统动态获取 HIS 医嘱内容、手麻手术安排、LIS 危急值、护理文书高危评定等内容后自动显示。系统也支持手动新增处置任务，弥补临床紧急情况，比如医生抢救过程中下的口头医嘱；大屏还包括一些功能助手如患者生日提醒、通讯录等。

4.3.3 输液动态监测与非接触式体征监测系统

输液动态监控系统通过采用物联网技术，结合输液监测终端、PDA、智能呼叫，实现输液智能化，从而全面实现输液患者安全输液的闭环管理，为患者提供连续、全程的输液护理服务。后台监测系统负责输液信息的存储、计算、分析；输液监测终端负责数据实时显示，主要包括患者信息、当前速度、输液状态、剩余容量等。

非接触式体征监测系统，采用一体化便携式床垫，可监测病人血压、脉搏、心率、是否离床等数据，然后通过蓝牙控制中心将病人体征数据实时传到后台，系统与 HIS 进行对接，输液监测终端实时显示病人体征数据，对于有异常的数据实时报警。整个病区输液信息、体征信息的集中显示，医护人员对整个病区病人输液体征情况一目了然，患者及家属可安心休息，极大提高患者就医体验。

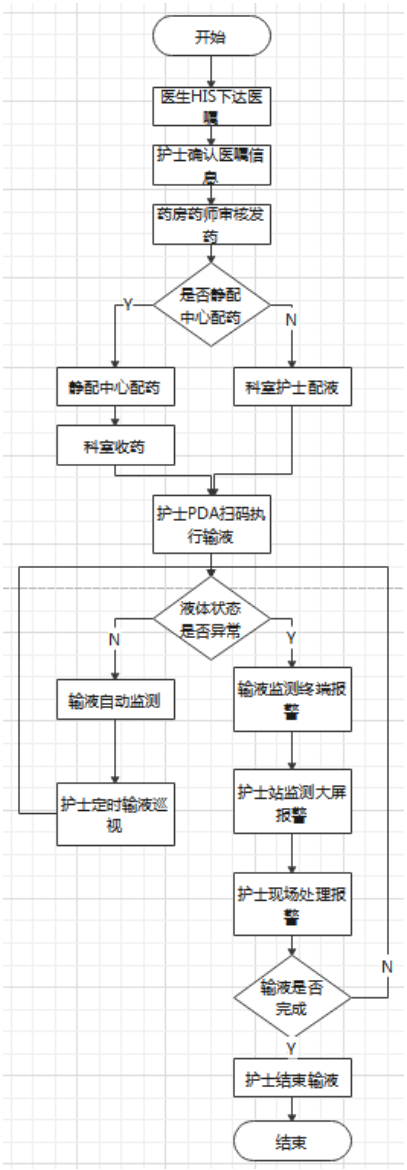


图 6 智能输液监测流程图

4.3.4 病房智能呼叫系统

病房呼叫系统基于物联网传输技术，实现护理人员与患者信息联络。患者通过床头屏上的呼叫器进行呼叫，护士站可视化终端、门口屏、走廊屏多屏同时接收到呼叫信息，及时联动响应，并且后台记录了呼叫记录，方便后续进行查询。

5 实施

5.1 移动端 APP 开发模式

护士站大屏、床头屏、门口屏、走廊屏、护士站呼叫系统、输液动态监测与非接触式体征监测系统都是基于安卓系统开发的 APP，采用 Cordova+Ionic+AngularJS 混合开发模式，Cordova 是一款开放源代码的 App 开发框架，旨在让开发者使用 HTML、Javascript、CSS 等 Web APIs 开发跨平台的移动平台应用程序；Ionic 是一个新的、可以使用 HTML5 构建混合移动应用的用户界面框架；Angular 遵循软件工程的 MVC 模式，并鼓励展现、数据、逻辑组件之间的松耦合，通过依赖注入，Angular 为客户端的 Web 应用带来了传统服务端的服务，例如独立于视图的控制，后端减少了许多负担，产生了更轻的 Web 应用[6]。

5.2 系统开发模式

系统采用 MVP(Model-View-Presenter)模式，Model 提供数据，View 负责显示，Presenter 负责逻辑的处理。MVP 由 MVC 演化而来，MVP 结构简化了 View 层的复杂度，提高了 View 层对模型数据的访问效率。简化了系统的规模和资源开销的同时，提高了系统的灵活性和扩展性，使得系统易于维护[7,8]。

5.3 无线网安装

智慧病房部署在内网中，但部分应用通过无线网传输，患者数据涉及个人隐私，因此要选择安全性高的通讯技术，同时患者体征数据多元化，要确保数据传输吞吐量及速率，以及信号的抗干扰性。蓝牙 5.0 具有传输速度快、功耗低、传输距离远的特点，并且使用快速跳频、扩频、调频技术及前向纠错，能提高数据传输稳定性和安全性，所以智慧病房无线网络选用蓝牙 5.0[9]。

6 系统应用效果

本院智慧病房项目按照病区分为 4 期建设，目前已完成 1 期建设，共 7 个病区，包括：硬件与网络安装、护士站大屏交互系统、床旁交互系统、智能呼叫系统、输液动态监测与非接触式体征监测系统等。

通过对比发现，智慧病房建设能有效的改善医患沟通，提升患者就医体验，提高医护人员工作效率和服务质量，减轻工作负荷，降低差错率，避免医疗纠纷，从而科学合理地优化病房诊疗流程。项目启用以来，这几个病区患者满意度上升 21%，护理人员工作效率也明显提升。

7 总结

智慧病房依托智能化硬件、软件、数据平台，使患者与医护人员、智能医疗设备紧密联接，构建了病房与护士站一体化闭环，实现了美化病房环境、改善患者体验、提高医护人员工作效率，使有限的医疗资源得到充分共享利用，为医院创造更多的医疗服务价值[10]。随着人工智能、物联网、大数据的发展应用，未来智慧病房还会延伸扩展，迎来更大的挑战机遇[11]。

PO-040

基于 PACS/RIS 的数字影像服务系统的设计及应用

赵治远 卫新乐 郝晓伟 许锋波
河南科技大学第一附属医院

1 建设背景

近年来，随着移动互联网和智能手机的普及，大量涉及人们衣食住行的互联网应用已成为人民群众生产生活中不可缺少的一部分。

反映到医疗服务上，越来越多的医院开始利用互联网手段为患者提供便捷的服务。具体到医学影像检查领域，传统的以“实体胶片+纸质诊断报告”的检查信息载体正在向以“互联网+移动终端”为特征的数字影像服务模式过渡。

河南科技大学第一附属医院是一家大型三甲医院。医院在 2017 年建成了院内影像与检查报告自助打印系统，为患者提供影像检查胶片与报告的自助打印服务。该服务虽然相较于人工分装胶片与报告的方式减少了影像科的人力投入，且为患者提供了便利，但依旧有部分患者需要医院提供检查影像资料数字格式的导出和打包服务。同时，全院级 PACS/RIS 系统虽然可以提供院内医生工作站上影像资料的浏览，但临床医生外出开会、非工作时间内的非正式会诊咨询需求需要参考患者影像资料，通过屏幕拍照微信转发的方式图像质量得不到保证，提供互联网端、移动端轻量级检查影像浏览访问服务一直是临床医生的需求。

基于此，医院进行了数字影像服务系统的建设和应用探索。

2 建设方法

2.1 建设模式

数字影像服务实质上是将过去存放在医院内部的诊疗数据以数字化的形式向患者全面开放，并以此来推动检查结果的互认和共享。系统将为每个患者建有一个专属的存储空间，其中保存了个人各项检查报告和影像资料，这份资料能够长期在线存储，并能为患者的会诊、复诊和异地就诊等诊疗操作提供极大的便利，同时兼顾临床医生在医院外部通过联网终端对检查影像的访问需求。

医院信息管理中心经过调研，发现国内其他医院数字影像服务系统建设有委托商业公司租用公有云建设和医院自行建设两种模式。前者具有成本低、安装部署快、无须医院管理维护的特点，但对于像我院这种体量的大型医院而言，在目前的环境下，数据和安全等因素尚不能全面掌控。

信息管理中心从数据管理、网络安全、系统开发集成和应用流程等多个方面进行了详细评估，由医院 PACS/RIS 系统提供数字影像服务功能将有助于降低系统集成难度、快速上线，并可直接使用医院自建的存储空间向患者提供数字影像服务，无须将影像数据向第三方云存储空间上转存，有利于加强医院对诊疗数据和数字影像服务的管控能力。因此，我院的数字影像服务系统建设，采用了由医院自主规划、建设和管理，并将数字影像服务系统纳入医院整体信息化建设的一个组成部分。

2.2 技术路线

（1）功能指标

1、患者影像检查各个状态手机短信实时提醒，以方便患者知晓检查进程，合理安排时间。

2、检查完成后无须在院内等待，检查结果和影像资料可在线上获取。

3、检查结果和影像资料可长期在线存储，随时随地通过手机和电脑查看，不用担心胶片和报告不便携带以及丢失损毁等问题。

4、当患者有会诊需求时可通过手机一键转发会诊，影像资料可以便捷分享给其他医师。

5、通过微信企业号，为临床医生分配影像浏览权限，允许其通过互联网联网终端访问与其关联的患者检查影像资料。

（2）系统架构

为将患者的检查登记信息、检查报告信息和三个月的检查影像转存到数字影像服务系统之中，数字影像服务系统通过视图获取 PACS/RIS 系统中的检查登记、检查报告和影像文件，将 Web 服务向临床医师和患者提供数字影像服务。如图 1 所示：

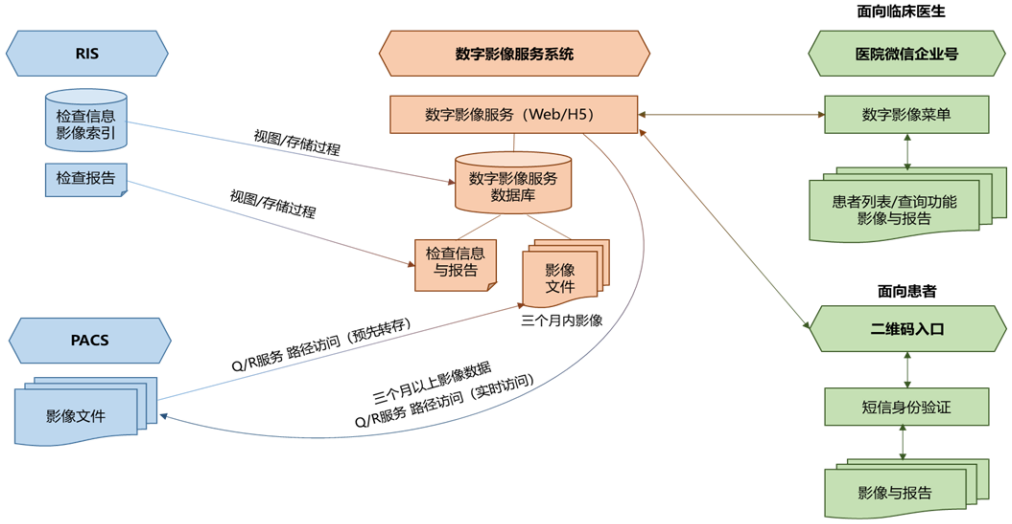


图 1 系统架构

系统采用 Html5 技术，支持 Android、IOS、Windows 操作系统，实现零插件安装、实现根据终端的类型提供自适应页面的大小。

（3）安全管理

数字影像服务系统部署在独立的服务器上，该服务器部署在医院信息中心的隔离区域内（DMZ），保护区内 PACS/RIS 服务器通过同步程序可以与数字影像服务器进行通信，但来自互联网的访问并不能只访问院内的 PACS/RIS 服务器，从而保障了院内核心业务的安全。

2.3 建设过程

（1）分工协作

为推动数字影像服务系统建设，信息管理中心成立了数字影像服务系统项目建设小组，委派两位工程师负责项目的调研和需求分析，并由四位工程师分别负责项目的软件开发商协调、安装部署、用户意见收集和管理维护，由 PACS 软件开发商负责软件的开发和集成，由硬件集成商负责硬件、网络平台的建设和安全策略的管控。

（2）上线过程

信息管理中心多次召开数字影像服务系统协调会，部署需求调研、软件开发和上线运行事宜，并安排多名工程师与软件开发商、系统集成商进行对接，全程跟踪数字影像服务系统项目的实施上线工作，尽可能在不影响影像科室工作的前提下，实现系统的平稳上线。经过为期两个月的需求调研、软件开发、安装部署和内部测试，2021 年 3 月数字影像服务系统上线运行，可提供 DR、CT、MR 三大影像检查的数字影像服务。

2.4 应用流程

数字影像服务系统分面向患者的应用和面向临床医生的应用，其应用流程如下。

（1）面向患者的应用流程

1）患者在我院 DR、CT、MR 做检查时，预约成功后影像科室会发给患者一个检查指引单，该单据上除告知了做检查的时间、地点及注意事项外，还印有患者访问个人数字影像服务的二维码。

2）患者做完检查后，可直接扫描预约指引单上的二维码，即可在线查看本次检查的检查影像，即使是成百上千幅的影像，也可以通过手机等智能终端在线浏览，系统提供放大、缩小、窗宽窗位调整等功能，还可以进行影像动态播放和对数值进行测量。

3）除了在检查完成后即可在线浏览检查影像外，还可以通过扫描二维码，随时关注检查报告审核发布的进度或在手机收到报告发布短信提醒后在线查看检查报告。一旦报告发布，还可以在线查看与实体胶片一致的电子版胶片和检查报告。系统提供两种在线查阅检查报告的方式，包括便于在手机上查看的纯文字版检查报告和与正式纸质版报告一致的电子版报告。

4）数字影像服务系统还提供微信端收藏和分享功能，可将个人的检查影像资料收藏至微信收藏夹

便于随时查阅，或通过微信分享功能发送给医师进行专业的诊断和会诊，提高诊断和会诊的效率。
5）患者通过自助报告打印机打印的检查报告单上也带有数字影像服务二维码，患者在看纸质报告时，也可通过手机扫一扫功能查看与报告对应电子胶片和影像资料。
（2）面向临床医生的应用流程
1）在临床医生的微信企业号中，增加了“院内数字影像”的菜单；
2）临床医生登录企业微信，点击院内数字影像菜单即可打开与其所在科室关联的在院患者的数字影像列表。
3）通过患者检查列表，临床医生可查询、选择某个患者，选择报告，查看其检查报告；选择影像，可在移动端打开影像浏览器，进行影像的查看，包括窗框窗位调整、播放、测量、放大、缩小等操作。

3 建设效果

数字影像服务系统自 2021 年 2 月上线运行以来，已经累计为患者提供 1 万余次的在线浏览服务，发送短信 2 万余条。2021 年 3 月份，每 10 例 DR、CT、MR 检查中，就有 4 例检查的患者使用手机扫一扫预约指引单或检查报告单上的二维码，访问数字影像服务。
在数字影像服务上线运行后，信息管理中心安排多名工程师到影像科室实地查看运行情况，并对患者进行随机调查，发现患者对数字影像服务功能普遍给予好评。据患者普遍反映：检查完成后不用再在院内排队等待取结果，可以通过手机在第一时间内查看检查影像。如果有认识的大夫，还可以把链接用微信发给大夫，及时进行会诊。一旦报告发布，手机端就能收到短信提醒，在手机上就能看检查结果，节省了大量的等待时间。
数字影像服务系统的上线运行，使来我院就诊的患者享受到更加便捷的服务，相较于传统胶片携带不便、容易损坏、信息量有限等不足，具有非常突出的优势：
1、携带更方便，不用担心传统胶片丢失、损坏或者携带不便等问题，患者转诊或者复诊，只需一部能上网的手机、平板或电脑，即可随时随地二次调阅。
2、影像信息更完整，随着高端影像检查设备的普及，动辄成百上千幅的影像检查资料已经不能通过几张传统胶片来传递，传统实体胶片的局限性往往导致患者外出就医时重复检查，进而加重患者的负担，而数字影像服务系统除了可以查看与传统胶片一致的电子版胶片外，还可以浏览患者全部的检查影像，更有利于诊断和会诊。
3、患者体验好，减少了患者在院内的等待，利用信息技术为患者获取个人检查报告和检查影像提供了最大的便利。
4、患者黏性的增加，向患者提供数字影像服务后，患者的历次检查记录均可通过个人联网终端查阅，能够提供我院范围内的连续的检查记录查询，有利于增加患者的黏性。
同时，该系统向临床医生开放在手机端、平板端、个人电脑端的院内影像和报告浏览服务，使得临床医师可以突破医院内网工作站的限制，在何时何地，通过互联网络即可访问患者在本院的检查影像，可以对 DICOM 原始影像进行浏览、缩放、窗宽窗位调整及重建，可以提高疑难病例的会诊效率。该应用已嵌入医院的微信企业号之中，丰富了微信企业号的应用内容。

4 讨论

顺应后疫情时代需要，减少不必要的接触，推进“互联网+医疗”便民惠民服务工作，努力做到让人民群众便捷就医、安全就医，为临床医生提供便利方便，数字影像服务系统有着巨大的优势，但目前还存在以下问题，需要以发展的眼光看待。

1、河南省暂无明确的收费标准

目前，类似我院的数字影像服务，在浙江、山西、贵州、辽宁等省份，省级物价部分已经有明确的收费标准。浙江省已于 2018 年 8 月 13 日出台了数字影像服务费收费标准，每例患者最高不超过 20 元，可提供门诊患者 15 年、住院患者 30 年的电子胶片无限次数查询、下载和分享服务[1],旨在推动数字影像服务的快速普及应用。
因河南省暂无明确的收费标准，我院目前是以“互联网+医疗”便民惠民服务的形式患者提供数字影像服务，全部建设和运行成本由医院承担，未向患者收取任何费用。因此，也并未向患者承诺数字影像服务可供查询的时限。

2、患者个人移动终端的差异性

受限于患者个人手机、平板等移动互联网终端的差异性，检查全序列 DICOM 影像的加载、浏览速度和显示效果存在着很大的差异。数字影像服务系统中所展示的信息只能用于轻量级的浏览、查看和临床参考，但不适用于影像诊断（经 SFDA 或 FDA 认证的移动设备除外）[2]。虽然数字影像服务提供的全序列 DICOM 影像浏览能够突破了传统胶片的局限，可以保留更多、更真实的检查影像信息，但相较于传统实体胶片+观片灯能够明确用于影像诊断的特点，数字影像服务适用于影像诊断，需要匹配满足医学影像诊断需要的显示设备,而患者个人移动终端显然是很难满足的。未来数字影像适用于影像诊断，还需要形成闭环，比如在患者持数字影像服务链接到其他医院就诊中，能够在医生端符合医学影像需要的诊断设备上访问数字影像服务，并加载全序列无损的 DICOM 影像。实现这一过程，还有很多工作要做。

3、影像调阅的速度

我院上线的数字影像服务，在联通、电信、移动稳定的 4G 网络稳定满格的信号环境中，CT/MR 全部序列缩略图加载速度小于 5 秒，单序列首幅图像加载速度小于 2 秒。但因为提供的是全序列无损 DICOM 影像的浏览服务，尽管图像经过无损压缩并支持边下载边浏览技术，但上百幅的 CT、MR 检查图像全部下载完成尚需数十秒甚至更长的时间，这是影响使用者使用体验的关键环节。

在数字影像服务系统上线后，我院正在部署 5G 基站，探索 5G+医疗的应用场景，推广 5G+医疗应用。经过初步测试，在 5G 网络下，影像的加载速度较 4G 网络有了 70%左右的提升。未来 5G+医疗的应用，海量影像数据的互联网高速访问，将是重要的应用场景,这也为数字影像服务全面用于影像和临床诊断的闭环提供了可及性。

PO-041

电子病历档案系统在医院病案标准化管理中的管理效果、管理规范情况探索

王刚

哈尔滨医科大学附属第四医院

所谓病历，是指出对患者疾病的发生、发展及转归等情况进行详细记录的一种临床资料，病历不仅仅可以作为实践工作的总结性报告，同时还可以方便临床对患者病情的发展规律进行有效的探索。对于医院而言，患者群体的病历资料是重要的信息资源，医院在医疗、科研、管理甚至法律等多个方面均离不开病历的帮助，病历资料本身就拥有着重要的信息价值[1]。随着我国医疗事业的不断发展进步，以及国民健康意识的增强，导致就诊的人数较过去出现了极大幅度的增长，就诊压力的加大在一定程度上增加了医院在病历档案管理方面的压力与难度，也导致病历管理不规范等问题频发。如今，为了能够更好管理海量的病历档案，有不少医院都实现了病历档案的改革与升级，应用电子病历档案后不仅提高了管理工作的效率，也拉高了整个管理部门的工作效率，实现了医院内部管理的革新与进步[2-4]。鉴于此，本研究将选取我院 120 份病历作为本次研究的对象，旨在分析电子病历档案系统在医院病案标准化管理中的管理效果，对电子病历档案的管理规范情况进行探究，现将研究结果整理如下。

1 资料与方法

1.1 一般资料

本研究对象为：120 份临床病历，样本选自 2020.1-2020.12 月期间，以随机数表法对研究样本进行分组处理，最终分为对照组与观察组，两组各为 60 例。对照组病历资料中，男女比例为 33：27，患者年龄为 20~64 岁，平均年龄为（40.41±8.45）岁，BMI 为 18~24kg/m^2，平均 BMI 为（20.21±1.24）kg/m^2；观察组病历资料中，男女比例为 32：28，年龄为 20~65 岁，平均年龄为（41.11±8.61）岁，BMI 为 18~24kg/m^2，平均 BMI 为（20.29±1.32）kg/m^2，两组间一般资料等对比无明显统计学差异（P＞0.05），可进行对比。

1.2 方法

对照组实施临床病历常规管理：对所有病历资料进行统一的日常管理，不许对病历资料进行任何的篡改或涂改，不能对病历信息进行任何的伪造、隐匿，不得随意销毁病历资料，若对病历档案有借阅需求，则必须向病案主任医生提出申请，申请同意后方可借用病案资料，但是任何病案资料都不能带出病案室。

观察组实施电子病历档案系统管理：（1）重视基础设备部署，并开展相关的培训工作。为充分发挥电子病历档案系统的管理效能，医院必须配搭软硬件的部署，电子病历档案系统的管理依赖于电子技术，需建立系统且全面的电子档案系统，完全实现电子化管理；其次，还要组织临床各个科室的人员参与电子病历档案系统的应用培训，让所有医生都可以在临床工作中熟练且规范地应用电子病历管理系统，培训过程中需强调信息的完整性与全面性；（2）重视电子病历的规范操作。针对各科室的电子病历使用情况必须统一展开培训工作，规范病历的书写要求，对不同科室的病历进行随机抽查，检查临床上电子病历的记录情况，若其中任何一个环节出现问题，都应在第一时间进行纠正与完善，比如对归档前的电子病历进行详细的查阅、质检，全面检查病程记录情况、病历记录情况以及终末病历首页情况等，对出现的问题进行及时的整改，杜绝电子病历出现质量低下的情况；（3）及时收集、合理应用临床病历。部署电子病历档案系统后，能够有效缩短病历档案的收集时间、整理时间以及归档时间，方便于医院对信息数据进行统计与科研；医院还可以建立数据中心，让医院内部实现信息的横向共享，让医疗机构之间实现有效协同，推动协同效能可以延伸至整个区域，让医院与公共卫生组织、保险机构、社会卫生服务组织等实现高效结合，让电子病历档案发挥重要的管理作用；（4）对病案的保管时限及介质进行明确。现阶段，医院对病案的储存情况有着明确的年限与介质要求，然而电子病历档案在现实中的铺设力度仍不强，因此我国目前还没有出台具体的政策划定储存的时间以及相应条件，而电子档案与传统的病案相比，则可以实现永久性的储存，尽管在多数情况下，并没有对此类资料进行永久性储存的必要，但医院仍可以借鉴纸质档案的管理办法来适当地延长电子病历档案的储存时间，同时积极开展电子病历档案系统的检查、备份以及升级工作，对电子病历档案进行规范管理；（5）电子档案与纸质档案管理制度并行。从病案的应用情况分析，即便是在信息化水平高度发达的现代社会，也不能让电子档案完全取代纸质

档案，电子档案具有纸质档案不具备的共享性与便利性，所以在医院的病案管理过程中，应当应用病历档案双轨制度，让两种不同的档案模式可以实现互补，让信息平台可以实现资源共享；（6）对档案进行有效的保护，在电子病历档案中，可合理应用软件技术及管理制度，让电子病历档案管理工作更具真实性、有效性以及系统性，针对重要的档案信息还可以应用加密软件，禁止非法访问者访问医院内部的档案系统；搭建相关的附录系统，对电子病历档案的输入、输出及访问等进行详细记录，推动电子签名管理，保证电子病历档案信息的完整性、真实性以及有效性。

1.3 观察指标

评估两组管理工作效果：评估病历资料的完整程度、真实程度、有效性以及安全方面等维度得分，每项分值为 0~20 分，最终分数值越高则说明管理的成效越好。

评估两组的管理规范性：评估依据为病历信息的完整性、涂改情况、填写情况以及是否存在缺页等情况，对管理规范进行综合评估。

1.4 统计学方法

采用 SPSS21.0 分析，计量资料以（$\bar{x}\pm s$）表示，经 t 检验，计数资料经 x^2 检验，以（%）表示，差异有统计学意义为 $P<0.05$。

2 结果

2.1 两组管理效果对比

经对比后发现，观察组的病历完整性、病历真实性、病历有效性及病历安全性等各个方面的管理效果评分均高于对照组（$P<0.05$），详见表 1。

表 1 两组管理效果对比（%）

组别	例数	病历完整性	病历真实性	病历有效性	病历安全性
对照组	60	14.45±2.24	13.62±2.78	13.56±3.84	14.58±3.85
观察组	60	18.96±1.89	19.65±1.54	18.98±1.47	18.79±1.65
t		11.919	14.671	10.210	7.785
P		0.001	0.001	0.001	0.001

2.2 两组管理规范情况对比

经对比后发现，观察组出现信息不完整、涂改、填写错误及缺页等问题的概率低于对照组（$P<0.05$），详见表 2。

表 2 两组管理规范情况对比（%）

组别	例数	信息不完整	涂改	填写错误	缺页	总发生率
对照组	60	7	4	2	2	15（25.00）

观察组	60	1	0	1	1	3（7.50）
x^2						9.411
P						0.002

3 讨论

我国作为人口大国，对医疗系统有着极高的需求，医院是保障民生质量的重要服务机构，而在医院的日常工作中，将产生大量的病案资料。所谓病案资料，是指详细记录了患者病情、临床治疗方法及患者临床反应等的资料合集，在医疗工作中病案资料发挥了极为重要的重要，它的出现不仅可以为临床提供客观、可靠的依据，还能成为医疗科研工作的关键数据。随着时代的变迁、医疗事业的发展，医院病案统计管理中存在的问题逐渐暴露了出来，对存在的问题及潜藏的问题进行改善，不但可以提高医院病案的管理效率，还能促进我国医疗事业的发展。从医院的内部管理角度分析，病历管理是日常管理中必不可少的环节，而在传统的管理模式中，病历管理以纸质方式为主，但是纸质病历档案所需时间较长、管理力度相对较大，记录过程容易出现失误导致记录结果出错，对医院的管理工作将造成负面影响。近几年来，伴随着我国医疗事业的不断发展与进步，医疗领域中的科技含量越来越高，电子信息技术因此在医疗领域中得到了充分的应用，有助于提高医院管理工作的效率以及质量。而电子病历档案是纸质病历与软件技术的结合，利用电子形式来记录、储存患者群体的就诊信息，可确保记录工作的客观性、真实性以及准确性，为医疗行业的发展提供了充足的依据，有助于提高医疗行业的效率与服务质量，还可以就此推医疗技术的发展[5-6]。

所谓病案，特指记录患者病情及临床治疗情况的详细记录，病案在疾病诊断与治疗中均有重要价值，因此加强病案的管理，是医院日常工作中的重要一环。随着我国社会经济的发展与进步，医院管理中逐渐涌现出了许多的新技术与思想观念，医院管理面临着重大的时代变革，而病案管理也应顺应变化趋势，对病案流程进行不断的优化，强化对病案的质量管控，以此来提升医院管理工作的效率与质量。本研究结果提示：经对比后发现，观察组的病历完整性、病历真实性、病历有效性及病历安全性等各个方面的管理效果评分均高于对照组（P＜0.05）；经对比后发现，观察组出现信息不完整、涂改、填写错误及缺页等问题的概率低于对照组（P＜0.05）。此结果反映，电子病历档案系统的应用将降低档案管理的差错率，同时还可以确保病历资料的完整性，有助于为临床诊疗工作、医疗科研工作等提供有效数据。电子病历档案系统所依托的是计算机系统软件程序，与传统将的纸质档案相比，全面记录了患者所有的就诊信息，尽可能地实现了病历资料的信息共享。应用电子病历档案系统后，可方便临床相关人员对病历信息进行随时的查阅，不再受时间与空间的限制，在很大程度上节省了医疗人员的时间，有利于开展诊疗工作的开展[7-8]。电子病历档案系统的应用，标志着医院管理工作得以实现现代化、多样化以及电子化，在很大程度上提高了医院管理工作的质量与效能。但是对于多数基层医院而言，电子病历档案系统仍旧处于初步部署及发展阶段，并未与医院实际情况融合，在实际应用中也面临着诸多问题，如管理职能落实不完善、监督力度不

够、病历回收与归档不及时等，数据的衔接与更新都存在局限性，导致电子病历档案系统的管理作用受到了影响，无法发挥出重要的管理作用。因此，部署或准备部署电子病历档案系统的医院，都应重视病案保护、构建附录系统、专人管理等措施，让电子病历档案系统发挥重要作用[9-10]。

综上所述，与传统的病历常规管理办法相比，在医院标准化管理流程中科学、合理地应用电子病历档案系统可有效缩短病历的借阅时间以及录入时间，可避免病历出现信息不完整等诸多方面的问题，具有理想的应用价值。

PO-042

面向社区居民的数字健康能力评估体系研究

谭政[1] 赵新蓉[1] 姚艺倍[2] 乔怡茹[2] 曹锦丹[2]
1.深圳市卫生健康发展研究和数据管理中心
2.吉林大学公共卫生学院医学信息学系

世界卫生组织（World Health Organization，WHO）发布的《数字健康全球战略（2020—2025）》[1]中，“数字健康”是指与开发和使用数字技术改善健康相关的知识和实践领域。涉及的内容包括更广泛的智能设备，使用智能连接设备的数字消费者，与物联网、人工智能、大数据和机器人技术相结合的健康服务等。显然，数字技术在健康管理方面的应用也涵盖在其中。从数字时代的发展来看，数字健康将深刻影响卫生健康行业和健康相关领域。

社区健康管理是基于管理理论和新健康理念，对社区健康人群、疾病人群的健康危险因素进行全面建档、监测、分析、评估、预测、干预、预防、维护，以及发展个人和家庭健康技能的全过程[2]。美国和日本的实践经验表明[3-4]，在社区实行健康管理，可有效防治80%的疾病。数字健康趋势下的社区健康管理和健康服务体系建设，需要以居民的健康信息素养和数字能力为基础，通过数字技术支撑，达到转变居民健康意识，提高健康决策能力的目的。然而，由于目前“数字鸿沟”的存在和社区居民类型的多样化，社区居民在使用数据技术的方式、获取健康信息的渠道以及认知健康信息的思维等方面都存在较大差异，进而影响其健康行为决策和健康管理。

本项目聚焦社区居民，在深入分析健康信息素养和数字能力要素基础上，以《数字公民素养框架》[5]和 DIGCOMP 2.2[6]为基本理论框架，构建面向社区居民的数字健康能力评价指标体系。旨在未来能将研究结果应用于社区居民数字健康能力评估提供参考，为增强社区在健康管理和数字化建设中的数据和信息协调能力、开展针对性的健康教育和健康干预、提高健康信息服务水平提供理论基础和建设启示。

1 资料与方法

1.1 社区居民数字健康相关要素分析

本文基于社区健康管理的概念和理论框架，结合现行常见的社区慢性病健康管理模式和全生命周期健康管理模式，特别是目前的疫情防控政策和措施等，深入分析探讨居民（健康消费者）与社区等多方联动共同助力健康管理和防疫的数字要素资源以及影响因素。参照社区健康管理逻辑和要点，结合健康领域的治理特点和社区居民群体特点[7]，从三个层面对居民数字健康要素进行提炼：

（1）目标：数字健康受数字化技术限定，其治理着力点并非宏观的健康、医疗、卫生等制度体系，而是这些制度体系中流动变化的信息、系统、数据，因此居民数字健康的主要方向是通过技术手段改善并提升健康服务的质量和效率。

（2）应用：数字健康主要是面向社区居民使用数字技术提升居民的健康素养及健康信息素养，应用的目的是提升居民的健康水平，促进人群主动健康素养。

（3）规范：规范是数字健康的底线[8]，社区居民要重视并妥善处理在数字技术与健康内容相结合的过程中所产生的伦理道德、信息安全、隐私保护等问题。

1.2 拟定社区居民数字健康能力评估指标

本研究对国内外社区居民数字健康能力评价指标的相关研究成果进行了全面收集和分析。①使用“健康素养、数字素养、信息素养、电子健康素养、健康信息素养、数字能力、媒体素养”（中文）及“Health literacy, digital literacy, information literacy, e-health literacy, health information literacy, digital competence, media literacy”（英文）为检索词，对中国知网、万方数据知识服务平台、Web of Science 数据库、PubMed 数据平台进行文献检索；②检索和查阅了中华人民共和国卫生健康委员会、中华人民共和国工业和信息化部、欧盟委员会网站、DQ 研究所等政府及有关机构网站；③参考和借鉴社区健康管理相关的文献内容；

④在此基础上对专家进行访谈。经过专家访谈和 5 次研究小组研讨，确定以《数字公民素养框架》和 DIGCOMP 2.2 为基本理论框架；⑥结合健康信息素养相关概念及评价工具以及社区居民数字能力和健康信息素养需求，对已有相关评估框架和能力标准进行了辨析，厘清相关概念之间的区别和联系，界定了面向社区居民的“数字健康能力”内涵和范畴，即“数字健康能力”是一种涵盖健康素养、数字素养、信息素养、电子健康素养以及数字能力、媒体素养的一种综合能力。在此基础上，构建了社区居民数字健康能力评估指标，共包括 4 个一级指标，12 个二级指标，33 个三级指标构成。

1.3 专家函询

（1）编制函询问卷

函询问卷包括四部分内容：1）引导语：包括本研究的背景及目的、问卷回收方式及时间、联系方式、致谢等。2）社区居民数字健康能力评估指标判定表：包括相关概念界定、填写说明以及三级指标的咨询表。3）专家情况调查调查表：包括专家的个人信息、联系方式、专业领域、技术职称等。4）专家对咨询内容的熟悉程度与判断依据调查表：包括专家对咨询内容的熟悉程度调查表以及专家对咨询内容的判断依据调查表。

（2）遴选函询专家

德尔菲法需要根据项目研究所需要的知识范围，确定专家、专门人员。专家人数的多少，可根据研究项目的大小和涉及面的宽窄而定，一般在 8-20 人左右为宜[9]。本研究专家纳入标准为：1）本科及以上学历，中级及以上职称；2）5 年及以上相关领域工作经验；3）自愿参加本研究，调查前对本研究领域有一定的了解，能提供具有参考价值的意见，并能够持续参与本研究直至专家函询结束。本研究共邀请了 15 位在相关领域具有一定知名度的专家及对社区居民的需求和能力比较了解的社区管理人员线上进行专家咨询。专家基本情况见表 1。

表 1　专家基本情况（n=15）

项目	人数	构成比（%）
性别		
男	1	6.7
女	14	93.3
年龄		
30~40	3	20.0
41~50	7	46.7
≥50	5	33.3
学位		
本科	2	13.3
硕士	2	13.3
博士	11	73.3
职称		
中级	2	13.3
副高级	6	40.0
高级	7	46.7
工作领域		
医学信息学	8	53.3
卫生事业管理	2	13.3
信息管理与信息系统	3	20.0
护理教育	1	6.7
工作年限		
≤10	4	26.7
10~20	3	20.0
≥20	8	53.3

（3）实施专家咨询

本研究于 2022 年 10-11 月对遴选专开展行了专家咨询。共进行了 2 轮专家咨询，第 1 轮的专家咨询问卷包括引导语、社区居民数字健康能力评估指标判定表、专家情况调查表和专家对咨询内容的熟悉程度与判断依据调查表四部分组成。本研究条目的纳入标准为[10]：同时满足重要性评分均数＞4.0、满分率＞20%、变异系数＜0.25，同时结合专家的修改意见，进行小组讨论后，进行条目的增加、修改或删减，研究小组根据函询结果修改问卷，形成第二轮专家咨询表。

1.4 统计观察指标

对两轮线上专家咨询过程中收集到的问卷的观察指标进行统计分析。专家积极系数用问卷的有效回收率和专家意见提出率表示。专家的权威程度用权威系数（Cr）表示，来自于专家的自我评价，由专家是学术造诣水平（q1）、专家对问题的判断依据（q2）和专家对问题的熟悉程度（q3）三个因素决定，一般认为专家权威系数 Cr≥0.8 即认为研究结果可靠[11]。专家的意见协调程度用变异系数（Cv）和协调系数(W)表示，通常认为 Cv＜0.25 是一个较好的指标，W 在 0-1 之间，W 越大表示专家的协调程度越好。一般地，经 2-3 轮咨询协调后，协调系数一般在 0.5 的范围波动，误差控制较好，Kendall's W 协调系数 P 值＜0.05 表明专家意见具有一致性[12]。数据采用 Excel 和 SPSS 19.0 软件进行处理。

1.5 计算指标权重

在得出最终版评估指标体系后，本研究使用 Yaahp12.9 软件绘制评估指标的层次模型图，根据专家函询对各个指标赋分的数据两两比较构建专家判断矩阵，得出各指标的权重。

2 结果

2.1 专家积极程度

本研究共进行了两轮德尔菲专家咨询，第一轮发出咨询表 15 份，回收 15 份，回收率 100%，意见提出率为 66.7% ；第二轮发出 15 份，回收 15 份，回收率为 100% ，意见提出率为 46.7%。这说明本研究所选专家关心本研究，参与的积极程度高。

2.2 专家权威系数

本研究专家学术水平系数（q1）平均值为0.993，表明专家学术水平较高，发表的意见较有价值；判断依据系数（q2）平均值为0.895，说明判断依据对专家的影响程度较高；熟悉程度系数（q3）平均值为0.583，表明专家对所咨询的问题较为熟悉。权威系数（Cr）为0.804，表明专家权威程度较高，咨询结果具有可靠性。

2.3 专家意见协调程度

第一轮专家函询的变异系数为0.00~0.43，第一轮所有条目的肯德尔协调系数W=0.292（X^2=196.204，P＜0.05），W值在0.2~0.4之间说明一致性程度一般。第二轮专家函询的变异系数为0.00~0.19，第二轮所有条目的肯德尔协调系数W=0.514（X^2=254.174，P＜0.05），W值在0.4~0.6之间说明一致性程度较高，误差控制较好。由此可见，经过两轮函询后，专家意见趋于一致，对各指标认可度高。

2.4 专家函询结果

（1）第1轮专家函询结果

根据条目纳入标准和专家修改意见，经过研究小组讨论，在第 1 轮专家函询后对评估指标具体做了如下改动：

一级指标：将“健康信息/数字素养”修订为“健康信息素养”；将“情感与沟通”“数字保护”合并为“健康数字素养”；将“健康问题解决”修改为“健康素养”。此外，根据专家意见对这三个一级指标的内涵和外延进行了具体界定。

二级指标：对一级指标“健康信息素养”下的三个二级指标名称进行进一步归纳概括；删除二级指标“数字身份管理”；合并二级指标“数字保障”到“数字安全”；合并二级指标“数字参与“到“数字交互”；将二级指标“确定健康需求”修改为“健康需求”，将二级指标“健康决策”拆分为“健康决策”和“健康行为改变”两个二级指标，最终共得出10个二级指标，此外，根据专家意见对这10个二级指标进行了具体界定。

三级指标：删除“能浏览检索结果/能随意（即：无具体信息需求目标）浏览软件或网页中的信息”、“能从检索结果中筛选出答案”；删除二级指标“数字情感”中“能重视人的尊严和人权，在数字社会同样尊重自己和他人”、“能重视在线文化多样性，尊重不

同在线文化”、“能重视在线民主、正义、公平、平等和法治”、“在线包容和同情心”。

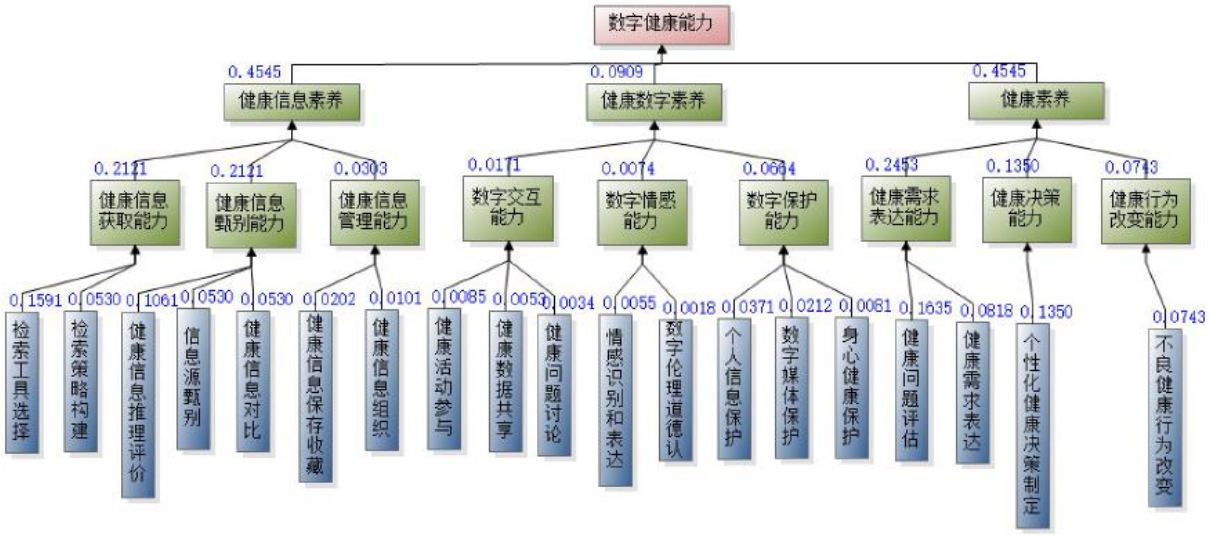
（2）第 2 轮专家函询结果

对第 2 轮的专家函询结果进行分析，根据本研究的评估指标纳入标准，二级指标“数字情感”“健康信息管理”和三级指标“社区居民在数字环境中寻求健康帮助时，能正确采用网络礼仪”、“社区居民在数字环境中能够合理、规范的使用数字工具，避免因姿势的不规范或者时间的不节制等对身体健康造成不良影响” 4 个条目不符合纳入标准，结合专家意见对该 4 个条目进行重点分析。专家指出“网络礼仪”一词表达不准确，通过小组讨论和参考数字公民素养量表，将“网络礼仪”修改为“数字伦理道德”。专家指出二级指标“数字安全”、“数字健康保护”中的数字安全和数字健康保护难以区分，根据专家意见结合组内讨论结果，合并二级指标“数字安全”、“数字健康保护”为“数字保护”。专家指出二级指标“数字安全”和“健康信息甄别”下面均涉及虚假信息识别，根据专家意见结合组内讨论结果，删除三级指标“社区居民在数字环境中能准确识别虚假信息，防止被骗取财物”。根据专家意见结合组内讨论结果合并三级指标“社区居民在数字环境中寻求健康帮助时，能够避免数字环境中可能的风险（如网络欺凌），保护自己和他人的心理健康”和“社区居民在数字环境中能够合理、规范的使用数字工具，避免因姿势的不规范或者时间的不节制等对身体健康造成不良影响”。根据专家意见对二级指标的命名表达进行调整——在二级指标后均加上能力。此外专家指出三级指标的表达需要精练，经过小组讨论后对三级指标的表达进行了统一精炼和释义。

2.5 层次分析结果

本研究使用 Yaahp12.9 软件绘制评估指标的层次模型图，根据专家函询对各个指标赋分的两两比较数据构建专家判断矩阵[13]，得出各指标的权重如图 1。

图 2 数字健康能力指标权重



最终形成社区居民数字健康能力评估指标，含一级指标 3 个，二级指标 9 个，三级指标 19 个，详见表 2。

表 2 社区居民数字健康能力评估指标

指标	重要性赋值		权重
	评分（$\bar{x}\pm s$）	变异系数	
1 健康信息素养	5.00±0.00	0.00	0.4545
1.1 健康信息获取能力	4.93±0.25	0.05	0.2121
1.1.1 检索工具选择	4.40±0.61	0.14	0.1591
1.1.2 检索策略构建	4.20±0.65	0.16	0.0530
1.2 健康信息甄别能力	4.93±0.25	0.05	0.2121
1.2.1 健康信息源甄别	4.87±0.34	0.07	0.0530
1.2.2 健康信息推理评价	4.93±0.25	0.05	0.1061
1.2.3 健康信息对比评价	4.87±0.34	0.07	0.0530
1.3 健康信息管理能力	4.00±0.52	0.11	0.0303
1.3.1 健康信息保存/收藏	4.07±0.57	0.14	0.0202

1.3.2 健康信息组织	4.00±0.52	0.13	0.0101
2 健康数字素养	4.33±0.47	0.11	0.0909
2.1 数字交互能力	4.13±0.81	0.19	0.0171
2.1.1 健康活动参与	4.27±0.68	0.16	0.0085
2.1.2 健康数据共享	4.20±0.40	0.10	0.0053
2.1.3 健康问题讨论	4.13±0.62	0.15	0.0034
2.2 数字情感能力	3.80±0.65	0.17	0.0074
2.2.1 情感识别和表达	4.00±0.63	0.16	0.0055
2.2.2 数字伦理道德认同	3.80±0.54	0.14	0.0018
2.3 数字保护能力	4.73±0.57	0.12	0.0664
2.3.1 个人信息保护	4.60±0.49	0.11	0.0371
2.3.2 数字媒体保护	4.47±0.50	0.11	0.0212
2.3.3 身心健康保护	4.20±0.40	0.10	0.0081
3 健康素养	5.00±0.00	0.00	0.4545
3.1 健康需求表达能力	5.00±0.00	0.00	0.2453
3.1.1 健康问题评估	5.00±0.00	0.00	0.1635
3.1.2 健康需求表达	4.87±0.34	0.07	0.0818
3.2 健康决策能力	4.87±0.34	0.07	0.1350
3.2.1 个性化健康决策制定	4.80±0.40	0.08	0.1350
3.3 健康行为改变能力	4.80±0.40	0.08	0.0743
3.3.1 不良健康行为改变	4.67±0.47	0.10	0.0743

3. 讨论

3.1 社区居民数字健康能力指标体系的必要性

社区健康管理以健康为中心，目的是改善社区居民健康状况，提高社区人群健康水平[14]。社区健康管理涉及多个层面。其中，最核心也是最基本的要素就是社区健康管理和服务的对象，即社区居民。社区居民的数字健康素养和能力直接关系到数字健康社区的建设以及社区健康管理水平和质量。然而，由于社区居民类型的多样性，数据健康方面的接受度、以及数据技术采纳和获取健康信息技能、健康信息的认知和评价等方面都会有较大差

异。如何准确把握不同社区居民的数据健康能力，通过数据技术为社区居民健康提供精准的健康管理和服务，是目前社区数字化服务和建设中应面对的一项工作。本研究构建的社区居民数字健康能力指标体系，目的是将其结果应用于社区数字化健康管理，从而提高其管理水平和社区居民的数字健康能力，为社区在健康管理和数字化建设中面向社区居民开展针对性[12]的健康教育和健康干预提供理论基础和建设启示，研究具有重要的现实意义和应用前景。

3.2 社区居民数字健康能力指标体系的构建具有科学性

本研究在实施专家函询前系统分析了大量的国内外文献、政府政策文件等，结合专家访谈、6 次小组会议结果，确定了社区居民数字健康能力指标体系结构和条目。共进行了 2 轮专家函询，所选专家涵盖医学信息学、社会医学与卫生事业管理、信息管理与信息系统等领域，覆盖学科范围广，专家建议具有良好的全面性和指导性。专家平均权威系数（Cr）为 0.804，学术水平、判断依据、熟悉程度均较高，表明专家对条目的评价科学、客观，研究结果可靠。两轮专家函询的问卷回收率均为 100%，表明专家参与此次研究的积极性较高。同时，本研究运用层次分析法，通过比较同一层级指标的相对重要性来综合计算指标的权重系数，将决策者的经验判断定量化，增加了决策依据的准确性。

分析两轮的专家函询结果可知，第二轮函询的肯德尔和谐系数为 0.514（$P<0.01$），显著高于第一轮的 0.292（$P<0.01$），说明专家的协调程度较高，专家意见基本趋于一致。在分析各个指标的权重时本研究使用 Yaahp12.9 软件，根据所有专家对每项指标打分的平均值确定。所有判断矩阵的一致性比例数据均小于 0.1，表明指标构建具备可靠性。

3.3 社区居民数字健康能力指标体系的内容及权重

社区居民数字健康能力评估体系的一级指标中，“健康信息素养”、“健康素养”所占权重相同且较高（0.4545），“健康数字素养”所占权重较低（0.0909）。

“健康信息素养“下有三个二级指标，其中“健康信息获取能力”、“健康信息甄别能力”所占权重相同且较高（0.2121）。目前网络以及相关的新媒体已经成为社区居民获

取健康信息的主要来源，在面对海量的健康信息时，社区居民对健康信息的甄别能力尤为重要。

“健康数字素养“下包含三个二级指标，其中“数字保护能力”权重相对较高（0.0664），该条目对社区居民在数字环境中寻求健康帮助时提出了新的要求：要充分保护个人信息，防止因操作习惯不当或对对信息源的甄别不当等原因导致个人信息的泄露，给自己和他人带来困扰；在使用数字媒体进行健康服务时要保护数字媒体，适当具备清内存、软件查杀等能力；在数字环境中要保护好自己和他人的身心健康，避免因姿势的不正确或者时间的不节制对身体造成伤害，避免在网络环境中网络欺凌以及被欺凌等。

“健康素养”中包含三个二级指标，其中“健康需求表达”权重相对较大（0.2453），在网络环境中进行健康服务时，能准确评估和表达健康需求是前提更是基础，只有准确的评估和表达了健康需求，才可以根据需求进行针对性的健康服务，才能制定个性化健康决策，从而促进健康行为的改变。

4.局限性与展望

本研究运用德尔菲法，基于数字健康能力相关领域专家的函询意见，确定了社区居民数字健康能力评估指标。但是，由于专家人数、专业限制以及德尔菲法本身的缺陷，可能会对社区居民数字健康能力评估指标的构建形成人为偏倚，后续研究中将进一步扩大范围，展开更为深入的调查分析。此外，本研究通过文献分析、访谈、专家咨询以及层次分析法等质性和量性方法制定出了较为严谨的评估指标体系，构建过程和结果均具有一定的科学性和可靠性。但是受研究时间和疫情的影响，本研究尚没有应用评估指标进行实证调查，后续将通过一定样本的调查，进一步检验此指标的可行性和敏感性，从而进一步提高评估指标的应用性。

PO-043

OBE-CDIO 理念下面向医学信息人才创新实践能力培养的教学模式优化路径研究

张新颖[1] 张治国[1,2]沈丽宁[1,2]
1华中科技大学同济医学院医药卫生管理学院
2 湖北省卫生技术评估研究中心

1 引言

在当今数字化和智能化的时代，信息技术与医疗领域的融合不断推动着医疗服务的创新与升级。由此引发了医学信息领域高等教育人才培养理念、方式和治理体系的系统性变革，对学生所拥有的知识、技能与能力提出了新的要求。教育部的新医科建设以及《国务院办公厅关于加快医学教育创新发展的指导意见》（国办发〔2020〕34 号）[1]、《“十四五”卫生健康人才发展规划》[2]等文件中，对医学信息人才培养提出了适应时代和国家发展的新期望和新要求。在健康中国和数字中国建设大背景下，在面向人民生命健康新方针的引领下，高校急需培养大批具有创新思维、批判思维和应用实践能力的医信交叉的高质量复合型人才，使之能够熟练运用交叉学科知识，解决现实世界中医学领域前沿问题，以更好地应对“以医文、医工、医理等交叉学科支撑的医学教育新模式”的转变[3]。

医学信息人才作为集医学、管理学和计算机学于一体的交叉学科复合型人才，培养其创新实践能力和批判性思维在高校人才培养中有着越来越重要的地位。有院校提出“双金字塔”型医学信息人才全链条培养模式，从“本-硕-博-职”四阶段和“医学知识、信息素养”双金字塔层面进行人才培养体系构建，以期打破学术型与专业型医学信息人才教育壁垒问题[4]。有院校通过构建“三进两融合”应用型人才培养模式，以达到将信息技术与医学、管理学知识进行深度融合,校企协同育人的目标[5]。有的院校开展基于 CIPP 模型（Context evaluation-Input evaluation-Process evaluation-Product evaluation）的医学院校信息管理专业“T”型人才培养模式构建，使得医学信息人才具备扎实的医学知识和计算机技术能力，同时在实践操作和动手能力上表现突出[6]。目前，医学信息人才培养相关研究多从培养体系

和培养模式展开，极少有研究将侧重点集中在教学模式上，同时也并未强调创新实践能力。但在新工科和新农科领域，有针对提升学生创新实践能力的教学探究[78]，而在医信交叉学科领域，尚缺乏在相关理念指导下系统性的研究，尤其是将对象定位为本科生的研究。教学模式作为一种相对固定的教学活动结构和程序，其建立基于特定的教学思想或教育理论[9]，用于组织和传递教育内容以实现特定的教育目标，通常包括课程体系、教学资源、教学方法、评估方式以等，不同的教学模式适用于不同的教育场景和教学目标。在人才培养中，教学模式发挥着至关重要的作用，它有助于将教育理念转化为切实可行的教育活动，确保学生能够达到培养方案所设定的目标和要求。因此，本文拟进行基于 OBE-CDIO 理念的面向本科医学信息人才创新实践能力培养的教学模式优化路径研究。

OBE-CDIO 理念是将成果导向教育（Outcomes-Based Education，OBE）理念融入进代表构思-设计-实现-运作(conceiving-designing-implementing-operating，CDIO)的工程教育模式理念，用以构建高校创新型、实践应用型专业人才培养模式。OBE 强调结果，以学生达到预期毕业水平为核心任务，开展以学生为中心的课程体系教学改革[10]；CDIO 强调教学过程和架构，以项目为载体，采用“学生主导、教师引导”的教学方式，解决理论知识与实践技能学习之间的不平衡问题，将知识积累、个人能力培养以及综合素质提升紧密结合，使得理论学习、实践操作和创新思维融为一体[11]。OBE-CDIO 理念以成果为导向，通过将教学目标与教学过程紧密结合，使学生能够独立地构思、设计、实施和运营真实项目，并以教学效果反馈不断完善教学目标，这一理念在培养创新实践能力和批判性思维方面具有巨大的潜力，见图 1。

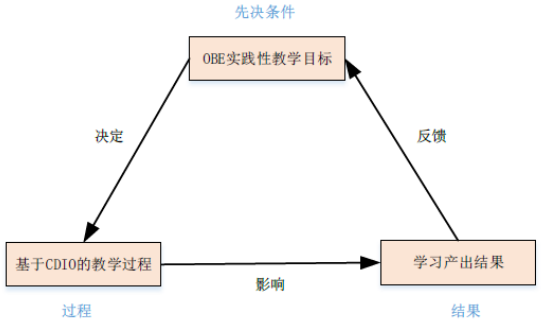


图 1 OBE-CDIO 理念设计

本文旨在通过深入分析医学信息人才培养现状和存在问题，探讨基于 OBE-CDIO 理念优化适应医学信息人才培养的教学模式，寻找更加贴近实际、更具创新性的教学路径，以提升医学信息相关专业本科生的创新实践能力，从而满足医药卫生行业对于人才创新实践能力的迫切需求。

2 医学信息人才创新实践能力需求分析

2.1 医学信息人才培养现状与问题

我国目前已有 40 余所高校开展本科医学信息人才培养相关教育。2003 年，中南大学首次获批医学信息学本科专业。目前国内开设医学信息学本科专业的高校较少,大多是通过设立相近专业的方式开展医学信息专业相关教育 [12]。根据教育部最新发布的《2022 年普通高等学校本科专业目录》,相近专业命名多样，如信息管理与信息系统(120102)、信息资源管理(120503)、健康服务与管理（120410T）、生物信息学(071003)、生物医药数据科学(101012T)、公共事业管理(医学方向)(120401)等 [13],主要培养跨学科的医信管复合型专业人才，使之不仅拥有医学基础知识和信息科学基础知识，还具备广泛的计算机科学技术知识、实践应用能力和创新批判思维。

在信息技术蓬勃发展，医学信息领域处于快速的数字化和智能化转型背景下，医学信息人才培养显现出一些问题。传统的教学模式往往以传授知识为主，学生在技术操作和实际应用能力上存在较大差距，对于培养学生的创新实践能力和跨学科协作能力存在一定局限。卫生健康行业对人才的需求已经不再仅仅局限于技术层面，而是更加强调学生能够将所学知识灵活应用于实际场景，解决医疗实践中的问题。尤其是在数字化医疗、健康管理、医疗大数据等领域，越来越需要具备创新实践能力的医学信息人才，他们能够从医学与信息学的双重视角出发，通过独立思考和团队合作等方式，有效解决现实世界中的复杂医疗信息相关难题。

2.2 医学信息人才创新实践能力培养的迫切需求

随着社会老龄化以及人们对健康管理的不断关注，5G、大数据、云计算、人

工智能等新兴技术衍生出的影像诊断、远程医疗在医疗行业的不断应用，智慧医院建设的不断推进，医学信息领域迎来了前所未有的机会和挑战，对医学信息人才的能力要求重点逐渐从理论知识层面转移到创新实践层面。基于此背景，医信管交叉学科有机结合，发挥多学科优势，培养具备创新实践能力的医学信息人才显得尤为迫切。首先，创新实践能力能够使他们在面对复杂多变的医疗问题时能够灵活应对，提供创新性解决方案，应是医学信息人才必备的核心能力之一。其次，创新实践能力的培养有助于弥补医学信息领域教育中的实际应用短板，使学生能够更好地将理论知识转化为实际操作。此外，创新实践能力的培养也能够提升医学信息人才在跨学科合作中的表现，增强其与医生、技术专家等不同领域人才的有效沟通与合作能力。要想培养医学信息人才的创新实践能力，需从教学的根本——教学模式着手，以往“重理论，轻实践”的教学模式在创新实践能力培养方面还很薄弱，如何利用有效的教学模式将教学内容与学生实际生活、社会现实等连接起来，真正激发学生学习的兴趣和创新实践能力将起着关键作用。

因此，我们迫切需要优化医学信息人才的培养教学模式，从而更好地培养创新实践能力。基于 OBE-CDIO 理念的教学模式可以有力地弥补当前医学信息教育的不足，使学生能够更加全面地掌握医学信息知识，更加熟练地运用技术工具，更加灵活地解决实际问题。这种教学模式将有助于培养具备创新思维和实践能力的医学信息人才，为医学信息领域的发展注入新的活力。

3 基于 OBE 理念的教学目标明确

教学目标对于教学具有极其重要的作用，是教学过程的核心，影响着教学的质量和效果。教学目标能够引导教师选择适当的教学模式，以确保对学生的培养具有针对性。同时通过定期检查和评估教学目标的实现情况，教师可以更好地了

解教学模式的效果，发现问题并进行改进。

明确教学目标是基于 OBE-CDIO 理念的教学模式构建的第一步。以 OBE 理论为指引，明确遵循医学信息行业需求导向，以学生职业生涯发展规划为宗旨，最大程度保证教学目标与教育结果的一致性[14]。基于学生、教师、用人单位三方视角，广泛调研专业学生就业需求，医院、医疗软件企业、高校、卫生行政部门等用人单位对医学信息专业人才的创新实践能力要求，并对收集的资料进行整理和分析，拟归纳总结出四个方面的专业能力分解：**基础能力、实践能力、创新能力和思政能力**，据此进一步细化制定科学合理的医学信息人才培养目标，并生成服务于培养目标的教学目标：知识目标、技能目标、创新目标和素质目标，为后续构建教学模式起到指导作用。

4 基于 OBE-CDIO 理念的教学模式优化路径探索

教育领域不断发展和变化，教学模式在其中扮演着至关重要的角色。选择合适的教学模式可以提高学生的学习效果，不同的教学模式可以培养学生的多样化技能和综合能力，创新的教学模式可以激发学生的兴趣。教学模式的选择和实施需要根据不同的教学场景和教学目标进行调整，以确保教学活动的有效性和成果。总之，教师应认识到教学模式的重要性，积极探索和采用创新的教学方法，以满足学生的需求，提高教学质量，培养具备综合能力的学生，促进教育领域的发展。

如何针对医学信息人才创新实践能力培养目标，构建相适应的教学模式，是一个值得深思的议题。OBE 所强调的成果导向和 CDIO 所倡导的“做中学”的教学方法与学生创新实践能力培养高度契合。以培养医学信息人才的创新能力和实践能力为出发点，基于 OBE 理念反向设计教学目标，并融合 CDIO 工程教育模式，围绕课程体系、教学资源、实施过程和效果评价等方面，优化构建规范化、系统化的“四位一体、四要素、三方面、一中心”教学模式，见图 2。

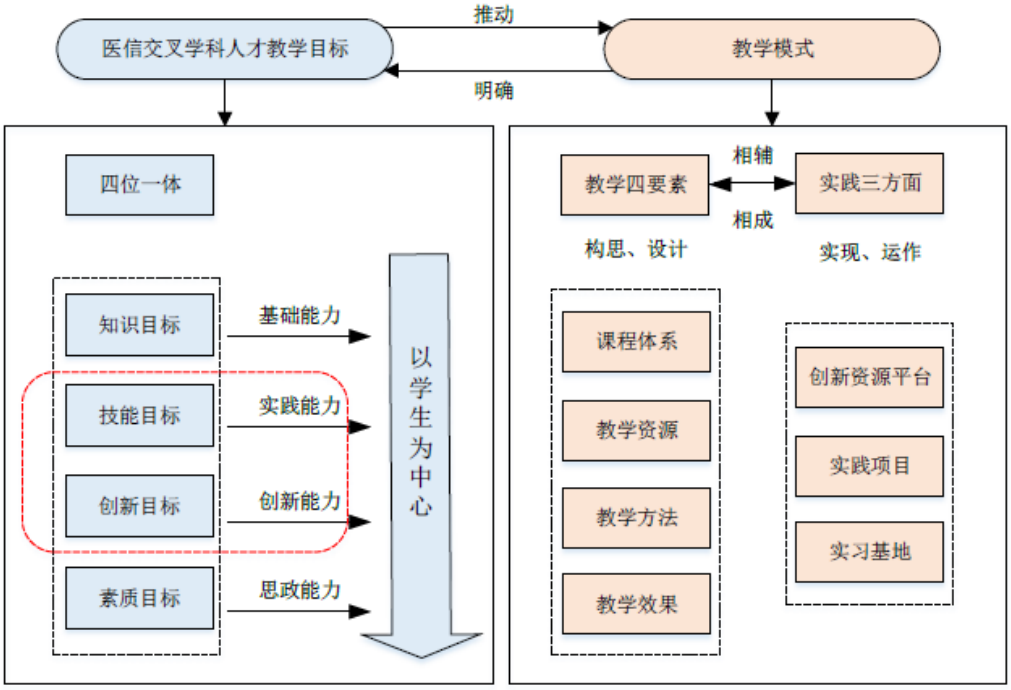


图 2 教学模式优化路径

4.1 重构实践课程体系

完善课程体系的设置，是实现教学目标的重要保障，是优化教学模式的基石。针对每类课程在课程体系中的定位要求,首先以规范化课程教学大纲的形式，明确每类课程的教学目标及其对创新实践能力培养要求的支撑关系,进而规定课程内容、教学安排、教学方法、课程考核与成绩评定等各项细节[15]。好的教学大纲能够传达完整的教学模式，将教学目标和教学内容有机融合，让学生明晰需要参与哪些活动以及完成哪些任务[16]。

以培养医学信息人才创新实践能力目标为导向，面向行业发展需求，基于 OBE-CDIO 理念完善实践课程和各类课程的实践教学大纲，深化课程的产教融合设计。对于实践课程来说，首先，明确课程的总体目标，然后细化为具体的学习成果，基于学习成果，设计具体的实践项目或任务，在大纲中明确每个学习目标和实践项目的时间分配，同时设计反馈机制，以便学生在实践过程中获得指导和反馈。对于课程的实践教学环节来说，针对编程类课程，如计算机及程序设计基础，强调项目驱动的学习，引入医药软件企业或医院的可用编程项目锻炼学生的实操能力；针对信息管理类课程，如管理信息系统，强调结合现实问题，引入医

院健康信息管理真实案例激发学生创新和批判性思维；针对交叉学科类课程，如卫生管理统计学，强调运用多学科知识解决问题，引入案例分析提升学生将理论知识与实际应用相结合的能力；针对实践环节，如毕业实习，强调专业技能融入实际工作，提供实习基地以及相对应的岗位，在实习工作中巩固专业技能、拓宽专业视野、明确职业规划。

4.2 建立教学资源支持体系

4.2.1 教材和教学案例

教材作为教学过程中最基本的工具，发挥着知识载体的作用，也是深化教育教学改革、全面推进素质教育、培养创新实践人才的重要保证[17]。在进行教学模式优化时，要注重教材与新教学模式的适配，促使教材适应新型教学模式并发挥应有作用[18]。针对医学信息人才创新实践能力的培养，在进行教材建设时，首先考虑多学科协同，引入跨学科内容，帮助学生掌握基础知识；其次整合医学信息领域实际案例，展示行业内的真实问题和挑战；然后强调实践操作，采用项目驱动的任务或项目提案，配套操作指南、步骤和示例，以提升学生实操技能，并鼓励其提出创新方案；最后定期更新内容，引入行业相关最新技术和成果，紧贴行业能力需求，避免与行业脱节。

教学案例可以为教学模式优化提供教学内容支撑[19]。教学案例通常基于现实问题，学生通过分析和解决案例问题，能够将理论知识应用到实际情境中，同时能够激发学生的创新思维，鼓励他们提出新的解决方案，挑战传统观念，促进实践能力和创新精神的培养。可构建教学案例库，广泛吸收优秀教学案例，或自行撰写特色教学案例，用以支撑案例教学，达到 “授知识、讲案例、重实践”的效果，促进学生主动吸收理论知识并结合现实应用场景实现知识融会贯通、职业素养内化构建，为创新实践能力的提升奠定扎实基础。

4.2.2 师资队伍

在面向医学信息人才创新实践能力培养的教学模式中，师资队伍至关重要。通过不断提升教师的专业素养、教学水平和实践经验，能够更好地满足学生的需求，有效地引导和培养医学信息人才的创新实践能力。对于教师个体，应积极跟踪医学信息领域最新发展，参与相关项目和研究，积累实践经验，使其具有能够指导学生完成实际项目和案例研究的能力，后续便可开展以“学生为主体，教师

为引导”的项目导向的教学，在教学中培养学生创新实践能力；对于教学团队，应该具备多学科背景、实际工作经验、、专业技能和教学能力。在教学团队建设过程中，一是要发挥老教师的“传、帮、带”作用，注重青年教师的培养；二是要强化“双师”教师队伍的建设，将校外导师制度与校内教师团队有机融合，校外教师掌握行业内的最新动态、前沿技术和研究成果，能够让学生接触到最新的知识和趋势，带来不同的视角和观点，拓宽学生的思维和知识领域，同时校外教师拥有丰富的实践经验，能够与学生分享真实的行业情况和实际操作技巧。有效推进教学团队与行业共同制定教学方案和开设课程的实践环节，符合新医科、新文科、医管及医信交叉的专业发展需求，是教学改革创新理念的实践路径。

4.2.3 实践、实习基地

实践、实习基地能够为学生提供现实世界中的实践机会，让他们将课堂知识应用到实际情境中，提升解决实际问题的能力。通过实践实习，学生可以与行业专业人士接触，了解医学信息领域的最新趋势和实际工作流程，积累宝贵的行业经验。同时实践实习也能鼓励学生在面对实际挑战时寻找创新解决方案，提升实践技能水平和创新应用能力。以创新实践能力培养要求为依据，明确实践、实习基地的建设要求，积极打破校内和校外资源壁垒，协同高校、医院、卫生行政部门和医疗软件企业等多方力量，构建校内外发展互补、共享增效的实践、实习基地，进行实景式体验教学，实现多方协同育人，推动行业企业全方位参与人才培养，着力提升医学信息人才创新实践能力，使其更好适应行业就业需求。

4.3 优化教学实施过程

教师在备课过程中围绕教学目标，研究制定切实可行的教学方法。通过引入新颖的教学方法、充实课程内涵、优化教学设计以及改进课堂管理等手段，把“提升学生实践应用能力”与“创新批判性思维”贯穿人才培养的全过程，坚持价值塑造、知识传授和能力培养有机融合，实现全员、全程、全方位的人才培养。教学实施过程从以下三个步骤展开。

（1）教师在基于 OBE-CDIO 理念的教学模式指引下，以产出成果“创新实践能力培养”为导向，为学生分配实践任务，按照“构思-设计-实现-运作”四个阶段，实施项目式教学，见表 1。例如在开展管理信息系统课程实践教学时，以医院挂号系统功能为导向，利用案例情境引出课程教学目标，教师发挥引导作用，

学生进行实际问题的分析和解决，期间可进行小组讨论和互动，也可利用“雨课堂”等线上学习平台提供补充资料和相关知识的讲解，最后可启发学生提出创新解决方案。

表 1 CDIO 教学过程四阶段

阶段	目标	方法
研究与构思	相关概念引入	教师案例情境教学
独立设计	培养创新思维和团队协作精神	小组讨论+教师指导
自主开发与实施	培养实践动手能力	小组合作+教师指导
操作与应用	培养独立解决问题能力	小组汇报+教师评议

（2）教师团队在学期末根据教学实施情况及效果反馈进行总结分析，形式包括讨论会议、专家研讨会等，基于教学实施过程中发现的问题进行讨论，同时邀请业内专家，对教学模式展开论证和评估，在相关建议下优化教学模式。

（3）教师根据优化后的教学模式继续开展实践教学，收集整理学生产出成果和实践实习反馈意见，如学生参加各类学科竞赛和创新创业项目所获奖励、医院或企业实习个人评价等，通过分析相关产出评判教学目标的达成度。

教学实施过程中的关键点是将 OBE-CDIO 理念融入课程的实践环节中，以产出为导向，秉持以学生为中心的理念，以案例引导、项目驱动为侧重点，突破传统课堂教学方式，采用翻转课堂(Inverted Classroom)结合 O2O(Online To Offline)教学的混合教学方法，同时充分利用第二课堂和实践实习基地特点，开展实践教学，达成教学目标。翻转课堂将课堂时间用于讨论和实践，而将课程内容的传达和理解放在课前，学生课前通过在线学习资源（如视频、文献等）预习课程内容，课上便可与老师和同学们讨论、解决问题，增加了更多的互动交流和实践机会。O2O 教学将线上学习和线下实践相结合，教师可以利用在线平台提供课程材料、练习和讨论，同时在线下进行实际操作、项目工作和小组讨论，这种方法可以增强学习的灵活性和便捷性。第二课堂能够丰富学生的知识面，实践实习基地能够拓宽学生的行业视野，二者均能激发学习热情和创新思维，鼓励学生参与创新实践项目，保障基于 OBE-CDIO 理念的以创新实践能力培养为目标的教学模式的有效实施。

4.4 重视教学效果评价

基于 OBE-CDIO 理念的面向医学信息人才创新实践能力培养的教学模式，应采用多元化的评价方式，综合考量和评价学生的学习成果[20]。对于实践课程，根据教学目标，建立与之相适应的“多元多维多层”的课程考核评价体系：“多元”即课程成绩组成多元化，包括出勤情况、分组汇报、实践报告、期末考试等；“多维”指在评定学生各项成绩时,从知识、能力、素质多个维度进行考核，即从知识掌握程度、实践应用能力和思想素质等方面进行综合分析,尊重学生的个性，培养学生的创新实践精神和批判性思维；“多层”指的是在考核评价时结合学生的课堂参与度、发言积极性、问题解决贡献和课后自主学习情况等方面综合分析。在课程考核方式上摒弃以往以试卷为载体的单一理论考试，更多的是考核学生在课堂上对案例的分析、小组讨论的结果、实践项目完成的情况等，以此来评估学生对理论知识掌握程度、创新实践能力、团队协作精神以及批判性思维和表达能力。采用这种多维度的考核方式，不仅可以更全面地反映学生的学习效果和能力，帮助学生更好地理解和应用所学知识，还可以激发学生学习的积极性，提升他们的创新实践能力，培养具备综合素质的高质量人才，也能为教学模式的优化提供更细节的反馈。

除此之外，对于实践环节，与实践实习单位共同制定实习考核评价体系，使用多种评估方式，包括实际项目成果、口头报告、书面报告、同事评价、自我评价等，以全面评估学生的实践应用能力。基于此，教师也能更加明确用人单位对学生创新实践能力的期望和需求，以及学生自身对创新实践能力培养的反馈，以此不断优化完善教学模式，使人才培养更符合市场需求。

5 结语

在数字中国和健康中国的背景下，医学信息领域正迎来前所未有的机遇和挑战，对于医学信息人才的需求逐渐迫切，其培养显得尤为重要。本文旨在探讨基于 OBE-CDIO 理念的教学模式优化在培养医学信息人才创新实践能力方面的应用，为培养具备跨学科、创新思维、实践应用能力和团队协作能力的高质量医学信息人才提供有力支持，以适应医学信息领域快速发展的需求。

在基于 OBE-CDIO 理念的教学模式中，教师的角色从传统的知识传授者转变为学生的引导者和支持者，教师提供适当的指导和支持，帮助学生克服困难，解决问题。同时，教师通过鼓励学生自主学习，培养他们独立思考和解决问题的

能力。基于 OBE-CDIO 理念的教学模式将更加全面地培养医学信息人才的创新实践能力，这种教学模式不仅能够使学生从理论到实践无缝衔接，还能够让他们在解决实际问题的过程中不断成长和发展，为数字化医疗、健康管理等领域的发展注入新的动力。此外，教学模式的持续改进和创新也将为医学信息领域的教育提供新的思路和方法。

书面交流

PU-001

DRG 支付制度背景下医生的非预期行为及有效应对

颜新宇[1,3,4,5]，常维夫[1,3,4,5]，罗爱静[2,3,4,5]

1. 中南大学湘雅第三医院，湖南长沙，410013

2. 中南大学湘雅第二医院，湖南长沙，410013

3.医学信息研究湖南省普通高等学校重点实验室(中南大学)，湖南长沙，410013

4.中南大学生命科学学院，湖南长沙，410013

5.湖南省心血管智能医疗临床医学研究中心，湖南长沙，410013

摘要：基于疾病诊断相关分组（Diagnosis Related Groups, DRG）的医保支付制度能有效调节医疗服务和医保支付中的信息不对称，控制医疗费用不合理增长。但是，信息不对称问题并未彻底解决，经扎根访谈与数据分析结合研究发现，医生在 DRG 支付制度框架下仍会实施非预期行为（unintended response），包括高编码（up-coding）、挑选患者（cherry-picking）、分解住院（readmission）。承担着更多的社会责任的大政府管理模式虽然具备主观能动性的优势，但在治理环节存在劣势。为有效规避医生的非预期行为，应积极建立多方参与、平等沟通的协商机制，并增强技术手段以加强对 DRG 支付制度下非预期行为的识别。

关键词：DRG 支付、非预期行为、信息不对称、高编码、挑选患者、分解住院、协商机制

背景

全球各国普遍实施的医疗保险制度，是为了应对公民个人无法承担过高的医疗成本而设计的，它通过筹集全体公民的资金形成一个巨额资金池，并在公民个人产生医疗需求时为其提供资金支持。医疗保险制度在世界范围内普及以来，有效地降低了公民个人承担巨额医疗费用的风险，提高了公民的生活质量。但是，由于医疗服务、医保支付过程中存在的信息不对称，患者和医生因此产生了从医保基金中获取更多个人利益的想法，最终呈现出的结果是医疗费用飙升，医保基金使用效率降低[1-2]。

为了抑制医疗费用的高增长，各国尝试采取措施规避医疗服务和医保支付中的信息不对称。其中，根据消耗的医疗资源对疾病进行合理分组的 DRG（Diagnosis Related Groups）分组方式、将疾病分组结果关联到医保基金支付的 DRG 支付，是规避医疗服务和医保支付中的信息不对称的有效方案。这项诞生于美国的医保支付方式，经过实践验证可以有效减少医疗服务和医保支付中的信息不对称，降低医疗费用的增长速度[3-4]，也因此，DRG 支付方式在全球得到广泛应用。

但与此同时，DRG 支付方式并不能完全消除医疗服务和医保支付中的的信息不对称。处于信息不对称优势方的医生仍可能在 DRG 支付制度背景下，实施基于 DRG 支付制度框架下的非预期行为（unintended response）[5-8]，例如未治愈出院（early discharge）（在患者尚未痊愈但处于低风险的时令其提前出院），挑选患者（cherry-picking）（在收治病人阶段倾向于挑选医保支付额减去医疗成本后盈余更高的病例）等，这些非预期行为造成了诊疗服务效果下降、医保基金浪费等负面后果。研究发现，非预期行为广泛地存在于实施 DRG 支付制度的国家中。

区别于那些率先实施 DRG 支付的国家，中国在文化背景和医疗体系上具有不同的特征，一是长期作为发展中国家的中国，医疗系统及医保支付系统的发展相对缓慢，上个世纪末 DRG 支付制度在欧美国家开始普遍实行的时候，中国尚处于从无到有地建立全民医疗保险制度的阶段。DRG 支付制度的正式实施也是在近几年才开展，直到 2014 年，北京才首次正式应用 DRG 分组用于绩效评价，2019 年国家医保局在全国 30 个城市开展 DRG 支付制度改革试点工作，标志着中国开始大规模地推进 DRG 支付制度普及工作。二是中国的社会管理具有政府高权限高责任的特点，相比欧美国家，政府具有更大的管理幅度，同时也承担了更多的社会责任。那么中国 DRG 支付制度的实施效果如何？存在哪些医生的非预期行为(unintended response)？目前相关学术研究较少，除了高编码(up-coding)行为外，其它类型的医生非预期行为（unintended response）鲜有研究。鉴于此，本文对中国某市的医生非预期行为类型展开研究。

材料与方法

研究对象

研究选择湖南省湘潭市（DRG 试点城市）开展调查研究，该城市截止 2020 年底，常住人口为 273 万（中国第七次全国人口普查结果），共有二级及以上医院 32 家，其中共有 8 家医院于 2019 年参与了第一批 DRG 支付制度试点，故选择此 8 家医院作为研究对象。

数据

数据取自 DRG 国家试点城市 8 家医院 2021 年 1 月 1 日至 2021 年 9 月 30 日参加基本医疗保险的住院患者的病案首页数据和医保结算数据，共计 188256 例。数据均来源于湘潭市统一的医院诊疗管理 EHR，由监管方湖南省卫健委基于医保基金监管的目标进行信息收集。经过伦理审查、征得卫健委与湘潭市医院同意和数据脱敏后，研究团队取得这一数据并基于数据对医保基金使用的效率和

合理性进行研究。

取得数据后，对字段缺失、不准确的数据以及病案首页数据与医保结算数据无法关联度的数据进行预处理，最终纳入分析的数据量为 169778 例，占总数据量的 90.18%。包括病案首页数据和医保结算数据，具体的数据类型包括：入院日期、出院日期，住院天数、出院主要诊断编码、出院次要诊断编码、主要手术和操作编码、次要手术和操作编码、是否有出院 31 天再入院计划、医疗总费用。住院总费用、基本医保补偿金额、患者自付金额。

研究方法

为了更加全面的了解湘潭市 DRG 支付制度中医生策略性行为的类型，本研究首先运用扎根理论方法对医生策略性行为的主要类型开展质性研究，以便更加准确地了解 DRG 支付制度实施后医生行为的具体变化。然后，采取定量研究的方法来分析 DRG 支付制度实施后医生策略性行为的表现形式。通过将定性与定量研究方法相结合，探讨中国湘潭市实施 DRG 支付制度改革后医生非预期行为的主要类型。

（1）扎根理论与半结构化访谈

本研究选取某 DRG 支付制度国家试点城市的 3 家医院作为研究现场，对部分科室的医生、护士以及医院的管理人员开展半结构化访谈。抽样样本数量遵循资料饱和原则，对每个样本进行多频次的深入访谈，本研究从 2022 年 3 月下旬开始，2022 年 4 月低结束，对选定了 23 名受访者进行了深度访谈。访谈现场具备选择无外界干扰的室内，遵循立场中立、个人信息保护、避免价值判断等原则，在取得被访谈者的同意后，配备 1 名记录员对重点的访谈内容进行记录。本研究按照以下访谈内容的递进关系进行访谈：

1、医院或科室的 DRG 支付制度实施情况；

2、DRG 支付制度实施后，对医生的行为产生了哪些积极影响；

3、您觉得 DRG 支付制度实施后，医生非预期行为类型包括哪些。

访谈结束后，将访谈资料按照扎根理论的编码步骤，对访谈资料进行三级编码，分别为开放式编码、主轴编码和选择性编码。编码的基本原则是：同一资料中含有类似含义的条目只计一次，不同资料在计算时进行合并；与研究对象就编码结果进行讨论，对于编码不确定的地方或新发现的编码，经讨论后再进行修改和完善。遵循以上编码原则，以期获得医生在 DRG 支付制度中策略性行为的主要类型。

表 1 受访者样本特征

项目	分类	人数	占比（%）

性别	男	14	60.87
	女	9	39.13
年龄	30 周岁以下	5	21.74
	30-45 周岁	10	43.48
	45 周岁以上	8	34.78
职业岗位	临床医生	11	47.83
	护士	8	34.78
	医院管理人员	4	17.39
医院性质	三级	14	60.87
	二级	9	39.13
	其它	0	0

（2）中断时间序列分析法

间断时间序列的分段回归模型分析适用于评估政策干预纵向效果的实验研究设计。这一方法能够使用广义差分法解决时间序列的自相关问题，消除因变量随时间本身变动的趋势，只保留虚拟变量和自变量对因变量的影响，有利于准确地评估虚拟变量与自变量在研究中的作用[9]，本文的研究主题十分契合间断时间序列的分段回归模型分析。

运用间断时间序列的分段回归分析研究方法的步骤可描述为：首先，基于收集到的湘潭市 8 家医疗机构实施 DRG 支付制度前后参保患者的住院病例的数据，将实施 DRG 支付制度的时间点作为间断点，将其设置为虚拟变量。随后，对 DRG 支付制度实施前后的指标数据以半个月为一个标的，进行标记。最后，分别对 DRG 支付制度干预前后时间段的指标数据（包括 DRG 支付制度实施前后伴严重合或并发症的 DRG 组中病例数占比、Rw<1 的病例数及占比、14 天内非计划重返住院率等指标进行分析，从而了解湘潭市 DRG 支付制度背景下医生存在哪些非预期行为）进行多元同归分析，从而估算干预前后时间分段指标值的水平和趋势变化。

模型表达公式如下：

$$Y_t = \beta_0 + \beta_1 \times time + \beta_2 \times DRG + \beta_3 \times time_{DRG} + e_t$$

式中各参数的意义如下：

Y_t 为各因变量指标的平均水平，本研究包含伴严重合并症或并发症 DRG 组中病例数占比、Rw<1 的病例数及占比、14 天内非计划重返住院率等指标。

β_0 为基线值，即因变量指标在 DRG 支付制度实施初期的水平估计值；β_1 为基线斜率估计值，即 DRG 支付制度实施前，观察指标随时间变量 t 变化的趋势估计值；*time* 为时间的连续性变量，本研究以每半月为一个观测周期，共设立 18 个数据时间标的，具体的取值为 1-18；β_2 为观察指标的水平变化估计值，即观察指标在 DRG 支付制度实施后的分段起始端与实施前时间分段的末端的水平值差值，它是一种瞬时的取值变化，表示 DRG 支付制度实施时的瞬时作用；β_3 为观

察指标的趋势变化估计值，即实施 DRG 支付制度后时间分段趋势值与实施前时间分段趋势值的差值，表示实施 DRG 支付制度后因变量变化趋势的变动；$time_{DRG}$ 在政策实施前赋值为 0，政策实施后赋值为 1-8；e_t 为误差项，即无法用模型里上述参数来解释的随机误差。

采用 SPSS23.0 统计软件中的中断时间序列分段模型对 DRG 支付制度实施前后相关指标的变化进行分析，检验水平为 0.05，均采用双侧检验。

结果

访谈结果

对非结构化的访谈记录进行逐条整理，将相似的事件与概念进行合并归类，得到统一后的概念及其编码，最后得到 14 个类别。

表 3 概念归纳后的医生策略性行为

编号	概念	原始代表性语句	频次
a1	挑选盈余病例	对门诊患者进行评估，优先把“优质”患者（盈余多的患者）收入院，对于不盈利或者容易出现超支的患者，以“现在没床位”为借口建议其到其他医院住院。	11
a2	将轻症患者转移至门诊治疗	一些轻症患者通常会被分入费用极低 DRG 组，如果收治住院后，医院或科室亏损的风险很大，这时候就让患者在门诊进行治疗，不再办理住院。	3
a3	劝导患者	部分疑难危重病例的治疗成本高，容易出现超支。DRG 支付制度实施后，通常会劝患者转上级医院治疗。	12
a4	低标准入院	很多疾病原本可以在门诊进行处理，但是因为门诊很多费用无法报销，因此，有些患者为了能节省开支就强烈要求住院，这时候医生就只好为其办理住院手续。	2
a5	主要诊断编码高靠	有时患者的诊断不明确，在住院过程中没有确诊，但是在出院时会先选择一个更为严重的疾病编码作为主要诊断编码，这样可以使病例分入权重跟高的 DRG 组。	16
a6	主要手术或操作编码高靠	有时候患者在住院过程中仅开展了诊断性操作，而在出院病案首页中会尽量将诊断性操作编码填写为相似的手术编码，从而使病例分入权重更高的手术类 DRG 组。	5
a7	合并症或并发症编码多（高）编	对于部分病情危重的患者，通常其合并症或并发症多，无论诊疗过程中是否对这些合并症或并发症进行了处理，都要把所有的诊断编码填写进去，这样可以尽可能的使这些病例分入权重更高的 DRG 组，从而避免费用超支带来的亏损。	8
a8	提供使病例分至权重更高 DRG 组的服务	部分 DRG 组设置不合理，如 AH19 组，只有当患者使用呼吸机时长大于或等于 96 小时才能分入该 DRG 组。如果患者使用呼吸机时间接近 96 小时但又没到，这时医生会延长患者使用呼吸机的时长，从使其分入权重高的	3

		AH19 组。	
a9	分解住院	为了节省一次住院中的成本、增加住院例数，将患者原本通过一次住院就能解决的问题拆分为两次或以上，从而增加科室的收入。	5
a10	未达标准出院	为了控制成本，有时会让没有达到出院标准的手术患者提前办理出院手续，然后建议患者回家进行康复治疗或转到其他医疗机构进行康复治疗。	6
a11	减少规定的检查检验项目	在部分专家共识、指南或者临床路径中，明确了部分病种需要开展的检验检查项目，但是医生根据自己的临床经验判断患者不需要做这类检查检验项目，那么为了节约成本，医生就不会开相应的医嘱。	3
a12	诊疗不足	尽管卫生行政部门制定了 1000 余种常见疾病的临床路径，但是对于一些慢性病的住院患者，有时候医生为了控制成本，会减少一些治疗项目，防止超支。	1
a13	完善大型检查后再入院	在患者收入院前，会尽量让患者在门诊完成大型检查，这是控制成本的较好方式，在 DRG 支付实施后可以防止费用超支带来的亏损。	4
a14	让患者到院外购药	为了控制成本，有时会让患者去院外购药，尤其对于医保双通道药品，反正患者到院外去买也能报销，这样患者住院的费用就大大的降低，该疾病在 DRG 支付下就能有很好的盈利。	2

进一步地对这 14 个开放式概念进行分类，将内涵相近的概念归纳得到更加规范的主轴概念，明确主轴概念与开放式概念之间的包含关系。将主轴概念进行编码后得到主轴编码。具体结果见表格 3。

表 3 医生策略性行为的主轴编码表

编号	主轴概念	概念	频次	占比（%）
A1	选择患者	a1 挑选盈余病例 a2 将轻症患者转移至门诊治疗 a3 劝导转诊 a4 低标准入院	28	34.57
A2	高编码	a5 主要诊断编码高靠 a6 主要手术及操作编码高靠 a7 合并症并发症编码多（高）编	29	35.80
A3	过度服务	a8 提供使病例分至权重更高 DRG 组的服务	3	3.70
A4	分解住院	a9 分解住院 a10 未到标准出院	11	13.58

A5	缩减服务	a11 减少规定的检查检验项目 a12 诊疗不足	4	4.94
A6	转嫁费用	a13 完善大型检查后再入院 a14 让患者到院外购药	6	7.41

主轴编码得出的 6 个策略性行为包括高编码、选择患者、缩减服务、转嫁费用、分解住院、过度服务，将策略性行为类型的结果与文献研究结果作对比后发现，本研究调研中出现频次较高的医生策略性行为，相应的研究也较多，且本次调研中出现的策略性行为类型，也占到文献研究中 DRG 支付制度背景下普遍存在的医生策略性行为的多数，这表明本次调研结果大体符合现实情况及过往的研究结果。

描述统计分析结果

（1）费用变化

8 家医院在实施 DRG 支付之前的 5 个月中，所有病例的次均费用为 9260.84 元，在实施 DRG 付费后的 4 个月，所有病例的均次费用为 7525.56 元，次均费用下降了 1735.28 元。具体情况见表 4。从分析结果可以看出，实施 DRG 支付后，大多数医院次均费用出现了较大幅度的下降，这说明医院在 DRG 支付政策下已经开始改变其医疗服务行为。

表 4 家医院 DRG 付费前后次均医疗费用情况

医院名称	DRG 付费前次均费用（元）	DRG 付费后次均费用（元）
韶山市人民医院	5515.74	4566.36
湘潭市第一人民医院	11646.05	9664.62
湘潭市第二人民医院	6738.20	5529.43
湘潭市第三人民医院	7258.41	5875.55
湘潭市妇幼保健院	4986.38	4642.38
湘潭市中心医院	13644.90	10546.62
湘潭县人民医院	7044.22	5880.16
湘乡市人民医院	6005.51	5520.29
总计	9260.84	7525.56

（2）各 DRG 组中病例数及占比分析

8 家医院在实施 DRG 支付制度之前的 5 个月中，病例总数为 97192 例，这些病例被分入 393 个 DRG 组。在实施 DRG 支付制度改革后的 4 个月，所有病例数为 72589 例，这些病例被分入 391 个 DRG 组。剔除组别病例数较少的 DRG 分组后，现将 DRG 支付制度实施后病例数排名前 50 的 DRG 组情况列出，总病例数 98843 例，占总病例数的 58.2%，具体情况见表 5。

从表中可以看出，实施 DRG 支付制度后，部分 DRG 分组病例数及病历占比发生了较大变化，50 个 DRG 分组中也有 35 个变动幅度超过了 15%。部分 DRG 组中病例数出现了较大的增长。例如，RE19 组，其权重为 0.5，在 DRG 支付制度实施前共有 773 例，占全部病例数的 0.80%，然而在 DRG 支付制度实施后共有 3573 例，占比达到 4.92%。其增幅达到了 518.92%。也有部分 DRG 组中的病例数出现了较大下降。例如，FV23 组，其权重为 0.7，在 DRG 支付制度实施前有 1065 例，占全部病例数的 1.10%，而在 DRG 支付制度实施后病例数为 415 例，占比为 0.57%，降幅达到了 47.82%。随后可以进一步从 MDC、ADRG 的视角研究病例数的变化情况，从而分析医生在 DRG 支付制度实施前后的行为变化。

表 5 DRG 支付制度实施前后各 DRG 组中病例数及占比分析

DRG 组	权重	总数		实施前		实施后		占比增幅
		病例数	占比	病例数	占比	病例数	占比	
BR23	0.8	9415	5.55%	5006	5.15%	4409	6.07%	17.93%
RE19	0.5	4346	2.56%	773	0.80%	3573	4.92%	518.92%
FR33	0.9	5590	3.29%	2954	3.04%	2636	3.63%	19.48%
FR19	1.1	4431	2.61%	2114	2.18%	2317	3.19%	46.76%
IU29	0.7	4489	2.64%	2511	2.58%	1978	2.73%	5.48%
ES25	0.4	4473	2.63%	2612	2.69%	1861	2.56%	-4.60%
BR21	1.2	3418	2.01%	1862	1.92%	1556	2.14%	11.89%
ER39	1.3	1912	1.13%	505	0.52%	1407	1.94%	273.06%
ES23	0.8	4549	2.68%	3186	3.28%	1363	1.88%	-42.72%
DT15	0.3	2650	1.56%	1360	1.40%	1290	1.78%	27.01%
KS13	0.9	2141	1.26%	859	0.88%	1282	1.77%	99.84%
FM35	1.2	2852	1.68%	1687	1.74%	1165	1.60%	-7.53%
ET25	0.8	3650	2.15%	2530	2.60%	1120	1.54%	-40.72%
ET21	1.1	3483	2.05%	2458	2.53%	1025	1.41%	-44.16%
RU19	0.8	2380	1.40%	1488	1.53%	892	1.23%	-19.73%
EJ11	2.1	1343	0.79%	545	0.56%	798	1.10%	96.06%
LK19	2.1	1167	0.69%	401	0.41%	766	1.06%	155.78%
GW13	0.7	2417	1.42%	1685	1.73%	732	1.01%	-41.83%
RW23	0.7	1834	1.08%	1125	1.16%	709	0.98%	-15.61%
GW15	0.4	2189	1.29%	1482	1.52%	707	0.97%	-36.12%
GF19	1.4	1520	0.90%	827	0.85%	693	0.95%	12.20%
KS11	1.2	839	0.49%	259	0.27%	580	0.80%	199.85%
GZ13	0.8	1544	0.91%	964	0.99%	580	0.80%	-19.44%
GV13	0.7	1382	0.81%	813	0.84%	569	0.78%	-6.29%
BX23	0.8	1980	1.17%	1414	1.45%	566	0.78%	-46.40%
BR25	0.6	963	0.57%	412	0.42%	551	0.76%	79.07%
EJ13	1.5	1043	0.61%	497	0.51%	546	0.75%	47.10%
LL19	1.2	763	0.45%	222	0.23%	541	0.75%	226.30%
EX25	0.3	1559	0.92%	1026	1.06%	533	0.73%	-30.44%

RW21	0.9	1153	0.68%	629	0.65%	524	0.72%	11.55%
IU19	0.7	1429	0.84%	906	0.93%	523	0.72%	-22.71%
PU15	0.4	973	0.57%	523	0.54%	450	0.62%	15.21%
FM13	6.9	1237	0.73%	808	0.83%	429	0.59%	-28.91%
FV23	0.7	1480	0.87%	1065	1.10%	415	0.57%	-47.82%
LU13	0.5	845	0.50%	431	0.44%	414	0.57%	28.62%
IJ15	1.2	669	0.39%	260	0.27%	409	0.56%	110.63%
NF19	0.7	1086	0.64%	680	0.70%	406	0.56%	-20.05%
LR15	1	781	0.46%	391	0.40%	390	0.54%	33.56%
SR11	1.4	570	0.34%	189	0.19%	381	0.52%	169.92%
RW25	0.6	918	0.54%	559	0.58%	359	0.49%	-14.01%
GZ15	0.5	862	0.51%	506	0.52%	356	0.49%	-5.79%
FM31	1.5	744	0.44%	405	0.42%	339	0.47%	12.08%
GS13	1	805	0.47%	466	0.48%	339	0.47%	-2.59%
CB19	1.1	813	0.48%	475	0.49%	338	0.47%	-4.72%
NC19	2.1	655	0.39%	323	0.33%	332	0.46%	37.63%
RC19	6.2	634	0.37%	305	0.31%	329	0.45%	44.44%
KS15	0.7	531	0.31%	203	0.21%	328	0.45%	116.35%

（3）伴有严重合并症或并发症的 DRG 组病例数及占比分析

在实施 DRG 支付制度后，伴有不同严重程度的合并症或并发症的 DRG 组别之间，病例数和病例数占比也发生了变化，其中伴有严重合并症或并发症的 DRG 疾病代码中第四位数为 1，伴有一般合并症或并发症的 DRG 疾病代码中第四位数为 3，不伴有合并症或并发症的 DRG 疾病代码中第四位数为 5。

将与实施前进行比较出现显著变化（P 值<0.05）的数据筛选出来，并对各 DRG 组的病例数变化情况进行统计分析，具体见表 6。结果显示，筛选出的 10 个 ADRG 组中，有 7 个 ADRG 组中，伴有一般合并症或并发症的病例数占比减少，伴有严重合并症或并发症的病例数占比增加，这一大的趋势变化需引起重点关注。除了者 7 个 ADRG 组外，ED1 组中 ED11 与 ED15 组病例占比增长率相近，ED13 组病例占比减少。FM3 组中病例占比变化率较小。LV1 组中，伴有一般合并症或并发症的病例数占比增加，伴有严重合并症或并发症的病例数占比减少。

表 6 各 DRG 组中伴严重合并症和并发症的病例数及占比

ADRG 组	DRG 组	实施前		实施后		增幅	χ2	P
		病例数	占比	病例数	占比			
ED1	ED11	151	44.41	115	52.04	7.62	15.891	0.000
	ED13	166	48.82	75	33.94	-14.89		
	ED15	23	6.76	31	14.03	7.26		
ER1	ER11	174	50.88	98	60.49	9.62	4.092	0.043
	ER15	168	49.12	64	39.51	-9.62		

ES1	ES11	194	29.94	141	36.72	6.78	7.004	0.030
	ES13	342	52.78	171	44.53	-8.25		
	ES15	112	17.28	72	18.75	1.47		
ET1	ET11	34	36.96	60	63.16	26.20	12.835	0.000
	ET15	58	63.04	35	36.84	-26.20		
FM3	FM31	405	19.36	339	22.54	3.18	5.394	0.020
	FM35	1687	80.64	1165	77.46	-3.18		
FW1	FW11	172	37.97	66	47.14	9.17	11.285	0.004
	FW13	257	56.73	59	42.14	-14.59		
	FW15	24	5.30	15	10.71	5.42		
LV1	LV11	175	29.86	148	31.69	1.83	15.759	0.000
	LV13	388	66.21	274	58.67	-7.54		
	LV15	23	3.92	45	9.64	5.71		
QS4	QS41	45	39.47	60	50.85	11.37	6.242	0.044
	QS43	55	48.25	38	32.20	-16.04		
	QS45	14	12.28	20	16.95	4.67		
RW2	RW21	629	27.19	524	32.91	5.72	14.880	0.001
	RW23	1125	48.64	709	44.54	-4.10		
	RW25	559	24.17	359	22.55	-1.62		
SR1	SR11	189	49.22	381	57.55	8.33	44.936	<.0001
	SR13	88	22.92	204	30.82	7.90		
	SR15	107	27.86	77	11.63	-16.23		

（4）DRG 支付制度实施前后各 MDC 中 Rw<1 的病例数及占比分析

实施 DRG 支付制度后，各 MDC 组中 Rw<1 的病例数占比出现了一定的变化。将各 MDC 组中病例数大于 50 例纳入分析，具体见表 7。结果显示，DRG 支付制度实施后，19 个 MDC 组中有 15 个 MDC 组 Rw<1 的轻症病例占比在下降。其中，部分 MDC 组中 Rw<1 的病例数占比有小幅降低，例如 MDCB、MDCC、MDCG 和 MDCI 组中 Rw<1 的病例数占比分别降低 1.86%、4.05%、2.53%、4.43%。但是，部分 MDC 组出现了较大幅度的下降，例如 MDCE 组，实施 DRG 支付制度前，其 Rw<1 的轻症型病例数为 11661 例，占 MDCE 组总病例数的 70.05%，而在实施 DRG 支付制度后，其 Rw<1 的轻症型病例数为 6002 例，占 MDCE 组总病例数的 56.4%，其占比下降了 13.65%。类似的还有 MDCM、MDCS 组，其 Rw<1 的病例数占比分别降低 27.24%、12.77%。

表 7 各 MDC 中 Rw<1 的病例数及占比

MDC 组	实施前		实施后		χ^2	P
	病例数	占比	病例数	占比		
MDCB	9624	74.42	7399	72.56	10.146	0.001
MDCC	843	63.96	505	59.91	3.603	0.058
MDCD	3579	73.79	2913	75.27	2.470	0.116
MDCE	11661	70.05	6002	56.4	529.653	<0.001
MDCF	7935	57.16	4751	47.91	199.136	<0.001
MDCG	8229	76.74	5289	74.21	14.921	<0.001
MDCH	1639	51.01	1532	58.05	28.934	<0.001
MDCI	4669	77.42	3461	72.99	28.147	<0.001
MDCJ	1376	80.09	955	70.38	39.032	<0.001
MDCK	1616	76.48	1994	74.51	2.456	0.117
MDCL	2804	56.18	2398	58.25	3.928	0.048
MDCM	535	74.72	414	47.48	121.351	<0.001
MDCN	1246	61.68	755	56.01	10.795	0.001
MDCO	279	75.82	179	81.00	2.143	0.143
MDCP	1078	90.13	729	87.62	3.193	0.074
MDCQ	530	89.53	497	83.25	9.953	0.002
MDCR	8125	94.71	6610	92.33	36.943	<0.001
MDCS	690	69.49	755	56.72	39.387	<0.001
MDCX	575	82.38	361	64.58	51.712	<0.001

（5）ADRG 组中 31 天非计划再住院病例数及占比分析

在实施 DRG 支付制度后，各 ADRG 组中 31 天再住院的病例数出现了一定的变化。将 ADRG 组病例数小于 50 例、31 天计划再住院的病例以及、肿瘤放化疗患者、康复、随访等以外的病例剔除后，对 DRG 支付制度实施前后各 ADRG 组的 31 天非计划再住院率进行统计分析，具体见表 8。结果显示，DT1、FM3、GZ1

等 ADRG 组中，其 31 天非计划再住院病例占比出现了明显的上升。表中占绝大多数的 14 个 ADRG 组的 31 天非计划再住院病例数占比都在增加，但较小的绝对值并不一定具有说服力，在间断时间序列的分段回归模型分析将进行更进一步的研究。

表 8 各 ADRG 组中 31 天非计划再入院数及占比

ADRG 组	实施前		实施后		χ^2	P
	31 天非计划再住院病例数	占比	31 天非计划再住院病例数	占比		
BR2	31	0.43%	35	0.54%	0.919	0.338
BX2	12	0.59%	7	0.84%	0.539	0.463
CB1	10	2.11%	10	2.96%	0.599	0.439
DT1	9	0.56%	19	1.21%	3.771	**0.049**
ER3	7	1.40%	27	1.93%	0.598	0.439
FM3	12	0.57%	18	1.21%	4.195	**0.041**
FR3	21	0.70%	30	1.06%	2.146	0.143
FR4	10	0.72%	4	1.38%	1.288	0.257
GS1	11	2.24%	10	2.77%	0.243	0.622
GV1	18	1.83%	16	2.18%	0.265	0.607
GZ1	6	0.41%	16	1.72%	10.641	**<0.001**
HS2	13	2.57%	13	3.45%	0.576	0.448
HU1	4	1.19%	10	2.45%	1.599	0.206
KS1	6	0.46%	11	0.51%	0.040	0.841
LL1	3	1.35%	14	2.64%	1.172	0.279
LQY	18	2.76%	3	4.29%	0.524	0.469
LR1	11	1.42%	15	2.23%	1.346	0.246
LU1	9	1.19%	13	1.55%	0.381	0.537
PU1	24	2.23%	22	3.03%	1.115	0.291
QS3	8	6.78%	10	7.25%	0.021	0.884
RE1	21	10.50%	65	12.24%	0.424	0.515
SR1	7	1.86%	18	2.80%	0.890	0.346

DRG 支付制度中医生的策略性行为变化

（1）伴严重并发症或合并症的 DRG 组中病例数占比及其他诊断数量

根据前述卡方检验分析，在 DRG 支付制度实施后，在 ED1、ER1、ES1、ET1、FM3、FW1、LV1、QS4、RW2 和 SR1 等 ADRG 组中，伴有严重合并症或并发症的病例数出现了增长，其中 ED1、ET1、FW1、QS4、SR1 等 5 个 ADRG 组变化较大，超过了 10%，因此对上述变化较大的 5 个 ADRG 组进行间断时间序列分析。具体分析结果如表 9 所示。

从表中可以看出，ED11 组的病例数占比基础斜率值 β_1 为 -0.005

（P=0.647>0.05），表示在实施 DRG 支付制度实施前，ED11 组病例数占比呈下降趋势，但无统计学意义。ED11 组病例数占比的水平变化量β_2的估计值为-0.053（P=0.611>0.05），表示 ED11 组占比在实施 DRG 支付制度的瞬时降低了 5.3%，但无统计学意义。ED11 组病例数占比的趋势变化量β_3的估计值为-0.010（P=0.622>0.05），表示 ED11 组占比在实施 DRG 支付制度后呈下降趋势，每半个月占比降低 1.5%，无统计学意义。ET11 和 FW11 趋势变化量β_3的估计值分别为-0.023 和 0.022，但是 P>0.05，其他统计量也无统计学意义。

QS41 组的病例数占比基础斜率值β_1为-0.004（P=0.768>0.05），表示在实施 DRG 支付制度实施前，QS41 组病例数占比呈下降趋势，但无统计学意义。QS41 组病例数占比的水平变化量β_2的估计值为-0.044（P=0.680>0.05），表示 QS41 组占比在实施 DRG 支付制度的瞬时降低了 4.4%，但无统计学意义。QS41 组病例数占比的趋势变化量β_3的估计值为 0.050（P=0.027<0.05），表示 QS41 组占比在实施 DRG 支付制度后呈增长趋势，每半个月占比增长 4.6%（见图 1），有统计学意义。

SR11 组的病例数占比基础斜率值β_1为-0.027（P=0.009<0.05），表示在实施 DRG 支付制度实施前，SR11 组病例数占比呈下降趋势，且有统计学意义。SR11 组病例数占比的水平变化量β_2的估计值为 0.044（P=0.586>0.05），表示 SR11 组占比在实施 DRG 支付制度的瞬时增长了 4.4%，但无统计学意义。SR11 组病例数占比的趋势变化量β_3的估计值为 0.061（P=0.001<0.05），表示 SR11 组占比在实施 DRG 支付制度后呈增长趋势，每半个月占比增长 3.4%（见图 2），有统计学意义。

表 9 DRG 支付制度实施前后各伴严重合并症或并发症的 DRG 组中病例数占比变化

DRG 组名称	基础截距β_0		基线斜率β_1		水平变化量β_2		趋势变化量β_3	
	估计值	p 值	估计值	p 值	估计值	p 值	估计值	p 值
ED11	0.519	<.0001	-0.005	0.647	-0.053	0.611	-0.010	0.622
ET11	0.623	<.0001	0.005	0.742	-0.243	0.080	-0.023	0.388
FW11	0.530	<.0001	-0.024	0.008	0.190	0.016	0.022	0.128
QS41	0.414	<.0001	-0.004	0.768	-0.044	0.680	0.050	0.027
SR11	0.655	<.0001	-0.027	0.009	0.044	0.586	0.061	0.001

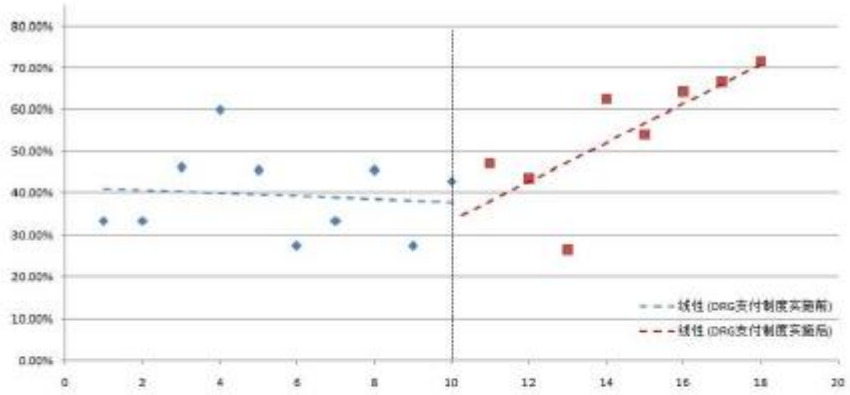


图 1 DRG 支付制度实施前后 QS41 组病例数占比变化折线图

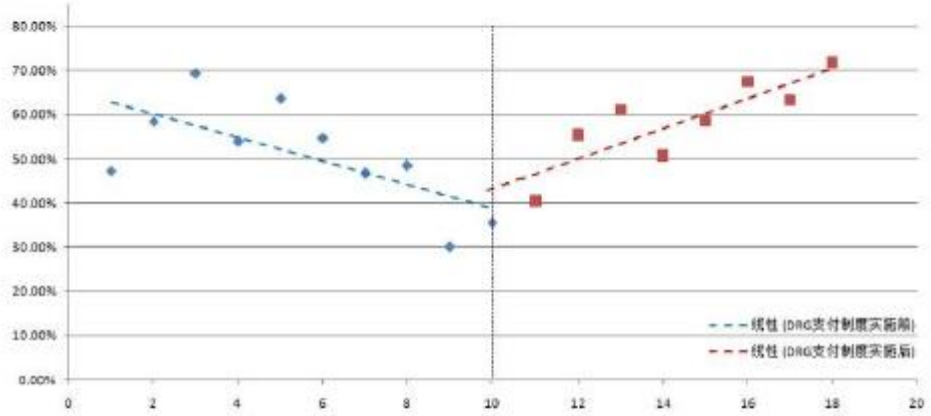


图 2 DRG 付费前后 SR11 组病例数占比变化折线图

（2）各 MDC 组中 Rw<1 的病例数收治占比

根据前述卡方检验分析，MDCE、MDCM 和 MDCS 组在 DRG 支付制度实施后，Rw<1 的轻症病例数出现了较为明显的变化。运用间断时间序列方法分析 DRG 支付制度实施后 MDCE、MDCM 和 MDCS 组中 Rw<1 的病例数占比变化，具体分析结果如表 10 所示。

从表中可以看出，MDCE 组中 Rw<1 的病例占比基础斜率值β_1为 0.005（P 值=0.241>0.1），表示在实施 DRG 支付制度实施前，MDCE 组中 Rw<1 病例数占比呈上升趋势，但无统计学意义。MDCE 中 Rw<1 的病例数占比的水平变化量β_2的估计值为-0.073（P 值=0.087<0.1），表示 MDCE 中 Rw<1 的病例数占比在实施 DRG 支付制度的瞬时减少了 7.3%，有统计学意义。MDCE 中 Rw<1 病例数占比的趋势变化量β_3的估计值为-0.027（P 值=0.004<0.1）,表示 MDCE 中 Rw<1 病例数占比的趋势变化率每半个月降低 2.2%。

MDCM 组中 Rw<1 的病例数趋势变化量β_3的估计值为 0，实施 DRG 支付制度前后无趋势变化。

MDCS 组中 Rw<1 的病例数占比呈现每半个月 2.8%（P 值=0.001<0.1）的增长趋势。在 DRG 支付制度实施的瞬时，Rw<1 的病例数占比下降了 2.2%，但无统计学意义，趋势变化量β_3的估计值为-0.079（P 值=0.000<0.1），表示 MDCS 中 Rw<1 病例数占比的趋势变化率每半个月降低 7.9%，即实施 DRG 支付制度后 MDCS 中 Rw<1 病例数每半个月下降 5.1%（7.9%-2.8%）。

表 10 DRG 支付制度实施前后各 MDC 组中 Rw<1 病例数占比变化

MDC 组	变量	基础截距β_0		基线斜率β_1		水平变化量β_2		趋势变化量β_3	
		估计值	p 值	估计值	p 值	估计值	p 值	估计值	p 值
MDCE	RW<1 占比	0.665	0.000	0.005	0.241	-0.073	0.087	-0.027	0.004
MDCM	RW<1 占比	0.737	0.000	0.004	0.843	-0.236	0.157	0.000	0.000
MDCS	RW<1 占比	0.505	0.000	0.028	0.001	-0.022	0.716	-0.079	0.000

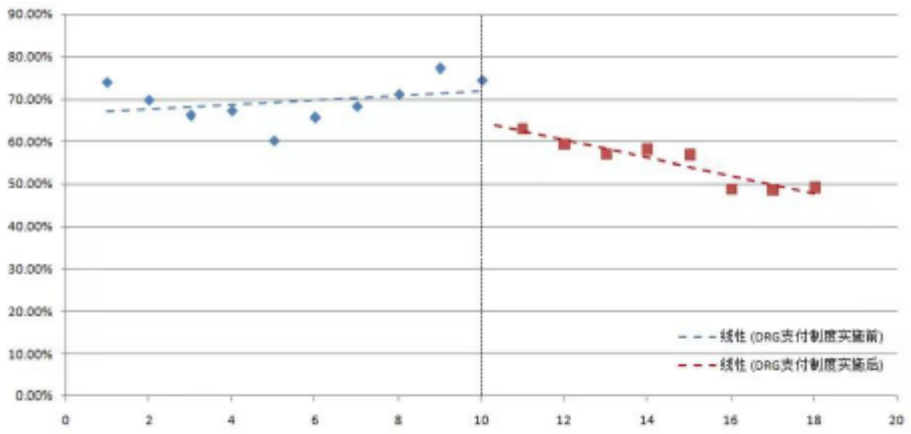


图 3 DRG 支付制度实施前后 MDCE 中 Rw<1 病例数量占比变化折线图

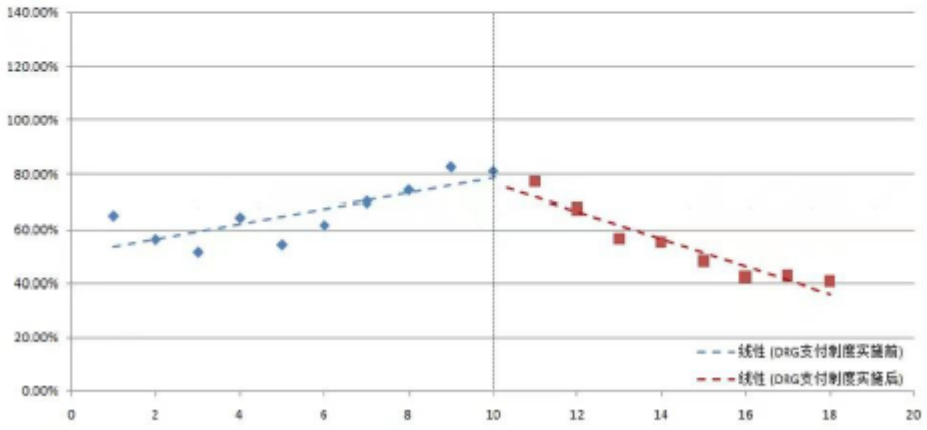


图 4 DRG 支付制度实施前后 MDCS 中 Rw< 1 病例数量占比变化折线图

（3）DRG 支付制度改革前后再住院率变化分析

根据前述卡方检验分析，DT1、FM3 和 GZ1 等 ADRG 组在 DRG 支付制度实施后，31 天内非计划再住院率出现了较为明显的变化。采用间断时间序列方法分析 DRG 支付制度实施前后 DT1、FM3 和 GZ1 等 ADRG 组中 31 天内非计划再住院率，具体分析结果如表 9 所示。从表中可以看出，在 DRG 支付制度实施前，DT1 组中 31 天非计划再住院率的基础斜率值β_1为 0.05%（P 值=0.115>0.05）,表示在实施 DRG 支付制度实施前，DT1 组中 31 天内非计划再住院率呈上升趋势，但无统计学意义。在 DRG 支付制度实施的瞬时，DT1 组中 31 天内非计划再住院率下降了 0.3%，P 值为 0.042<0.05，有统计学意义。在 DRG 支付制度实施后，DT1 中 31 天非计划再住院率的趋势变化量β_3的估计值为 0.15%（P 值=0.035<0.05）,表示 DT1 中 31 天非计划再住院率的上升趋势增快了 0.15%（见图 6）。而 FM3 和 GZ1 由于其基线斜率β_1、水平变化量β_2、趋势变化量β_3的 P 值都大于 0.05，均无统计学意义，因此不做进一步讨论。

表 10 DRG 支付制度实施前后各 ADRG 组中 31 天内非计划再住院率变化

ADRG 组	变量	基础截距β_0		基线斜率β_1		水平变化量β_2		趋势变化量β_3	
		估计值	p 值	估计值	p 值	估计值	p 值	估计值	p 值
DT1	31 天内非计划再住院率	0.0034	0.000	0.0005	0.115	-0.0030	0.042	0.0015	0.035
FM3	31 天内非计划再住院率	0.0028	0.000	0.0005	0.142	0.0027	0.072	0.0000	0.097
GZ1	31 天内非计划再住院率	0.0063	0.000	-0.0004	0.148	0.0096	0.071	0.0013	0.065

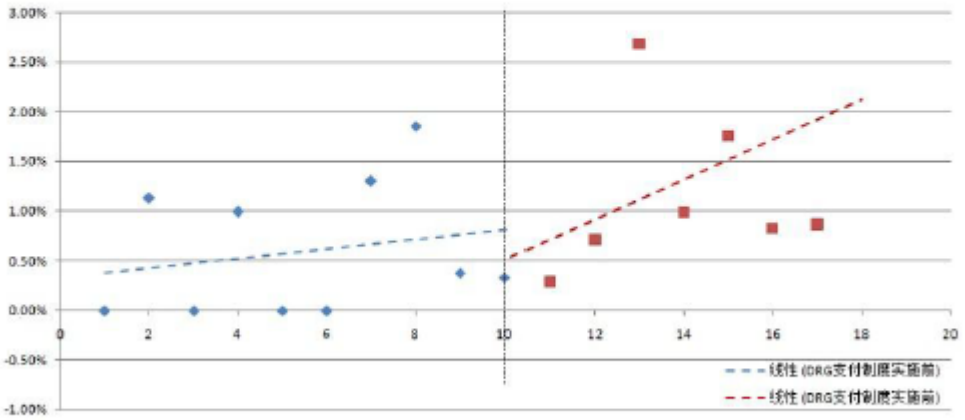


图 5 DRG 支付制度前后 DT1 组中 31 天非计划再住院率变化折线图

讨论与结论

医生的非预期行为

（1）高编码行为

通过现场访谈了解到，医生为了控制超支风险或者获得更多的费用结余，会通过调整疾病诊断编码来使得病例分入权重更高的 DRG 组。医生谈到“有些 DRG 组的权重是根据历史费用确定的，但随着以来新技术的开展，相应的疾病治疗成本也逐渐上升，但 DRG 组的权重还是对应技术进步之前的治疗成本，因此为了补偿开展新技术所导致亏损，会选择多填写其他诊断，这样使得病例分入权重更高的 DRG 组，以此获得更多的医保基金补偿”。另外，也有临床科室主任提出“相比医保管理部门对高编码的处罚，我更担心科里高码低编，因此要求科室的医生要尽可能的多把疾病诊断编码都填写进去，按照 DRG 组的分组规则，只要有其中一个疾病编码属于严重合并症或并发症列表中的编码，那该病例就可以入更高权重的 DRG 组，这样科室的 CMI 指标也升高了，科室的医疗收入也增加了，既能提升在院内的考核排名，又可以得到更多的绩效收入”。通过访谈再次证实了在湘潭 DRG 支付制度实施后，部分医生为了获取更多费用结余，会通过改变主要诊断编码或者多填其他诊断编码的方式来使得病例分入权重更高的 DRG 组。

间断时间序列分段模型的结果研究显示，实施 DRG 支付制度后 QS41 的病例数量占比的上升趋势增快了 5%，最终呈现了每半个月占比增长 4.6%的变化趋势。而 SR11 组占比在实施 DRG 支付制度后也呈较大增长趋势，每半个月占比增长 3.4%。在该 DRG 国家试点城市居民的各类患病率未受到其他明显内部因素和外部因素影响的前提下，QS41 和 SR11 的病例数量占比应当保持在一个较小的波动范围内。然而分析发现，在 DRG 支付制度实施后，QS41 和 SR11 组的病例数量及占比有明显的上升趋势，这主要是由于医院为获得更高医保基金补偿而主动地对疾病编码进行高编码操作，通过改变编码使得病例都分入权重更高的 DRG 组。

通过实地调研，了解到在该试点城市的 DRG 分组方案中，QS41 组（其他贫血且伴严重并发症或合并症）的权重为 1.01，QS43 组（其他贫血且伴一般并发症或合并症）和 QS45（其他贫血且不伴一般并发症或合并症）权重分别为 0.78 和 0.52，较 QS41 组的权重低。SR11 组（败血症伴严重并发症或合并症）的权重为 1.45，SR13 组（败血症且伴一般并发症或合并症）和 SR15（败血症且不伴一般并发症或合并症）权重分别为 0.98 和 0.40，较 SR11 组的权重低。

医生高编码的策略性行为容易发生在 DRG 分组不够细致、并发症和合并症考虑不充分的 DRG 组中，这主要是由于 DRG 分组中疾病编码的设置本身存在一定的局限性，这就为医生高编码提供了可操作空间；另外，考虑到医生对疾病诊

断、治疗有信息优势，其他相关人员难以对编码是否合适进行有效评估。再加上目前 DRG 支付制度的开展还处于初始阶段，缺乏对医生高编码这一策略性行为的应对经验与监管方式，使得高编码的现象较为隐秘地存在于某些疾病编码过程中[10-11]。

（2）选择患者行为

在 DRG 支付制度实施后，由于部分 DRG 组的权重非常低，医生提到，有些轻症病例被分入到支付标准偏低的 DRG 组后，医保常常无法覆盖治疗成本，有医生谈到“轻症患者诊疗后都会亏钱，因为 DRG 组的权重太低，都无法覆盖检查……”，也有医生提出“有些轻症病例，本来可以住院治疗，但是由于将其收治住院后，科室存在超支风险，而这种风险最终会使得医生自身的收入降低，因此会让这些轻症病例去门诊治疗，不再收治住院”。另外，访谈中有一名呼吸内科主任提到，“普通型肺炎且不伴有严重并发症或合并症的病例通常被分入 ES25 的 DRG 组，该 DRG 组的权重为 0.44，而城乡居民医保的费率为 5203 元，这样支付标准仅为 2289 元，当这类患者收入进来，需要做呼吸道分泌物或血病原学检查、CT、胸片、心电图、血常规、尿常规等检验检查项目，同时治疗期间还需要用抗菌药物等，因此尽管非常尽力的控制治疗成本，但是还是非常容易超支，因此这类患者大多数时候就只好转移到门诊进行治疗”。通过访谈证实了在 DRG 支付制度实施后，部分科室或医生会拒收或少收轻症患者，选择性的收治可以盈利的病例。

实施 DRG 支付制度前，MDCE 组中 Rw<1 的轻症型病例数 MDCE 组总病例数的 70.05%，而在实施 DRG 支付制度后，其 Rw<1 的轻症型病例占比 56.4%，其占比下降了 13.65%。实施 DRG 支付制度前，MDCS 组中 Rw<1 的轻症型病例数 MDCE 组总病例数的 69.49%，而在实施 DRG 支付制度后，其 Rw<1 的轻症型病例占比 56.72%，其占比下降了 12.77%。同时，运用间断时间序列方法分析 DRG 支付制度实施前后 MDCE 和 MDCS 组中 Rw<1 的病例数占比变化，结果显示，MDCE 中 Rw<1 病例数占比在实施 DRG 支付制度后呈下降趋势，每半个月占比降低 2.2%。而 MDCS 中 Rw<1 病例数占比在实施 DRG 支付制度后也呈下降趋势，且每半个月占比降低 5.1%。由此可以得出，实施 DRG 支付制度后，MDCE 和 MDCS 中 Rw<1 的轻症患者比例快速下降，结合现场调研的结果，医生有较大的可能性在实施在 DRG 支付制度后出现了拒收或少收轻症患者的行为。

该行为的出现主要是由于临床科室为降低超支风险而导致的。由于部分轻症患者被分入的 DRG 组的权重低，因此其支付标准也低，而收治这些轻症患在扣除医疗服务成本后，其医疗效益非常有限。另外，也非常容易出现治疗费用超过 DRG 组支付标准的现象，而超出 DRG 支付标准的费用通常需要由科室或医生自

行承担。在这种情况下，尽管医生收治 Rw<1 的轻症患者承担的疾病救治的医疗风险较小，但是其承担的经济风险却增加了，因此医院或医生就会选择少收治或者不收治这类轻症患者[12]。与危急重症相比较，轻症患者尽管病情平稳，但是由于医院或医生选择性的收治患者或者拒收、少收轻症患者，一方面会延误病情，损害患者的健康水平，另一方面也可能使得轻症患者发展为重症患者，使得患者的诊疗费用更高，造成医疗资源的浪费[13]。

而 DRG 支付制度时候后医生出现的选择性收治患者的策略性行为主要是由于在现有的医疗体系下，科室或者医生对是否收治患者具有很高自主权，这也为其挑选收益高、风险小的患者提供了操作空间。同时，目前在临床科室或医生层面对是否收治患者的决策往往依靠医生的主观经验判断，还没有其他参与决策的因素对这一环节加以控制，因此需要在临床科室或医生层面加强对患者收治环节的监管[14]。

（3）分解住院行为

通过现场访谈也了解到，分解住院也客观存在。例如访谈中有一名医生提到“尽管医保管理部门会对分解住院的行为进行处罚，但是由于分解住院的行为通常需要核查病历资料才能确认，而医保管理部门的监管力量有限，医院的出院病例数多，因此，其监管有一定的难度，也无法做到全覆盖，这就使得分解住院的策略性行为容易出现”。也有访谈对象提出“由于 DRG 支付制度的原理是对患者的一次住院过程进行付费，而患者无法在一次住院期间把所有的问题都解决，因此，只是对当次住院的主要疾病进行干预，待病情稳定后让患者先办理出院，等过一段时间再次办理入院后再对其他疾病进行治疗”。根据访谈的结果，分解住院的策略性行为需要通过大数据分析并结合病历现场核查的方式才能确认，这就使得监管的成本非常高，因此，分解住院的策略性行为也会存在于医疗服务过程中。

在 DT1、FM3、GZ1 等 ADRG 组中，其 31 天非计划再住院病例占比出现了明显的上升。DRG 支付制度实施前，DT1 和 FM3 组中 31 天非计划再住院病例数占比分别为 0.56%和 0.57%，而在 DRG 支付制度实施后，其占比都上升为 1.21%，提高了近 2.16 倍。而 FM3 在 DRG 支付制度实施后，31 天非计划再住院病例数占比提高了 4.19 倍。而以 DT1 为例，通过运用间断时间序列的方法对医院在实施 DRG 支付制度前后的 31 天内非计划再住院率进行研究，发现在实施 DRG 支付制度后呈上升趋势，且每半个月上升 0.2%。

实施 DRG 支付制度后，医院会设法降低每个患者的医疗成本，具体方法有包括优化内部临床路径等具有积极作用的方式，但也会出现为了控制成本而分解住院的现象，医生会通过将单次住院分解，将治疗患者的成本分摊到两次或多次

住院，即两次获多次 DRG 医保基金补偿[15]。分解住院的行为既会影响患者的医疗质量，也造成了医保基金的浪费。当医生的收入受到 DRG 支付制度较大影响时，医生就有更加强烈的让患者分解住院、提升自身收入的动机。同时，在某些医疗水平相对较低的科室或医生中更容易出现让患者分解住院的现象，这主要是由于这些医疗水平较低的科室医生在医疗费用控制上能力有限，从而很容易出现超标现象，因此只有选择通过分解住院来降低费用支出。

医生非预期行为的出现原因

实施 DRG 支付制度后，研究的 8 家医院中的医生出现了非预期的医疗服务行为，包括高编码、选择患者与分解住院的非预期行为，结合上文涉及到的医生访谈，医生在实施非预期行为时常常处于受迫性状态，即如果不实施非预期行为，自身会蒙受损失。DRG 支付制度确实有规范医生诊疗服务行为的目的，但更多的是希望引导医生采取更加合理的诊疗举措，而不是由于经济利益和精神压力被迫改变自己的行为。同时，虽然访谈中医生谈到非预期行为的出现常常是受迫性的，但仍然存在为了经济利益主动实施非预期行为的可能性。

中国医保基金监管体系是大政府的中国管理模式，国家医疗保障局具备了很大的管理权限，同时承担着更多的社会责任。区别于行政权和司法权分离制衡的西方制度设计，中国的医保基金监管更注重主动性，降低掣肘，使得医疗支付方式的改革能顺利推进。但同时也有治理上的天然弱势，国家医疗保障局作为管理医保基金的最高行政单位，在体系中更加强势，不可避免地采取了更多的惩罚性措施而非更加复杂的激励性措施，同时面对医疗机构提出的建议与反馈，反应也会更加迟缓。

目前，在已实施 DRG 支付制度试点的城市中，医疗保险经办机构开展谈判的积极性不高，并且缺乏积极谈判的动力。其主要原因，一是医疗保险经办机构已经习惯于以往的行政管理方式，认为行政力量对医疗机构更有制约力。二是开展医疗保险谈判意味着医疗保险经办机构自身的革命，医疗保险经办机构需要转换角色（由管理者转变为购买者），转变管理理念（从单方面自上而下的命令强制转变为平等的协商谈判），医疗保险经办机构并不愿意主动作出这种强化责任却削弱权力的改变。三是缺乏相应的激励和监督机制来促使医疗保险经办机构积极开展谈判来为参保人争取最大利益（购买到质优价廉的医疗服务）。

DRG 支付下的非预期的应对

因此，首要需要重新确立医保经办机构和行政管理部门在医保基金治理中的自我定位，医保经办机构需要观念转换和角色转换，通过协商、沟通、讨价还价，而非行政命令强制、单方面决定来处理与医疗机构之间的相关事务，这样只会导

致交易成本增高，即使签订协议执行力也会降低。同时引入参保人群体参与谈判，建立可持续的、和谐的医、保、患三方参与的医疗保险谈判机制[16]，医疗保险管理或经办方与医疗服务提供方及患方通过谈判，就参保人所享受的医疗服务范围、待遇、质量、及支付制度、医疗服务的管理等达成协议，以明确各利益相关方的责、权、利[17]。

除了考虑医、保、患三方的协商参与，还可以通过整合医院协会力量增强医院整体谈判力量，实现与医疗保险方的力量平衡。单个的医疗机构与医保经办机构谈判是没有主动权的，必须通过集体谈判才能取得对等的地位，整合相关利益主体，探索将医院和医生代表（包括医院协会、医师协会等）纳入医保支付方式和支付标准谈判的协商程序，更大程度实现参保人、医保基金和医疗行业多方共赢[18]。

而面对主动实施非预期行为的现象，则需要加强对 DRG 支付制度下非预期行为的识别[19]。一是提高监管技术，降低监管成本。以往的监管方式较为简单，除了医院内部自己实施的举措，主要以实地监管与人工核查为主，成本相对较高。而中国幅员辽阔、人口众多，地域差别较大，监管面变得更广，监管指标也更加复杂，如仍然按照过去的方式对医生的行为进行监管，要么达不到监管效果，要么投入过高。

因此，需要通过改善监管方式来降低监管成本。而基于信息化的监管方式是重要的监管手段，需要在现有的医保信息系统上进行完善。首先，可以在系统前端建立智能化的病案校验系统。通过完善病案数据质量监管规则库，保障病案首页数据的合规性、合理性和真实性，提升病案质量，从数据源头防范高套诊断、挤掉不合理和不规范医疗行为“水分”，实现医保源头监管[20]。

其次，在 DRG 支付制度的信息系统中台建立医疗服务行为监管子系统，添加对策略性行为的监管逻辑，以实现对不合理诊断、不合理用药、不合理耗材使用、大型检查阳性率、重复住院、非计划再入院等行为的过程监管[21]。通过相关规则的内嵌，还可以实时对各 DRG 组、各医疗机构、各医保医师等进行全方位的智能监控，及时预警和纠正医生的策略性行为。

最后，在后端融入医疗机构绩效评价体系。通过智能化的方式采集大数据，对推诿患者、分解住院、选择轻症、病例组合指数（Case-mixIndex,CMI）、三四级手术占比、低风险死亡率、高低倍率病例占比等数据进行分析，并将这些指标的分析结果纳入绩效考核体系中。

PU-002

电子病历应用水平分级机遇与挑战

张红坤
河南大学淮河医院

以电子病历为核心的医院信息化建设是公立医院改革的重要内容之一，也是“智慧医院”发展中“智慧医疗”的基石，推动电子病历信息化建设，对建立健全现代医院管理制度，保障医疗质量和安全，提高医疗服务效率，改善群众就医体验，加强医疗服务监管等，具有重要意义。电子病历系统功能应用水平分级评价工作的开展，是国家全面实施健康中国战略其中重要的一环，通过学习《电子病历系统功能应用水平分级评价方法及标准》，可以让医院各部门清楚认识到信息化的重要性，改变了对医院信息化的理解，提升了医院信息化的战略高度，同时也指导了医院信息化发展的方向，明确了工作思路，这是医院信息化的大机遇，应该抓住机遇，提高信息化水平。在有机遇的同时，也是对医院的挑战，信息化的发展医院要投入巨大的资金、精力，如果没有一个正确的方法可能会得不偿失。

1、机遇

1.1、主导地位的转变是医院信息化发展的机遇，很多医院信息化建设完全是在信息科的主导下完成的，没有明确目标，没有战略规划，不能站在全院发展的角度看待信息化发展，这就造成了在信息科主导下的医院信息化发展乏力，缺资金、缺支持、不理解、推进缓慢；研读《关于进一步推进电子病历为核心的医院机构信息化建设工作的通知》中，明确指出“医疗机构主要负责同志是电子病历信息化建设的第一责任人”，改变了医院信息化职能架构，这一改变将彻底将医院信息化发展放在了院级战略发展层面，由医院主要负责人角度提出战略规划、发展方向，也为医院信息化发展提供了更多的资金、权益，也是医院信息化建设开展工作的强有力后盾。

1.2、以医务人员推动的医院信息化建设的转变是医院信息化发展的机遇，医院信息化建设的主角由信息科转换为医务人员，医务人员是医院信息化的发起者，也是医院信息化成果的既得受益者，将医务人员的被动接受、不愿使用转变成主动要求、主动使用，这是医院信息化成功的关键。

1.3、互联互通建设是医院信息化发展的机遇，《电子病历系统功能应用水平分级评价方法及标准》明确指出医院信息化系统的标准，各软件供应商也要依据标准开发系统，以往各软件供应商各有各的标准，各功能之间接口众多，多个功能之间融合困难、数据挖掘不完整，通过信息平台建设，将各个系统进行标准化改造融合，解决了信息孤岛、信息烟囱问题，为医务人员提供了完整可靠的病历数据，也加强了医疗服务质量控制、效果和效率指标的统计分析和评价。

1.4、信息平台建设是医院信息化发展的机遇，信息共享是电子病历水平分级评价的重要一环，要实现各个电子病历各信息系统的互联互通，应该建设医院信息平台，将分布在不同部门的不同信息系统由分散到整全再到嵌合融合，形成基于平台的整体统一的电子病历信息系统，信息平台的作用不仅仅是数据的融合，还能利用数据进行挖掘分析，提高医院运作效率，提升医院的整体服务能力，有效控制医疗成本，减轻病人医疗负担，提高病人满意度。

1.5、临床决策支持系统建设是医院信息化发展的机遇，临床决策支持系统的运用，是提升医疗质量的重要手段，因其根本目的是为了评估和提高医疗质量，减少医疗差错，从而控制医疗费用的支出。临床医生可以通过临床决策支持系统的帮助来深入分析病历资料，从而做出最为恰当的诊疗决策。临床医生可以通过输入信息来等待临床决策支持系统输出"正确"的决策进行选择，并通过简单的输出来指示决策。

2、挑战

1.1、战略架构的设置，全院级规划是电子病历水平分级评价的挑战，电子病历水平分级评价工作必须是以院领导为核心，信息部门参与决策制定战略规划，医务、护理部门为主，医技科室为铺，软件提供商提供技术支持的架构运行，不能偏离这个架构运行，否则将事倍功半、得不偿失，架构是否合理，规划是否正确将是电子病历水平分级评价成功与否的关键。

1.2、资金投入的增加是电子病历水平分级评价的挑战，信息化建设的设入是不能产生直接效益的，但任何人都不应该忽视信息化带来的变改，信息化建设不但能够提高效率，还能优化流程，提高工作质量，提供决策分析，这些稳性效益对医院发展是不可估量的效益。

1.3、信息化系统应用质量是电子病历水平分级评价的挑战，电子病历水平分级评价工作决不仅仅是实现功能的问题，实现功能是基础，真正将功能大范围的使用、高质量的应用才是目的，将这些功能真正运用到实际工作中是电子病历分级评价工作中一大挑战。

电子病历应用水平分级评价是医院信息化发展的机遇，它为医院的信息化发展提供了正确的方向和目标，引导医院利用信息化促进医疗管理水平提高、改善医疗服务体验、促进智慧医院发展，同时也带来了很多挑战，克服困难、转变思想才能真正确保电子病历信息化建设成功。

PU-003

基于 SWOT 分析法浅析电子病历系统应用水平分级实践中的问题与对策

邵 伟

濮阳市安阳地区医院

1 前言

电子病历系统应用水平分级能够使医院能够客观、动态地了解本院电子病历应用和医院信息化建设情况，引导医院的电子病历建设沿着科学、实用、有助于提高医疗管理水平和质量的方向发展[1]。而其应用水平分级越高，标志着医院整体信息化应用的融合度越强，越有利于各科室、各部门信息交互畅通。

同时，实施电子病历系统应用分级，也是推动医院管理层充分重视电子病历信息化建设的催化剂，有助于实现信息多跑路，医务人员、患者及其家属少跑路，从而进一步提高医疗服务效率，保障医疗质量和安全，改善群众就医体验，还有助于医院管理层对医疗质量的监管，并助推“智慧医院”的建设与发展。

2020 年初，医院提出了电子病历系统达标升级的目标，全院各部门立足现状并结合工作实际，认真梳理系统功能与应用问题，找出不足和短板，积极协调解决。通过 SWOT 分析，从优势、劣势、机会、威胁 4 个维度进行综合权衡，不断的对系统进行整合和优化升级，加强培训力度，强化应用推广，于 2021 年 10 月顺利通过了四级评审验收，为电子病历系统应用水平分级的提升奠定了坚实的基础。

2 材料与方法

SWOT 分析法是通过综合考虑内外部环境的各种因素，客观而准确的进行系统评价一个组织现实情况的方法。利用这种方法分析医院在电子病历系统应用过程中的内部的优势和劣势，以及外部面临的机会与威胁，找出有利因素和不利因素，明确应用过程中存在问题和对策，为医院更好、更快推进电子病历系统应用水平分级提供有益支持。

3 具备的机会和优势

3.1 国家宏观政策的有力推动，为电子病历系统应用水平的提升奠定了牢固的基石与指引。

目前，电子病历系统建设作为推进医院信息化建设的关键环节越来越受到国家重视[2]。2018 年 8 月，国家卫健委颁布了《关于进一步推进以电子病历为核心的医疗机构信息化建设工作的通知》（国卫办医发〔2018〕20 号），对医疗机构以电子病历为核心的信息化建设提出了更加详细的要求与举措。

同年 12 月，国家卫健委发布了《关于印发电子病历系统应用水平分级评价管理办法（试行）及评价标准（试行）的通知》（国卫办医函〔2018〕1079 号）,再次要求医疗机构持续推进电子病历信息化建设，并明确提出电子病历应用水平分级的时间表与路线图，极大推动了电子病历向更高水平建设迈进的步伐。随后，国家卫健委把电子病历应用功能水平分级标准评估结果纳入公立医院绩效考核，进一步夯实了电子病历应用水平提升的基石和指引。

3.2 电子病历应用水平分级是医院信息化建设的关键抓手。

信息技术的发展，推动医院建设不断走向深入，尤其在 5G 时代，智慧医院、智慧医疗已成为医院信息化发展的趋势。在移动互联网时代，电子病历系统是移动医疗的基础工程和枢纽工程，可以广泛应用于临床、教学、科研、社会、远程教育与远程医疗，电子病历系统的建设具有多重意义，将对医疗卫生事业产生深远的影响[3]。

在既往医院信息化建设中，由于技术上的制约以及管理层存在认知和理解的不同，医院信息化建设呈现着系统林立、软件多样，但大多缺乏系统、软件之间的融合贯通。目前，电子病历系统应用水平分级明确提出了医院信息化建设的方向及评价标准，为医院信息化建设提供了根本遵循，指导着医院进行务实有益的探索。

电子病历系统承载着各种医疗行为并贯穿疾病诊断、治疗的全过程，而以电子病历为核心的信息化建设，是打通医院信息断头路的关键抓手。功能完备而又交互的电子病历系统能够顺畅地把医院门诊、住院、检查化验及医疗质量控制等各个环节贯通起来，有利于临床医技各部门开展医疗相关工作，并方便对医疗数据精准提取、充分发掘、智能分析，为医院决策层提供详实的客观依据；同时，也能够为患者提供更加清晰、便捷的诊治流程，提高医院与患者的互动性，不断提升患者及其家属的就医体验。

3.3 智慧医院建设离不开电子病历系统的支撑。

在国家提出“互联网+医疗健康”战略之后，智慧医院建设也走上了快车道，极大推动着医院信息化建设的步伐，更是带动电子病历系统不断优化升级。

在医院中，电子病历系统是医疗工作的核心与纽带，所有的医疗记录都要通过电子病历系统来记载和呈现，建设智慧医院离不开电子病历系统的支撑。而建设高度信息化的电子病历系统，能够发挥互联网、大数据等有关信息技术在医疗救治工作中的优势，让患者及其家属在就医过程中得到更便捷、更安全、更舒适的诊疗体验，进一步助推智慧医院功能的发挥与应用。与此同时，把人工智能应用到电子病历系统，也能够充分利用其在智能分诊导诊、辅助信息采集、辅助诊疗决策支持等方面的作用，不断提高医务人员工作效率，保障医疗质量与安全，更好的服务患者。

3.4 实行疾病诊断相关分组（Diagnosis Related Groups,DRGs）医保付费的需要。

目前，国家大力推进 DRGs 医保支付试点工作，加强医保基金支付科学化管理，使得疾病分组与医保给付水平密切联系在一起。安阳作为国家 DRGs 医保支付 30 个试点城市之一，先行先试，对医疗机构来说，在疾病分组过程中，必须通过电子病历系统提取疾病诊断和手术操作相关诊疗信息，进一步确定疾病的难易程度并赋予相应的权重和点数。电子病历系统分级水平的高低则会

直接影响数据提取的便捷性、准确性，进而决定着医保支付盈亏。因而，医院必须加大电子病历系统建设，不断提升其应用水平分级，更好的应对医保支付制度改革等外部环境变化。

3.5 改善医疗服务，保障医疗安全的需要。

电子病历系统应用水平分级与医院信息化建设密不可分，通过信息支持，有助于医院提高医疗安全性、提高医疗服务质量和提高医疗工作效率[4]。电子病历系统应用水平分级越高，说明其角色越丰富、功能越强大，能够更好的辅助医务人员开展医疗活动，对医疗行为进行提醒、建议，甚至是阻止不合理医疗行为，实现安全诊疗、合理用药，降低医疗风险。同时，电子病历系统提供友好的应用界面，实现简便、智能化的操作，有利于医护人员节省书写医疗文书的时间，把诊疗时间更多的还给患者，不断改善患者就医体验。

3.6 多年信息化建设，积淀了相应的架构和储备。

自 2011 年原卫生部印发了《电子病历系统功能应用水平分级评价方法及标准（试行）》以来，医院越来越清晰的认识到信息化建设的重要性，逐步加大经费投入，购置和升级医院信息系统（Hospital Information System, HIS）、医学影像信息系统（Picture Archiving and Communication Systems，PACS）、实验室信息系统（Laboratory Information System, LIS）、电子病历系统（Electronic Medical Record System，EMRS）等信息系统，完善和整合了系统功能应用。同时加强人才引进和培养，优化人才梯队，为信息化建设积淀了智力支持。经过十多年的信息化建设，逐渐完善了电子病历相关信息系统功能应用架构，储备了相应的人才，为医院电子病历系统发展应用夯实了基础。

4 面临的威胁与劣势

4.1 医院信息化建设需要较高的投入。加强信息化建设，需要不断购置及升级完善 HIS、PACS、LIS、电子病历系统等相关信息系统，并实现各系统之间的互联互通、交互融合。而各类系统的购置、完善和对接，需要大量资金支持，更需要长期的维护与更新，对医院来说，费用支出大，积极性就会受到影响，已成为限制信息化建设的首要因素。没有足够的资金支撑，信息化建设就如空中楼阁，无论是在建设质量还是在效果上都得不到保障[5]。

4.2 专业化人才缺乏。当前，医院信息化管理科专业人员配备不足，缺乏复合型人才，而且现有人员自主研发积极性不高，对厂商依赖性比较大，不能及时对照评价标准进行有效自主改造，已成为制约医院信息化建设的重要因素。此外，医院信息化人才队伍建设滞后问题将制约“十四五”期间国内医院信息化的建设发展[6]。

4.3 电子病历系统构建水平不高。基于节省支出，在电子病历系统建设时，往往会有选择性筛掉一部分功能架构，如缺少智能化监控，不能及时发现病历存在的问题，也不能进行病历质量自检，往往会影响电子病历系统的使用效果。另外，电子病历系统运行速度慢，以及可操作性不友好等问题，更会影响医护人员使用体验。同时，医院各部门采用不同的软件系统，又尚未做到交互融通，势必影响着患者就医和医务人员应用的便利性、连贯性，也影响电子病历系统效用的发挥。

4.4 电子病历系统功能应用不充分。存在系统已实现功能，但临床应用不足[7]。在实际工作中，电

子病历系统已具备的功能，却不能在临床得到充分的应用，如电子病历质控等，需要临床进行多步操作，影响了系统功能应用的积极性。

5 对策与建议

5.1 高度重视，积极谋划，明确职责。医院决策层要建立健全长效工作机制，把电子病历系统建设作为重点，按照国家卫健委要求，加强电子病历在医院信息化建设中的核心地位，持续深入推进相关工作，不断提升电子病历系统应用水平分级。明确医院一把手作为电子病历建设的第一责任人，统筹负责电子病历建设，协调信息技术部门、医药护技以及有关职能科室等，加强管理与质量控制，确保电子病历系统更友好的服务临床工作，保障医疗质量和医疗安全；医药护技以及有关职能科室等要以服务临床为导向，以患者为中心，立足实际，提出电子病历信息化建设需求，并在应用过程中不断改进和完善；医院信息技术部门要积极主动的与各相关科室进行沟通，根据需求加强系统开发、维护、运行和技术支持。同时，加大资金投入力度，依靠走出去与引进来相结合的方式，提高信息化建设水平[8]，提升有自身特色的信息化建设新方向。

5.2 积极引进培养复合型人才，持续强化人才梯队建设。医院信息化建设不仅需要大量资金保障，还需要打造一支高水平的复合型人才队伍，并应根据发展需要积极引进技高层次人才，不断完备信息技术人才梯队建设。同时，注重内部人才培养，定期培训和进修学习，不断更新业务技术知识，提升信息支持保障能力，实现信息化建设创新可持续发展。

5.3 推进系统整合和交互贯通。医院在进行电子病历系统建设过程中，应注重细分需求、统筹推进，对分布在各科室的信息系统持续整合直到交互贯通，逐步解决信息断头路，建成互联互通的电子病历系统，实现院内各诊疗环节按照权限顺畅调阅相关信息。

5.4 加强人员培训，提高应用实效和水平。电子病历系统作为临床信息系统的核心，承载着医疗质量与安全的关键流程与环节[9]。为了发挥电子病历系统应用的便捷性、可靠性、规范性等优势，医院应定期加强系统应用功能培训，持续提升临床医技人员的熟练程度，并对应用过程中存在的问题及时予以解决和优化，不断提升临床医务人员的依从性。同时，加强管理，转变理念，立足于电子病历系统现有的功能，积极推广应用，不断推进电子病历系统应用实效和水平。

6 小结

随着信息技术的迭代发展，智慧医疗、智慧医院建设将成为突出亮点，电子病历系统作为临床应用的核心，借助国家对电子病历基本功能、技术标准规范以及应用水平分级评审的东风，将会得到高度重视并不断完善，为电子病历信息化建设注入强劲动力。

PU-004

医疗保障结算数据质量优化研究

王佳莹
北京大学人民医院

1 引言

2019 年 5 月北京市作为国家首批 30 个 DRG 付费试点城市之一，已开始模拟 DRG 付费。随着国家医保局先后发布《国家医疗保障 DRG（CHS-DRG）分组方案》和《DRG/DIP 支付方式改革三年行动计划》，2022 年 3 月 15 日北京市城镇职工医疗保险住院患者正式启动 CHS-DRG 实际付费。如何面对 CHS-DRG 实际支付，提高医疗保障结算数据质量成为新时期医院需要解决的重点问题。医疗保障结算清单投入使用，当病案首页或医保结算清单上传数据不符合医保结算系统校验逻辑，则不能完成上传；主要诊断选择不当，或主要诊断与主要手术操作不匹配时，则为歧义病例[1]。北京市某三级甲等综合医院在模拟运行中有部分病例不能正常入组，为提高医疗保障结算数据质量实施一系列改善措施，本研究旨在分析该措施的有效性。

2 数据与方法

2.1 数据来源

本研究纳入北京市某三级甲等综合医院 2021 年 10 月 1 日至 2022 年 2 月 28 日及 2022 年 5 月 1 日至 2022 年 9 月 30 日期间出院的北京市城镇职工医疗保险患者进行分析。CHS-DRG 分组结果来自北京市医疗保险医院端业务组件，歧义病例包括分组结果为 0000 组和 QY 组的病例。

2.2 研究方法

为提高医保结算数据质量，该医院在 2022 年 3 月至 4 月期间实施一系列改善措施。

2.2.1 系统建设

DRG 事中管理系统嵌入医生端电子病历系统。医生站可实时获取在院患者诊断与手术操作信息，提示预分组结果，包括 DRG 病组名称、支付标准。医生可监控患者诊断和手术操作信息、住院天数及实际费用情况（包括：药占比、耗占比、费用项目、费用明细）。对预分组不能入组病例，系统提示 QY。

建立医保结算清单审核系统。按照医保结算清单数据上报要求，设置数据传送接口，生成统一满足 DRG 结算要求的医保结算清单数据。医生、编码员和医保质控员具有不同的修改与审核权限。

DRG 分析系统用于结算后分析。完成结算后可查看各科室总体入组情况，包括出院人次、入组率、支付标准和盈亏金额；还可查看具体 DRG 病组明细情况，包括病例数、平均住院日、支付标准和次均盈亏等 DRG 相关指标数据，并可导出相应 DRG 结算报表用于临床科室分析 DRG 入组情况。

2.2.2 医疗精细化管理

对 2021 年 10 月 1 日至 2022 年 2 月 28 日模拟 CHS-DRG 支付期间存在歧义病例的科室逐个开展培训，培训内容包括：CHS-DRG 分组方案介绍、病案首页诊断与手术操作填写规则、医保结

算清单填写要求、病案首页与医保结算清单区别以及歧义病例分析。

对培训临床科室的 DRG 专管员，利用表单大师开展问卷调查。问卷包括单选题和简答题。单选题：（1）通过培训是否对 DRG 分组有了更深入的理解。（2）是否清楚了此次分析 QY 病例和 0000 病例的入组异常原因。（3）本次歧义病例分析是否及时。（4）是否对科室后续的 DRG 相关工作有帮助。（5）是否希望持续开展歧义病例分析工作。简单题为对歧义病例分析工作的建议。每个单选题设置五个选项，按照满意程度由低到高，依次赋分 1-5 分，问卷总分 25 分。

对歧义病例逐一核对，分析是否符合《住院病案首页数据填写质量规范(试行)》要求。如不符合编码规则，按照规则修改后上报；如符合编码规则，病案首页转化成医保结算清单，按照《医疗保障基金结算清单填写规范(试行)》要求上报。

2.2.3 统计分析方法

由于该医院在 2022 年 3-4 月实施了提高医保结算数据质量相关干预措施，将 2021 年 10 月 1 日至 2022 年 2 月 28 日设定为干预前，2022 年 5 月 1 日至 2022 年 9 月 30 日设定为干预后。利用 Excel 2016 建立数据库，对数据进行描述性分析。运用 χ^2 检验分析干预前后入组率变化，分析软件采用 SPSS 26.0 版，以 $p<0.05$ 为差异有统计学意义。

3 结果

3.1 总体入组情况

本研究纳入分析的病例总数为 36178，其中歧义病例数为 594（即 45 例为 0000 组，549 例为 QY 组）。干预前入组 15824 例，异常入组 373 例，干预后入组 19760 例，异常入组 221 例。干预前入组率为 97.70%，干预后入组率提高到了 98.89%，经卡方检验分析，χ^2 值为 79.347，$p<0.001$，差异有统计学意义（见表 1）。22 个科室的入组率均有所提升，且其中创伤骨科，风湿免疫内科，骨肿瘤科，老年科，乳腺外科，神经内科，心血管内科和医疗美容科的干预前后入组率差异经卡方检验有统计学意义（p 值均小于 0.05）。医疗美容科入组率增幅最大，入组率提高了 21.38%。胸外科和血管外科干预后入组率分别降低了 0.67 和 0.98，但差异无统计学意义。

表 1 干预前后入组率分析

项目	干预前 入组率(%)	干预后 入组率(%)	入组率 差值(%)	χ^2 值	p 值
全院	97.70	98.89	1.19	79.347[a]	<0.001
产科	96.25	99.31	3.06	3.172[b]	0.075
创伤骨科	99.21	99.30	0.09	<0.001[b]	1.000
创伤救治	93.20	96.83	3.62	0.930[b]	0.335
耳鼻喉科	98.05	99.39	1.33	4.636[a]	0.031
风湿免疫内科	92.97	99.51	6.54	24.248[a]	0.001
妇科	98.97	99.38	0.41	2.156[a]	0.142

肝病科	97.62	99.57	1.95	0.384[b]	0.283
肝胆外科	97.74	97.89	0.15	0.274[a]	0.601
骨关节科	99.12	100.00	0.88	2.292[b]	0.130
骨肿瘤科	93.58	97.58	4.00	7.885[a]	0.005
呼吸与危重症医学科	96.06	97.57	1.50	1.944[a]	0.163
老年科	91.94	100.00	8.06	19.93[a]	0.001
泌尿外科	98.85	98.99	0.15	0.066[a]	0.797
皮科	90.00	94.32	4.32	2.001[a]	0.157
乳腺外科	99.30	99.76	0.46	4.380[a]	0.036
神经内科	98.24	100.00	1.76	4.690[b]	0.030
肾内科	96.28	98.99	2.71	3.165[b]	0.075
疼痛医学科	98.54	99.50	0.96	0.116[b]	0.734
消化内科	99.16	99.21	0.04	<0.001[b]	1.000
心血管内科	96.82	99.08	2.26	13.178[a]	<0.001
胸外科	99.75	99.08	-0.67	0.903[b]	0.342
血管外科	96.63	95.65	-0.98	0.244[a]	0.621
血液科	98.41	98.70	0.29	0.372[a]	0.542
医疗美容科	63.93	85.31	21.38	16.256[a]	<0.001

注：a 代表皮尔森卡方值，b 代表连续校正卡方值

3.2 分科室培训满意度调查结果

对模拟 CHS-DRG 支付存在歧义病例的科室 DRG 专管员分别进行培训，培训后发放满意度调查问卷，回收 20 份有效问卷。 满意度调查结果见表 2，培训后临床科室 DRG 专管员总体对 DRG 分组方案有了更深入的理解，清楚了产生歧义病例原因，对后续 DRG 相关工作有帮助，并且希望持续开展异常入组病例分析工作。

表 2 满意度调查结果

项目	平均分	选项	
1.通过培训是否对 DRG 分组有了更深入的理解	4.65	非常理解(65.00%)	理解(35.00%)
2.是否清楚了此次分析 QY 病例和 0000 病例的入组异常原因	4.65	非常清楚(65.00%)	清楚(35.00%)
3.本次异常入组病例分析是否及时	4.60	非常及时(60.00%)	及时(40.00%)
4.是否对科室后续的 DRG 相关工作有帮助	4.90	非常有帮助(90.00%)	有帮(10.00%)
5.是否希望持续开展异常入组病例分析工作	4.95	非常希望(95.00%)	希望(5.00%)

总分	23.75	-	-

4 讨论与建议

4.1 深入临床科室交流有助于提高入组率

根据以上干预前后入组率分析和歧义病例分析培训满意度问卷调查得分，开展歧义病例分析培训有助于提高入组率。8 个临床科室入组率提高了，且有统计学意义。在问卷卷调查中，大部分科室 DRG 专管员对异常入组培训非常满意，并且全部 DRG 专管员希望持续开展异常入组病例分析工作。因此，建议医院一方面在临床科室阶段性持续培训，内容包括病案首页填写规范、DRG 分组规则和歧义病例原因分析，从源头提高病案首页填写质量；另一方面，提高编码员的编码准确性，对病案首页编码结果进行智能审核与人工监管，增加编码员与临床科室交流，对于疑难编码病例定期归纳总结[2-4]。

4.2 合理使用医疗保障结算清单有助于提高入组率

根据《北京市医疗保障局关于开展国家医疗保障疾病诊断相关分组（CHS-DRG）付费改革工作有关问题的通知》（京医保发【2022】10 号）要求，定点医疗机构应按照规定上传医疗保障基金结算清单或住院病案首页与医保部门完 CHS-DRG 结算。然而，医疗保障基金结算清单与住院病案首页所使用的编码字典库不同，上传病案首页后北京临床版诊断与手术操作编码被映射成医保版编码，医保版编码字典库在部分章节分类更加细化，个别编码在映射时不够精准，而被分在 QY 组。CHS-DRG 分组方案中入组规则是根据医疗保障基金结算清单设定，上传病案首页后，无论病案首页选择的主要手术操作是哪一个，所有病案首页的手术与操作均会与主要诊断进行匹配，有的病例入组结果并不是所选择的主要手术操作所在组别，有的病例被分为 QY 组。当上传医疗保障结算清单时，可选择与主要诊断相符的主要手术操作，以上问题可以避免，分组过程中主要诊断与主要手术操作进行匹配，分组结果既符合预期，又能顺利入组。同时，医保结算清单审核过程中应严格按照医保结算清单要求选择主要诊断及主要手术操作，避免“低码高编”情况。

4.3 DRG 智慧管理系统应用是提高工作效率的途径

DRG 事中系统与 DRG 分析系统在应用时数据传送过程的仍需人工参与，临床科室获得 DRG 结算报表时间相对滞后。提示在今后工作中，需探索数字化时代医院统计工作新模式，整合电子病历系统、医保结算清单系统、DRG 分析系统、医保分析系统和住院结算系统为 DRG 智慧管理系统，多部门协同合作，做到动态 DRG 分组、事中提示和事后分析，减少人工传递数据过程，从而提高信息传送效率，提升医院科学管理水平[6-7]。

PU-005

医院高质量发展下的手术操作编码现状与需求分析

刘立煌△，黄锋△*，张霞△，张宏宇△，李楠△，霍琍君△
北京大学人民医院

1. 引言

统计是政策科学制定、政策实施监督、政策效果评价的重要依据，特别是在数字化浪潮下，高质量统计对于宏观、中观、微观治理提升的赋能愈加明显，而统计的基础是编码，手术操作编码数据作为卫生统计的核心数据源之一，为卫生行业管理提供了有力的抓手。本文系统分析手术操作编码体系的现状以及公立医院绩效考核对统计指标和数据要素的需求，以完善手术操作统计的决策支撑功能，充分释放其数据应用价值。

2. 高质量发展下手术操作统计的现状与需求分析

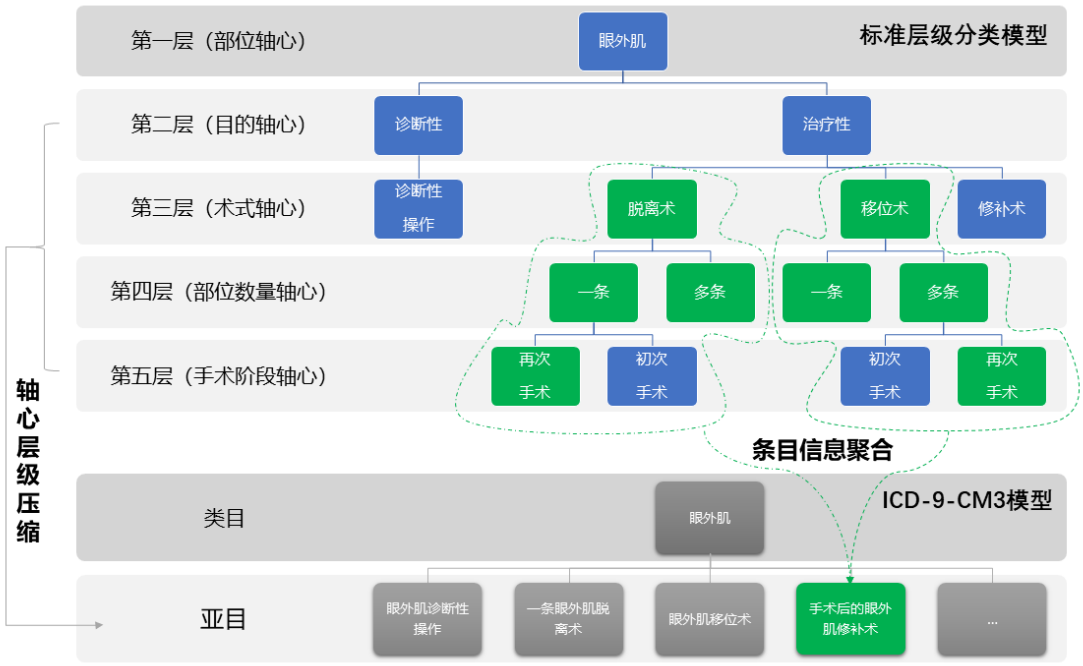
2.1 高质量发展下手术操作统计指标现状与需求分析

目前我国在国家层面对于手术操作的统计应用分析主要集中在公立医院绩效考核（后文简称国考）中的功能定位和医疗质量安全管理等维度中。这些指标能够反映在一定程度单体医疗机构的技术和管理水平，但在整体卫生体系方面有待进一步完善，例如对高质量发展内涵的结构维度的“体系完整”、“连续协同”、“富有韧性”以及在过程维度的“公平性”、“可及性”等目标都缺乏针对性的分析指标。

结合美国、英国、澳大利亚、加拿大等发达国家对于手术操作统计应用的经验，在卫生服务体系的可及性评价方面可以借鉴手术等待时间来反映，对于体系完整性方面可以通过手术谱来反映整体技术概况，对于公平性方面可以通过手术的地区人群分布来反映健康公平性。在指标准确性方面，目前单病种是直接采用住院死亡率进行比较评价，但由于各医院收治患者疾病难易程度不同，直接比较干扰因素较多，因此可以借鉴美国风险调整后的住院死亡率或者和加拿大的 CMG+分组后的手术住院死亡率，提高指标的科学可比性。

从手术操作自身特征属性和价值供给角度，还可以挖掘支撑高质量发展其他维度的指标，例如在连续协同方面，对于紧密型医联体可以采用转诊手术比例，在人民群众健康水平方面可以采用术后生存时间（OS/PFS/DFS），以及术后生存质量评分等，充分满足高质量发展的多方位全面准确的画像评价（见表 1）。

图 1 ICD9-CM3 编码分类底层模型解析



2.2 手术操作统计数据要素的现状与需求分析

高质量发展对数据要素本身也提出了新的需求，特别是在数据质量方面，目前存在较多问题，包括底层标准模型缺陷、编码指南尚未统一、上层应用干扰上报等问题，影响对高质量发展评价的需求满足。

在底层标准方面存在单层多轴心但缺乏注释导致歧义编码、代码位数不足从而聚合分类导致编码信息丢失等问题，目前采用的是 ICD-9-CM3 作为各个版本的基础，其基础架构是一个扁平的 2-3 层的层级分类模型（见图 1），而手术操作的国际上常用的分类轴心包括部位、术式、入路、设备方法、操作部位数量、手术阶段状态、具体解剖细节等 7 个轴心，因此 ICD-9-CM3 扁平的层级模型无法做到一个层级一个轴心，从而产生一个层级多个轴心的现象，这就要求在分类书工具书中对分类优先级必须进行明确注释，否则将导致编码的不确定进而影响数据的一致性，而目前的手术操作分类工具书有多处存在单层多轴心但缺乏注释的情况，例如类目 15-眼外肌手术下的亚目层同时存在手术目的、手术术式、操作部位数量、手术阶段、疾病性质等多个轴心而缺乏明确的分类注释，残余性内斜视的患者再次入院行“一条眼外肌的后徙术”，从手术术式的轴心角度应该分入亚目-15.1 下的 15.11-一条眼外肌的后徙术，而从手术阶段轴心角度看应该分入 15.6-眼外肌手术后的修补术，但分类工具书上在两个亚目下面均无明确注释，实际编码的可能产生歧义，影响手术操作上报数据的质量。

其次，ICD-9-CM3 采用杜威十进制代码结构，一个代码点位容量限定在阿拉伯数字 0~9 共 10 个，一个层级的分类条目除去残余分类最多为 9 个，其余的要么聚合条目信息，要么只能放在残余分类，造成手术操作信息的细节损失（见图 1），例如上述类目 15-眼外肌手术下的亚目，由于点位容量不足，亚目 15.1~15.5 体现了术式和部位数量但丢失手术阶段细节，亚目 15.6 体现了手术阶段细节但丢失了其他细节，因此残余性内斜视再次手术的患者，无论归入哪个亚目都有信息损失，影响手术操作数据的完整性。总之，ICD-9-CM3 的扁平层级分类模型导致轴心层级压缩、条目信息聚合，不利于手术操作的准确完整信息表达，即高质量编码。

在编码指南方面，目前我国缺乏全国层面统一的编码指南，导致在工具书上没有注释明确的问题缺乏共识，影响上报数据的一致性。例如术式是影响手术难度进而影响手术分级的重要轴心，但 ICD-9-CM3 对术式的内涵定义缺乏明确的说明，如“部分切除”与“病损切除”的内涵外延、区分边界、典型举例说明，在临床实际中为了完整切除肿物，经常会扩大边界切除，这种情况是属于部分切除还是病损切除目前仍无共识，而不同的编码产生可能产生不同的手术分级结果，例如

54.4x12-腹腔镜下网膜病损切除是国考的四级手术，而 54.4x14-腹腔镜下网膜部分切除术则不是四级手术，若没有统一的编码指南和明确的术式内涵定义，将可能产生“编码寻租”的现象，影响高质量评价的公平公正性。

在上层应用干扰方面，由于医院手术操作数据被多个政府部门收集应用，不同渠道的数据采取的标准和上报原则有所差异，而基层医疗机构只有一个统计上报部门和一套编码人员，因此多种上报规则可能会互相干扰手术操作上报数据的一致性和准确性。以心外科升主动脉夹层患者为例说明医保 DRG 付费对 HQMS 上报的影响，此类患者经常累及主动脉瓣，因此手术除了进行全主动脉弓人工血管置换并支架象鼻手术（Sun's 手术）外，还需要进行主动脉瓣修补术、主动脉窦修补术，而后面两个术式如果正确完整上报，按照 CHS-DRG 的分组规则此类患者将被分入 FB29-心脏瓣膜手术，这显然不符合以主动脉夹层为主要治疗目的的临床实际，正确的分组应该是 FE2-大血管手术，而如果要入大血管手术组则需要“漏报”主动脉瓣修补术、主动脉窦修补术，此时就产生了医保 DRG 上报与 HQMS 完整准确上报之间的现实冲突。

3. 讨论

手术操作指医疗机构及其医务人员以诊断或治疗疾病为目的，在人体局部开展去除病变组织、修复损伤、重建形态或功能、移植细胞组织或器官、植入医疗器械等医学操作的医疗技术。可见手术操作一方面是重要的医疗服务过程，另一方面也是医疗技术水平的体现，因此手术操作数据既能够反映医疗服务过程，可用于公平、质量、效率评价，另又能够反映医疗技术水平，可用于功能定位等评价，是卫生服务行业高质量发展评价的核心数据源。

但是由于底层标准模型先天不足、统一编码指南缺乏、多个上层应用干扰等现实因素影响手术操作编码的质量，为进一步满足高质量发展需求，建议在国家层面采用先进的面分类架构为主的手术操作编码分类模型，例如国际手术操作分类 ICHI，理由如下 1）利于准确统计：各分面条目定义清晰，减少歧义编码，提高收据准确性/唯一性，2）利于统计细节：手术操作分类要素齐全，信息表达充分，提高数据完整性（完整性）3）科学性：体现了最核心的手术属性，相当于元数据 4）可扩延性更高：面分类体系架构弹性好，采用字母数字编码，单个代码点位容量为 34（不含字母 I 和 O），容量大，5）开发主体为世界卫生组织，相比 PCS 更具资源可及性，且便于与国际统计对接，以上特点使其更符合高质量发展对于高质量数据的需求。其次建议在国家层面编写手术操作指南，对国内目前高频出现的争议问题、灰色地带问题组织专家讨论，例如淋巴结切除范围界定、根治术的另编码、骨科融合术的内涵等问题进行统一共识，提高手术操作数据的一致性。最后，对于应用干扰数据质量的问题，建议医疗机构针对不同的上报部门建立独立的上报流程，尽量减少交叉上报，避免互相干扰，以满足国家对高质量发展的数据要素需求。

PU-006

基于互联网医院平台电子处方流转的云诊室服务

陈文豪[1] 褚震宇[1]
长治市人民医院

1. 引言

自 2019 年山西省省开展药品零售企业电子处方服务试点工作以来，山西已有数千家药品零售企业和近百家省外第三方机构参与试点工作[1]，一定程度上缓解了药品零售企业处方来源不足、执业药师审方不合规的问题。但同时也暴露出远程诊疗不规范、处方来源不合规等问题。对此，各地开展了多形式的零售药店违规销售处方药专项整治行动，涉及的药店数量也较多。山西省和长治市市场监督管理局通过对药店监督检查发现药店销售处方药存在网络销售处方药[2]时未严格落实“四类药品”购药实名登记、信息上报制度缺失、不凭处方销售处方药、AI 虚假医生开方、审方流程如同虚设等问题。

国家卫健委发布《关于开展改善就医感受，提升患者体验主题活动的通知》[3]指出：转变药学服务模式。鼓励医疗机构开展“互联网+药学服务”，推动处方流转等服务。国家医疗保障局发布

的《做好定点零售药店纳入门诊统筹管理的通知》[4]指出“加强处方流转管理，实现定点医疗机构电子处方顺畅流转到定点零售药店。”山西省省药监局在《药品零售企业开展电子处方服务试点工作的通知》中，明确鼓励在本省药品零售企业推行电子处方服务，做好处方用药审核、电子处方流转等工作，确保电子处方的合法性和真实性。

为贯彻落实政策要求，长治市人民医院互联网医院上线了云诊室项目。采用执业医师视频开方形式，为社会零售药店开具电子处方,建立智慧监管平台，电子处方全过程记录同步上传长治市市场监督管理局，以确保电子处方合规流转，改善患者用药依从性[5]、改善医疗环境[6]、保障群众用药安全。

2. 云诊室服务流程

2.1 患者开方申请

患者在药店购买处方药需要开具电子处方时，首先手机打开微信，扫描开方功能二维码，点击开方申请，进入开方申请页面（见图 1）。注：依据国家法律法规，必须是复诊患者[7]。

图 1 开方申请界面

2.2 患者信息维护

患者自主填写身份证作为唯一标识码，绑定好就诊人信息（见图 2），填写患者主诉，全部信息都填写完成（见图 3），点击提交申请。

图 2 患者信息绑定

图 3 填写主诉和添加药品界面

2.3 医患视频对话沟通（患者端）

提交后，系统分配医生，医生可直接用手机端点击进入视频问诊，同时也可以进行图文咨询。进入视频，可以进行麦克风、摄像头关闭开启、摄像头翻转的操作（见图 4）。

图 4 医患沟通界面

2.4 医患对话沟通（医生端）

选择患者的开方订单，进入订单详情，可进行图文沟通和视频沟通，为患者开具电子病历和药品处方（见图 5），确定后授权自动签名，审方通过后药品处方发送给患者。

图 5 医患对话沟通（医生端）

2.5 后台审方（药店方）

打开微管理后台。点击左侧导航栏互联网医院-审方，选择待审核订单，点击查看详情，检查处方笺无误后，进行审核操作（见图 6）。

图 6 后台审方（药店方）

3. 关键技术与支持

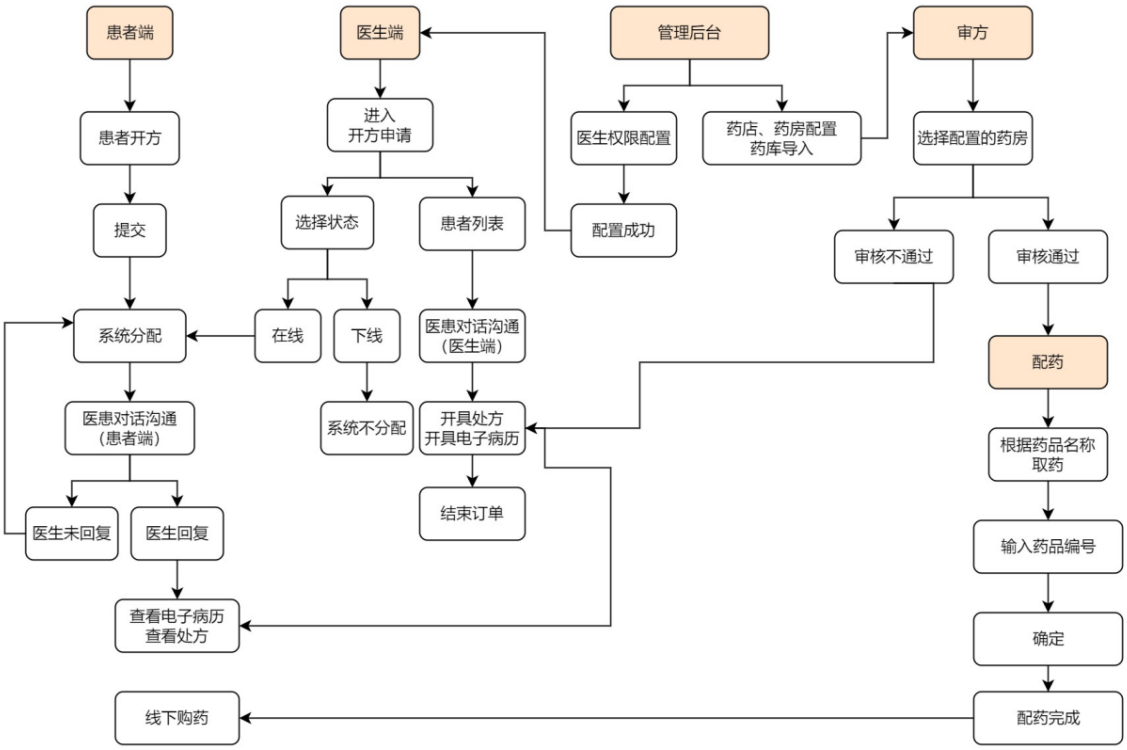
3.1 互联网医院执业资质

长治市人民医院互联网医院已于 2021 年 9 月正式获得执业资格证。具备完善的互联网诊疗制度和信息三级等保，能够为药监市场监管部门提供实时的电子处方数据。符合《互联网诊疗管理办法（试行)》、《互联网医院管理办法（试行）》等法律法规要求，具备《山西省药品监督管理局关于在药品零售企业开展电子处方服务试点工作的通知》开展条件。

3.2 系统软件建设

建设了规范化电子处方系统和审方软件系统（见图 7）。互联网医院使用的信息系统已实施第三级信息安全等级保护，具有合法的 CA 电子签名，确保了云诊室所开具的电子处方的合法性和真实性[8]。建立了智慧监管平台，电子处方数据同步上传市场局监管，对接药品追溯系统，实现药品来源可查、去向可追。

图 7 云诊室系统业务流程



3.3 专业化本地医师团队

长治市人民医院实体医院为互联网医院云诊室提供坚实后盾。作为一家实体医疗机构，有数量众多的执业医师和执业药师，资质齐全，专业知识过硬，所涉学科众多，能为患者提供更专业的远程诊疗服务和便捷的惠民服务，更好地解决患者合理用药和购买处方药的问题。

云诊室服务除电子处方流转业务外，逐步完善云诊室功能，增加许多增值服务[9]，如线上挂

号、线上缴费、线上检查检验报告查询、线上复诊、体检预约、护理服务到家等一系列便民服务。通过云诊室增值服务，群众直接在家门口的药店挂实体门诊专家号，在药店取出医院检查报告，在药店买药时，还可以连线我院医生咨询病症、指导用药和预约检查[10]等等（见图 8）。

图 8 云诊室增值服务

4. 试运行情况

4.1 试运行模式优化

互联网医院云诊室项目于 2023 年 1 月 29 日正式开始试运行，药店参与数量为 33 家。截止 2 月 24 日，开具处方单数量为 7643。经过调研市场现有开方软件的接单时间、药店营业时间和试运行情况，最终确定，医生接单时间为 8：00-13:00、14:00-20:00。由图 9 可知，药店购药高峰时间段为上午 9:00-12：00、下午 16:00-19:00。

图 9 药店开方单量分布

测试初期，采用医生兼职开方模式，即医师利用手机端在工作碎片时间完成开方工作，但不能实时响应药店实时视频开具处方的申请。试运行二周后，统计数据表明，药店开方申请及时接通率仅为 10.7%，远不能满足零售药店及时开具处方的需求。

就阶段性测试成果召开专门会议进行研究，决定采用固定医师排班，匹配零售药店的销售高峰时段。改进服务形式后，二周后运行的数据的表明，实时接通率达到 96.7%，其中 10 秒内接通率达到 87.2%，可以满足零售药店的实时开方需求。

在测试中，我们根据不同的维度进行了数据收集，较有意义的数据见表 1。

表 1 试运行处方单量数据统计

参与药店数量(单位：家）	33
测试时段（（2023/02/21-2023/03/12）（单位：天）	21
完成处方开具（单位：张）	13246
平均每天开具处方（单位：张）	631
平均每天每药店申请量（单位：张）	19
峰值时段	10:00-12:00/16:00-18:00
峰值时段每小时申请量（最高）	149
峰值时段每小时申请量（平均）	115
完成单张处方开具时间（单位：秒）	25

在测试中，我们通过软件使用分析，结合药店实际使用情况，通过优化，提升工作效率的两个方面，见表 2：

表 2 优化内容清单

序号	**用户端**	**医生端**
1	处方申请人录入方式	开方医生 PC 端
2	增加排队人数提醒	增加排队人数提醒
3	优化填写所需内容	优化订单状态（处方审核完成自动结束）
4	增加会员信息导入功能	
5	增加审方后台移动端	
6	增加审方后台选择配药人	
7	优化订单状态（处方审核完成自动结束）	

4.2 试运行人员优化

于 2 月 20 日起，要求 33 家药店所有处方必须全部使用处方平台进行开单，日均订单从 180 增加至 750。我们根据订单量不断调整医生人数并统计接通率，见图 10。

从图中可知，总体接通率稳定在 88%，采用医生 1 主 1 辅模式（此模式为日常时间医生 1 名，药店高峰时间，增加 1 名医生作为辅助接通。）下，10 秒内接通率在 48%，单量日均 750，最高为 22 日，单量在 924，此日采用医生 2 名同时工作。结合 10 秒接通率和排队占比可看出，在单量在 700 时，1 人 1 辅模式可满足需求，缓解排队情况，单量至 900,时，2 人同时在线模式虽接通率降低，但可满足需求，不过排队占比较高，分析原因为药店高峰时段患者较多，医生开单时间虽控制在 20 秒内，但仍会造成排队。未接通订单集中在排队及患者超时方面，电脑及软件问题基本没有。日均 700 单，1 人 1 辅模式可满足需求，在 900 单时，2 名医生可满足需求，之后单量再增加可依据此数据支撑类推：1 名专职医生可承担日均 750 单量，接单时间为 8：00-13:00、13:00-20:00（分两班，每班一人）。在药店购药高峰时间段为上午 9:00-12：00、下午 16:00-19:00，需增加 1 名医生作为辅助接单，以缓解排队压力。

图 10 日均线上开方量分析

日期	订单量	10秒接通率	接通订单			未接通订单						
			10秒内接诊	超10秒接诊		排队	医生超时	非问诊时间	患者超时、未进入	小鱼欠费	电脑/网络	系统问题
				排队接通	超时接通							
1月29日	116	23.28%	27	14	63	4	8	0	0		0	0
1月30日	135	35.56%	48	33	39	2	5	4	1		0	3
1月31日	150	56.00%	84	25	30	1	4	4	1		0	0
2月1日	167	61.08%	102	8	45	0	1	4	6		0	1
2月2日	155	79.35%	123	18	7	1	0	1	5		0	0
2月3日	141	53.19%	75	37	15	4	1	3	6		0	0
2月4日	180	81.67%	147	19	10	0	0	2	2		0	0
2月5日	127	39.37%	50	44	12	4	0	0	5		2	10
2月6日	185	63.78%	118	35	11	4	0	9	6		0	2
2月7日	189	54.50%	103	54	20	0	0	9	3		0	0
2月8日	252	77.78%	196	26	17	0	0	6	7		0	0
2月9日	205	39.51%	81	66	47	1	5	2	3		0	0
2月10日	235	77.87%	183	26	7	3	0	5	11		0	0
2月11日	317	42.27%	134	68	89	10	1	1	14		0	0
2月12日	167	69.46%	116	11	28	3	3	1	5		0	0
2月13日	211	60.66%	128	46	25	2	2	1	7		0	0
2月14日	198	76.26%	151	20	22	0	0	1	4		0	0
2月15日	224	66.96%	150	36	27	1	1	4	4		0	0
2月16日	251	70.52%	177	23	22	15	4	6	4		0	0
2月17日	215	59.53%	128	40	33	2	2	4	6		0	0
2月18日	285	61.75%	176	64	35	0	1	1	6		1	1
2月19日	207	60.39%	125	53	19	2	2	1	2		0	3
2月20日	572	47.73%	273	179	53	17	3	21	26		0	0
2月21日	767	48.37%	371	275	43	45	8	5	20		0	0
2月22日	924	58.66%	542	229	30	39	7	29	46		0	2
2月23日	747	38.42%	287	305	65	51	13	4	22		0	0
2月24日	722	49.31%	356	239	51	28	9	7	30		1	1
2月25日	809	33.75%	273	408	57	48	1	9	13		0	0
2月26日	849	56.30%	478	214	85	22	5	9	31	4	0	1
2月27日	861	49.13%	423	276	55	54	8	9	28			8
2月28日	793	44.89%	356	298	54	30	4	17	33			1

10秒接通：10秒内医生接诊，且连线成功；排队接通：排队订单，超过10秒接诊；超时接通：非排队订单，超过10秒接诊；排队：排队订单，未连线订单；医生超时：非排队订单，超过10秒接诊，未连线成功；患者未进入：10秒内医生接诊，患者未进入，连线不成功

5. 结论

以实体医院互联网医院作为开方服务的主体医院，利用医院已有的软件平台做支撑，完备的数据备份系统，开发扩展云诊室药店终端，逐步推进如线上问诊、检查检验预约、慢性病管理等应用场景，使得药店终端成为医院扩大服务范围的有利抓手。使得原来很多需要到医院现场完成的环节，直接在药店就可以完成，增加了患者就医可及性和便捷性。

PU-007

电子病案智能化在健康中国建设中的应用与前景

周益权
河南亚太骨病医院

引言

病案是动态医疗工作的静态记录资料，对于社会医疗行业来讲，病案能够客观、全面、真实、系统地记录生命安全维护理论、方法与技术，是一种科学价值极高的医疗档案。随着信息技术的不断深入，医院开始普遍使用电子病案以促进医疗改革，提升医院管理工作质量[1-3]。本文将简单介绍电子病案的定义与作用，论述电子病案的三大优势，举例分析电子病案的应用效果，并浅谈电子病案的发展前景。

一、电子病历病案的概念与作用

电子病历(EMR,Electronic Medical Record)也叫计算机化的病案系统或称基于计算机的病人记录(CPR,Computer-Based Patient Record)。它是用电子设备(计算机、健康卡等)保存、管理、传输和重现的数字化的病人的医疗记录，用电子设备(计算机、健康卡等)保存、管理、传输和重现的数字化的病人医疗记录，取代手写纸张病历。电子病历具有主动性、完整和正确、知识关联、及时获取等特征，是医疗机构对门诊、住院患者(或保健对象)临床诊疗和指导干预的、数字化医疗服务工作记录。

二、数字化医院人员和经费的投人方面

1.数字化医院的建设是需要持续不间断的发展，需要不断的新知识、新技能更新。为了满足医院日常工作需要，需要对网管中心人员定期进行专业知识、专业技能的学习培训。可以采取借鉴和学习大型医院信息化建设经验，以适应新形势下工作的需要，同时还要保证技术人员队伍的稳定。

2.我国 1000 床左右的大型综合医院，网管中心人员可
能在 5-20 人之间。而在欧洲，如此规模的医院，网管中心
的人员可能会 50-100 个;在美国的医院，人员可能更多。

三、完善相关信息管理制度

要保证系统能在一个相对稳定的网络环境中运行,网管中心需认真对待相关信息管理制度.院里已经针对医院信息化管理方面制定许多管理制度，现要做的是对现有的管理制度进行一次认真梳理完善。然后组织学习，实行内部责任落实到位。

四、电子病案发展前景

与传统病案相比，电子病案具有良好的发展前景，这种病案具有动态性、关联性和主动性，随着信息技术的不断深入，电子病案的发展也日趋智能化与信息化。此外，理想的电子病案与传统病案的差别主要体现在以下七个方面:

(1) 传统病案是被动的、静态的、孤立的，电子病案是主动的、动态的、关联的。

(2) 传统病案完全不具有电子病案的第二方面功能，即没有主动性和智能。

(3) 纸质病案放在那里，可以被阅读，也可补充新内容，但其内容与内容之间无法建立有机联系，病案内容与患者的实际状态完全脱节，而且，病案内容与其相关知识没有连接、病案只能完成顺序不变的记载作用。电子病案则不同，其内容与患者的实际状态密切相关，涉及到了各方面的专业知识，不仅具有记载作用，而且具备极高的参考价值。

(4) 电子病案所储存的信息不再是孤立的、静态的，而是关联的、动态的，不再是单一的块状知识，而是知识的集合。新补充的信息会与已存在的所有信息建立必要的联系，变换结构，根据现有的知识、规律、规则、先例，对患者的状态进行综合分析判断，主动提示相关医师或者患者，提出相应的治疗计划等。

(5) 传统病历无法保证数据的完整性，电子病案则可保证完整、准确、及时获得信息资料。

(6) 电子病案能够全面管理各种信息资料，传统病案则不具备这种功能。

(7) 电子病案可以根据自身掌握的信息和知识，主动进行判断，在个体健康状态需要调整时做出及时、准确的提升，并给出最优方案和实施计划，传统病案则无法实现主动判断，难以制定最优解决方案。

（**8**）无纸化病历电子签名、时间戳的实施。

引入电子签名(CA)认证，推进病历无纸化。电子病历中的电子认证与签名,所有医疗记录处理系统

产生的最终医疗档案具有可靠电子签名，可信时间戳是解决医疗电子数据法律效力的有效手段。在医患纠纷中既降低了医患间举证的成本又保障了双方的合法的权益。
综上所述，加强医院病案管理工作，提升医院工作效率、准确率及医院服务质量，则需要全面提升电子病案的应用价值，加强病案信息化动态管理工作，重建电子病案管理制度，则需要电子病案管理系统的及时性，有效性，有助于对于电子病案的管理及为电子病案管理工作培养高素质职业人才。

PU-008

某医院 2022 年住院病案终末质量分析

岳馨培

病案是医院管理中最为重要的档案材料之一，是关于患者和疾病的第一手重要医学信息资源[1]。它既是医疗、教学、科研、预防的原始资料，又是医疗保险及医疗纠纷鉴定的法律凭证，它既能反映医疗技术水平，又能评价医院管理质量[2]。随着公立医院绩效考核和 DIP/DRG 医保支付方式改革，病案的书写质量变得越来越重要。本文通过对住院病案进行质量检查，分析病案中存在的主要缺陷及其原因，探讨相应的管理对策，为提高病案终末质量提供参考。

1 资料来源与方法

1.1 资料来源

随机抽取某院 2022 年 1 月 1 日-2022 年 12 月 31 日出院的病案，涵盖医院所有的临床科室，每个病区每月抽取 30 份病案，共抽查病案 61521 份。

1.2 检查方法

每月从各病区抽取质控医师 213 人，职称要求为主治医师及以上。各病区之间进行交叉互查（A 查 B，B 查 C，C 查 A），每人每月审查 30 份病案。质控医师按照国家《病历书写基本规范》和医院制定的《医疗文书规范与管理》等文件的相关规定，进行质量检查并评分。每份病案总分为 100 分，其中入院记录 25 分，病程记录 40 分，出院记录 10 分，知情文件 10 分，首页 5 分，医嘱及辅助检查 5 分，病情评估 5 分。单项否决项每项扣 10 分，包括首页空白，入院 24 小时无入院记录，无首次病程记录，无有创操作记录，无手术记录，无术前讨论，无麻醉记录，无出院（死亡）记录，无手术、麻醉、输血知情同意书等。发现单项否决项，直接定为乙级或丙级病案。病案得分＞90 分为甲级病案，75 分＜得分≤90 分为乙级病案，得分≤75 分为丙级病案。使用 Excel 对病案质控结果进行汇总和统计分析。若 1 份病案中多次出现同一缺陷项目，在计算缺陷频次时进行叠加计算。

2 结果

2.1 抽查病案缺陷情况

共抽查病案 61521 份，其中，满分病案 30862 份，缺陷病案 30659 份，占比 49.84%；甲级病案 61060 份，甲级病案率 99.25%；乙级病案 461 份，乙级病案率 0.75%；无丙级病案。30659

份缺陷病案共统计出缺陷项目 147 项，发生频次合计 48058 次，平均每份病案有 1.6 项缺陷。在抽查的 61521 份病案中，发生频次最多的是首次病程记录照搬入院病史，未归纳提炼，条理不清，占总缺陷频次的 7.50%。发生频次排名前 20 的缺陷项目中，共有 8 项出现在入院记录中，5 项出现在病程记录中，见表 1。

表 1 抽查病案中排名前 20 的缺陷项目发生频次及其占总缺陷频次的比例

序号	缺陷项目	发生频次	占比/%
1	首次病程记录照搬入院病史，未归纳提炼，条理不清	3605	7.50
2	缺出院须知	2975	6.19
3	入院记录患者基本信息缺项或写错或不规范	2739	5.70
4	上级医师首次查房记录无分析讨论或分析讨论不充分	1995	4.15
5	知情同意书签字不完善	1725	3.59
6	无抗凝治疗出血风险评估表	1225	2.55
7	现病史中主要症状、体征及伴随症状描述不清楚	1177	2.45
8	婚姻、月经、生育史缺项或不规范	1099	2.29
9	缺上级医师同意出院的记录	1055	2.20
10	主诉不规范或用诊断代替	1011	2.10
11	无初步诊断或初步诊断不合理、不规范	999	2.08
12	现病史与主诉不相关、不相符	923	1.92
13	手术患者无手术风险评估表	886	1.84
14	内容前后矛盾	883	1.84
15	首次病程记录诊疗计划简单且套话、无针对性、无具体内容	878	1.83
16	首程无分析讨论、无鉴别诊断或分析讨论不充分	877	1.82
17	有辅助检查结果未记录或记录有缺陷	840	1.75
18	知情同意书非患者签名无授权委托书	816	1.70
19	现病史缺有鉴别诊断意义的阴性症状与体征	700	1.46
20	首页三级医师未体现	661	1.38

2.2 乙级病案缺陷情况

461 份乙级病案共统计出缺陷项目 97 项，发生频次合计 1324 次，平均每份病案有 2.9 项缺陷。所有乙级病案的平均分为 87.86 分。乙级病案中，发生频次最多的是缺少术前讨论记录，占总缺陷频次的 9.52%。其次是知情同意书签字不完善，无上级医师首次查房记录或未在患者入院后 48 小时内完成。出现乙级病案较多的科室为肝胆胰外科, 胃肠外科，泌尿外科，消化内科，内科的病案质量普遍高于外科。乙级病案中排名前 10 的缺陷项目发生频次及其占总缺陷频次的比例见表 2。

表 2 乙级病案中排名前 10 的缺陷项目发生频次及其占总缺陷频次的比例

序号	缺陷项目	发生频次	占比/%
1	缺术前讨论记录	126	9.52
2	知情同意书签字不完善	122	9.21
3	无上级医师首次查房记录或未在患者入院后 48 小时内完成	107	8.08
4	手术、麻醉、输血或有创操作病例无知情同意书或未签字	59	4.46
5	无手术记录或未在患者术后 24 小时内完成	58	4.38
6	无术前小结或有缺项、漏项等	46	3.47
7	缺出院须知	31	2.34
8	首次病程记录照搬入院病史，未归纳提炼，条理不清	30	2.27
9	缺术后首次病程记录	28	2.11
10	无有创检查操作记录或未在操作结束后 24 小时内完成	28	2.11

3 讨论

3.1 病案缺陷的主要原因

临床科室方面，一是临床医师对病案书写重视不足。临床医师重诊疗轻管理，存在超时书写、缺项漏项、复制粘贴的现象，在提交病案之前没有进行逐级审查。很多临床医师忽略了病案的法律意义，他们往往在发生医疗纠纷之后，才意识到病案的重要性。二是科主任对病案质量疏于管理。科主任对病案的重视程度直接影响到病案的质量。通过对病案质量较好的科室进行调研，发现这些科室的科主任对病案质量非常重视，会定期组织全科人员进行病案点评活动，评选优秀病案并进行全科展览，极大地提高了科室医师对病案书写的积极性。三是过度依赖病案模板。统计结果显示，在抽查的病案中，“首次病程记录照搬入院病史，未归纳提炼，条理不清”占比 7.50%，“上级医师首次查房记录无分析讨论或分析讨论不充分”占比 4.15%，这些缺陷都与病案模板有关。临床医师过度依赖病案模板，缺乏思考，复制粘贴现象普遍，会导致较多的医疗记录错误，缺乏有意义和个性化的病情分析[3]。四是科室质控流于形式。科室质控专员在提交病案之前，并未认真审查病案，导致知情同意书签字不完善，缺少术前讨论、术前小结、手术记录等缺项漏项问题反复存在。

职能部门方面，一是对环节质控重视不足。长期以来，病案质控模式一般仅重视终末质控，忽略了环节质控的重要性[4]。环节质控可以通过信息系统进行实时监控，加强对运行病历的质量控制，才能及时发现缺陷，真正达到提高终末病案质量的目的[5]。二是质控追踪监管不足。我们将每个月的质控结果，通过信息系统以整改通知单的形式发送到临床科室，由责任医师限期整改后再次提交。对于没有整改的病案，管理部门没有及时进行定责和追踪，导致部分科室反复出现同一病案缺陷。三是奖惩制度落实不到位。虽然医院制定的《医疗文书规范与管理》中明确规定了对乙、丙级病案的处罚措施，但在实施的过程中，执行不足，没有引起临床医师的重视。四是病案质控系统功能有待完善。目前统计的缺陷项目中，书写超时和缺项漏项等问题依然占有较高的比例，这些缺陷可以通过优化病案质控系统实现自动监控。

3.2 提高病案质量的对策

(1) 建立全过程病案质量管理体系 病案质量管理是医院质量管理评价、创建等级医院、国家绩效考核等的关键环节，提高病案质量管理规范化和科学化水平，是医院管理工作中的一项重要内容[6]。病案质量需要临床科室、质控部门和信息部门三方配合，多管齐下，形成合力，建立覆盖病案书写全过程的质量管理体系，通过全覆盖和全过程的培训、考核、评定、绩效评价、奖惩等系列举措，充分提高科主任和临床医师对病案质量的重视程度，严格落实病案书写规范相关要求，提升以病案质量为重点的基础医疗质量内涵，促进医疗核心制度落实，规范临床诊疗行为，保障医疗安全，实现医疗质量的持续改进和提升。

(2)加强病案书写培训 一是组织对临床医师的培训，详细讲解医疗文书的书写规范。加强科主任、质控医师、上级医师对于病案审签的责任心，并通过剖析典型纠纷案例，提升医师的法律意识[7]，提高临床医师对病案书写的重视程度，让临床医师充分认识到病案质量与医院和医生的利益是息息相关的。二是开展对质控医师的培训，详细讲解质控流程和细则，保证质控医师正确理解每一条质控细则的内涵，提高病案质控的效率和准确性。三是对乙级、丙级病案的责任医师开展专项培训，根据乙级、丙级病案中出现的主要缺陷，进行针对性的培训，培训结束进行现场考核，考核不合格者继续参加下一轮的培训，直到考核合格为止。四是针对病案质量比较差的典型科室，每月组织专家和病案管理人员到科室进行督导培训，分析查找病案出现缺陷的根本原因，从源头上提高病案书写质量。

(3) 充分利用信息化管理手段 一是利用信息系统强化环节质量管理。终末质控是对出院已提交的病案进行质量控制，由于病人已出院，病案中的部分缺陷，如知情文件签字不完善等问题已经无法修改，而环节质控可以及时发现问题，及时纠正，因此将病案质控关口前移，将重点放在环节质控上显得尤为重要。要重点加强以“电子病历应用”为核心的信息系统建设，实施医疗质量的系统实时监控，提高运行病历的书写质量，将质量安全管理关口前移，将事后控制转化为事前预防[8]。二是加强对病案系统模板的管理。合理利用电子病案模板可以帮助临床医师快速完成病案书写，有效节约工作时间，良好的模板设计不仅能提高工作效率，还具有管理、查询和统计分析等功能。针对病程大量使用模板，复制粘贴的现象，临床科室要根据实际工作情况建立具有专科特色的病案模板[9]，并进行定期修改、持续改进[10]。

(4)强化病案质控闭环管理 通过病案系统的升级改造，医院从传统的纸质质控模式,到现在实现病案质控闭环管理，可以在终末质控系统中，完成电子病案的审查、打回、整改、追踪、复核和统计分析。病案管理科每月将缺陷病案全部打回，临床医师可在病案系统中查看病案并进行限期整改。对于多次、反复出现的缺陷，按照 PDCA 原理[6]，由病案管理科联合临床科室进行统计分析，查找出现缺陷病案的根本原因，根据原因制定整改方案，落实方案后再次进行检查、反馈，周而复始，持续改进病案质量。对于优秀病案，进行全院展览并给予奖励；对于乙级、丙级病案和未及时整改的病案，进行全院通报并处罚，严格落实奖惩制度，提高临床医师对病案书写的重视程度。

PU-009

电子病历风险防范策略研究

董一超
郑州大学第一附属医院

在医疗卫生信息化建设的推动下，开发并使用医院信息系统（HIS）成为各医疗机构实现信息化的必经之路。我国卫生部于 2011 年 5 月发布了关于推进以电子病历为核心的医院信息化建设试点工作的通知，自此，电子病历成为全国各医疗机构信息化建设的重要内容。近几年，随着国家健康医疗大数据政策的推行，电子病历作为其中的基础数据库之一，在医院信息系统中的地位不断攀升。电子病历在技术上，涉及到网络安全、界面集成、功能集成、数据集成等技术。电子病历在管理上目前还不尽如人意，尤其是风险防范存在一些问题。因此，解决电子病历风险分析与防范的问题，迫在眉睫。

1 国内研究现状及意义

黄云，周敏等（2007），从安全性与隐私性角度阐述了电子病历的医疗风险管理。认为针对我国的实际情况，在一定时间段内电子病历的安全性管理会是风险管理的重点，随着电子病历信息的网络化，隐私性将逐渐成为人们关注的焦点。邹玲玲（2015）认为电子病历的风险管理应该从外部环境、电子技术、管理手段等三方面进行分析。孔庆玲（2017）认为电子病历的风险主要存在于信息缺失、信息失真、信息泄密等三个方面，应从这三个方面分析风险防范的措施[1]。劳颖琨（2017）认为医疗机构关注较多的是电子病历管理中的安全问题，如，数据传输安全、用户访问安全、存储安全等方面[2]。钟敏（2012）认为电子病历在信息生成、信息结构、信息存储和信息传播等方面的特征，是电子病历安全风险产生的根源，主要表现在电子病历载体和电子病历的信息安全等方面[3]。李丽芳（2017）认为电子病历的风险管理主要体现在对以下三个方面的管理：载体的技术风险，存储介质的存储风险，人为管理存在的信息风险[4]。

目前，国内针对电子病历风险防范的研究较少，在风险防范中，我国还存在一些问题，比如，风险防范的意识不足，风险防范的思想并没有深入到项目管理者的心中，对风险并没有很强的防范意识，一般都会在风险发生时才去考虑应对策略。另外，管理者没有成熟、体系化地学习风险管理的相关知识，所以，无法科学地防范和治理风险，尤其对于电子病历这类复杂项目，复杂度高，风险发生率也较高，而可参考的风险防范经验却很少。基于以上结果，针对电子病历风险因素开展相应的分析研究，进而有针对性指出相关解决方案与保障手段。旨在为电子病历的防范提供理论借鉴。

2 电子病历的风险识别

2.1 电子病历的可用性对电子病历长期保存的影响因素分析

电子病历的质量，特别是电子病历存储介质的质量越高，电子病历的存储环境也就越安全，电子病历长期保存工作的风险就越低。对长期保存的电子病历执行正确、合理、及时的操作活动，可以确保电子病历不被破坏。如数字迁移操作可以将存储介质格式过时的电子病历迁移到新的存储介质上，确保其可用性；周期性的病毒检测可以避免由于病毒攻击而导致的电子病历被破坏。在此基础上，可以使电子病历在长期保存过程中具有可用性，在一定程度上可以规避风险。

2.2 电子病历的可识别性对电子病历长期保存的影响因素分析

正确的、完整的电子病历元数据可以确保电子病历能够被发现和检索到。如一个电子病历若缺少唯一标识符则该病历就无法被检索到。另外，电子病历保存系统还需配备合适的软硬件环境，才能将检索到的电子病历呈现出来，从而被用户使用。如针对患者的影像病历，需要使用图像处理系统（PACS）才能够查阅[5]。在此基础上，可以使电子病历在长期保存过程中具有可识别性，在一定程度上可以规避风险。

2.3 电子病历的持续完整性对电子病历长期保存的影响因素分析

完好的电子病历存储介质是保证电子病历内容安全的前提，合理的、正确的操作和存储工作可以确保存储介质的安全。在此基础上，可以使电子病历在长期保存过程中具有持续完整性，在一定程度上可以规避风险。

2.4 电子病历的真实性对电子病历长期保存的影响因素分析

电子病历的真实性要求电子病历没有被改变，或者虽有改变，但改变被如实地记录。这很大程度上取决于电子病历的管理流程和保存政策。正确的电子病历管理流程可以保证针对电子病历的所

有操作都有正确的依据，可以规避错误的操作，也可以保证对电子病历的改变能够及时地被记录，而合理完善的保存政策又是所有管理活动有据可依的基础。在此基础上，可以使电子病历在长期保存过程中具有真实性，在一定程度上可以规避风险。

3 电子病历的风险防范保障策略

3.1 技术保障措施

3.1.1 电子病历的访问权限控制

医院的流程较为繁琐，文件的传输是要经过多人的传递，所以，对于病人的私密病历加密有着至关重要的作用，这就限制了什么人能看，什么人不能观看，只对于个人隐私的保护是极为重要的，所以，这就需要查阅病历是，个人拥有相对应的权限。如，部门主管和主任医师，其不仅可以查阅自己病人的病历，还可以查阅手下管理的医生所填写的病历，但是，他的所有操作流程，都会被系统记录下来，进行保存。

3.1.2 用户管理

对于管理系统来讲，用户的管理主要考虑的是安全性，所以，采用了实名制的医疗卡和指纹识别，对文件进行多重保密。当两次进入系统的时间间距过久时，就会出现自动退出的现象，只有重新输入指纹和面才能顺利登陆，这就有效的防止有人对病人信息的恶意修改和泄露。

3.2 人员保障措施

3.2.1 增强责任意识

确定医院内部各个部门的管理人员和领导的素质以及能力，采取定定期进行培训和考核，严格管控其自身的工作态度和对病历的填写能力，要对病人的信息严格进行保密，不允许进行泄露，不断的对医院全体人员进行素质培养，使其的素质的到质的提升，增强其对病人的责任心，使其对工作引发加大的激情，做好自身的工作。普遍情况下，各个部门的任务都较为多，在每个节点上，员工都要认真对待，等级较高的医生要对新手医生进行不断的提点。电子病历，不仅能与看病一块进行，还能对其提供帮扶，监督医院的工作是否能顺利进行，进一步提升了医院的工作效率。另外，各个部门可以通过病历对其自身的工作进行了解，及时的反映情况，对病历进行及时的管理，这有利于医院管理工作的进行和系统的完善。

3.2.2 关注专业技能指导和培训

医院的领导者对于员工要时常进行系统的培训和增强其关注度，这就极大的促进了系统在员工中间的应用频率。在领导者的管理下，医院内部的各个部门、各个员工进行不断的培训和，不断的操作，才能使员工对系统更加的熟悉。医院对于病历的填写和录入要给予更多的关注度，潜移默化的影响各个员工，对所有病历采取统一进行管理。严格的规章制度和认真的工作态度可以使医院的效率得到极大的提升。

4 总结与思考

我国关于电子病历的发展较短，各个层次的技术都较为稚嫩，系统中存在着较多的隐患。所以本文聚焦电子病历，对其进行风险防范研究。归纳总结了医疗机构在长期保存电子病历过程中存在的问题以及可能出现的风险，发现目前医院在电子病历上存在安全性风险、可用性风险、可识别性风险、真实性风险、持续完整性风险五大类风险，并进一步提出风险防范保障措施。力争在提出医院电子病历风险管理应对解决措施的基础上，对电子病历管理优化提供实践借鉴经验。

PU-010

基于五级电子病历的互联网医院系统信息化建设与实践

瞿泽虹 张佳
重庆医科大学附属第一医院

1．引言

近年来国家卫建委发布《国家卫生健康委办公厅关于印发医院智慧服务分级评估标准体系（试行）的通知》，医院应用信息技术改善患者就医体验，加强患者信息互联共享，提升医疗服务智慧化水平。互联网医院让居民足不出户即可享有优质、快捷的医疗服务。电子病历书写成为互联网医院就诊必不可少的项目，质量管理已成为医疗质量管理的重要内容之一，更是医院管理工作的重要环节[1]。重庆市某院互联网医院基于五级病历的要求基础上设计了互联网医院电子病历，并将其整

合进入医院信息系统（hospital information system，HIS）中，本文拟就此院互联网医院电子病历建设的重要内容、质量标准等展开探讨，为未来互联网医院电子病历的建设提供借鉴和参考。

2.五级电子病历系统应用水平分级评价标准

2.1 局部要求：

各部门能够利用全院统一的集成信息和知识库，提供临床诊疗规范、合理用药、临床路径等统一的知识库。

2.2 整体要求：

2.2.1 全院各系统数据能够按统一的医疗数据管理机制进行信息集成，并提供跨部门集成展示工具。

2.2.2 具有完备的数据采集智能化工具。

2.2.3 利用知识库实现决策支持服务，并能为医疗管理和临床科研工作提供数据挖掘功能。

3.某院互联网医院五级电子病历的建设实践

3.1 预问诊系统与电子病历的结合

智能预问诊从患者和医生双方的角度来优化就医流程，实现诊前问诊和电子病历的结合。在预问诊过程中，系统根据患者的主诉症状，运用医学知识库和 AI 推理引擎，全程模拟医生进行相关问询。根据患者的年龄、性别、主要症状等不同特征生成不同的问诊策略，并通过引导的方式全面采集患者病史信息，包括患者的主诉症状、发作时间、部位、频率和诱因等信息。在预问诊过程中同时根据问诊内容生成结构化的电子病历，在患者就诊的时候呈现给医生。从而帮助医生快速全面了解患者病情，提升医生的工作效率。

3.2设置部署数字证书及认证中心（certificate authority，CA）电子签名系统

电子病历作为贯穿患者诊疗全过程的重要记录和具有法律效力的辅证文书，其在医疗质量管理中具有不可替代的作用[2]。该院互联网医院平台通过采用和第三方CA认证中心合作实现可信行为管理，通过采用证书认证及数字签名服务，可以见证、绑定、生成反映行为真相的电子证据。并与权威司法鉴定机构合作，提供事后的司法鉴定服务，为事后的纠纷处理提供具备高采信力和证明力的电子证据，回溯行为发生的事实真相，因此CA认证对保证电子病历的可溯源性、安全性和合规性是非常重要的。

3.3 电子病历数据与HIS深度对接，信息互联互通

患者的门诊就诊记录、住院病历、院内检查/检验报告，都通过his接口实现实时查询，供医生更全面了解患者病情背景；医生开具的电子病历及处方会回传his系统，形成完整的病历记录，完成院内&线上互联网医院信息的互联互通。

4．讨论

信息化医院的核心就是利用互联网技术提高就诊质量、提升就诊效率，满足患者多元化需求[3]，推进电子病历信息化建设，对保障医疗质量和安全，提高医疗服务效率，改善群众就医体验等具有重要意义。信息化发展电子病历系统可从以下几个方面完善：

4.1 继续提高医疗服务质量和效率

互联网医院医师必须按照《病历书写基本规范》相关要求规范书写互联网医院病历。病历内容应当包括：就诊时间、科别、主诉、病史、必要的体格检查和辅助检查结果、诊断、治疗处理意见和医师签名等。电子病历利用其特有的一些辅助系统，利用信息化的优势，辅助医务人员进行诊疗工作，自动为其提供高性价比、优质的诊疗方案和用药剂量，提高了临床医疗质量，让医生将时间还给患者。

4.2 发展结构化电子病历

结构化电子病历是指从医学信息学的角度将以自然语言方式录入的医疗文书按照医学术语的要求进行结构化分析，并将这些语义结构最终以关系型(面向对象) 结构的方式保存到数据库中［4，5］。结构化电子病历为临床研究者提供了结构化、数字化的临床信息。

4.3 加强医疗质控管理

电子病历的实施能够有效对医疗质量进行全面、多角度的监控管理。设置互联网医院电子病历质量控制岗位，专人专岗，从而实现了病历实时监控和管理，便于对病历质量进行管理。

4.4 电子病历信息共享

利用互联网平台实现不同医疗机构之间的资源共享。在患者就诊时，可通过网络平台查询其在其他医疗机构的检查结果，以便于更加快速、准确的做出合理诊疗方案，有效节约医疗资源，降低了患者经济负担。

4.4.1 院内信息共享

互联网医院系统接入层对接线下业务系统，包括分院区以及托管医院的 EMR、HIS、LIS、PACS 等系统，实现互联网医院和院内系统的信息对接。

4.4.2 院外医疗数据管理系统

与外院的数据互通可采取专线建设模式。相比互联网信息传输，专线网络具有保密性强、数据安全、传输误码率低、传输时延小等优点。患者的就诊信息属于患者的个人隐私信息，数据安全就显得尤为重要。互联网数据在公网上传输容易被不法分子窃取，信息安全得不到保障。而专线网络很好的克服了这一网络传输劣势，使用光纤专线或者VPN让用户的数据在专有的通道上进行传输，实现了物理隔离从而保障院方业务数据安全。

5.结语

医疗信息化的核心是医生和医疗机构，诊疗行为产生的数据均发生在医生把控下的医疗机构内。而信息化将转变为以患者为核心，向患者的健康、诊疗、用药等信息来源也将扩展，从院内到院外、从线下到线上、多种数据采集监测路径等。在诸多的探索中，整个数字医疗行业正逐渐完善“筛、诊、治、养”的闭环。伴随着人们主动健康意识的逐步觉醒，健康消费需求增长，以患者为中心进行信息化建设、优化诊疗流程对医院运营效率的提高起到关键性的作用[6],数字医疗将由“以医院为中心”逐步向“以患者为中心”转变。打破了传统医疗时间和空间的限制，提高居民就医获得感。

电子病历系统作为医院信息管理现代化的重要标志,是医院提高工作效率,改善医疗质量,降低医疗成本的重要途径,是现代医院科学管理发展的必然趋势[7]。互联网医院作为医院综合信息化能力体现的重要部门,建立完善的信息化电子病历系统为出诊医师提供高效便捷应用体验，同时也规范了医师诊疗行为,实现医疗质量的全程管控。电子病历系统的建设一定要具备安全管理体系，具备对电子病历创建、修改、归档等的追溯能力，并且应当符合相关行业标准和规范的要求；在保障信息安全的前提下，促进电子病历信息有效共享，发挥电子病历相对于纸质病历的优势[8]。

PU-011

浅谈住院电子病历质控系统的应用和体会

陈丹
洛阳市东方人民医院医务科

电子病历 EMR(Electronic Medical Record)也叫计算机化的病案系统或称基于计算机的病人记录,它是用电子设备保存、管理、传输和重现的数字化的病人的医疗记录[1]。它将传统的纸张病案电子化，不仅包括纸张病案的所有内容，而且包括声像图文等信息。因其自身特点，电子病历在给医生带来病历书写效率提高的同时，其弊端也日渐显露，由于医生过度依赖模板，病历复制粘贴的滥用，带来了病历质量下降。病历书写的改变，病历质控也随之改变，甚至给医疗质量控制带来了新的机遇和挑战。我院从 2012 年开始运行电子病历系统，同期开始开展电子病历质控工作，2016 年经历了电子病历系统、质控系统的更换，在系统的更新中，质控也不断完善。下面浅谈一下我院电子病历质控系统的应用和体会。

1 电子病历质控系统的应用

1.1 质控标准的确定

根据卫生部《病历书写基本规范》、《河南省病历书写基本规范实施细则》、《医疗质量管理办法》、《医疗质量安全核心制度要点》、《临床输血管理制度》、《抗菌药物临床应用指导原则》等管理规定，结合我院在日常病历质控中发现的问题，制定我院的《病历质量评分标准》，嵌入电子病历质控系统。并根据政策、法规、制度等的调整，不断修订、完善。

按照入院记录、出院记录、日常病程记录、上级医师查房记录、医患沟通、手术相关记录、医嘱等大的分类项目设定缺陷类型，每个大的分类项目下设定更具体的评分标准若干条，根据问题的严重程度对应不同的分值。每份病历满分为 100，终末病历评分分为甲级病案（≥90 分），乙级病案（80 分～90 分）及丙级病案（＜80 分）。

现在根据管理需要，又在质控系统中增设了胸痛中心、卒中中心、VTE 病历专项质控，现已形

成涵盖时效性、逻辑判断、内涵性的质控条目近 300 条。

1.2 质控过程

1.2.1 系统自动质控

病历书写中有时限性要求的、逻辑性项目可实现系统自动质控，并对质控出的问题在医生电子病历书写界面提醒，实现了对全部在院病历的适时监控。

借助病历质控系统，已实现国家卫健委《病案管理质量控制指标（2021 年版）》中病历书写时效性指标完成情况系统自动判读，精确至每份病历超时具体分钟及责任医师，可生成全院及各临床科室各项指标完成率报表，并能提供具体超时病历，为医院病历质量管理工作提供有力的支撑。目前我院将病历书写时效性指标作为专项管理，纳入绩效考核。在规则的设定上，我院要求以电子病历系统中医师签名提交时间为病历完成时间，凡是仅建立书写模板、书写后不签名、不提交均视为未完成，凡是开放后再次提交的，以最后一次完成提交时间为完成时间，以防止有些医生只建立模板提交，实际未完成，从而保证真正意义上的完成，以达到通过病历书写管理实现医疗质量安全管理的目的。截止今年 8 月，我院病历书写时效性指标每项均达到了 95%以上，入院记录 24 小时内完成率更是达到了接近 99%，全院近三分之二的临床科室可达到完成率 100%。

在临床上，一些评估对规范诊疗具有指导意义的，例如胸痛中心质控指标中的 GRACE 评分、VTE 风险评估、营养风险筛查评估等，我院也通过设定病历质控规则，通过系统判断是否进行了评估及评估的及时性，从而督促诊疗行为的规范。

1.2.2 人工质控

1.2.2.1 质控人员的确定

人工质控重点是内涵性，是系统通过时效性、逻辑性判断不能质控的那部分内容，是诊疗过程中更能体现医生诊疗思维的内容，因此对质控人员的临床专业要求和对病历书写规范的掌握，包括医保政策、医疗质量管理要求的了解都有更高的要求。我院在人员配置上实行两种管理方式 ：一是各临床科室间的病历质控由各科室主任担任，每月以“推磨式”的方式互查；二是由临床经验丰富，副主任医师职称以上人员组成的院级病历质控专家，多数是退休的科主任。

1.2.2.2 质控模式

1.2.2.2.1 科室间病历质控

按照内科系、外科系，以“推磨式”互查的方式，医务科每月通过质控系统随机分配质控的科室，由各个临床科室主任通过自己的工号登陆病历质控系统，在规定的时间内完成指定科室的病历质控。外科系质控的重点是手术病历，尤其是四级手术病历，内科系质控重点是疑难病历、死亡病例。

1.2.2.2.2 院级病历质控

院级病历质控专家一是负责行政性的质控任务，例如优秀病历评审，试用期医师病历书写质量评审，以及其它病历评审工作；二是对科室间病历质控得分异常高或低的病历进行复评，以确保质控的准确和公平。

1.2.2.2.3 质控规则

医务科管理员有分配质控科室的权限，每月在质控系统随机分配科室间病历质控互查的科室，科主任只能查看被授权质控科室的病历。院级病历质控专家分配专用的质控工号，每月也有医务科分配需质控的科室，在质控界面可筛选出要求质控的具体病历。

科室主任或院级病历质控专家登陆质控系统后，在质控列表看到的科室即本月需质控的科室，按照系统中的质控操作步骤质控即可，发现问题后点击质控大类下的具体质控标准，就会有相应的扣分，并可以针对该份病历的具体问题做缺陷描述，缺陷描述以消息的的方式即时发送给段落书写医师，责任医师只要登陆电子病历书写系统就能收到该消息，并以发消息的形式将整改情况发送给质控专家，专家收到消息查看是否整改到位，完成系统内质控闭环。

为避免因病历质控产生误解，保障质控的质量和公平，无论是科室间病历质控还是院级病历质控，责任医师收到的消息发送人均为病历质控委员会，仅能看到被质控病历的缺陷内容、质控时间。

1.2.2.2.4 专科质控

在质控系统内设置了胸痛中心、VTE 病历专项质控等专科质控，把各中心的质控标准及要求嵌入质控系统中，生成专科质控指标报表，通过数据分析促进各中心进一步规范化管理，促进医疗质量进一步提升。

1.2.2.2.5 病历评分

系统根据时效性、逻辑性判断可对每份病历质控，扣除相应的分值，得出每位医师每份病历的时效性评分，并可根据需要按照不同的时间段统计出某位医生的病历平均分。

内涵性的人工质控目前还达不到份份质检，每月均为抽查，被抽查到的病历内涵性扣分和系统判断的扣分可自动相加，得出每份病历的最后分值，同样可根据需要按照不同的时间段统计出某位医生的病历平均分及甲级病历率。这样的病历是质控较为完整的整份病历。

科室的病历书写质量得分主要由内涵性的人工质控病历组成。每月被抽查到的该科室医生书写的病历质量平均分即为该科室当月的病历质量平均分。目前，我院将病历书写质量做短板管理，将每月抽查到的该科室病历书写质量最低分病历作为该科室当月的病历质量得分纳入考核。

1.3 结果运用

1.3.1 科室精细化管理

主要从两方面加强科室的精细化管理，一是按照需求和管理要求统计出科室病历质量得分及排名，作为科室医疗质量管理指标，纳入绩效考核。二是科室主任作为本科室质量管理的第一责任人，该项得分作为对科室主任管理能力的评估的一项指标。

1.3.2 人员管理

可统计出任意时间段每位医师病历书写质量得分及排名，为全院临床医师建立了病历质量档案，为晋升、评优等评定提供量化指标。对参加质控的科室主任、院级病历质控专家也建立的质控档案，根据需要可统计任意时间段每位质控者的工作量，包括质控的病历份数及查出问题的条数，为质控专家的能力评定提供量化指标。

1.3.3 提升医疗质量

问题导向，对病历缺陷进行分类汇总、分析，针对问题发生的科室、人群、类型等目标明确的制定有效整改措施，通过提升病案内涵质量全面加强管理，不断提升医疗技术能力和医疗质量水平。

2 电子病历质控系统应用体会

2.1 便捷性。不受时、空限制，无论是运行病历、归档病历，无论是什么时间，质控者只要被授权，可随时随地登陆质控系统完成质控，极大的方便了质控工作的开展 。

2.2 效率高。一是时效性、逻辑性问题，可对全部病历进行质控，系统自动判读，即时发现问题并提醒医生。二是因提前将质控标准嵌入质控系统，便于人工内涵性质控对标准的掌握，系统内勾选或描述缺陷问题，提高了病历质控的速度和效率。

2.3 全程质控。通过病历质控系统可即时向临床医师反馈质控问题，便于及时对病历中的缺陷作出修正。将传统事后质控转变为实时质控、事前质控，从终末质控转化为环节质控中，既保证了诊疗活动的独立性和完整性，又实现了诊疗活动全程质控，把医疗隐患降到最低。

2.4 管理高效。电子病历质控均在系统内完成，不仅可根据需要生成不同的质控报表，而且可对各种数据进行多角度的分析，为日常医务管理提供决策支持。

2.5 局限性。在电子病历系统提供各种模板的便利条件下　格式和时限已不是病历质量的主要矛盾　而病历内涵质量已成为主要矛盾[2]。但是，目前电子病历质控系统还不能完全达到医疗质量管理所要求的，特别是鉴于人工智能和文本语意识别的限制，病历的内涵性质控很大程度上还依赖于人工，还需要医院的管理措施和质控人员水平的提升。

3 结束语

病历质量是医疗质量管理的重点，也是医疗质量的重要组成部分，是反映医疗质量高低的重要指标，对病历质量及时、有效的监管，是提高病历整体质量和保障医疗安全的关键环节。电子病历质控利用其强大的电子计算机网络技术, 以结构化电子病历为依托, 以病历时间、医嘱、模板调用为质控节点[3], 综合应用, 实现病历质量监测控制自动化功能[4]。通过信息及智能技术的不断发展, 逐步完善电子病历质控管理　不断提升病案内涵质量，持续改进医疗质量。

PU-012

病案无纸化管理实践及思考

王崇梁

郑州大学第一附属医院

1 前言

近年来国家卫生健康委积极推动智慧医院建设工作，先后发布一系列文件强调应用信息化技术优化医疗服务流程，提高医疗质量和效率，提升医院精细化、智能化管理水平。智慧医院建设的第一个领域即是面向临床医务人员的以电子病历为核心的智慧医疗。而电子病历的普及，无纸化管理，受到影响的群体远远不止临床医务人员。作为重要的医疗辅助系统，正处于由传统模式向无纸化模式转型的医院病案管理部门是如何完成自身的工作职责，以及如何能更好地为一线医务及患者提供配套的病案服务值得关注。

随着信息技术的快速发展，以各种信息手段进行的记录和存档取代传统纸媒载体，这一变化在病案领域中的应用尤为显著。我国对无纸化病案信息管理进行探索和尝试始于上世纪末，首先在一线城市的大中型医院中试运行了具有部分数字化功能的电子管理系统，如在药品管理和医疗收费等流程中对患者信息进行结构化管理，但由于此时缺少系统性和关联性，零散数字化的部分远远无法承担起病案的主题内容。进入 21 世纪以来，越来越多的医疗机构认识到无纸化病案对自身快速发展的重要性，无纸化病案的发展开始加快进程。2010 年左右，我国出台了电子病历的基本构架和数据标准，并在全国多个地区开展无纸化病案的试点工作[1]，我国病案无纸化的发展开始进入正轨。如今，电子病案虽然已在我国得到广泛应用，无纸化管理模式已成为医院病案信息管理的发展趋势[2]，但在具体执行的细节和实际应用中还存在诸多空白和有待解决的瓶颈问题，导致多数医院只能暂时保留电子版和纸质版双轨制管理办法，且病案信息只局限在医院内部，无法实现互通共用，总体上仍处于初级发展阶段[3]。想要实现病案全面无纸化，还有较长的路要走。

以电子病历为核心的智慧医疗的强力推行，使病案无纸化管理箭在弦上。本文结合实际工作经验，定性分析医院病案无纸化管理的应用优势、现阶段存在的问题，探讨相应的优化策略。

2 病案无纸化管理突显的优势

2.1 针对临床医生的病案信息生成层面

综合性医院日常往往要接诊数以万计的患者。每位患者在诊疗过程中往往需要流转于多个医务环节，接受数项到数十项不等的检查检验，其诊疗数据经由多人处理后汇总于病历中。使用病案无纸化管理系统，在系统内书写完成绝大多数记录和检查结果，可以最大程度在病历生成环节降低纸质材料的产生，无论是召集多病区医生会诊或将患者转诊其他病区，此前的相关信息均以数字化模式跟随患者，既无需为传统病历的流转工作耗费过多精力，也降低了后续病历手工整理、核对的麻烦，方便后期系统内的首页质控、调取等步骤。

2.2 针对管理者决策者的病案信息使用层面

无论是在日常工作中还是面对突发事件时，上级卫生部门和医院的管理层随时可能需要即时准确的相关患者信息，如患者人数、检查结果、病情转归情况等。传统纸质病历对上述情况的结果统计具有明显的延迟性，无法保证每次上报的准确，而使用无纸化管理，可以对以上数据进行结构化一键提取。可以确保有关上级在最短时间内将临床一线的救治情况，及时进行有关命令的下达和决策的调整。

2.3 针对患者的病案信息服务层面

患者的病案信息服务主要是出院患者为慢病治疗或医保报销而申请复印住院病历。现阶段的无纸化病案分为两部分：一是直接在系统内完成记录的电子病历，二是暂时无法直接转化为电子病历而进行的纸质病历高拍，需要时均只需要在系统内操作即可查看和提取。但对患者对病案资料的需求仍然存在，针对这种情况，线上病案提取申请显著增多。无纸化病案管理系统使得患者及其家属可以利用自助打印机器以及线上申请来自主提取病历，不仅分散了病案管理部门人工服务的压力，也降低了人群密度，减少了聚集性传染的风险。

3 存在问题

3.1 人力消耗及病案质量问题

无纸化病案虽然免去了纸质病历搬运和保存的麻烦，为病案在院内各部门之间流通提供了便利，但是在生成过程中却变相增加了一些额外的工作量，如耗材条码，个别手签文件的高清拍摄。这些无疑增加了临床医生的工作负担，尤其是在工作量骤增的疫情时期，容易出现延迟拍摄，甚至漏拍的情况。但在无纸化进程的初级阶段，这些都是不可避免的，需要得到临床医生的理解和配合。

3.2 无纸化病案中相应内容的法律效力

病案是判断患者行为能力以及受害情况的重要依据，特别是在医疗纠纷的处理过程中，病案将会直接关系到患者及医院的利益。尤其是像新冠疫情这样新发现的病毒，其治疗手段无疑是在摸索中前进，后续是否会因此发生医疗纠纷，暂未可知。虽然早在 2017 年 4 月，原国家卫生计生委医政医管局颁布了《电子病历应用管理规范(试行)》并开始施行[4]，文件规定：可靠的电子签名与手写签名或盖章具有同等的法律效力。但除此之外，无纸化病历管理中仍存在许多法律盲点，如输血记录、高值耗材等医疗产品的序号条码的高拍影像是否与原始材料具有同等的法律效力，至今没有明确的法律条文规定。在这些疑问落实前，任何医疗机构都不可能弃掉原始的实物条码，而继续保存这些原始材料使“无纸化”仍仅是徒有其名。

4 改进策略

4.1 不断根据实际需要改进系统设置

由于病案无纸化管理尚在探索阶段，无论采用哪个软件公司提供的病案管理系统，都不能一次性尽如人意，因此需要在使用过程中不断根据不同医院、不同科室的实际情况进行调试。因此，在选择病案管理系统时，医院应选择能够灵活进行改进升级、能够兼容多种后续子系统的信息系统。在实际操作中，对不同病区，应根据其实际需要、实际功能，对需要加入病历的检查结果及记录的模块能连尽连，以减少高拍工作量。此外，完善各模块关联功能也是一项值得关注的问题，如果能够合理处理各具有关联意义的模块间的关联功能，往往可以事半功倍的。

4.2 适当设置专职的高拍岗位

临床医生日常工作繁忙，遇到疫情则更是需要连续加班，往往没有多余的时间对额外产生的耗材条码等材料进行高拍，所以相应医疗机构能考虑在未能实现完全无纸化管理前设置专职人员对特定文件进行高拍，受过相应训练的专职人员不仅能够减轻一线医生的工作负担，也更能保证高拍的质量。

4.3 完善无纸化病案的法律条文

纸质版的病案受我国相关法律法规保护，可以作为医疗纠纷处理的有效证据，但是当前我国无纸化病案的相关法律法规还不完善。对此，我国应当进一步健全无纸化病案管理法律法规，促进我国无纸化病案的发展。同时，我国应当制定无纸化病案的管理规范，进一步规定和完善无纸化病案在存储及使用过程中的数据安全保护措施，并对危害无纸化病案安全的违法行为制定详细的处罚细则。

5 结论

综上所述，推进病案信息管理无纸化在现今智慧医院体系要求下显得尤为必要，当前暴露出的在相应法律法规、行政制度和管理惯性上的漏洞也亟待解决。总的来说，实现病案无纸化管理将是医院信息管理工作的一次具有里程碑意义的变革，其带来的便利不仅能惠及现阶段的医院管理层、一线工作人员和患者，也会为将来的医疗工作带来深远且深刻的影响和红利。因此，虽然在探索阶段不可避免地会遇到诸多波折，但仍应大力鼓励条件允许的医院积极推进病案管理无纸化的进程，探索路径，积累经验，形成可复制易效仿普适性强的运行模式，争取早日形成科学高效的无纸化病案信息管理体系。

PU-013

中文医学本体数据库建设现状与建设发展趋势分析

戴维[1] 赵文龙[2]
1. 重庆医科大学医学信息学院
2. 重庆医科大学医学信息学

1 引言

近年来，人工智能（Artificial Intelligence，AI）的迅猛发展为医学领域带来了翻天覆地的变化。从病患数据的收集，到大数据（Big Data）的集中处理、自动化的知识发现（Knowledge Discovery）；从早期疾病预防[1]，到临床辅助诊断[2,3,4]、术后延续护理[5]，医用人工智能呢都得到了长足应用。其中，自然语言处理（Natural Language Processing，NLP）是最为重要的应用场景之一，包含实体识别（Named Entity Recognition， NER）、文本分类（Text Classification）、文本生成（Text Generation）等多种多样的下游任务。由于医学信息特性，我们拥有海量的或结构化（Structured Data）或非结构化（Unstructured Data）数据供以科研目的使用，如何从中更有效率地挖掘出高质量数据，一直以来都是医学信息学领域重要的问题之一。

为获取更高效、准确的 AI 工具，积累高质量标注数据是必不可少的一环。然而，目前国内的相关科研工作较为分散，往往集中针对单一病种和需求，构建对应的语料库（Corpus）、数据集（Dataset），如药物不良反应[6-8]、冠心病[9]、肝癌[10]等。当需要解决的医学问题改变，往往没有一套统一的标准将他方数据转移至新的应用场景上来。且由于医学数据的保密性质，许多数据集无法对外公开，导致科研过程中的复现与探索更加困难繁杂。即使有预训练模型（Pre-trained Models）等工具减轻开发压力，在实际解决临床问题时，数据收集、清洗、标注仍然是最为费时的工作部分。进一步提高数据处理效率，解决不同数据集信息共享问题，是未来医用人工智能发展过程中必不可少的一环。目前，建立一个统一标准、覆盖面广、及时更新的中文医学本体数据库，是较为可行的方向之一。

2 中文医学本体数据库

2.1 本体与本体数据库

本体（Ontology）作为一种语义数据模型，用于定义某个指定领域范围中各种概念的类型、属性及各概念间的关系。作为数据库内容存储时，各本体拥有唯一性，以保证其概念准确，无法与其他本体混淆。其最为直观的应用便是近年来收到广泛关注的知识图谱（Knowledge Graph）技术。知识图谱是结构化的语义知识库，以符号形式描述实例化的本体及其相互关系，构成网状的知识结构，在解决临床知识发现[11]和构建智能医学问答系统[12]等需求中起到了关键作用。另一方面，本体数据库也可用于多种术语标准的对齐，如将国际疾病分类标准（International Classification of Diseases，ICD）、GB/T 14396-2016 疾病分类与代码，乃至更加复杂的临床常用俗称等内容综合起来，建立具有唯一性的本体的数据库。本体数据库是在知识图谱、语料库、术语标准等内容之上进一步概念抽象、整合所得到的综合结果。此类数据库使得各医用人工智能开发时的数据能够拥有统一标准、本体唯一的知识网络，进而节省多语对齐、同义词替换、指代消解等自然语言数据处理的成本与可能存在的模糊、歧义等问题。

2.2 国内外研究现状

目前，最为全面的医学语料库和术语标准当属美国国立医学图书馆研发的 UMLS （Unified Medical Language System），该系统属于本体数据库的一种，汇集了许多健康和生物医学词汇和标准，以实现计算机系统之间的互操作性。UMLS 语料库知识源的核心和使用最多的是来源于涉及 19 个语种的 224 个词表源的超级叙词表(Metathesaurus),其使用量约占总使用量的 94%[13]。 最新版本(2022AA),超级叙词表包含有 455 万多个概念和 1680 多万个唯一概念。UMLS 采用术语组织的方法，将表达同一事物的不同来源的不同表达集中在一起，形成一个概念，同时选择一个较为通用的词作为优选词(Preferred Term)来表达这个概念，并对这个概念分配一个概念唯一标识(Concept Unique Identifier)。医学系统命名法-临床术语（Systematized Nomenclature of Medicine -- Clinical Terms，SNOMED CT ）则是当前临床应用最为广泛的术语标准，涵盖如疾病、症状、操作、微生物、药物等方面的临床信息，支持实现临床数据的标引、存储、检索和聚合，便于计算机处理；同时，它还有助于组织病历内容，减少临床照护和科学研究工作中数据采集、编码及使用方式的变异。中文 SNOMED 电子版含 145,856 词条，包括解剖学、形态学、功能 、活有机体、

化学制品/药品和生物制品、物理因素/活动和力、职业、社会环境 、疾病/诊断、操作、连接词/修饰词等 11 个模块，通过层次结构的树型构造表达。每个词条的内容包括: 编码、中文名、英文名、类别符、层次、与该词条相关的外部编码、ICD-9-CM 码、药品编码、药厂编码、酶编码及 SNOMED 相关词条的交叉参照列表。

语料库、医学术语标准等建设工作在国外已经有多年研发历史，但国内医学领域的相关工作或是较为滞后、规模不足[14, 15]，或是实际应用在临床问题时难以推广和准确实施[16]。另一方面，直接以构建本体为目标的大型数据库则更鲜少有探索，大多为多语种翻译对齐、写作教学等尝试，中文医学本体的专用数据库仍有相当大的发展空间。目前，国内较为成熟的医学相关本体数据库是由粤港澳大湾区数字经济研究院（福田）与清华大学统计学研究中心合作研发的生物医学信息本体系统（Biomedical Informatics Ontology System，BIOS)[17]。该项目是一项长期系统工程，由 AI 工具辅助和大量的专家人工校验完成，其成本耗费巨大，且不可避免的存在错误、模糊数据等问题。同时，其源数据主要来源于 Pubmed 等英文为主的权威医学文本，对中文内容处理仍稍显薄弱。

3 医学本体数据库发展建设趋势

3.1 众包合作与数据脱敏

AI 的迅速发展离不开开源合作模式与数据共享，而海量的医学数据应用到 AI 中的只有冰山一角。众包（Crowdsourcing）模式作为一种协作机制，引入医学专家和志愿者，能在数据共享、开发合作过程中发挥关键作用。尤其面对体量庞大的中文医学本体数据库，由合作者协助处理大规模的数据集，提高工作效率，降低成本；由专家抽查审核，或培育更多合作者，能够在可接受的范围内更快、更好地构建一个强大的中文医学本体数据库。众包模式在医学数据处理方面已有部分实践[18,19]。

由于医学数据的特殊性，目前仍存在另一显著问题，即数据脱敏（De-identification）。脱敏在当今医学信息学和数据隐私保护领域十分重要，旨在在维护患者隐私的同时，确保敏感医学数据的可用性与合规性。通过脱敏技术，可有效剥离医疗数据中的患者身份信息，如姓名、地址和社会安全号码等，以避免未经授权的访问以及隐私侵犯，同时仍然保留着包括疾病诊断、治疗方案和病例历史等关键的医学信息，用以 AI 学习使用。在构建中文医学本体数据库的进程中，制定严格的数据脱敏标准显得尤为必要。标准应涵盖哪些信息需要被脱敏、脱敏的具体方法和程度、如何验证脱敏操作的有效性等。这一规范性框架的建立不仅有助于确保医学数据的安全性和隐私保护，还有助于遵守法规和伦理规范，以维护患者的隐私权。

此外，积极鼓励医疗机构和研究机构参与数据脱敏和数据共享，对于中文医学本体数据库的构建至关重要。通过数据共享，不仅可以扩大数据库的规模，还能增加数据的多样性、代表性，提升相关 AI 模型的准确性、泛化性，从而提升医学本体的整体质量。

3.2 AI 助力与反馈推进

在面对数量极为庞大的医学数据时，AI 可以参与并助推大数据处理与数据挖掘的每一步流程。以基于 Transformer 架构[20]的框架为例，可在构建中文医学本体数据库的过程中，较为准确地从医学文本中提取实体，如疾病、药物、症状等。此过程的自动化显著降低了数据标注的人工工作量，提高了构建速度[21]。在获得大量可用的医学实体数据后，进一步构建知识图谱[22]，并最终结合 ICD-10 等标准进行实体对齐[23]，抽象后构建中文医学本体，以确保数据的一致性和互操作性。AI 的参与不仅降低了时间成本，还减少了依赖专业领域专家进行手动标注和对齐的需求。这意味着专家可以将更多的时间用于专业的数据校正和质量控制，提高数据库的可信度。同样的，利用逐步扩建的中文医学本体数据库，其数据可以反过来用于 AI 模型的训练，提高模型的准确性和泛化性。通过将模型训练于更多的高质量数据和标准规范之上，AI 能够更好地理解和处理医学信息，为临床决策、研究和教育等领域提供更可靠的支持。这种互惠的关系使得 AI 和医学本体数据库相辅相成，为医学研究和临床实践带来巨大的价值。

3.3 应用场景交叉

中文医学本体数据库的建设不仅局限于为医用 AI 提供高质量数据支持，同时也能与多学科、多应用场景展开交叉合作，如生物信息学、心理学、医学教育领域等。这种跨学科合作趋势将推动医学本体的综合性发展，促进多领域知识的有机整合。特别以医学教育领域为例，中文医学本体数据库能够结合自然语言生成技术（Natural Language Generation），生成更高质量、用词准确的教育材料、课程大纲和测验题目等[24]。这些资源可以以医学本体的知识网络为规范，确保内容一致、标准一致，与最新的医学知识保持同步，从而减轻教育资源制作的工作量，提高医学教育的效

率。此外，医学本体数据库的知识结构可用于构建虚拟实验室和模拟环境，为学生提供实际临床经验的机会，加强实际操作技能的培养。这种交叉应用不仅有助于医学教育的提升，还为医学知识的传播和应用提供了新的途径。未来，中文医学本体数据库与自然语言生成技术的结合将为医学教育领域带来更多创新，支持医疗从业人员的培训和发展。这也强调了医学本体数据库在不同领域之间的关键作用，为多学科合作和知识整合提供了坚实的基础。

4 结论

在本文中，我们深入探讨了中文医学本体数据库的建设现状和发展方向。当前国内医学领域的进一步信息化建设相关工作相对不足，对于医用人工智能，尤其是自然语言处理相关领域，在标准化、数据共享方面仍存在艰巨挑战。面对这些问题，建立一个统一标准、广泛覆盖、及时更新的中文医学本体数据库是可行的道路之一。中文医学本体数据库未来的建设与发展将受益于脱敏医学数据的众包合作、数据建设与 AI 工具迭代的相辅相成和跨学科、跨应用场景的交叉合作等方面，在现有的科研基础上集中资源，为未来医学信息学的发展提供高质量、高标准的数据支持。综上所述，中文医学本体数据库的建设具有广阔的前景和潜力。希望未来会有更多的研究和合作致力于推动这一领域的发展，以更好地满足医学信息学的需求。

PU-014

智慧医疗背景下全椒县医院信息系统一体化建设研究

陈 旭

全椒县人民医院

信息化建设有利于提高医院的医疗服务水平，利用信息化平台整合信息和资源，合理分配医疗资源，提高医院的工作效率和工作质量。医院的信息化建设发展要与地区发展模式相契合，全椒县人民医院按照国家信息化建设标准规范发展信息系统，并和县域各级医院联合起来共同发展，通过信息系统一体化建设，打造一个全县系统化、一盘棋的信息化建设发展模式。在信息化技术的帮助下，充分利用区域内的医疗资源，提高医院内部各类信息管理的效率和质量，加强不同科室之间的协作与沟通，提升医院服务水平和患者体验。同时，医院信息系统一体化建设也为全椒县的卫生信息化建设提供了支持，有助于推进全面的健康信息化进程。

1 建设背景

我国一直非常重视智慧医疗的发展，并出台一系列相关政策，以推动和规范智慧医疗的发展。国家发展改革委员会、国家卫生健康委员会等多个部门联合发布的《互联网+医疗健康服务行动计划（2018—2020 年）》提出，通过互联网、大数据、云计算等技术手段，推动传统医疗机构与互联网企业深度融合，发展远程医疗、移动医疗、健康管理等新模式。国家发展改革委员会等部门联合发布的《人工智能三年行动计划（2018—2020 年）》明确提出，鼓励在医疗领域推广应用人工智能技术，包括医学影像诊断辅助、智能辅助决策、个性化医疗等方面，提高医疗服务效率和质量。国家卫生健康委员会发布的《国家健康信息化建设规划（2019—2025 年）》明确提出，推动健康信息化与智慧医疗深度融合，加快推进电子病历、医学影像、远程医疗、健康档案等领域的信息化建设。《健康中国 2030 规划纲要》将智慧医疗和健康信息化列为重要内容，提出要加快健康信息化、健康大数据建设，促进医疗资源共享和医疗服务协同发展。

全椒县人民医院积极响应国家发展智慧医疗的相关政策，提出“5431”信息化建设工程，通过信息化手段联通医院和全椒县乡镇卫生院，包括各家医院各个业务部门，形成全县医疗信息化一张网，促进全椒县人民医院及医共体内医院的信息化保持高度一致。建立和完善以电子病历为核心的医院信息系统，是实现现代化医院管理目标，促进智慧医院建设的重要举措，也是医疗质量提质增效的有力抓手。全椒县人民医院于 2019 年开始将电子病历五级评审和互联互通四级甲等评测建设工作纳入医院的重点工作目标并开始实施，2021 年医院通过信息互联互通标准化成熟度四级甲等测评，2023 年顺利通过电子病历系统功能应用水平五级评审，信息化建设迈入了一个新的台阶，电子病历应用水平迈入国家高级别医疗机构行列。

2 建设内容

2.1 建设目标

全椒县人民医院领导班子非常重视信息化建设，在院党委、院领导班子的筹划和指导下，提出“5431”信息化建设工程，“5”是指电子病历 5 级，“4”指互联互通四级甲等，“3”指智慧管理、智慧服务三级，“1”指全椒县区域信息平台建设。即按照国家的信息化建设标准规范，通过信息化手段联通县人民医院和乡镇卫生院，包括各家医院各个业务部门，形成全县医疗信息化一张网。

全椒县医院信息系统一体化建设的目标是实现县级医疗机构之间的信息资源共享、业务协同和管理整合，提升县域医院的服务能力和效率。通过一体化建设，实现县域内医疗机构之间的信息互通共享，包括患者基本信息、电子病历、医学影像、检验检查结果等数据的共享，方便患者就医和医生诊疗。通过整合各个医疗机构的业务流程和工作模式，实现业务的协同配合，提高医疗资源的综合利用效率，实现患者可以在县域内的不同医疗机构之间就医、转诊，医生可以共享患者的病历和检查结果，提供更加连贯和高效的医疗服务。通过一体化信息系统，实现对县域内医疗机构的整体管理和监控，提高管理效率和决策支持能力，包括人力资源管理、财务管理、药品物资管理等方面的优化。通过信息系统的支持，提升县域的医疗服务质量和安全水平。如智能化的临床辅助决策系统可以为医生提供更准确的诊断和治疗建议，避免患者因为信息不一致或重复检查带来的不便和风险。

总的来说，县域信息系统一体化建设的目标是优化县域医疗资源配置，提高患者就医体验和医疗服务效率，推动医疗质量的提升，这将有助于构建更加协同、高效和质量可控的县域医疗服务体系。

2.2 建设举措

2.2.1 电子病历五级和互联互通四甲

全椒县人民医院于 2019 年开始实施以电子病历 5 级和互联互通四级甲等为核心的信息化建设，并将该项工作纳入医院的重点工作目标，组织人员到省内和省外学习相关政策标准，同时邀请省内外专家来院内授课指导，开展各类研讨会、培训会、协调会共达上百场。

为稳步扎实推进电子病历应用评审工作，成立以院长为组长、分管院长为副组长的电子病历应用评审工作领导小组，下设电子病历五级办公室。办公室专职人员三人，一人负责临床医生组、一人负责临床护理组、一人负责协调信息科及信息管理员。信息科按照电子病历五级条款中的十大角色分工，每人负责一个。办公室每周组织相关人员针对电子病历五级条款中需要推进临床应用、强化数据质量等相关问题进行讨论，并制定工作进展报表以两周为一个周期向院长汇报，同时听取下一步的工作安排。当有全院级信息化培训或上线新软件时，电子病历五级办公室将通知全院各科系统管理员参加并由信息科负责人配合，真正形成上、中、下三方协同，夯实推进电子病历五级建设工作。

2.2.2 医共体内信息系统互联互通和医共体中心药房

全椒县人民医院在医共体内加快推进信息化建设，以电子病历 5 级建设为核心，以区域信息平台建设为中心，构建医共体内信息系统互联互通，并加强乡村医疗决策诊疗能力和远程医疗服务能力。

同时积极探索医共体中心药房建设，全椒县人民医院和县中医院牵头和所辖 10 家乡镇卫生院形成“7+3”模式。2018 年，全椒县人民医院成立医共体中心药房，实现了五统一，分别是：统一用药范围、统一带量采购、统一集中配送、统一支付结算、统一药学服务。2019 年 2 月 26 日，国家卫生健康药政司时任正司级监察专员汪建荣一行莅临全椒县人民医院调研县域医共体中心药房工作，充分肯定了医共体中心药房建设。

2.2.3 医共体内信息一体化建设和无纸化建设

2014 年，全椒县人民医院在老院区上线门诊一卡通系统时探索是否可以通过上线一卡通带动门诊无纸化建设，并开始构思医共体内信息一体化建设。2017 年，全椒县人民医院下辖的七家乡镇卫生院全面实现了门诊无纸化、门诊结构化电子病历。通过在全椒县人民医院实现门诊无纸化建设倒逼医生规范病历书写，实现医共体各镇卫生院标准、统一、同质化的信息化建设。过去医共体/医联体通过手工纸质进行“双向转诊”，存在统计困难、转诊管理不规范、追溯困难、信息不完整、资金流失等严重问题，经过无纸化改造，实现通过电子转诊单申请转诊，审批流程简化、效率更高、方便快捷，提高了患者满意度。此项工程曾荣获江苏省南京市卫生信息学会授予的医疗信息化建设优秀案例三等奖。同时，通过门诊无纸化加诊间支付等各项信息化建设来提高门诊患者看病的便捷性。据抽样调查统计，2021 年全年患者单次就诊结算时间平均减少 40 分钟左右。

2.2.4 银医通自助服务和远程医疗

2016 年，全椒县人民医院启动银医自助服务项目，共投入 75 台多功能一体自助设备，分布于本院、城东分院、二郎口、武冈、古河等乡镇卫生院，功能涵盖人脸识别、电子健康卡、医保卡、门诊病历打印、检验检查报告打印等。患者通过自助设备挂号缴费，住院病区内也可以通过桌面自助机实时缴费、查询，不仅提高了患者的就诊效率也减少了患者排队的时间。

2015 年起，全椒县人民医院与南京东南大学附属中大医院和滁州市第一人民医院、各乡镇卫生院实现了医联体、医共体内远程诊断、教学。通过上下联动加快信息化建设的步伐，提升乡镇卫生院服务水平和患者就医满意度。

2.2.5 医共体内县级与乡镇信息化建设具体举措

原全椒县二郎口乡镇卫生院只有单机版收费系统和医技科室的设备电脑，网络单一且闭塞，医保的报销需到县医保中心，给患者造成不必要的麻烦。全椒县人民医院接管二郎口分院后加强网络和信息化建设，投入大量的硬件和软件，给广大的患者带来非常便利的服务。

二郎口分院主干网由光纤直连全椒县人民医院中心机房，实现全院数据的共享。门诊区域、住院区域和手术室直接由光纤接入二郎口的交换机房，交换机房到各楼层由光纤连接，各楼层弱电间到各诊室都由千兆的网线连接，确保了数据传输的及时性和稳定性。

门诊区域的门诊收费处和住院收费处采用门急诊收费管理系统和出入转院收费管理系统，患者从门诊办卡、挂号、缴费和报销，到住院的入院登记、预交金和出院结算报销等，不管自费病人、医保慢性病人、贫困人口等，一站式服务减少患者的来回奔波。综合药房采用医院信息管理药房和药库管理系统，实现和全椒县人民医院的药库系统的无缝对接，实时查询药品的库存。

门诊区域各诊室使用门诊医生工作站系统，病人只需手持就诊卡，从检验检查到处方拿药，全程便捷地完成，避免了患者拿错相关单据的问题。同时病人的相关检验检查结果和处方单留存在就诊卡内，是病人下次就诊的宝贵参考资料。

门诊的检验科实现和全椒县人民医院检验科无缝对接，县医院所有的收费项目下沉，门诊医生可以开具任何地方的检验项目，在任何地方都可以查询和打印病人的检验报告，方便患者的上下级转诊的检验结果的共享。

门诊检查科室全部实现了和全椒县人民医院的无缝对接，县医院所有的收费项目下沉，门诊医生可以开具任何地方的检查项目。使用 PACS 系统在任何地方都可以查询病人的检查报告，方便患者的上下级转诊的检查结果的共享。同时影像中心的 CT 和 DR 分院检查完成后由全椒县人民医院的影像中心完成报告的书写和审核，解决了分院医疗资源不足的难题，切实给患者带来优质的资源和服务。

门诊自助机服务，分院采用和全椒县人民医院门诊相同的自助机服务，从门诊挂号、充值、门诊病历打印、检验结果打印、处方打印和影像报告打印一站式。

住院内科病区和外科病区医生站采用住院医生工作站系统，与全椒县人民医院患者信息资源共享，包括临床路径模板、科室病历模板、医生文书、医嘱项目等。病情严重患者只需转诊而无需转院，相关的出院结算可以在任何地方进行，大大节省患者的时间。

住院内科病区和外科病区护士站采用住院护士工作站系统，实现和全椒县人民医院的护士工作站资源共享，护理文书的模板、护理项目、检验条码打印实现了统一的规范，严重患者转诊到全椒县人民医院，护士只需点击转科就可以完成操作，给患者带来优质的服务。

2021 年 10 月 26 日，安徽省卫健委陶仪声主任一行莅临全椒县人民医院、二郎口、武冈等乡镇卫生院实地考察了医共体各成员单位通过信息化建设服务药事管理、规范医生诊疗、提升病历质量、引入自助服务等全椒县、乡镇信息一体化的建设成果。全椒县、乡镇信息一体化建设将为未来全县医疗信息整合和发展打下了坚实的基础。

3 建设效果

全椒县医院信息系统一体化建设自实施起历经四年左右的时间，2020 年 7 月，初步建成了以 CDR 临床数据中心为核心的集成信息平台，围绕服务患者、服务临床、服务管理、医共体/医联体内信息系统互联互通的建设目标，实现了一卡通用、患者信息高度整合、医共体成员单位门诊无纸化等建设目标，提升了全椒县人民医院及医共体所有乡镇卫生院整体医疗品质与服务。

互联互通四甲、电子病历五级的建设给医院发展及全县医院信息化水平带来的益处无法用一句话和标准答案来回答。可能要追溯到医院拥有的第一台电脑、第一个简易收费系统、第一个药房划价、第一台服务器，全椒县人民医院发展至今已有各类硬件设备 2000 多台，各类系统模块超百个，信息化设备由无到有，信息化水平由弱变强。医院信息化建设历经 20 多年，经过无数人的不

断付出、砥砺前行，信息系统打破了信息孤岛，不断迭代升级优化直至通过互联互通四级甲等成熟度测评，实现医院部门间业务系统环环相扣、星状网联，为医院的精细化管理提供了平台，打下了坚实的基础。

电子病历 5 级的建设实现了全院和部门间信息系统的无缝对接和互联互通，提高了全院人员对电子病历系统的掌握和应用水平，落实了合理用药、合理检查检验、诊疗闭环、临床路径、危急值等数据质量建设。医院信息化建设迈上一个崭新的台阶，提升了医院管理，提高了医疗质量，保障医疗安全，改善了患者的就医体验，真正实现“信息多跑路，患者少跑腿”，让广大人民群众在优质的就诊服务中，切身感受信息化、智能化变革带来的便捷、安全、高效。

全椒县人民医院发展本院信息化的同时，通过带动本县医院及乡镇卫生院发展，逐步形成全县信息系统一体化建设模式。实现了县域内医疗机构之间的信息共享和协同工作。通过信息共享和协同，更加合理地配置和利用县域内的医疗资源，提高了医疗服务的质量和安全水平。信息系统一体化建设可以提升医院的管理效率，同时为医疗数据的应用和科研提供了平台。

4 总结和展望

随着医院的长足发展，在更高级别的电子病历等级评审和互联互通成熟度测评目标促进下，医院将以电子病历六级、互联互通五乙、智慧管理四级、智慧服务三级为目标，利用信息化手段，不断优化业务流程，改善医疗服务质量，提高便民服务水平、改善就医体验，把信息化建设落到实处，促进全椒县人民医院及各医共体成员单位通过信息化建设在管理、应用、服务上持续改进，真正为了老百姓服务，为全县人民的健康保驾护航。

PU-015

关于“互联网+医疗健康”下的全民电子健康记录的思考

尚小平
郑州大学第一附属医院

前言

随着信息技术的快速发展和普及，全民健康信息化已成为当今社会的一种发展趋势。同时，"互联网+医疗健康"作为新兴的医疗模式，也为传统的医疗行业带来了新的机遇和挑战。如何把握这一机遇，解决当前存在的问题和挑战，提高医疗服务的质量和效率，是当前亟待解决的问题。

1.电子健康档案

1.1 建立健康档案

建立健康档案是电子健康档案的第一步，它是通过收集个体基本信息、家族史、生活习惯、疾病史、体检报告等健康相关信息，建立一套完整的电子化个人健康档案。这有助于为个体提供全方位的健康管理和医疗服务[2]。

1.2 记录健康信息

电子健康档案可以实时记录个体的健康信息，包括生命体征、运动数据、饮食数据等，这些信息对于个体健康管理和疾病预防具有重要意义。通过智能设备、医疗仪器等手段，可以实现对个体健康信息的实时监测和记录。

1.3 分析健康数据

健康档案可以对收集到的个体健康数据进行深度分析和挖掘，从而为个体提供更加精细化、个性化的健康建议和服务。例如，通过分析个体的生命体征数据，可以预测其未来可能患某种疾病的风险，并提供相应的预防措施。

1.4 提供健康建议

基于对个体健康数据的分析，电子健康档案可以向个体提供针对性的健康建议，包括饮食、运动、生活习惯等方面的建议。这些建议可以帮助个体更好地管理自己的健康，预防潜在的健康问题。

1.5 预约就诊服务

电子健康档案可以与医疗机构的信息系统进行对接，实现线上预约就诊、在线咨询等功能。这可以大大缩短个体就诊等待时间，提高医疗服务的效率和质量。

1.6 记录医疗信息

电子健康档案可以记录个体的医疗信息，包括诊断结果、治疗方案、药品使用情况等。这些信息不仅可以为个体提供更加全面、连续的健康管理服务，还可以为医生提供更加详细、准确的病人信息，有助于提高医疗质量和安全性。

1.7 智能健康设备

智能健康设备是电子健康档案的重要支撑技术之一。通过智能设备，可以实时监测个体的健康状况，收集个体的健康数据，并将数据传输至电子健康档案中。此外，智能设备还可以根据个体的健康数据和状况，提供个性化的健康管理和服务。

综上所述，电子健康档案是未来医疗领域发展的重要方向[3]。通过建立电子健康档案，可以实现个体健康信息的全面数字化管理和利用，提高医疗服务的效率和质量，为个体提供更加精细化、个性化的健康服务。同时，电子健康档案的发展还需要加强相关法律法规和隐私保护措施的制定和实施，确保个体健康信息的合法权益和安全。

2.电子处方

电子处方系统是一种基于信息技术（IT）的电子医疗记录系统，用于存储、处理和管理处方信息。这种系统可以大大提高处方的准确性和效率，减少处方错误，缩短患者等待时间。电子处方系统还可以提供实时的药物咨询和药物相互作用检测等功能，为患者提供更安全、高效的医疗服务。

2.1 电子处方的优势和挑战

电子处方的优势主要表现在以下几个方面：提高处方效率和准确性，减少处方错误。提供实时的药物咨询和药物相互作用检测，为患者提供更安全的医疗服务。实现患者医疗信息的共享和整合，提高医疗协同性。降低纸质处方的成本和环境负担。

然而，电子处方也面临一些挑战：电子处方系统的标准和规范不统一，不同系统之间的互操作性有待提高。电子处方系统的安全性和隐私保护问题，需要保障患者个人信息的保密性。电子处方系统的培训和维护成本较高，需要投入大量的人力物力。

2.2 电子处方与药师工作

电子处方系统的实施对药师工作产生了积极的影响。首先，药师可以更加便捷地获取患者的处方信息，进行药物咨询和药物相互作用检测，提高药物使用的安全性和有效性。其次，电子处方系统可以提高药师的工作效率，减少手动处理纸质处方的繁琐流程。然而，电子处方系统的实施也需要药师掌握相关的信息技术技能，进行系统的培训和学习。

2.3 电子处方在不同医疗场景的应用

电子处方系统在不同医疗场景中都具有广泛的应用前景。在门诊部[4]，电子处方系统可以缩短患者的等待时间，提高处方的准确性和效率。在住院部，电子处方系统可以提供实时的药物咨询和药物相互作用检测，保障患者的用药安全。在社区卫生服务中心，电子处方系统可以提高药师的工作效率，为居民提供更加优质的医疗服务。

2.4 电子处方的法规和政策

电子处方的法规和政策是推动电子处方系统发展的重要保障。目前，国内外都已经出台了一系列相关的法规和政策，规范电子处方的合法性和标准，推动电子处方的广泛应用。然而，由于不同国家和地区的法规和政策存在差异，电子处方的全球推广和应用也面临一定的挑战。

综上所述，电子处方是未来医疗发展的重要趋势，具有广泛的应用前景。然而，电子处方的推广和应用还需要解决一系列的挑战和问题，包括系统的标准规范、安全性隐私保护、培训和维护成本等。随着相关法规和政策的不断完善，相信电子处方将会在未来的医疗领域中发挥更加重要的作用。

3.电子病历

电子病历（Electronic Medical Record，EMR）是一种以电子化方式记录、存储和管理病人医疗信息的系统[5]。它包含了病人的基本信息、就诊记录、诊断结果、治疗方案、药品信息等，是医生进行临床决策的重要依据。电子病历具有存储方便、查询快捷、易于共享等优点，已经成为现代

医疗管理的基础。

3.1 电子病历系统的技术

电子病历系统采用了一系列先进的技术，主要包括以下几个方面：网络技术：电子病历系统通过互联网技术实现不同医疗机构之间的信息共享。数据库技术：电子病历存储在数据库中，以便进行快速查询和数据分析。数字签名技术：电子病历的修改和使用都需要进行数字签名，保证数据的完整性和安全性。加密技术：电子病历中的敏感信息需要进行加密处理，确保数据不被非法获取。人工智能技术：电子病历系统可以通过人工智能技术进行数据挖掘和分析，为临床决策提供支持。

3.2 电子病历在医疗管理中的应用

电子病历在医疗管理中具有广泛的应用，主要包括以下几个方面：提高医患沟通效率：电子病历可以快速传递病人的医疗信息，方便医生与病人之间的沟通。
减轻医务人员工作压力：电子病历系统可以实现自动化的医嘱处理和药品管理，减轻医务人员的工作压力。提高医疗质量：电子病历系统可以对医疗过程进行实时监控，及时发现和解决问题，提高医疗质量。优化医疗资源配置：电子病历系统可以实现医疗资源的统筹管理，优化资源配置，提高医疗效率。

3.3 电子病历在远程医疗中的应用

电子病历在远程医疗中具有广泛的应用，主要包括以下几个方面：优化诊疗流程：电子病历系统可以实现不同医疗机构之间的信息共享，优化诊疗流程，提高诊疗效率。提供实时医疗服务：电子病历系统可以实时传递病人的医疗信息，为远程医疗提供支持。提高医疗质量：电子病历系统可以对远程医疗服务过程进行实时监控，提高医疗质量。

3.4 电子病历的挑战与未来发展

虽然电子病历具有很多优点，但是在实际应用中也面临一些挑战：数据安全和隐私保护问题：电子病历中包含大量的敏感信息，如何保证数据的安全性和隐私保护是一个重要的问题。系统集成问题：不同的医疗机构可能使用不同的电子病历系统，如何实现系统之间的集成是一个挑战。用户培训和习惯改变问题：电子病历的使用需要医务人员进行相关的培训，同时需要改变传统的医疗习惯。

未来，随着技术的不断发展，电子病历将会实现更多的功能和更高的效率。例如，通过人工智能技术，电子病历系统可以实现疾病的预测和预防，为个体化医疗提供支持。同时，随着物联网技术的发展，电子病历系统将会实现与更多医疗设备的连接，实现更加全面的医疗信息采集和管理。

结论

全民健康信息化和"互联网+医疗健康"是当前医疗行业的重要发展趋势。未来，随着信息技术的不断发展和完善，全民健康信息化和"互联网+医疗健康"将在医疗行业中发挥更加重要的作用。同时，我们也应该认识到其中存在的各种问题和挑战，并采取有效措施加以解决，推动医疗行业的持续发展。

PU-016

门诊电子病历建设的国际经验与中国实践

吴倩[1] 陈瑶[1] 卓丽军[2] 陈芝芳[3] 李俊毅[4] 曾华堂[1*]
[1]深圳市卫生健康发展研究和数据管理中心
[2]华中科技大学同济医学院医药卫生管理学院
[3]深圳市康宁医院 医院办公室
[4]中山大学肿瘤防治中心 医疗质量管理科

完整的门诊病历能够真实反映门诊患者疾病的发现、诊断、治疗等整个医疗服务的全流程[1]，良好的病历质量在医疗、科教、医院管理、医疗保险、医疗纠纷等方面发挥着十分重要的作用[2]。以美国为首的发达国家建设电子病历系统已有 50 年之久，并率先将电子病历应用于门诊，在数据信息系统建设、电子病历质量管理、电子病历的数据应用上的探索颇多[3, 4]。鉴于电子病历的规模化部署，数据治理是运用电子病历开展卫生服务研究的基础[5]。随着医疗信息技术的不断发展，我国自 2004 年起开始推广使用电子病历。不同于国外，我国的电子病历主要应用于住院服务，后逐渐应用于门诊服务。2018 年，《电子病历系统功能应用水平分级评价方法及标准》出台，门诊电子病历系统建设是评价等级达到 4 级的基础，进一步推动门诊电子病历标准化的发展。2021 年，全国二级以上医疗机构门诊人次为 34.8 亿次，由于我国医院门诊工作量巨大且单个患者门诊就诊时间短的特点，相对于住院电子病历建设情况，门诊电子病历的发展相对缓慢[6]。传统的手写门诊病历具有病历资料的延续性较差、手写病历字迹潦草不利于患者阅读、病例由患者单向保管不利于医院职能部门监管等缺点，然而门诊电子病历也不仅仅是停留在通过“无纸化”来提高门诊服务效率，还是提高门诊医疗质量管理的有效工具[2]。本文系统梳理了国外门诊电子病历的发展特点、应用情况，并对国内门诊电子病历的政策演进、建设现状、质量管理及应用存在的问题展开讨论，为门诊电子病历建设的下一步发展方向提供参考。

1. 门诊电子病历建设的国际经验

1.1 应用范围广泛

20 世纪中末期，国外的门诊电子病历应用日渐成熟。电子病历的概念由电子医疗记录（Electronic Medical Record，EMR）、电子患者记录（Electronic Patient Record，EPR）发展为电子健康记录（Electronic Health Record, EHR）。国际上通用的 EHR 更适合于大健康领域，包括患者参与健康教育活动、保健服务、免疫接种等全生命周期的记录内容[7]。20 世纪 60 年代，电子病历系统应用始于美国麻省总医院的门诊服务。在英国的基层医疗领域，电子病历的覆盖率非常高，全科医师诊所（GP Practice）基本上使用了电子病历系统。荷兰的医疗卫生主管部门建立全国医院电子病历的接口平台，通过集成平台形成全国范围内的国家转接点（National Switching Point），实现各医疗组织之间能够安全保密的快速交换患者信息。该项目是全国信息框架的结构基础，建成后的集成的信息系统，医师能够随时随地访问到患者的医疗信息。

1.2 政府重视

持续性政府财政投入等支持性政策极大促进了门诊电子病历的发展。日本政府尤其重视电子病历建设，财政投入力度较大。1995 年，日本成立电子病历开发委员会，到 2006 年，财政共投入约 500 亿日元用于开发电子病历系统。并从上个世纪 90 年代初开始，日本厚生省开始组织研究电子病历的构造和设施间的数据交换的标准化、使用医学用语的标准化技术、安全技术灯基础研究工作。除了加大政府财政投入，卫生行政部门也会采取法律及经济激励措施来规范门诊电子病历的应用。2009 年美国出台《卫生信息技术促进经济和临床健康法案》，美国医疗保险与医疗救助服务中心通过经济奖罚并举的手段激励医师使用电子病历系统。调查显示[8]，2012 年至 2016 年，美国医师的电子病历使用率由 74%增加至 91%。英国政府规定所有的 NHS 医疗机构统一使用 SNOMO 临床术语集，保障采集的患者病例数据结构化[9]。

1.3 助力医疗质量管理

门诊电子病历通过信息化手段持续性归档患者门诊就诊记录，其迅速发展对医疗机构、医师、患者三方均产生正向影响。**一是提升医疗质量。**海量的门诊电子病历数据库为临床诊断决策提供强有力的支持，能够有效改善医疗质量，提高临床指南的遵守率，药物不良数量显著减少[10]。美国退伍军人管理局[11]将部分专科患者的数据成功应用于临床决策，如预测艾滋病患者的死亡率、辅助神

经性疼痛患者的临床治疗等。**二是促进患者慢性病健康管理。**大部分患者是在门诊完成慢病管理及健康促进的，一项基于 39209 名心血管患者的调查显示，基于电子病历的医嘱有效提高心血管内科患者的随访率[12]。

1.4 门诊电子病历应用存在的问题

门诊电子病历的广泛应用也带来一些问题。**一是增加门诊流程复杂性。**部分电子病历系统的界面不符合临床诊疗流程，检索与录入信息可能造成既定诊疗流程的中断，降低门诊工作效率，导致患者量减少[13]；**二是推行互联互通的电子病历系统成本较高。**2012-2016 年，美国的电子病历系统增长至 1100 个，为适应不同临床专科特点，电子病历系统的个性化较强，因此其兼容性较差，难以体现电子病历系统普及实现互联互通的优势[8]。2010 年英国实施 SCR（Summary Care Record）政策，即在患者同意下，患者的用药、过敏以及药物过敏等基本信息会在医院、社区医生（General Practitioner）诊所、社区药房等机构共享，但也出现了不同医疗机构间的信息系统不兼容、信息孤岛等问题；**三是增加患者安全风险。**门诊电子病历涵盖了患者隐私，其广泛的应用与推广可能诱发潜在的患者隐私泄露的风险[14]。目前，国外主要通过法律层面完善监管制度、技术层面加强信息加密手段、道德层面提高医师自律与患者自保意识等措施保障电子病历数据安全。

2. 我国门诊电子病历建设的政策演进

我国门诊电子病历建设主要历经以下三个阶段，**一是起步阶段**，即 2004-2017 年，该时期的电子病历建设主要应用于住院病案，大部分医院不重视门诊电子病历。2009 年新医改提出，要以医院管理和电子病历为重点，推进医院信息化建设次年开始，国家卫生部陆续发布一系列与电子病历建设相关的规范与标准，包括《病历书写基本规范》、《电子病历基本规范》、《电子病历工作试点通知》、《电子病历系统功能规范》等，电子病历开始在试点地区及医院试行。**二是标准化建设阶段**，即 2018-2020 年，该时期，仅约 5 成的医院逐步使用门诊电子病历系统，但对于其质量管理关注度不够。2018 年，《电子病历系统应用水平分级评价管理办法及评价标准》要求，到 2020 年，三级医院要达到评级 4 级以上。2019 年起，电子病历应用功能水平评价纳入国家三级公立医院绩效考核评价体系。**三是高质量发展阶段**，即 2021 年至今，2021 年《国务院办公厅关于推动公立医院高质量发展的意见》要求持续推进电子病历建设，强化信息化支撑作用。2022 年 2 月，国家卫健委强调高度重视国家健康医疗大数据服务，逐步建立全国统一的电子病历信息标准体系，实现互联互通、信息共享和业务协同。同年 6 月，国家卫健委出台的《医疗机构门诊质量管理暂行规定》强调，促进门诊病历等医疗文书管理规范化、标准化，“提高门诊电子病历使用率”首次纳入 2022 年度质控工作目标。此阶段，门诊电子病历建设与其质量管理受到高度重视，相关政策的出台进一步加强医院对门诊电子病历的质控管理。

3. 我国门诊电子病历建设面临的问题

3.1 门诊电子病历应用时间短，书写率偏低

国内门诊电子病历应用起步较晚，其应用主体局限于医疗机构。长期以来，建设与推广门诊电子病历是我国大部分医院的短板。因此，国家层面发布一系列政策，如 2017 年《电子病历应用规范（试行）》、2021 年《病案管理质量控制指标（2021 年版）》、2022 年《国家医疗质量安全改进目标》等，促进门诊电子病历使用率的提高。目前，门诊电子病历更适合于“大门诊小病房”的专科医院，如口腔医院等，更显其优势。门诊电子病历书写率偏低。国家医疗服务与质量安全报告数据显示[15]，2020 年在全国 2197 家综合医院门诊医疗质量调查中，门诊电子病历使用率仅为 51.15%，其中，三级公立医院门诊电子病历使用率为 67.09%，委属委管医院最低，仅为 40.89%。电子病历书写延长门诊工作时间，存在降低医师及患者满意度的风险。相关研究显示，所在科室是影响门诊医师电子病历使用意愿的主要因素[16]，内科、妇产科等患者流量较大的科室医师对门诊电子病历的使用意愿较低，不流畅的信息系统或不熟练的操作都可能会降低门诊服务效率，延长患者就诊等待时间。

3.2 门诊电子病历质量有待提高

门诊电子病历质量不高的原因主要有以下三点，一是门诊病历书写标准尚未统一，以致于各专科各级医师书写病历质量差次不齐。不同于住院电子病历的书写时间较长、患者主诊断明确等特点，门诊电子病历要求接诊医师面诊结束后即完成该患者病历的书写[6]，书写时间较短。我国门诊医师工作负荷重，单个患者接诊时间短，短时间内准确快速完成病历书写对门诊电子病历系统的成熟度及医师的熟练度要求非常高。由于我国公立医院门诊医师年龄结构整体呈现倒三角形，部分年龄较大的门诊医师受传统诊疗模式的影响，更习惯在患者的纸质病历上翻看病史、药方并作出详细记录，对于电子病历系统的操作并不熟练，且门诊诊断的书写习惯不符合医学诊断术语规范化表

达，不适应当前系统中门诊标准诊断库的使用。二是电子病历智能监测系统具有局限性。智能质控监测系统解决了门诊电子病历必要元素缺项、病历书写及时性等结构完整性的问题，但无法识别患者主诉不清、现病史描述不符或不相关、门诊手术等其他记录的完整性和合理性。三是门诊电子病历书写的内容质控管理缺乏有效监督。目前全国仅约有 1/3 的省份成立了省级门诊质控中心。2018 年《国家医疗服务与质量安全报告》数据显示，开展住院病案终末形式质控的医院占调查总数的 85.81%，而开展门诊病历质控的医院仅占调查总数的 32.01%[17]，门诊电子病历的质量管理重视程度仍存在不足。在医院层面，湖北省某医院成立了电子病历质控工作组，由工作组按季度随机抽取 100 份门诊病历，并建立门诊电子病历质量考核标准，分别对病历的完整性、合理性进行打分。近两次的季度抽检结果显示，满分病历率约为 20%，大多问题为病史描述不全、诊断、用药、检查不合理等主观评价。

3.3 门诊电子病历数据应用不充分，数据共享壁垒高

国内对于住院电子病历，即病案首页数据的分析与研究较为充分，主要集中在住院费用等疾病经济负担[18]，住院服务质量[19]的相关研究等。门诊电子病历数据应用不足的原因有以下两点：一是现行门诊电子病历的数据量较大，海量的诊疗记录散落在各家医疗机构的信息系统。2017-2021 年，我国二级及以上医院门诊人次数由 30 亿增至 34.8 亿，按照门诊病历电子化 50%[15]的比例测算，每年约有 18 亿条门诊电子诊疗记录；二是医疗机构的门诊病历系统与数据结构不统一，尚未实现医院间的门诊患者的检查、用药等病历数据共享，病历数据质量难以保障。

4. 现行门诊电子病历管理优化策略

4.1 提高门诊电子病历书写率是保证电子病历质量的前提

在建立门诊电子病历质量管理体系前，首先应保障门诊电子病历的书写率。在国家医疗质量安全改进目标、三级公立医院绩效考核及电子病历评级等硬性考核指标的要求下，绝大部分医院开始引入并更新门诊电子病历系统。国内门诊电子病历使用率仍处于较低水平，其质量监管更是处于尚未开发的阶段。孙艳艳[20]等通过院内 PDCA 循环发现门诊电子病历书写率偏低的原因主要是未落实医疗核心制度、医师责任心问题、质控部门监管力度不够、法制观念淡薄，通过院内加强质控监管力度等方式，门诊电子病历书写率由 78.07%提高至 100%。研究显示，医院门诊电子病历推广与院领导的重视程度[21]息息相关，而医师作为门诊电子病历的第一责任人，其对电子病例的接受与熟悉程度也是提高书写率的关键所在[22]，因此，需要通过落实制度建设来加强医师对门诊电子病历的认同感。

4.2 构建前中后全流程质控监测体系，优化门诊电子病历质量

构建结构化门诊电子病历实现前端质控。由于门诊疾病的复杂性、诊断更新迭代速度较快等特点，半结构化门诊电子病历的应用更为广泛。一是有机整合门诊电子病历系统与医院信息系统（HIS）、检查信息系统（LIS）、影响系统（PACS）、排队叫号系统等，门诊医师通过数据接口自动导入患者相关信息[21]，病历填写效率大大提高；二是通过结构化编辑器录入病历信息，将 WHO 公布的最新版 ICD-11 编码植入门诊电子病历系统，考虑到门诊病例的复杂性，定期将难定义的门诊诊断更新入库，标化门诊诊断及诊疗路径，门诊电子病历填写进一步规范；三是通过嵌入知识库、常规疾病规范病历模板、合理用药系统等智能功能，为门诊医师提供诊断及治疗提醒，辅助临床决策，降低诊疗差错率，提高门诊电子病历书写质量。针对部分专科，国内已有医院通过回顾既往门诊病历设计并制作结构化门诊电子病历模板，如膝骨性关节炎[23]、糖尿病全科门诊[24]等，进一步优化门诊的诊疗质量与效率。

智能质控系统过程监控与末端人工抽检相结合。在中末端门诊电子病历的质量管理上，绝大多数医院采取智能质控系统与人工抽检相结合的方式[25]。智能信息质控系统实现过程质量管理，人工抽检实现结果质量管理。但是，现行智能信息系统的审核机制仅能实现电子病例的完整性，并找出明显的逻辑错误，如缺项则无法提交、性别错误系统自动报错等。对于研判门诊诊断、用药、检查等操作的合理性等功能，智能信息质控系统尚未能实现。门诊电子病历的内容质控主要是通过人工抽检的方式完成，由门诊办公室、质控科、临床科室等相关部门共同成立门诊电子病历指控组，制定门诊病历质量评价指标体系，对各科室进行月度、季度抽检，对比各科室的电子病历质量，并将结果应用于科室与临床医师的绩效考核。人工抽检作为一种事后干预的手段，由于受到时间等现实因素的限制，无法对所有的病历进行抽检，同时也对门诊电子病历抽检制度的规范性及完善性、质控小组成员的专业水平提出较高的要求，具有一定的局限性。

4.3 扩大门诊电子病历数据的应用范围

基于良好的门诊电子病历数据质量，质控部门应充分发挥门诊电子病历在医疗质量管理中的作

用。通过门诊电子病历系统，可以了解门诊工作流程、诊疗行为及服务质量，识别门诊服务质量风险，建立事前防控监测体系。其次，门诊电子病历发展到一定阶段后，其在部分专科积累的医疗数据具有样本量大、反映临床实际需求等优势[26]。通过大数据手段定向采集并记录某一门诊专病病例的全过程诊疗路径，建立门诊专科数据库，为临床诊疗反哺基础科研提供数据支持。

4.4 门诊电子病历建设需契合医保支付方式改革大方向

未来的门诊支付方式改革将会对门诊电子病历质量管理提出更高的要求。DRG/DIP 改革后，住院费用得到有效控制。医疗费用不可避免向门诊端流动，因此，医院下一步的控费重点是门诊服务。国内门诊支付方式改革试点地区金华探索的门诊按人头包干结合门诊病例分组（Ambulatory Patient Groups，下简称：APG）与住院病例 DRGs 分组逻辑相似，将门诊病人临床特征、资源消耗和成本相似的归为一组，差异大的归为另一组。相关文献显示，APG 在门诊费用控制上具有明显优势，同时也对门诊电子病历的质量管理提出挑战[27]。一是需要大量门诊病历等数据资源支持。APG 分组主要是能通过门诊电子病历及信息系统提取关键信息，病历书写越规范，分组器分组越准确。二是探索开发门诊病历数据结合 APG 的门诊医疗服务绩效考核工具，通过门诊病历组合指数值、门诊时间消耗指数、门诊费用消耗指数等衡量门诊疾病的复杂程度[28]。

PU-017

基于人工智能的病案首页智能编码技术研究与应用

温丹

开封市人民医院病案统计科

本文旨在研究基于人工智能的病案首页智能编码技术，并探讨其在医疗领域的应用。病案编码是医院病案管理中非常重要的一环，但传统的手工编码方式存在效率低下、易出错等问题。因此，本文提出了一种基于人工智能的病案首页智能编码技术，旨在提高编码效率和准确性。

首先，本文分析了传统手工编码的问题。通过对医院病案编码过程的观察和数据分析，发现手工编码存在耗时长、易出错等问题。然后，本文介绍了人工智能在医疗领域的应用和发展情况。人工智能技术，尤其是自然语言处理和机器学习算法，在病案编码中具有广阔的应用前景。本文详细介绍了基于人工智能的病案首页智能编码技术的原理和方法。该技术利用自然语言处理算法对病历文本进行分析和理解，提取出关键信息，然后通过机器学习算法进行编码分类。通过不断的学习和优化，该技术可以自动识别和编码病历信息，提高编码效率和准确性。

最后，本文通过实际案例验证了基于人工智能的病案首页智能编码技术的应用效果。通过在某医院实施该技术，编码效率大幅度提升，同时编码准确性也得到了显著提高。此外，该技术还提供了编码质量的可视化分析和评估，有助于医院管理人员监控和评估编码工作的质量。

综上所述，本文通过研究和应用，提出了一种基于人工智能的病案首页智能编码技术。该技术能够有效解决传统手工编码方式存在的问题，提高编码效率和准确性，为医院病案管理提供了一种高效、智能化的编码方式。

PU-018

基于病案首页系统的数据挖掘研究

杨林朋

郑州大学第一附属医院

病案首页包括患者基本信息、住院过程信息、诊疗信息和费用信息。在大数据应用与患者的维权意识不断提升等影响下，病案首页数据的正确性显得尤为重要。病案首页信息的完整与规范填写可以影响医疗服务监督的整体质量，且能促进医院重点学科建设，尽量减少医疗纠纷，更高效的处理医疗保险相关事宜。数据挖掘是运用统计学、数学、情报检索、机器学习等方法，从大量的数据中通过算法搜索隐藏其中的信息。“互联网+5G”时代的到来，信息技术得到了快速发展，各种现代化技术广泛用于医疗领域，促进了医学信息的信息化、数字化，很多医院都提出了数字化医院的概念，并建立了大数据平台。随着病案首页的发展，数据规模变大，复杂度增加，形成了很多医院的病案首页。但是尽管有了大数据，真正能体现这些数据的价值，并用到医院临床中的非常少。病案首页只是体现在“大”的方面，并没有实现数据的有效利用。我们运用数据挖掘理论应用于病案首页，对病案首页进行分析，提取有意义、有价值的信息，对医院管理层做出决策，医生对病人进行正确地诊断治疗及其他临床研究都起重要的作用，目前，国外已有不少这方面的成果案例，显示出数据挖掘技术用于病案首页及疾病辅助诊断的广阔前景。本文在以往研究的基础上，探讨了医学数据挖掘的过程，说明了基于病案首页的数据仓库构建的关键技术及需要解决的问题。

1 医学数据挖掘的过程

数据仓库的目的就是在多个维度整合和归纳数据，包括数据清洗、数据集成、数据转换、数据加载（ETL 过程）等。数据仓库可以看作是数据挖掘的一个重要过程。另外，数据仓库还提供 OLAP 工具用于交互分析：多维数据在不同粒度上的分析——方便高效的数据整合和挖掘。可以继承到 OLAP 操作中的数据挖掘工具包括：关联分析、分类、聚类和预测等（这些操作都可以增强知识的挖掘）。因此数据仓库是医学数据挖掘的一个非常重要的平台。

2 数据仓库构建的关键技术

2.1 数据建模

数据的建模主要解决数据异构和异质的问题。其主要任务是将多个异构数据源，多种业务的数据，采用合适的数据拓扑承载数据，使其在业务系统中能够更好的组织和展示，这是解决数据异构的问题。而对于数据本身，则需要准确的抽象，例如，如何抽象和表示数据，一个业务数据可以抽象为多少个 measurement，多少个 counter，这是解决数据异质的问题。另外，还需要对数据的维度建模，根据医学数据的特点，可以抽象出多少种数据的维度，例如时间、业务对象等，以及在这些维度上如何做数据的聚合。

2.2 ETLoad 设计

ETLoad 是指数据的抽取（Extract）、转换（Transform）和加载（Load），主要完成原始数据向模型数据的转化工作。ETLoad 首先是要处理海量的数据流，能够在多个进程之间如在均衡。在获取文件之后要能够解析多种格式的文件并加载入库。在多层的数据集成系统中，ETLoad 还需要上层数据集成系统传递数据。ETLoad 还需要完成数据按照多个维度的汇聚工作。多维数据立方体操作已有大量研究。

2.3 OLAP 设计

OLAP 主要完成数据上卷（Roll-up）、下钻（drill-down）、切片（Slicing、Dicing）等工作。上卷（Roll-up）又称为整合（consolidation），即是数据的聚合，使得数据在一个或者多个维度积累和计算。上卷（Roll-up）就是为了预测趋势。相反，下钻（drill-down）是使得用户查看数据细节的技术，Slicing 是指抽出 OLAP 空间中特定集合的数据； Dicing 是指从不同视角查看数据片。

2.4 医疗数据挖掘平台服务框架

数据仓库是数据挖掘的基础数据存储平台，是面向多维数据分析的基础库。在数据仓库之上，可以构建各种数据挖掘算法库和个性化医疗服务的推荐系统等。

3 医学数据挖掘问题探讨

3.1 多学科交叉

医学数据挖掘涉及医学、统计、计算机等多个学科，研究内容偏向前沿问题和尚未解决的问

题，在数据尺度增加的情况下，问题更加复杂。生物医学工作者往往不能掌握复杂的分析工具，计算机科学工作者缺乏相应的医学背景，因此，培养能够在多个学科内长期积累和沉淀的数据人才至关重要。另外，从工具的角度来讲，把数据挖掘的流程模块化，把数据挖掘的工具组件化，把数据挖掘的过程智能化，可以大大的提供医学工作者的效率。

3.2 数据的集成和共享

数据的集成和共享，避免信息孤岛是行业信息系统的共同目标，在医疗行业实现该目标的挑战则更大。其原因主要是业务本身的复杂性和数据的多样性，各个医疗部门在构建系统之初，不可能采用统一的平台，这势必为后期数据集成带来困难。

3.3 统一数据接口

各个医疗部门采用不同厂商的硬件和软件，在内部数据格式上更是千差万别。为了开展有效的数据挖掘，制定统一数据接口标准至关重要。

3.4 数据挖掘与统计分析

HIS 中的数据挖掘与统计分析功能主要完成医院管理、医疗服务等领域的各种统计信息的生成和查询，数据仓库中数据的组织方式为进行这种分析与查询提供了可能，但仅仅依靠数据仓库并不能完成一些复杂的分析与查询，因此，为了完成上述功能，可以采用定期和不定期地对历史数据进行加工汇总。根据统计需要，统计汇总包括单项统计、复合统计。

数据挖掘实现对样本库的多因素、多层次、多导向的统计分析。统计分析一般包括统计描述和统计推断两部分。（1）统计描述是用统计指标、统计表和统计图描述资料的分析规律及其数量特征；（2）统计推断是以概率论与数理统计为基础,把数理解析过程作为主要统计方法，包括总体参数估计和假设检验两个部分。参数估计的范围是用样本统计量估计总体参数；假设检验则是用样本来检验事先对总体某些数量特征是否成立所作的假设。医院经济效益、社会效益的评价是建立在大量统计信息分析基础上的指标评价体系，这个体系包括：患者主要疾病疗效分析、门诊和住院工作量分析、床位和设备利用率分析、病人费用负担程度分析、手术情况及手术并发症发生情况分析、医疗质量分析、职工绩效分析分析等。

3.5 数据的安全性

为保护医院数据的安全性，应该从网络链路、数据传输、数据应用等层面，为数据安全提供保护。（1）网络链路的安全保护，通过虚拟专用网络（VPN）、流量控制器、防火墙等网络安全技术和设备，将对数据的各种操作限制在封闭的通道和环境内，内网操作可以有效防止数据泄露或被窃取。（2）数据传输的安全保护，由于安全环境不同，对数据安全威胁较大。因此，对传输的数据进行安全加密，将数据泄露影响降低到最低程度。（3）数据应用的安全保护对医疗数据中的敏感信息进行脱敏处理，实行授权登录和高强度密钥来保护账号安全问题，不同账号的数据使用权限通过角色管理来控制。

4 医学数据挖掘的应用

本应用针对医院 2018 年住院患者的数据进行分析，根据临床医生提出的课题建模要求，共入库 200 万个就诊病人，合并相同病人标识为同一个人后，共 20 万个不同的病人，采集 HIS、LIS、EMR 数据库中临床相关数据，完成从数据导入、集成、检索、样本队列生成、样本数据统计检验等，取得了很好的分析成果。实践说明数据挖掘能很好支撑医生的临床科研，特别是既能对医院各个业务系统结构化数据进行处理，又能对半结构化电子病历文本进行处理，在功能、性能上都非常符合要求，为对有科研需求的医生在较短时间内完成某个项目数据分析提供了坚实的基础。

5 结语

本文主要对病案首页挖掘的相关概念与技术做出归纳与总结，病案首页数据是卫生健康统计工作的根基，卫生健康统计工作是反映卫生健康发展现状、进程和规律的重要基础性工作，对于宏观决策、行业治理和服务社会具有重要支撑作用。2020 年 08 月 27 日，国家卫生健康委印发《关于加强卫生健康统计工作的指导意见》（国卫规划发〔2020〕16 号），对新形势下卫生健康统计工作提出新的要求，即必须牢固树立大卫生大健康理念，坚持依法统计、规范统计过程管理，坚持质量优先、保证数据真实准确，坚持应用导向、服务行业实际需求，坚持创新发展、加强新兴技术应用，进一步拓展覆盖范围，转变工作方式，强化服务效能。同时提出了两个阶段的目标，分别是力争到 2025 年，实现卫生健康统计调查体系、队伍建设、数据资源、方式方法日臻完备，统计数据真实性、准确性、完整性不断增强，统计工作法制化、规范化、信息化水平明显提高，有效支撑卫生健康事业高质量发展；到 2030 年，建立健全科学治理、权威统一、创新驱动、安全高效的统计工作体系，为实施健康中国战略提供重要支撑。下一步工作是在目前工作的基础上研究和实施数据

挖掘在病案首页中的应用。在实现临床数据互通共享的情况下便捷地开展远程医疗会诊，提高医整体的医疗技术水平，促进医疗水平与研究成果开放共享。

PU-019

互联网医疗患者满意度测量量表研究：系统综述

杜怡菲 谷雨

首都医科大学

互联网医疗是以互联网为载体、以信息技术为手段（包括移动通讯技术、云计算、物联网、大数据等）与传统医疗健康服务深度融合而形成的一种新型医疗健康服务业态的总称[1]。2015 年国务院发布《关于积极推进“互联网+”行动的指导意见》后，国家卫健委又陆续出台多项配套政策文件，2018 年国家卫生健康委和国家中医药管理局发布《互联网诊疗管理办法（试行）》《互联网医院管理办法（试行）》，2022 年发布《互联网诊疗监管细则（试行）》，为规范互联网诊疗活动、防范化解互联网诊疗安全风险、保障医疗服务安全和质量营造了良好的政策环境[2]。近年来，我国互联网医院数量呈现出爆发式增长，由 2018 年末的 4 家发展到 2022 年末的 1700 多家[3,4]；互联网医疗用户规模极速扩大，截止到 2022 年末，达 6.87 亿[5]；互联网医疗行业市场规模也显著扩张，截至 2022 年末，突破 635.5 亿元[5]。目前，互联网医疗已经成为医疗服务的重要组成部分之一[6]，是优化医疗资源配置、促进优质医疗资源下沉、推进区域医疗资源整合共享、提高医疗服务可及性和降低患者负担等的有效方式，在满足人民群众就医需求、改善医疗服务等方面发挥了积极作用[7-9]。

患者满意度反映医疗服务模式认可程度和接受程度[10]，是考察医疗服务质量及患者对医疗服务提供者的忠诚度的工具,更是医疗服务机构质量改进措施和服务发展策略制定的常用参考依据[11,12]。因此，在互联网医疗日益普及、视频诊疗模式越来越普遍的大背景下，本研究旨在梳理互联网医疗视听模式患者满意度的相关测量因素，制定一个更全面、更系统的患者满意度量表，为促进互联网医疗服务水平提供参考依据。

1 资料与方法

1.1 文献检索策略

本研究遵循 PRISMA 准则进行[13,14]，在中国知网（CNKI）、万方（Wanfang）和维普（VIP）三个中文数据库及 PubMed、EMBASE 和 Web of Science 三个英文数据库中进行文献检索。检索式由两个主要部分组成：“互联网医疗”和“满意度”。根据这两个主要部分，确定中文检索词：“互联网医疗”“互联网医院”和“满意度”，英文检索词：“telemedicine”“telehealth”“mobile health”“mobile medicine”“eHealth”“mHealth”和“satisfaction”。基于上述检索词，为每个数据库定制搜索策略，并在可用情况下使用主题词（MeSH）、自由词、检索过滤器和布尔逻辑检索[15]。检索时限为从 2020 年 1 月到 2023 年 2 月。

1.2 纳入与排除标准

纳入标准：①研究对象为使用视频/音频访问的互联网医疗的患者；②研究内容为患者对互联网医疗的满意度；③研究设计为横断面研究（定量或定性）；④评价视角为患者满意度。

排除标准：①非中、英文文献；②文献无法获取或不完整；③综述类文献；④无患者满意度问卷；⑤非视频/音频访问的互联网医疗。

1.3 文献筛选

根据上述纳入和排除标准阅读文献的题目和摘要，排除明显不符合纳入标准的文献，并对可能符合的文献进行全文阅读，以确定其是否完全符合纳入标准。

1.4 资料提取

阅读纳入文献的全文后，在 Excel 中整理相关数据和信息，包括作者、出版年份、国家、疾病类型、问卷维度和指标数量。

1.5 质量评价

为确保本研究的可信度和有效性，参照澳大利亚乔安娜循证护理中心（JBI）横断面研究偏倚

风险评价标准对纳入文献的质量进行评估[16]。

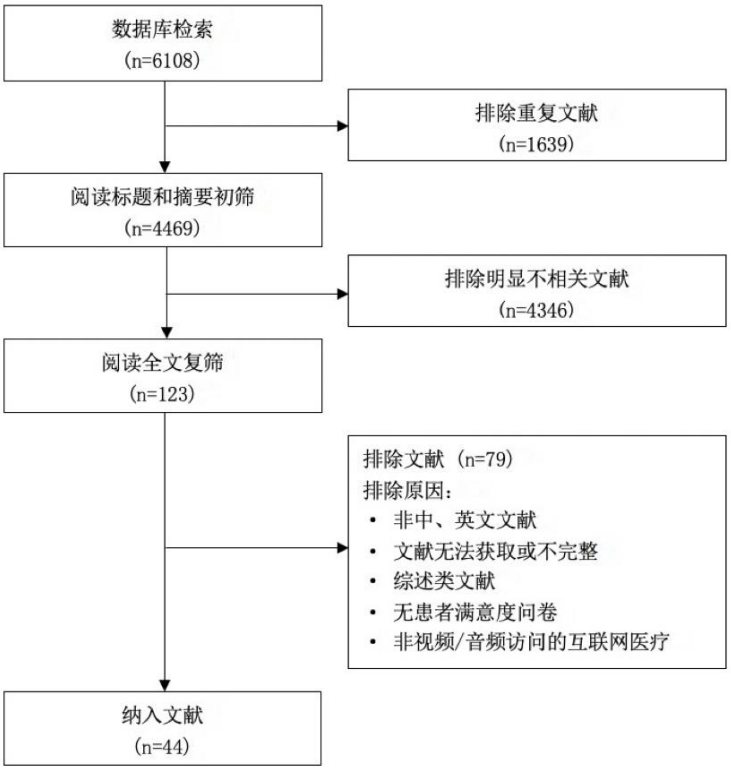
2 研究结果

2.1 文献纳入结果

文献筛选流程和文献检索结果如图 1 所示。使用上述检索策略，从 6 个数据库中检索到 6108 项研究。删除 1639 个重复项后，4469 项研究按标题和摘要进行筛选，4346 项研究被排除在外。其余 123 项研究进行了全文阅读，79 项被排除在外，最终纳入 44 篇文献。

图 1 文献筛选流程与结果

2.2 纳入研究的基本特征



44 篇纳入文献的基本特征如表 1 所示。纳入文献均为横断面研究，每篇文献都是围绕通过视频/音频访问提供的互联网医疗的研究。大多数纳入研究（n=32）在美国进行[17,19,22-31,33-39,42,43,47-56,60]，少数在意大利（n=3）[40,41,59]、西班牙（n=2）[18,44]、埃及（n=1）[20]、澳大利亚（n=1）[21]、印度（n=1）[32]、英国（n=1）[45]、加拿大（n=1）[46]、法国（n=1）[57]和哥伦比亚（n=1）[58]进行。对于疾病类型，绝大多数（n=38）提及了参与者的疾病类型，其中包括头颈部耳鼻喉科疾病（n=1）[17]、儿科疾病（n=1）[18]、需物理、职业和言语治疗类的疾病（n=1）[19]、整形外科（n=1）[22]、小儿肺部疾病（n=1）[23]、癌症（n=4）[24,34,41,46]、神经疾病（n=1）[25]、小儿泌尿科疾病（n=1）[27]、鼻科疾病（n=1）[28]、神经肌肉疾病（n=1）[29]、过敏（n=1）[30]、小儿糖尿病（n=1）[31]、癫痫（n=1）[32]、化疗前患者（n=1）[33]、小儿风湿病（n=1）[36]、神经外科疾病（n=1）[37]、囊性纤维化疾病（n=1）[39]、儿童和青少年 1 型糖尿病（n=1）[40]、需肩关节镜检查的疾病（n=1）[42]、儿童外科疾病（n=1）[44]、血管外科疾病（n=1）[45]、产妇心理健康和药物使用障碍（n=1）[47]、转诊患者（n=1）[48]、皮肤科疾病（n=1）[49]、血管内神经外科疾病（n=1）[50]、妇科癌症（n=1）[51]、骨科疾病（n=1）[52]、泌尿外科疾病（n=1）[53]、眼科疾病（n=1）[54]、肠易激综合征（n=1）[55]、镰状细胞病（n=1）[56]、肥胖病（n=1）[57]、儿童颅缝术（n=1）[58]、慢性神经疾病（n=1）[59]和结直肠手术（n=1）[60]患者，而其余的研究（n=6）并没有限制患者类型[20,21,26,35,38,43]。部分文献（n=18）的问卷分维度设计[17,20,21,23,26,27,29,36,38,40,41,43-45,47,53,56,59]。

表 1 纳入研究的基本特征

纳入研究	国家	疾病类型	问卷维度	指标数量
Layfield (2020)[17]	美国	头颈部耳鼻喉科疾病	5：有用性、易用性、有效性、可靠性、满意度	21

López Seguí (2020)[18]	西班牙	儿科疾病	-	10
Tenforde (2020)[19]	美国	需物理、职业和言语治疗类的疾病	-	7
Abdel Nasser (2021)[20]	埃及	-	2：参与者对远程医疗的满意度、参与者对远程医疗的态度	13
Bate (2021)[21]	澳大利亚	-	4：自信、咨询的整体质量、节省的成本、节省的时间	4
Bisson (2021)[22]	美国	整形外科	-	9
Capusan (2021)[23]	美国	小儿肺部疾病	4：技术、访问体验、总体满意度、再次使用远程医疗平台的可能性	30
Chang (2021)[24]	美国	癌症	-	7
Dratch (2021)[25[	美国	神经疾病	-	6
Drerup (2021)[26]	美国	-	3: 工作人员的友好性、预约时间的便利性、与医生的沟通	9
Gan (2021)[27]	美国	小儿泌尿科疾病	3: 就诊对获得医护的影响、患者/家庭体验、有效性	6
Hentati (2021)[28]	美国	鼻科疾病	-	7
Hooshmand (2021)[29]	美国	神经肌肉疾病	7：沟通、医生的及时性、图片质量、音质、隐私保护、体检的舒适性、接受远程医疗的便利性	8
Lanier (2021)[30]	美国	过敏	-	6
March (2021)[31]	美国	小儿糖尿病	-	12
Nair (2021)[32]	印度	癫痫	-	14
Sathiyaraj (2021)[33]	美国	化疗前患者	-	8
Shaverdian (2021)[34]	美国	癌症	-	20
Volcy (2021)[35]	美国	-	-	3
Waqar-Cowles (2021)[36]	美国	小儿风湿病	4：有用性、易用性、有效性、满意度	14
Yoon (2021)[37]	美国	神经外科疾病	-	8
Zimmerman (2021)[38]	美国	-	3: 医生的态度和行为、办公环境和员工、总体满意度和改进期望	14
Ahmed (2022)[39]	美国	囊性纤维化疾病	-	6
Bassi (2022)[40]	意大利	儿童和青少年1型糖尿病	4: 医疗保健的充分性、心理影响、可能优势和未来使用、远程护理	15
Cascella (2022)[41]	意大利	癌症	6: 有用性、易用性和可学习性、界面质量、交互质量、可靠性、满意度和未来使用	22
Cha (2022)[42]	美国	需肩关节镜检查的疾病	-	8
Chen (2022)[43]	美国	-	3：准入、医护提供者、总体评估	9

Cockrell (2022)[44]	西班牙	儿童外科疾病	9：提供者评级、办公室推荐、解释、倾听、问题、理解、病史知识、尊重、时间	9
Contractor (2022)[45]	英国	血管外科疾病	3：远程咨询的可接受性、远程咨询的好处、虚拟诊所未来的作用和可接受性	17
Gondal (2022)[46]	加拿大	癌症	-	8
Guille (2022)[47]	美国	产妇心理健康和药物使用障碍	4：整体医护质量、与面对面医护的相似性、获得医护的机会、MH护理	24
Jones (2022)[48]	美国	转诊患者	-	2
Kaunitz (2022)[49]	美国	皮肤科疾病	-	4
Majmundar (2022)[50]	美国	血管内神经外科疾病	-	7
Mojdehbakhsh (2022)[51]	美国	妇科癌症	-	11
Omari (2022)[52]	美国	骨科疾病	-	7
Sansone (2022)[53]	美国	泌尿外科疾病	5: 预约、技术、提供者、个人需求、总体满意度	19
Summers (2022)[54]	美国	眼科疾病	-	4
Yu (2022)[55]	美国	肠易激综合征	-	13
Zhang (2022)[56]	美国	镰状细胞病	4: 人际沟通、关怀、护理提供、熟练程度	24
Daouadji-Ghazou (2023)[57]	法国	肥胖病	-	9
Kilipiris (2023)[58]	哥伦比亚	儿童颅缝术	-	9
Rosellini (2023)[59]	意大利	慢性神经疾病	3：对当前远程就诊的满意度、对未来远程就诊的意见、医患关系的质量	11
Yao (2023)[60]	美国	结直肠手术	-	7

2.3 质量评价结果

根据澳大利亚乔安娜循证护理中心（JBI）横断面研究偏倚风险评价标准，42 篇纳入文献具有较低的偏倚风险[17-22,24-53,55-60]，而其余 2 篇纳入文献被评为中度偏倚风险[23,54]。

2.4 评价维度分布

本研究总结了 44 篇纳入文献的互联网医疗患者满意度的测量因素，通过整合发现互联网医疗患者满意度的测量涉及多个维度。为确保研究科学合理，有必要建立一个测量患者满意度的概念框架。如表 1 所示，有 18 篇纳入文献对问卷划分了维度[17,20,21,23,26,27,29,36,38,40,41,43-45,47,53,56,59]。基于技术接受模型[61]、整合型科技接受模型[62]和患者满意度问卷[63]等理论模型和现有经典问卷，以及纳入文献的问卷内容，本研究将互联网医疗患者满意度测量量表划分为 8 个维度：人文关怀、医患沟通、服务效率、可靠性、易用性、系统质量、有用性和总体满意度。

2.5 量表编制

根据上述维度划分，在 44 篇纳入文献中，14 篇文献涉及人文关怀，其中包括医生对患者的礼貌、友好和关心[24,26,32,34,37,38,40,43,44,47,51-53,56]。28 篇文献涉及医患沟通，其中包括医生对患者的倾听、医生向患者解释和医患之间的交流[17,18,20,22-24,26,29-32,34,36,38,41-44,46,47,49,51,53,55,56,58-60]。9 篇文献涉及服务效率，其中包括互联网医疗就诊过程中的准时性和及时性[17,29,41,42,47,53,58-60]。23 篇文献涉及可靠性，其中包括患者的问题、担忧和需求是否得到满足[17-19,23-25,27,28,32,36,38,40,41,45-47,49,50,52,53,55,56,58]。19 篇文献涉及易用性，主要关于医疗服务和系统技术的容易程度[17,20,23,26,30,32-

34,36,41-43,45,47,53-56,58]。27 篇文献涉及系统质量，即互联网医疗访问期间的系统支持情况，以及患者的个人隐私安全[17,20,21,23,25,27,29,30-33,36,37,41,42,45-47,50,51,53,54-58,60]。19 篇文献涉及有用性，即互联网医疗所带来的益处[17,20,21,23,25,27,29,32,36,38,40-43,45,47,53,55,56]。39 篇文献涉及总体满意度，即患者对互联网医疗的总体评价[17-20,22-43,45-48,50-53,56-60]。

基于上述分析，最终构建的互联网医疗患者满意度量表如表 2 所示，包含 8 个维度和 37 个指标，包括 36 个客观问题和 1 个主观问题。

表 1 互联网医疗患者满意度量表

维度与指标	参考文献	频率
人文关怀	24,26,32,34,37,38,40,43,44,47, 51-53,56	14
我的医生很有礼貌。	44,51,53,56	4
我的医生很友好。	26,34,38,51,56	5
我的医生关心我。	24,32,34,37,40,43,47,52,56	9
医患沟通	17,18,20,22-24,26,29-32,34,36, 38,41-44,46,47,49,51,53,55,56, 58-60	28
我的医生认真倾听。	26,44,47,53	4
我的医生详细解释。	20,22,26,30,31,34,43,44,46,51, 59,60	12
我的医生详细解释我的诊断和治疗。	26,34,38	3
我的医护人员技术娴熟，知识渊博。	18,32,38,47,51,53,56	7
我的医生问我是否还有别的问题。	17,23,38,41,44	5
我与医生的沟通很顺畅。	36,38,41,42,49,58	6
我与医生的沟通时间很充足。	18,22,24,26,29,44,51,55,60	9
服务效率	17,29,41,42,47,53,58-60	9
我的互联网医疗访问按时开始。	29,42,47,53,58-60	7
我的处方和订单被及时处理。	42,53	2
我相信使用互联网医疗系统可以提高效率。	17,41	2
可靠性	17-19,23-25,27,28,32,36,38,40, 41,45-47,49,50,52,53, 55,56,58	23
互联网医疗可以解决我的每个医疗问题。	18,24,38,40,47,49,52,53,55	9
互联网医疗可以解决我的每个医疗担忧。	19,28,50,53	4
互联网医疗可以满足我的每个医疗需求。	17,23,25,27,28,32,36,41,45-47, 53,56,58	14
易用性	17,20,23,26,30,32-34,36,41-43, 45,47,53-56,58	19
我预约看病很容易。	20,26,34,43,53,58	6
互联网医疗系统很容易学习。	17,23,41,56	4

互联网医疗系统很容易使用。	17,23,30,32,33,36,41,42,45,47, 53-56	14
系统质量	17,20,21,23,25,27,29,30-33,36, 37,41,42,45-47,50,51,53,54-58, 60	27
互联网医疗系统质量良好。	17,21,37,41,42,53,55,56,58	9
我可以很清楚的看到医生。	17,20,23,27,29,31-33,36,41,42, 45-47,51,54,56	17
我可以很清楚的听到医生说话。	17,20,23,27,29,31-33,36,41,42, 45-47,51,54,56	17
使用系统看到医生并与医生沟通，我感到很舒服。	17,20,23,25,27,30-32,36,37,46, 47,56, 57	14
我不担心我的隐私问题。	29,41,45,50,51,53,57,60	8
有用性	17,20,21,23,25,27,29,32,36,38, 40-43, 45,47,53,55,56	19
互联网医疗节省我的出行时间。	17,21,23,27,32,36,40-42,47,55, 56	12
互联网医疗是一种可以接受的医疗服务方式。	17,23,25,32,36,45,47,56	8
互联网医疗提高了医疗服务的可及性。	17,20,23,25,29,32,40,41,45,47, 55,56	12
我很容易找到我所需要的互联网医疗医生。	47,53	2
我被告知当我的症状恶化时应该怎么办。	38,43	2
总体满意度	17-20,22-43,45-48,50-53,56-60	39
我对医疗质量感到满意。	20,31,40,46,47,53,56	7
我喜欢使用互联网医疗系统。	17,23,36,41,42	5
总的来说，我对互联网医疗系统很满意。	17,23,25	3
总的来说，我对互联网医疗访问很满意。	19,20,24,27,32,33,35-38,42,45, 46,48,50-53,57-60	22
我对互联网医疗访问和面对面访问一样满意。	17,28,30,31,33,37,41,45,59	9
我会再次使用互联网医疗。	17,18,20,23-25,28-33,35-37,40, 42,45,47,48,50-53,56-59	28
我将会把互联网医疗推荐给其他患者。	22,26,27,30,33,34,42,43,51,59	10
改进期望。	24,33,39	3

3 讨论

3.1 量表注重调查对象普适性，测量维度多元化

本研究对用于不同疾病类型的互联网医疗满意度问卷中的评价维度进行总结提炼，依据多种理论模型和现有经典问卷指导量表维度设计，可供各类患者的互联网医疗满意度测量。目前，患者满意度评价方法多为以调查问卷形式为主的定量分析或以访谈形式为主的定性调查[10]。本量表运用定量指标确保大规模研究的标准化和精确化，再结合定性指标给调查对象发表个性化意见的空间，以

弥补定量指标的局限和缺陷。本研究打破现有研究的零散设计，通过构建互联网医疗满意度测量的结构框架并运用混合方法，建立更加科学和系统的互联网医疗患者满意度评价体系。

3.2 国外经验丰富，国内积极借鉴

本研究纳入的文献均来自美国和意大利等发达国家。我国互联网医疗发展起步较晚[64]，现阶段以图文咨询方式为主，图文模式可能存在回复速度较慢和信息不太准确等问题，已不能完全满足患者需求[65]，因此，不少患者开始选择视频或音频模式进行线上问诊。国外视听模式远程医疗服务患者满意度测量研究为国内测量问卷的设计提供了参考。

3.3 传统医疗与互联网医疗满意度的异同

对于传统医疗而言，诊疗服务是影响患者满意度的关键因素[66]，就诊环节、候诊时长和医疗费用等是影响患者满意度的主要指标[67]。对于互联网医疗而言，其依托互联网医疗平台开展服务，因此平台的系统质量是区别于传统医疗患者满意度评价的重要维度[10]。与线下医患面对面交流不同，在互联网医疗服务过程中，有可能出现网络连接不佳造成的视频或音频卡顿问题，这导致医患信息交流不畅、信息出现疏漏[68]，进而影响医生对患者的诊断和治疗效果。此外，互联网医疗高速发展的背后，一些问题也随之出现[69]，如患者隐私泄露等[70,71]，该问题因触及患者个人切身利益和涉及医疗社会信任而备受瞩目[72]，直接影响患者对互联网医疗服务的使用意愿。尽管互联网医疗患者满意度的测量变量与传统医疗存在差异，但无论是哪种医疗模式，医患沟通和医疗质量始终是影响患者满意度的核心因素。

4 小结

本系统综述研究了 44 篇涉及各类患者对互联网医疗满意度的文献。通过全面获取患者满意度的多维测量因素，制定了一个测量互联网医疗患者满意度的量表。该量表为我国互联网医疗患者满意度研究提供参考。本研究仍存在一定局限性：（1）本研究由于语言的限制，仅纳入中英文文献，可能会遗漏有效文献；（2）本研究重点关注通过视频/音频访问的互联网医疗满意度研究。

PU-020

基于手机 APP 延续性护理对心脏机械瓣膜置换术后患者抗凝治疗的影响

丁娜娜[1]，周家梅[2]，罗弦[3]，江雪[4]，王晓华[5*]
1.贵州医科大学附属医院护理部
2.遵义医科大学附属医院心外科
3.遵义医科大学附属医院神经外科
4 普定县人民医院
5 遵义医科大学医学信息工程学院

心脏瓣膜置换术是治疗中重度心脏瓣膜病的常规手术，可以提高患者的生存率[1]。通常情况下，患者在心脏机械瓣膜置换术（Mechanical Heart Value Replacement，MHVR）后需要终身抗凝治疗，以预防血栓和栓塞[2]。华法林作为一种口服维生素 K 拮抗剂，是 MHVR 后患者抗凝治疗的首选药物[3]，但华法林治疗窗口非常窄、半衰期长[4]，且抗凝治疗的效果容易受到药物、生活习惯（如饮酒）、情绪、饮食等的影响，如果抗凝不佳，会导致严重的并发症（血栓栓塞和出血），甚至危及生命[5]。有研究[6]表明延续性护理能提高 MHVR 后患者的治疗依从性，减少术后抗凝治疗相关并发症的发生。但传统的延续性护理方式都存在不同程度的缺陷，比如：形式单一、缺少实时的医患互动、人力物力资源消耗大以及信息量小等，已不能满足患者的延续性护理需求[7, 8]，因此需要探寻一种新的延续性护理模式。所以本研究研发了一款手机 APP—“瓣护”APP，对 90 例 MHVR 术后患者采用基于手机 APP 延续性护理干预，取得了良好的效果，报告如下。

1 研究方法

1.1“瓣护”APP 的开发

成立“瓣护”APP 的研发团队（心血管外科临床专家、心血管外科护理专家各 1 名、医学信

息学专家 1 名、软件设计师 1 名、1 名研究者）。构建的团队和某软件公司共同开发 MHVR 后延续性护理的“瓣护”APP，心血管外科临床专家和护理专家把控 APP 的设计方向，医学信息学专家给予指导，软件设计师负责 APP 的开发及系统维护，研究者在医学信息学专家的指导下进行 APP 的设计，并负责就“瓣护”APP 的设计及调试与软件开发设计师进行沟通交流。

1.2“瓣护”APP 的患者端模块和功能

①健康教育模块：患者出院后可随时了解相关的健康知识和医疗信息。②健康管理模块：患者记录运动、脉搏、血压等日常指标，医务人员可以通过医护端实时监测并给予反馈。③用药记录模块：患者记录服用药物的名称、剂量、用法、次数，设置用药提醒，提醒的方式有铃声或震动；华法林的剂量根据 INR 值由医生进行调控后，调控后的华法林剂量反馈给患者后由其在用药记录模块输入并保存，删除上一次的剂量，并设置用药提醒在每天晚上 21：00 提醒患者服药，提醒的方式有铃声或震动，服用后进行打卡。④化验记录模块：患者出院后复测的 INR 值、凝血酶原时间、钾、钠值手动输入记录，其他的化验单拍照上传，医护人员可以通过医护端随时查看并给予反馈。⑤护患咨询模块：患者出院后遇到或存在问题，可通过每周周二、五的互动答疑进行咨询，或随时留言，护士不定时查看留言，给予解答，如有不能解决的问题，及时咨询医生并给予正确的解答。⑥复查提醒模块：患者可以根据医生提示设置复查日期，提醒方式有铃声或震动。

⑦个人设置：包括患者的姓名、年龄、性别、手术名称等。

1.3“瓣护”APP 的医护端功能和模块

①护患咨询模块：设计用于与患者沟通。**②查看模块**：可以查看患者各个模块的健康动态。

2 “瓣护”APP 的应用

2.1 应用对象

纳入 2020 年 11 月至 2021 年 6 月贵州省某三级甲等综合医院心血管外科择期行 MHVR 的 90 名患者，纳入标准：（1）结合临床表现和心脏彩超诊断为 VHD 的患者，并于全身麻醉体外循环下首次进行 MHVR；（2）年龄 18-65 岁；（3）患者本人或者家属能有效地使用智能手机 APP 并愿意配合者；（4）小学及以上文化程度；（5）患者或监护人对本研究知情同意并签署知情同意书。排除标准：（1）认知功能障碍、精神障碍的患者；（2）患有严重慢性病的患者（肾、呼吸、肝衰竭、恶性肿瘤、冠心病、糖尿病、脑出血等）；（3）严重心力衰竭（心功能分级Ⅳ级以上）；（4）不配合研究。脱落标准：（1）干预途中自愿退出的患者；（2）干预期间死亡的患者；（3）未完成全部干预和数据收集的患者。中途干预组死亡 2 例、失访 2 例；对照组死亡 1 例、失访 1 例，最终全程参与研究的研究对象共 84 例，采用 SPSS 统计软件生成随机数字，以随机不透明信封实行分配隐藏并在信封表面依次进行编号，患者按照进入本研究的顺序开启相应的信封，研究对象就按照信封内的分组计划被随机分为对照组（45 例）或干预组（45 例）。本院医学伦理委员会已批准本研究的开展，研究人员向患者及家属讲解本研究的目的和内容，患者及家属充分了解后签署了知情同意书，从而参与本研究。

2.2 对照组

采用 MHVR 后常规护理。出院前一天进行健康宣教，出院后随访：出院后第 1 个月每半个月电话随访 1 次，之后每个月 1 次直至第 6 个月；结合患者复诊时间在出院后 1 个月、出院后 3 个月、出院后 6 个月进行门诊随访。

2.3 干预组

在对照组干预的基础上成立延续性护理小组，由 1 名心外科医生、2 名主管护师、1 名营养师、1 名研究者组成的护理小组。住院期间，延续性护理小组成员向患者讲解本次研究的目的，取得患者及其家属的配合和信任。延续性护理小组成员定期参与病区的查房，掌握患者的疾病发展情况、治疗现状、术后康复情况、患者及家属对疾病知识的认知程度，发现并预见患者需求。并在患者确定手术后的第 1 天，帮助患者或家属下载“瓣护”APP 并注册账号，同时告知患者或家属 APP 的使用目的。在安装后到出院这段期间教会患者或者家属能娴熟地使用 APP，提高患者的使用依从性。患者出院后，可学习 APP 中的健康教育模块中的内容，如遇到或存在问题，每周星期二和星期五可通过 APP 向护士咨询，或及时在 APP 上留言，护士不定期的检查是否有留言，并给予相应的解答。及时了解每个患者的健康状态，对于患者共同存在的问题，制作对应的解答内容在 APP 中公布；对于患者存在的个人问题，按照具体情况指引患者进行个体化的健康宣教知识阅读。同时读取患者后台使用数据，如超过 1 个星期，患者没有使用 APP，研究者本人则通过电话联系患者，告知积极使用 APP 软件的意义，建议其坚持使用。动态掌握患者康复进程。延续性护理小组成员通过 APP 对出院后 MHVR 后的患者进行 6 个月的随访观察。

2.4 评估工具

2.4.1Morisky 服药依从性量表

该量表由 Morisky 等[9]人制作，研究表明[10]：Morisky 服药依从性量表-8 中文修订版 Cronbach's 系数为 0.81，评定者间的一致性系数为 0.92，重测相关系数为 0.95，表面效度较好，CVI 为 1.00。量表共有 8 个问题，仅含一个维度。1-7 题的选项为二分类，即“是”和“否”，答“否”计 1 分，答“是”计 0 分，其中第 5 题反向计分，第 8 题选项为 likert5 等级评分，即“从不”、“很少”、“有时”、“经常”、“总是”，依次计分 1 分、0.75 分、0.5 分、0.25 分和 0 分。量表满分为 8 分，得分越高，依从性越好。

2.4.2 华法林抗凝知识问卷

是由司在霞等[11]人编制，条目 CVI 为 0.78～0.95，问卷 CVI 为 0.93，问卷 Cronbach's 系数为 0.89。该问卷共 18 个条目，其中前 17 个条目的选项为二分类，即“是”和“否”，答“是”计 1 分，答“否”计 0 分，得分越高，表明患者抗凝知识掌握越好；最后 1 个条目为多项选择，用于评估患者了解抗凝知识的途径。

2.5 资料收集及录入

为了避免因为资料收集人员不同造成的信息偏倚，所有资料均由研究者收集，在患者出院前一天，收集患者的一般资料、Morisky 服药依从性量表、华法林抗凝知识问卷、INR 值达标率以及抗凝治疗相关并发症发生率，在出院后 1、3、6 个月电话、门诊随访时或者通过“瓣护”APP 分别再次收集患者的各项指标。由双人录入所有资料并交叉核对。

2.6 统计学方法

本研究采用 SPSS 17.0 统计软件进行统计分析。呈正态分布的计量资料以（ $\bar{x}\pm s$）表示，组间比较采用独立样本 *t* 检验；呈非正态分布的计量资料以 M（QR）表示，组间比较采用 Wilcoxon 秩和检验；计数资料以频数和百分比来表示，组间比较采用 χ^2 检验。组内不同时间点比较采用重复测量方差分析。以 P＜0.05 表示差异具有统计学意义。

3 结果

3.1 两组患者的一般资料比较

对干预组和对照组的一般资料进行组间比较无明显差异（*P*>0.05），详情见表 1。

表 1 两组患者一般资料比较

项目	干预组	对照组	$\chi^2/t/Z$	*P* 值
个数（例）	41	43		
性别[例（%）]			3.063[a]	0.080
男	19（46.34）	12（27.91）		
女	22（53.66）	31（72.09）		
年龄（岁）	51.32±0.77	52.00±1.08	0.515[b]	0.608
民族[例（%）]				0.354
汉族	38（92.68）	42（97.67）		
少数民族	3（7.32）	1（2.33）		
居住地[例（%）]			0.017[a]	0.897
城市	11（26.83）	11（25.58）		
农村	30（73.17）	32（74.42）		
文化程度			-0.178[c]	0.859
小学	26（63.40）	30（69.77）		
初中	14（34.15）	7（16.28）		
高中或中专	1（2.44）	4（9.30）		
大专	0	1（2.33）		
本科	0	1（2.33）		

婚姻状况[例（%）]				0.235
已婚	39（95.12）	43		
离异或分居	2（4.88）	0		
心功能[例（%）]			-1.427[c]	0.153
Ⅰ	5（12.20）	6（13.95）		
Ⅱ	6（14.63）	14（32.56）		
Ⅲ	20（48.78）	15（34.88）		
Ⅳ	10（24.39）	8（18.60）		
手术方式[例（%）]			1.381[a]	0.501
机械主动脉瓣置换术	7（17.07）	4（9.30）		
机械二尖瓣置换术	18（43.90）	23（53.49）		
机械双瓣置换术	16（39.02）	16（37.21）		
病因			0.178[a]	0.673
RHD	21(51.22）	24（55.81）		
VHD	20（48.78）	19（44.19）		
EF(%)	54.90±1.41	52.30±1.45	-1.283[b]	0.203

注：民族、婚姻状况、居住地的比较采用 Fisher 确切概率法，a 为 χ^2 值，b 为 *t* 值，c 为 *Z* 值

3.2 两组患者华法林服药依从性、华法林抗凝知识问卷比较

两组患者不同时间点各指标比较，华法林服药依从性、华法林抗凝知识问卷得分时间效应、组间效应及交互效应均具有统计学意义（P<0.05），两组同时间点进一步比较，干预前各指标差异均无统计学意义（P>0.05），干预后，华法林服药依从性、华法林抗凝知识问卷出院后 1、3、6 个月得分比较，干预组均优于对照组，差异有统计学意义（P<0.05），分别见表 2、3。

表 2 两组患者华法林服药依从性比较（ $\bar{x} \pm s$，分 ）

分组	出院前一天	出院后			时间效应	组间效应	交互效应
		1 个月	3 个月	6 个月	*F*	*F*	*F*
对照组	5.13±0.22	4.89±0.19	4.49±0.17	4.14±0.23	29.654**	126.338**	72.373**
干预组	5.12±0.23	5.69±0.19	6.43±0.18▲#	9.29±0.23▲#⋆			
t	0.012	-2.999	-7.873	-15.894			
P	0.991	0.004	<0.001	<0.001			

注：**代表 *P*<0.001；▲代表出院前一天比较，*P*<0.05；#代表出院后 1 个月比较，*P*<0.05；⋆代表出院后 3 个月比较，*P*<0.05

表 3 两组患者华法林抗凝知识认知比较（ $\bar{x} \pm s$，分 ）

分组	出院前一天	出院后			时间效应	组间效应	交互效应
		1 个月	3 个月	6 个月	*F*	*F*	*F*
对照组	3.74±0.34	3.84±0.31	4.09±0.23	4.02±0.28	73.116**	447.475**	60.719**
干预组	4.12±0.23	7.17±0.24▲	9.63±0.31▲#	11.37±0.27▲#⋆			
t 值	-0.925	-8.416	-14.391	-19.170			
P	0.358	<0.001	<0.001	<0.001			

注：**代表 *P*<0.001；▲代表出院前一天比较，*P*<0.05；#代表出院后 1 个月比较，*P*<0.05；⋆代表出院后 3 个月比较，*P*<0.05

3.3 两组患者 INR 值达标率比较

两组患者在出院前一天 INR 值达标率之间无明显差异（*P*>0.05），在出院后 1、3、6 个月，干预组 INR 值达标率高于对照组（*P*<0.05），见表 4。

表 4 两组患者 INR 值达标率比较 n（%）

分组	出院前一天	出院后		
		1 个月	**3** 个月	**6** 个月
对照组	34（79.07）	22（51.16）	19（44.19）	18（41.86）
干预组	33（80.49）	30（73.17）	33（80.49）	35（85.37）
χ^2	0.26	4.311	11.728	24.838
P	0.872	0.038	0.001	<0.001

3.4 两组患者抗凝治疗相关并发症发生率比较

干预组中没有发生抗凝治疗相关并发症；对照组中有 6 例出现抗凝治疗相关并发症，其中 2 例发生牙龈出血，分别发生在出院后第 2 月内和第 4 个月内；另外 1 例在出院后第 6 月以内出现鼻出血，还有 2 例大便隐血阳性，分别发生在出院后第 3 月内和第 6 个月内；最后 1 例是在出院后第 6 个月内发生胃肠血栓。在出院后 6 个月以内，干预组抗凝治疗相关并发现发生率低于对照组（*P*<0.05），见表 5。

表 5 两组患者抗凝相关并发症发生率比较 n（%）

项目	干预组	对照组	***P***
抗凝相关并发症[例（%）]			0.026
是	0	6（13.95）	
否	41（100.00）	37（86.05）	

注：抗凝相关并发症发生率比较采用 Fisher 确切概率法

4 讨论

4.1 手机 APP 延续性护理能提高 MHVR 后患者华法林服药依从性

本研究结果显示：出院后 6 个月内，干预组华法林服药依从性显著上升，而对照组则是缓慢下降，且干预组华法林服药依从性高于对照组，说明手机 APP 延续性护理能有效的提高 MHVR 后患者抗凝依从性，这与张永春等[12]研究结果相似。其原因可能是：①APP 上有相关医务人员会在每周星期一和星期五进行互动答疑，患者可随时留言，护士不定时查看留言，进行解答。特别是相关药物的使用是否会影响华法林药效等问题上及时给予回答，降低患者不正确的联合用药、私自停药等行为的发生，提高了华法林如何用药的认知程度；②APP 不仅可以设置每天定时提醒患者及时服用华法林，服药后进行打卡，研究者也可以通过后台数据了解患者服药的情况从而督促患者按时服用华法林，患者忘记服用华法林的情况明显减少，而且 APP 还能设置复查日期，提醒患者按时复查及时调整华法林的用量；③因手机 APP 延续性护理小组成员和患者两者间建立了良好的医患关系，促使患者拥有较高的就医依从性。综上，将手机 APP 延续性护理应用在 MHVR 后患者的抗凝治疗管理上，不仅能让患者及时与医务人员沟通，获取抗凝治疗相关的专业指导，而且还能提高其华法林服药依从性。

4.2 手机 APP 延续性护理能提升 MHVR 后患者抗凝治疗相关知识得分

要提高患者及家庭照顾者的认知水平，专业的健康宣教是不可或缺的步骤，而医护人员的职责之一就是对患者及家庭照顾者进行健康宣教。但医护人员因临床工作负担重，所以健康宣教流于形式，导致患者及家庭照顾者不能获取细致的疾病相关康复知识和指导。且在我国心脏外科医生往往更注重住院治疗，会忽视后续治疗，特别是在我国中西部地区，缺乏专业的心脏手术后相关的医疗保健[13]。运用 APP 进行健康宣教可打破传统健康宣教模式时间、地域的限制，具有快捷、便利、时效性等特点，便于医护人员更好的进行健康宣教和患者获取抗凝治疗相关知识，特别是对出院后的患者，可弥补这期间医护人员健康管理的缺失，这将是完善和提升传统延续性护理的有效手段[14]。本研究结果显示：两组患者在出院前一天华法林抗凝知识认知得分无明显差异，出院后 6 个月内，干预组华法林抗凝知识认知得分显著上升，且干预组华法林抗凝知识认知得分高于对照组，这与王晶晶等[15]研究结果一致。作为医务人员，应拒绝常规填鸭式的健康宣教，而是根据患者自身对抗凝治疗相关知识需求给予针对性的宣教，所以本研究利用 APP，根据患者填写的问卷项目和后台阅读的数据，对于低分条目和阅读时长较短的内容，有针对性的提醒患者学习相应的内容，存在不理解的地方可通过 APP 咨询医务人员或者留言来取得专业的解答，从而巩固抗凝治疗认知薄弱的环节，最终形成牢固的抗凝治疗认知体系。对于对照组的患者只能通过短暂的门诊、电话随访进

行健康宣教，不能全面的了解患者需求，缺乏个性化健康宣教和指导。虽然对照组的患者可通过各种健康类 APP 进行学习，但是 APP 内囊括的内容参差不齐，不仅浪费时间，还容易混淆视听。所以通过强化健康宣教的干预方案促进患者获取相关内容，使患者获得积极的态度去改变自身行为。因此，手机 APP 延续性护理在帮助 MHVR 后患者掌握抗凝治疗相关知识方面比传统延续性护理模式更有优势。

4.3 手机 APP 延续性护理能提高 MHVR 后患者 INR 值达标率

维生素 K 拮抗剂因其治疗窗口非常窄，所以针对它的治疗相对复杂，则需要定期监测 INR 值来保证患者维持在正常的抗血栓范围内，不会因太高而出现出血风险或者太低而出现血栓形成风险[16, 17]。本研究结果显示：两组患者在出院前一天 INR 值达标率之间无明显差异，在出院后 1、3、6 个月，干预组 INR 值达标率高于对照组，首先可能与手机 APP 延续性护理在干预实行的全过程中，始终秉承着“以患者为中心”的核心理念，为患者的抗凝行为、认知方面提供“私人定制式”的护理服务，如患者把复查的 INR 值记录在化验记录模块中，便于医务人员及时查看以后做出反馈，从心理、生理上起到积极推动作用，提高了 INR 值复查率和华法林用药剂量准确性，从而提高 INR 值达标率；最后 INR 值达标率与华法林服药依从性成正比[18]，而干预组的华法林服药依从性高于对照组，则 INR 值达标率也随之增加。综上，手机 APP 延续性护理能有效的提高 MHVR 后患者 INR 值达标率。

4.4 手机 APP 延续性护理能降低 MHVR 后患者抗凝治疗相关并发症的发生率

本研究中干预组和对照组抗凝治疗相关并发症发生率（0%vs.13.95%）相比较后发现，该数据和其他临床研究[15, 19]数据相符，出院后 6 个月发现，干预组的发生率低于对照组，证实了手机 APP 延续性护理能有效的减少抗凝治疗相关并发症的发生。原因可能是：首先因为本研究利用 APP 与出院后的患者维持联系，进行需求答疑、健康宣教、用药和复查提醒及抗凝指导等服务，与其建立以维持健康为目标、互信互助、协调合作的医患关系，从而对终身抗凝治疗和术后恢复起到决定性作用；其次因为 Horstkotte 等[20]研究表明，不管 INR 值高值还是低值，并发症总是发生在 INR 值波动的时候而很少发生在稳定时，因干预组 INR 值达标率高于对照组，则抗凝治疗相关并发症的发生率低于对照组；最后有研究[21]表明：因药物使用不当而引起的出血和栓塞占 MHVR 后患者长期并发症的 75%。因干预组华法林服药依从性高于对照组，则抗凝治疗相关并发症的发生率低于对照组。

5 小结

本研究结果表明，手机 APP 延续性护理能提高 MHVR 后患者华法林服药依从性、华法林抗凝知识认知得分、INR 值达标率以及降低其抗凝治疗相关并发症的发生率。但本研究研发的“瓣护”APP 属于初步应用阶段，功能模块还有待完善和进一步改良；并且受人员、时间、经费等因素的制约，最终只纳入了 90 例患者，样本量相对较小，干预时间也只有 6 个月，随着下一步功能模块的更新，扩大样本量、延长干预时间以便于观察长期效果。

PU-021

互联网+下健康医疗的融合与发展

李振浩[1, 2]，龚万里[1, 2]，张梦超[1, 2]，张清泉[1, 2]，于琦[1, 2]
1. 山西医科大学管理学院
2. 临床决策研究大数据山西省重点实验室

当前我国城乡卫生资源布局、医疗卫生服务体系、医疗人才队伍、服务费用等方面,还存在着明显的结构性矛盾,呈现城乡医疗二元化特征[1]，而如何弥补两者之间的差异，使得医疗资源分布更加公平合理，更好的保障我国人民的健康，确保“健康中国”的顺利实施，是现阶段我国面临的主要问题。伴随着互联网以及通讯技术的持续升级。这一问题正不断被改善。近年来,国务院办公厅相继出台了《关于促进“互联网+医疗健康”发展的意见》和《关于促进和规范健康医疗大数据应用发展的指导意见》。山西省积极响应国家政策出台了《关于促进“互联网+医疗健康”发展的实施意见》提出夯实“互联网+医疗健康”基础、健全“互联网+医疗健康”服务体系、加强行业监管和安全保障等一系列措施更好的保障“互联网＋健康”的平稳落地。

根据《医改蓝皮书：中国医改发展报告(2020)》统计，截至 2018 年底，有开展远程医疗服务

的三级公立医院 1452 家，二级公立医院 2746 家，分别比上年增加 302 家和 504 家。其中，基层医疗卫生机构建立远程医疗服务网络的三级公立医院 1124 家，二级公立医院 1 386 家。三级公立医院、二级公立医院开展远程医疗服务分别为 608.6 万人次、368.1 万人次，比上年分别增加 406.9 万人次、222.6 万人次[2]因此通过以上数据得知，无论是基层医院还是二级、三级公立医院在开展“互联网＋”的数量上还是医疗门诊服务人数上都呈现出显著的增长。除此之外“互联网＋健康医疗”系统提出健康信息“惠民计划”、“智慧医疗行动”、公共卫生信息化、健康医疗大数据应用、医疗行业监管与信息标准及继续医学教育模式等新任务，并结合了相关保障措施,充分发挥互联网等技术在医疗卫生领域中的作用,加大拓展互联网与医疗卫生行业融合的广度和深度, 实现健康信息的有效整合,推动医疗行业互联互通和业务协同[3]。其次从微观环境来说，传统的医疗卫生服务获取过程耗时耗力，当前的传统医疗领域，医院、患者和企业的痛点长期存在，互联网的兴起，提供了更丰富的解决方案。特别是对于慢性病患者来说，互联网慢性病管理可以提醒患者用药、查看不良习惯、告知回访或在线回访时间，对慢性病患者的科学管理起到重要作用[3]。“互联网+健康医疗”服务模式是以互联网等技术为依托，结合医疗行业发展现状，整合线上线下医疗资源，推动优势医疗水平和医疗资源的均衡发展，充分发挥优质医疗资源引领作用，推进智慧医疗建设。因此相较于传统医疗模式医疗资源过于集中、就以方式效率低下、缺乏院后服务、财政支持不足，医保缺口明显等条件下互联网＋的医疗模式优点更加突出[4]。

对于人口数量庞大的中国，“互联网＋健康医疗”仍是新兴事物，尽管当前处在高速发展的阶段，但不可避免的存在诸多问题和挑战。[5]首先互联网＋医疗配套的法律和制度保障并不完善，没有形成独立体系和普通医疗行为法律共用。其主要原因是远程医疗相关的法律和制度 数量较少、时间严重滞后、且法律位阶较低。[6]其次于应用“互联网＋医疗”而导致的权责分配问题，由于新技术的应用导致的纠纷问题。如整个分级诊 疗体系中，电子病历从基层医疗机构到三级医疗机构的互通互认，从而引发的隐私保护等伦理问题，以及取证和“网络诊断”的合法性保护等都尚未有规范[7]。因此应以明确的政治制度架构解决“互联网＋ ”的瓶颈问题，做好国家层面的制度规范，创造良好的政策环境，更加侧重就医人员和医务人员的需求，不断完善医疗领域的政策法规，从而促进“互联网＋医疗”的健康发展。

随着经济和社会的发展与转型，我国健康治理面临着多重因素交织叠加的复杂局面，人们对健康服务的需求日益迫切，并有了新的更高期盼，因此我国从国家层面推动健康中国战略，切实解决人们最关心的健康问题。疾病谱的变化为健康服务和健康保障带来了新的要求，慢性非传染性疾病已经成为目前影响我国国民健康的主要病种；生态环境及生活方式的变化带来了健康生活、健康环境的压力。来自这些方面的压力促使我国卫生健康领域发生政策范式转移[8]因此未来医疗领域将发生重大变革包括民众层面、医院层面、诊疗层面。等等因此互联＋医疗要适应这三方面的转变，更好的服务群众便需要：不断重视“互联网＋”提升“互联网＋”的效率和可及性，同时要利用互联网的便捷性提升居民自我健康管理意识，提升自我健康管理水平，做自身健康管理第一责任人。前瞻产业研究院发布的《2015—2020 年中国医疗信息化行业市场前瞻与投资战略规划分析报告》数据显示，我国医疗信息化市场 2020 年有望突破 1000 亿元；CCW Research（计世资讯）的研究显示，我国远程医疗市场 2019 年达到了 27.5 亿元，预计 2020 年市场规模将达 172 亿元。互联网+医疗健康以互联网平台为载体，将新兴科技与传统医疗产业相结合而形成一种新型医疗服务业态。[9]对传统医疗行业进行创新, 改变健康管理方式、创新就医方式,改善就医体验,重构购药方式,重构医患生态;提高医疗服务效率,提升医疗水平 和健康管理能力,降低医疗费用,使患者享受安全、便利、优质的诊疗服务。医疗器械、养老服务、医药电商和移动医疗等产业将迎来新的发展机遇。

未来,互联网+医疗的模式将是线上线下相结合,整合各类服务资源,囊括用户患者的各个环节,把医院、医生和患者三者打通,让资源和信息有效流动,提供更加智能的医疗服务[10]。因此,医院所提供的服务势必要以患者为中心,改变医院服务模式,改变患者就医模式,解决医院资源紧缺问题,提升医疗诊断准确度和效率,让更多人享受到专业可靠的医疗服务。具体包括：要着眼于互联网医疗服务模式,进行核心业务流程重组，不断优化诊疗过程，着重提升患者就医便捷性。以优质持续服务为纽带构建医院独特的网络服务品牌，在提供同质化网络医疗服务的同时还应根据自身技术特点与专科实力,建立品牌专科,为患者提供独特的差异化服务。适应互联网医疗服务的即时性、有效性规则,合理调配服务资源。

经济的发展最终目的是为了改善人们的生活条件，提高生存质量。相应的随着通讯技术的不断升级得以让智能手机迅速普及，也让绝大多数人有机会接触到互联网并且享受到互联网带来了的便利，这也让“互联网＋医疗健康”有了发展的空间。在这中模式下，传统医疗结构将被打破，优质

的医疗资源也将不会被地区所限制，患者也逐渐掌握主动权。同时医疗服务显著迁移以及后延。帮助患者进行全生命周期的健康管理，最终实现高效、个性、精准的医疗服务。不仅体现了我国科技的显著进步，也同时体现了 “健康中国”战略的前瞻性和实践性。

PU-022

基于元分析的用户数字健康素养影响因素研究

姚艺倍 曹锦丹* 钟玉骏 蔡林 冯傅昭
吉林大学

目的 针对现有用户数字健康素养影响因素相关实证研究中存在结果不一致现象，对该领域相关研究进行梳理验证，旨在明确影响用户数字健康素养的真正因素以及关键调节变量，揭示影响用户数字健康素养的直接原因，为完善用户数字健康素养影响因素及相关领域的后续研究提供借鉴。
方法 选取 24 篇符合标准的中外文献进行元分析，探究影响因素及其影响程度，从国内外研究情境和健康状况两个维度探讨影响因素与用户数字健康素养间的调节效应。
结果 结果表明，19 个因素与用户数字健康素养相关，其中年龄、文化程度、社会支持、网络使用频率、搜索健康信息频率和计算机素养与用户数字健康素养呈现中等程度相关，居住地、婚姻状况、职业、个人月收入、家庭经济状况、年级、健康观念、自评健康状况、对待网络健康信息态度、每日使用网络时间、健康信息获取渠道以及学习压力与用户数字健康素养呈现低程度相关；同时验证了国内外研究情境、健康状况对数字健康素养具有调节作用。
结论 本文采用元分析方法，通过对前人研究结果的再分析，得到更加精准的研究结论，是对数字健康素养领域现有研究的有力补充。但本文未对各个影响因素自变量之间的影响作用进行研究，有待后续进一步研究。

PU-023

“互联网＋医疗健康”新业态

张梦超
山西医科大学

一、我国互联网医院的发展现状

随着互联网的迅猛发展及其对各个行业的不断渗透, 众多行业正不断被互联网颠覆与重构。这已是不争的事实, 关乎全民健康的医疗产业也不例外。近几年, 医疗行业与互联网渐趋融合, 正焕发出新活力, 发展前景广阔。互联网医疗是互联网在医疗行业的新应用, 是以互联网为平台和技术手段的挂号预约、医疗信息查询、医疗过程跟踪监测、在线疾病咨询、 电子处方、远程会诊及康复等多种形式的医疗健康服务。医疗行业正日益显示出医疗服务网络化、医疗信息数字化、可穿戴医疗设备普及化的变革趋向, 搭载了互联网平台及应用的医疗产业在医疗服务与管理方面越来越高效化、智能化、便利化, 病人则得到更加优质周到而又个性化的医疗服务, 提高了健康水平。前不久在国家会展中心 (上海) 举办的第 75 届中国国际医疗器械博览会, 展出了各种各样的医疗网络信息化服务方式及数字化医疗可穿戴设备, 让人进一步感受到互联网医疗的强大魅力。互联网医疗提供高效智能的信息化工作平台, 将一部分线下的医疗活动转移为线上的服务与管理, 为医务人员的工作提供了便利, 节省了就医人员的时间成本。以阿里巴巴和腾讯为代表的互联网科技企业纷纷抢滩互联网医疗, 积极与医院达成合作, 为互联网医疗的发展提供技术支撑与平台构建。资本市场的大力追捧更为互联网医疗的发展提供了强大的驱动力。与此同时, 移动医疗 App 如雨后春笋般大量涌现, 不断满足了人们多样化的医疗健康需求。据不完全统计, 国内目前的移动医疗 APP 已达 2000 多款。随着全国智能手机普及率的提高以及移动宽带网络和服务的拓展, 移动医疗无疑将在未来的医疗保健业发挥重要的作用。

二、我国互联网医疗健康产业发展趋势分析

（一）互联网医疗健康服务的产业实践主要包括两种发展路径:

1.是互联网企业从政策“弱监管”的领域入手，逐步向医疗服务的“深水区”渗透。例如，好大夫在线、微医集团 (挂号网) 、贵阳朗玛等企业在推出健康咨询、健康管理、预约挂号等服务后，开始联合医疗机构开办“互联网医院”，提供远程诊疗、电子处方、在线医嘱等医疗服务。

2.是医疗机构从内部信息化开始向更广泛的互联网健康服务扩展。例如，郑州大学第一附属医院、中日友好医院等大型甲级医院联合基层医疗机构形成网络“医联体”，建设“甲级医院+底层医疗机构”的远程医疗网络，为基层医疗机构提供远程会诊、病理诊断、远程手术指导、远程双向转诊、远程医学教育培训等服务，促进优质医疗资源下沉到基层。同时，大型医疗机构在远程医疗网络的基础上也开始向院外的预约挂号、健康管理、健康咨询、院外康复等领域扩展。

未来，互联网企业与医疗机构将在技术、模式、资源等各个层面进行多种形式的合作，不断推动技术深入与模式再造，重塑医疗健康服务业态。在融合过程中，技术创新、标准完善、产业融合、安全保障、政策牵引是影响其发展的五大关键要素。

（二）“端、云”协同创新构建互联网健康技术基础

互联网健康服务发展有待于新型智能健康终端、云计算、大数据及人工智能等领域的核心技术突破。

1.智能健康终端技术是精准量化人体的关键

智能健康终端包括专业医疗检测设备、家用/社区医疗健康终端、可穿戴健康设备等，通过与应用软件、云服务相结合实现对各种人体体征数据的采集、传输等功能。智能健康终端对人体指标的精准量化是健康大数据分析的基础，其亟待突破高性能高可靠生物体征感知技术、低功耗轻量级底层软硬件技术、低功耗广域智能物联技术等。

2. 信息系统架构向云计算迁移，服务质量与信任问题备受关注

云计算已经成为互联网健康服务发展的重要基础设施，改变了传统医疗信息化发展模式，具有按需使用资源、建设运营成本低、可扩展性强等优势。当前，医疗机构、行业主管部门等均积极推动信息系统向云计算架构迁移，以提高工作效率和管理效率，最终惠及广大患者。但医疗健康云计算在其发展过程中也存在服务质量参差不齐、数据安全信任体系尚未建立等问题，需要相关标准和监管措施推进行业健康发展。

3. 大数据与人工智能技术促进医疗健康服务潜力充分释放

医疗健康数据的采集、处理、分析以及可视化技术将在医疗机构管理、医疗教学科研、医疗服务质量提升、患者安全保障、新药研发、重大疾病防控等方面发挥重大作用，充分挖掘医疗健康服务能力。同时，IBM 的沃森分析工具、百度医疗大脑等人工智能技术将在治疗复杂疾病等领域广泛应用。

三、我国互联网健康服务发展的策略建议

当前，“互联网+”正在推动我国医疗健康行业迈向更加广阔的发展空间，我国应抓住机遇从技术产品研发、标准机制创新、应用示范、政策牵引等层面全面推进互联网医疗健康服务的发展。

1.促进医疗卫生行业和信息通信产业协同创新和技术攻关

切实对接医疗卫生行业需求，加强跨行业的技术交流与产业合作，引导医疗机构和信息通信企业凝聚共识、发挥合力，共同推进对互联网医疗健康共性关键技术和发展模式的研究，进一步优化产业创新资源。与国家有关部门项目加强沟通，加速医疗健康可穿戴设备、医疗云计算、健康医疗大数据、人工智能、区块链等关键共性技术领域的重点布局，带动并促进全产业链的互动发展，促进整个医疗健康行业的转型升级。

2.依托联盟推进标准制定和评测验证

依托联盟、协会建设适应行业发展的团体标准工作机制，推进互联网医疗健康终端产品、云服务、大数据等标准制定，并形成与行业标准、国家标准的衔接机制。构建互联网医疗健康评测体系，推动医疗健康终端产品、云服务及大数据的评测验证，促进互联网健康服务标准完善与产品成熟。2016 年 6 月，中国信息通信研究院与国家卫生计生委联合组建“互联网医疗健康产业联盟”，首批吸纳医院、科研机构和企业成员 80 余家，正在推动互联网健康服务相关联盟标准制定。

3.基于大数据平台和共享机制开展应用示范

以打造开放共享的健康医疗大数据平台为着力点，整合多部门、多行业、多企业的信息资源，提高共享能力，促进互联互通，有效提高公共服务水平。推动大数据平台关键技术的研究，包括大数据关键技术研究、数据开放机制研究、中间件技术及开放工具研究、健康数据可视化研究等。建立完善的数据共享机制，加强数据保护和信息安全，促进资源的安全有序开放。

4.协调相关主管部门突破政策瓶颈

协调国家相关主管部门, 共同建立完善的互联网健康服务监管体系。加强互联网医院、健康服务、医药电商等新兴业态管理, 包括准入资质、服务质量、收费标准、责任划分等制度制定。促进互联网健康服务与国家基本医疗保险、商业保险的对接, 促进闭环业务模式的形成。推进电子处方流转, 促进处方药网上销售政策逐步有序开放。加快公共卫生、医疗服务、医疗保障、药品管理等信息系统的互联互通和数据的开放共享。

当前, 大力推进基于互联网的医疗服务、健康管理、养老、社会保障、药品供应等创新服务模式, 已成为推进我国医药卫生事业改革的重要抓手, 将有效促进医疗健康服务的创新供给和信息资源的开放共享, 大幅提升医疗健康服务能力和普惠水平。我国信息通信产业与医药卫生行业正加速融合, 重塑医疗健康服务业态, 在融合过程中技术创新、标准完善、产业融合、安全保障、政策牵引是影响其发展的五大关键要素。我国应抓住机遇从技术产品研发、标准机制创新、应用示范、政策牵引等层面全面推进互联网健康服务的发展。

PU-024

健康管理特征视角下移动健康用户类型及其对移动健康满意度的影响研究

王 桐[1]，雷健波[2]，王 伟[1]
1.吉林大学公共卫生学院 医学信息学系
2.北京大学医学部 医学技术研究院

1 引言

移动健康（mHealth）产业拥有一个超过 500 亿美元的全球市场[1]，其作为健康管理的主要方式正在蓬勃发展，产品数量庞大，用户众多[2]；然而移动健康正面临着持续使用不佳的挑战。用户满意度是持续使用重要的前驱变量，得到了移动健康领域学者的关注[3]。当前研究多为移动健康 app 的整体满意度评估，少数研究评估了移动健康 app 的某方面特征的具体满意度[4–7]。整体满意度评估很难揭示移动健康不同功能的具体实现效果及可用性，但当前缺少系统全面地以移动健康不同功能及设计为研究对象的满意度差异分析。移动健康 app 的用户满意度受多方面因素影响，从使用主体及客体考虑可将影响因素分为移动健康功能或服务特征、使用效果和用户特征三大方面[8–10]，围绕改善系统性能和服务质量的用户满意度研究相对较多，缺乏从健康管理角度系统分析相关特征对满意度的影响，为弥补上述局限，本研究将从健康管理行为相关因素入手，确定移动健康 app 用户健康管理特征，基于多个特征不同水平的组合情况划分用户类别，探究各个健康管理特征对移动健康 app 功能设计满意度的影响，分析不同类别用户对移动健康 app 功能设计层面满意度的差异。

2 理论基础

以患者为中心的健康管理模式中，认知与行为是健康管理核心的两个组成部分，评估这两方面的表现可以揭示用户在使用移动健康过程中的健康管理特征。本研究基于健康信念模型与行为改变轮探究用户对疾病和健康的信念认知及使用移动健康进行健康管理行为的驱动因素，以此为基础构建用户健康管理特征框架（见图 1）。

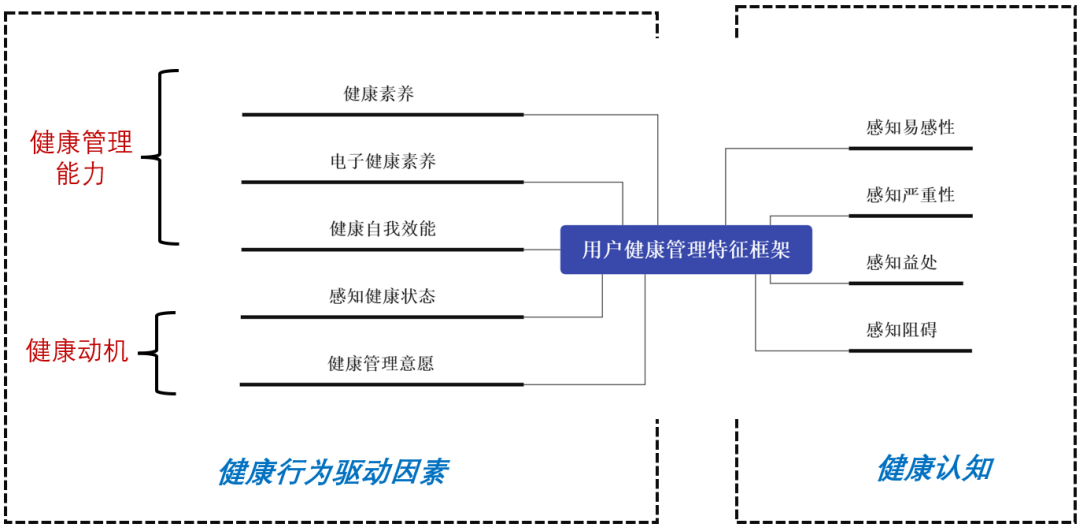


图 1. 用户健康管理特征框架

2.1 健康信念模型

健康信念模型（HBM）是使用最广泛的健康教育模型，从认知层面解释健康行为产生的原因，也可以作为健康行为干预的指导框架[11]。HBM 认为人们采取某种健康行为最直接的影响因素是意识、态度和信念[12]。具体而言，意识是指人们认识到容易产生严重健康问题及该健康问题会引起疾病并造成严重后遗症的程度。态度是指人们相信遵守具体健康建议有助于降低健康风险和疾病威胁。信念是进行健康管理的稳定的心理倾向和健康行为意图[13]。HBM 具有 6 个核心结构：感知易感性、感知严重性、感知益处、感知阻碍、健康行为的自我效能和行为指示（cues to action）[12]。感知易感性、感知严重性、感知益处和感知障碍这四个概念揭示了个人对疾病或健康的意识和看法，体现了用户在健康管理过程中认知层面特征，因此本研究将这四个概念归纳为用户健康管理特征框架中的健康认知维度。

2.2 行为改变轮

行为改变轮是解释行为改变的理论模型，被广泛用于个体健康行为促进和疾病预防等方面[14]。理论中解释了行为的三个来源，认为当个体具备能力、动机和机会时，才能实现某种行为的改变[15]。个人在进行健康管理时，健康管理能力与健康管理动机是来源于用户自身的行为来源，因此本研究以这两个概念为基础构建用户健康管理特征框架中的行为驱动维度。首先，用户在使用移动健康进行健康管理过程中，健康素养是指个体获取、理解和使用信息以维护和促进自身健康的能力[16]，电子健康素养是个体评估来自电子资源的健康信息并用于解决健康问题的能力[17]，二者反映了个体进行健康管理所具备的知识。健康管理自我效能是指个体对自身能否利用所拥有的技能去完成健康管理任务的自信程度[18]，反映了个体具备的健康管理技能，因此本研究将健康素养、电子健康素养和健康管理自我效能归纳为行为驱动维度中能力方面的子特征。其次，个体对其健康状况的认知水平有助于判别是否具有健康管理需求，进而激发健康管理意愿，可视为反省性动机。个体变得更健康的意愿和准备是其采取健康管理行为的直接激励信念，可视为自发性动机。因此本研究将感知健康状况和健康管理意愿归纳为行为驱动维度中动机方面的子特征。

3 材料与方法

3.1 确定研究对象

本研究旨在从功能层面探究各用户类别对移动健康 app 不同设计的满意度和偏好的差异，因此需要了解移动健康 app 具备哪些主要功能设计。本研究以既往研究为基础，对各个文献中给出的功能特征进行重新梳理，提出本研究要进行分析的移动健康 app 功能设计分类（见表 1）。移动健康 app 中一半以上涉及体重、饮食和运动[19]，因此本研究以运动减肥类 app 为代表进行分析。

表 1. 移动健康 apps 的功能分类和特征

功能主题	定义	特征
健康指导	根据健康计划和状态提供实时反馈和建议[20]，或提供有助于改善健康状态的行为改变策略[21]	智能虚拟助手、与保健人员沟通[22]；与医护人员沟通[20]；健康计划制定、目标设定[21]；

健康教育	提供必要的知识来实现健康行为改变[23]。	个性化教育、常识教育[24]；图片教育、视频教育、音频教育[23]
健康监测	记录过去和当前的健康状态[23]。	健康记录、健康日记[22]；跟踪器、数据录入、数据导出[24]；自我跟踪[21]；
社交功能	提供平台实现用户间的交流与互动并保持联系[23]。	社区论坛、社交媒体[24]；社交分享[23]
游戏化	将游戏机制或元素整合到移动健康 app 中，赋予用户游戏体验[25]。	个性化头像、挑战、任务、健康奖励[22]；积分、徽章、等级[24]；

3.2 测量工具及数据收集

本研究通过问卷调查收集数据，为了合理设计测量量表，本研究根据研究目的对既往经过验证的量表稍作调整。测量采用 5 点李克特量表，从 1（非常不同意）到 5（非常同意），量表由 24 位测试者进行预调查，根据评价和建议进行修改，形成最终问卷。

调查对象是近三个月使用过运动减肥类 app 的用户。于 2022 年 6 月 28 日至 9 月 8 日，通过微信进行滚雪球抽样调查。共收回问卷 2518 份，其中 902 个参与者 3 个月没有使用体重管理类 app 的经历，591 份由于中止作答、答题时间过短或没有通过质量控制题被判定无效作答，因此纳入分析 1025 份问卷。

3.3 数据分析

本研究利用 k-means 聚类确定用户类型，聚类依据是健康管理的 9 个特征（见图 1）。根据簇内误差平方和（SSE）与聚类数的关系确定最优聚类数[26]，SSE 低越低聚类效果越好，根据 SSE 与聚类数目曲线，发现当聚类数目为 6 时曲线出现拐点，因此聚类数确定为 6 个。通过偏最小二乘法分析健康管理特征对用户满意度的影响效应，秩和检验探究不同类型用户对移动健康 app 功能满意度和偏好的差异，检验的显著水平为双侧检验 $P<0.05$。

4 结果

4.1 用户健康管理特征的聚类分析及用户类型确定

根据 9 个健康管理指标的 Z-score 标准化结果对样本进行聚类分析，6 个类别的样本分布较为均匀（226，184，120，154，176，165），图 2 揭示了聚类结果中 6 个类别各个指标的分布情况。Z 分数接近于 0 表示该集群中对应的指标趋近于平均水平，Z 分数为正数代表该集群中对应的指标高于平均水平。

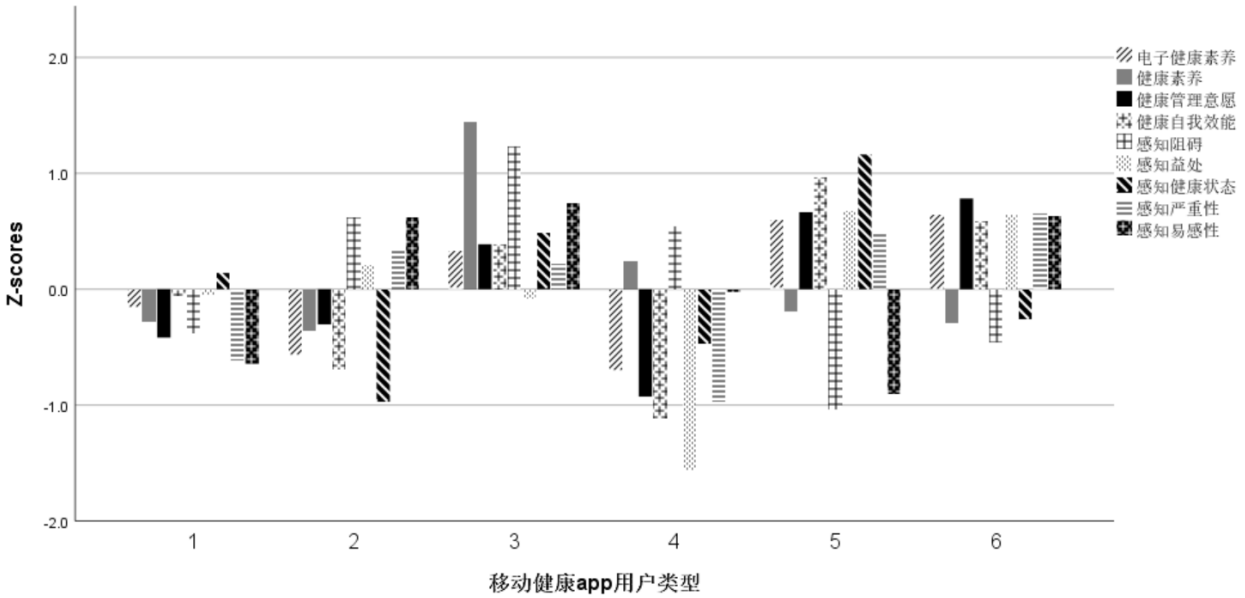


图 2. 6 个集群的特征分布

通过 Kruskal-Wallis 检验进一步分析 6 个集群中各个指标差异，确定各个集群的独特性。每个集群中各个指标得分的 Z-scores 存在统计学上差异，6 个集群各指标的 Z-scores 及差异情况见表 3。

表 3. 6 个类别中各个变量的 Z-Scores.

聚类变量		类别 1 (n=226)	类别 2 (n=184)	类别 3 (n=120)	类别 4 (n=154)	类别 5 (n=176)	类别 6 (n=165)
	感知严重性	-0.614$_{a}$	0.358$_{b,e}$	0.215$_{b,c}$	-0.966$_{d}$	0.501$_{e,f}$	0.653$_{f}$
	感知易感性	-0.643$_{a,e}$	0.625$_{b,c,f}$	0.744$_{c,f}$	-0.023$_{d}$	0.909$_{e}$	0.633$_{f}$
	感知阻碍	-0.378$_{a,f}$	0.616$_{b,d}$	1.229$_{c}$	0.547$_{d}$	-1.042$_{e}$	-0.462$_{f}$
	感知益处	-0.057$_{a,c}$	0.206$_{b}$	-0.083$_{c}$	-1.561$_{d}$	0.679$_{e,f}$	0.641$_{f}$
能	健康自我效	-0.055$_{a}$	-0.694$_{b}$	0.382$_{c,f}$	-1.117$_{d}$	0.965$_{e}$	0.585$_{f}$
态	感知健康状	0.140$_{a}$	-0.974$_{b}$	0.487$_{c}$	-0.473$_{d,f}$	1.163$_{e}$	-0.259$_{f}$
养	电子健康素	-0.161$_{a}$	-0.573$_{b,d}$	0.331$_{c,e,f}$	-0.705$_{d}$	0.598$_{e,f}$	0.639$_{f}$
	健康素养	-0.278$_{a,b,d,f}$	0.362$_{b,e,f}$	1.448$_{c}$	0.242$_{d}$	-0.192$_{e,f}$	-0.289$_{f}$
	健康管理意	-0.417$_{a,b}$	-0.303$_{b}$	0.388$_{c,e}$	-0.926$_{d}$	0.664$_{e,f}$	0.783$_{f}$
愿							

注: 下标为 “a”的组在统计学上与下标为“b”、“c”、“d”、“e”、“f”的组不同。然而，下标为“b”的组与下标为“b，c”的组在统计上没有区别，因为它们共享下标“b”。不同的分组标签表明集群之间的异质性。

由于各个集群具有独特性，我们根据每个集群中各个指标的组合情况命名了 6 个集群，用以揭示各个集群中用户健康管理特征的主要特点。集群 1 命名为**健康管理低需求者**，该类集群中的用户健康状态趋于平均，感知易感性、感知严重性和健康管理意愿都低于平均水平。集群 2 命名为**健康管理潜在需求者**，该类集群中的用户健康状态远低于平均水平，且感知益处、感知易感性和感知严重性都高于平均水平，但其健康管理意愿较低，且感知障碍较高。集群 3 命名为**积极意向者**，该类集群中的用户健康状态、健康素养和健康管理意愿都高于平均水平，但感知障碍和感知易感性也较高。集群 4 命名为**消极态度者**，该类集群中的用户健康状态低于平均水平，感知益处、健康管理自我效能、感知严重性和健康管理意愿均很低。集群 5 命名为**正向积极高能者**，该类集群中的用户健康状态、健康管理自我效能和健康管理意愿较高，且与积极意向者相比，感知障碍和易感性很低。集群 6 命名为**负向积极高能者**，与正向积极高能者相比，该类集群中用户健康状态低于平均水平，且感知易感性较高，但具有较高的健康管理意愿、健康管理自我效能和电子健康素养。

4.2 用户健康管理特征对移动健康 app 功能设计满意度的影响分析

以 9 个健康管理特征为自变量，5 个特征设计满意度为因变量，通过 PLS 分析健康管理特征对移动健康 app 特征设计满意度的影响。R^2 是指可以用自变量解释因变量的方法，用于解释回归方程的拟合度，β表示变量间的相关性。表 4 显示了自变量与因变量间的 β、和 R2 统计显著性。

表 4. 移动健康 app 不同功能满意度的偏最小二乘法分析结果 (β 值).

自变量	健康指导满意度	健康教育满意度	健康监测满意度	社交功能满意度	游戏化满意度
感知严重性	0.062	0.053	0.123**	0.040	-0.004
感知易感性	0.062	0.039	0.041	0.074	0.096*
感知阻碍	0.150**	0.116*	0.047	0.110*	0.049
感知益处	0.075*	0.020	0.100**	-0.179***	-0.150***
健康自我效	0.105*	0.111*	0.044	0.152**	0.097*

能						
	感知健康状态	0.101*	0.115**	0.077	0.114**	0.106*
	电子健康素养	0.116**	0.102**	0.148***	0.084*	0.090*
	健康素养	0.012	0.009	0.027	0.097*	0.081*
	健康管理意愿	0.147**	0.175***	0.079	0.132**	0.090*
	R^2	0.167	0.153	0.137	0.079	0.138

Note: * p < .05, ** p < .01, *** p < .001.

4.3 不同类型健康用户的移动健康 app 功能设计满意度差异

首先，通过 Kruskal-Wallis 检验分析所有用户对移动健康 app 的 5 个功能设计满意度的差异，结果表明所有样本中，用户对不同功能设计的满意度存在显著差异，H(K)=445.388，*P*<0.001，具体而言，所有用户对健康监测的满意度最高（M=4.00，IQR=1.00），其次为健康指导（M=4.00，IQR =1.00）和健康教育（M=3.67，IQR =1.00）的满意度，二者没有显著差异（*P*=0.885），随后为社交功能的满意度（M=3.33，IQR =1.00），用户对游戏化的满意度最低（M=3.00，IQR =1.00）。其次，进一步以健康管理用户的 6 个类别为分组依据，移动健康 app 的 5 个功能设计满意度为因变量，利用 Kruskal-Wallis 检验分析不同用户类型对同一功能设计满意度的影响差异。6 类健康管理用户对移动健康 5 个功能设计的满意度分布的中位数（M）和四分位距（IQR）以及不同用户对同一功能设计满意度的差异性结果见表 5。

表 5. 6 种用户对移动健康 app 不同功能涉及满意度的中位数和四分位距.

变量	**中位数 (四分位距)**					
	类别 1 (n=226)	**类别 2 (n=184)**	**类别 3 (n=120)**	**类别 4 (n=154)**	**类别 5 (n=176)**	**类别 6 (n=165)**
健康指导满意度	3.67_{a} (1.00)	4.00_{a} (1.00)	4.00_{b} (0.67)	3.33_{a} (1.00)	4.00_{b} (1.67)	4.00_{b} (1.00)
健康教育满意度	3.67_{a} (1.00)	3.67_{a} (1.00)	4.00_{b} (1.25)	3.33_{a} (1.00)	4.00_{b} (1.58)	4.00_{b} (1.33)
健康监测满意度	$3.67_{a,b,d}$ (0.67)	$4.00_{b,c}$ (0.67)	$4.00_{c,e}$ (0.67)	3.67_{d} (1.00)	4.00_{e} (1.33)	4.00_{e} (1.33)
社交功能满意度	$3.00_{a,d}$ (0.67)	$3.00_{a,d}$ (0.67)	4.00_{b} (1.25)	$3.33_{a,c,d}$ (0.67)	3.33_{c} (1.00)	$3.33_{c,d}$ (1.00)
游戏化满意度	3.00_{a} (0.67)	3.00_{a} (1.25)	4.00_{b} (1.33)	3.00_{a} (0.33)	3.00_{a} (1.00)	3.00_{a} (1.33)

注: 下标为“a”的组与下标为“b”、“c”、“d”、“e”组的差异具有统计学意义。然而，下标为“b”的组与下标为“b，c”的组在统计没有显著差异，因为它们共享下标“b”。

5 讨论

5.1 移动健康 app 的用户类型及用户画像

移动健康用户可分为 6 类。健康管理低需求者占比最大（22.05%），该类用户代表着当前健康状态良好，对健康管理没有给予过多重视，对健康风险对健康损害的意识不足，因此没有较高的健康管理意愿的群体。健康管理潜在需求者（17.95%）代表着当前健康状态不佳，但其意识到健康风险对健康的危害，同时也认可健康管理对状态改善的益处，然而实施健康管理的阻碍较高，进行健康管理的意愿也较低的群体。积极意向者（11.71%）是占比最少的用户，该类用户代表着健康状态、健康素养和健康管理意愿都高于平均水平的群体，该类群体具有健康管理的主观能动性，认为自己可能存在健康状态健康的风险，在状态较好的基础上积极进行健康管理，但也存在一定感知障碍。消极态度者（15.02%）与健康管理低需求者有相似之处，感知益处、健康管理自我效能、感知严重性和健康管理意愿均很低，但不同之处在于消极态度者的感知健康状态低于平均水平，且其感知阻碍较高。这类群体已经感受到健康状态的降低但仍没有意愿改善，有很高的患病风

险。正向积极高能者（17.17%）和负向积极高能者（16.10%）均代表了健康管理认知水平、素养和能力较高的群体，并有较高的健康管理意愿，是当前移动健康 app 可能受益的主要群体，二者的区别在于正向积极高能者的感知健康状态较高，而负向积极高能者的感知健康状态低于平均水平。

5.2 健康管理特征对移动健康 app 功能设计满意度的影响

首先，既往研究多从移动健康整体角度探究用户健康管理特征与满意度间的相关性[27]，但本研究表明，感知严重性对移动健康的健康监测的满意度具有显著正向影响，感知易感性对移动健康的游戏化设计的满意度具有显著正向影响，说明健康管理特征对移动健康满意度的影响应从功能设计角度考虑，总体满意度难以为个性化服务提供更细粒度的建议。例如感知易感性仅对移动健康 app 中的游戏化满意度具有显著影响，那么对于远程医疗这类基本不存在游戏化设计的移动健康工具，服务提供者应改善用户感知易感性以外的因素以提高移动健康 app 的满意度。其次，结果表明移动健康 app 中 5 个主要的功能设计满意度的影响因素并不相同，应从主要功能设计角度完善移动健康 app。如移动健康 app 的健康指导和健康教育的满意度受健康管理意愿、感知健康状态和电子健康素养的显著正向影响，既往研究表明较高的电子健康素养与感知健康状态相关[28]，因此以健康指导和健康教育为主要功能设计的移动健康 app 应努力提高用户的电子健康素养，为其提供利用电子资源查找健康相关信息与知识的简要教程，使其能有效利用健康指导与健康教育功能，提高对其满意度。

5.3 用户类型对移动健康 app 功能设计满意度的影响

首先，所有用户样本的移动健康 app 功能设计满意度中，健康监测的满意度最高（M=4.00，IQR=1.00），社交功能（M=3.33，IQR =1.00）和游戏化（M=3.00，IQR =1.00）的满意度最低，说明当前移动健康对用户健康信息的收集和健康状态的监控和管理较好，但社交功能和游戏化设计还有待完善和改进。其次，对每个健康管理特征进行 PLS 分析有助于理解用户的健康认知和行为能力对移动健康功能设计满意度的作用，但用户具有多个特征，单一特征的影响作用难以从个体层面解释用户对功能设计综合评价的结果。例如，健康管理意愿对社交功能满意度有显著积极影响（β=0.132，*P* <0.01），积极意向者、正向积极高能者和负向积极高能者的健康管理意愿均高于平均水平，但正向积极高能者与负向积极高能者对社交功能的满意度显著低于积极意向者，并且与其他三类用户对社交功能的满意度的差异不显著，出现该现象的原因可能是正向积极高能者与负向积极高能者的感知障碍低于积极意向者。由此说明分析用户对移动健康 app 功能设计满意度时，不应仅考虑单一用户健康特征的影响，应考虑用户多个健康特征不同水平的组合情况，以用户个体层面为研究对象，探究用户类型对功能设计满意度的影响，以提供针对不同类型用户的个性化移动健康服务。

PU-025

“互联网+医疗健康”背景下我国互联网医院管理模式及应用现状分析

刘尊尊
郑州大学

“互联网+医疗健康”是国家战略实施的一部分，党和国家领导人也十分重视互联网医院的建设，多次提出要大力发展互联网医院建设，为人民众提供更加便利的医疗服务。2016 年 10 月，中共中央国务院出台“健康中国 2030”规划纲要，高度体现了与联网技术对于医疗卫生事业未来发展的重要性。2018 年 4 月，国务院办公厅印发《关于促进“与联网+医疗健康”发展的意见》[1]，该《意见》的出台具有里程碑性重要意义，明确了“写联网+”与医疗卫生业务融合发展的重点领域和支撑体系，划出了“与联网+医疗健康”的监管和安全底线，体现了医疗卫生机构注重医疗安全与质量的行业特殊性。2018 年 7 月，国家卫生健康委、国家中医药管理局《互联网医院管理办法(试行)》明确了互联网医院发展模式、开展形式、服务范围[2]。

互联网模式下的医疗健康管理发展推动了传统医疗行业模式转型，为医疗卫生健康行业提供巨大的发掘潜力和探索空间，同时也是发展的必然趋势。

1 互联网医院的定义和发展现状

1.1 互联网医院的定义

互联网医院即通过互联网远程为患者提供导诊、预约挂号、常见病和慢性病的诊疗，并开具处方和配送药物的医疗服务平台[3]。

互联网医院通过互联网线上平台，将医院的各种功能通过结合人工智能、大数据、物联网等技术手段应用到线上，让患者在家就能享受到医院的医疗、护理、药事、心理健康、康复等诊疗服务。

1.2 我国互联网医院发展现状

目前，大多数的互联网医院还只能提供预约挂号、线上缴费、报告查询、图文咨询、视频问诊等基础功能。只有少数互联网医院可以实现在线问诊、处方流转、医生辅助和医院流程优化。仅极少数自建的互联网医院能够真正做到实现结合线上、线下配置更多优质资源，快速、便捷服务基层和贫困地区的群众，做到线上预约问诊、线下检查治疗，院后健康管理智能随访，拥有远程医疗和智能穿戴设备，实现诊前、诊中、诊后全流程覆盖，其中黑龙江省医院互联网医院做到了全国前列[3]。

2 我国互联网医院运营管理模式及机遇和挑战

2.1 互联网医院的的运营模式

我国互联网医院运营模式分为三种：实体医院的医疗资源线上服务模式（H+I 模式）、医联体共同线上融合服务模式（H 与 I 融合模式）、集聚医生资源的平台服务模式（I+H 模式）。“H+I 模式”的服务内容是面向公众的互联网服务。主要形式是线上预诊、线下确诊、线下治疗、线上复诊。包括、慢病在线延方、药师在线审方、药品物流配送、居家用药指导、家庭延续护理、院后在线随访等。武汉市中心医院就是“H+I 模式”的典型代表。“H 与 I 融合模式”通过多家第三方互联网服务平台进行线上问诊，线下服务依靠实体医院硬件平台和医疗服务。居民通过互联网平台实现本地挂号就医，也可以与其他地区优秀医生进行医疗沟通，在网上进行专家会诊，并为患者提供转诊服务。该模式的典型代表是银川互联网医院。“I+H”模式则由互联网企业发起，以多点执业集聚各地医生资源，医生在通过互联网企业医疗平台进行注册，并为全国各地的患者提供诊疗、检查复诊、手术预约、药品配送等服务。该模式典型为乌镇互联网医院、微医等。

2.2 互联网医院的机遇分析

2.2.1 国家政策支持

2018 年 4 月发布的《国务院办公厅关于促进“互联网+医疗健康”发展的意见》指出，允许依托医疗机构发展互联网医院，常见病和慢性病可通过互联网进行复诊的服务，医生掌握病历资料后可在线开具处方；2018 年 7 月，国家卫生健康委员会和国家中医药管理局组织印发《互联网医院管理办法（试行）》等 3 个文件，对互联网诊疗行为作出进一步规范；2019 年 8 月，国家医疗保障局下发《国家医疗保障局关于完善“互联网+”医疗服务价格和医保支付政策的指导意见》，对“互联网+”医疗服务的项目价格管理及医保支付提出具体要求；2020 年 2 月，国家医疗保障局办公室下发《国家医疗保障局办公室关于全面推广应用医保电子凭证的通知》，提出全面推广应用医保电子凭证，助力提升医保治理现代化、科学化、精细化能力和水平，促进医疗健康产业发展。

随着各部门政策文件的不断出台更新，体现出国家对互联网医疗支持与信心。这些政策的出台也大力助推了“互联网+”医疗健康行业迅速发展。

2.2.2 市场需求扩大

我国人口老龄化问题持续加剧，国家统计局公布的数据表明，2022 年末我国 60 岁及以上人口有 28 004 万人，占全国人口的 19.8%，其中 65 岁及以上人口 20 978 万人，占全国人口的 14.9%[4]。庞大的老年人群为医疗和健康需求和互联网医疗的开展提供了良好的用户基础和广阔市场[5]。

“健康中国”下，国民对医疗健康的的意识日趋增强，医疗健康需求指数也成倍增长。前瞻产业研究院发布的《2015—2020 年中国医疗信息化行业市场前瞻与投资战略规划分析报告》数据显示，我国医疗信息化市场 2020 年有望突破 1 000 亿元；CCW Research（计世资讯）的研究显示，我国远程医疗市场 2019 年达到了 27.5 亿元，预计 2020 年市场规模将达 172 亿元[3]。

2.2.3 技术革新进步

新形势下，大数据、物联网、人工智能等信息技术的兴起，带动“智慧医院”的建设升级，为

互联网医院的发展提供了强有力的技术支撑。5G 技术的升级换代促进了远程医疗和移动医疗的更广泛运用，人工智能、物联网的发展推动了个人健康数据的全面采集和充分利用，云计算、大数据等促进了医保管理的精细化和监管的智能化[6]。这些新技术的快速发展与应用为互联网医院的数字化改革攻坚破除了技术壁垒，提供了技术保障。

2.3 互联网医院的挑战

2.3.1 相关政策法律仍需健全

当前，互联网医疗领域存在国家法律监管滞后、行业监管相关的专门性立法少、法律位阶不高等问题。如相关领域的许可准入、对医师执业的限制、互联网医疗的法律关系和责任问题以及诊疗记录、病例等医疗信息的管理均未规范界限，这些均是互联网医疗行业常见的问题，与该行业蓬勃发展的现状不相匹配[5, 7]。因此，为维护了消费者权利、顺应和支持我国互联网医疗行业发展，相关政府部门应该加大行业监管力度，完善监管机制，建立行业标准。

2.3.2 互联网诊疗服务定价和医保报销标准不统一

远程医疗服务定价和医保报销缺乏统一标准。当前，互联网医疗存在线上线下同一医疗服务的价格不一、各地互联网医疗服务价格与收费标准不统一、互联网医疗患者在线支付费用医保报销问题及互联网医疗保险种类单一等现象[8-10]。因此，互联网医疗服务的价格制定与医保支付方面还需制定具体行动方案，提供保障措施，以此推动互联网医疗可持续发展。

3 互联网医院未来的发展趋势及建议

3.1 互联网医院发展趋势

互联网医院和可穿戴设备都是我国医疗卫生体系中的新元素，两者联合应用对促进医疗服务从在院的诊断和治疗环节向入院前的疾病预防、健康管理和出院后的康复监测等环节拓展和延伸的过程中发挥重要作用。

一方面，可穿戴医疗健康设备扮演着一个健康数据采集者的角色，将正常医疗流程无法获取的数据转换为实时数据流，便于患者(用户)进行自我健康管理[11]， 及时更正不良的生活习惯，筛查、预防疾病;也可以借助云计算、大数据技术，通过数据挖掘寻找日常生活行为与身体各项体征数据、疾病的关联，并可接受专业医护人员的各种健康咨询、筛查、预防、诊断、治疗、监护和干预服务。

另一方面，构建融合可穿戴设备监测结果反馈功能和互联网医院线上问诊、处方结算以及药品配送功能为一体的集成性网络平台，实现结果查询、线上问诊和费用支付等环节的一站式操作[12]，使患者能够获得优质便捷的医疗服务，一定程度上缓解我国看病贵、看病难的问题，推动我国医疗卫生事业高质量发展。

3.2 互联网医院发展相关建议

3.2.1 针对互联网医疗发展完善政策法律

互联网医院是医疗领域中出现的“互联网+医疗健康”新发展模式，具有非常大的市场空间，也是未来发展的趋势。国家通过出台各项政策来鼓励、支持互联网医疗的发展，但目前我国在互联网医疗领域暂时还没有制定专门性的法律监管进行市场监管，现有的法律制度对互联网医疗的监管规制相对滞后。互联网医疗本身具有特殊性，在互联网+健康服务中参与主体、行为内容和责任承担都与传统医疗有着巨大的差别，因此需要完善互联网医疗监管领域的相关立法，及时跟上互联网医疗行业发展的步伐，让互联网医疗在法律的框架下有序健康发展。

3.2.2 统一线上线下医疗服务和医保支付标准

互联网医疗相对于传统问诊、医疗方式打破了时间、地域的限制，节约了时间和经济成本，提高了医疗服务效率。但存在线上线下同一医疗服务的价格不一、各地互联网医疗服务价格与收费标准不统一、互联网医疗患者在线支付费用医保报销问题及互联网医疗保险种类单一等问题。国家医疗保障局表明：“国家层面负责明确‘互联网+’医疗服务立项原则、项目名称、服务内涵、计价单元及计价说明等的规范，指导各省做好医疗服务价格项目工作”。互联网医疗服务的价格制定与医保支付方面还需制定具体方案，来解决当前问题，促进互联网医疗可持续发展。

4 结论

面对“互联网+医疗健康”的机遇与挑战， 传统医院当以自身为基础， 把握机遇，借助互联网的优势实现管理升级、服务优化、 业务外延，不断拓展市场， 实现医院运营管理、社会效益、经济效益等多方面的快速发展。

PU-026

关于云胶片服务在医院中的应用与现状分析

杨恺熠[1,2,3]，赵永根[1,2,3]，齐国强[1,2,3]，朱志文[1,2,3]，俞刚[1,2,3*]
1.浙江大学医学院附属儿童医院数据信息部
2.浙江-芬兰儿童健康人工智能联合实验室
3.儿童健康医工信交叉研究中心（国家儿童健康与疾病临床医学研究中心）

引言

云胶片是一种存储在云平台服务器上的电子胶片，是移动互联网时代出现的一种“互联网+医疗”新技术[1]。它以影像数据的集成作为基础，用户可通过移动端的智能设备进行浏览和下载。云胶片服务包含了各种影像检查（如 DR、CT、MRI、US 等）的报告以及所有的原始图像（即 DICOM 格式的电子影像），它所展现的信息是完整的，可以在终端对图像进行直接地调窗或者测量，也能够满足其他分析应用的需求[2]。患者和家属只需通过扫报告上的二维码，或者通过 app 或微信小程序上绑定的就诊信息即可调阅云胶片和电子报告。

2020 年 12 月 31 日，国家卫健委印发《关于进一步规范医疗行为促进合理医疗检查的指导意见》，要进一步规范医疗行为，促进合理检查。在机制创新层面，要求医疗机构通过建立检查资料数据库或“云胶片”等形式，推动检查资料共享[3]。

在传统的就诊模式中，医用干式胶片作为影像检查后影像资料的载体，有着种种不便：如不便于携带与储存；CT 和磁共振之类的图像，在胶片上都是小图的形式，不方便阅片等。而电子胶片虽然可以存储在 U 盘等移动存储设备中，但其原始格式 DICOM 格式需要专业软件才能打开，在患者跨院就诊中也甚是不便[4]。因此，在移动互联网越来越发达的如今，急需一种新的介质来解决这些矛盾。而“云胶片”的出现很好地解决了这些问题。

云胶片服务作为“互联网+医疗”的新产品，极大地提高了患者获取自己影像资料的效率，也免去了患者在之后为了获取自己的电子胶片而在各个医院部门奔波的麻烦。因此云胶片服务是一种值得推广的“互联网+医疗”新技术[4]。为响应浙江省物价局、浙江省卫生和计划生育委员会《关于核定数字影像服务费等有关事项的通知》的号召，为切实减少患者的看病费用，规范医院的收费行为，削减医院的耗材采购成本，我院于 2019 年开始着手准备云胶片服务，并于 2019 年年底正式上线。

1. 云胶片的架构与操作流程

1.1 云胶片的架构

浙大儿院目前实行的云胶片服务的硬件由以下设备组成：影像设备、PACS 服务器、PACS 存储服务器、RIS 服务器、信息集成平台服务器、云胶片服务器、医生工作站、自助报告机、移动终端。

1.2 云胶片服务的流程

该院的云胶片服务具体操作流程见下图 1。①由影像设备采集到对应的影像，并生成电子胶片，再上传至 PACS 服务器。②PACS 服务器将电子胶片归档存储至 PACS 存储服务器中。③报告医生在医生工作站中通过调取 RIS 服务器的病人信息和 PACS 服务器中的电子胶片，进行阅片、诊断、书写报告，再将书写完的报告上传并存储至 RIS 服务器中。④云胶片服务器通过调取信息集成平台的数据，获得病人的电子胶片和影像报告。⑤患儿家长通过扫纸质报告上的二维码直接访问云胶片，也可以通过浙大儿院的 APP 或微信小程序，访问绑定的患儿就诊帐号，从而阅读该患儿的云胶片和影像报告[5,6]。

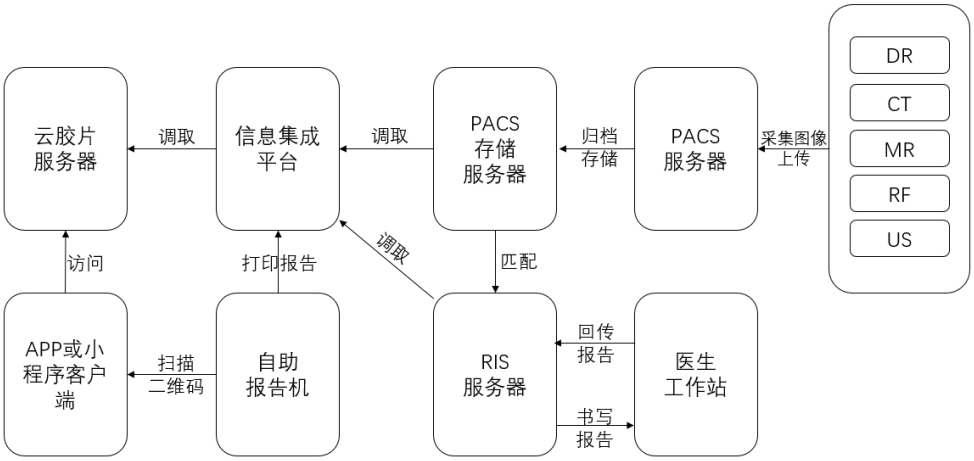


图 1 云胶片操作流程

2. 云胶片服务现状与优势

2.1 云胶片服务的现状

浙大儿院于 2019 年 12 月正式上线“云胶片”服务，由以传统胶片为主、“云胶片”为辅的初期探索模式，迅速发展，至 2021 年，已经基本全面落实云胶片，云胶片占总的胶片开单比例 85%以上，累计服务 31 万人次，详情可见表 1 表 2（其中 2023 年数据统计至 7 月）。其中，2020 年为“云胶片”发展最为迅速的一年。经过我院临床一线医护人员的大力推动、信息中心积极配合下，云胶片的开单比例逐月上升，从年初 19.77%的“云胶片”占比，历经 8 个月，占比到达 80%以上，彻底改变了人们的胶片获取习惯。

表 1 2019 年～2023 年传统胶片与云胶片开单量

年份	开单数（例）	无需胶片（例）	传统胶片（例）	云胶片（例）	传统与云（例）	云胶片比例
2019	260 170	26 027	231 989	723	1 431	0.28%
2020	416 353	87 309	12 316	298 344	18 383	71.66%
2021	538 869	51 018	6 978	477 883	2 990	88.68%
2022	563 509	86 049	6 125	469 745	1 590	83.36%
2023	422 983	49 979	5 097	367 138	769	86.80%

表 2 2020 年度各月传统胶片与云胶片开单量

月份	开单数（例）	无需胶片（例）	传统胶片（例）	云胶片（例）	传统与云（例）	云胶片比例
1 月	37 872	24 627	4 519	7 488	1 238	19.77 %
2 月	11 318	4 593	694	5 202	829	45.96 %
3 月	26 019	9 022	1 226	13 870	1 901	53.31 %
4	33	8 788	1 078	20	2 506	62.89

月	339			967		%	
5月	38 092	6 609	834	27 775	2 874	%	72.92
6月	34 018	6 363	871	24 644	2 140	%	72.44
7月	41 308	6 885	992	31 243	2 188	%	75.63
8月	41 437	4 474	625	34 724	1 614	%	83.80
9月	32 190	3 684	424	27 079	1 003	%	84.12
10月	38 284	3 327	410	33 474	1 073	%	87.44
11月	40 462	4 202	462	35 009	789	%	86.52
12月	42 013	4 735	181	36 869	228	%	87.76

2.2 云胶片服务的优势

自云胶片上线以来，不仅为患者的医疗费用带来了实质性的降低，同时也有效地减少了医院在耗材采购方面的开支。具体的优势如下：

（1）看病费用降低。

由国家卫计委于 2017 年在《关于全面推开公立医院综合改革工作的通知》中提出，“百元医疗收入（不含药品收入）中消耗的卫生材料降到 20 元以下”[6]。云胶片的推出成功实现了这一目标，为患者提供了实际的经济优惠。传统的医学胶片费用因不同检查项目和数量而异，一次检查的费用范围从 14 元到 84 元不等，一次开多个部位的项检查胶片价格甚至还能超过 100 元。同时，胶片还属于丙类项目，医保无法报销，需要患者自费处理。如下表 3 所示，以 2022 年该院开单最多的三类检查为例，DR 检查每个患儿平均需要 1.52 张胶片，平均胶片费用为 21.28 元；CT 检查每个患儿平均需要 3.48 张胶片，平均胶片费用为 48.72 元；MR 检查每个患儿平均需要 5.22 张胶片，平均胶片费用为 73.08 元。而云胶片服务，每个检查项目都只需要 20 元，极大缩减患者就诊费用。

表 3 2022 年云胶片与传统胶片患者支付费用对比

检查项目	平均打印胶片数（张）	平均胶片费用（元）	云胶片费用（元）	降费比例（%）
DR	1.52	21.28	20	6.02
CT	3.48	48.72	20	58.95
MR	5.22	73.08	20	72.63

（2）采购成本降低。

从医院层面来说，云胶片服务的推广，减少了放射科的耗材支出[7]。在云胶片实施前，浙大儿童医院的年胶片打印量约为 54 万张。在推广云胶片后，2022 年，浙大儿童医院的年开单数约为 53 万例，预期的胶片打印量为 120 万张（平均每个患儿需要的胶片数为 2.13 张）。但实际的胶片打印量约为 1.64 万张，仅为原先预期的 1.37%。云胶片服务的推广，极大地降低了放射科的耗材成本，每年可为医院节省传统胶片的采购成本约 1600 万元。

表 4 2022 年预期胶片采购费用与实际胶片采购费用对比

开单数	预期胶片数量	预期胶片采购费用	实际胶片数量	实际胶片采购费用

56 3 509	120 万（张）		1 680 万（元）	1.64 万（张）	22.96 万（元）

（3）落实最多跑一次。

在云胶片服务推广前，如果患者有外院转诊或进行伤残鉴定等情况，就必须获得原始 DICOM 格式的电子胶片。患儿家属需要前往医院的信息中心填完影像资料拷贝申请单，然后前往医务科盖章批准，最后返回信息中心拷贝电子胶片，这一流程相当繁琐。而有了云胶片服务后，患儿家属无需亲自前往医院拷贝电子胶片，他们可以直接在云胶片服务中向医生展示影像和报告，或者在云胶片服务中下载原始 DICOM 格式的电子胶片提交给相关机构。云胶片服务的推广为浙江省的“最多跑一次”改革提供了有力的支持，给看病的群众提供了极大的便利[8]。

3.云胶片目前存在的问题及建议

当前云胶片服务仍存在一些问题，如数据调取和服务器自身的不稳定，会给正在使用云胶片服务的患者产生一定的影响，还需要不断地改进云胶片服务的服务质量。

3.1 云平台服务器调取数据仍不稳定

由于电子胶片和报告都涉及到很多的服务器，服务器之间数据传输的不稳定会导致极个别的患者访问云胶片时，客户端上无法访问到云胶片数据[9]。这时患者如需反馈，则得找到放射科或者信息中心，才能继续处理问题，反馈的流程比较漫长繁琐。因此建议可以在云胶片服务上添加一个客服反馈窗口，在患者访问不了云胶片时，直接向客服反馈，由客服及时联系后续处理。

3.2 云平台服务器自身仍不稳定

由于云平台服务器自身也有不稳定的情况，会导致客户端所有的云胶片都无法访问。这时会对正在跨院就诊的病人造成困扰。因此建议在客户端提供本地缓存服务，可以让患者在有网络的情况下提前将需要的云胶片保存在客户端上。这样在服务器无法联络或者没有网络的情况下也能继续访问云胶片。

3.3 不同医院间云胶片无法共享

由于不同医院使用的 PACS 系统与云胶片服务的厂家不统一，各自产生的电子胶片的编码也不统一，所以患者无法在一个客户端内访问自己在不同医院的云胶片与报告[9]。所以需要有相关部门主导建立一个数据集中平台来统一收集区域内各家医院的电子胶片和报告[9]，方便患者在自己的客户端归档生成自己的历史档案，也能方便患者在医院就诊时更好地展示病程记录。

4.小结

随着以智能手机为代表的移动互联网的飞速发展，国家也在积极地推动“互联网+医疗”的发展，而云胶片服务也成为了“互联网+医疗”的典型应用,云胶片逐步取代传统胶片已是大势所趋[1,10]，代表着医院的影像管理找到了新的方向。云胶片服务的推出，不仅降低了患者的就诊费用，还减轻了医院耗材的采购成本，是对患者和院方都有利的一项新技术。云胶片服务在提供电子胶片的同时，也提供了电子报告，这样还能免去排队等候打印胶片和报告的时间，给就诊的患者提升就诊体验[11]。另外，它还为患者的跨院就诊提供了便捷，省去了在不同医院之间奔波的麻烦。

与此同时，云胶片服务中“云端”的数据安全也是值得关注的一个问题[12]。在云胶片服务的推广过程中，不能忽视患者影像数据的安全性。确保“云端”数据的安全、准确性以及数据访问时的网络畅通，将成为云胶片服务未来发展中需要特别关注的方向[9,13]。

PU-027

优化营商环境背景下互联网医院发展研究—基于 SWOT-PEST 模型分析

严 恺 于 琦
山西医科大学管理学院

1 背景

互联网医院是以互联网技术为基础的医疗服务模式，旨在提供更便捷、高效的医疗服务[1]。近年来，随着我国政府大力推动优化营商环境，互联网医院得以迅速崛起。然而，互联网医院仍然面

临着政策、市场和技术等多重挑战。本论文通过 SWOT-PEST 分析，旨在深入研究互联网医院在政府优化营商环境政策下的市场发展前景。

1.1 优化营商环境背景下互联网医院发展的意义

营商环境是指市场主体在准入、生产经营、退出等过程中涉及的政务环境、市场环境、法治环境、人文环境等有关外部因素和条件的总和[2]。营商环境的优劣水平决定了生产要素资源的集聚与流向。中国特色社会主义已进入新时代，经济已经由高速增长阶段转向高质量发展阶段。营造一流营商环境，是政府提供公共服务的重要内容，也是实现高质量发展的重要基础和关键一环。2020 年国务院办公厅印发《全国深化“放管服”改革优化营商环境电视电话会议重点任务分工方案》中明确提出：要优化互联网医疗发展环境，鼓励支持医疗机构对接互联网平台，开展互联网复诊、互联网健康咨询、远程辅助诊断等服务，研究制定互联网健康咨询流程和标准，探索根据病种分级分类实施互联网问诊管理，实现线上咨询、线下诊疗衔接[3]。

总的来说，通过优化营商环境，旨在降低制度性交易成本，对市场主体特别是中小微企业提供更宽松的准入门槛，推动互联网医院能够蓬勃发展，为更多患者提供便捷、高效、负担得起的医疗服务，有助于提高整个医疗卫生服务保障体系的效率和质量。同时，它也为经济增长和就业创造了机会，对于社会和经济的可持续发展具有重要意义。

1.2 互联网医院发展现状

“互联网医院”是指以实体医院为依托，集问诊、处方、支付及药物配送为一体的一站式服务平台[4]。截至 2022 年 10 月，全国已有 2700 余家互联网医院，开展互联网诊疗服务超过 2590 万人次。2022 年，互联网医疗用户规模达 3.6 亿人，增长率为 21.7%[5]。互联网医院逐渐盛行并被人们广泛关注。当前的互联网医院发展模式主要分为三种：

（1）线下实体医疗机构为主体，通过微信公众号、微信小程序、手机 APP 等方式，将既有线下医疗服务向线上扩展延伸，包括但不限于在线咨询、预约挂号、远程问诊、药品配送等;

（2）实体医疗机构提供医疗服务内容，医院与企业共建互联网医院，由共建企业负责线上服务运营;

（3）完全由第三方搭建互联网诊疗平台，患者和医护在第三方平台上实现服务全过程。

2 互联网医院发展 SWOT-PEST 模型分析

2.1 互联网医院的优势（Strengths）

2.1.1 便捷的医疗服务

从患者角度考虑，实体医院就医流程复杂，等待时间长，知名专家的预约和看诊更是难上加难。而互联网医院突破了传统就医模式的局限性，提供了线上预约、咨询，在线问诊、药品配送、线上缴费、健康管理等便捷的医疗服务，患者无论身处何时、何地，只要在有网络的环境下，可以随时随地享受医疗服务，大大节省了患者的时间和精力。

从医生角度考虑，医生可利用碎片化的时间提供在线诊疗服务，实现多点执业。一方面可积累诊断经验，另一方面也可合理增加收入，提高知名度和影响力[6]。同时，互联网医院执行的是准入制管理，有专门的监管部门对机构资质、医生资质和诊疗行为进行监管，能进一步规范和激励医生不断提高诊疗及咨询服务质量，为患者提供更为优质的服务，减少医患矛盾的发生。

2.1.2 合理的就医成本

良好的营商环境可以促使互联网医院提供更加经济实惠的医疗服务，为患者省去了的传统医疗带来的交通、食宿、误工等附加费用，节约了患者的时间成本和经济成本，同时提高了就医效率和就医体验。

2.1.3 优质的医疗资源

互联网医院扩宽了医院的服务半径，辐射和联动区域内的医疗资源，实现跨区域诊疗服务，方便快捷地把优质的医疗资源覆盖至医疗卫生资源分布不足的偏远地区，使优质的医疗资源实现下沉，让百姓在基层医疗机构也能接收到三甲医院医务人员带来的服务，推动分级诊疗的落地。

2.1.4 先进的技术支持

良好的营商环境鼓励创新，使互联网医院能够更容易地引入新技术和新模式，提供更多样化、高效率的医疗服务。包括远程医疗、人工智能、医疗大数据分析等都有助于提高医疗服务的质量和便捷性。

2.2 互联网医院的劣势（Weaknesses）

2.2.1 服务方式的改变带来的诊断问题

由于网络的虚拟性，患者无从知道为其在线问诊的医务人员是否有相关的执业证书，这会对医

疗服务质量造成一定的影响;而医生对于患者在线提供的资料，也无法保证其真实性。缺少传统医疗服务方式面对面进行的视触叩听等诊疗环节[7]，只通过患者所提供的信息进行医疗诊断，可能会出现误诊或漏诊。

2.2.2 业务投资带来的资金分配问题

互联网医院是一个新兴领域，需要在基础设施建设、技术研发和更新、医生培训和认证、市场推广和普及等多个方面的大量资金投入，而多数医疗机构仍然对互联网医院持保留态度，不愿意将过多的资金投入到这个领域中。

2.2.3 市场版图的争夺带来的竞争问题

互联网医疗市场竞争激烈，已经存在着众多的医疗服务提供者，新参与者需要应对激烈的竞争压力，包括价格竞争和市场份额争夺。建立规范有序的营商环境能够有效缓解此类风险，但在竞争的过程中或多或少都会存在服务提供者为了个人利益“打擦边球”的情况，随之带来的是患者在选择互联网医疗服务时的不确定性和不舒适的体验感，进一步会造成医患关系的恶化。

2.3 互联网医院的机会（Opportunities）

2.3.1 政治政策环境（Politics）

我国政府积极支持互联网医院发展，自2015年国务院发布《关于积极推进“互联网+”行动的指导意见》和《关于推进分级诊疗制度建设的指导意见》后，互联网医院如雨后春笋般逐渐在全国各地兴起。从2018年至今，中央发布了《互联网医院管理办法(试行) 》[8]以及《关于促进“互联网+医疗健康”发展的意见》等一系列文件，明确了积极的政策导向，规范、引领互联网医院的发展。

表1 互联网医院相关的国家政策

序号	发文单位	发文时间(年)	政策文件
1	国家卫健委 国家中医药管理局	2018	《互联网诊疗管理办法(试行) 》《互联网医院管理办法(试行) 》《远程医疗服务管理规范(试行) 》
2	国务院办公厅	2018	《关于促进“互联网+医疗健康”发展的意见》
3	国家医保局	2019	《关于完善“互联网+”医疗服务价格和医保支付政策的指导意见》
4	国家医保局 国家卫健委	2020	《关于推进新冠肺炎疫情防控期间开展“互联网+”医保服务指导意见》
5	国家卫健委	2020	《关于在疫情防控中做好互联网诊疗咨询服务工作的通知》
6	国家卫健委	2020	《关于进一步推动互联网医疗服务发展和规范管理的通知》
7	国家卫健委	2020	《关于做好公立医疗机构“互联网+医疗服务”项目技术规范及财务管理工作的通知》
8	国家卫健委	2021	《医院智慧管理分级评估标准体系（试行）》
9	国务院	2021	《“十四五”数字经济发展规划》
10	国家卫健委 国家中医药管理局	2022	《互联网诊疗监管细则(试行) 》
11	国务院办公厅	2022	《“十四五”国民健康规划》

2.3.2 经济环境（Economic）

（1）医疗卫生支出持续增加

《2021年我国卫生健康事业发展统计公报》[9]显示，2021年我国卫生总费用占GDP的比重为6.5%，对比2010年的5.15%有明显增长。随着国民健康意识觉醒升级特别是对高质量医疗服务的持续追求，医疗卫生消费将迎来长期稳步发展阶段。医疗机构和政府部门会有更多的资金投入到医

疗信息化建设和互联网医院建设中来。

（2）就业机会持续增加

互联网医院作为医疗服务的新兴模式，为多个领域提供了就业机会，从医疗专业人员到技术人员和客户服务代表，以及管理和研究人员等。这些职位不仅有助于提供高质量的医疗服务，还有助于促进经济增长和就业机会的创造。

2.3.3 社会环境（Social）

（1）市场需求持续增长促进互联网医院的发展

我国人口老龄化问题凸显、慢性病患病率逐年上升，老年人由于行动受限，不方便到实体医院就医，互联网医院可以帮助老年人通过在线问诊、健康监测等服务[10]；对于慢性病患者，提供慢性病复诊、随访等服务。尤其是在疫情期间，人们已经感受到远程医疗带来的便捷服务，互联网医院的市场需求不断增长。

（2）医疗资源分配不均衡促进互联网医院的发展

我国医疗资源存在地区分配不均的问题，且城乡差异较大。总体来说，我国优质医疗资源大多分布在一二线城市。另外，医疗机构间的资源差异也比较大，三级甲等医院，拥有先进的医疗设备，优质的医疗人力资源。互联网医院的发展有助于平衡医疗资源。

2.3.4 技术环境（Technological）

互联网技术的不断发展为互联网医院提供了强大的技术支持，例如：互联网医院可以利用人工智能来进行疾病诊断、药物推荐、病历分析和患者跟踪；互联网医院可以利用大数据分析来处理和分析患者的医疗数据，提供更精准的诊断和治疗建议。大数据分析还有助于识别健康趋势、疫情爆发和药物疗效，提供有针对性的医疗服务；区块链技术可以提供医疗数据的安全和隐私保护，确保患者的个人健康信息得到安全存储和传输[11]。它还有助于建立可信的医疗数据共享系统，促进医疗信息的互通性。这些技术因素不仅提高了互联网医院的效率和质量，还创造了更多的医疗服务和治疗选择。随着技术的不断进步，互联网医院将继续受益于这些技术。

2.4 互联网医院的威胁（Threats）

2.4.1 推广程度不足

根据国家远程医疗与互联网医学中心、健康界联合发布的《2021 中国互联网医院发展报告》数据显示，超过 90%的互联网医院处于“建而不用”的僵尸状态[12]。经调查发现主要问题两方面：（1）医疗机构管理部门和物价管理部门未形成整合政策，导致互联网医院线上诊疗费用规范和条目缺失，大多数医院给与的激励制度不够，医生不愿意花更多的时间进行线上诊疗；（2）目前互联网医院开展的业务范围多数只停留在预约挂号、慢病复诊等，患者对互联网医院认可度不够，尤其是老年患者，本身可以有充足的时间花在排队候诊上，加之对智能设备操作的不熟练更加降低了其在网上看病的意愿。

2.4.2 数据共享不足。

各医疗机构之间存在信息壁垒，数据共享程度不高，数据多元分析、整合利用的能力不强，数据孤岛现象仍存在。导致互联网医院难以实现诊前一诊中-诊后体系的贯通，特别是诊后的监测、指导等。如何形成医院间信息数据共享的局面仍然是发展的一大难题[13]。

2.4.3 市场竞争压力

市场竞争对手不断涌入互联网医疗行业市场，可能导致价格竞争和利润下降。有可能引起一些小微企业放弃原有市场版图，导致了市场的不稳定性。

3 互联网医院发展的对策建议

3.1 完善政策保障，创新监管方式

医疗机构应明确互联网医院发展定位，不断优化服务流程、创新服务模式，打破线上线下界限，为患者提供全场景、全流程的畅通诊疗路径。在省级互联网医疗服务监管平台监管下[14]，规范医务人员线上诊疗行为；线上线下实行同质化医疗行为监管，确保医务人员线上诊疗质量和安全，形成事前准入、事中监管、事后追溯的全流程监管体系[15]，切实推动医疗服务高质量发展，提升人民群众就医满意度。

3.2 加强宣传引导，转变就医观念

加强对互联网医院的宣传推广，提高医患认知水平，拓展互联网医院的使用广度和深度。政府加快制定互联网医院各类在线诊疗定价标准，根据专家级别和工种设置分级定价，充分体现资深专家的医疗价值，调动医疗专家上网诊疗的积极性。逐步推广医保网上支付，引导常见慢性病患者到互联网医院平台复诊，优化医疗资源配置。创新多种在线诊疗方式，满足患者个性化诊疗需求。

3.3 优化营商环境，降低交易成本

持续放宽市场准入，并实行全国统一的市场准入负面清单制度。加大对民营企业、中小企业的支持力度，进一步破除隐性门槛，推动降低市场主体准入成本；加快建立高效便捷、优质普惠的市场主体全生命周期服务体系，全面提高线下“一窗综办”和线上“一网通办”水平，推动降低市场主体办事成本；健全大数据、人工智能、基因技术等新领域、新业态知识产权保护制度，切实保护市场主体合法权益。

4 总结

通过 SWOT-PEST 分析，我们可以看到互联网医院在优化营商环境下有望迎来更广泛的发展。政府的政策支持、市场需求增长以及技术创新都为互联网医院提供了机会。然而，互联网医院仍然面临政策风险、市场竞争和技术创新等挑战。在不断优化的营商环境下，互联网医院需要制定合适的战略，确保其可持续发展，为患者提供更便捷、高质量的医疗服务。

PU-028

基于实证数据的互联网医院线上线下融合模式的价值评估

张萍 乔一鸣*
大同市第五人民医院

随着信息技术和互联网的快速发展，医疗服务逐渐从传统的面对面咨询模式转向更加便捷、高效的在线服务模式。互联网医院作为这一变革的产物，通过提供在线问诊、电子处方、远程会诊、护理上门等服务，极大地提高了医疗服务的可达性和效率。近年来互联网医院在国内逐渐形成了线上线下相结合的融合发展模式。互联网医院借助互联网，人工智能和大数据技术,实现患者就医过程的线上化,与传统医院达成合作,形成线上线下融合的新模式。

这种线上线下融合模式最大的优势在于融合了互联网医院的便捷高效与实体医院的专业服务，不仅能够让患者在线上得到及时、准确的医疗咨询，还能在必要时引导他们到实体医院进行更详细的检查和治疗。这一模式的成功实施，不仅能提高医疗服务的效率和质量，还能在很大程度上解决基层医疗机构资源不足的问题。同时一些慢性病的患者也可以获得线上长期健康管理。

其次，本研究从数据层面评估这种新兴模式的科学性与先进性，为模式的广泛推广提供理论依据，积极引导医患关系和就医行为的变化，推动医疗行业信息化建设和模式创新。

一、数据和方法

本文数据来源于大同市第五人民医院互联网医院，数据范围涵盖医院互联网医院上线至今的问诊、处方、护理上门、线上线下联合义诊的医疗服务数据，在经过数据脱敏建模和整理，进一步分析这种融合发展模式在整合优势资源,提高医疗服务效率和体验的效能和价值,为互联网医院线上线下融合模式的发展方向提供一定的研究性支撑。

1 问诊服务

基于从“大同市第五人民医院互联网医院问诊订单”中获得的数据，我们可以看到互联网医院已经开展了多科室线上问诊服务，不仅可以方便了患者，还使医生能够更灵活地提供服务。

数据预处理结果

总记录数：6119 条

涉及医生：126 人

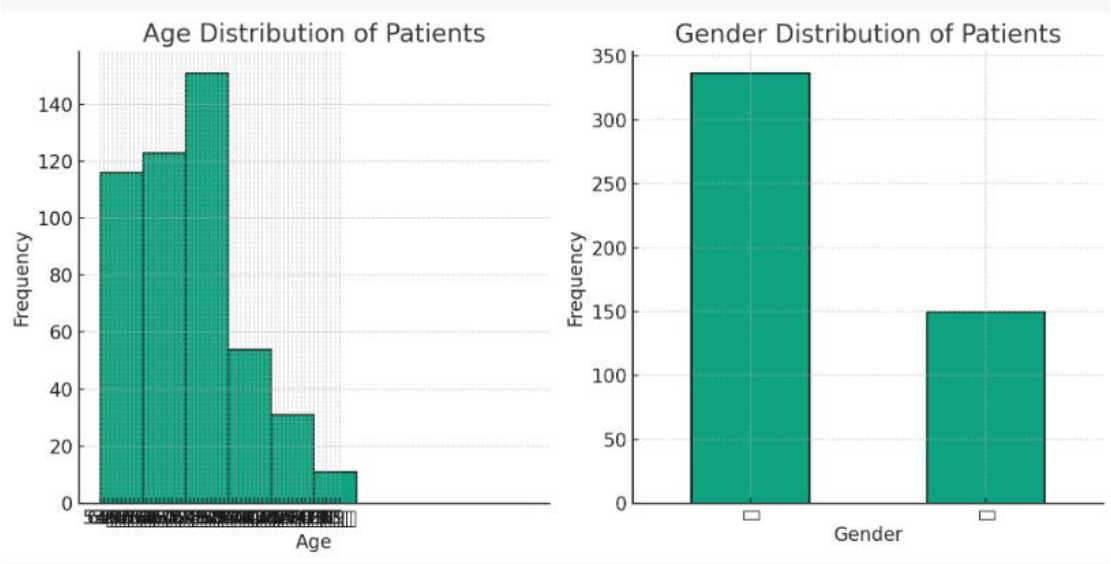
平均价格：约 12.08 元

交易状态：主要为“已付款”，具体数量为 5042 条

主要从需求分布、价格可承受性、交易成功率三方面展开。

2 线上线下融合远程会诊方面

基于从“大同市第五人民医院互联网医院线上线下融合会诊患者统计表.xlsx”中获得的数据，本研究从患者群体特征与医疗、专家参与和科室分布、诊断结果和建议需求、时间分布 4 个方面进行模式效能分析。



不难看出会诊服务受到了各个年龄段患者的欢迎，其中大部分患者的年龄集中在 30 至 60 岁之间。还有一部分儿童和青少年在监护人陪同情况下利用这项服务。患者性别分布：男性患者使用义诊服务的频率稍高于女性患者。

3 护理业务方向

基于从“大同市第五人民医院互联网医院护理上门服务登记表.xlsx”中获得的数据，从月度需求变化、服务类型和费用分布、用户特性方面进行分析。

（1）服务类型和费用分布

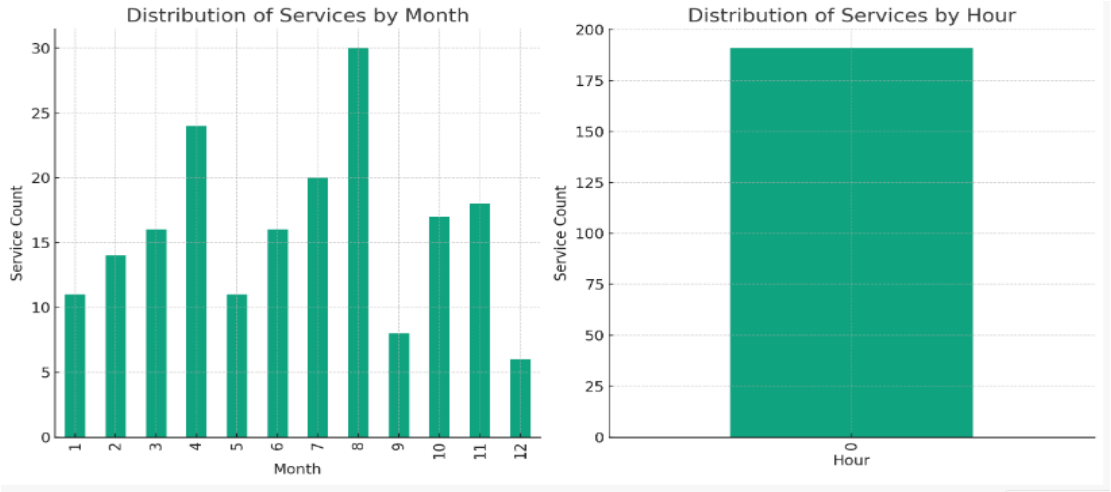
最常见的上门护理服务是“留置、更换尿管”，共有 81 次，其次是“静脉采血”，共有 51 次，平均服务费用约为 131.48 元，最低费用为 83.5 元，最高费用为 245.76 元。

（2）用户特性

平均年龄约为 72.22 岁，最大年龄为 93 岁。

性别分布方面，男性用户（129 次）明显多于女性用户（70 次）。

（3）季节性和月度需求分析：月度需求数据表明，可能存在季节性因素影响上门护理服务的需求。这一点值得进一步研究。



（4）用户需求多样性：不同类型的护理服务需求表明，未来上门护理服务应更加多样化和个性化。

（5）用户特性和服务定位：由于用户年龄偏高，上门护理服务似乎主要针对老年人群，这在未来的服务优化和资源分配中是一个重要因素。

（6）性别差异与文化因素：男性用户明显多于女性，这可能是文化或社会因素影响的结果，值得进一步研究。

在以上数据分析过程中，我们使用了一些统计学方法和建模技术：

（1）描述性统计分析

中心趋势分析：计算平均值、中位数和众数来描述数据的中心趋势。

离散程度分析：通过计算标准差和方差来描述数据的离散程度。

（2）探索性数据分析 (EDA)

数据可视化：使用条形图、折线图、箱线图等来可视化数据并识别任何模式或趋势。

相关性分析：通过计算相关系数来探索变量之间的关系。

（3）时间序列分析

季节性分析：识别数据中的任何季节性模式或趋势。

趋势分析：使用时间序列分析来识别长期趋势和模式。

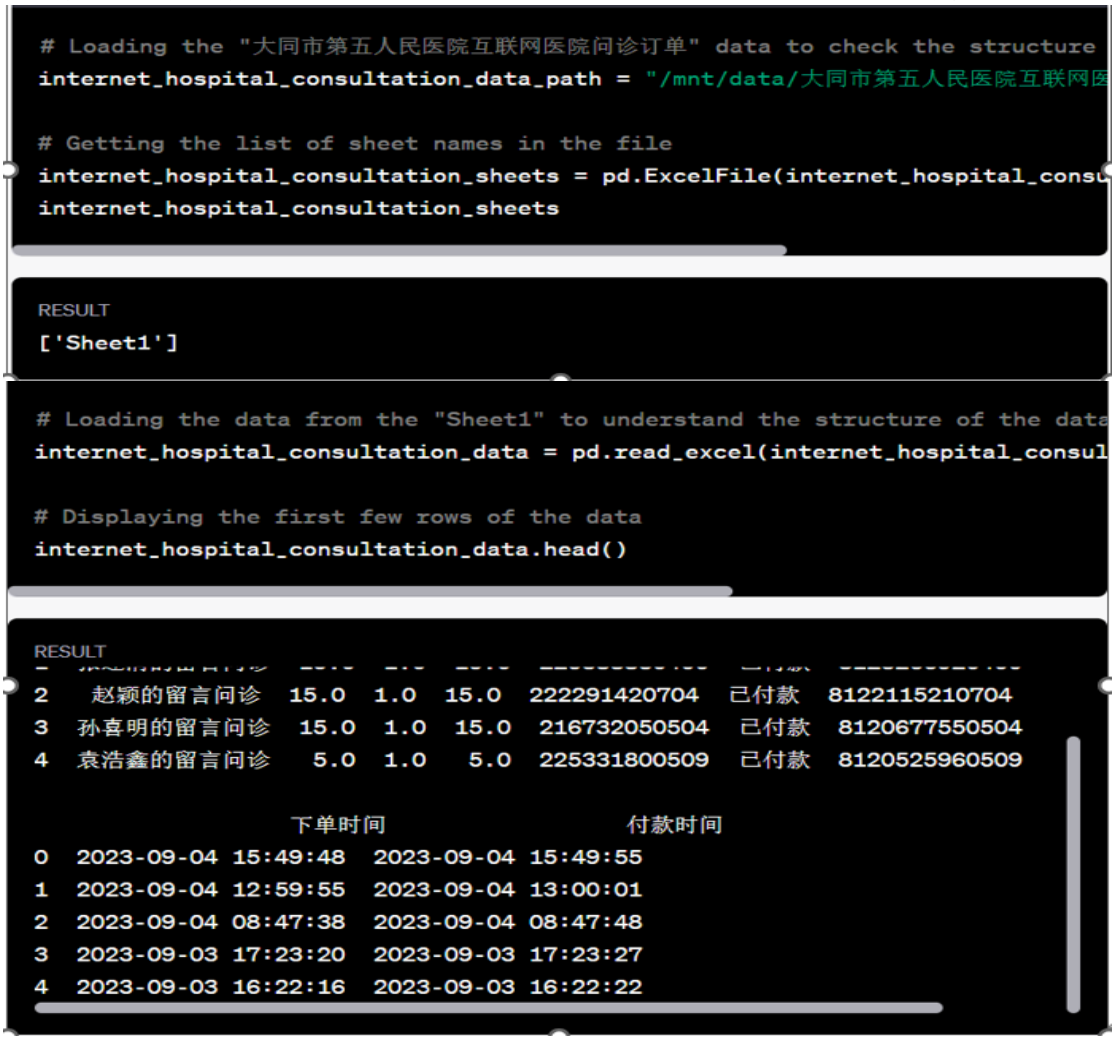
（4）机器学习和预测建模

聚类分析：使用机器学习算法来识别数据中的群体或聚类。

分类和回归分析：根据历史数据来构建预测模型和分类器。

在进行数据分析时，我们将首先清理和整理数据，并将数据进行脱敏处理，然后使用上述技术来探索和分析数据。分析的最终目的是识别线上线下融合模式的价值并提供有根据的结论。

其中模型的建立和算法训练是一个关键过程，以下是我们分析过程中的一些代码示例：



```
# Loading the "大同市第五人民医院互联网医院问诊订单" data to check the structure
internet_hospital_consultation_data_path = "/mnt/data/大同市第五人民医院互联网医
# Getting the list of sheet names in the file
internet_hospital_consultation_sheets = pd.ExcelFile(internet_hospital_consu
internet_hospital_consultation_sheets
```

```
RESULT
['Sheet1']
```

```
# Loading the data from the "Sheet1" to understand the structure of the data
internet_hospital_consultation_data = pd.read_excel(internet_hospital_consul
# Displaying the first few rows of the data
internet_hospital_consultation_data.head()
```

```
RESULT
2   赵颖的留言问诊  15.0  1.0  15.0  222291420704  已付款  8122115210704
3  孙喜明的留言问诊  15.0  1.0  15.0  216732050504  已付款  8120677550504
4  袁浩鑫的留言问诊   5.0  1.0   5.0  225331800509  已付款  8120525960509

                 下单时间                付款时间
0  2023-09-04 15:49:48  2023-09-04 15:49:55
1  2023-09-04 12:59:55  2023-09-04 13:00:01
2  2023-09-04 08:47:38  2023-09-04 08:47:48
3  2023-09-03 17:23:20  2023-09-03 17:23:27
4  2023-09-03 16:22:16  2023-09-03 16:22:22
```

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
# 1. 数据获取和清洗
data = pd.read_csv('your_data.csv')
data = data.dropna()  # 删除缺失值
```

```
# 2. 特征工程
X = data[['feature1', 'feature2', 'feature3']]
y = data['target']  #
# 3. 数据分割
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# 4. 数据标准化
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
# 5. 模型选择和训练
model = LinearRegression()
model.fit(X_train, y_train)
# 6. 模型评估
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
# 7. 结果解释
# 通过可视化或其他方式来解释和展示结果
```

算法训练过程简述：

1 定义模型

2 损失函数

3 优化算法

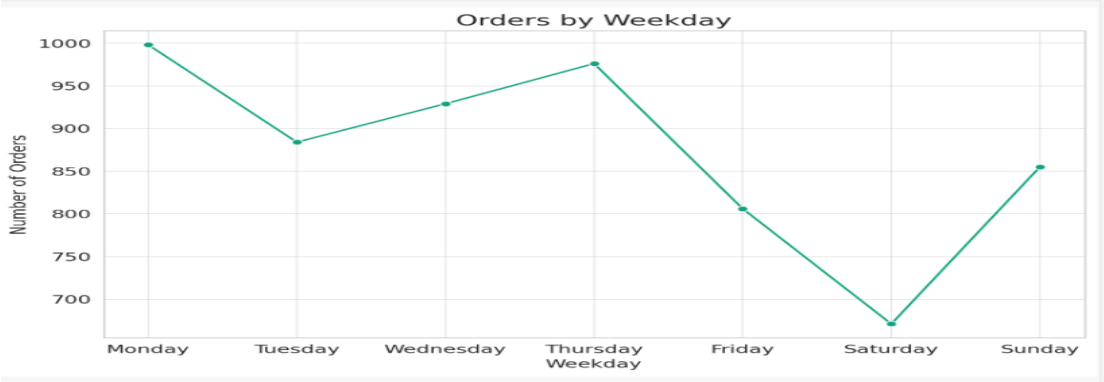
本次我们使用了最常用的优化算法梯度下降

```python
from sklearn.linear_model import LinearRegression

# 创建一个线性回归模型实例
model = LinearRegression()

# 使用训练数据来训练模型
# X_train: 训练数据的特征
# y_train: 训练数据的目标变量
model.fit(X_train, y_train)
```

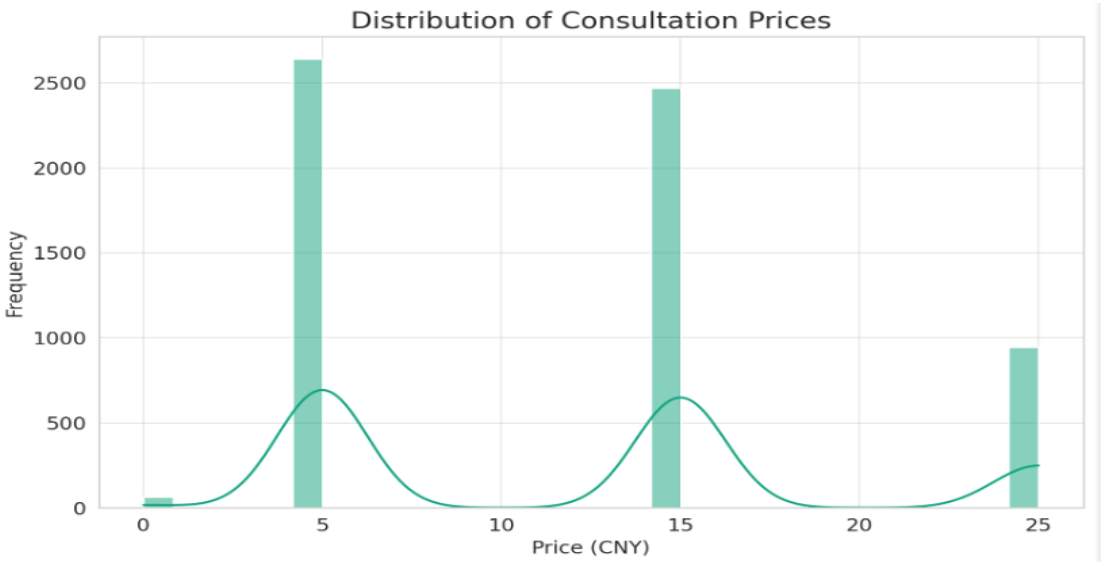
二、结论和分析

1、问诊方面

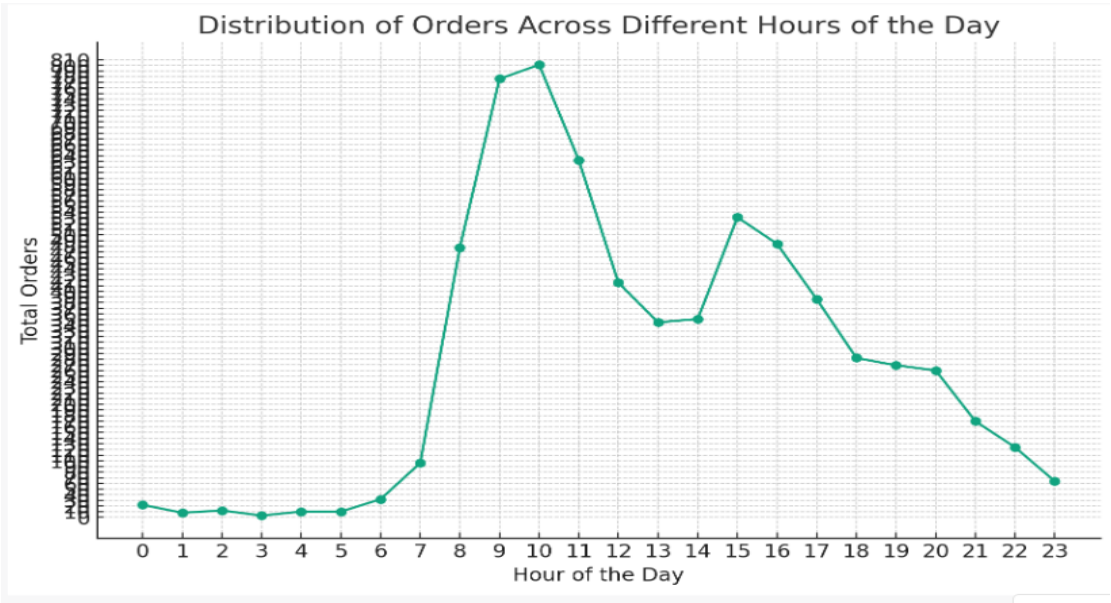
我们可以看出，按周天分布：订单数量在周一至周五相对稳定，周末略有下降。



价格可承受性与交易成功率分析结果：问诊价格主要集中在 5-15 元之间，相对较低。交易成功率：已付款订单数量为 5042，总订单数量为 6119，交易成功率约为 82.4%。



时间分布方面: 在凌晨 0 点至 4 点之间的订单量相对较少，订单数量在白天达到峰值，晚上逐渐减少。



从实证数据不难看出互联网医院线上线下融合模式具备多方面重要价值：

高效和可接近：根据统计问诊订单平均响应时间是 15 分钟，这远远低于一个人在实体医院等待的时间，这突显了线上服务的高效性。工作日的白天，订单数量最多，在高需求时段快速处理大量订单，展示了其高效性。大部分的订单价格都集中在 5 元和 15 元，这表明服务非常便宜和可接近，可以吸引更多的人使用，可以使得优质医疗资源覆盖到一个更广泛的地理区域，尤其是那些远离医院或行动不便的人群。

明显的经济效益：对患者而言，可以减少交通费用，假设一次去实体医院的往返交通费用平均为 50 元，那么通过 6119 次的在线咨询，用户共节省了 305,950 元的交通费用。节省住宿和餐饮费用：对于那些不得不因为就医而留宿的人来说，他们可以通过在线咨询来节省住宿和餐饮费用。假设每次实体医院访问的平均住宿和餐饮费用为 200 元，那么通过在线咨询，用户可能已经节省了 1,223,800 元。对于医生而言，通过线上问诊服务，医院可以创建一个新的收入流。平均每笔订单收入为 10 元，则一个月的收入可以达到 61,190 元。

显著的社会效益：通过线上问诊服务，可以减轻实体医院的压力，从我们的 6119 笔订单，可以推断，这可能帮助减少了大约 6000 次实体医院的访问，从而为那些需要紧急或面对面医疗服务的患者腾出了空间。并且基于 6119 次订单在一个月中的分布，我们可以简单得出平均每天都有 200 个额外的就医机会并且不受时间和空间约束。其次还有高效的资源分配: 通过线上平台，医院可以更有效地分配其资源。例如，可以通过分析数据来识别高需求时间段，并相应地调整资源分

配。假设在工作时间有最高的订单需求，医院可以安排更多的医生在这段时间在线。

全天候服务：从早上 7 点一直到晚上 23 点这段时间订单数量最多，确保了患者可在工作日间隙也能方便地使用该服务，而且服务仍然在深夜保持活跃，使得人们在任何时候都能够寻求帮助，特别是在传统医疗服务不可用的时间，极大的体现了该模式的重要性。

2、线上线下融合远程会诊方面

通过分析最常见的诊断词，我们可以看到每个年龄和性别组的主要健康问题。模型中，我们选取了每个组最常见的 3 个诊断词来描绘他们的主要健康问题。

主要健康问题：

15-20 岁：主要面临的问题是特发性面神经麻痹、胸痛和头痛。

21-30 岁：双膝关节疼痛和腰部问题成为主要的健康问题。

31-40 岁：糖尿病、高血压和脂血症成为这个年龄段的主要问题。

41-50 岁：慢性胃炎、功能性消化不良和高血压病是这个年龄段的主要健康问题。

51-60 岁：高血压、糖尿病和头晕是这个年龄段的主要健康问题。

61-70 岁：冠心病、慢性胃炎和高血压病是这个年龄段的主要健康问题。

71 岁以上：高血压、心律失常和冠心病是这个年龄段的主要健康问题。

性别相关的特定健康问题：

男性：糖尿病、高脂血症和高血压病是男性主要的健康问题。

女性：2 型糖尿病、糖尿病和高血压病是女性主要的健康问题。

以上是对各个年龄段和性别的诊断词进行深入分析的结果。从这些数据中，我们可以得到一个初步的了解，即高血压和糖尿病是最常见的健康问题，而这些问题随年龄的增长而变得更加严重。这为医院提供了有关需要加强哪些健康教育和预防措施的信息，可以提升医疗服务效能，高效地提供针对性的基层医疗服务和健康指导。

3、护理业务方向

总体来说互联网医院线上线下融合模式的护理上门的价值集中表现：

高需求与广泛覆盖：护理上门服务对于有身体不适、行动不便或缺乏时间去医院的人来说是非常有价值的。此服务提供了一个方便的方式，使他们能够在家中接受必要的医疗护理。

经济性与可承受性：合理的费用结构使得更多的人可以负担得起此项服务。此外，线上预订和支付方式提供了额外的便利性，使整个过程更加高效。

高交易成功率：高交易成功率表明服务具有很高的可靠性，用户对该服务的认可度高。

三、综合讨论

汇总分析互联网医院线上线下融合模式在问诊、护理和会诊方面已经展现出显著的价值和巨大的潜力。对于患者而言，这种模式确保了更高的便利性和可承受的费用。全天候的问诊服务消除了时间和地点的限制，使得医疗咨询变得更为简单和方便。在护理方面，它提供了个性化和定制的服务，同时也减少了患者的等待时间。会诊方面则通过提供专家的意见和建议来确保更高质量的医疗服务。对医院来说，这种模式提供了一个新的收入来源，同时也优化了资源的分配和管理，提高了服务的效率和效果。而从社会的角度来看，线上服务可以减轻实体医院的负担，提供更多的空间给需要紧急或面对面服务的人，从而提高了整体的医疗服务质量。

然而，这种模式也带来了一些挑战，包括患者在将来可能过分依赖线上服务，以及医疗人员可能面临的增加的工作压力。其次还有数据安全和隐私问题，服务质量的不一致性，数字鸿沟问题以及对实体医院和基层医疗的一定冲击，因此，我们必须解决这些问题，确保模式的成功和可持续性。

PU-029

在线慢性病社区用户的健康信息需求研究——以百度贴吧高血压吧数据为例

赵梦圆[1]，应峻[1]

1. 复旦大学图书馆

慢性病已成为 21 世纪危害全球居民健康的公共卫生问题，2018 年世界卫生组织（WHO）发布的数据显示，慢性病引起的死亡率已占全球死亡率的 70%。慢性病通常具有发病率高，但治愈率和控制率低的特点[1]，随着我国人口结构以及生活习惯的变化，慢性病在我国也已成高发态势，这不仅是一个健康问题，而且日益成为一个严重的社会问题。当前居民的健康知识知晓率普遍偏低，不健康的生活方式比较常见，由此带来的疾病问题也在日益显现。随着互联网技术的快速发展与智能手机的普及，越来越多的用户接触并使用在线健康信息，但我国目前尚未成熟的在线健康平台仍有着很大的发展空间。

在线健康社区能够将专业医务人员、病人及其家属等用户聚集在一起，允许他们在社区中自由发表或分享观点并寻求支持。尤其对于慢性病用户来说，在与疾病长期的抗衡过程中，需要获取信息、情感、评估等多方面支持来帮助他们进行有效的自我健康管理。本文主要探究在线慢性病社区用户的健康信息需求特征，利用 LDA 模型分析社区用户的交流文本数据，从而识别慢性病社区用户热议话题，为开拓“互联网+医疗健康”新业态提供借鉴。

1. 研究设计

1.1 数据爬取

当前我国互联网上的在线健康社区可大致归纳为三类，第一种是包含专业医务人员和患者的平台，提供家庭医生、健康测评、挂号、在线问诊、健康科普等多种功能，比如平安健康、好大夫在线；第二种是专业医务人员之间进行专业领域知识交流讨论的平台，比如丁香园；第三种是包含类似健康问题的患者的平台，允许患者之间相关交流沟通并分享经验，如百度贴吧、慢友帮。其中，百度贴吧是结合搜索引擎建立的一个在线交流平台，基于关键词的主题交流平台，用户可以通过搜索满足需求，找到对同一话题感兴趣的网友，是全球领先的中文社区。同时考虑到中国是高血压的高发国家，成人高血压患病率已经达到了 27.5%。尤其随着生活水平的提高，高血压的发病率也在增加。大约四分之一的成人患有高血压，这是一种常见且多发的疾病。本研究主要关注慢性病患者在日常生活中如何搜寻健康信息来预防和控制疾病，所以笔者选取百度贴吧的“高血压吧”为例，以高血压吧中患者之间的交流数据作为研究对象，识别在线健康社区中慢性病患者关注的健康信息主题，进而初探其健康信息搜寻行为的特征。

本研究以百度贴吧的“高血压吧”下社区用户的交流数据作为研究对象，实验过程中通过 Python 爬取高血压吧中的话题、帖子和回复内容等数据。图 1 给出了高血压吧中用户生成内容示例。爬取数据过程中去除带有广告、本吧吧主火热招募等标注的数据，爬取日期截止至 2021 年 9 月 20 日，最终实验数据包括 9268 条话题、6512 个帖子和 6182 条回复，共 2 万多条交流记录。

图 1 高血压吧中用户生成内容示例

1.2 研究方法

考虑到直接统计文章出现的词语作为主题来源的方法，一方面忽略未提取词语的语义特征，另一方面又比较依赖作者或者数据库等给定的主题范围[2]，因此，许多学者采用主题模型来获取文章的研究主题。唐晓波[3]、李湘东[2]、曲佳彬等学者[4]均利用 LDA 模型进行了主题演化分析，较好地分析了研究内容随时间变化的规律。因此，本文利用 LDA 模型进行主题聚类分析，进而探讨用户的在线健康信息需求特征。

2. 研究结果

首先，通过网络爬虫方法采集高血压吧中用户生成内容，即话题、帖子和回复。然后，将用户发表的话题、帖子和回复内容提取出来，进行文本预处理操作——符号去除、分词、停用词过滤，再对文本进行 LDA 建模，通过不同主题下的困惑度值来确定最优主题数。如图 2 所示，当主题数为 12 时，困惑度值达到极小值，表明此时模型的泛化能力较强。

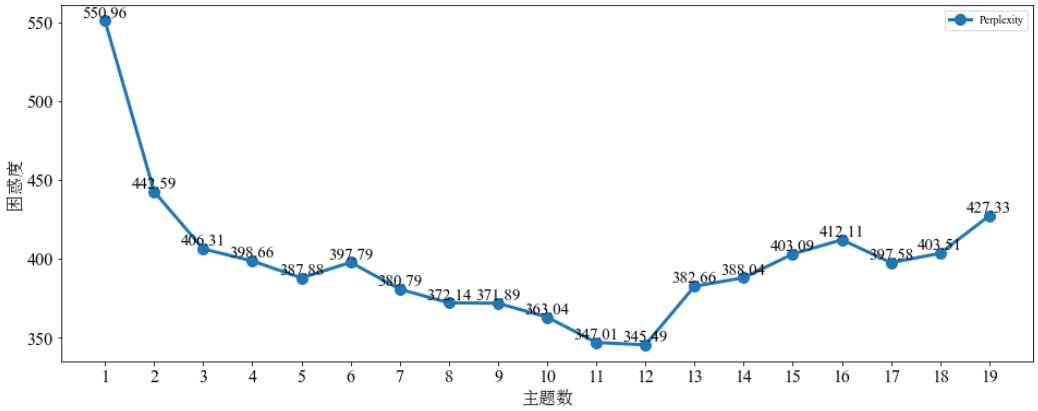


图2 不同主题数下困惑度值的变化情况

通过不同主题数下的困惑度值确立了最优主题数为 12，表 1 展示了实验得到的 12 个“主题-词”概率分布，每个主题给出了对该主题贡献度较大的前 12 个有含义的关键词及其权重来区分主题。基于此，本研究借鉴盛姝等人构建的结直肠癌主题分类体系[5]对识别的主题进行归纳整合，具体结果如表 2 所示。从主题归纳结果可以看出，在线健康社区的慢性病用户主要搜寻疾病、治疗方案、诊断检查、症状描述和社会生活这五种类型的健康信息。

表1 主题关键词及其概率分布

主题	**主题关键词和概率分布**
主题 1	血压（0.156）、正常（0.033）、医院（0.029）、医生（0.023）、血压高（0.018）、问题（0.015）、检查（0.014）、体验（0.013）、发现（0.011）、测量（0.009）、测血压（0.008）、紧张（0.007）
主题 2	患者（0.098）、非洛地平（0.036）、中风（0.035）、尿蛋白（0.025）、遗传（0.015）、鱼跃（0.010）、普利（0.010）、住院（0.009）、范围（0.009）、科技（0.006）、复查（0.006）、胶囊（0.005）
主题 3	血压计（0.072）、电子（0.037）、跑步（0.018）、水银（0.017）、恢复（0.016）、记录（0.013）、下降（0.010）、老人（0.010）、低血压（0.009）、缓解（0.009）、不适（0.009）、打卡（0.008）
主题 4	降压药（0.073）、效果（0.049）、注意（0.041）、请问（0.031）、比较（0.029）、怎么办（0.028）、情况（0.026）、年轻（0.015）、副作用（0.014）、推荐（0.014）、失眠（0.012）、严重（0.008）
主题 5	服用（0.042）、服药（0.034）、沙坦（0.031）、氨氯地平（0.030）、缬沙坦（0.026）、强的松（0.025）、洛尔（0.024）、阿利（0.021）、美托（0.021）、琥珀酸（0.019）、服片（0.017）、血液（0.017）
主题 6	引起（0.043）、血管（0.036）、耳鸣（0.034）、症状（0.033）、疾病（0.023）、导致（0.021）、耳聋（0.010）、心脑血管（0.009）、常见（0.007）、并发症（0.007）、心血管（0.007）、肥胖（0.006）
主题 7	*治疗（0.077）、降压（0.063）、中医（0.035）、病友（0.018）、年轻人（0.016）、食物（0.016）、有效（0.014）、帮助（0.013）、西药（0.011）、请教（0.009）、维护（0.008）、病情（0.007）、*
主题 8	*交流（0.070）、调理（0.057）、预防（0.048）、降血压（0.045）、换药（0.035）、中药（0.032）、养生（0.026）、糖尿病（0.011）、食疗（0.010）、加油（0.010）、维生素（0.009）、注意事项（0.006）*
主题 9	*运动（0.041）、头晕（0.039）、心率（0.029）、本人（0.024）、产品（0.019）、体重（0.016）、减肥（0.015）、焦虑（0.015）、熬夜（0.012）、喝酒*

	（0.012）、睡眠（0.011）、难受（0.009）
主题 10	*低压（0.106）、高压（0.083）、左右（0.052）、早上（0.034）、晚上（0.031）、偏高（0.019）、欧姆龙（0.019）、下午（0.015）、每次（0.012）、绿色（0.011）、一片（0.009）、尿酸（0.007）*
主题 11	*健康（0.062）、根治（0.030）、三高（0.021）、咨询（0.020）、免费（0.019）、管理（0.016）、舒张压（0.014）、知识（0.011）、功效（0.010）、专家（0.009）、经验（0.008）、科普（0.008）*
主题 12	*吃药（0.079）、控制（0.043）、饮食（0.037）、生活（0.020）、身体（0.019）、长期（0.014）、药物（0.014）、检测（0.014）、锻炼（0.013）、建议（0.013）、停药（0.012）、生活习惯（0.005）*

表 2 主题归纳

类目	子类目	对应主题
1.疾病	1.1 确诊	主题 1
	1.2 成因	主题 2、主题 9
	1.3 并发症	主题 6
2.治疗方案	2.1 疾病护理	主题 7
	2.2 药物	主题 2、主题 5
	2.3 非药物治疗	主题 9、主题 12
	2.4 治疗效果	主题 4、主题 7
3.诊断检查	3.1 指标数据	主题 10、主题 3
	3.2 血压检测	主题 3、主题 10
4.症状描述	/	主题 9
5.社会生活	5.1 情感支持	主题 8
	5.2 健康科普	主题 11
	5.3 疾病预防	主题 8

2.1 疾病

慢病社区用户在高血压吧经常交流分享关于疾病的确诊、成因和并发症等健康信息。确诊对应着识别的主题 1，该主题下血压的权重最高。主题 2 中的关键词有提到和高血压成因相关的要素，比如遗传；以及主题 9 下的熬夜和喝酒等都有可能诱发高血压。高血压可能引起的并发症对应着主题 6。由此可见，慢性病患者愿意使用互联网去搜寻与健康相关的信息，将自己的健康情况分享到在线健康社区平台，并积极参与社区讨论。

2.2 治疗方案

关于高血压的治疗方案也是患者在高血压吧热议的话题，由于高血压是一种慢性病，目前没有方法能够治愈高血压，对于高血压患者来说，治疗目的是可以将血压控制在正常范围内，以及不引发其他疾病。用户在健康社区中讨论的高血压治疗方案主要包括疾病护理、药物、非药物治疗和治疗效果四个方面。此外，一些主题下还出现了请问、推荐、怎么办、请教等寻求帮助的关键词，由此可见，在线健康社区为慢性病用户进行自我健康管理提供了信息支持。

2.3 诊断检查

许多患者会将高血压吧作为日常健康管理工具或在线问答平台，用来分享或记录自己的日常血压范围，甚至在提问时会附上自己具体的血压情况，所以在高血压吧关于诊断检查的健康信息也很常见，主要体现在指标数据和血压检测两个方面。从关键词可以看出有些用户会将在线健康社区当

作健康管理工具或者在线问答平台来记录自己的健康情况以及寻求健康信息支持，从而更有效地预防和控制疾病。

2.4 症状描述

用户可能在症状没有显现的情况下就患上了无症状的疾病，特别是高血压患者。高血压典型症状表现在头痛、心律失常、疲倦不安、心悸耳鸣等方面。有些高血压患者愿意将个人疾病的症状表现发布至在线健康社区上，引发具有类似情况的用户进行讨论和分享经验，从而获取有用的健康信息以及心理安慰。

2.5 社会生活

用户在高血压吧中还经常讨论一些关于社会生活主题的信息，即他们通过社交活动而产生的日常交互共享信息，主要体现在情感支持、健康科普和疾病预防三个方面。如疾病预防对应着主题8，其中的调理、预防、降血压、养生、食疗、注意事项等关键词都可以看出患者关于疾病预防的相关讨论。

3. 结论与展望

本研究利用 LDA 主题模型对高血压吧用户的交流文本进行主题聚类，基于主题分布及对应关键词归纳了 5 种健康信息需求类型，总的来说，可以将用户的健康信息需求特征归纳为四个方面。

（1）用户更多搜寻与疾病本身相关的健康信息。从高血压吧中提取的主题大多聚焦在高血压的确诊、成因、药物治疗、治疗效果、症状描述等方面，而且体现这些主题的相关关键词的占比也比较高。虽然用户在就医过程中可能已经从医生等渠道获取健康专业知识，但高血压是一种慢性病，需要长期且持久的治疗过程，他们不断从各种便捷渠道搜寻信息来补充着健康知识，以满足个性化的健康需求。

（2）用户从慢病社区中搜寻健康信息来满足个人的情感需求。对于慢性病患者来说，由疾病带来的身体反应相对于心理变化更容易被感知到，而心理反应在疾病治疗过程中往往被忽略或者难以表达，而在线健康社区允许每位用户可以匿名表达或揭示困难及各种想法。

（3）用户将在线健康社区当作日常健康管理工具。许多用户每天将自身的健康状况在健康社区中详细记录，不仅能够追踪和监控自身的健康状况变化，更加掌握自身的健康情况；而且能够引发具有同种情况的用户的讨论，获取更多关于疾病管理的经验和信息。

（4）用户越加关注健康的生活方式来控制或预防疾病。慢性病患者因长久处于疾病的困扰中，对健康信息会更加关注，而且慢性病患者除了必要的药物治疗，最重要的是生活方式干预。从高血压吧中识别的主题可以看出，多个主题下包含对“食疗”、“养生”、“运动”等方面的讨论。

本研究聚焦百度贴吧高血压吧用户的信息需求特征，为探究在线健康社区平台中不同用户群体行为及信息需求的深入研究提供参考，为卫生部门、互联网平台、医院等专业机构为慢性病患者给予精准服务提供启发。如政府相关卫生部门可根据慢性病患者最关切的需求开展系列活动和工作部署，优化慢性病综合防治工作机制和服务网络，有助于中国防治慢性病中长期规划（2017—2025年）的实施。互联网平台可根据患者需求进一步优化功能属性，为其提供更加便捷可靠的在线健康社区平台。医院、社区卫生服务中心等专业机构可针对患者需求定期组织健康科普活动，普及患者所关注的疾病、治疗方案、诊断检查、症状描述和社会生活等健康信息，有效减少慢性病疾病困扰，提高疾病治疗的依从性，优化个体治疗方案减轻患者疾病负担，改善患者生活质量。。

本研究的局限性在于仅选取百度贴吧中高血压吧的用户交流数据来识别慢性病患者的健康信息搜寻特征，没有参考更多的数据去分析，比如用户信息、回复数量、话题对应帖子数量等其他特征值，只报告了慢性病用户关注的信息类型，未进一步明晰不同疾病类型、不同年龄组用户的搜寻特征。今后可从不同类型疾病的健康社区出发，并融合患者的个人信息、帖子情感表征等数据，进行用户画像分析，关注不同疾病不同用户的差异性，以提供更加全面、精确的研究结论。

PU-030

以患者为中心的互联网医院平台建设实践与思考

万宁[1] 刘阳[2]
1.海南医学院第一附属医院 信息中心
2.中南大学湘雅医学院附属海口医院（海口市人民医院）信息管理处

1 引言

随着互联网技术的不断创新，“互联网+医疗健康”行业扬帆起航，医疗服务圈逐步扩大，为用户提供线上挂号、健康咨询、音视频问诊、在线结算、药品配送、移动端查看报告等便捷功能。2018 年，国务院办公厅印发《关于促进“互联网+医疗健康”发展的意见》，提出允许医疗机构开展部分常见病、慢性病复诊等互联网医疗服务，为“互联网+医疗健康”明确了发展方向[1]。2020 年，国家卫健委印发《关于深入推进“互联网+医疗健康”“五个一”服务行动的通知》，总结推广在疫情防控期间，各地创新的线上服务模式，持续推动“互联网+医疗健康”便民惠民服务向纵深发展[2]。因此，代表着医疗行业新发展方向的互联网医疗，不仅有利于解决中国医疗资源不平衡和人们日益增加的健康医疗需求之间的矛盾，更是卫生主管部门积极引导和支持的医疗发展模式，能有效加强常态化疫情防控的信息技术支撑，为老百姓提供足不出户的健康咨询和线上就诊服务。

与此同时，海南省也积极贯彻落实国家政策精神。2020 年，海南省卫生健康委员会为规范和推进省内互联网医院持续健康发展，规范互联网医院管理，印发了《海南省互联网医院管理办法（试行）》琼卫规〔2020〕3 号[3]，明确互联网医院准入规则、执业规则和监督管理规则。另外，海南省卫健委在《海南省互联网医院医疗服务事前监管指南(第一版)》琼卫医函〔2020〕193 号[4]中，列出具体的互联网医院核查方法，从执业规则、人员资质、诊疗科目设置、管理体系、规章制度、信息安全、设备部署、系统软件、诊疗流程、监管对接、数据全程留痕、知情同意、患者满意度等多维度规范互联网医院医疗服务管理。

2 明确以医院为主体的互联网医院建设目标

互联网医院通过开展线上业务可以与线下实体医疗机构相辅相成、有机整合，提供可信的线上线下一体化、闭环式医疗服务，同时利用线上平台在不同等级实体医院之间建立转诊通道，以患者流动激活基层医疗资源，从而达到优化医疗资源结构，提升患者就医体验，缓解医生工作压力，提升医院运行效率的目标。

3 互联网医院建设的方式与方法

互联网医院建设的三种模式：①实体医院为主体的互联网医院的自建模式。②第三方平台型互联网医院的运营模式。③联合共建互联网医院的代建模式。海南医学院第一附属医院互联网医院积极探索，明确了自行建设的模式，以实体医院海南医学院第一附属医院（以下简称：海医一附院）为主体，由医务部牵头成立互联网医院运营管理办公室，明确各相关部门在互联网医院中的工作职责。

3.1 互联网医院平台整体模型

海南医学院第一附属医院互联网医院平台简称“海医互”，是以“统一平台、统一门户、统一终端、统一服务体系”的四个统一模型的方法构建整体业务架构的平台体系。

3.1.1 线上线下统一的“海医互”平台

“海医互”平台是以海医一附院混合云数据中心为基础[5]，对下与海医一附院院内 HIS 系统进行对接，实现对基础医疗服务的数据支撑。对上与海南省区域医疗卫生信息平台进行数据上报，形成医院、上级行政主管部门、医务人员、患者、药品物流配送等第三方相关机构的综合服务平台。

3.1.2 互联网医院平台的统一门户

“海医互”平台为医师和药师构建了统一的门户网站，提供互联网在线排班、线上问诊、医嘱处方开单、书写门诊电子病历、线上审方、在线随访等一体化的功能。

3.1.3 移动终端统一接入

统一终端也就是统一 App：“海医互”平台为医务人员和患者量身定制了移动端 App，方便医患进行沟通交流。其中患者端 APP 提供如 AI 健康咨询和智能导诊、预约挂号、在线复诊、线上结算、报告单查询、药品配送等功能，另外患者也可通过微信小程序的应用入口进行线上问诊。

3.1.4 统一“研发+运营+维护”的服务体系

海医一附院根据自身特点，建立了统一的研发+运营+维护服务体系，以保障互联网医院平台的实际运行效果，同时保障互联网医院平台的不断升级优化和信息安全。

3.2 互联网医院系统架构

互联网医院系统平台由应用层、服务层、业务层、数据层、管理层等 5 层构成。

3.2.1 应用层负责与平台用户的业务交互

“海医互”平台由互联网医院后台管理系统、医生端 PC 网站、医生端 APP、患者端 APP、患者端小程序和药师端 PC 网站组成，为医生用户提供网上看诊功能，为患者用户提供网上就诊服务，主要包括在线问诊、在线健康咨询、处方病历查询、接收检验检查报告等业务功能。

3.2.2 服务层采用身份识别与数字签名保障接入安全

“海医互”平台服务层采用 API 技术为终端用户提供细颗粒度的业务服务。对医师、患者、药师进行用户身份识别，通过调用腾讯云人脸核身接口，传送身份证号获取用户信息，人脸识别进而实现身份验证。线下院内 HIS 系统使用物理介质 KEY 进行数字签名，线上平台数字签名通过医网信 APP 实现。医生和药师通过手机验证码核实身份登录医网信 APP，在医疗文书确认时通过自动调用数字签名完成病历和处方开立。

3.2.3 业务层与医院集成平台进行业务对接

近年来，海医一附院已通过国家医疗健康信息互联互通标准化成熟度四级甲等测评，线下系统已经全部接入集成平台。目前，通过互联网医院平台在集成平台的接入，使数据达成了共享，线上与线下业务得到了有效联动。

3.2.4 数据层采用有效的数据保护措施

海医一附院已建立全量的临床数据中心（CDR），由 Hadoop 分布式集群构建而成，以数据流为基础，采用大数据技术进行清洗、转换加载到数据中心，优化数据质量；通过数据中心的主数据映射功能，实现各个业务系统数据间的标准统一，将不同厂商异构数据进行高度整合；并利用大数据应用技术，建设基于数据中心的应用，如患者全息视图、运营决策支持等系统。

3.2.5 管理层

通过管理层能够提高后台业务管理和运维管理层的效率。

3.3 线上、线下一体化医院系统平台服务架构

通过复用集成平台构建的互联网开放平台，打通线上线下系统应用服务，形成一体化的系统服务架构，其中部分功能仍在不断完善中。

3.3.1“海医互”平台基于混合云架构

如图 1 互联网医院系统平台网络拓扑图所示，部署在海南省内公有云平台和医院本地私有云机房中，共计 9 台公有云服务器+4 台院内私有云服务器以及相关云安全服务。公有云计算平台通过 20M 专线点对点连接医院本地私有云服务器，链路之间根据业务需求和安全管理需求配置访问控制策略，具备入侵防御和恶意代码防范能力，在网络节点处对重要的用户行为和重要的安全事件进

行审计，从而保障数据及应用服务安全。作为核心业务，海医互平台的信息安全水平符合等保三级的要求。

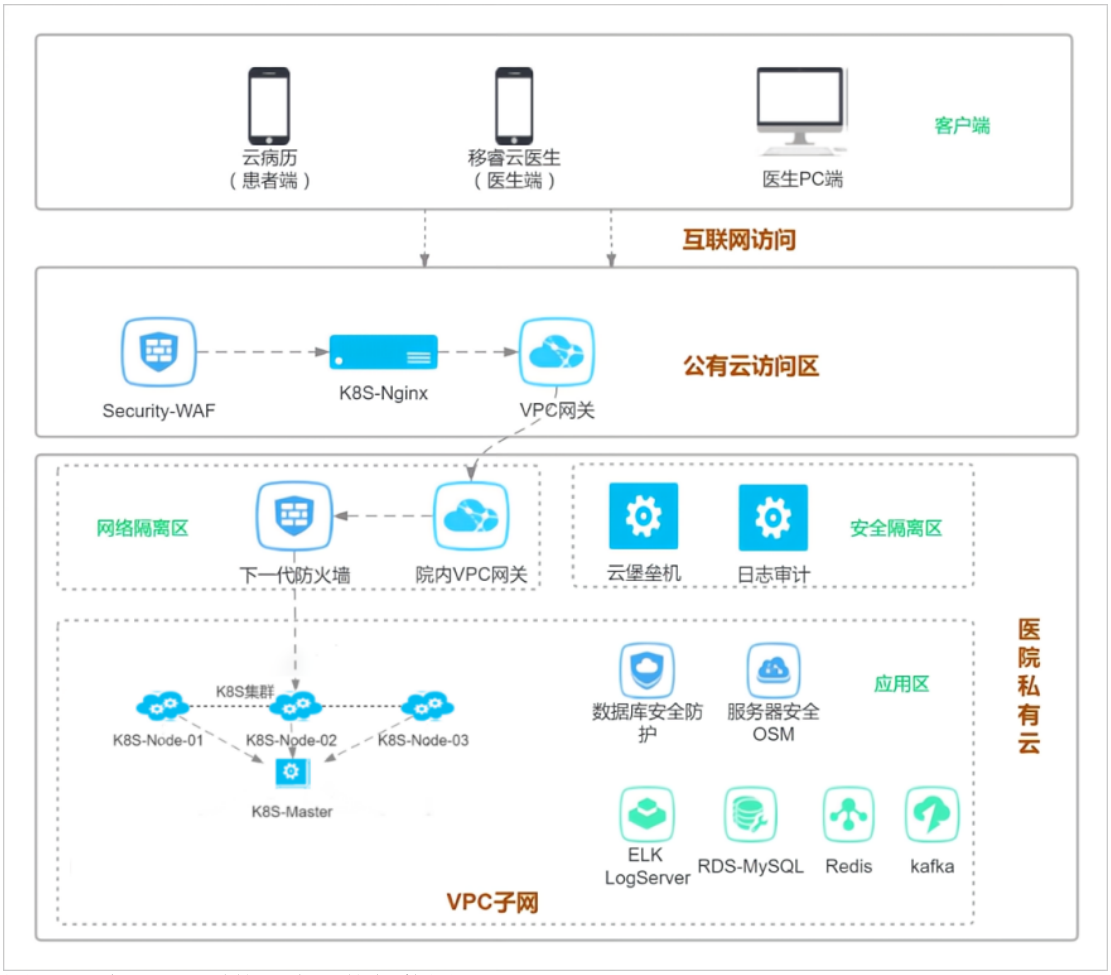


图 1 互联网医院系统平台网络拓扑图

3.3.2“海医互”平台技术架构

“海医互”平台基于统一的标准规范体系和安全保障体系，通过数据中心将各业务数据进行整合，再由互联网医院基础平台提供消息日志等服务，对接各项互联网医疗应用及海南省互联网医疗服务监管平台，如图 2 互联网医院系统平台技术架构图所示。

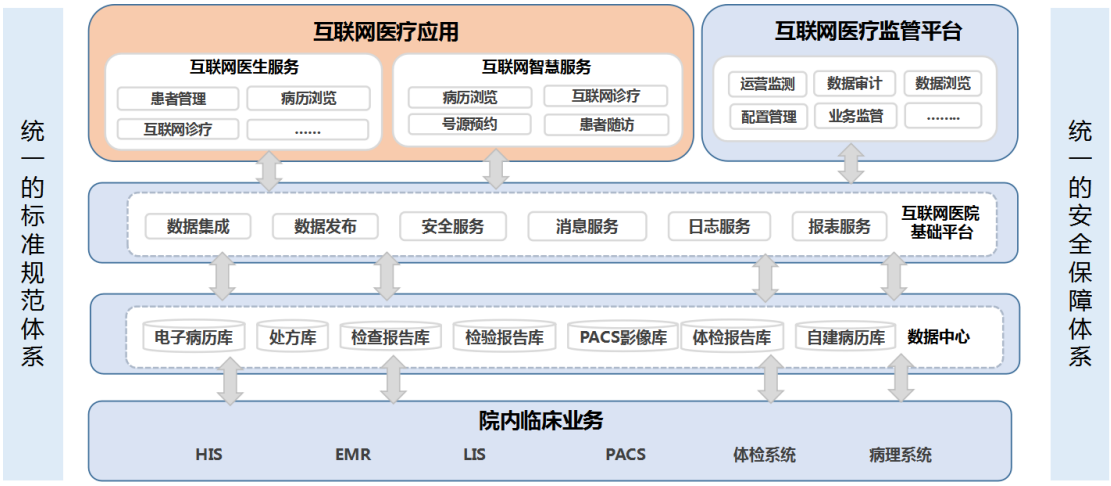


图 2 互联网医院系统平台技术架构图

4 结果

海南医学院第一附属医院互联网医院平台是实体医院互联网化的延伸和补充，在新冠肺炎疫情的持续影响下，线上诊疗优势突显。如图 3“海医互”患者端 APP 部分界面所示，通过移动互联网、实时音视频、在线支付、病历资料聚合等技术，在现有医疗机构设施基础上，向患者提供便捷就医、健康咨询、在线复诊、慢病配药、候诊提醒、用药指导、AI 智能分诊等一系列涵盖就医全流程的智慧服务体系。院内院外为一体、线上线下相融合，数据统一标准化、信息互联互通，优化了就诊服务流程，改善就医体验，提高患者满意度，真正实现便民惠民；同时提高了医务人员接诊效能，提升医院管理和运营水平。借助互联网医院平台，构建实体医院与患者之间线上沟通的桥梁，实现了复诊患者“线上、线下”一体化联动的全新就医模式。

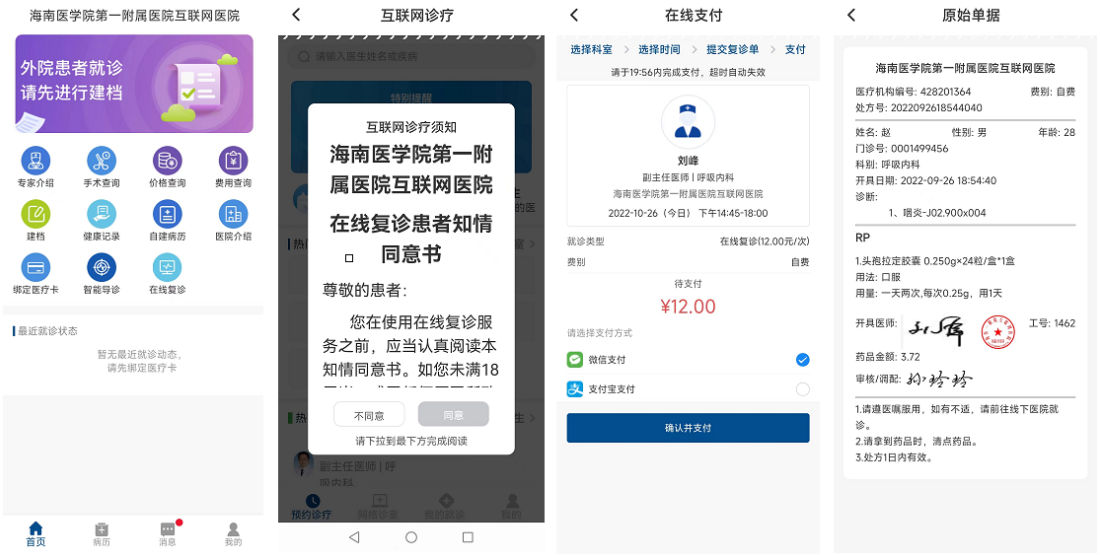


图 3 “海医互”患者端 APP 部分界面

5 讨论和思考

5.1 互联网医院建设成效

5.1.1 在线复诊，方便患者就医取药

首先，“海医互”平台支持手机 APP 和小程序两种就诊途径，实现了患者线上就诊流程的优化，足不出户就能完成复诊和取药，真正实现优势医疗资源下沉，普民惠民。另外，“海医互”平台已实现云病历，打通既往诊疗资料，患者实名认证绑定后，可查看本院近五年的所有门诊和住院记录，包括处方、病历、各项检验检查报告等就诊信息，能够加强患者的自我健康管理，提高患者使用平台的粘度。

5.1.2 在线问诊，提高医生接诊效率

“海医互”平台出诊医生可在 PC 端和移动 APP 端进行在线问诊和健康咨询，有效提高接诊效率。医生严格把握复诊患者准入，诊疗全程留痕可追溯，通过查看患者统一视图，可全面回顾患者在本院的就诊记录。患者在使用线上服务的过程中，也可主动提供院外健康数据和日常血压监测等指标。此外，在线问诊不受诊室场所限制，医生随时可以看诊，有序高效的就诊模式，提高了医生的接诊效率，提升了医疗服务水平。

5.1.3 节约成本，提升医院管理水平

海医一附院通过互联网医院平台的建设，加强了医院线上服务能力，减轻实体医院线下门诊压力，延伸了医院的诊疗服务半径。医生在移动端接诊全程无纸化，并减少占用诊室场地空间，节约了医院运营成本。此外，通过收集和展现诊疗相关数据，使医院管理层可以更为直观、快速地了解诊疗情况，进一步为医院配置医疗资源、组织开展诊疗研究，提供精确、可靠、实用的数据支持和

决策工具，从而提高医院的管理水平和运营水平。互联网医院平台后续实现医联体医疗数据的集成与交换、跨院转诊资源共享，能维护忠实患者关系，增强医院公信力和竞争力。

5.2 遗留问题思考和解决方法

5.2.1 医保线上结算政策有待进一步明晰

互联网医院是独立运营的线上医疗机构，需要申请成为医保定点医疗机构，才能进行医保在线结算。目前，海南省内的互联网医院仅支持纯自费结算，有待政策的进一步明晰。为推动互联网医疗、医改便民等政策执行，建议落实国家医保局《关于积极推进“互联网+”医疗服务医保支付工作的指导意见》（医保发〔2020〕45 号）要求，坚持线上线下一致，对线上线下医疗服务实行公平的医保政策，保持待遇水平均衡[6]。完善线上医疗机构成为医保定点医疗机构的审批流程。参保人在本统筹地区“互联网+”医疗服务定点医疗机构复诊，可以按照统筹地区医保规定支付[6]。

5.2.2 电子处方流转探索有待深入

目前互联网医院处方药品全部从院内药房发出，随着线上复诊人数的增多，药品配送工作量同步加大，配药环节给药房人员带来的压力与日俱增。建议探索定点医疗机构处方信息与定点零售药店互联互通，支持“互联网+”医疗复诊处方流转，患者可选择就近的基层卫生服务机构或定点零售药店进行药品配送[7]，节约运输时间，提高取药效率。

5.2.3 互联网医院运营推广需要加强

大部分患者仍遵循传统就医模式，到线下实体医院就诊，对于互联网医院线上复诊模式存在一定的观望心理，信任度不够。另外，在目前已制定的互联网医疗收费项目中，互联网诊疗收费较低，诊疗项目较为单一[8]。其三，互联网医院运营缺乏持续稳定的盈利模式，收付费标准体系需要进一步完善。建议在互联网医院平台构建的同时，统筹资源规划，加大宣传推广，在管理制度和绩效方面给予倾斜，鼓励医生在线问诊。结合实体医院的特点，发挥自身优势，总结学科特点和患者需求，分类制定运营方案，创造可持续的盈利模式[8]。例如：生殖医学科门诊大部分是纯自费项目，可引导患者在线复诊，缓解线下就诊压力。另外，建议相关主管部门改进完善互联网医院收付费标准体系[8]，推动互联网医院健康长效发展。

6 总结与展望

海医一附院通过互联网医院的建设,提高了实体医院的线上接诊能力，拓宽集团化医院的服务范围，医院内部信息化逐渐向外部延伸[9]。新冠疫情期间，互联网医院平台为患者居家就医提供了有力保障。“海医互”平台后续可与医生集团、医师协会、上级专家、医联体、医共体单位等建立远程会诊、远程教学指导等业务协同，形成了独具特色的“海医互”服务模式。

在互联网医院建设大潮中，医疗质量是诊疗服务管理的核心，互联网诊疗应把医疗质量放在首位，健全医疗质量责任保证体系，将质量管理纳入互联网诊疗的整个流程[10]。根据法律法规要求和互联网诊疗的工作特点，建立可行的互联网诊疗质量管理方案及评价、监控、改进体系[10]，推动质量管理不断完善。另外，随着一系列鼓励“互联网+医疗”的政策文件出台，应尽快落实医保线上结算，从门诊慢特病开始，逐步扩大医保对常见病、慢性病“互联网+”医疗服务支付的范围[6]。其三，互联网医院向纵深化、专业化发展成为必然趋势，逐步扩大接诊科室和病种[9]，让更多患者享受到便捷的医疗服务。

天下大同是中国传统文化智慧，医疗共同体也是当代构建人类命运共同体的重要组成部分。在互联网线上建立一个挂号机会平等、诊疗资源共享、医疗质量可靠、医保实时结算的大同平台，跨越地域的阻碍，让人民群众切实享受到“互联网+医疗健康”创新成果带来的实惠与便利[9]，值得广大医疗信息化人为之努力。

PU-031

数字医疗新风口 互联网医院发展按下“加速键”

谢琳艺 何慧敏* 翁梦晴 蒙嘉嘉
广西医科大学

目的 随着人工智能、区块链、大数据、云计算等数字技术的日益成熟，以及“健康中国”战略深入推进，中国互联网医院迈入创新发展期并已成为新兴的医疗服务业态。数字技术与医疗行业的深度结合催生数字医疗，数字医疗具有夯实医院“智慧医疗”建设、赋能医院高质量发展的医疗应用价值。

方法 本文以互联网医院为载体，深入分析互联网医疗业务场景、数字医疗领域延伸及智慧服务新业态、新模式，探讨数字医疗与互联网医院联合发展新趋势。

结果 以“数字疗法”为例，提出互联网医院作为数字医疗实践的重要载体，在疫情推动下，正逐步击碎传统医疗服务极限。数字医疗作为医疗实践的延伸，补齐传统医疗短板，与数字技术平台多边结合引入新型商业模式为数字医疗的发展注入新的活力。从基础诊疗入手，提出数字医疗连接各个角落，通过不断提高在线服务，释放医院物理空间，从而扩大医院接诊能力提高医疗资源利用率。切入医疗数据应用，数据挖掘、网络分析、图像可视化等数字技术技术促进非结构化数据和结构化数据的重建与开发，而数据的应用是保持互联网医院核心竞争力的本质。从医院业务探析，提出数字医疗技术也在加速构建互联网医院医教研防“四位一体”数字化协同体系建设。医疗技术为核心，数字技术为导向，临床科研教学为抓手，有效推动优质医疗资源纵向流动的同时提高医疗可及性。

结论 数字医疗的应用和落地夯实数字底座，拓宽互联网医院适用范围和应用范围，实现优质医疗共建共享。

PU-032

医院数字化转型的探索与实践

王建磊 杜建芳 张晓皓 孙媛
北京京都儿童医院

目的 医疗机构的发展目标在于提高医疗服务质量、管理效率以及患者满意度，同时提升医院的可持续发展能力、竞争力及品牌影响力。我们希望借助信息化进行数字化转型，更好的助力医院实现发展的目标。
首先医院数字化转型的核心是数据化，让数据说话。
其次，数字化技术可以提升医疗质量。
此外，数字化技术提高医院运营能力及效率，主动监控业务，调配资源。
方法 借助数字化技术，以患者为中心，以患者数据仓库做为信息基座，以“信息多跑用户少跑”做为指导思想，通过互联网的工具结合院内信息系统进行整合，与现有平台集成，并与已有的 HIS 、CIS、LIS、PACS 等系统连通，目标是汇集患者院前、院中、院后的所有数据，实现各业务流程闭环，形成全生命周期的健康档案。其中业务中台是高内聚、低耦合的面向业务领域的能力中心,提供丰富的共享服务，包含体系化建设医院能力域的方法和机制。是生产上层应用，包括配置、编排和扩展业务对象、业务能力、业务规则及业务流程，为上层应用系统的稳定运行提供了高并发、高可用的执行环境的医院资源运营管理的平台。数据核心以数据中台为中枢整合院内医疗数据和线上用户的行为数据（包括多源异构数据）提供数据产品、数据服务，包括实现数据接入、清洗，提供标准的数据接口，供数据订阅、数据消费，决策智能分析等，同时将就医过程中的挂、缴、查、候等业务线上化，通过 SCRM+CDP 实现服务与运营的协同。
结果 1、患者数据仓库：
2、就医流程消息闭环
3、门诊无纸化【2】：
4、特定场景的流程再造：
1）诊前检查
2）输液平台改进
3）候诊时长监控平台
5、在线咨询与复诊：
6、私域运营服务模式：
1)私域体系建设
2)用户画像和标签
3)素材库
4)社群&好友&专属健康管家
通过数字化，为医院服务品质、医疗流程制定了统一的管理标准，通过数字化的严格管理，保证各科室的规范行医，赢得了患者的信任和口碑。
结论 1、患者就诊时长改善
通过数字化的工具使用和对就医流程的不断优化，医院大流量科室“1 小时内完成就诊”的用户，大幅上升，由 60%上升至 80%。最大候诊时长也得到有效的改善，由最高的 255 分钟下降至 160 分钟。
2、患者满意度
患者满意度 NPS【3】，是衡量医院服务质量的最重要的维度，通过对流程的不断改善和更及时、贴心的服务，用户的满意度在持续提升。
3、服务效率的提升
通过企业微信、用户画像、素材库，大幅提升服务的效率，服务人员可服务的用户也大幅增加。
4、在线活跃用户提升
随着线上提供的服务越来越丰富，活跃用户也在不断的提升，用户停留的时长也越来越长（服务的多样性）。
5、医院品牌力的建设

我们的患者群体是谁？在哪里？我们要做到知己知彼，是需要以数据来衡量和优化医院的发展策略、品牌推广策略。通过就诊数据与地图的整合，可以了解医院覆盖区域的用户情况，以便更好的设计服务流程、义诊活动，为患者提供更优质的服务，提升患者对医院品牌的认知和信任。

PU-033

基于太赫兹时域光谱及数据库的中药鉴别现状浅析

彭树岩
重庆医科大学医学信息学院

响应国家中医药管理局所印发《“十四五”中医药信息化发展规划》的号召，夯实“十三五”时期以来中医药信息化发展建设基础，推动新时代背景下的中医药与信息科学的一体化建设，深化贯彻习近平总书记关于中医药工作的重要论述。中药作为我国自古以来所传承和发扬的文化瑰宝，具有维护中华传统文化源远流长发展现实意义的同时，饱含了理解古人智慧、促进中药应用到现代医疗药物工作的建设中的医学意义。在新时代医疗发展的大背景下，中医药的适应化发展，离不开信息化技术的驱动：推动中医药信息化建设的进一步提高，医工一体化建设水平不断上升，解决中医药信息化发展建设工作中存在的发展不均衡、不协调、浅层次的主要矛盾与未能完成数字化中国医疗、中医药创新继承发展、全面推动居民健康信息化建设的目标建设的现实问题。从中药鉴别的视角出发，介绍新医科背景下信息化技术为中药鉴别工作赋能新状况，以太赫兹时域光谱技术（THz-TDS）对多种中药药物的识别鉴定应用为论述对象，就目前蓬勃发展的中药鉴别行业而言，通过阐释传统物理鉴定方法和信息化技术有机化结合实例，浅析新医科背景下中医药鉴别路径，为相关行业提供新标准、新思路，以期未来建设呈现出高精度、高效率、多元化的行业发展状况，为推动中医药文化和技术现代化、为人民医疗健康事业多次元化发展提供有力支持。

PU-034

基于太赫兹时域光谱及数据库的中药鉴别现状浅析

彭树岩
重庆医科大学医学信息学院

1 引言

2022 年 12 月国家中医药管理局正式发布《“十四五”中医药信息化发展规划》通知文件，赓续了“十三五”时期的中医药信息化发展的薪火，中医药走向现代化、信息化的进程逐步发展和完善。伴随着国家政策有序地推进，中药鉴别方法呈现出百家争鸣的盛态，在新医科的背景下积极向着中医药数字化建设的目标靠拢。

目前行业中流行的中药鉴定方法：显微切面鉴别[1]、热分析技术与红外分光光度法[2]。相较于其他方法，受到了鉴别准确度低、前期化学预处理繁杂、鉴别效率不高等难题的限制，太赫兹光谱技术有着灵敏度强、效率高、操作简便、无损性检测等优势。在中药鉴别的领域，太赫兹光谱技术具有广泛的研究空间和应用价值。本文首先简述太赫兹时域光谱技术，接着介绍了新医科背景下太赫兹光谱技术等物理方法如何实现与信息化技术的联结，最后提出了总结和希望。

2 中药鉴别路径介绍

随着我国社会经济与科研实力的飞跃式发展，中医药不再是隐于市井的家族独学，中医药逐渐走到了新时代新医科发展的道路上，中药鉴别作为中医药传承发扬的关键，传统的望、问、触的方式已然满足不了日渐增长的鉴别需求。为保障中医药产业可持续发展，广大研究者积极投入相关研究之中。经历了长期的实践研究，紧跟着激光深入研究的热潮，太赫兹光谱技术逐渐脱颖而出，饱受众多中药鉴定行业从事者的青睐。

2.1 太赫兹时域光谱技术简诉

太赫兹时域光谱技术是一种新兴的、行之有效的相干探测技术，太赫兹时域光谱技术的研究开始于 20 世纪 80 年代，跟随着激光技术跨越性的发展提升的脚步，稳定、强度适宜的太赫兹光源得以诞生[3-4]。太赫兹光源主要应用太赫兹波（一种介于微波与红外辐射之间，频率为 0.1~10THz 的电磁辐射）。太赫兹光具备优良穿透性能的同时，兼顾了光子能量低所具有的鉴别安全特性与梯度分区精细灵敏的指纹特性。凭借以上特性优势，太赫兹时域光谱技术可以广泛应用到物质真伪鉴别与成分检验分析等生物、医药、材料领域中。时至今日，太赫兹时域光谱技术也逐步走出了实验室，进一步达到商业化应用状态。

2.2 太赫兹时域光谱在中药鉴别中的应用

中药作为我国传承五千年的文化瑰宝，蕴含着深厚的文化意义，由于本身存在成分繁多复杂、各成分振动吸收的波段跳跃性较大等难题，很难进行系统化的检测。大量的实验研究结果表明：太赫兹光谱波段的分子内振动模式与中药代谢产物中晶格内低频振动模式具有颇高的重合度[5]。因此通过采用太赫兹时域光谱技术来进行常温常态下的中药结果、形态的鉴别工作具有可观的应用价值和前景。

国内运用太赫兹时域光谱技术开展了多种中药的鉴别检测，获得了多方面相关实验研究数据和成果。

在真伪鉴别方面，李凯瑞等[6]的研究中，采用太赫兹时域光谱技术获得了巴戟天，粉防己，猪苓三种中药药物及其伪品的太赫兹光谱图像，反映了太赫兹时域光谱技术针对中药鉴别具有可行的科学意义；杨玉平[7]等的研究中，应用太赫兹光谱技术测量藏红花、牛黄及其伪品，实现了高准确性的鉴别。

在产地鉴别方面，华南理工大学的学者余静孝等[8]的研究中，建立基于太赫兹光谱技术的陈皮鉴别模型，结果呈现出 100%准确率的鉴别效果；魏枭[9]等的研究中，捕捉各产地大豆太赫兹光谱差异，结合化学计量学等方法优化，实现了国内外不同产地大豆样品的鉴别。

在品种鉴别方面，王芳等[10]的研究中，针对 14 种外表类似的不同苜蓿牧草品种，利用太赫兹时域光谱技术结合多元统计学方法，得以快速识别不同苜蓿牧草差异；刘燕德[11]等研究中，以四种行业中常见的贝母品种为对象，利用太赫兹时域光谱系统收集和比较各样品谱图，表面太赫兹光谱技术在针对同类中药不同品种的鉴别具有普适性。

基于以上研究，研究者从不同种类的中药材真伪品类鉴别、中药材的主要成分分析和主要形态识别、中药材的特征性光谱数据收集处理等层面，证明了太赫兹时域光谱技术在中药材的鉴别工作中具有普遍性应用的实际意义。目前，大多数的中药材进行太赫兹时域光谱技术鉴别的方式具有简洁、高效的特点，基本的操作为：原材料粗处理，实现去除大颗粒杂质；粗处理后的样品进行粉末化操作后进行干燥处理并装袋；粉末化样品进行压片操作，样品切片尺寸控制在 0.5-1.5mm 的厚度、直径在 1cm 左右；搭建机器视觉平台（动态和静态）。其中动态平台主要以传送装置为核心，将样品切片传动至检测机器并进行相机记录，静态平台以可以进行连接的计算机为核心，便于收集和处理检测后得到的光谱信息；光源射出稳定的太赫兹光，检测机器进行特征性光谱区段提取后，将成像出的光谱粗图像传送至连接的计算机中；导入光谱粗图像采用对应的图像处理模型，通过提前设计好的算法实现进一步的特征化图像区段锁定与精细化，并将不同样品产生图像进行分类；最后对比不同样品谱图的波峰、波谷等特征性点位，实现不同产品的区分和鉴别。基本的流程详见图 1。

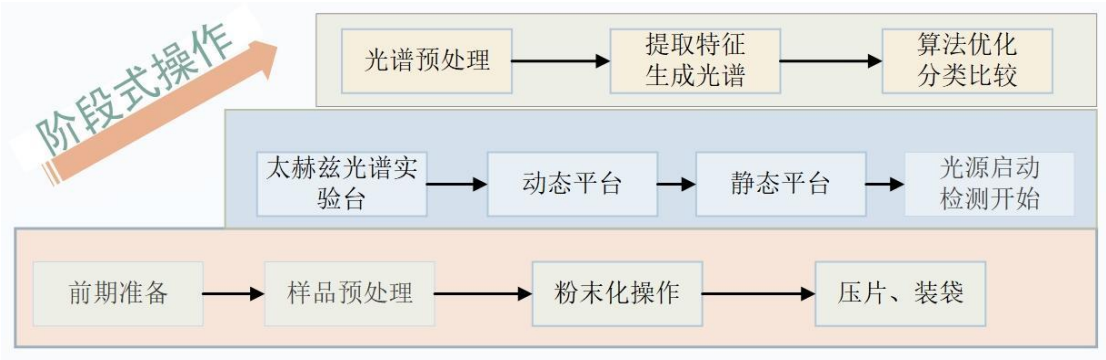


图 1 太赫兹时域光谱技术中药鉴别流程

2.3 新医科背景下的中药鉴别的研究进展

我国进入全面建设社会主义现代化国家新征程阶段，“十四五”时期作为国家信息化、数字化引导中医药高质量、高水平发展的井喷时期，“互联网＋中医药”的融合发展思维具有了进一步的贯彻。中药鉴定的方法也不只是局限在物理层面，口口相传的中医药知识传播方式逐渐淡出主流，以大数据、云计算为代表的信息化力量，赋能中药鉴定的智能化发展，为中药鉴定铺就坚定的技术基石。

响应国家《“十四五”中医药信息化发展规划》文件的号召，推进中医药数据资源创新应用，大量学者积极投入中医药数据的整理与应用工作。

在质谱数据库方面，刘潇潇[12]等收集多种岭南地区有毒的中草药进行统一化鉴定，并建立对应精准识别的质谱数据库，实现了在数据库中高效、准确的进行相应中草药的鉴别工作。中南大学的学者负栋[13]等,通过进行中药质谱库深度数据挖掘，机器学习方法智能化数据修正和优化，达到了有效进行不同中药药品特征化学信息提取，完善了中药质谱库检索匹配机制。

在指纹图谱库方面，华中科技大学的学者杨灏[14]等，利用太赫兹时域光谱技术通过选取多种药品进行测量，获得每一种药品独有的指纹图谱，并进行收集和整理工作，建立相关药品的指纹图谱数据库，从而实现通过对比不同药品指纹图谱信息实现特异化鉴别；高梅[15]等研究基于太赫兹时域光谱技术建立的中药指纹图谱数据库，分析总结太赫兹光谱技术在中药复杂成分鉴别和质量评估方面的显著的优越性。

在数据库方面，河南中医药大学的学者鲁轮[16]等基于 2020 年版的《中国药典》，通过提取、归纳、分类中药关键信息，进行了中药数据库的开发，实现了功能涵盖广、实用价值高、操作程序简易的中药数据库，为中医药行业人员提供了数字化检索平台；天津大学的学者郭帅[17]等利用太赫兹光谱技术进行牛膝等 10 余组中药的鉴别和谱图收集，采用 Matlab 软件进行数据的加载与优化，建立中药数据库及其管理系统，充实了系统化中药数据库的研究开发。

基于以物理鉴定手段和中药数据库开发的中药鉴别新思路，满足了《“十四五”中医药信息化发展规划》文件中实现中药鉴别质量高水平、高效率发展的议题，积极把握好“十三五”时期打下来夯实的基础条件优势，进一步发展和完善中医药信息化的建设，让新医科的理论不只是局限于西医药。中医药在推陈出新的发展进程中，贯彻好新时代新医科的发展理念，抓紧抢牢信息革命的发展机遇，乘着信息化发展的春风，实现新医科时代背景下高水平、高质量发展，激发中药鉴定乃至整个中医药行业的活力，完成自身的创新性进步，不落后于外来的西医药技术，深刻贯彻健康中国重要战略，为全体中国人民的健康事业提供坚实支撑。

3 结语

本研究系统化的介绍太赫兹光谱技术在中药鉴别领域广泛应用，展望了其在新医科背景下与互联网技术融合的巨大潜力。通过太赫兹光谱技术结合大数据等信息化技术的方式，从药品鉴定筛查角度实现了行业一体化提高升级。“十四五”时期正是中医药产业进行数据资源整合，实现建立中医药信息化发展多元化投入机制的关键阶段。中医药鉴别将传统物理方法与互联网技术结合的行为模式，焕发了自身行业的全新面貌，产生了良好的示范和创新意义。为实现全面小康目标下的人民健康一体化进程给予了可靠的助力。

PU-035

从政策方向看关键技术与“互联网+医疗健康”新业态的形成

王叶 伍丽群 曾华堂
深圳市卫生健康发展研究和数据管理中心

1 引言

“互联网+医疗健康”概念强调以互联网为载体，以信息技术为手段，与传统医疗健康领域的深度融合，是以移动通信、物联网、云计算和大数据等信息通信技术为手段，与传统医疗健康服务深度融合而形成的一种新型医疗健康服务业态的总称[1]。《国务院关于积极推进“互联网+”行动的指导意见》（国发〔2015〕40 号）提出，顺应世界“互联网+”发展趋势，充分发挥我国互联网的规模优势和应用优势，推动互联网由消费领域向生产领域拓展。因此，为支撑并完善“互联网+医疗健康”的发展，国务院办公厅于 2018 年出台《关于促进“互联网+医疗健康”发展的意见》，进一步将“互联网+医疗健康”纳入国家重要发展战略。此后，全国各级政府相继出台“互联网+医疗健康”相关指导意见及配套政策，全面促进各地医疗卫生服务体系的完善和“互联网+医疗健康”产业发展。“互联网+”与医疗服务行业的深度融合形成转变了传统服务模式，不断催生出新的应用场景。

习近平总书记在中央全面深化改革委员会第十四次会议上指出，要高度重视新一代信息技术在医药卫生领域的应用，以此重塑医药卫生管理和服务模式，优化资源配置，提升服务效率。《国务院办公厅关于印发全国医疗卫生服务体系规划纲要（2015—2020 年）的通知》（国办发〔2015〕14 号）也强调，云计算、物联网、移动互联网、大数据等信息化技术的快速发展，为优化医疗卫生业务流程、提高服务效率提供了条件，必将推动医疗卫生服务模式和管理模式的深刻转变。在国家系列政策促进下，移动通信、物联网、大数据、云计算、区块链等关键技术逐渐在医疗健康服务领域融合与应用，催生出“互联网+医疗服务”、“互联网+护理服务”、“互联网+药品供应保障”等医疗服务体系新业态。

2 现有医疗服务体系的发展限制

随着经济社会发展与人民生活水平的提高，传统医疗服务模式不能满足人民群众多样化、差异化的医疗健康服务需求。医改进程的不断深化解决了部分传统医疗服务模式缺陷，但现有医疗服务体系仍存在一定的发展限制。

首先，优质医疗资源分配不均，就医观念仍需进一步转变。传统的医疗服务模式导致优质医疗资源过度集中于大城市，边远、贫困地区对于医疗服务的可及不足。患者及医疗机构对就医观念转变较慢。相对于互联网等信息技术企业，大部分医疗机构对互联网等新兴事物开放和接受程度迟缓。此外，人工智能等新兴技术尚未发展成熟，未能解决医疗卫生健康领域更多实际问题和需求，因此居民对传统就医观念的转变效果不佳。互联网医疗有利于促进医疗资源的公平合理配置，但是医生及患者端访问不便捷、接受度不高、使用率低、未在基层医疗机构广泛应用等问题对其在分级诊疗和远程医疗上的作用未及预期[2]。

其次，医疗服务体系与公共卫生服务体系长期割裂。目前，我国医疗服务体系与公共卫生服务体系长期处于割裂状态，体制机制不健全、资源要素配置不均导致一系列问题。首先，重治轻防仍是我国医疗卫生体系的现状。预防是最经济有效的健康策略，且预防投入产出效益比治疗投入产出效益更佳[3]。但是我国疾病预防控制与救治体系发展不协调，预防与医疗之间在理论和应用方面缺乏沟通和交叉。第二，医疗机构与疾控部门与人员之间缺乏紧密合作沟通，信息无法互联互通，其

中管理条块分割、服务协作不足“碎片化”问题突出，由此导致居民健康问题无法做到早期的健康风险研判和识别管理。第三，我国居民慢病死亡已占到总死亡的 88%，疾病负担占我国总疾病负担的 70%以上，疾病负担向慢病转变。但当前门诊医生和公卫医师对慢病患者仍处于“分段管理”模式，此种模式难以适应慢病防控形势和群众健康服务需求，对我国提供“以居民健康为中心”全生命周期、综合连续的闭环服务带来一定挑战。

第三，信息互联互通仍是医疗行业所面临的关键问题。一体化的分级诊疗体系的建立及医联体建设，都需要强大的信息互联互通能力。不同类别、级别医疗机构之间存在信息壁垒，医疗信息数据共享困难等问题成为“互联网+医疗健康”发展的阻碍[4]。现有互联网医院由于依托于不同的实体医疗机构，彼此之间的信息壁垒并未打通，本质上并没有解决信息互联互通的问题。只有借助于大数据等关键技术打通医疗健康数据，才能最大限度的整合多来源数据, 挖掘多应用场景, 培育多业态模式, 全面提升人口健康科学决策和服务管理水平。[5]

3 政策视角下关键技术在“互联网+医疗健康”的应用方向

3.1 政策视角下关键技术在“互联网+医疗健康”的应用方向

2018 年《国务院办公厅关于促进“互联网+医疗健康”发展的意见》发布以来，国务院办公厅、国家卫健委等国家级政府部门共发布 15 个提及互联网、大数据等关键技术的“互联网+医疗健康”相关文件，以此推动互联网诊疗、互联网医院、远程医疗服务以及预约诊疗等互联网医疗服务快速健康发展。对这 15 个文件（附表）进行主题词分析发现，2018 年至今，关键技术在“互联网+医疗健康”中的主要应用方向为医疗服务和医疗机构管理，占据所有引导方向的一半以上（75/109）。其次分别是药品供应保障、中医领域、公共卫生、分级诊疗、医学科普教育、医联体建设等其他医疗健康领域。因此，从政策引导层面，短期内关键技术在“互联网+医疗健康”的应用仍集中于医疗服务和医疗机构管理。

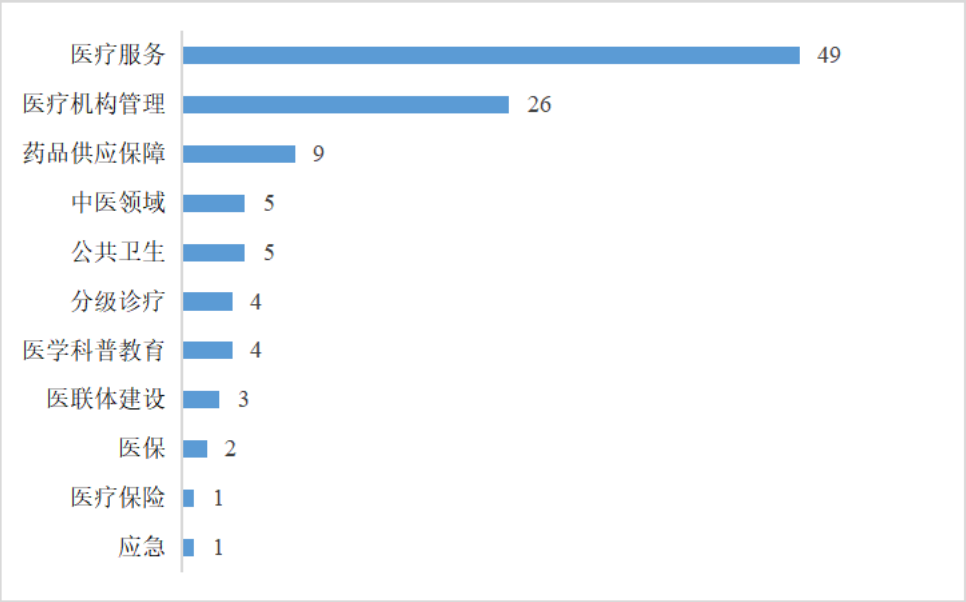


图 1 2018 年至今 15 个国家级政策文件中主题词分类分析

3.2 新一代信息技术的发展催生新的需求

由于传统医疗卫生体制的限制及人工智能等新技术尚未成熟等因素，“互联网+医疗健康的发展”仍处于初步阶段。但随着经济社会发展和生活水平提高，人民群众更加重视生命安全和健康质量，人民群众多样化、差异化的健康需求不断增长。同时，针对传统医疗服务模式不能解决的痛点问题，人工智能、大数据、云计算等关键技术在医疗健康领域也催生出新的应用场景。例如，上海探索利用大数据及物联网技术探索实施医用设备的智能化管理，通过建设“市级医院医用设备全生命周期管理智能数据平台信息系统”，对医用设备全生命周期管理的智能分析与评价，引导医院精准规划大型医用设备配置，实现提质增效。同时，互联网巨头企业、医疗信息化企业和传统医药企业也探索将不同关键技术的特点与医疗服务体系中的痛点相结合，不断发现新的应用场景与催生出新的需求。

4 多种技术融合构建多元化“互联网+医疗健康”新业态

随着关键技术与医疗科技的发展，互联网医疗服务与产业融合上存在广阔的发展空间。互联网医疗完整产业链应包括互联网医院与互联网诊疗服务、互联网器械可穿戴设备等、医药电商与药品流转等，但在企业独立设置的互联网医院、互联网器械可穿戴设备、医药电商与药品流转的方面发展仍相对滞后，互联网医疗产业链尚未形成。目前，浙江杭州借助于其互联网产业聚集优势和“互联网+”元素的聚合效应，已基本完成“互联网+医疗健康”产业的规划布局[6]。基于经济强市的示范引领作用，北京、上海、深圳等地拥有较强的互联网产业配套资源，可以探索建立聚集医药健康全产业链资源、线上线下联动、高效运转的互联网医疗体系，充分满足市民就医需求。

同时，借助于互联网+、人工智能等关键技术，可以实现多种场景下建立新的医疗服务体系：一是将“互联网+医疗服务”进一步拓展至基层医疗机构和社区，技术和政策双重引导医生和患者共同共建多元化互联网医疗服务体系。二是将人工智能、大数据等技术深度融入医学科普教育，充分聚集互联网资源助力患者健康教育和医学生规范化培训。三是开展互联网+家庭医生签约、AI 随访等实践，释放大型医疗机构医护人力，使其更好投入复杂疾病的诊治中。四是借助互联网和大数据平台整合区域资源，推动医联体建设，促进分级诊疗体系的建设。五是发挥区块链优势，保障医疗大数据和健康信息安全。六是利用物联网技术发展进一步重构医疗健康服务模式。

附表 2018 年-至今国家“互联网+医疗健康”相关文件及关键技术应用方向

颁布时间	发文单位	政策名称	涉及新一代信息技术	
2023 年 6 月 15 日	国家卫生健康委、国家中医药局	《关于印发进一步改善护理服务行动计划（2023—2025 年）的通知》	互联网、人工智能、5G、物联网等新一代信息技术	● 护士培训、临床书写 ● 延续性服务 ● 远程护理会诊 ● 优化护理服务流程，提高护理工作效率，减轻护士工作负荷
2023 年 5 月 23 日	国家卫生健康委、国家中医药局	《关于开展改善就医感受提升患者体验主题活动的通知》	互联网、人工智能	● 预就诊模式建立 ● 处方流转、药品配送等 ● 改善就医体验，为临床诊疗服务提供高质量辅助，提升医疗服务质量和效率
2022 年 2 月 8 日	国家卫健委办公厅、国家中医药局办公室	《关于印发互联网诊疗监管细则（试行）的通知》	人工智能、大数据等新兴技术	● 互联网诊疗合理性的分析和监管
2020 年 12 月 4 日	国家卫生健康委、国家医疗保障局、国家中医药管理局	《关于深入推进“互联网+医疗健康”“五个一”服务行动的通知》	互联网、大数据、人工智能等信息技术、智能物联及其终端设备	● 拓展服务空间和内容 ● 慢性病患者和高危人群的特征指标数据的监测跟踪和管理 ● 赋能流调 ● 下沉专家资源至基层医疗机构
2020 年 10 月 24 日	国家医疗保障局	《国家医疗保障局关于积极推进“互联网+”医疗服务医保支付工作的指导意见》	大数据、互联网等技术手段	实时监管“互联网+”医疗服务费用结算明细、药品、耗材、医疗服务项目和门诊病历等信息
2020 年 10 月 12 日	规划发展与信息化司	《关于加强全民健康信息标准化体系建设的意见》	互联网、大数据、人工智能、区块链、5G 等新兴信息技术	● 互联网医院、远程医疗等应用标准化建设 ● 智能临床辅助诊疗、医用机器人、人工智能药物研发、智能公共卫生服务、智能医院管理、智能医疗设备管理、智能医学教育等 ● 应急救治、远程会诊、远程手术示教、远程超声检查、远程内窥镜检查、移动重症监护、移动医疗设备管理、移动医护、智慧养老、远程机器人手术等 ● 医疗联合体、个人健康档案、电子处方、药品管理、医疗保险、智慧医院管理、疫苗管理、基因测序等建设

2020年5月21日	国家卫生健康委办公厅	《国家卫生健康委办公厅关于进一步完善预约诊疗制度加强智慧医院建设的通知》	互联网、物联网等信息技术	● 下沉医疗资源、方便群众就近就医 ● 部分常见病、慢性病复诊、药品配送等服务， ● 医院内部信息系统的互联互通、实时监管
2020年04月07日	国家发展改革委中央网信办	《关于推进“上云用数赋智”行动培育新经济发展实施方案》的通知	大数据、人工智能、云计算、数字孪生、5G、物联网和区块链等	● 互联网医疗医保首诊制和预约分诊制、互联网医疗的医保结算、支付标准、药品网售、分级诊疗、远程会诊、多点执业、家庭医生、线上生态圈接诊等
2020年03月02日	民政部办公厅	《新冠肺炎疫情社区防控工作信息化建设和应用指引》	互联网、大数据、人工智能等信息技术	● 社区疫情监测、信息报送、宣传教育、环境整治、困难帮扶等
2020年02月08日	国家中医药管理局办公室	《关于加强信息化支撑新型冠状病毒肺炎疫情中医药防控工作的通知》	互联网技术、移动终端、移动互联网	● 在线开展部分常见病、慢性病复诊、中医远程会诊、远程查房等 ● 优化就诊服务流程：智能分诊导诊、智能候诊、在线支付、药品配送、结果查询、信息推送等智慧健康服务、优化电子病历系统功能 ● 搭建社区医护人员与管理对象信息互动平台：在线用药指导、咨询、复诊患者方药随症调整以及疗效随访等服务 ● 网上义务咨询、居家医学观察指导、用药咨询指导和心理健康疏导等
2020年02月03日	国家卫生健康委办公厅	《关于加强信息化支撑新型冠状病毒感染的肺炎疫情防控工作的通知》	大数据、5G	● 对疫情发展进行实时跟踪、重点筛查、有效预测 ● 网上义务咨询、居家医学观察指导等 ● 提高定点救治医院网络稳定性和传输质量
2019年9月	国家发展改革委	《促进健康产业高质量发展行动纲要（2019-2022年）》	互联网、物联网、云计算、大数据、移动互联网、人工智能	● 优化医院服务流程：分时段预约诊疗、区域内检验检查结果互认、智能导医分诊、免（少）排队候诊和取药、移动端支付结算、检查结果自动推送、智慧中药房等 ● 药品流通：简化流通层次，优化流通网络，提高供求信息对称度和透明度 ● 医学影像辅助判读、临床辅助诊断、多维医疗

				数据分析等 ● 医疗资源薄弱地区、基层医疗机构提高诊疗质量，促进分级诊疗
2019 年 2 月	国家卫生健康委员会	《关于开展“互联网+护理服务”试点工作的通知》	人脸识别等人体特征识别技术	● 加强护士管理
2018 年 7 月 17 日	国家卫生健康委	《互联网医疗诊疗管理办法》	人脸识别等人体特征识别技术、互联网	● 加强护士管理 ● 医联体建设，向下转诊患者
2018 年 4 月 25 日	国务院办公厅	《国务院办公厅关于 促进“互联网+医疗健康”发展的意见》	互联网、人工智能，人群流动、气候变化等大数据技术	● 部分常见病、慢性病复诊 ● 远程会诊、远程心电诊断、远程影像诊断等，检查检验结果实时查阅、互认共享 ● 预测疾病流行趋势，传染病等疾病的智能监测 ● 研发临床诊疗决策支持系统 ● 提升基层中医诊疗服务能力 ● 个人健康实时监测与评估、疾病预警、慢病筛查、主动干预 ● 促进实现分级诊疗和各级各类医院信息共享

PU-036

医疗卫生领域商用密码安全评估与建设的实践与思考

刘阳[1] 康亚西[2] 冯标斌[1]
1 中南大学湘雅医学院附属海口医院（海口市人民医院）信息管理处
2 海南省人民医院（海南医学院附属海南医院）信息工程部

0 引言

近年来，在《网络安全法》颁布施行之后，《密码法》《数据安全法》《个人隐私保护法》相继出台。密码法承前启后，在先后出台的一系列法律所形成的多元法益保护下，商用密码安全建设不仅确保了信息系统的网路安全和数据安全，更保障着我们每一个人的隐私安全[1]。商用密码已经在政府、通信、能源、金融等重要领域得到了一定的应用[2]。但是，在医疗卫生这个关乎人民生命健康的领域的部分地区，依然存在无法有效落实密码评估制度，造成密评和商用密码建设无法循环迭代的问题，导致医疗业务系统无法落实持续改进，致使网络安全的最后一道防线形同虚设，无法应对网络安全新威胁[3]。因此，在医疗卫生领域普及和推广商用密码迫在眉睫[4]。通过本文的说明，让读者充分了解密码评估和建设的作用，躲坑避雷、少走弯路。

1 概念和含义

《中华人民共和国密码法》是我们国家密码安全工作所遵循的根本法律。定义了密码的概念是包括密码算法、密码技术、密码产品、密码服务，可以采用特定的变换方法对信息和数据进行加密保护、安全认证，实现身份真实性、数据机密性、信息完整性、操作不可否认性。国家对于密码施行分类管理，分为核心密码、普通密码、商用密码三类。其中商用密码用于保护不属于国家秘密的公民、法人和组织的信息[5]。《信息系安全技术 信息系统密码应用基本要求》GB/T39786-2021[6]等标准规范，明确了同步规划、同步建设、同步运行、定期评估的商用密码建设要求[7]。商用密码应用安全性评估即密评，是指在采用商用密码技术、服务集成建设的网络和信息系统中，对密码应用的合规性、正确性、有效性进行评估。由此可知，商用密码，是网络安全防护的最后一道防线，起到保底作用；密码评估意在找出商用密码建设存在的漏洞，持续增强网络安全防护水平。

2 密码安全评估与建设的原则与方法

商用密码在智慧医院建设中，是不能孤立于临床业务应用单独存在的，要想密码应用起到成效，我们结合在电子病历系统应用评级第五级的要求中明确的患者病历安全要求，必须以 CA 电子签名的应用来保障医院的 HIS、EMR、LIS、PACS 等相关业务应用系统的数据安全，才能让法律法规在保障医疗过程安全上形成多元法益保护[8]。医疗行业智慧医院建设的评价标准体系中包含智慧医疗、智慧服务、智慧管理三个组成部分，智慧医院“以评促建、以评促改、以评促用”也同样是保障密码安全所需要遵循的基本原则。以外“三同步一评估”的原则，也是贯穿医院信息系化建设全生命周期重要保障[9]，目的就是为了促进密码安全落到实处。

2.1 同步规划先决条件

信息系统的规划，首先要根据网络安全保护的等级，依据《信息系安全技术 信息系统密码应用基本要求》GB/T39786-2021[6]中的相应等级要求，结合业务应用系统的安全需求进行规划。如果规划的系统还没有进行网络安全保护定级，其密码应用的等级至少应参照三级系统的标准。

密码应用规划设计方案包含密码应用解决方案、实施方案、应急处置方案三个内容。其一，解决方案主要包含系统现状分析、系统风险分析、密码应用需求、密码技术方案、密码建设方案、密码管理方案、密码产品清单等内容，是设计阶段的主要方案。其二，实施方案包含项目概述、项目组织及人员、实施内容、实施计划、部署方案及部署图、保障措施等内容，是系统部署阶段必须明确的方案。其三，应急处置方案包含密码安全事件分类分级、应急处置组织机构、应急处置流程等内容，是在系统的生命周期中的运维阶段的应急预案[10]。在医疗机构特别是大型三甲医院中，围绕以电子病历系统为核心的智慧医疗建设的业务系统繁多、结构复杂，要想建设的好，规划设计就非常重要。此外，密码方案编制完成后需要进行密码评估和专家论证，2 种评价的方式，只有通过评估的密码应用方案才能够作为密码建设的依据[11]。在医院实施方案的重要性也毋庸置疑，应急处置方案更是重中之重，需要根据不同业务场景，如：门诊开药闭环管理、门诊检验检查闭环管理、住院手术闭环管理等相关管理流程，在系统出现应急故障的时候，按照应急处置方案，根据故障影响分为的分级采取不同的应急方案执行。医院的网络安全工作需要贯彻“统筹规划，分步实施，循序渐进，持续改进”的理念，因此，同步规划是先决条件[12]。

2.2 同步建设是基础

在制定好通过方案评估的密码规划设计方案后，密码应用建设从密码技术和密码管理两个方面展开，这两个方面的建设离不开密码基础的底座支撑，实现密码产品的部署。详见图 1 商用密码应用建设体系图。

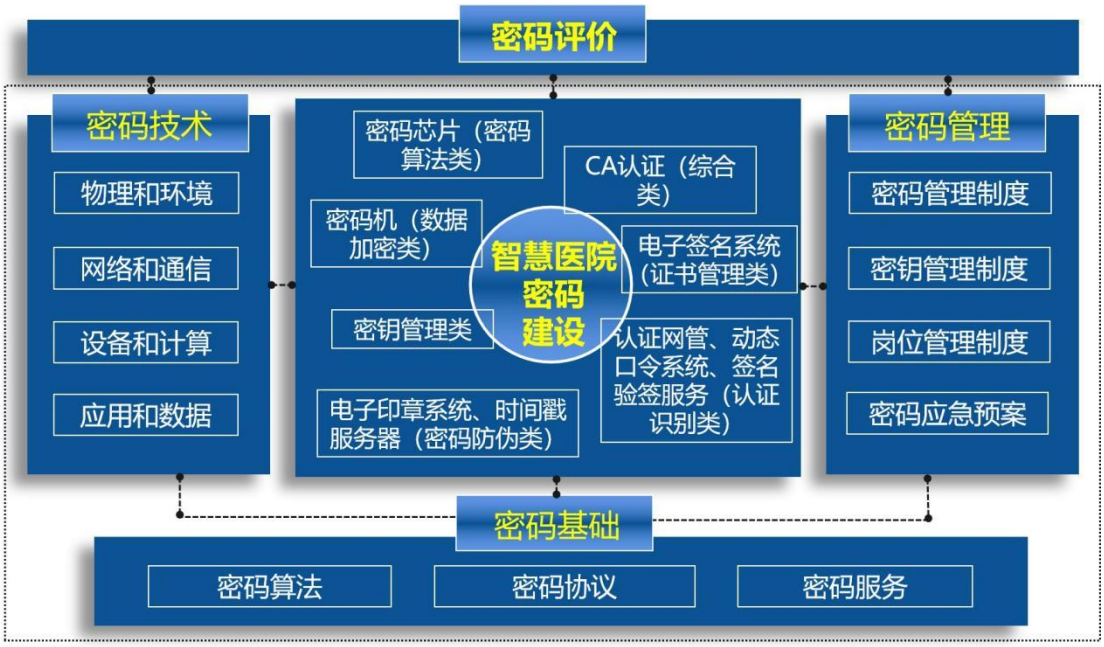


图 1 商用密码应用建设体系图

2.2.1 密码基础

密码基础建设方面，医疗业务信息系统需要使用符合法律、法规以及相关国家标准，结合与医疗行业标准有关要求的密码算法、密码协议、密码服务，加强密码技术应用和密码管理统筹。信息系统应使用 SM2、SM9 等高等级的商用密码算法[13]，应使用 TLS2.0、TLS3.0 等高版本的密码协议，应使用通过商用密码产品检测机构认证的密码产品，应使用具备相关资质的密码服务机构，以避免密码算法、密码协议、密码服务可能存在的安全风险[14]。

2.2.2 密码技术

密码技术建设方面，信息系统需要在物理和环境、网络和通信、设备和计算、应用和数据四个层面均采用密码技术进行防护[15]。首先，在物理和环境层面，物理环境访问人员的身份真实性可以由电子门禁系统或者视频监控系统的存储记录来实现。第二，在网络和通信层面，网络空间访问人员的身份真实性可以由网络接入认证系统、IPSec VPN、SSL VPN 等认证系统，来保证信息在网络中传输的完整性和机密性。第三，在设备和计算层面，主机访问人员的身份真实性可以由智能密钥、服务器密码机、安全浏览器等密码技术，来加强设备身份、访问控制、关基设备的安全性。第四，在应用和数据层面，用户登录系统的身份真实性可以由 CA 认证的签名验签服务器、数据库加密机、智能密钥等密码技术，来系统运行的稳定性和数据存储的安全性[16]。

2.2.3 密码管理

密码管理建设方案，信息系统需要从管理制度、人员管控、运行管控、应急处置等方面对密码的使用和管理进行规范。其一，制定密码相关的管理制度，包括密码人员管理制度、密钥管理制度、项目管理制度、密码安全培训制度等。其二，制定密钥管理制度，包括密钥产生、分发、存储、使用、更新、归档、撤销、备份、恢复、销毁等密码生命周期管理制度。其三，制定密码人员岗位管理制度，包括密钥管理员、密码安全审计员、密码操作员等关键岗位管理制度，达成各关键岗位之间的相互监督、相互制约机制。其四，制定系统运行和密码安全事件应急预案、安全事件报告制度及流程规范[17]。

2.2.4 密码产品的部署

医疗场景的密码安全建设中，为了保证密码技术的真实性、机密性、完整性、不可否认性，需要部署密码产品作为商用密码建设的支撑[18]，全面覆盖医疗环境中各业务系统和各业务流程。在医疗卫生领域进行密码产品的部署中，需要结合业务场景，选择适合的密码产品，发挥密码技术在网络安全中的保护作用。

密码产品按照功能划分为密码算法类、数据加密类、认证识别类、证书管理类、密钥管理类、密码防伪类以及综合类，总共七大类。（一）密码算法类产品主要是指提供密码运算功能的产品，包括密码芯片等。目前密码芯片被广泛应用于各类密码产品和安全产品，如：智慧医院必须使用国密算法功能的门禁管理系统。（二）数据加密类产品主要是指提供数据加密功能的产品，包括服务器密码机、云服务器密码机、VPN 密码设备、加密存储介质（加密硬盘、加密 U 盘）等，在医疗业务应用计算和存储、网络通信和传输等方面起到至关重要的作用。（三）认证识别类产品主要是指提供身份识别功能产品，包括认证网管、动态口令系统、签名验签服务器等，应用于互联网医疗环境中，为医生、护士、药师、患者提供实名身份认证，保障互联网医疗安全的顺利开展。（四）证书管理类产品主要是指提供数字产生、分发、管理功能的产品，包括数字证书认证系统等。医务人员使用数字证书认证系统为电子病历等医疗文书进行电子签名，确保医疗电子文书的真实可信，包括用户注册管理、证书生成和签发、证书存储和发布、证书状态查询等。（五）密钥管理类产品主要是指提供密钥产生、分发、更新、归档、恢复等功能的产品，包括密钥管理系统等。（六）密码防伪类产品主要是指提供密码防伪验证功能的产品，包括电子印章系统、时间戳服务器等，使卒中、胸痛等急危重症病人的抢救时间点得到准确记录和不被篡改，形成电子病历质控的闭环管理和全程可追溯。（七）综合类产品主要是指提供含有包括医疗业务系统普遍使用的 CA 认证系统和电子水印等防篡改功能的产品。在医院，住院电子病历归档，在病案系统中患者历史病历需要进行至少 30 年的保存周期，因此使用电子印章、时间戳、甚至是区块链技术，就可以保障数据的长期存放不被篡改和抗抵赖。

2.3 同步运行需要与密码评估深度结合持续改进

信息系统建设完成后进入系统运行阶段，医疗机构和组织需要严格执行既定的密码安全管理制度按，定期委托密评机构对信息系统开展密码评估。《密码法》中明确，商用密码应用安全性评估，可与网络安全等级保护测评等相关工作相互衔接、统筹协调[19]。信息系统运行期间，根据医疗机构的安全需求、系统脆弱性、风险威胁程度、系统环境变化状况、管理人员安全认识程度，进行及时的检查和整顿，调整密码应用安全措施，确保密码技术和密码管理措施的落实到位。因此，需要将系统运行与密评深度结合，促进密码防护措施需持续改进。此外，在约束条件发生重要变化时，如：近年来医疗机构信息系统的运化部署调整、基于信创要求的国产化部署调整等，可以对密码应用方案进行修订，对商用密码系统进行升级改造[20]。

3 结果

综上所述，海口市人民医院结合自身特点，在构建密码体现之前，运用密码基础和密码技术手段，结合密码管理方法，部署了存储加密系统和传输加密系统、基于电子病历系统的 CA 认证系统等密码产品。通过省市两上级 WX 主管部门举办年度网络安全实战化攻防演练，时长约 20+10 个工作日，对密码传输保护、安全边界、存储和传输加密策略、密码安全管理进行了演练，找出医院网络中相关需要加固的安全设备点，如 2023 年度，根据演练结果，我们对安全传输加密中客户端、服务器等主机存在的密码漏洞，对加密机设备中 3 个核心模块：弱口令监测预警、终端安全授权防护、应用数据传输加密保护进行应用系统和安全防护的策略调整，使得密码管理体系得到了加强。如图 2 密码安全管理监测图所示，在海口市人民医院部署的密码安全管理监测平台，覆盖医院 HIS、LIS、PACS 等业务系统，通过白名单的维护，促进密码安全保护的持续改进，实现医疗业务的可信互联、安全互通，保障系统稳定运行和数据的不泄露。对于 Web 弱口令、管理员弱密码登录的行为进行监测和定位溯源；安全意识方面，通过在管理机制中加入密码管理的处罚措施，以及与相关维护人员签署数据保密协议的方式，促使全院职工的密码安全保护意识有较大提高，对密码安全事件进行全面监测，使得密码安全管理得到有效提升。

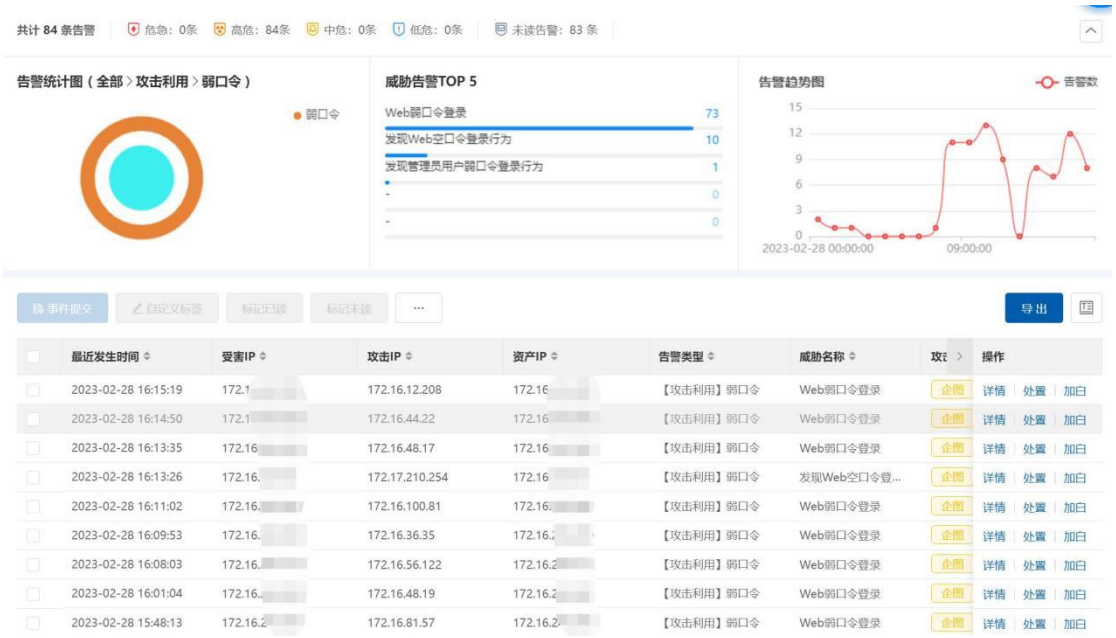


图 2 密码安全管理监测图

4 讨论与思考

从 2021 年开始，国内医疗卫生行业开始推荐商用密码评估与建设的时间较短，虽然在北上广等发达省份的医院已经进行了一些实践，但是还存在不少的困难和阻力。

4.1 医疗行业密码安全意识不强、投入资金不足

组织内部决策层对于密码评估和建设的意识有待加强；商用密码建设的经验和投入经费不足。解决以上存在问题，除了必须尽快提高各级医疗组织领导的重视程度以外，还需要提供经验、教训的相关培训和经费投入的支持，增强密码安全意识夯实安全基础。

4.2 医疗行业密码安全人员的专业技术能力有待加强

2023 年 6 月 15 至 17 日，从上交会的相关分析了解到，现阶段我国对密码人才需求人数约 30 万人左右，实际人才缺口约 20 万人。医疗行业 IT 技术和管理人员的密码安全能力尚有一定差距，熟悉密码技术和相关管理的人才短缺，还需要一个逐步提升的过程。因此，要想解决如上问题，需要加强密码管理人员的培养和技术队伍的锻炼。

4.3 与其他行业以及传统安全方案相比，在医疗行业开展密码安全保护非常必要

结合金融、政府、能源等其他行业使用密码安全评估和建设的方法，医疗行业应该结合医院的系统多、复杂度高的现状，密码安全与系统性能不能两全其美的传统方法，使用前述的密码技术和管理相结合、基础和部署相辅助的方法，提高密码安全的保护能力。

5 结论与展望

展望未来，密码技术作为网络空间安全的核心技术，在网络安全体系中起着不可替代的作用，尤其是在医疗卫生领域，密码安全与网络安全、数据安全等存在无法分割的关联关系。只有从根本上改变商用密码建设和密评难落实的应用现状，才能发挥商用密码在医疗行业网络安全防护中压舱石的重要作用。

PU-037

大数据时代下医学信息安全与隐私保护

龚万里[1,2]，侯雅琪[1,2]，李振浩[1,2]，于 琦[1,2*]
1.山西医科大学管理学院
2. 临床决策研究大数据山西省重点实验室

“新医科”背景下，推动高等教育主动适应人类社会从工业文明逐步进入信息文明社会对人才需求的转变已成为时代趋势。通过医工理文融通以人工智能、大数据等为背景，融合传统医学，发展精准医学、转化医学、智能医学等医学新专业[1]。在以海量数据为前提的新医学领域中，良好的数据素养有助于医学信息如医学文献、医学数据库、电子病历等的保护。但是，由于数据管理机制缺

失、数据化思维不足等仍需关注并以满足在高质量发展下医院大数据时代到来之时医务人员的应对能力。

大数据是指以巨大的、复杂的数据集合为基础，利用计算机技术和数据分析方法进行收集、存储、处理和分析的过程[2]。医疗领域产生的大规模、多样化和复杂的数据经过汇集生成医学大数据。它包括从临床实践、研究、医院信息系统、医学影像、基因组学、生物传感器、移动健康应用等多个来源收集到的大量数据。具有规模庞大、时效性、多维度等特点。通过挖掘医学大数据中隐藏的关联、趋势和模式，可以为医学领域的创新和发展提供有力支持。同时，医学大数据的应用还需要解决隐私保护和数据安全等重要问题。

一、大数据在医学信息领域的应用

1.1 疾病预防与公共卫生

大数据在疾病预防与公共卫生领域的应用具有重要意义。通过收集和分析大量的医学数据，可以识别和监测疾病爆发的趋势，早期预警和采取相应的公共卫生措施。通过收集和整合大量的疫情数据、人口流动数据、社交媒体数据等，可以实时监测和分析疾病的传播趋势，提前预警和快速响应突发公共卫生事件。通过大数据分析，可以评估医疗资源的分布情况和利用效率，为公共卫生决策提供依据。大数据还能帮助揭示与疾病相关的风险因素，以及推动更有效的预防策略。

1.2 个性化治疗

大数据可以为医生提供更全面的、基于证据的决策支持，根据患者的特定情况、基因组信息和疾病模式等数据，制定个体化的诊断和治疗计划，提高治疗效果和患者满意度。真正实现精准治疗。通过个体基因组学、生物标志物分析、临床数据挖掘、科学文献分析以及医生与患者的互动，为患者提供个性化、针对性的治疗方案，提高治疗效果和个体健康结果。

1.3 医疗研究与药物开发

通过大数据分析有助于医学研究和药物开发过程中的样本选择、结果分析和临床试验设计。通过整合和分析多个数据源，可以提供大规模、多维度的数据资源，帮助科研人员更好地理解疾病机制、发现新的治疗目标，并加快药物研发过程，降低成本，推动创新。总之，大数据在医疗研究与药物开发中的应用可以加速药物发现和开发过程，优化临床试验设计，提高药物安全性监测，为疾病治疗提供个性化和精确化的方法。这有助于推动医疗科学的进步，提高药物研发的效率与成功率，并为改善人类健康做出贡献。

1.4 病例管理与医疗质量改进

通过整合和分析大量的临床数据，可以评估医疗服务和患者结果的质量，发现和纠正潜在的问题，并制定改进措施，提高医疗质量和安全性。帮助医疗机构优化医疗服务流程，提高医疗质量和效率，便于顺畅的医患交流，提高人民群众医疗卫生水平和健康水平。

1.5 健康监测与远程护理

利用传感器、智能设备和移动应用程序等技术收集个人健康数据，结合大数据分析，可以实现实时的健康监测和远程护理[3]。这可以帮助提前发现健康风险，进行早期干预，提高慢性病管理的效果。为患者提供更加方便、快捷的医疗服务，促进慢性病管理和健康管理水平的提升。

1.6 药物重定位

药物重定位也被称为药物再利用或老药新用，是指将已经存在且已经获得批准用于其他疾病治疗的药物，重新应用于新的疾病治疗领域[4]。利用海量数据挖掘分析可以实现利用已有的临床数据、安全性数据和制剂工艺等信息，省去了许多研发新药时间和资源。

二、当前医学信息领域使用大数据存在的风险

目前使用大数据的过程中，由于客观条件等因素确实存在一些风险和挑战，包括但不限于以下几个方面：

2.1 隐私泄露和数据安全

医学大数据包含了大量患者的敏感健康信息，如病历、基因组数据等。如果未能妥善保护数据的隐私性，存在泄露个人身份、疾病情况以及其他敏感信息的风险。其存储和传输需要采取高度安全的措施，以防止数据被非法访问、篡改或损坏。数据安全问题可能导致信息完整性和可靠性的问题，甚至危及患者的安全。

2.2 数据质量

医学大数据的质量对于数据分析和应用的结果至关重要。由于医学数据的多样性和复杂性，数据质量问题可能包括错误、缺失、噪声等，这会对数据的解释和应用产生负面影响。同时数据的一致性和数据源的可靠性也至关重要。

2.3 偏差和歧视性

医学大数据可能存在采样偏差，例如经济地位偏差、性别偏差、未能涵盖某些特定人群的数据等，这可能导致结果的误解和不公平性。此外，如果数据分析和应用过程中存在潜在的歧视性，可能导致不公平的照顾和偏见。

2.4 从业人员信息素养

从业人员是医学信息安全与隐私保护的重要参与者，但是可能存在安全意识不足的问题，对医学信息的重要性和敏感性认识不足。缺乏相应的技术能力和知识，不了解数据安全和隐私保护的基本知识，无法处理复杂的数据安全问题等困境。同时从业人员的工作需要获取医学数据和患者信息，如果没有合理的数据授权和访问控制机制，数据泄露的风险将大大增加。

2.5 法律和伦理问题

医学大数据的使用需要遵守相关的法律法规和伦理准则，如个人隐私保护、知情同意等。同时，还需要解决数据分享和合规性的问题，以确保数据的合法使用和共享。

三、如何实现医学信息安全与隐私保护

针对风险一提出以下举措：（1）通过在医学数据存储和传输过程的中应使用加密技术，确保数据在被传送或保存时无法被未经授权的人访问或解读。采用强密码和加密算法是加强数据安全的重要步骤。Almalawi A 等人提出了一种新型的基于 Lionized remora（LRO-S）优化的加密方法对敏感数据加密，减少未经授权的用户和黑客的隐私泄露和网络攻击[5]。（2）通过匿名化和脱敏处理，将医学数据中的个人身份信息去除或替换为不可逆转的标识符，以保护患者的隐私。这样可以使数据在分析和共享过程中更加安全。（3）医学数据应存储在安全的服务器和数据库中，采用物理安全措施（如防火墙、入侵检测系统等）和网络安全措施（如安全传输协议、虚拟专用网络等），确保数据在存储和传输过程中不受到未经授权的访问和攻击。

针对风险二提出以下举措：（1）医疗数据中可能存在数据录入错误、传输错误或仪器测量误差等问题。为了解决这个问题，可以通过提供培训和技术支持，确保医务人员正确记录和输入数据。此外，还可以引入自动化的数据采集和验证机制，减少人为错误。（2）为了降低噪声的影响，可以通过数据清洗和预处理操作，删除异常值、去除重复数据或使用滤波技术，提高数据的质量和准确性。（3）医学数据通常涉及多个来源、多个系统或多个时间点，可能存在不一致的情况。为了确保数据的一致性，应建立标准化的数据模型和数据字典，明确数据的定义和规范。此外，采用数据整合和校验方法，对数据进行交叉验证和比对，消除不一致性问题。（4）医疗数据可能来自不同的数据源，其可信度和质量也会有所差异。因此，在使用医疗数据时，需要对数据源进行评估和验证，确定数据的可靠性和准确性，并采取必要的措施来纠正或排除不可信数据。

针对风险三，为了消除医疗数据偏差和歧视性，需要采取综合的措施，包括公平的数据收集和采样、多元化的参与者招募、数据标准化和整合、开放透明的数据共享和分析，并加强相关机构和决策者对数据偏差和歧视性问题的认识与关注。同时，倡导伦理意识和合规性，确保医疗数据的使用符合法律法规和道德准则。

针对风险四，开展相关的安全培训和教育，加强对从业人员的安全意识培训，提高其对数据安全和隐私保护的重视程度。利用各种培训机会和学习平台，加强从业人员的相关技术培训，提高其数据安全和隐私保护的相关知识和技能水平。建立合理的个人隐私保护机制，规范从业人员处理个人隐私信息的流程和标准，确保数据被妥善保护。

针对风险五，首先相关法律法规（如《个人信息保护法》）对个人数据的收集、存储、处理和共享等环节提出了严格的要求，医疗机构和从业人员需要遵守这些规定[6]。同时在收集和使用医疗数据时，应尊重患者的知情权和自主决策权。患者应该明确知道他们的数据将如何被使用，并有权拒绝或控制其数据的使用方式。医疗机构和研究机构需要征得患者的知情同意，确保合法合规。还需要采取严格的安全措施来防止未经授权的访问、泄露或滥用。这涉及到数据存储和传输过程中的加密、权限管理、安全审计等技术和管理措施。

总之，应对这些风险，医学界和政府部门需要共同努力制定和执行严格的数据安全和隐私保护措施。同时，加强数据的质量管理和标准化，提高数据分析和挖掘的能力，以确保医学大数据的可靠性和有效性。此外，促进公众教育和参与，加强对医学大数据应用的监管和治理也是必要的措施。

PU-038

零信任技术在医院信息化应用及展望

彭浩明 仇成凤*
湖南医药学院总医院

1、背景分析

随着信息化技术的不断发展，医院信息化的应用不断加强，为了提高患者就医满意度，在互联网应用方面，大力推广网上预约挂号、微信支付、检查报告单在线浏览等服务，方便了患者的就医就诊，提升医院的整体的竞争能力。在国家卫建委的响应与号召下，各个医院逐渐实现了院内业务系统的互联互通，打破了传统医疗信息的壁垒。随着“互联网+医疗”的广泛应用，医院核心业务系统不再封闭，系统与系统间的数据共享及互访打破了传统的数据孤岛。我们在享受信息化带来的便利同时，也发现了信息化暴露出很多安全问题。

2、传统安全防护及问题

当前医院网络结构采用内外网隔离区域划分的方式，核心业务被部署于内部网络中，利用防火墙、IPS/IDS 等安全设备根据业务需求进行边界防护，如图 1 所示。随着远程医疗、微信预约、手机支付、互联网医院等等互联网应用的普及，内外网之间的隔离边界已经逐步消失。早期医院内网终端原本只需要访问医院的 HIS\LIS\PACS\EMR 等核心业务系统，但随着互联网的普及，给患者带来了极大的便利和体验感，但内网终端的访问对象不再局限于内部系统，便捷的方式直接增加了导致内网电脑感染外网病毒几率，从而直接对内部核心系统产生了威胁。

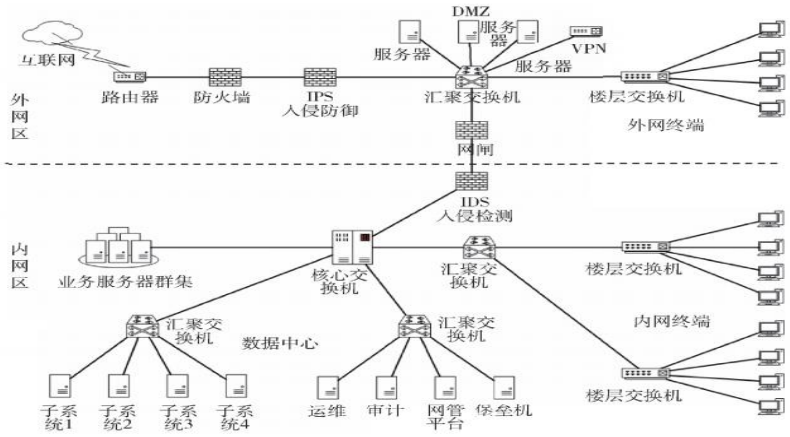


图 1 传统网络架构图

近年以来，病毒通过远程扫描端口，侵入外网前置主机或者终端，通过暴力破解或者蠕虫等传染性病毒的横向扩展，在潜伏期以后通过内网主机对核心服务器发起攻击，从而导致业务终端、文件加密、系统蓝屏等各种故障，严重影响到医院各项业务的开展。而医院内外网业务的融合、医疗数据的互联互通、国家数据网上直报等业务的需要，物理上区分内网外网的架构在现代医院信息化建设中越来越难以实现，我们需要一个开放、安全、便捷的网络架构，能够在安全可控的情况下实现资源互联互通。

3、零信任理念框架

要随着云计算、移动办公等技术的普及，网络结构变得复杂，传统的网络安全模型基于边界防护的思想，已经逐步无法应对目前严峻的网络安全局势，一种零信任架构应运而生。

零信任网络中的一切 ，不区分外网与内网，均默认为是不可信的,都需要经过认证和授权。对网咯实体的认证与授权以每次连接为单位进行策略随状态变化动态调整。

在 BeyondCorp 中，用户首先需行身份认证，认证方式由地点决定。在谷歌大楼内，用户使用 RADIUS 行认证；在公网使用单点登录 SSO 方式进行认证，通过认证后还需要授权才能获得访问权限，授权通过“访问控制引擎”赋予“访问代理” 授权信息而获取。BeyondCorp 会在身份认证后把结果（人和设备）关联至基于具体业务构建的 vnet 网段，形成一个个不同安全等级的隔离域，用户如果想跨域须遵守相关安全策略.

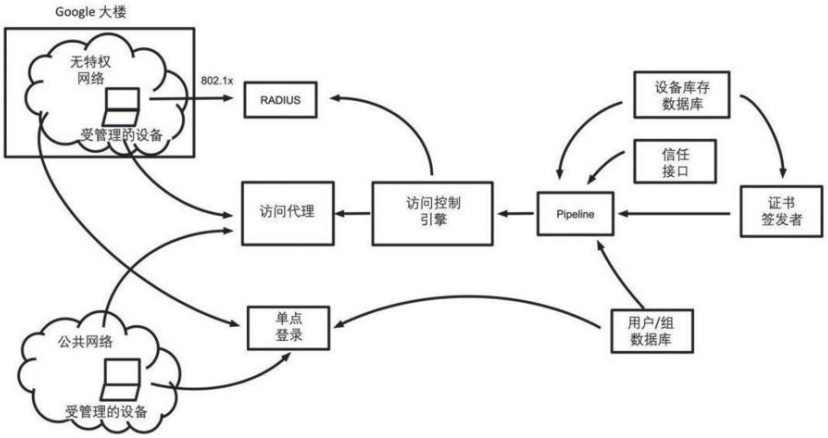


图 2 零信任 BeyondCore 架构

在美国国家标准技术研究院（NIST）发布的零信任架构标准中，明确列出了当前实现零信任的三大技术路线，可以归纳为“SIM”组合：SDP，软件定义边界；IAM，增强身份管理；MSG，微隔离。这三种技术路线可以单独实现零信任方案，也可以配合起来实现更加全面的零信任架构。零信任削弱了网络位置的优势地位，消除了过渡的隐式信任，代之以显式的基于身份的信任。从某种意义上说，零信任创建了个性化的“虚拟边界”，该边界仅包含用户、设备、应用。零信任还规范了用户的体验，消除了在医院网络之中访问存在的差异。

4、零信任技术在未来医院信息化中的落地实践

4.1 远程运维的应用

过去我们使用 VPN (虚拟专用网)在公网上面建立专用网络，通过对数据包的加密和数据包目标地址的转换实现远程访问，VPN 奉行的是“内部都可信”、“外部不可信”的安全模式，也就是说可信的人员在内部，不可信的人员在外部。

基于零信任安全理念在医院远程运维场景下实施安全接入解决方案。落地实践业务流如下所示：

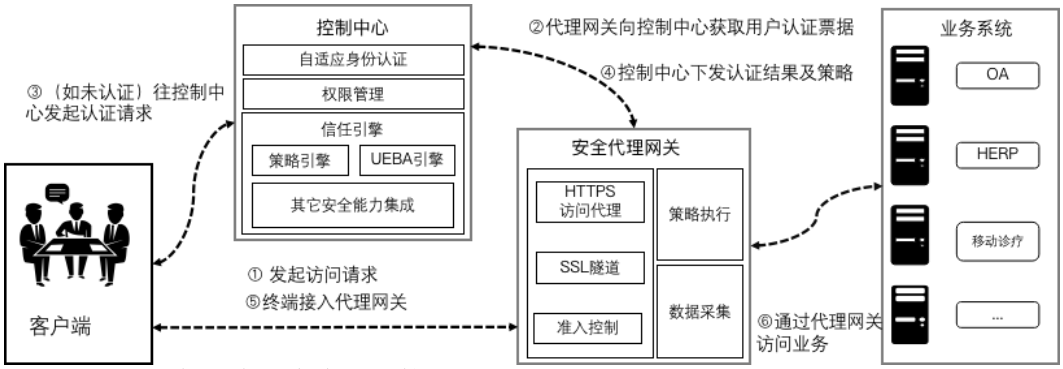


图 3 基于零信任理念的安全访问接入

安全代理网关部署在业务系统前面，当用户访问业务系统时，符合管理员设定的安全基线才允许访问，具体应用中要求远程运维终端必需打齐操作系统补丁才允许访问业务应用；要求终端必需安装防病毒软件；要求终端必需安装制定终端准入客户端才允许访问 OA、HERP、ORALCE 数据库等系统；要求终端必需加入域控在内网才允许访问财务系统数据库。

通过 SPA 单包授权安全机制实现『网络隐身』，保护业务系统服务器（服务器不会响应来自未授信客户端的任何连接），同时可以缓解对 TLS 的拒绝服务攻击。基于敏感业务设置 WEB 水印；事先震慑与事后追溯有效控制数据外泄。

权限与用户关联，对于外部运维人员仅开放链接数据库客户端，不开放服务器访问权限，通过动态访问权限控制，常用 IP 控制、白名单控制、终端进程控制等多源信任评估机制实现灰度处置，对可疑行为进行有效处置，兼顾安全与体验。

4.2 内外网无边界访问应用

随着“互联网+医疗”业务的发展，以及云计算、移动互联等技术让医院的人和业务、数据“走”出医院的内网的边界；另一方面，大数据以及物联网等新技术开放协同需求导致了外部的人员，平台

和服务“跨”过了医院的数字护城河。例如，移动诊疗业务采用多种物联网技术采集患者信息，提供给临床人员与此同时提供给患者在互联网环境下可以进行随时访问。所以说，复杂的现代化的医院网络基础设施已经不存在单一的、已识别的、明确的安全边界，或者说医院的安全边界正在逐渐瓦解。

为适应未来医院业务全面互联网化，通过零信任架构，构建内外网无边界安全访问，整体设计如下所示：

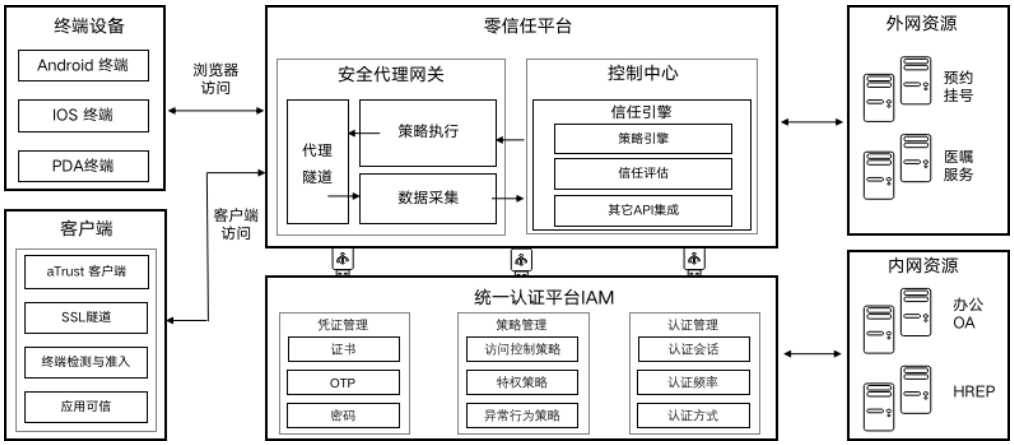


图 4 基于零信任架构内外网无界访问

通过以上架构设计可以实现，医院内网环境与互联网环境下通过统一认证平台 IAM 实现同一个 Web 页面下访问内外网业务系统，内外网无界安全访问整体组件设计如下：

安全代理网关

零信任访问通道的基础组件，负责将内部的 Web 应用，使用反向代理的方式映射到外网。用户访问应用时会首先检查回话是否进行了认证，与统一认证管理平台 IAM 进行联动，完成用户的认证和授权。

统一认证平台 IAM

统一身份认证平台可实心账号的认证、授权。审计、审批和应用管理的，统一认证平台 IAM 与安全代理网关设备进行集中管理，控制策略的及服务端的下发功能

分析与控制中心

实现信任策略的管理（用户、设备、行为、网络）安全上下文属性的处理（用户认证、设备、UEBA），策略的执行效果评估等安全分析能力。

5、零信任技术框架在医院应用展望

本次基于零信任技术在医院信息化建设环境之中进行实践，基于零信任技术理念，使用 SDP 组件，实现了远程接入运维接入的终端检测、动态授权、可信验证、安全可控。最大程度上缩小了内部服务器的暴露面，实现了安全边界重构。

基于 IAM 与零信任平台实现了基于终端环境、身份基用户行为进行持续风险评估，威胁进程终端禁止访问，基于身份 ACL 进行自适应身份安全，实现动态鉴权基 PC 与移动终端统一授权接入，提升了用户的访问体验。

未来，结合医院安全能力扩展如堡垒机等安全设备使用 IAM 机制，从安全与运维两个角度实现第三方能力集成，与各安全组件进行联动，实现医院网络建设价值最大化，完善与优化零信任架构在院内的应用落地。

PU-039

病案管理中患者隐私的泄露风险及保护措施

赵 敏

郑州大学第一附属医院 病案管理科

尊重和保护公民的隐私权是社会文明和进步的表现。随着社会发展，公民对保护自身隐私权的要求越来越高。病案涉及大量患者隐私，在对其管理过程中，可能由于种种导致患者信息泄露，从而使患者的隐私权受到侵犯。因此，如何保护好患者的隐私权益，是病案管理人员必须要考虑的问题。本文通过探讨病案中所涉及的患者隐私内容，分析患者隐私泄露途径，制定病案管理相关制度，从可能造成患者隐私泄露的各环节入手保护患者隐私，从而维护医患双方的权益，构建和谐医患关系，使病案信息更好、更全面地为社会服务。

1.患者隐私包含的具体内容及泄露风险和相关法律规定

隐私是一种当事人不愿他人知道或他人不便知道的个人信息。隐私权属于人格权的范畴，是指公民享有的不公开私人生活相关的事实和秘密的权利。这是一个公民的最基本的权利。只要一个人的隐私权与公共利益并不冲突，它就属于个人隐私的范畴，可受法律保护。病案中包含的患者信息涉及诸多隐私。一旦患者隐私泄露会给患者本人带来不可预测的损失或者伤害。尊重和保护患者隐私权不受侵犯是对病案管理工作者的法律要求和职业道德体现。

1.1 患者隐私包含的内容

病案是医务人员按规范记录病人疾病表现和诊疗情况的档案文件，它客观、完整的记录了病人的疾病诊断、诊疗经过、及转归情况，是医疗、教学、科研的基础资料[1]。病案不仅有纸质的，还有电子文档、医学影像胶片、病理切片等形式。病案中涉及的患者隐私内容包括个人基本信息和疾病诊疗信息。个人基本信息包括姓名、出生日期、户籍地址、身份证号码、家庭地址、职业、工作单位、联系方式、婚姻状况等。疾病诊疗信息不仅包括患者的主观描述的发病情况、疾病诊断、检验结果、影像资料、治疗经过、病程转归、手术记录、麻醉记录、会诊记录等，还包括患者既往疾病史、传染病史、婚育史、特殊个人史、家族史等。病案属于卫生信息档案，涉及患者的诸多隐私，因此具有高度的保密性。病案管理的保密工作是对公民隐私权的法律要求和基本尊重。医疗机构及相关人员要做好病案的保密工作。

1.2 相关法律法规

虽然目前我国暂未制定专门的医疗隐私信息保护法，但现行的法律法规中对保护患者隐私也有相关规定。《中华人民共和国民法典》第一千二百二十六条：医疗机构及医务人员应当对患者的隐私和个人信息保密，泄露患者隐私和个人信息，或者未经患者同意公开起病历资料的，应承担侵权责任。《中华人民共和国侵权责任法》第六十二条：医疗机构及其医务人员应当对患者的隐私保密。泄露患者隐私或者未经患者同意公开其病历资料，造成患者损害的，应当承担侵权责任。《医疗机构病历管理规定》第六条第二款规定：医疗机构及其医务人员应当严格保护患者隐私，禁止以非医疗、教学、研究目的泄露患者的病历资料。《执业医师法》第二十二条规定医师在执业活动中，有尊重患者的隐私的义务。《传染病防治法》第四十三条规定医疗机构及相关工作人员必须保护好患者隐私，不得用患者隐私获取非法利益[2]。《护士管理条例》 《医务人员医德规范及实施办法》等多部法律法规中也都有保护患者隐私的明确规定。

1.3 患者隐私泄露之后带来的风险

病案中包含的患者隐私内容一旦泄露，会给患者本人带来诸多风险和损失。包括身份盗窃，金融损失，虚假信息，网络钓鱼，个人声誉受损，社交攻击，个人隐私侵犯等[3]。个人身份信息被泄露，不法分子可利用并冒充受害者进行欺诈活动，如打开信用卡账户、申请贷款、购买商品或者服务等，从而导致受害者遭受巨大的经济损失。此外当个人信用记录受到破坏时，还可能导致个人信贷能力和信用评级，从而影响个人未来的财务状况。不法分子还可以还可以利用个人敏感信息来制造虚假信息，来欺骗受害者提供更多的个人信息。隐私泄露还可能导致个人声誉受损，使受害者受到骚扰或者恶意攻击，影响其社交和职业生涯，不法分子窃取个人敏感信息，并将其发布到互联网上，从而侵犯受害者的隐私权和人格尊严。

2. 患者隐私泄露的途径

病案管理工作流程中的任何一个环节出问题都可能会导致患者隐私对外泄露，造成侵犯患者隐私的行为发生。其原因是多方面的，既有相关人员的主观意识因素，也有数字存储及网络技术存在安全隐患，也有病历采集、回收、保存、复印、借阅等环节操作规范不够完善和工作流程存在缺陷等因素。不仅涉及到书写病案的临床医务人员和病案管理者，也可能涉及到有权限的病案查阅及使用者。根据相关案例分析，病案管理中侵犯患者隐私权的行为主要表现为以下几方面。

2.1 有意泄露患者隐私

部分医疗从业人员缺乏职业道德，且法律意识淡薄，受非法利益的诱惑和驱使，利用工作之便获取并主动泄露患者的隐私内容，如姓名、家庭住址、联系方式、职业、工作单位、生理缺陷、传染病史、不正常的婚姻关系、特殊治疗、特殊的个人史等，或者利用系统漏洞或后门等窃取或更改相关档案数据，或留下漏洞以便他人的恶意攻击。医学研究者或他人未经患者同意擅自让他人或组织翻阅患者的病案资料，引用患者的病案信息，或使用患者的照片和真实姓名进行临床病例研究、行业学术交流、个人科研论文撰写等[4]。目前，有意泄露的形式在医院患者隐私泄露案件中占了很大的比率，而且呈逐步增长的趋势。

2.2 医务人员操作不规范造成的隐私泄露

医护人员在诊疗过程中因疏忽大意等原因造成患者病案的丢失或将患者病案隐私内容泄漏给他人等[5]。如在医院公开场合将患者检查检验结果随意放置，或者离开办公室未关闭病历系统，这都给无关人员随意翻看查阅病历内容提供了可乘之机。病案管理人员在收藏、整理病案的过程中的不当行为导致病案的损坏，或因责任心不强导致病案丢失、被盗等。因病案借阅权限未严格限制，病案追踪不明确，比如疑难病案、护理大查房、带教示范等借出病案未及时收回；部分医务人员疏于核对患者家属身份就向其透露患者病情病况，致使病案还未归档就已出现隐私泄露的风险；未归档病案在病区随意传阅或被其他违规使用；病案质控小组等管理部门对病历填写情况进行常规性的评估检查，对不规范病历给予公开点评并定期通报，这些都可能导致患者隐私外泄事故的发生。对于个人代理复印病案资料的，往往存在着亲属、同事或朋友在无患者本人授权或者手续不全的情况下获知患者医疗信息的风险。对于单位或者集体体检申请复印的，如果审查不严违规操作，不仅会导致成批量的患者隐私安全不保，还极有可能激发各方冲突，引起不必要的纠纷。

2.3 信息与网络安全存在的风险

网络安全泄露 基于数字网络的普及和发展，医疗构信息化建设不断推进，电子病案、远程医疗、医疗信息系统、实验室信息系统等飞速发展[6]。网络技术的普遍纵深性应用在给医疗活动带来便捷的同时，也不可避免地为患者隐私安全带来了信息泄露的问题。用户在对病案进行调阅查看时，易被他人恶意利用或钻空档进行攻击，导致数据的丢失和被窃，以及在利用 U 盘或硬盘时拷贝时也存在数据丢失的安全隐患。另外，病案存放的数据库存在被攻击者利用病毒、木马、蠕虫等攻击并导致盗取复制、外泄篡改的巨大风险[7]。区域范围内的医疗信息共享、信息传递交流可能存在医疗信息泄露的问题。

3. 患者隐私的保护措施

针对病案管理中患者隐私泄露的各种途径，应制定一系列患者隐私保护的措施，减少或杜绝患者隐私泄露。首先尊重和保护患者的隐私权是医务人员的职业操守、道德规范和法律要求。应开展继续教育和培训，学习相关的法律法规规章，提高病案管理人员的职业素质和信息安全意识。其次应建立健全病案管理制度，制定严格的病历借阅、复印制度，建立医院病案监督领导小组，严格控制病案管理规范流程，加大处罚力度。最后还要加快医院病案信息化建设，建立信息网络安全规章制度，加强技术改进和网络监督，保障电子病历在存储、传输、查阅方面的保密性、安全性。

3.1 加强医德教育提高医务人员修养

希波克拉底在医学誓言中说: “行医处事所见所闻，永当保密，绝不泄密。”尊重和保护患者的隐私权是医务人员的职业操守、道德规范和法律要求。应开展继续教育和培训，学习相关的法律法规规章，提高病案管理人员的职业素质和信息安全意识。严格执行相关规定和制度，严格保密，不随意传播、不外泄，保护患者隐私权不受非法侵犯。绝对不能为了谋取不正当利益，将病历违规借阅、复印、外传，导致患者隐私泄露[8]。

3.2 建立健全管理制度

第一，制定严格的病历借阅、复印制度。借阅或复印病历必须经过患者本人同意或委托他人（有法律承认的委托函）且填写申请单签字或按手印。第二，所有病历不得以任何理由外借，如需进行科研教学的，经主管部门审批后可在病案阅览室阅读。第三，加强病案管理人员管理，未经主管部门同意，不得擅自同意借阅或复印患者病历。第四，建立医院病案监督领导小组，联合质控

处、医务处、病案科、感染科、护理部等行政职能部门，负责全院病案的监督管理，严格控制病案管理，防止患者隐私泄露。第五，严格规范流程，加大处罚力度。对于违法违规行为，及时纠正改错，并按情节严重程度给予不同的处罚；对给单位或患者造成严重损失的要从重处罚，甚至追究其刑事责任[9]。

3.3 加快医院病案信息化建设

随着医院电子病历系统建设水平不断提高，提升医疗信息安全与保护患者医疗信息是当前医院加强病案管理的重要方向。制定电子病历管理制度，建立电子认证机制，设置使用者的权限。通过程序设置严禁复制拷贝病历。规定电子病历只有患者主管医护人员才有权限调取查阅[10]。限制电子病历接触人员，从而减少患者隐私扩散的可能性。还应当建立信息网络安全规章制度，加强技术改进利用数据加密，身份识别，反病毒等先进技术手段，加强网络监督，保障电子病历在存储、传输、查阅方面的完整性、保密性、安全性，保障电子病案及远程医疗的安全，保护患者隐私的安全。

综上所述，病案具有严格的保密性，它关系到患者的切身利益。病案的保密工作是一项复杂细致的工作，这要求病案管理工作人员具有较高专业素养和法律意识，自觉遵守病案管理法律法规，制定并严格执行病案工作各项规章制度，加强病案管理系统信息化建设，加强技术改进和网络监督，在病案管理的各项环节中防止泄密、失密及病历资料丢失，保护患者隐私权，维护患者利益。

PU-040

医院信息安全和隐私保护的政策研究

黎颜
广德市中医院

一、研究背景、目的和重要性

随着科技的不断发展，医疗健康的信息化日益增强，医院信息系统也在各种医疗服务和管理环节中扮演着重要的角色。然而，这同时也向我们凸显出一个令人担忧的问题，那就是医疗信息的安全和隐私保护问题。医疗数据中包括患者的身份信息，病史，诊疗信息等，一旦这些数据被非法获取，或发生泄露，都将对患者的权益和医疗系统的正常运行造成严重的影响。因此，医院信息安全与隐私保护问题必须引起高度重视。

本研究的目的在于探究和提出有效的医院信息安全和隐私保护政策，以应对当前正在面临的挑战。其重要性主要体现在以下几个方面：首先，这对于维护患者的合法权益具有重要作用，可以防止其敏感信息的滥用和泄露；其次，该研究有助于增强医疗机构的信誉和公众对医疗服务的信任，进而促进医疗行业的健康可持续发展；最后，研究结果还可为医疗信息系统的合理设计和规范运行提供理论依据和推动力。

二、医院信息安全存在的问题

2.1 医院信息系统的脆弱性

医院信息系统是一个复杂的网络结构，其中包含大量的敏感和私有信息，如病人的健康记录、诊断、治疗方案、医生和护士的个人信息等。这些信息系统的脆弱性可以分为以下几类：

1、技术脆弱性：由于信息系统的设计和实施缺陷，可能会出现技术性的脆弱性，如系统安全配置不足、过时的安全补丁、未修复的漏洞等。

2、人为脆弱性：包括内部人员的错误操作、恶意行为，或者外部攻击者利用内部人员的无知或欺诈等行为。

3、管理脆弱性：包括信息安全政策的缺失或执行不力，缺乏有效的风险管理策略，以及医院对信息安全的投资不足。

2.2 威胁医院信息安全的常见因素

威胁医院信息安全的因素众多，常见的包括：

1、网络攻击：包括病毒、恶意软件、勒索软件、钓鱼攻击等，这些都可能导致信息系统的损害，数据丢失或被盗。

2、内部威胁：员工的误操作或恶意行为可能导致信息泄露。例如，员工可能因为缺乏足够的安全培训，不小心点击了恶意链接，或者泄露了密码。

3、物理威胁：包括设备损坏、自然灾害等，这些都可能导致信息系统的中断或数据丢失。

2.3 医院信息泄露的风险

医院信息泄露的风险非常大，可能会导致以下后果：

1、病人隐私泄露：医院信息系统包含病人的个人信息、疾病历史、治疗方案等敏感信息，一旦泄露，可能会对病人的隐私权造成严重损害。

2、法律风险：医院可能因为未能妥善保护病人信息，而面临法律诉讼、罚款等问题。

3、信誉风险：信息泄露事件可能会影响医院的公众形象，降低患者和公众的信任度，这可能会对医院的运营产生负面影响

三、医院隐私保护政策

3.1 国际和国内隐私保护法律框架

国际和国内的隐私保护法律框架为医院隐私保护政策提供了基础。在国际层面，联合国《人权理事会关于隐私权的保护在数字时代的重要性》决议明确了个人隐私权的普遍适用性和保护的重要性，倡导各国政府采取措施保障公民的个人信息安全。在国内层面，这方面的法律包括《中华人民共和国网络安全法》、《中华人民共和国个人信息保护法》等。这些法律对个人信息的收集、使用、储存、共享、转让等行为进行了规范，旨在保护个人隐私和信息安全。

3.2 医院隐私保护政策制定的关键考虑因素

在制定医院隐私保护政策时，需要考虑多种因素。首先，需要考虑的是法律规定，任何政策都必须符合法律要求。其次，需要考虑的是患者的隐私权益，政策制定者必须确保患者的个人信息得到充分保护，防止被未经授权的人员获取和滥用。此外，还需要考虑到医院的运营需求，比如医疗研究，医疗服务提供等。

3.3 医院隐私保护政策的实施和合规性监管

在医院隐私保护政策制定之后，实施和监管就成为了关键。首先，医院需要制定出具体的实施计划，并对所有员工进行培训，确保他们了解并遵守隐私保护政策。然后，医院需要设立专门的合规部门或者人员，负责监督并检查政策的执行情况。对于违反政策的行为，需要有明确的处罚措施。此外，医院还需要定期对隐私保护政策进行审查和更新，以应对法律和技术的变化。

四、医院信息安全管理

医院信息安全管理是一个关键问题，尤其是在我们的医疗系统日益依赖技术的当今世界。这需要一个多方面的方法，包括建立有效的安全管理框架，制定和实施安全策略和措施，以及进行员工培训和教育。

4.1 医院信息安全管理框架

医院信息安全管理框架是一个组织的基础，它定义了如何管理和保护信息资产。这个框架应该包括明确的策略、程序和工具，以确保医疗信息的安全。例如，它可能涵盖数据加密、访问控制、备份策略、安全审计以及应对安全事件的应急计划。此外，该框架应当符合相关法规要求，如《中华人民共和国个人信息保护法》。

4.2 安全策略和措施的制定与实施

安全策略和措施的制定与实施是信息安全管理的核心部分。这些策略和措施应明确定义如何处理和保护敏感信息，以防止数据泄露或被非法访问。例如，可能需要实施强密码策略、定期更新软件和硬件、使用防火墙和防病毒软件、限制对敏感数据的访问，以及定期进行安全评估和审计。所有这些措施都应与整个安全框架一致，并确保其实施。

4.3 员工培训和教育

员工培训和教育是一个重要的组成部分，因为员工往往是安全威胁的主要目标。医院应提供定期的安全培训，使员工了解当前的威胁，如网络钓鱼和恶意软件，以及如何防止这些威胁。培训还应包括正确的数据处理程序，如什么信息可以分享，什么信息需要保密，以及如何安全地存储和处理病人数据。通过这样的培训，医院可以减少由于员工误操作导致的数据泄露风险。

五、医院信息安全技术应用

5.1 医院信息安全技术的分类和特点：

1、访问控制技术：确保只有授权用户才能访问医院的信息系统。特点是通过用户身份验证和访问权限设置实现信息的保密性和完整性。

2、数据加密技术：通过对数据进行加密，使得即使数据被非法获取，也无法被读取。特点是能够保护数据的安全性和隐私性。

3、安全防火墙技术：用来防止外部的非法访问和内部的非法外发。特点是可以有效防止网络攻击和数据泄露。

4、安全审计技术：通过记录和分析系统的活动信息，来发现可能的安全威胁。特点是能够及时发现和处理安全事件。

5、入侵检测和防御系统：能够检测并阻止网络攻击和恶意软件。特点是可以提高系统的抗攻击能力。

5.2 医院信息安全技术在实践中的应用案例分析：

北京协和医院将智慧科研的建设作为创新切入点，通过全系统数据整合，完善临床信息系统标准化建设，不断探索科研模式创新、构建智慧科研新生态，建设了多层次、多领域覆盖的全院级综合性医疗大数据平台，以及以专科为核心的专病数据库平台等。医院搭建了从数据到应用的多层架构体系，其中包括：标准体系、多模态数据体系、多模态数据融合与处理、科技创新应用等

六、医院信息隐私保护措施

1、数据加密和访问控制 数据加密是通过将数据转换为无法读取的形式，以防止未授权用户获取和理解这些信息的一种方法。在医疗环境中，这可能包括使用诸如高级加密标准（AES）或其他加密协议的技术，将患者的个人健康信息（PHI）转换为加密的形式。访问控制则涉及到确定哪些用户可以访问特定的信息。这可能包括使用用户名和密码系统，以及更高级的生物识别技术，如指纹扫描或面部识别。此外，还可能包括设置权限级别，以确定哪些员工可以访问哪些信息。例如，护士可能只能访问他们的病人的信息，而医生可能能够访问更广泛的信息。

2、医疗记录的机密性和完整性保护 医疗记录的机密性是指这些记录只能由特定的授权人员访问和查看。这通常通过设置严格的访问控制和使用加密技术来实现。此外，所有有权访问这些记录的人员都应接受关于保护隐私和数据安全的培训。医疗记录的完整性则是指保护记录不被未授权的修改或删除。这可以通过使用特定的软件和协议来实现，比如使用版本控制系统来跟踪对记录的所有修改，以及使用备份系统来防止数据丢失。

3、医院隐私检测和违规处理 医院隐私检测是一种程序，用于定期或者连续地监视和检测医疗环境中的隐私违规行为。这可能包括使用特定的软件工具来监测潜在的数据泄露，或者使用人工智能技术来识别可能的隐私侵害行为。违规处理则是指当检测到隐私侵害行为时，医院如何采取行动。这可能包括调查疑似侵犯隐私的行为，制定和执行相应的纪律处分，以及在必要时报告这些行为给相关的监管机构。同时，医院还需要采取措施来防止类似的侵犯行为在未来再次发生。

七、未来的发展和挑战

7.1 医院信息安全和隐私保护的发展趋势

1、强化网络安全：随着医院信息化程度的提高，医院信息系统面临的网络安全威胁也越来越多。因此，医院将加强网络安全措施，包括加密通信、防火墙、入侵检测系统等，以保护医院信息系统的安全。

2、提升员工安全意识：医院将加强员工对信息安全的培训和教育，提升员工的安全意识。员工必须了解信息安全的重要性，并掌握信息安全的基本知识和技能，以避免社会工程学攻击、密码泄露等安全风险。

3、加强数据隐私保护：医院将加强对患者和医疗数据的隐私保护。医院将采取严格的权限管理措施，确保只有授权人员可以访问患者和医疗数据，并采取必要的技术手段加密和保护数据传输过程中的安全性。

4、建立信息安全管理体系：医院将建立完善的信息安全管理体系，制定和执行信息安全政策和控制措施。医院将进行信息安全风险评估，并定期进行安全演练和安全审计，以识别和解决安全问题。

5、加强云安全保护：随着云计算在医疗行业的广泛应用，医院将加强对云安全的保护。医院将选择可信的云服务提供商，确保数据在云平台上的安全存储和传输，并加强对云平台的监控和管理。

7.2 面临的挑战和解决方案

1、安全威胁的多样化: 随着网络技术的发展，安全威胁也在不断变化和升级。除了传统的网络攻击和数据泄露，还有来自内部人员的安全威胁，以及由于设备和系统的漏洞引起的安全问题。解决方案：建立完善的安全防护体系，包括网络安全、终端安全、数据安全等多个层面。同时，需要定期进行安全审计和风险评估，以便及时发现和处理安全威胁。

2、法规合规的压力: 随着医疗数据保护法规的不断完善，医院也面临着越来越大的法规合规压力。 解决方案：建立专门的合规团队，负责跟踪最新的法规要求，制定和实施合规策略。同时，需要进行定期的合规培训和审计，确保医院的各项操作都符合法规要求。

3、技术更新的挑战: 随着科技的快速发展，医疗信息系统的更新换代也在加快。然而，新的技术和系统往往会带来新的安全威胁。 解决方案：在引入新技术和系统时，需要进行全面的风险分析，了解可能存在的安全问题。同时，需要制定和实施相应的风险管理策略，以确保新技术和系统的安全。此外，定期进行软件和操作系统的补丁和更新，以及员工的安全培训，也是保障新技术和系统安全的重要手段

4、病人数据保护的挑战：由于医疗数据的敏感性，如何在确保医疗服务的便利性和效率的同时，保护病人的数据隐私，是医院面临的一个重要挑战。 解决方案：建立和实施强有力的数据隐私保护政策，包括对数据的收集、存储、处理和传输进行严格的控制。同时，需要提高病人对数据隐私的知识和意识，让他们了解自己的数据是如何被保护的，以及他们可以如何参与到这个过程中来。

5、网络安全威胁的挑战：随着网络的连接性和设备的智能化，医院面临的网络安全威胁也在增加。 解决方案：首先，需要了解医院网络中所有的电子受保护健康信息（ePHI）的位置，包括软件、连接设备、遗留系统等，以便对这些信息进行全面的保护。其次，需要定期进行网络扫描，以识别和处理可能的漏洞。此外，还需要提高员工的网络卫生习惯，如定期更改密码，不在公共网络中访问敏感信息等。

总的来说，医院信息安全和隐私保护的未来发展趋势和挑战都体现了科技发展和法规要求对医院的影响。医院需要不断更新和完善自己的安全和隐私保护策略，以适应这些变化和挑战。未来的研究应该继续关注医院信息安全和隐私保护的重要性和挑战，探索有效的应对策略，并紧跟科技发展和法规要求的影响，为医院信息安全和隐私保护的实践和政策提供更好的理论依据和指导。

PU-041

内外网文件安全交换管理系统

冯东建
南阳市第二人民医院信息科

前言

医疗信息系统由于其特殊性，安全性，为了防止病毒、木马的攻击，我院在建设过程中采取了内外网隔离模式。医护人员仅在医院内部局域网内办公，不仅不能通过访问互联网获取数据资料，也无法做到医院内网和互联网之间文件的便捷传输。各科室人员如需要跨网传文件时，要拿着申请单到信息科排队处理。受限于科室人员不够充裕，以及繁琐的传输流程，不仅传输效率极低，传输的时效性也无法保证。特别是体检中心、财务科、检验科等业务科室经常会有特别紧急的资料需要及时传输，非常影响工作效率。再者是医院内网部署的 HIS，LIS，PACS 等业务系统，运维人员需要把执行脚本拷贝到内网，但不允许自行拷贝其他类型文件到内网，更不允许把内网数据拷贝到外网。运维工程师经常需要在内外网之间传输更新代码、系统升级补丁和日志文件等，这类文件为确保其安全性及隐私性往往需要审核、杀毒后传输，并留档备查，这一系列流程由人工处理特别费时、费力，而且非常影响系统问题的解决效率。内外网隔离这种模式虽然满足了数据的安全需求，但随着互联网、物联网、云计算、大数据等信息技术的发展及普及不足之处也很明显：如何将内网的数据和外网的数据进行安全的传输是医院所有工作者的诉求，也是医院信息工作者一直在努力探寻解决的实际问题。

通过人工处理还存在一些问题：病毒查杀只能依赖工程师手动进行，一旦工程师使用的终端电脑上病毒库没有及时更新，或者因文件传输紧急未执行手动杀毒，就很容易造成带毒文件传输引起内外网之间的病毒扩散；同时在数据交互过程涉及了大量内部敏感数据或病患个人隐私数据，一旦发生数据泄漏很难溯源，缺少技术审计的技术手段。

南阳市二院联合南阳理工学院南阳大数据研究院、奇安信科技集团股份有限公司等单位联合攻关，成功解决内外网传输安全、快捷、方便、可追溯问题。

一、痛点难点

1.人工控制数据交互，效率低下，越来越不适应医院高速的业务增长需求。

之前内外网数据交互采用的方式是先填写申请表，经过科室主任签字同意后，再递交到信息科，由信息科工作人员人工处理。各科室在传文件时，要拿着申请单到信息科排队处理。特别是

（体检业务、PCR 检验、财务报表等）业务科室特别紧急的时候，也得等信息科工程师一个一个处理。

尤其是同时有多个科室需要上报大量材料的时候，可能同时需要发送数十个文件，受限于人工处理的流程，病毒查杀时间较长，经常会出现人员排队等待的情况。

此方式效率很低，各科室也反应等待时间较长，信息科和业务科室都深受困扰。

2.医院有大量业务系统涉及的内外网文件交互无法有效管理。

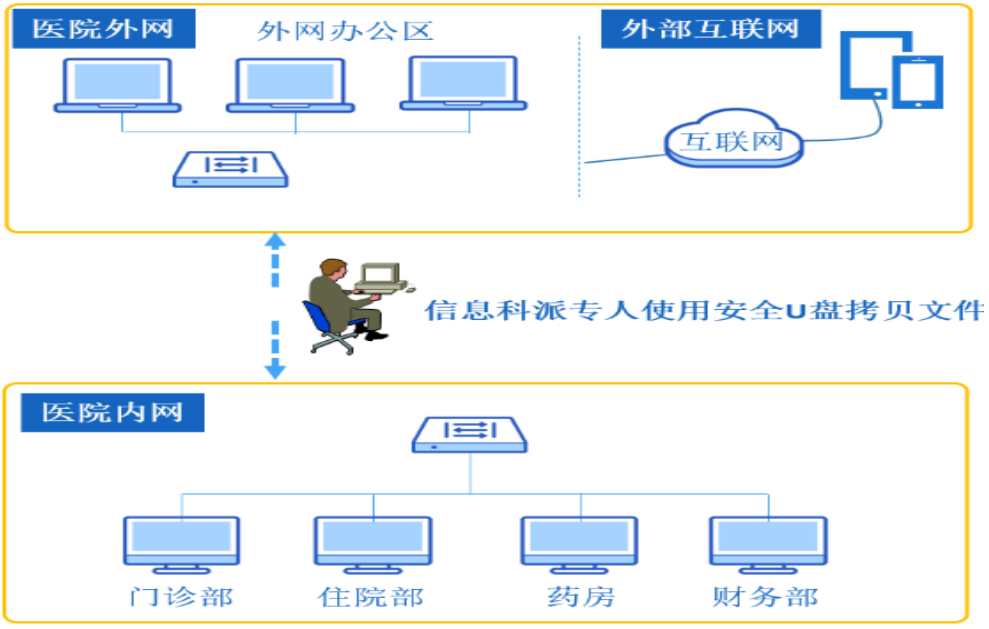
医院内网部署的 HIS，LIS，PACS 等业务系统，运维工程师经常需要在内外网之间传输系统升级补丁和日志文件等，这类文件为确保其安全性及隐私性往往需要审核、杀毒后传输，并留档备查，这一系列流程由人工处理特别费时、费力。尤其是在出现故障排除问题需要核查日志时，此方式非常影响业务系统问题的解决效率。

3.数据交互过程，缺少自动化的手段进行安全检查，内外网间病毒扩散风险较高。

之前的人工处理过程，只能依赖工程师手动进行病毒查杀，一旦工程师的操作终端上病毒库没有及时更新，或者疏于执行手动杀毒，就很容易造成带毒文件的遗漏，可能引起内外网之间的病毒扩散。

4.数据交互过程缺少技术审计手段，一旦发生数据泄漏很难溯源。

之前在体检、PCR 检验、质管、药学病案、绩效、财务等科室进行内外网传输的文件中，涉及大量医院内部敏感数据或病患个人隐私数据，虽然有科室主任的审批单据，但是对于在人工进行 U 盘拷贝的过程中仍然有很大的数据泄漏风险，对于文件内容无法检验、对于 U 盘使用终端的去向无法审计追溯。而且由于缺乏技术手段，一旦发生数据泄露，很难从源头查清数据是从哪里泄露的，也很难确定影响范围到底有多大。



医院原有文件传输流程

二、思路方法

为了解决以上这些问题，我院信息中心提出了内外网‘文件交互系统既要方便医务人员的实际需求，又要满足于数据安全的要求’，这样就需要建立一套从数据的发送、安全检查、审批、传输，到全面的审计日志记录、原始数据留存归档这样一整套数据安全交互的管理措施，从而实现内外网数据交互过程的全生命周期管理。对于这些问题，我院信息中心和奇安信集团及南阳理工学院进行合作，共同建立了“跨网文件交换技术创新联合实验室”。实验室的建立旨在结合我院医疗业务场景进行不断的探索创新，不断完善南阳二院数据安全交换的规范化、流程化、自动化。我院和奇安信集团克服了重重困难，投入大量研发人员不断完善逐步满足了跨网文件交互的需求，经过一段时间的测试和优化，现已充分满足了我院的需求，在实现跨网文件的高速传输的同时，又满足了防病毒和权限审批、事后审计等方面的要求。从使用近两年的实际效果来看，医院在数据交互安全方面有了三重提升，完美解决了之前内外网文件传输的问题：

1.简化流程，提升工作效率。

跨网文件安全交换管理系统上线后，在设置审批流程及权限后，业务人员只需在平台登录自己账户，然后通知科室主任审批，审批通过后即可传到内（外）网。各科室的跨网传递文件可并行自动处理，无需等待。不仅大大提升了业务科室的工作效率，也解决了原有繁琐的传输流程,对信息科来说也释放了一个工程师，可投入到科室其他工作中，缓解了信息科人员不足的问题。

从文件交换系统生成的报表来看，自上线以来我院内外网传输次数达到了 4730 次。整个过程中无需信息科干预按预设自动执行，效率提升非常明显。得到了医院工作人员的一致好评。

2.数据交互过程自动化安全检查，提升数据交互的安全性。

跨网文件安全交换管理系统自带的杀毒引擎和可自动更新的病毒库，使信息科无需担心外网电脑的木马病毒通过拷贝的文件带入内网，极大的减少了各类因此造成的安全事件。同时，针对不同的业务数据交互类型，还可设置不同的安全策略和审批流程，进一步夯实网络安全责任。

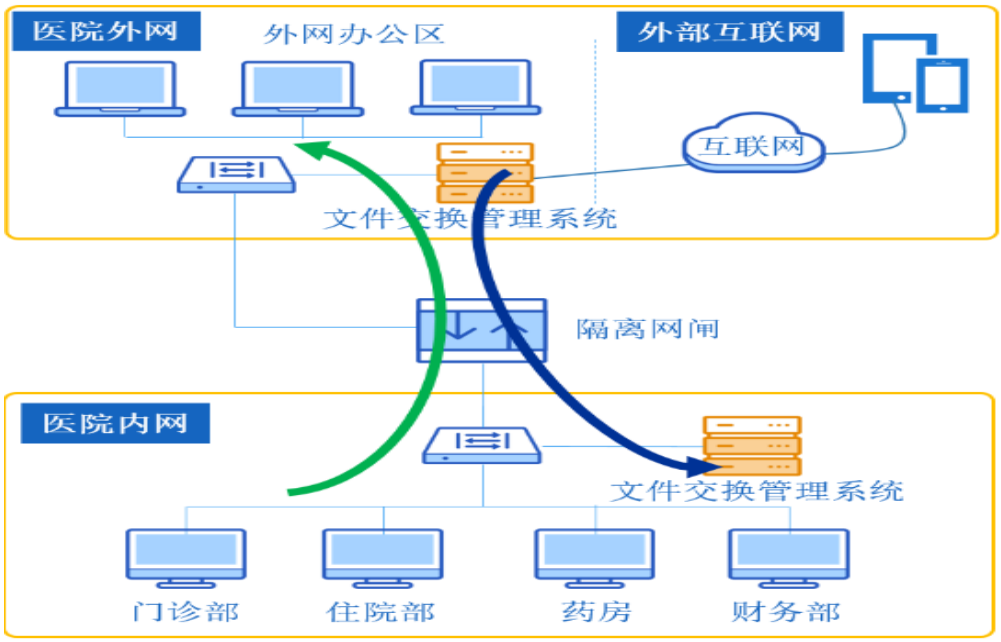
3.对文件内外网交互全过程的审计，提升了发生数据泄漏事件的溯源追查能力，保护了医院内部和个人信息数据的安全性。

在跨网文件安全交换管理系统上线后，业务科室专人登录个人账户到专用空间，内网到专人外网空间，非业务科室人员无法接触隐私数据，很好的保护了隐私数据。不同软件的运维人员在使用跨网文件安全交换管理系统传输文件时，会触发不同的安全审批策略，系统自动进行初步检测，对违规文件自动告警隔离，优化了工作流程，提高了传输效率，并且传输成功后会自动归档保存做到了留档留痕。假使出现违规泄密事件，也可通过强大的日志审计功能，审计原始操作纪录和审计原始文件，快速定位问题所在。

项目建成后一段时间内，文件流转的效率相比原来 U 盘的方式提升 70%左右，一个文件从上传、安全检查、审批、提取只需要几分钟即可完成。大大提升了办公效率。原来文件流转方式需要信息中心频繁介入，目前采用此方案实现文件自动化，大大节省信息中心运维效率以及人员投入。

通过文件交换方案， 医疗客户实现了患者影像资料拷贝带走、医疗检验结果数据流转审批。大大提升了效率。受到业务部门、人民群众一致好评。

此项目创新的解决了不同网络之间、不同安全域之间文件流转管控场景下管控颗粒度粗，易发生信息泄露、传输文件无法审计，出现问题无法追责、不同部门协调难，传输效率低下、文件防病毒等问题。



实施内外网文件安全交换管理系统后文件传输流程

三、技术创新

1.采用前置机模式，支持多网多站点部署，不改变原有网络隔离架构，未经审核许可，数据绝不越界，符合等级保护要求。

2.基于内容与行为的安全检查引擎

可根据文件包创建和发送的行为进行安全策略检查。系统支持对发送文件符合文件名称、扩展名、单体文件大小、总文件大小、操作人及目标人范围等内容条件时，自动触发审核流程。内置防病毒引擎发现病毒后通知管理员自动将病毒转入隔离区。

3.基于动态条件的审批流程引擎

跨网文件安全交换管理系统提供内置的审批审核功能。按照业务的需要建立传输审核审批流程。系统支持审核审批触发条件配置管理，管理员可预制审核流程，当文件名称、扩展名、单体文件大小、总文件大小、操作人及目标人范围等触发了审核条件时，系统将转向指定的人工审批流程。

4.数据防泄漏技术

文件类型识别，不受文件扩展名影响，支持多种文件格式，支持压缩包；内置 OCR 可识别图片的内容，对特定文件类型进行阻断，防止数据泄露，内置 DLP 系统进行内容检查，识别敏感文件，根据识别结果自动阻断或转为人工审核，保护数据安全。

5.高性能传输技术

超大文件智能分段：对大文件进行传输时，并行传输。支持断点续传、支持自动重传、支持文件安全校验。

结语

该技术方案可以为医疗行业提供文件流转管控方案，解决内外网文件传输无法做到事前审核、事后审计的共性难题。该跨网文件交换管控解决方案改变客户原来不同网络之间文件流转的模式，采用全新的文件流转管控模式，替换原来安全 U 盘、共享 FTP 等方式的流转。

通过跨网文件交换项目的成功实施，我们感受到安全和效率并不矛盾，并且还都有很大的提升空间。医疗数据安全作为医院最重要的安全红线之一，需要在复杂的安全形势下不断创新，从而满足在更加复杂的场景下数据交互安全的需要。特别是在信息高速发展的今天，医院信息系统要打破原有观念的束缚，既要满足信息安全工作，也要适应医疗行业高效发展的“互联互通”的需求。

PU-042

新冠疫苗犹豫及其影响因素：系统综述

李晨露[1] 吴泰来[2]
[1]华中科技大学医药卫生管理学院
[2]华中科技大学医药卫生管理学院

1 引言

2019 年 12 月 30 日，世界卫生组织（WHO）宣布严重急性呼吸系统综合症冠状病毒 2（SARS-CoV-2）作为一种新型人类病原体和 2019 年冠状病毒病（COVID-19）的致病因子出现[错误!未找到引用源。]。该病原体迅速在全球蔓延，导致大流行，到 2022 年 6 月造成超过 5 亿病例和 600 万人死亡[错误!未找到引用源。]。目前，抵御 COVID-19 的疫苗已经开发出来， 2020 年 8 月 11 日在俄罗斯、2020 年 12 月 2 日在英国、2020 年 12 月 11 日在美国、2020 年 12 月 30 日在中国授权使用[错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。]。初步估计 70%的人口需要接种疫苗才能结束这场大流行，然而，对 COVID-19 疫苗的犹豫成为实现目标接种水平的巨大障碍[错误!未找到引用源。]。在 COVID-19 疫苗被授权和传播之前，公众舆论普遍存在对已有的 COVID-19 疫苗的接受疑虑和担忧。随着时间的推移，新冠疫苗和相关信息普及开来，目前研究关注的重点已经从评估对 COVID-19 疫苗的预设想法转移到观察行为、测量疫苗的吸收情况以及描述与 COVID-19 疫苗接受度相关的因素。了解与 COVID-19 疫苗的接受和拒绝有关的因素，将对新冠疫苗犹豫这一健康信息行为的研究有所裨益。

2 方法

2.1 搜索策略

本文严格遵守系统综述和 Meta 分析优先报告项目 PRISMA(preferred reporting items for systematic reviews and meta-analyses)进行文献的检索、收集、筛选和信息提取。为了保证数据的检全率，通过中国知网和 Web of science 两个数据库进行检索，并纳入参考文献的施引文献，进而提高文献的检全率。检索时间为 2020 年 1 月 1 日至 2023 年 12 月 21 日。中文检索策略为：

（新冠疫苗犹豫）AND(新冠疫苗)；英文检索策略为：(COVID-19 vaccine hesitancy) AND (COVID-19 vaccination rate)。由于包含研究类型的异质性，无法根据 PRISMA 指南评估偏倚。

2.2纳入与排除标准

根据研究目的，纳入文献标准：(1)文献语言类型：纳入明确文献为中、英文；(2)文献来源：类别选择核心期刊；(3)研究对象：新冠疫苗犹豫的因素；(4)文献主题：内容要涉及 3 个方面的至少一项(新冠疫苗犹豫的决定因素、新冠疫苗犹豫的影响因素、对新冠疫苗犹豫的作用机制)，才能被纳入综述。(5)研究类型：纳入的研究类型包括综述、观察性研究和试验性研究。

排除标准：(1)评论观点、书信、新闻、社论、文献目录、会议摘要、项目介绍、报纸、专利等文章；(2)研究对象单纯是新冠疫苗，未涉及到评价的 3 个维度；(3)无法获取原文的文献；(4)对象是新冠疫苗犹豫的阶段、解决新冠疫苗犹豫的策略等其它问题。(5)在数据库中检索到的非中文和非英文的文献。

2.3 文献筛选与信息提取

首先进行文献预筛查，提前发现筛选标准中的问题及某些筛选人员的理解错误，并及时作出调整和纠正，完善筛选标准；正式筛选时遵循 PRISMA 准则，即依次去除重复、标题、摘要、质量评估标准和内容不符合条件的文献[错误!未找到引用源。]。根据按照排除标准进行筛选，对纳入的文献进行全文阅读进行二次筛选，确认文献与研究内容主题的相关性(新冠疫苗犹豫及其影响因素）。

2.4 文献质量评估

在对文献主题和文献全文进行排除筛选后，需要对研究结果相关规范性和文献的质量进行再次检查，以确定新冠疫苗犹豫及其影响因素。本文所选文献的研究方法为定量研究和定性研究，其中定性研究主要是对文献进行观察以及对患者进行访谈；定量研究主要包括研究目标、研究设计、抽样、结果和结论等流程，依据以上研究方法的标准和质量内容进行综合确定纳入的文献量[错误!未找到引用源。]。

2.5 数据提取

数据提取同样进行预实验，从所有文献中提取用于分析的关键信息，提取的主要内容包括：(1)文献基本信息的描述(2) 新冠疫苗犹豫在各国的情况和影响因素(3)方法学内容(4)主要结果和结论等，然后对所提取的资料进行比较。

3 研究结果

3.1文献纳入结果

如图 1 所示，本文对从两个数据库中检索到的 599 篇文献进行研究，删除重复文献后，仍对文献的内容主题(主要以标题和摘要)进行再次筛选，进一步排除。最终，将 187 篇文献纳入本文的研究范围内并进行了全文阅读，之后又进行了文字质量的评估，从而得到了 13 篇文献作为本文的研究对象。其中中文文献 9 篇，英文文献 4 篇。

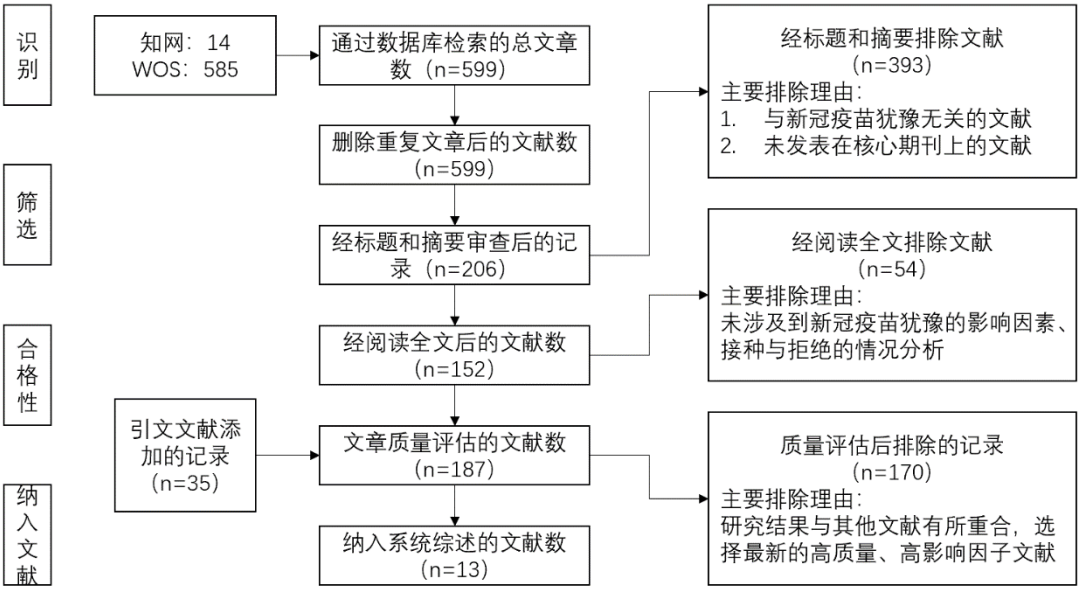


图 1 研究选择过程

3.2 文献概况

纳入的 13 篇文献 3 篇包括新冠疫苗犹豫的决定因素的研究、8 篇包括新冠疫苗犹豫的影响因素的研究，2 篇包括对新冠疫苗犹豫的作用机制的研究。其中 4 篇为定性研究，9 篇为定量研究。新冠疫苗犹豫的决定因素主要由疫苗有效性、疫苗犹豫的患病率、疫苗的可及性和成本三个维度衡量，表 1 为新冠疫苗犹豫的决定因素的描述性统计。新冠疫苗犹豫的影响因素主要为疫苗、个人、认知、群体、信息等多方面，表 2 为新冠疫苗犹豫的影响因素的描述性统计。对新冠疫苗犹豫产生作用机制的主要有社交媒体信息和信息疫情视角，表 3 为对新冠疫苗犹豫的作用机制的描述性统计。

表 1 新冠疫苗犹豫的决定因素的描述性统计信息

纳入文献	研究方法	作用维度	主要发现
Zhang J, et al[错误!未找到引用源。]，2022	定量研究	疫苗有效性	医疗保健专业学生对 COVID-39 疫苗的接受程度更高。疫苗安全性和有效性问题仍然是影响学生接受的主要因素。
Ouyang H, et al[错误!未找到引用源。]，2022	定量研究	疫苗犹豫的患病率	缺乏信心和风险被确定为 COVID-19 疫苗犹豫的因素。年龄、职业状况和收入水平、社交媒体的使用频率、社交媒体使用的多样性、媒体信任和健康信息素养与 COVID-19 疫苗犹豫显著相关。
Truong J, et al[错误!未找到引用源。]，2022	定量研究	疫苗的可及性和成本	发现了影响疫苗接种决策的七个主要因素：影响疫苗接受度的人口因素、疫苗的可及性和成本、个人责任和对疾病的风险认知、在疫苗接种方面采取的预防措施、对卫生当局和疫苗的信任、疫苗的安全性和有效性， 以及缺乏有关疫苗的信息或错误信息

表 2 新冠疫苗犹豫的影响因素的描述性统计信息

纳入文献	研究方法	作用维度	主要发现
刘翰谕等人[错误!未找到引用源。]，2021	定性研究	决定因素矩阵	疫苗犹豫的影响因素可用决定因素矩阵归类,包括疫苗或免疫规划影响、个人/群体影响和情景影响三个方面。
刘晓曦等人[错误!未找到引用源。]，2021	定量研究	"3Cs"模型	感知到的疫苗安全性、有效性以及对提供疫苗的系统的信任度越高，接种疫苗越不犹豫；感知到疫苗接种的必要性以及感染新冠肺炎的严重性越高、越不自满，接种疫苗越不犹豫；感知到的接种疫苗便利程度越高，接种疫苗越不犹豫。
王志伟等人[错误!未找到引用源。]，2021	定量研究	疫苗犹豫的危险因素	居住在城市里的居民更易出现疫苗犹豫。
刘春梓等人[错误!未找到引用源。]，2021	定性研究	疫苗犹豫研究现状	新冠病毒疫苗犹豫受疫苗、个人、认知等方面因素影响
黎泽明等人[错误!未找到引用源。]，2021	定量研究	健康信念模式	政府等相关部门应注意性别、学历、社会心理学因素如公众感知严重性、感知行为障碍及外在客观因素等对疫苗犹豫的影响作用
龙思贵等人[错误!未找到引用源。]，2022	定性研究	不同国家患者的态度和偏好	苗犹豫受多种因素的影响，其中自信度，自满，便利，公众对疫苗、政府、媒体和卫生服务工作者的信任，疫苗可及性、可负担性和可获得性，集体责任感及疫苗来源也会影响疫苗犹豫
宁婧等人[错误!未找到引用源。]，2022	定性研究	不同国家家长的态度和偏好	家长疫苗犹豫主要受家长个人因素、儿童因素、疫苗因素、信息因素四个方面的影响。

He K, et al[错误!未找到引用源。]，2021	定量研究	手机	智慧医院建立健全安全管理制度、完善信息安全防护技术、落实信息安全等级保护制度，可提高患者的就医体验。

表 3 对新冠疫苗犹豫的作用机制的描述性统计信息

纳入文献	研究方法	作用维度	主要发现
崔家勇等人[错误!未找到引用源。]，2022	定量研究	社交媒体信息	关乎个人得失的叙事性信息比说教性信息能更大程度地缓解疫苗犹豫。
李锦辉等人[错误!未找到引用源。]，2022	定量研究	信息疫情视角	负面信息是国内家长新冠疫苗犹豫的核心因素，个体对疫苗信息的风险感知对于疫苗犹豫的影响作用仅次于负面信息。

3.3 新冠疫苗犹豫

（1）新冠疫苗犹豫的现状定义

疫苗犹豫，是指在疫苗接种服务可及的情况下拒绝或延迟接种疫苗。面对全球新冠感染疫情的大流行，现在的疫苗接种情况并不乐观，无论是普通人群还是医护人员，其新冠疫苗犹豫程度一直维持在较高水平[错误!未找到引用源。错误!未找到引用源。]。

（2）新冠疫苗犹豫的原因

犹豫的主要原因为[错误!未找到引用源。]：（1）怀疑疫苗的有效性和安全性；（2）认为接种疫苗是不必要的；（3）没时间接种。另外，面对保护效益明显的加强针接种，公众仍存在较高的疫苗犹豫[错误!未找到引用源。]。

3.4 新冠疫苗犹豫的影响因素

多种因素影响着新冠疫苗的疫苗犹豫。促进疫苗犹豫或接受的七个主要因素：影响疫苗接种的人口因素（种族、年龄、性别、怀孕、教育和就业）、可及性和成本、个人责任和风险认知、基于疫苗接种决定采取的预防措施、对卫生当局和疫苗的信任、新疫苗的安全性和有效性，以及缺乏信息或疫苗错误信息[错误!未找到引用源。]。生物医学研究界从病原体鉴定到成功的疫苗开发和生产的过程不仅在速度方面而且在公众知名度方面都是前所未有的。人们认为疫苗 “仓促”、技术 “太新”，SARS-CoV-2 可能会自行“烧毁”其提供持续的自然免疫力，这加剧了新冠疫苗犹豫的现象产生。的研究发现，反疫苗运动在疫情期间愈发壮大，而反疫苗运动的流行会显著增加公众的疫苗犹豫[24]。此外，媒体大量的虚假新闻报道、卫生工作者表现出的新冠疫苗犹豫和政府相关部门不当的疫苗策略制定与管理均损害了公众对疫苗的信心[错误!未找到引用源。]。而盲目自满以及较低的收入等也增加了公众的疫苗犹豫[错误!未找到引用源。]。

4 讨论

4.1 主要调查结果

本文进行了一项系统综述，以了解在疫情期间促进疫苗犹豫或接受的因素。通过阅读文献整理分析，总结了影响疫苗接种决策的七个主要因素：影响疫苗接受度的人口因素、疫苗的可及性和成本、个人责任和对疾病的风险认知、在疫苗接种方面采取的预防措施、对卫生当局和疫苗的信任、疫苗的安全性和有效性， 以及缺乏有关疫苗的信息或错误信息[12]。这些因素既不能简单地以因果关系来解释，也不能相互独立。这些因素中的每一个在影响新冠疫苗犹豫的决策方面都起着至关重要的作用。此外，新冠疫苗是在反复出现的大流行浪潮中开发、制造、测试、授权、分发和首次接种的。与一般疫苗接受或犹豫相比，新冠疫苗接受或犹豫相关的一些因素被夸大了，包括使用较新的制造技术，大流行控制的许多方面的政治化，以及公众和媒体对新冠疫苗研究和实施的各个方面的紧密关注。

（1）疫苗的新颖性

对于疫苗 “仓促”、技术 “太新”等问题，疫苗生产者应公开新冠疫苗开发流程，反复验证疫苗的有效性，以消除公众的疑虑。

（2）安全问题

除了对罕见但直接或近似副作用的担忧外，一些新冠疫苗犹豫源于对未知的长期副作用的担忧。其中包括对社交媒体上广泛传播的将新冠疫苗与女性和男性不育症联系起来的谣言对生育能力的影响的担忧。疫苗生产者应针对不实言论做出处理，拿出有力证据打消公众的疑虑。

（3）关于易感性、严重性和有效性的观念

众所周知，这种疾病在老年人群中往往更为严重，尤其是与儿童相比，并且有证据表明儿童被感染的可能性较小[错误!未找到引用源。]。因此，公众普遍认为新冠是儿童中相对良性的疾病，导致家长对儿童进行疫苗接种的存在抵制或疑虑。针对这种情况，建议政府有关部门通过发布相关推文等手段鼓励家长让孩子接种新冠疫苗，同时要求防疫站和医院等疫苗提供者不仅有能力储存疫苗，而且还要积极向患者家属推荐疫苗接种。

（4）健康素养

随着全球 COVID-19 变异株病例随着疫苗供应而持续上升，接种新的新冠疫苗比如加强针显得尤为重要。向不同群体宣传疫苗接种的重要性，提升公众的健康素养，广泛提升疫情防范意识至关重要。

4.2 结论

本系统综述研究了新冠疫苗犹豫及其影响因素，通过梳理 13 篇文献了解了新冠疫苗犹豫产生的原因、决策因素和影响因素。疫苗犹豫降低疫苗接种率，破坏免疫规划成果，危害人群健康。接种新冠疫苗是大势所趋，为了保护自己，也为了保护身边的朋友家人，我们都应该以更广阔的胸怀接纳它。政府和疫苗生产者可以从不同层面采取针对性措施，改善新冠疫苗犹豫，早日形成群体免疫。

4.3 局限性

本研究存在一些不足之处。一是纳入的文献是近 4 年的研究，可能有些文献有所遗漏，会存在检索不全的情况。二是纳入应用效果的评价标准不全面，本文只是纳入了新冠疫苗犹豫的决策因素、影响因素和对新冠疫苗犹豫产生作用机制的因素，其作用维度涵盖并不全面。三是纳入的文章质量可能不高，后期可以扩大并精确纳入标准，纳入更高质量的文献进行综述分析。

PU-043

医院信息安全与患者隐私保护研究

陆杨 邓成生
镇江市精神卫生中心

中国自古有讳疾忌医的传统，患者不愿意让别人知道自己得了病、进了医院、做过手术，尤其是一些特殊病症，如精神类疾病和传染病，也不愿意让别人知道自己的亲属生了病。根据《中华人民共和国个人信息保护法》和《中华人民共和国民法典》的相关条款要求，个人隐私保护越来越重要，尤其是患者的就诊治疗情况。因此对于医院而言，如何保护患者隐私，确保重要信息不被第三方恶意窃取和使用，让人民群众满意，是医院运营的一项重要工作。

医院的信息化相对于其他行业，起步晚发展快，各类系统使得医疗机构便捷了工作，然而技术是双刃剑，电子病历中包含了大量敏感的个人基本信息、就诊记录、用药情况等。这些信息如果泄露则可能导致患者隐私被严重侵犯，会造成相关第三方的不当得益，也会造成患者本人面对来自社会上的歧视以及经济上的损失，更造成医院遭到不必要的起诉和卫生系统的行政处罚，从而导致信誉受损。因此医院有必要通过各类技术建立高效、可靠、经济的信息安全防护体系，建立有效的管理制度和规章制度提升全院职工的信息安全意识。

1 通过先进和有效的信息化技术保障数据安全和患者隐私

1.1 医院信息安全和患者隐私

医院的各类信息系统主要有如 HIS、LIS、CIS、RIS 等，这些系统记录着海量的患者挂号、诊断、治疗数据，同时这些数据也是患者的隐私。有别于早期的纸质病历时代，如今随着信息化时代向数字化时代的过渡，大量此类数据得以被广泛收集和长期储存，甚至实现了大区域的规模化数据交换，这些数据不仅涉及个人的方方面面，如个人家庭住址和医保信息，还能通过数据关联覆盖到个人的所有诊疗信息乃至家族成员的信息。

医疗信息最核心的电子病历数据是医院医护人员进行诊断、治疗、护理乃至医院管理层和上级部门决策的重要凭据，其完整性、准确性和可信度极其重要。任何篡改和非法访问窃取都有可能导致医院自身不恰当的治疗乃至影响其他医院获取患者健康信息的误解，从而让医院面临内部的责任纠纷和医患纠纷乃至外部的司法诉讼和行政处罚，因此医院信息安全关乎到医疗质量、患者安全和医院的日常运营。保护患者个人隐私是可以防止患者的个人信息被恶意盗用，避免他们可能遭受的

身份名誉损失、电信金融欺诈或其他不法行为[1]，所以患者对他们个人的健康信息是否被不法人员获取有着合理的担忧，只有确保他们的个人隐私信息安全，才会让患者放心，这对于患者、医院乃至整个医疗行业都有着非常重要的现实意义。

1.2 终端安全和系统访问控制

终端安全是医院信息录入和防护的第一道防线，该终端不仅限于医护人员的台式机、平板和手机，更涉及各类电子医疗器械和对外访问或数据交换的堡垒机。关闭非必要端口，用精简化、必要化、最小化原则只保留业务端口，所有终端纳入统一管理机制，即漏洞扫描、病毒库升级和准入管理等，确保所有终端是经过授权并只能访问所允许的网络资源。

系统访问控制同样也是遵循最小化原则，医院各类信息系统品种繁多，权限复杂，信息部门应当汇合其他所有部门定期对所有服务器的工号和应用系统进行筛选，及时停用外出人员和离职人员的工号，对所有人的工号权限进行检查，对弱口令强制进行更改并实行定期更改制度，完善身份认证授权系统，部分重要系统应当使用电子密钥或手机动态口令。

1.3 数据传输安全和容灾备份

信息孤岛已是过去式，目前医院避免不了与外界单位的数据交换，尤其是和医保中心、上级数据质量监管部门、软件运营商等相关单位的日常化、规模化数据交换。数据传输安全的核心是网络安全，当根据现有网络拓扑图，合理设定线路，在设定的网络信息安全防护体系下对各层级进行优先级调度，集系统防护和智能化管理等进行多层级协同防护，确保信息流不被截取和篡改。做好网闸的数据流向设定，使用商业密码保护数据库数据安全。外部人员访问内网应当通过授权进行，非必要不得使用 U 盘等外界设备，相关文件进入内网必须经过文件交换站进行杀毒检查，堡垒机访问应当仅开启必要的端口，维护人员操作时采用一次性符合标准密码进行访问，任何修改记录应当进行日志记录。

数据备份是被动应对网络攻击的重要制度，医院应当做好数据备份工作，重要业务系统进行热备、双备、冷备和异地备份，全备份和增量备份按照实际情况设定，数据备份的磁盘应当用商业密码加密并将解密口令分开存储，定期验证备份数据是否完整，是否可以及时恢复。

1.4 做好主动防御工作和态势感知

在医院信息部门日常工作中，应当化被动为主动，主动了解外界网络安全行情，做好安全审查工作，对于平时网络上突然出现的漏洞和病毒，应当进行系统性的安全审查，评估和验证医院所有信息系统的安全性，包含系统配置、访问控制、数据加密等领域，确保所有网闸、防火墙、网络策略处在最佳实践状态。在有必要的基础上，可以将主动防御接入人工智能技术，减轻网络运维的压力[2]，提高医院的网络安全防御等级。

态势感知是一种基于环境的，动态、整体地洞悉安全风险的能力，是以安全大数据为基础，从全局视角对网络中的数据流进行实时监控，对安全威胁进行发现识别、响应处置。在实践的过程中，应当设定启动高级防护的各类触发点，自动和手工相结合，识别和清理各类非正常操作和安全风险威胁。定期对计算机网络信息安全风险做好评估工作，并制定相关响应处置措施。

1.5 建立独立的审计系统

经事实证明，建立独立的审计系统是非常有必要的，独立的审计系统在即便信息系统遭受攻击和篡改时，依然能独立于主系统调出事件的整个记录过程，是信息系统安全的重要组成部分。审计系统应该部署独立的线路，由专人负责，此举可以有效查询是谁访问了哪些信息、何时访问、访问的终端设备等相关信息。该类记录可以有效地用于事后安全事件的分析和调查，更好地了解相关事件的发生经过和影响关联，通过可追溯性，恢复数据的完整性和真实性。此外审计系统还能记录系统的异常情况并进行分析，如登录失败多次、越权访问、数据泄露风险等，定期分析此系统有助于对院内信息系统安全进行修复和防范。

1.6 对患者隐私数据做分类层次化管理

医疗大数据中的患者隐私权涵盖了人身权和财产权，其数据意义不言而喻，如何保障数据既能够被保护又能被合理研究和使用，对数据进行分类层次化管理即是重要的解决方案之一。其分类可以按照对于个人而言和对社会而言进行横向和纵向进行划分，横向是按照患者个人隐私的不同类型，分为一般数据、特殊数据、敏感数据[3]。纵向是对不同患者以及需要数据的不同机构进行分级，即低敏感数据、较敏感数据和高敏感数据。在保护原则上，首先是采集源进行限制，即和诊疗无关的数据不予采集，从源头上对患者隐私的损害降到最低，其次是公益优先原则，如院感数据，要及时上报给指挥中心，在内容展示上，软件对不同权限人员进行重要数据的脱敏展示。总体上搭

建一个定义化、统一化的数据管理模式，并形成医疗行业内标准，最大化保障患者隐私数据仅供特定人员参阅和研究。

1.7 安全技术和成本经济达到平衡

信息安全防御等级并非越高越好、越贵越好、越先进越好，应当针对医院的未来发展情况，有效地制订发展规划，而非不顾实际，盲目全方位采用国内乃至国际最先进的技术和设备。医院应当合理有效地根据本院实际情况，查漏补缺，在经济成本允许的情况下，实现信息安全的有效保障，以期安全技术和成本经济达到一个合理的平衡。

2 通过有效的管理制度让全院职工高度认识信息安全和患者隐私的保护

2.1 完善信息安全工作机制和规章制度

医院是一个公共服务的特殊机构，承担为实现健康中国的战略任务，承载着让人民远离病痛的崇高使命，俗话说：三分技术，七分管理。一切技术需要人工来管理去执行。在信息安全的问题上，应当加强各科室职能划分，信息安全不仅仅是信息科的事情，更需要全院职工的努力，一切为了患者，保障他们的隐私信息不被泄露，在对患者进行诊断、资料、护理的过程中，应当格外关注患者隐私保护，不在治疗和护理过程中谈论患者病情[4]。在管理上，需要建立严格的奖惩制度，强化责任落实机制，制订有效的突发事件预案，定期举行安全知识竞赛和网络攻防演练，不定期由院领导牵头进行安全审查，将相关成绩纳入个人绩效和年终评优机制中。

2.2 提升全院职工的信息安全意识

医疗信息安全意识不仅是医务人员的事情，行政和后勤人员等其他类型人员也同样不应该轻视其责任意识。各医院应当定期开展全员信息安全培训，并形成记录，以便交付给新员工阅读和执行。培训内容包含医院信息安全的基本内容和边界，患者隐私的重要性和相关的法律法规要求，播放相关医院信息安全和保护患者隐私的宣传片、教育片，从源头上制止内部人员有数据泄露和数据篡改、伪造的想法。同时针对不同类型医院职工展开分别的教育，尤其是信息科、医务部等与医院数据密切相关的科室，坚持安全意识宣传不动摇，安全管理常抓不懈。

3 结语

综上所述，各类医疗信息系统已经是医院运营不可或缺的工具，其应用也变得更加广泛和深入，医院信息安全与患者隐私保护也成为当下医院亟需关注的重要领域。医院需要在信息安全防御等级和经济成本相适应的基础上建立有效的终端安全设置和系统访问控制、数据传输安全和容灾备份机制、主动防御和态势感知部署、独立的审计系统、对患者隐私数据做分类层次化管理等，并完善院内各项工作机制和规章制度，提升全院职工的信息安全意识。这样可以最大限度地保护各类信息系统主要是电子病历信息的安全性和隐私，确保患者个人隐

PU-044

实验室生物安全风险及信息应对策略的探讨

刘婵桢 陈池梅
重庆医科大学附属第一医院

随着传染性非典型肺炎、禽流感、新型冠状病毒等新发传染病的流行，各传染病原菌在医务人员中感染中时有发生，生物安全实验室[1]的管理显得尤为重要。目前我国生物实验室数量多、分布广、监管流程繁琐，可能出现的风险环节较多，依靠传统的人工管理方式无法实现精准化管理，如何进行实验室生物安全信息化建设，是建立标准化管理和预防风险发生的有效手段。

1 实验室生物信息安全风险分析

实验室生物安全指的是当实验室作为科研和工作的场所时，防止危险生物因子造成实验室人员暴露、向实验室外扩散等引起的一系列危害，而采取的相关综合措施[2]。如何确保检验、检测实验室生物安全，保障实验室研究顺利进行，是目前管理层面对的主要问题，医学检验的内容主要为新陈代谢物，而这个代谢物都部分携带细菌和病毒，若再进行检验时发生操作失误，可能会发生交叉感染，造成生物安全问题[3]。

1.1 生物安全管理制度

随着我国生物实验室的建设与应用，我国相继出台了一系列相关法律法规，对生物实验室的安全进行管理。2002 年我国颁布了关于实验室生物安全方面的第一个行业准则《微生物和生物医学实验室生物安全通用准则》[4]，国家卫生健康委员会于 2005 年与 2006 年分别发布了《可感染人类

的高致病性病原微生物菌（毒）种或样本运输管理规定》与《病原微生物实验室生物安全环境管理办法》[5]，2022 年颁布了为生物安全实验室安全管理的工作实践提供了有力法律支撑的《生物安全法》[6]。

但这些法律法规和条例基本只是宏观的指导和规定，对于实验室的日常管理和流程以及相应的规范性要求较少。通过查阅文献发现，国内有关高校实验室安全管理已颁布了相应的法律法规，但这些法规都是指导性文件，并不细化、不具体，需要修改完善，增强可操作性[7]。科研实验室生物安全管理责任不够明确，缺乏专职或兼职生物安全员进行生物安全各项管理制度的建设、完善和落实，没有管理体系作为保障，很难保证实验室的安全运行[8]。卫生主管部门并未出台相应的细则，造成在行政执法和管理过程中对实验室的级别、安全评估等结果各地存在较大差异[9]。

同时实验室工作人员大都缺乏实验室生物安全相关法律与法规的教育和培训[10]，对生物安全防护[11]重视程度低，因此，必须加强工作人员教育和培训，提升生物安全防护意识，严格按照生物安全规定操作，做到按流程、按规章制度进行实验室活动，从而提高实验室的生物安全治理能力。

1.2 监管流程

实验室生物安全一般涉及采集、运输、实验、保藏、消毒等众多环节，监管流程比较长，可能出现的风险控制点[12]比较多，从事生物实验相关活动的人员较多并复杂，管理难度较大，依靠传统的逐级分级分类管理的模式很难实现全流程管理。实验室生物安全风险在运输、实验活动、消毒灭菌等环节尤为突出，目前靠行业专家或卫生监督管理部门的常规检查或飞行检查，只能发现制度或设备等问题，无法实现全程监管及实效性，而且一旦出现安全隐患无法实现预警预测。在运输和实验活动过程中，生物样品一旦丢失或被盗被抢，如果无法追溯，将引起社会恐慌。

随着近年来高致病性呼吸道病毒的爆发，检验的需求大幅增加，PCR(Polymerase Chain Reaction，聚合酶链式反应)实验室大规模的建设，造成 PCR 实验室地域上较分散，监管难度较大，有必要将各地区 PCR 实验室联网并纳入统一管理，实现对实验室工作环境、区域、工作流程、人员防护、实验室检测等各个环节及风险点进行全方位管理和监控，对可能带来生物安全隐患的各个风险点进行预警，及时发现生物安全隐患并排除隐患，保证实验室的生物安全[13]。

1.3 风险指标

当前各省、市、自治区的医疗监管机构对于生物安全实验室安全管理涉及的指标较为零散，以定性为主，评价过程中主观人为因素较多，若要实现客观、真实的评价，必须依靠科学的决策依据[14]。风险点的管控确定有多种方法，可依据 ISO15189 中的风险管理、CLSI EP18 中的识别和控制实验室错误来源的风险管理技术、CLSI EP23-A 中基于风险管理的实验室质量控制计划等相关要求。还可引入 ISO31010 中风险管理-分析评估技术中推荐的头脑风暴法、危险与可操作性分析、危险分析与关键控制点法、风险矩阵等相关分析手段。同时还可以参考相关文件，根据中国疾病预防控制中心制订的《突发事件公共卫生风险评估技术方案（试行）》及澳大利亚和新西兰制订的《风险管理标准》AS/NZS 4360-2004，确定实验室生物安全的风险点。

对实验室进行生物安全管理时，可能涉及到风险点包括实验室设施布局及相关配置、实验室生物安全设备、实验室医疗废弃物处理、实验室感染控制、实验室菌毒种及样本管理、实验室人员、实验室个人防护及应急装备等，这些分险点又可细化为二级风险点，形成几级风险体系。

1.4 生物安全数据

实验室生物安全数据来自多个渠道，围绕实验室人、机、料、法、环样本全生命周期，需对涉及到的生物安全数据进行实时采集，整合采集到的数据并进行数据标准化治理，形成生物安全监控和分析数据，通过数据分析和建模运算，构筑多元生物安全信息数据底座，从而实现生物安全相关因素的多点触发预警，多渠道 BI(Business Intelligence，商业智能)可视化监控和运控，多维度预测、分析和评估。

同时生物安全数据还涉及定级定密的日常管理，软件系统需安全运行与维护并进行权限管理，涉密数据还需做好备份和销毁管理机制，防止数据意外泄漏[15]。

2 实验室生物安全信息应对策略

随着现代技术的发展，利用信息技术加强实验室的生物安全管理已成为重要的管理措施之一，通过系统建设与技术手段，提醒实验室管理人员及时发现潜在的安全隐患，并做好相关的处理，可较好地评估与预防风险，保障实验室生物安全。

2.1 风险预防和管理

对实验室的温度、温度、有毒有害气体、烟雾浓度及生物安全柜等重点设备参数进行远程联网监控，实时显示实验室环境监测数据，实现实验室环境的网络化、可视化、智能化监控和远程管

理，一旦实验室发生安全事件或风险，可以通过短信等方式发出预警信息，以便第一时间采取应对措施，同时利用系统中实验室相关设施设备提供商，在线为实验室提供技术指导和物质保障，实现防护口罩、防护服、检测试剂等应急物资的快速调配，有效减少安全事故的发生和事故造成的损失。

2.2 全流程闭环管理

建设可追溯的实验室生物安全活动监控平台，对样本采集、使用、运输、保存、销毁全流程进行智能监管，确保一样本一码，全程可追溯，实时监控去向，对样本设立安全区与逾越安全区并可快速识别报警。利用智能监控摄像头建立覆盖开展生物实验室活动的的联网系统，应用人脸识别技术，对实验室人员进行准入控制，并利用 AI（Artificial Intelligence，人工智能）技术对实验室活动或危险进行智能识别，实现实验操作等关键环节的实时监测。通过 GPS（Global Positioning System，全球定位系统）定位技术、RFID（Radio Frequency Identification，射频识别）技术、远程温度传感器、5G（Fifth-generation，第五代移动通信技术）通讯等技术，改进样本的运输装置，建立覆盖区域的运输监控系统，掌握实时动态，防止被盗和丢失。

2.3 信息技术

信息化管理在实验室生物安全中应用范围广泛，便捷程度也不断提高，如系统架构采用基于物联网的智慧实验室设计，并采用视频图像的人工智能识别技术实现实验室安全智能化管理，通过无线传感网络技术实现对实验室环境的智能监测[16]。运用区块链和云存储等信息化新技术，对实验室重要安全要素、控制区域、检测流程进行模块化设计与应用，实现实验室活动记录和见证材料全程留痕可追溯，结果不可篡改，并具备统计分析和预警功能，提高实验室安全管理的自动化水平[17]。

适用于实验室活动各个环节生物安全的信息技术较多，各实验室也在不断完善系统和探索新信息技术的应用，从 GPS 定位技术和 GIS 地理信息管理技术、RFID 电子标签技术，到角色访问控制技术，最近的物联网技术、AI 技术、数字中台和区块链技术等，使实验室生物安全的信息化程度越来越高，解决了信息化手段不足等问题，提升安全风险预警能力，推进实验室生物安全信息化管理水平，减少资源浪费和人力成本，提升安全管理效能。在信息技术应用效果方面，还有许多值得研究的关键问题，如多维数据融合风险分析技术、AI 融合实验室安全警报研究、自主可控实验室生物安全集成技术、实验室生物安全风险数据阈值分析、根据生物安全法律法规与实验室活动管理规范建立风险识别和评估系统等。

2.4 风险自动识别、监控和评估系统

建设实验室内部的应急准备、监测、研判、处置及再准备的闭环管理平台，全面提升实时监测，定期评估、研判，实时通报预警，快速应急处置和精准干预指导的应急能力。平台可分为风险设定、风险处置、风险评估、改进计划四大部分，并与相关信息技术相结合，并打造决策知识库，支持实验室生物安全知识推荐、支持预案处置推荐和支持相似事件参考，并生成生物安全评估报告并用于指导风险点设置，从而实现实验室生物安全的闭环管理。通过闭环管理不仅从信息化角度解决实验室生物安全问题，还可促进实验室人员加强生物安全法律法规的意识与知识，通过决策知识库不断完善，形成实验室自己的生物安全风险预防细则。

3 结语

我国开展的生物检测的实验室数量逐年新增，涉及区域广，检测数量大，生物安全风险防控事关国家安全和发展，事关社会大局稳定，且关乎科学研究的安全运行。有必要充分利用先进的信息技术，切实解决手工与全流程监管等困难，实现安全风险预防、预警、预测和评估的自动化，并实现安全事件快速处置等，推动生物实验室高质量可持续发展。

PU-045

大数据时代的医疗信息隐私保护管理 SWOT 分析

侯梦薇①*
西安交通大学第一附属医院

1 引言

在现代化的步伐中，信息技术依靠其迅猛的发展经常给人们的生活带来了巨大的改变，这其中又以信息系统大数据的个性化带来的便利更为显著。社会各单位与机构利用个人上网行为记录、消费购买记录、就诊记录等大数据信息开展相关研究并制定决策[9]，这些研究与决策将有利于社会的进步和发展，这一点在医疗行业尤其显著。随着医药卫生体制改革的逐步深化，医疗信息系统将分散在医疗机构各部门的数据进行整合，并集中进行存储，使提取医疗数据的过程更加方便，从而满足用户对于医疗数据的使用需求。医疗信息化改变了医疗机构的现有工作流程，提升其工作效率和服务质量[10]。我国大部分三甲医院目前已经形成了一套以费用管理为中心的医院管理信息系统（HIS），医疗信息化已成为医疗活动中必备的工具。

医疗信息系统中存储了患者的基本信息、健康状况、疾病发展、诊疗过程等信息，包含了大量患者隐私数据，而因其易传播、易复制等特征，也使患者的隐私保护出现了新的问题。现阶段，我国医疗信息隐私保护的相关管理仍处于探索阶段，主要运用在访问控制和数据加密等方面，但在数据共享、数据发布等方面的隐私保护还涉及不足。此文在目前存在的医疗信息隐私保护研究进展和现状的基础上，运用 SWOT 分析法，对我国医疗大数据隐私保护的内外环境、机遇挑战进行分析旨在探讨适宜于我国医疗大数据隐私保护管理工作持续有效发展的新策略。

2 医疗大数据的隐私保护

2.1 医疗大数据中的隐私风险

医疗大数据具有“5V”特点，即大容量（Volume）、高速率（Velocity）、多类型（Variety）、高价值（Value）和真实性（Veracity）[3]。为了有效分析和利用数据，研究人员通过数据挖掘、数据分析等技术对大量医疗数据进行分析，这也随之带来了隐私泄露的问题。

在大数据概念出现之前，大部分隐私保护方法主要面向普通规模数据，而普通数据隐私保护方法在被应用到医疗大数据隐私保护时存在许多局限，如医疗大数据的多类型性使以访问控制、匿名化为代表的的传统隐私保护技术难以生效，医疗大数据的实时性需要令传统的加密技术收效甚微。此外，由于医疗大数据新技术的运用，使医疗大数据中的隐私信息泄漏变得越来越容易。因此，医疗大数据时代下的患者信息隐私保护管理工作将面临更大的挑战。

2.2 医疗大数据隐私保护管理模型

怎样在保护患者隐私信息的前提下从医疗数据集中提取有价值信息是隐私保护管理需要解决的关键问题[11]。针对这一问题，研究者们提出了各种保护隐私模型，模型包含各类算法保护患者的隐私信息和他们的隐私标准 。如图 1 所示，隐私保护模型介于可信任的数据持有者与不可信任的用户及数据挖掘模型之间，隐私保护模型通过特定的算法对医疗敏感数据进行处理，在满足患者信息隐私保护的条件下保证发布数据的可用性和获取性能最优的数据挖掘模型。

图 1 医疗大数据隐私保护管理模型

2.3 医疗大数据隐私保护管理现状

在医疗大数据应用管理中，医疗服务提供者特别是大三甲医院是管理着大量患者敏感信息，在对个人隐私保护中发挥着至关重要的作用[8]。医院在医疗大数据隐私保护所采取的的管理措施，主要包括：患者知情同意原则，调阅者权限限制、调阅内容限制、调阅地址局限于医疗卫生服务场所等。然而，很多医院隐私保护管理仍然存在“缺乏相关经费支持”、“法律法规依据欠缺”、“安全技术手段不足”等问题[12]。

3 SWOT 分析

SWOT 分析法，又叫态势分析法[13]。SWOT 分析法的分析属性包含了受自身内部因素影响的优势和劣势，和受外部因素影响的机遇和挑战。医疗大数据隐私保护管理本身是一种新理念、新理论，是时代发展的需要，也是医学发展的需要，其管理重点在于保障患者隐私的前提下对医疗大数据进行有效利用[14]。近年来许多大型医院将创建研究型医院作为自己的发展目标，因此，利用信息化手段对医疗大数据进行有效的隐私保护管理渐渐成为医院管理者关注的焦点[1]。本文以“医疗大数据”、“医疗隐私保护”和“SWOT”为检索词，在知网和万方数据库进行检索，查找并收集中国医疗大数据隐私保护 SWOT 分析相关文献，共检索 76 篇，经过筛选之后选取其中的 63 篇。在收集文献的过程中发现关于医疗大数据隐私保护的 SWOT 分析文献很少，相关文献主要集中在对大数据发展现状和存在问题的分析。本文对 63 篇文献中关于中国医疗大数据隐私保护管理的优势（Strength）、劣势（Weakness）、机遇（Overlook）、挑战（Threat）进行分析，获取有价值的策略信息，并发现解决问题的方法，得到综合评价结果，为医院决策层的管理工作提供有效依据。

3.1 优势（S）

3.1.1 大型医疗机构医疗数据存储方式较为规范

基于数据安全性的考虑，大型医疗机构通常在独立内部局域网的环境下部署数据库存储服务器，存储服务器上安装企业级专业版杀毒软件，且在内网入口和出口地址部署防火墙。专业的杀毒软件可以查找潜伏在系统深处的病毒并进行彻底性杀除，而防火墙就如同在网络闸口构建了一堵高强，杜绝了很多网络黑客的恶意入侵，屏蔽和隔绝病毒等危险程序，基本能够保证数据存储的安全。

3.1.2 隐私保护管理技术发展较为成熟

医疗数据隐私保护技术针对执行方式的不同主要包括数据匿名化技术和访问控制保护技术。采用匿名化技术保护医疗信息时，主要通过对属性值的泛化或抑制，达到隐私保护的目的，代表技术包括 *k-anonymous*、*l-diversity* 和 *t-closeness* 等[4,5]。访问控制隐私保护技术根据事先分配好的用户权限级别对不同的用户选择恰当的身份认证手段，主要代表技术包括基于角色的访问控制（*Role-based Access Control*）等。以上技术目前已经发展比较成熟，能够保证医疗数据内容的真实可用性和隐私安全性。

3.2 劣势（W）

3.2.1 缺乏统一隐私保护标准

虽然各家医疗单位都在逐步完善适应自身的数据隐私保护管理办法，但是目前的现状是医疗大数据的隐私保护管理在很大程度上取决于各家医疗机构的重视程度，却没有一个统计的隐私保护数据标准对现有医疗大数据进行规范要求。目前大多数医疗机构在医疗数据发布和共享时仅将标识患者标识符信息（姓名、身份证号等）进行去除[6]，但是这种方式并不能很好地达到隐私保护的目的。

3.1.2 缺乏相关技术人才

医疗数据隐私保护是隐私保护领域知识和临床医学领域知识的深度融合，这项技术起步时间较晚，且对领域人才的专业性要求很高。长期以来，医院的信息化建设工作得不到重视，缺乏系统的人才培养和团队建设体系，大量信息化开发依赖于外包公司，信息管理部门工作人员忙于对医院现有信息系统进行运行维护，对隐私保护技术不了解，对医疗大数据隐私保护管理不重视。人员知识结构的缺陷会影响医疗大数据的隐私保护管理能力。

3.3 机遇（O）

3.3.1 国家政策大力支持

近年来健康医疗大数据的管理和应用受到了高度重视，国家出台了大量医疗数据管理相关政策文件，从国家层面给予了极大的支持。2016 年 6 月，国务院办公厅发布了《关于促进和规范健康医疗大数据应用发展的指导意见》，对医院信息安全、人员资质、监督管理等进行了规范，明确提出了医院要实施第三级信息安全等级保护等规定。2018 年 5 月 1 日，全国信息安全标准化技术委员会正式发布的规范《信息安全技术个人信息安全规范》正式实施，标准针对个人信息面临的安全问题，规范个人信息控制者在收集、保存、使用等信息处理环节中的相关行为，保障个人的合法权益和社会公共利益。

3.1.2 医疗机构对于医疗大数据隐私保护管理需求迫切

医疗大数据的深入挖掘和研究对于加速医疗技术发展有着重要的意义[7]。但是，组成医疗数据的基础为患者的隐私数据，若此类医疗数据被泄露，将给患者的尊严、人格、声誉和生活造成极大影响，甚至会引起严重的道德甚至伦理问题。因此，科学有效的隐私保护管理一方面对患者的隐私数据进行保护，另一方面能够为医疗技术的发展提供有力的保障和支持，使医疗工作者和临床科研人员充分利用医疗大数据，对推动医疗技术发展有着重要的意义。

3.4 挑战（T）

3.4.1 缺乏相应的法律法规

从法律角度来看，医疗信息是患者个人信息的重要组成部分，包含患者诊断、治疗、用药、体检等医疗过程中涉及到个人健康状况、健康状况、遗传基因、病历病史、治疗过程等方面的信息。由于国情的特殊性，我国在患者隐私保护方面的政策制定起步较晚，目前只有《执业医师法》、《护士管理办法》等很少内容涉及隐私保护相关内容，且并无有针对性的隐私保护政策办法。

3.4.2 隐私保护理念尚未普及

大数据时代下的社交性网站使人们借助网络平台分享日常生活，各种文字、图片、视频等形式的个体化表演成为获得社会关注的重要方式，这无疑淡化了全民的隐私保护意识。医学事故信息的上报统计、医学专业知识的交流与分享、医疗科研数据的共享等[2]大量医疗大数据由于个人隐私保护意识的薄弱会导致了大量患者信息碎片被累积、关联，最终导致医疗信息隐私的暴露。

4 建议

目前我国医疗大数据隐私管理处于迫切需求和起步发展时期。基于上文中内部优劣势、外部机会和威胁的分析，不难看出医疗大数据隐私保护管理目前处于机遇与挑战并存的初期发展阶段。因此，及时确定当前和未来发展方向，对于我院未来的发展意义重大。结合前文分析结果，我认为医疗机构的医疗大数据隐私保护管理，应该从以下几个方面入手，寻求突破：

4.1 明确医疗数据使用过程的规章制度

医疗机构管理部门应出台完善的医疗数据共享标准及数据保护办法和规章制度，明确工作人员对信息的使用权限，避免出现医疗数据泄露或数据用于不正常途径等现象的发生。在医疗数据统计申请过程中，医疗机构应针对自身情况建立一套“事前授权、事中控制、事后审计”的医疗数据控制管理流程。这些管理办法需要契合医疗大数据庞大集中的特点，能够防范非法访问等安全隐患，提升隐私保护管理有效性。

4.2 推进医疗数据统一管理平台建设

医院信息系统以医疗业务系统为主，且系统繁杂，且由不同开发商负责建设，各业务系统产生的数据存放在不同服务器，不利于统一管理。因此，医疗机构信息部门需尽快推进医疗数据统一管理平台的建设，将多信息系统产生的分散存储的数据进行整合，从数据库层面打通系统之间数据的通道，实现数据按照安全级别分级存储管理、数据细粒度访问控制管理、数据处理过程日志记录管理、数据输出时脱敏处理、数据输出日志审计管理以及数据泄露防护技术。

4.3 加强信息部门技术研发能力

医疗大数据的隐私保护管理水平与医院信息部门的技术水平密切相关。医院管理部门在制订预算时需加大数据安全技术研发方面的投入，建立专门的医疗数据信息安全管理部门，并引进隐私保护领域的专家针对医院具体情况制定医疗数据隐私保护技术方案。此外，定期举办数据隐私保护技术相关培训，提升医院现有信息技术人员的相关技术能力和水平。

4.4 提升医务人员的医疗数据隐私管理意识

医疗机构工作人员是医疗大数据使用的主体，也是实施医疗大数据隐私保护的主要责任人。医疗机构应加强对医务人员隐私保护相关知识的培训和宣传，充分发挥医院宣传部门的作用，加强医疗数据隐私保护政策和制度的推广与宣传，转变医院管理人员和医务人员的观念。在编制的医疗数据相关的规章制度时应落实医院内医疗数据安全管理相关要求，与医务人员绩效挂钩，充分调动全体医务人员的数据隐私保护管理意识。

5 总结与展望

本文基于 SWOT 分析法，对大数据时代下我国医疗信息隐私保护管理的优势、劣势、机遇和挑战进行了分析，并通过分析结果对未来的医疗数据隐私保护管理的发展提出了相关建议，从规章制度层面、技术措施层面、人员配置层面和实际应用层面提出了四点医疗数据隐私保护对策和建议，为医疗数据隐私保护管理的未来发展提供参考。

针对医疗大数据隐私保护，未来有几个方向值得进一步研究。首先，从理论和实际两方面对医疗大数据发布时构成的隐私威胁进行研究是非常重要的，这需要识别和建模隐私攻击，并考虑患者医疗信息的医学背景数据，并对其可行性进行评估。其次，医疗数据集的复杂性和规模大幅增长，也为现有隐私保护算法的效率构成了重大挑战，如何在 GB 甚至 TB 级别的数据集上降低计算开销也值得研究。最后，目前并没有一种适用于复杂数据共享场景的隐私保护方法，无法以一种高效且保护隐私的方式集成不同患者医疗数据集且无法以安全的方式满足不同用户的查询请求。一种未来的发展前景是：以底层的健康医疗大数据服务为基础，各个医疗机构之间组成相互依赖、相互支撑的隐私保护服务体系，继续优化和创新针对健康医疗大数据的隐私保护技术，并通过技术手段和相关政策法规相结合，建立一套与之相适应的保障体系。

PU-046

医疗数字化背景下的信息安全建设探析

敬力源 邹小红
川北医学院附属医院

1 引言

随着计算机科学技术的快速发展，医院业务规模的不断扩张，医院实现了高度信息化，医疗信息系统成为了优化就医流程、提升患者就医满意度、提高经济效益和管理水平的强有力支撑[1]。众所周知，信息化水平越高，信息安全就越重要，信息安全的需求就越大。因此，本文将对医院信息安全相关建设进行讨论，以期为医院信息安全保障体系提供相对健全的理论支撑，推动医院高质量发展。

2 数字医疗信息安全问题剖析

本节将结合 2023 年中国医院协会信息专业委员会(CHIMA)发布的《2021-2022 年度中国医院信息化状况调查报告》（后续简称为 CHIMA 报告）、2022 年 9 月发布的首部《网络安全人才实战能力白皮书》[2]（后续简称为白皮书）、瑞数信息发布的《2021 Bots 自动化威胁报告》和《2022 Bots 自动化威胁报告》展开叙述。

2.1 院方重视程度不足

信息科属于行政后勤科室，地位相对弱势，信息安全建设很大程度上受到医院领导的安全意识影响[3]。站在领导的角度，相较业务系统，信息安全对医院实际业务水平、患者就医满意度的提升没有直观感受，抵御了多次来自内外部的网络攻击、解决了多少系统漏洞、防止了多少次数据泄露都没有直观呈现。因此在每年的预算中，领导更偏向对临床、科研、医技等方向的投入。从 CHIMA 报告中的两幅图（如图 2-1 和 2-2 所示）不难看出，信息安全建设在医院的重视程度还处于非常低的状态。

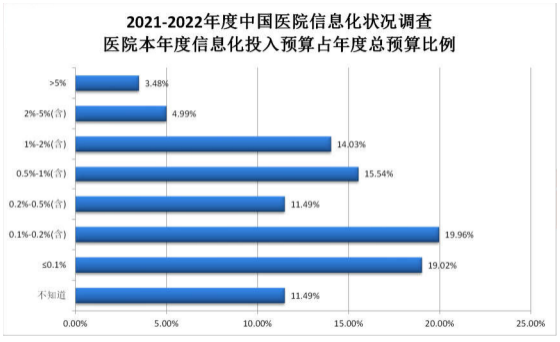


图 2-1 医院本年度信息化投入预算占年度总预算比例图

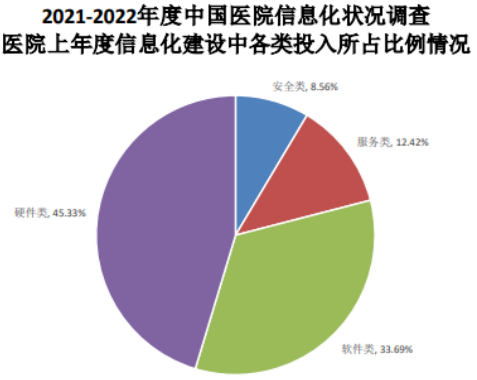


图 2-2 医院上半年度信息化建设中各类投入所占比例情况

2.2 信息安全管理体系不成熟

CHIMA 报告中还给出了 2-3 和 2-4 两幅对比图。从图 2-3 中可以看出，医院目前的信息安全建设主要还是围绕防火墙、入侵检测、网闸等方面，态势感知、WAF、防毒墙等安全性建设较差，数据库审计和堡垒机等安全防护手段也是近几年才开始大规模使用起来。结合图 2-4 不难得

知，大多数医院都只是为了满足对应等级保护制度的基本要求，从而增加对安全设备、密码机、终端管理软件的采购而已，缺乏整体的信息安全规划蓝图与建设思路。

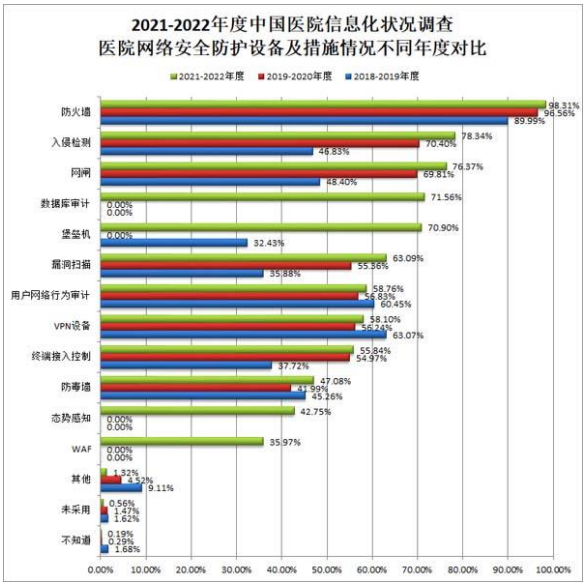


图 2-3 医院网络安全防护设备及措施情况不同年度对比

图 2-4 医院开展等级保护工作情况不同年度对比

2.3 医院内部安全人才缺失

白皮书中明确指出，我国网络安全人才年培养规模在 3 万名左右，到 2027 年人才缺口将达 327 万[2]。目前医院信息科的大多数职工来自信息系统与管理、生物医学工程、计算机科学与技术等相关专业，还有部分职工是医技护专业人员转岗过来的，他们对于信息安全的敏感性较低，对信息安全认知不足，容易出现信息安全事故处置失当的情况[3]。综上所述，医院内部信息安全人才的匮乏程度不言而喻。

2.4 相关信息安全法律政策匮乏、行业统一标准缺失

数字医疗作为一个新兴的互联网医疗产业，虽然近几年相继颁布了《网络安全法》、《数据安全法》、《个人信息保护法》等法律法规，但仍然缺乏专门针对医疗行业的信息安全法律法规。现行法律政策对医疗活动中产生的信息没有明确地进行定义、分类、划分权限和责任，因此在实际操作中存在着诸多疑惑[4]。

医疗行业是涉及个人隐私最多、最深的领域，没有统一的行业标准，容易导致各级医疗机构对数据安全、信息安全的认知不够，很难认识到信息安全对自身业务的深刻影响，导致领导和普通职工对该领域都不重视。

2.5 医院信息系统支离程度较高

医院信息系统大多采自不同的医疗软件公司，医疗数据需要在这些不同的信息系统之间进行流转。从病人挂号到出院的全流程来看，其个人基础数据和医疗数据需要在挂号系统、医生工作站、检查预约系统、排队叫号系统、检验系统、PACS 系统、护士工作站等多个系统之间进行流动，而信息的流动是通过应用程序接口(Application Programming Interface, API)来传递的，但 API 开放的数量越多，系统受攻击的概率就越大[5]。

据全球知名 IT 调研机构 Gartner 发布的研究报告称，2022 年大约有 90%的 Web 应用受到来自 API 的攻击，90%以上的数据失窃都涉及 API 暴露风险。从瑞数信息 2021 年和 2022 年发布的报告可以看出，医疗卫生行业成为了攻击的热点行业，API 的配置不当和相关漏洞使得 API 也成为了黑客优选的攻击手段。在医疗软件的版本更新迭代中，会产生一些僵尸 API 和影子 API，这些 API 便为攻击者留下了入侵的大门，从而导致数据泄露。

3 数字医疗信息安全建设探析

信息安全的五大领域分别是物理安全、网络安全、主机安全、应用安全和数据安全[6]，本节将探析在上述五大领域如何进行信息安全建设。

3.1 物理安全

物理安全也称之为实体安全，是信息安全中最基础的一环。在技术保障维度上需要满足 GB 4943-2001、GB/T 20271-2006、GB50174-1993 等标准，完善对中心机房的防雷击、防火、防水、防电磁干扰、防线路截获、防窃听等安全措施，确保信息实体不受到人为或自然因素的危害[7]。在管理保障维度上，采取基于最小权限范围划分的门禁访问控制策略和监控防护措施，并安排人员定期巡检。乙方工程师进入机房需进行严格审批，并在值班人员的陪同下方可进入机房活动[8]。制定有效的应急预案，在事故出现后能够第一时间响应，降低损失。

3.2 网络安全

保障网络安全本质上是保障网络中传输的数据流不受到窃取、篡改、丢失、监听等破坏，网络传输系统不受到攻击，网络服务不中止。医院借助链路聚合、虚拟路由冗余协议等技术增大网络鲁棒性的同时，还需做好安全区域的划分，根据其不同的功能大致划分为内网业务区、外网访问区、隔离区(Demilitarized Zone, DMZ)、安全管理区、灾备区等[7]。

内网业务区运行着医院的关键核心系统，包含大量的患者个人信息和医学影像信息，对网络的带宽要求较高，此外内网还通过专线连接着医保、卫健委、区域医疗平台等，因此还需要部署专线防火墙防止外部攻击。外网访问区主要是对访问终端进行限制和信息过滤，应配置 Web 应用防火墙、入侵防御系统、安全日志审计系统等设备实现对外业务系统的安全防护[2]。DMZ 区域通常部署的是医院官网、预约挂号等业务系统，可以采用流量探针、下一代防火墙、蜜罐等技术保障网络安全。安全管理区一般是部署日志审计、数据库安全审计、上网行为管理、堡垒机等安全设备或软件，方便出现事故后的追踪溯源。灾备区是核心业务系统的容灾备份节点，以便在相关业务数据丢失或者主节点故障时，能够快速完成主备切换，保障核心业务的高可用性。

3.3 主机安全

当前医院存在的主机安全问题主要包含以下几个方面，(1)业务系统采用弱口令；(2)远程管理采用向日葵、ToDesk 等；(3)用户权限不满足最小权限划分原则；(4)账号过期未删除。(5)端口暴露和共享文件夹暴露。主机安全的范围包括身份鉴别、访问控制、准入控制、入侵防范等[9]。其中针对身份鉴别，口令需满足强密码的复杂性要求，或者采用 PIN 码+动态码的形式；针对访问控制应细分用户权限，分配满足用户角色的最小权限即可；针对准入控制应禁止使用向日葵、ToDesk 等远程软件，若确需远程则使用堡垒机、IPSecVPN[10]、SSLVPN[11]等技术进行远程连接；针对入侵防范可以关闭危险端口和不常用的共享文件夹，安装杀毒软件，及时更新系统补丁等。

3.4 应用安全

应用安全是指应用程序使用过程中的安全，在医院主要体现在电子病历、PACS 系统、HIS 系统等方面。与主机安全类似，应用安全也涉及身份鉴别、访问控制、入侵防范、抗抵赖等方面。其中访问控制还可以采用零信任的方式来更好的保障信息安全；抗抵赖可以通过 CA 认证，例如电子病历必须插入带有个人信息标识的 USBKey 和账号密码才能完成签名[12]。

3.5 数据安全

站在技术保障角度，数据安全中最主要的是数据防泄露和数据容灾备份。针对医疗数据防泄露而言，可以采用数据隔离、数据加密、敏感数据管控、日志审计、访问控制、数据分级分类等方式[13]；针对数据容灾备份而言，可以采用全量备份和增量备份两种形式，或者采用云备份的方式将业

务数据备份到云存储上，实现异地容灾备份。站在管理保障角度，还应成立数据安全领导小组、科室数据安全工作小组、制定数据安全应急预案等，健全数据处理的规范化流程，确保数据安全的各项规章制度落实[14]。站在人员保障角度，还应定期开展信息安全相关知识讲座，提升对医院职工的数据安全意识。

4 结语

综上所述，在医疗信息爆炸的今天，医疗信息化不仅需要满足基础业务的需求，信息安全也是其中不可或缺的关键一环，对医院的长远发展可谓影响重大。建立完善的信息安全体系不仅需要提高网络设备的安全系数，还需要加强信息安全人才的培养和管理体系的完善，从技术、人员、管理三个维度覆盖上述五个领域，从而达到医院信息安全对机密性、完整性、可用性的三大需求。

PU-047

医院数据隐私保护模型建立及应用实践

乔阳阳 韩雪峰 程飞
哈尔滨医科大学附属第二医院计算机室

0 引言

根据《“十四五”全民健康信息化规划》内容，进一步推进新一代信息技术与卫生健康行业深度融合，随着医疗信息技术快速发展和电子病历广泛应用，医疗机构积累了大量的多维度医疗数据，医院数据已成为医疗机构不可或缺的医疗资源，针对其安全性问题，国家制定《网络安全法》《数据安全法》和《个人信息保护法》等法律法规，国家层面对数据安全及监管力度与日俱增，对于医院而言，保护患者隐私及医疗数据安全是基本责任[1-3]。医疗数据的潜在利益让隐私数据安全问题面临着极大的挑战，特别是近年来云服务的出现，大量个人隐私数据都存放于网络空间，增加了隐私数据泄露风险，使得医疗机构个人信息面临着内外部的双重安全威胁[4]。

隐私保护模型大多是针对整个领域的通用模型，专门针对医疗数据隐私保护模型相对较少，特别是现阶段医疗信息化加速、医疗应用终端普及、智慧医疗发展迅速使得患者隐私信息泄露更加多元化，需结合医院具体场景，从数据采集、数据传输、数据存储、数据使用、数据共享、数据销毁等多维度、多角度进行分析、评估和建立医院数据隐私保护模型，以确保医疗数据隐私安全，需从专项场景模式逐步向整体数据治理模式转变。

1 医院数据风险分析

1.1 数据风险

医院数据是医院的核心资源，包括患者信息、医疗科研数据、院内职工科室各类办公数据等，医疗数据中包含大量敏感信息和隐私数据，极易遭受黑客攻击、勒索病历记录、网络钓鱼等，患者、财务和人员数据，容易造成数据被盗取、未经授权使用、人为篡改、故意泄露、内部滥用、违反相关的法律法规和隐私保护规定使用医院数据等风险，各场景的数据在使用过程中存在极大安全风险[5]。

1.2 缺乏参考实践

目前医疗行业没有成熟数据模型可供参考，医疗行业数据涉及医疗、教学、科研、办公等诸多场景，各场景需依据数据级别和业务需求采取不同管控措施，存在场景多、应用广、数据安全相对复杂，且在安全治理和实施均属于探索阶段，需融合医院战略、应用场景、数据治理内涵等方面构建医院数据安全模型。

1.3 数据分类分级框架不完整

医院一般侧重关注某一场景下数据防控能力，没对医院数据进行梳理并建设数据全生命周期各阶段数据分类分级框架，无法实现数据整体防护，数据在使用过程中无法分级分类管理，无法依据分类分级对数据进行差异化处理，不知数据在何阶段需要审批、脱敏、水印等。

1.4 组织架构不配套，自动化工具不成熟

医院数据安全管理部门主要是信息部门，无医院级别多部门组织架构或存在组织架构但主要工作还在技术部门，作为技术部门关注点往往是某一场景安全防护，工作思路和方法都是通过某一安全产品解决某一安全隐患，加之自动化安全产品不成熟，需要大量人工完成数据审批、脱敏、审

计等，工作模式为救火式，很难从本质上改变医院数据安全。如何做到动态防护、精准防护、联防联控的目标，医疗机构面临着巨大的挑战[6]。

2 保护模型建立

各医疗卫生机构按照有关法律法规的规定，参照国家网络安全标准，履行数据安全保护义务，坚持数据安全与业务发展并重，医院数据隐私保护模型的建立需要从管理和技术手段共同应对风险，保障数据安全和数据应用的有效平衡，如图 1：

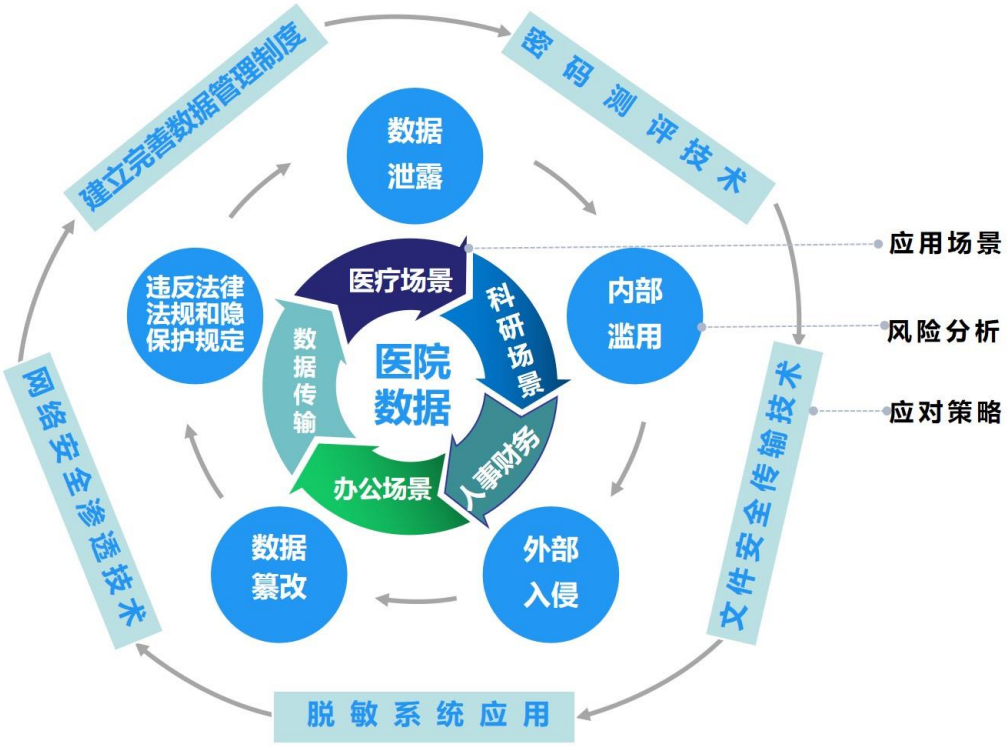


图 1 医院数据隐私保护模型

2.1 管理手段

2.1.1 建立组织机构及数据分级管理，完善数据管理制度

建立数据安全组织管理部，作为数据主管部门，主要工作为进行场景化数据梳理、数据分类分级、管理制度制定等。依据《信息安全技术网络数据分类分级要求》和《数据安全法》国家数据分级标准，依据使用场景、数据内容和安全风险，结合个人隐私、医院合法权益、公众权益和国家安全进行分类分级管理[7]，为数据安全和个人隐私、个人信息保护提供法律保障[8]，对于重要数据和高机密等级数据，采取严密措施进行保护，明确医疗数据、财务数据、人事数据为医院核心数据，按三级数据标准管理；明确教师、学生、监控数据为基础数据，按二级数据标准管理；明确科研教学所用报表等不涉及隐私数据为一级数据，并制定不同级别数据调取、审批流程以及主管部门。制定《数据管理制度》，包括数据采集、数据传输和存储、数据提取等方面，明确各部门责任、义务以及相关数据管理流程。

2.1.2 加强人才队伍建设及培养，完善组织保障体系

《数据安全法》规定：“重要数据的处理者应明确数据安全负责人和管理机构，落实数据安全保护责任”，《个人信息保护法》中设定“个人信息保护负责人”的岗位，负责对个人信息处理活动及采取保护措施等进行监督。在我国网络和数据安全人才紧缺严重，加强数据安全人才队伍建设刻不容缓，定期加强对从业人员数据安全和个人信息保护方面教育和培训，培养高水平数据安全人才，强化数据安全主体责任，增强数据安全意识和素养，提升医院数据安全全方位、全流程保障能力。

2.2 技术手段

2.2 .1 密码测评技术

信息系统普遍存在弱口令、口令和鉴别数据明码很容易被攻击侵入医院内部核心数据库，导致数据安全问题。密码测评技术以突破测试为主，是一种黑盒测试技术，攻击者尝试使用各种手段；如暴力破解、字典攻击等，尝试猜测用户或组织使用的密码，以确定密码的安全性[9]；渗透测试是一种全新的测试，可以模拟黑客恶意攻击，利用人工或工具对网络存在的弱点进行分析，发现传统安全防护方法不能发现的漏洞[10]。结合当前应用系统的运行状况和功能设计，梳理出应该使用密码技术场景；评估采用哪些密码技术和密码算法；建立包括人员管理、密码管理、建设运行、应急处置、密码软硬件及介质管理等制度，编写相关技术文档；根据方案，开展密码技术能力建设，全流程跟进，直至通过密评，帮助用户或组织评估密码的安全性以防止密码被攻击者盗取或猜测。

2.2.2 文件安全传输技术

医院信息网络除医疗业务外采用内外网隔离方式，通过安全认证管理系统禁止使用 U 盘等外接设备，以保证内网相对安全，但同时也无法满足医务人员进行文件调取、拷贝等需求，给员工办公和学习都带来极大不便，为解决医院办公文件传输过程中数据安全需求，通过采用服务器密码机和智能密码钥匙提供密码服务支撑，支持 Hash 算法和数字签名，Hash 算法可以对文件进行处理生成一个固定长度的字符串，通过比对两次计算的 Hash 值，可以判断文件是否被篡改；数字签名则是在文件使用私钥进行签名，接收方可以使用公钥进行验证签名的有效性，从而保证文件传输过程中的机密性和完整性。系统包括用户身份认证、信息加解密传输、文件签名验证、公文收发、统计查询、日志审计等功能。在保障数据安全的前提下，为医院数据传输搭建安全信息通道[11]。

2.2.3 脱敏系统应用

随着医院信息化建设的不断发展，海量的医院数据被采集、存储、传输和利用，一旦发生数据泄漏事件将对医院造成无法弥补的损失，甚至威胁国家安全[12]。如何在流转过程中保证数据的安全性，越来越受到信息管理部门的重视，根据数据隐私等级和风险评估结果，通过自动识别和定位敏感数据分布，并定时、周期、增量化对敏感数据进行发现及管控，提供全方位、高效敏感数据梳理。通过分析敏感数据流量，对数据的流转和使用进行监控，及时发现违规行为，针对性地采取合理安全防护措施。敏感数据分布位置、存在数量高效探查和精准定位，对于运维人员可进行数据全量脱敏、普通用户进行数据部分脱敏、高级用户进行数据全量显示；数据静态脱敏则用于非生产环境的测试、开发、分析使用，以满足不同场景下的数据安全需求。

2.2.4 网络安全渗透

医院网络面临的主要威胁在于黑客恶意攻击以及网络病毒，主要防范的内容在于网络安全威胁，网络安全渗透测试能够及时发现安全漏洞并及时修补，从而避免数据泄露、网络攻击和业务中断等安全事件的发生，提高安全意识，变被动防御为主动防御策略。渗透测试包括对网站和服务器的全方位安全测试，网站安全渗透包括 SQL 注入漏洞、文件上传漏洞、网面被篡改、数据库漏洞等；服务器安全渗透包括 DDOS 攻击测试、ARP 欺骗、远程桌面认证漏洞、FTP 漏洞等[13]。通过模拟黑客攻击的方法，提前检查网络漏洞，形成完整的安全报告及修复方案，及时进行整改，从而确保网站和服务器安全的稳定运行，保障医疗数据安全。

3 效果分析

3.1 规范数据管理体系

建立安全、合规的数据管理体系，制定数据分级管理和访问制度，降低由于外部威胁对个人信息泄露给医疗机构的安全带来严重的法律问题和社会影响，建立个人信息保护合规体系，推动国家医疗信息行业健康发展。

3.2 保证数据完整性

通过有效监管、加密技术、网络安全管控等措施，多方位多层次进行网络安全威胁监控[14]，发现数据传输过程异常活动快速应对问题，确保数据传输的安全性，有效防止数据泄露、网络攻击和未经授权的访问，帮助医院降低医院数据传输的隐私风险，保证数据传输过程中的完整性。

3.3 降低潜在风险

密码测评对于保护用户的隐私和数据安全具有重要意义。对信息系统密评整改 3 次，通过密码产品检测，可以评估产品的安全性，确保商用密码在网络和信息系统中有效使用，提高信息安全防护等级，使核心系统的数据资料等免受泄露、盗窃等攻击行为的破坏，为医院信息系统提供安全、可靠的网络环境。

3.4 强化数据合理使用

随着医疗信息需求的增长和多样化，根据数据使用的特点进行优化调整，针对不同类别、不同级别的敏感数据，采用不同数据分级管理制度，通过不同脱敏策略处理数据，满足医疗机构之间医疗信息的共享利用和科学研究，发挥医院数据隐私保护模型中的重要作用。

4 不足和改进

医院数据隐私在多场景应用中取得良好效果，需要注意的是隐私风险是不可能完全消除，因为技术和威胁都在不断发展。因此，需要进一步改进：1.持续关注最新安全技术和法律法规，及时更新和改进安全措施，以保持数据传输的隐私和安全；2.提升安全意识，通过持续的安全培训和教育，提升全体员工对数据隐私和数据安全的意识和理解；3.增强数据风险意识：云端存储大量患者影像数据，存在极高的泄露风险，建立院级 VPN 进行行为控制；建立数据加密与数据安全防范措施；建立数据访问权限管理，有效地提高医疗数据安全性和保密性；4.持续改进和监测：医院应实施定期的安全漏洞扫描和系统评估，了解当前安全风险状况，并及时采取相应的措施进行改进，持续监测、跟踪技术和威胁的发展，以保证系统稳定性和安全性；5.加强运维人员管理：医院需要加强对第三方运维人员监管，确保数据使用符合相关法律法规要求，防止数据的非法泄露与滥用[15]。通过以上改进措施，医院能够更好地应对数据传输隐私风险，不断提升数据安全保护效果。

5 结语

针对医院数据传输存在的隐私风险，我们采用了多重措施进行联防联控，有效降低数据传输的隐私风险，防止医院数据泄露和滥用，通过对多维度数据隐私进行分析，可以发现数据泄露的潜在风险和滥用问题，可以识别并弥补医院数据安全方面的薄弱环节，加强数据安全措施，提升数据传输、存储和处理过程的安全性。强化数据安全管理，确保符合相关法律法规的监管要求，从而提高合规性水平，获得患者认可和信任。综上所述，多维度医院数据隐私应用实践，为医院提供强有力的数据保护手段，保障患者隐私，提高数据安全性，强有效改进措施，实现医疗数据应用安全合规、风险可控的目的。

PU-048

给医院管理者的几点网络安全建议

韩雪峰 乔阳阳 程飞
哈尔滨医科大学附属第二医院 计算机室

.1 引言

以互联网为代表的医疗服务模式，已经成为医疗服务的重点，伴随着互联网医疗的快速发展，在不断提升医院信息化建设水平的同时，也带来了更大的安全风险[1]，针对医院的勒索、挖矿、医疗信息泄露等信息安全事件层出不穷，医院信息系统已经成为了不法黑客的重点攻击对象之一。在新的形势下，很多医院管理者们都认同网络安全的重要性及关键性，都在思索和探讨如何持续改进和完善医疗机构的网络安全。

2 建议

2.1 管理者应认清当前医院网络安全形势

从 2018 年开始，国内外医疗系统遭受攻击的频率就呈明显上升趋势，新加坡医疗系统遭遇严重 APT 攻击，黑客利用被恶意软件感染的计算机,窃取了 150 万新加坡人的健康记录数据, 其中包括新加坡总理李显龙。国内曾出现过多起医院遭遇勒索病毒的事件，攻击者入侵医院信息系统，导致部分文件和应用被病毒加密破坏，并暴力破解医院服务器的远程登录服务，下载多种挖矿木马以谋取利润，影响医院信息系统的正常运行。

在《2019 健康医疗行业网络安全观测报告》中，被观测的 15339 家医疗健康相关单位中近 400 家单位已经存在挖矿木马[2]。报告还对各省单位存在僵木蠕等恶意程序的占比情况进行了对比分析，其中很多医疗机构，“医疗行业网络安全现状”的风险级别为“较大风险”，即威胁种类较多、攻击频率较高，存在较多安全隐患。

伴随移动技术崛起、互联网+医疗、支付方式的改变、数据共享等需求，医疗机构都在充分发挥信息化手段，探索“互联网+医疗服务”模式，对医疗服务模式进行创新[3,4]。作为医院管理者应充分意识到，在新形势下对医院网络安全提出更高的需求，很多医疗机构的业务系统都成为潜在的攻击目标，同时外部威胁不断攀升，攻击者水平超过了防御水平[5]，医院信息网络如果不在这个时间点上做一个好的安全架构，未来补救的成本将是巨大的。

2.2 医院网络安全现状取决于医院管理者的网络安全观

伴随着《网络安全法》颁布、各种网络安全检查、各类泄密、攻击事件的发生，医院管理者网络安全意识普遍提高，大部分管理者都认同网络安全的重要性，同时也都认同网络安全是医疗机构必须面对和解决的问题。医院管理者的安全观主要有以下几方面特点：不清楚医院网络的风险点、什么样的安全因素会带来最大安全风险；对医院网络安全现状不了解导致自信心不足或过度自信；管理者与医院网络安全部门之间存在网络安全认知不一致，这种不一致体现对医疗机构安全认知程度及重视程度的不同,调查表明，管理者意识和支持的缺乏对网络安全的有效性产生了不利影响[6]。

现在文件要求信息安全是单位一把手负主要责任而不是信息科，《中华人民共和国网络安全法》第五十九条明确不履行相关义务所承担法律责任。医院管理者网络安全观需要与时俱进，彻底改变网络安全只是口头重视，需要切实深入了解医院网络安全状况，了解新形势、新需求下医院网络的安全风险点，真正意识到，网络安全并不只是隐私、财产安全，对医院网络安全来说还涉及人身、生命安全，要保持互联网医疗业务快速良好发展，首先要解决安全问题[7]，网络安全不只是一种防护，对医院管理者来说还是一种责任。作为医院管理者应从管理和技术上不断提升网络安全意识，只有医院管理者的网络安全认知不断升级，医院网络安全才能不断完善并坚不可摧。

2.3 医院管理者应重视网络安全持续投入，进行持续和理性安全建设

当前医院几乎所有业务都实现数字化，同时医院通过互联网与患者交互越来越多，医院在进行病患病情分析、诊断、治疗等多个环节都需要借助信息系统，以互联网+模式的主动医疗健康服务成为医疗健康服务模式的主流[8]，网络安全问题就越来越严峻。目前大多数医院网络安全基本停留在安全设备的堆叠，安全设备均采用“静态”的防御体系，不能满足和应对当前医院网络的安全态势。另外，医院管理者大多数缺乏信息安全意识，常依据主观认知来决定医院网络安全建设方向、重点及安全投入成本，由于其认知的局限性、偏差性和滞后等原因，在医院网络安全建设中常暴漏出以下现象。

2.3.1 符合规定，对医院防护能力过度轻信

一些医院管理者认为医院安全投入只是合规工程，只要满足合规性需求，医院系统防御能力就有保障了。医院管理者通常没意识到信息安全风险是“动态”变化的，攻击是未知的[9]，每天都会有新的攻击，合规性建设不能充分解决新的、动态的安全事件。建议通过专业安全机构，系统全面的对医院网络安全现状进行模拟攻击、渗透测试、检查和评估等方法，根据结果认清医院网络防护能力和明确网络安全建设方向，并采用基于风险评估的方法，应用最新的网络安全标准和业界最佳实践进行网络安全建设，有效地控制网络风险[10]。

2.3.2 认为医院网络受到攻击的几率非常低，不愿在网络安全方面投入过多

一些医院管理者不认为医院会遭受到攻击，认为医院已有一定的安全防护措施，可以起到防护作用，并且医院的服务器非关键基础设施，不会引起黑客的注意，所以对医院信息化建设不够支持。经专业机构调查，在受访的 390 家医疗机构中，2018 年网络安全投入超过 200 万元的医疗机构只占 32%，15.17%的医疗机构投入情况不清晰[11]，由此可见各医疗机构普遍安全投入偏低。在过去时间里，医院频频遭遇病毒勒索，从社区医院、三甲医院到卫计委都中过“招儿”，事实证明此类观点是错误的，医院管理者必须对医院网络安全引起足够重视。

2.3.3 医院网络安全建设误区

医院管理者大多数并非专业网络安全人员，在网络安全建设通常存在以下两种错误观念，一个是只要做好医院外网及边界安全防护，医院内网就安全了；一个是医院信息系统已经与互联网、医保、银行等网络实现物理隔离或逻辑隔离，因此医院网络绝对安全。而事实是，只有相对安全的网络，不存在绝对的网络安全，只有从技术、管理、内部网络、外部网络进行全方位的持续建设，围绕数据和业务生命全周期建立风险评估、主动防御、应急处置、持续检测及响应处置的医院网络安全技术体系，构建全方位自适应网络安全保障体系，弹性应对来自外部和内部的各种威胁[12]，才能保证医院网络相对安全。

医院管理者不但网络安全观要与时俱进，同时，应意识到医院网络安全不是一次投入一劳永逸，需要根据自身特点、国内外安全形势采取针对性投入，并依据等级保护相关标准进行定级、备案、测评，根据测评发现问题及漏洞及时整改，并把安全建设经费应列入年度预算，不应低于单位信息化建设的 10%。

2.4 缺乏复合型安全人才

网络安全人才是医院建设网络安全保障体系的重要保证，医院网络安全工作需要相关人员具备多项专业和知识技能，现代化的信息科要求既懂医学又懂计算机技术、网络技术的复合型人才，当前不仅非常缺乏，而且流动频繁[13]。

医院网络人员的安全技能、对设备网络的熟悉情况、人员的数量、工作的积极性和主动性等都将直接影响医院网络安全，各医院管理者应根据医院信息化规模和网络安全保障需求，不断进行网络安全人力资源评估，逐步增加网络安全专业力量，加强继续教育、交流学习机制，完善人才队伍培养[14]，以提升医院网络安全技术人员技能和专业素养，培养医院自己的网络安全人才。

3、小结

医院网络的安全还有很多方面，如建立良好的安全机制；积极应对、恰当处置和不断改进网络安全理念；不断持续建设网络安全技术体系和管理体系等。总之，医院网络安全对管理者不单纯是管理层面和技术层面问题[15]，可以说是关乎医院成败发展的战略问题。

医院管理者应对医院网络安全提供最大的支持并对网络安全有清晰、正确、全面认识，不断加强对网络安全工作的重视程度；不断完善工作机制；保持持续有效的网络安全建设投入。进一步完善网络安全框架，保障医院信息网络安全，建立安全稳定的医院网络服务平台，为“2030 健康中国”宏伟目标建设安全、可用、可管、可控的医院网络安全环境。

PU-049

大数据时代下医学信息安全与保护

彭 宇
哈尔滨医科大学附属第一医院 信息中心

伴随着医改的不断深化，推进医院高质量发展，医疗机构在发展过程中遇到一定程度困难，而医院的信息安全是医院平稳运行的保障，它的管理水平将会对医院的发展产生巨大影响。医院利用信息手段对患者数据进行统计和管理，所以信息系统的安全对于保障患者切身利益、提高医院工作效率以及维护医患关系都有非常重要的意义。随着信息技术的迭代发展，医院信息管理工作也随之发生改变与创新，要结合医院自身情况，探索新的信息管理模式，更好地适应新发展下的医院管理工作需要。

一、大数据的定义

“大数据”的定义最早是上个世纪 80 年代由美国首先提出，它是指随着现代计算机网络和信息技术的高度发展而出现的海量的数据，这些数据不能用现有的数据处理方法来及时有效地进行分析和处理。这一技术是时代发展的结果，它表明，在当今信息化时代，信息数据的处理方式已经不能适应形势发展需要，必须采用更高效、智能的方式，从数量庞大、类型多样的数据信息中寻找到有用的内容[1]。随着时代的发展，以信息技术为基础的互联网、物联网、移动 5G、云计算等新形势的信息技术应用，运行过程中，产生各种各样、形势复杂的数据。而大数据技术会对数据进行及时的挖掘，并进行分析、归纳和存储，从而产生更大的经济效益，同时为信息管理工作提供一个创新发展的依据。

二、医院统计信息管理的含义及现状分析

（一）医院统计信息管理的含义

医院统计信息通常是指医院内部两类主体信息，即医院自身数据和患者信息。主要内容有：患者的性别、年龄、检查检验结果、治疗方案、职工信息、医院绩效、财务信息等。

随着医院的运行，患者数据将会被实时记录下来，医院的统计数据会及时录入与更新。医院发展进程中，统计信息工作占有非常重要的地位，主要体现在：首先，统计信息管理是贯穿医院发展整个过程中，客观真实的统计信息既可以反映出医院发展现状，又可以帮助管理者及时地找到存在问题，并做出相应调整，对医务人员行为进行有效的规范，提升医疗服务质量。其次，利用统计资料的管理，可以使管理人员更好地了解自己的管理水平，并与外部环境相结合，更好地预测医院的未来发展，从而增强决策的科学性[2]。最后，医院是综合了人力、物力、财力等要素的综合体，进行统计信息管理的时候，可以最大限度地整合已有资源。

（二）医院统计信息管理的现状分析

为保证医院信息系统的安全性和高效运行，医院在多数业务主机上安装杀毒软件，这就要求有很多维护人员来进行安装操作与软件版本升级，这些繁重的工作导致了病毒数据库和杀毒软件无法

得到及时升级，因此，医院的信息系统依然存在着安全隐患。有些医院采用对 IP 地址和 MAC 地址直接连接方式，从而限制了外部用户的接入，但是绑定操作要求系统维护人员逐个进行，所以操作起来费时费力；同时，主机的 IP 地址和 MAC 地址的变更也比较简单，但是也很容易造成地址绑定失败。建立入侵检测系统，能够发现并预警，但由于不能及时采取相应的防护措施，只能作为一种辅助防护；信息数据库的安全稽查系统能够查找出与数据库相连的操作对话的 IP 地址，由于医院大部分主机都是公共的，同一个 IP 地址可能被不同人所使用，很难找到具体的人，也就很难进行问责。所以，上述的安全性维护措施并不能完全解决 HIS 的安全性问题[3]。

1. 信息管理观念不强

在医院统计信息的管理方面，从当前实际情况来看，医院管理者通常把注意力集中在医疗事务和医学研究上，对医院统计信息的日常管理缺乏足够的关注，常常忽视了对统计信息管理的跟踪。医院管理人员对于医院统计信息的管理意识比较薄弱，甚至根本没有信息安全防范意识和保护意识。建立健全有效的信息管理体系，在信息安全遇到风险，知道该如何处理，及时、有效阻断风险，降低损失。此外，医院对自身信息和病人所收集到的统计资料，在进行后续分析时，往往没有做好相应的管理工作[4]。

2. 计算机网络安全问题

在大数据环境下，将其运用于医疗统计信息管理过程，一方面可以加快医疗统计信息的处理速度，另一方面也会因其信息化、网络化的先决条件而带来网络安全风险。

基于计算机、网络技术的医院统计信息管理系统，具有计算机网络传播速度快、时效性强等优点，但也存在技术风险。网络安全问题给医院的信息化管理带来很大隐患。例如，一个病毒的攻击，不仅仅是摧毁了一个信息系统，而且由于它的传染性，如果数据被攻击，它的其它数据也会受到影响。这就给医院信息管理带来很大的风险，降低医院信息化管理效率，严重会造成医院治疗业务瘫痪。

三、大数据时代下的医院信息安全技术要点探讨

（一）提升大数据的信息化水平，明确信息管理目的

构建符合大数据特征的统计信息管理体系，对于提升大数据的信息化水平和实现医院统计信息的及时高效处理具有重要意义[5]。要想建立起一个医院安全大数据平台，首先要从勾勒出一个完整的信息安全平台的全貌入手，这不仅需要医院对自身信息展开有效的收集、获取和管理分析，还需要将原本无法获得的信息在大数据的平台下呈现出来。除此之外，要提升大数据的信息化水平，同时要对信息的安全进行监控，要对危险信息进行有效的拦截和封锁。只有通过大数据来进行信息安全监控，才能保证信息安全体系的全面构建和体系架构的建立[6]。最后，为提升信息化水平，还需要对信息系统中所采集到的大量数据信息展开对应的分析与处理，从这些大量的信息中，分析并总结出一些经验与规律，进而为改善与完善医院管理工作提供相关的数据支撑。

要想让统计信息管理更好地工作，医院从上到下都要转变传统的思想和观念，提高信息意识，对大数据时代特点有一个清晰认识，从管理层着手，这样才能做到上行下效。与国外发达国家相比，医院的信息化程度还存在着很大的差距，而且，由于信息化管理的起步较晚，缺乏足够经验，因此在转变观念之后，医院管理者要明确自己的信息管理目标，并根据医院目前的发展情况，积极地向国外先进管理经验学习，借鉴信息化应用的优点，在持续地学习和实践中，不断地对现行的信息化管理方法进行调整和优化。

（二）提高信息安全管理意识，提升专业化水平

由于信息统计工作的复杂性和广泛性，使得统计人员在面临大数据时代的同时，不仅要具有统计学、医院管理方面的知识，更要具有一定的计算机应用能力[7]。所以，医院应该积极地引进复合型管理人员，用严格的人才考核来选拔人才，实行薪酬制度，防止专业人才的流失，并对现有的工作人员展开计算机方面的专业知识的培训，将内部晋升机制贯彻到底，让每个员工都能充分发挥出自己的价值，提高工作人员对医院的忠诚，打造出一支高素质的人才队伍，为医院统计信息管理工作打下坚实的人才基础。

（三）采取物理安全措施，保障系统稳定运行

加强核心机房的维护工作，核心机房作为信息处理的中枢，必须对核心机房工作环境进行严格的维护，并根据各种信息设备工作情况，对其进行温度、湿度的调整；为核心机房配置多台电力装置，以达到多线供电的目的，防止中央机房停电或供电不稳定；应做好防雷工作，并设置防磁场干扰设备。服务器在信息系统中起着主要作用，如果服务器在运行过程中出现故障，那么就会引起服务中断或者数据丢失，甚至会引起系统瘫痪[8]。所以要保障医院的业务系统在任何时候都能保持运

作，以确保业务系统的安全性与有效性。采用双伺服器，在主机发生故障的情况下，从伺服器可以代替主机工作；配置高品质、高可靠性、可长期稳定工作的 UPS 电源，可确保伺服器的稳定运转。

医院各科室是医院信息化建设的主体，各科室有较多的个人计算机终端。每一个工作站都有各自的工作环境，在工作时要注意防止过热、灰尘、湿度等因素影响。还应该严格控制光碟、软盘等的使用，杜绝 USB 接口的访问，并在系统中设置网络管理客户端，以实现信息设备的监视。

在大数据时代背景下，医院在进行统计信息管理的过程中，要强化风险防范意识，并针对医院统计信息管理系统可能面临的风险，制定出相应对策。比如：在病毒方面，可以采用防火墙及安全评估技术，通过定期进行病毒查杀，可以及时地将病毒隐患清除掉，并且在系统与外界之间建立一道防火墙，对信息数据包中的信息进行有效地监控，保证各项信息与访问设置权限的一致性。在黑客方面，可以采用陷阱网络技术，建立一个网络陷阱，将入侵者引入到受控范围之中，从而降低正常系统被攻击的几率，并及时地排除安全隐患，提升系统风险防范能力，从而保证数据信息的安全[9]。

结束语：通过分析，可以看出大数据的发展为医院的统计信息安全提供了新的契机和新的挑战。医院管理者要提高信息安全管理意识，提升专业化水平，提升大数据的信息化水平，明确信息管理目的，并采取物理安全措施，保障系统稳定运行，从而促进我国医疗事业的高质量发展。本文对医院信息安全工作中所存在的问题进行了分析，并从技术的角度对加强医院信息安全的几种技术方法进行了讨论。以上只是作者个人的一点粗浅看法，希望可以为有关医院的信息安全工作起到一定的借鉴作用，从而保证各家医疗机构信息的安全性。

PU-050

医院信息系统数据资产分类分级探索与实践

王鹏 马现
郑州人民医院

一、基本情况

郑州人民医院始建于 1912 年，历经百余年发展，现已成为一家集医疗、教学、科研、预防、保健、急救、康复为一体，一院多区的集团化、现代化公立三级甲等综合医院。2021 年 7 月医院与北京奇虎科技达成战略合作，从 2022 年 3 月至 2023 年 2 月对我院 142 个信息系统中的 17 个核心系统开展了数据资产识别和梳理，依据《健康医疗数据安全指南》中的分级标准，得出适用于医院信息系统 1-5 级数据分类分级数据表和字段结果，并检测内网流量中数据传输的风险，探索医院数据分类分级的有效模式。

二、主要内容

（一）创新亮点

1.先工具自动识别，后人工审核修正的方式完成数据分类分级。组织医院 8 个职能部门（医务部、护理部、发展规划部、药学部、医学装备部、医保物价管理部、财务部、公共卫生与健康研究所）和相关软件公司，基于《健康医疗数据安全指南》，结合各个医疗软件系统数据库表结构，使用数据资产地图工具对医院关键业务系统数据表和字段进行分类分级打标，通过调研访谈、文件分析、工具探查，多维度了解数据资产，明确数据资产构成、特征、范围及流转情况，核查数据分类分级扫描结果，对其中识别错误、分类分级错误的模式规则进行修改，并二次扫描，最后人工修正分类分级结果，提高分类分级准确性，初步建立了医院数据资产地图。

2.开展内网网络全链路数据安全检测。以旁路部署方式对院内信息系统的数据交换流量进行分析，基于 http/https 网络、数据库通讯协议解析，从资产脆弱性、资产暴露面、账户共用、异常访问等维度，检测 API 接口资产风险，梳理庞杂的应用及接口，

检测敏感数据流动风险，识别接口调用的异常用户行为，分析 API 接口的变化，及时掌握存在的 API 风险隐患，从而实现对 API 数据暴露面的治理。

（二）解决的问题

目前对于医院的数据分类分级缺乏明确细致的指导和标准，为了帮助医院准确、有效的完成医院数据的分类分级工作，迫切需要形成一套可落地的、覆盖医院电子病历数据的分类分级建议和实践指导。我院通过近一年的实践，使用数据资产地图工具对医院关键业务系统数据表和字段进行分类分级打标，初步建立了医院数据资产地图，掌握了哪些是重要数据，哪些是敏感数据，为数据在不同生命周期和不同场景下的安全管控奠定了基础。

（三）技术解决方案

采购并部署了数据资产地图系统（图 1），对目标数据库进行表级和字段级的扫描，获得数据资产，通过自定义正则表达式策略来实现识别手机号，姓名，邮箱地址等个人敏感信息字段内容，通过《健康医疗数据安全指南》梳理出与医院数据配套的分类分级模板，导入系统中，通过算法加快策略优化效率，可以实时检索，同时将检索条件存储为识别策略，并提供元数据检索条件之间的相互转换，以增强实施现场策略运营沉淀的效率。数据资产地图系统支持各类主流数据库,我院信息系统数据库类型主要为 Oracle、SQL Server、MySQL。

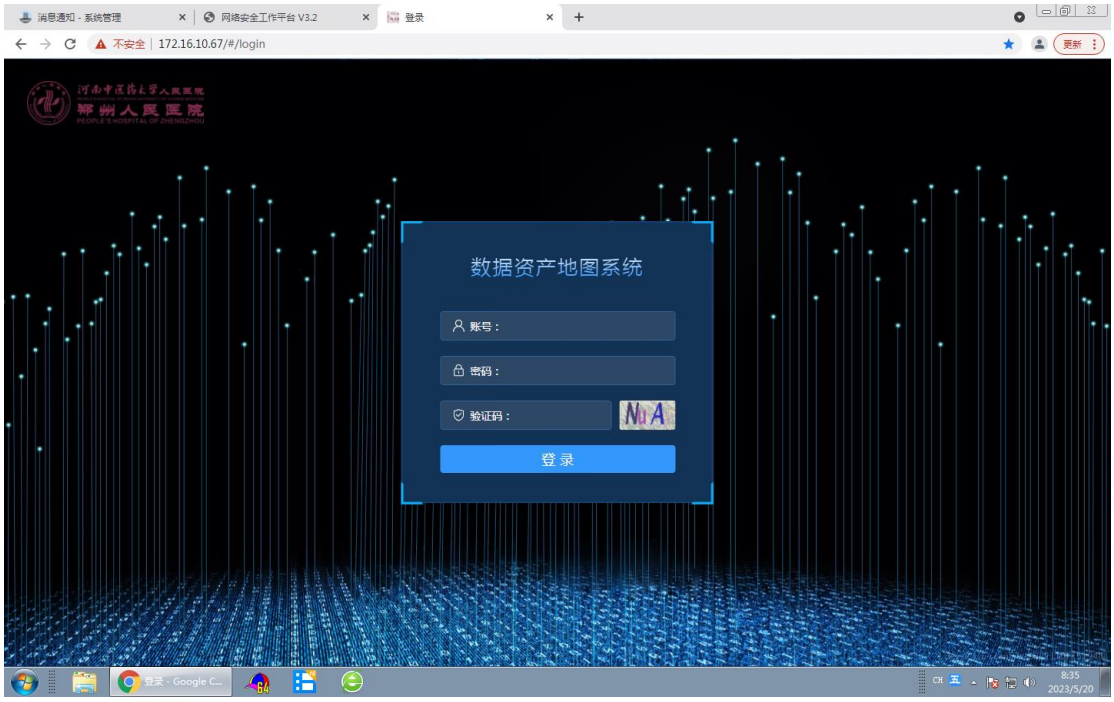


图 1 我院数据资产识别工具

（四）推广使用的机制设计

1、部署数据资产识别工具，数据资产地图系统对目标数据库进行表级和字段级的扫描，通过内置的正则表达式规则模型和自定义的检查、检验、诊断等医疗数据的匹配规则模型，实现了对医疗数据的自动化数据分类分级。

2、梳理全院信息系统资产清单，从系统重要性、数据共享性、厂家配合度三个维度对全院信息系统进行评价，确定开展数据安全治理的 36 个系统，包括 HIS、CIS、LIS、PACS、EMR、患者主索引、手麻重症、康复治疗、院感等系统。

3、我院以 LIS 系统为试点，首先对 LIS 系统进行了数据分类分级。数据资产地图系统系统内置有个人身份信息的规则模型，我们又添加了检查、检验、诊断等医疗数据的匹配规则模型，实现对医疗数据的自动匹配，探索数据分类分级的打标方法。
4、按照 LIS 系统的数据分类分级经验，建立剩余 16 个系统数据扫描计划，在凌晨非业务高峰时期进行数据分类分级扫描。同时为了确保核心业务不受性能影响，将 HIS、PACS 的扫描任务改为对测试库进行扫描。
5、软件公司提供信息系统数据库设计文档，并补充完善数据库表和字段文字的描述内容，梳理数据流向图。对照 GB/T 39725-2020《信息安全技术 健康医疗数据安全指南》对设计文档中的表进行人工数据分类分级。
6、组织医院相关职能部门检查数据分类分级结果，对其中识别错误、分类分级错误的模式规则进行修改，并二次扫描，最后人工审核并修正分类分级结果，提高分类分级准确性。
7、多个维度进行可视化呈现，提供了以数据清单维度的分级结果、数据分类维度的分级结果以及数据分级维度的识别结果，最终形成数据资产清单，形成《郑州人民医院数据资产梳理报告》《郑州人民医院数据资产管理分析报告》，包括：组织架构、数据安全相关政策制度和规范、业务特征、网络拓扑、数据存储情况、日常操作与管理等内容。

（五）取得的成效

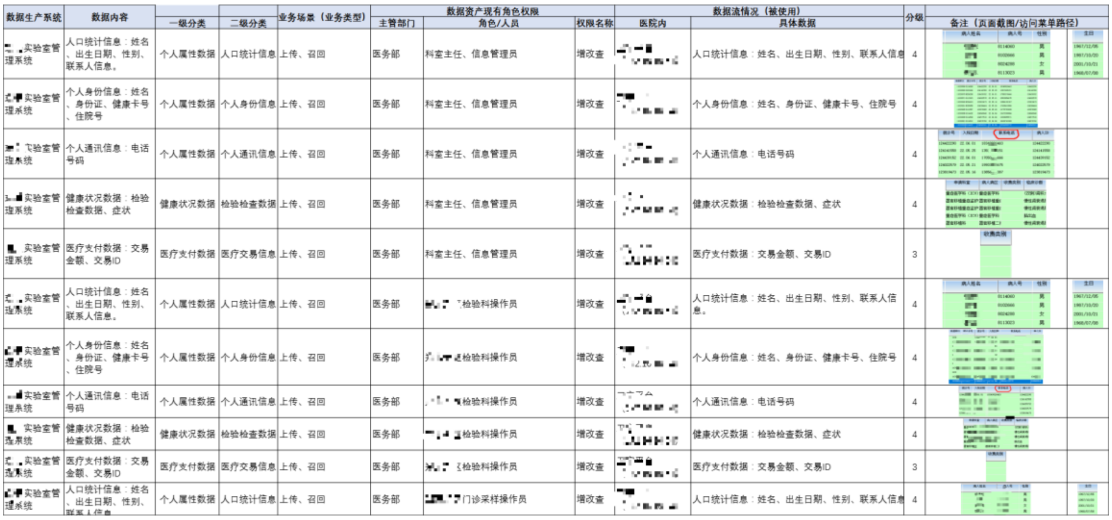
1.建立医院数据分类分级标准和清单

基于 GB/T 39725-2020《信息安全技术健康医疗数据安全指南》，初步建立了我院的医疗健康数据分类分级标准（图 2），形成了持续修订分类分级标准和定期完善数据资产清单的闭环管理机制。

一级分类	二级分类	三级分类	四级分类	* 所属分类
公共卫生数据	疾病检测数据	传染病监测	人口学资料	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	传染病发病和死亡及其分布	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	病原体型别、毒力、耐药性变异情况	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	人群免疫水平的测定	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	动物宿主和媒介昆虫种群分布及病原体携带状况	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	传播动力学及其影响因素的调查	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	防治措施效果的评价	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	疫情预测	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	专题调查（如暴发调查、漏报调查等）等	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	传染病监测	其它	数据分类分级标准-公共卫生数据-疾病检测数据-传染病监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	人口学资料	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	非传染病发病和死亡及其分布	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	人群生活方式和行为危险因素监测	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	地理、环境和社会人文（包括经济）因素等监测	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	饮食、营养因素等调查	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	基因型及遗传背景因素等监测	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	高危人群等确定	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	预防和干预措施效果等评价	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病检测数据	非传染病流行病学监测	其它	数据分类分级标准-公共卫生数据-疾病检测数据-非传染病流行病学监测
公共卫生数据	疾病预防数据		疫苗	数据分类分级标准-公共卫生数据-疾病预防数据
公共卫生数据	疾病预防数据		应接种人数	数据分类分级标准-公共卫生数据-疾病预防数据
公共卫生数据	疾病预防数据		实接种人数	数据分类分级标准-公共卫生数据-疾病预防数据
公共卫生数据	疾病预防数据		其它	数据分类分级标准-公共卫生数据-疾病预防数据
公共卫生数据	出生死亡数据		出生人数	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		出生率	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		死亡人数	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		死亡率	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		自然增长数	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		自然增长率	数据分类分级标准-公共卫生数据-出生死亡数据
公共卫生数据	出生死亡数据		其它	数据分类分级标准-公共卫生数据-出生死亡数据
医疗应用数据	手术记录		手术耗材	数据分类分级标准-医疗应用数据-手术记录
医疗应用数据	手术记录	手术耗材	使用记录	数据分类分级标准-医疗应用数据-手术记录-手术耗材
医疗应用数据	手术记录	手术耗材	其它	数据分类分级标准-医疗应用数据-手术记录-手术耗材
卫生资源数据	医院运营数据	耗材基础信息	耗材基础信息	数据分类分级标准-卫生资源数据-医院运营数据-耗材基础信息
卫生资源数据	医院运营数据	耗材基础信息	耗材产品字典	数据分类分级标准-卫生资源数据-医院运营数据-耗材基础信息
卫生资源数据	医院运营数据	耗材基础信息	耗材明细字典表	数据分类分级标准-卫生资源数据-医院运营数据-耗材基础信息
卫生资源数据	医院运营数据	耗材基础信息	其它	数据分类分级标准-卫生资源数据-医院运营数据-耗材基础信息

图 2 我院医疗健康数据分类分级

审核数据资产清单，确定各个系统字段的数据级别，形成医疗健康数据分类分级及数据资产清单（图 3）,包含 850 个数据表 6358 个字段。



数据生产系统	数据内容	一级分类	二级分类	业务场景（业务类型）	数据资产现有角色权限：主管部门	角色/人员	权限名称	数据流情况（被使用）：医院内	具体数据	分级	备注（页面截图/访问菜单路径）
[redacted]实验室管理系统	人口统计信息：姓名、出生日期、性别、联系人信息。	个人属性数据	人口统计信息	上传、召回	医务部	科室主任、信息管理员	增改查	[redacted]	人口统计信息：姓名、出生日期、性别、联系人信息	4	
[redacted]实验室管理系统	个人身份信息：姓名、身份证、健康卡号、住院号	个人属性数据	个人身份信息	上传、召回	医务部	科室主任、信息管理员	增改查	[redacted]	个人身份信息：姓名、身份证、健康卡号、住院号	4	
[redacted]实验室管理系统	个人通讯信息：电话号码	个人属性数据	个人通讯信息	上传、召回	医务部	科室主任、信息管理员	增改查	[redacted]	个人通讯信息：电话号码	4	
[redacted]实验室管理系统	健康状况数据：检验检查数据、症状	健康状况数据	检验检查数据	上传、召回	医务部	科室主任、信息管理员	增改查	[redacted]	健康状况数据：检验检查数据、症状	4	
[redacted]实验室管理系统	医疗支付数据：交易金额、交易ID	医疗支付数据	医疗交易信息	上传、召回	医务部	科室主任、信息管理员	增改查	[redacted]	医疗支付数据：交易金额、交易ID	3	
[redacted]实验室管理系统	人口统计信息：姓名、出生日期、性别、联系人信息。	个人属性数据	人口统计信息	上传、召回	医务部	[redacted]检验科操作员	增改查	[redacted]	人口统计信息：姓名、出生日期、性别、联系人信息。	4	
[redacted]实验室管理系统	个人身份信息：姓名、身份证、健康卡号、住院号	个人属性数据	个人身份信息	上传、召回	医务部	[redacted]检验科操作员	增改查	[redacted]	个人身份信息：姓名、身份证、健康卡号、住院号	4	
[redacted]实验室管理系统	个人通讯信息：电话号码	个人属性数据	个人通讯信息	上传、召回	医务部	[redacted]检验科操作员	增改查	[redacted]	个人通讯信息：电话号码	4	
[redacted]实验室管理系统	健康状况数据：检验检查数据、症状	健康状况数据	检验检查数据	上传、召回	医务部	[redacted]检验科操作员	增改查	[redacted]	健康状况数据：检验检查数据、症状	4	
[redacted]实验室管理系统	医疗支付数据：交易金额、交易ID	医疗支付数据	医疗交易信息	上传、召回	医务部	[redacted]检验科操作员	增改查	[redacted]	医疗支付数据：交易金额、交易ID	3	
[redacted]实验室管理系统	人口统计信息：姓名、出生日期、性别、联系人信息	个人属性数据	人口统计信息	上传、召回	医务部	[redacted]门诊采样操作员	增改查	[redacted]	人口统计信息：姓名、出生日期、性别、联系人信息	4	

图 3 医疗健康数据分类分级以及梳理数据资产清单

以 LIS 系统为例，经过资产梳理和评估，发现数据库和接口传输过程中存在敏感数据（图 4），针对 LIS 系统向郑州人民医院 APP、微信服务号、支付宝生活号开放数据的范围进行了梳理和管控，降低角色权限，进行数据脱敏，实现了线上医疗数据既安全又合理的流动，线上核酸检测预约已服务约 118 万人次。

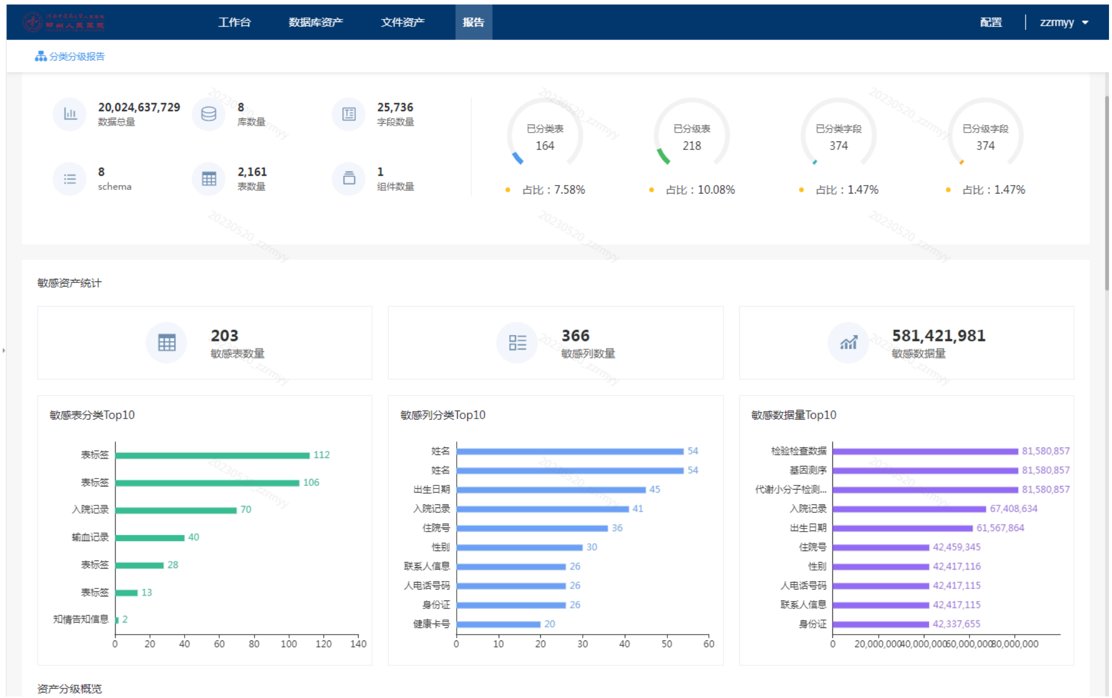


图 4 LIS 系统敏感数据

2.数据安全风险全链路检测

在等保 2.0 基础上，为进一步增强系统安全性，我院开展了敏感数据风险专项检测。

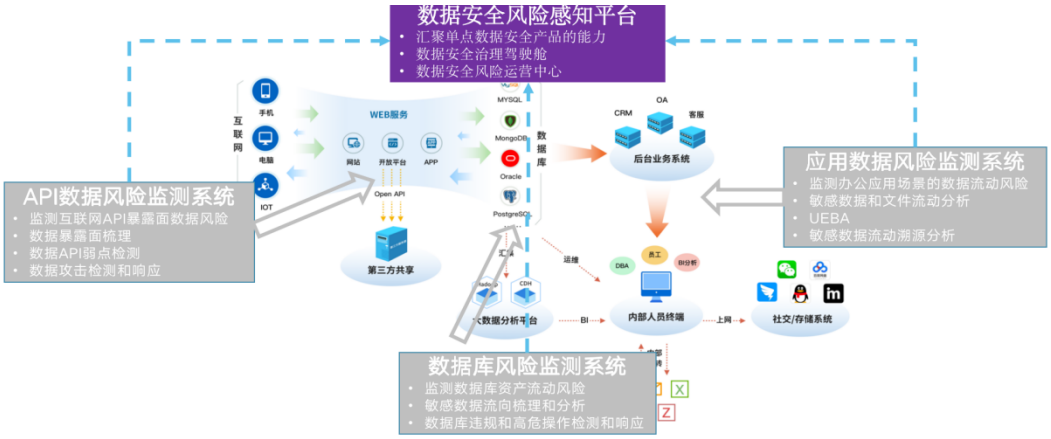


图 5 数据安全风险全链路检测示意图

通过医院网络流量全链路检测（图 5），实现了业务系统敏感数据、敏感文件在开放和共享场景中数据流动流向过程检测。通过流量解析，实现了敏感数据流转和操作异常行为预警；通过流量测绘，弄清楚了数据共享目的地，数据被哪个账号访问，传输到了哪些终端。2022 年 4 月 1 日至 8 日，共发现应用系统 70 个，敏感应用系统 6 个，敏感接口 8 个，存在账号安全、权限安全、数据暴露面方面的问题。流量监控中发现敏感数据 4065 条，其中姓名占比较大为 52.25%，违规敏感数据 11 条（图 6）。

排名	数据标签	数据量(条数)	占比
1	姓名	2124	52.25%
2	手机号	783	19.26%
3	身份证号码	646	15.89%
4	生日	466	11.46%
5	未成年人身份证号	21	0.52%
6	违规敏感数据	11	0.27%
7	IPv4	10	0.25%
8	性别	2	0.05%
9	密码	1	0.02%
10	公司	1	0.02%

图 6 敏感数据占比（去重后的敏感数据占比）

三、下一步工作

（一）存在问题

1、软件公司文档管理不规范，进行数据分类分级时，不能提供准确、全量的表结构，表结构与实际情况不符，甚至缺少中文描述，需要重新安排专人梳理修订表结构。

2、数据资产清单梳理工作涉及厂家多、系统多、流程多、数据种类多，牵涉面广、工作量大，医院和软件公司均存在人员不足的情况。

（二）下一步工作

1、梳理安全策略涉及的数据应用场景，协调软件公司逐步完成改造。针对不同场景、不同应用、不同角色划定适当的安全控制策略，对各个系统的接口进行微调，结合多因子认证、SSL 数字证书、动态脱敏等技术手段，确保数据传输的保密性、完整性。

2、按照 2023 年 5 月 18 日由中关村网络安全与信息化产业联盟数据安全治理专业委员会发布的《数据安全治理白皮书 5.0》提出的医疗数据安全治理路径，开展医院数据安全治理。

PU-051

我国个人健康信息保护现状研究

袁艺峰
中南大学湘雅二医院

1.引言

在大数据和人工智能时代，随着各类医院信息系统和智能终端的应用与普及，个人医疗健康信息的收集变得极其容易。这些信息集中存储和使用形成健康医疗大数据后，广泛应用于患者疾病管理中的各个进程。医疗健康信息已成为重要的数据资源，其应用和发展将深刻影响我国的医疗模式。在医疗过程中，通过医疗健康信息共享调阅可以整合医疗资源，有助于提高医疗服务效率，缩减患者医疗费用[1]。医疗健康信息的飞速发展使得潜在的系统性风险和法律风险逐渐增大[2]。如何在挖掘个人医疗健康信息数据价值的同时，合理保护个人权益和隐私，是当前关注的热点问题。

2.个人健康信息的法律法规要求

从20世纪70年代其，欧美国家已开始推动与个人信息保护相关的立法，现已有128个国家通过立法保护更个人信息[3]。我国根据自身医疗服务的特点制定了相关的法律法规。《中华人民共和国医师法》中明确确定医师需按规范书写病历。《病历书写基本规范》明确指出，门诊住院患者需收集患者的姓名、性别、婚姻、住址、联系方式等个人信息，及医疗过程中的诊断、检验、检查、治疗的详细信息。在患者就医过程中，有大量个人信息被用于提升医疗服务的质量和效率。

随着医疗信息化的不断普及和深入应用，传统的医疗信息保护手段已不在适用于当前环境，个人健康信息保护的问题越来越重要[4]。为适应医疗行业的发展需要，国家出台了一系列的相关政策法规，如《关于促进和规范健康医疗大数据应用发展的指导意见》、《《国家健康医疗大数据标准、安全和服务管理办法（试行）》、《国家医疗保障局关于加强网络安全和数据保护工作的指导意见》等。2021年7月，《信息安全技术健康医疗数据安全指南》发布，该文件从标准上指导如何对医疗建信息进行管理，推动了医疗健康数据保护的发展。

2021年，《网络安全法》、《数据安全法》和《个人信息保护法》相继出台并实施，这三部法律的出台为网络安全、数据安全、个人信息保护提供了坚实的制度保障。《个人信息保护法》中明确了网络中的个人信息的保护责任主体，知情同意等相关要求。《数据安全法》中规定，医疗机构作为患者个人健康信息的收集者，需要承担保护患者个人健康信息的责任和义务。而《个人信息保护法》中明确指出，个人健康信息属于敏感信息，需要对患者进行单独确认[5]。《个人信息保护法》参考了国外立法的相关经验，被认为是相关领域中严格的法律之一。而个人健康信息是患者医疗过程中的重要组成部分，收集和处理个人健康信息的环节繁杂。个人健康信息具有极高价值的同时，也为患者的个人信息保护带来了极大的风险[6]。

3.个人健康信息保护面临的挑战

3.1外部网络环境威胁日益增大

世界卫生组织在2020年4月发表声明：在新冠疫情期间，以医疗卫生机构为攻击目标的网络攻击数据同比增长5倍。卫生服务领域首次成为APT攻击的首要目标。与传统网络攻击模式相比，APT攻击具有攻击行为组织化、隐蔽性、攻击目标明确等特点。

《2021中国软件供应链安全分析报告》指出，通过分析国内各类软件项目，超过八成的软件存在已公布的漏洞问题。攻击者利用漏洞窃取个人健康信息，甚至攻击医疗设备。当前医疗软件厂商应对网络攻击的能力较弱，增加了系统被入侵的风险，个人信息保护更加困难。

3.2医疗行业规范指导有待完善

虽然三步法律相继出台，但是个人信息保护涉及的内容多为参考和借鉴国外立法经验，在国内并无参考，且个人信息保护涉及的范围广，如果没有国家层面对健康行业的权威指导，法律落实较为困难。在《个人信息保护法》公布之前，医疗领域已有各类法规和规定相继公布，这些文件对个人健康信息的收集和处理做出了规定。但部分规定已经不能适应新的法律要求和新的医疗环境。《个人信息保护法》规定，卫生行政部门需承担个人健康信息的保护工作[7]。当前并没有用国家层面的用以指导推进个人信息保护法落实的行业规范或实施细则。

3.3卫生机构的重视程度有待加强

医疗卫生机构作为收集患者个人健康信息的主要收集和处理的主体，对自身保护个人健康信息的责任的重视程度仍然有待加强。患者的医疗健康数据的数据价值对患者就医过程极为重要，收集、处理、存储环节复杂，个人信息保护的管理难度大。调查显示，32.30%的医疗卫生机构没有组织《个人信息保护法》的学习和宣传，59.00%的医疗机构没有根据个人信息保护的相关法律法规建立其适用于本机构的制度。近年来，我国医疗卫生领域对网络安全环境的重视程度和投入有所提高，但是这些投入主要体现在硬件方面。对于个人信息安全的管理制度、工作流程、员工培训等方面的重视程度仍然不够。

4.讨论

4.1 提升网络安全技术水平

网络安全是实现个人信息保护和个人数据安全的重要基石。但是医疗卫生领域的网络安水平仍然较弱，对患者的个人健康信息保护是潜在的威胁。各级医疗卫生机构应该在网络安全宏观规划、系统等级保护、数据分类分级等各方面开展工作，落实好各类规范的要求[8-10]。在人才队伍建设方面，应加强网络安全人才队伍的培养和管理，将数据管理和安全防护两者紧密结合在一起。确保患者的个人健康信息得到合理的保护。

4.2 建立行业规范和加强普法宣传

《个人信息保护法》并非针对某一特定行业的法律，其涉及到各行各业。具有较强的普遍性，且各项关于个人信息保护的规定都是新规定。各方对《个人信息保护法》的规定的理解各有差异，不利于《个人信息保护法》的推广和普及。如果没有统一的行业规范，会导致各类问题，增加患者理解难度和医院管理成本。由于医疗的独特性，必须对相关法律文本进行细化和解释。从国家层面制定和出台相关的行业法规、实施细则等文件，有助于《个人信息保护法》在医疗卫生行业更好的执行。大部分医疗机构也还未根据向相关法律的要求，重新制定和修改医疗数据相关的管理制度和工作流程。国家有关部门需要加强普法和监管，加强对个人信息保护相关法律的宣传。

4.3加强对个人健康信息的管理

医院作为医疗服务的提供者，会在各个环节采集和存储个人健康信息，医院有保护患者个人健康信息的责任。健全的个人信息安全管理制度有助于保障个人健康信息和医疗数据的安全。首先应该强化医院职工的安全意识，医院不止提供医疗服务，还应该保护患者的个人隐私。医院还需建立一整套保障信息安全的管理制度，明确各方

责任，对患者信息的获取、使用都有监控机制。通过健全医院管理制度，落实医疗数据的法律责任。

5.结论

随着信息化技术在医疗行业的深入应用，医院为患者提供了更加优质和高效的医疗服务。与此同时，就医过程中产生的医疗数据也飞速增长，个人信息保护的责任也越来越大。随着国家在个人信息保护方面的立法逐渐完善，医疗领域的行业规范和实施细则也会不断完善。医疗卫生机构作是处理患者个人健康信息的主体，需要在数据利用和隐私保护之间得到平衡，构建一套切实可行的体系来解决与患者个人健康信息的法律法规问题，对医院的发展非常重要。通过实现业务流程和管理的数字化闭环，推动医疗卫生机构在保护患者个人健康信息的同时，大幅提升其医疗服务质量和服务效率。

PU-052

医疗信息系统在生殖中心临床决策上的支持与应用

栗 晔
沈阳菁华医院

引言：随着信息技术的不断发展和应用，信息技术支撑下的各类医疗信息系统已经成为现代生殖中心的重要组成部分。医疗信息系统是指 HIS（医院信息系统 HospitalInformationSystem）、LIS（实验室信息管理系统 Laboratory Information Management System）、PACS（影像归档和通信系统 Picture Archiving And Communication System）、EMR(计算机化的病案系统 EMR,Electronic Medical Record)、HRP(医院资源规划系统 Hospital Resource Planning)、PASS（合理用药监测系统 Prescription Automatic Screening System）、CDSS（临床决策支持系统 Clinical Decision Support System）以及专业生殖实验室管理系统等医疗信息系统，包括数据获取、医疗记录、病人管理、数据分析以及决策辅助等各个方面功能。本文将详细介绍相关系统的功能、应用等方面，并探讨它们对生殖中心医疗质量和效率的提升作用。

一、医疗信息系统的主要功能

（一）病人信息管理

病人信息管理功能主要体现在 HIS 系统中，HIS 系统主要负责记录和管理病人的个人信息，包括姓名、性别、年龄、联系方式、家庭地址等基本信息。此外，系统还可以存储病历资料、检查报告、诊断结果、治疗方案等详细信息，以便医务人员随时查看和更新。这样，医生和护士可以更好地跟踪病人的健康状况，提供个性化的治疗和护理。

（二）医疗信息管理

医疗信息管理涵盖了多个关键功能，也是由不同的信息系统来支撑的。除了 HIS 系统，EMR、LIS、PACS 等系统也承担了病人医疗信息的管理工作，从病历管理到医嘱管理，再到检查和检验结果管理，以及医疗知识库和医疗数据分析等功能。首先，病历管理是医疗信息管理的基础，要完成记录和存储病人的电子健康记录，包括病历、诊断结果、治疗方案和手术信息等详细内容。通过病历管理，医务人员可以更加便捷地查阅病人的过往病史，提供更准确的诊疗建议和制定个性化的治疗方案。其次，医嘱管理是医疗信息管理中的重要环节。HIS 系统可以支持医生生成和管理医嘱，包括处方药品、治疗计划、检查和实验室检验项目等。通过电子化的医嘱管理，可以减少因手写引起的错误，并提高医嘱的准确性和完整性。此外，LIS、PACS 等系统还能够记录和管理病人的检查和检验结果。医务人员可以随时查看和比对这些结果，以便进行进一步的诊断和治疗。这种便捷的查阅方式能够节约时间，并有助于提高医疗服务的效率和质量。同时，医疗信息管理还包括建立医疗知识库的功能，可以为医务人员提供临床指南、科研论文、药物信息等医学知识资源。通过这些知识库，医务人员可以快速查找和获取相关的医学知识，以便做出正确的诊断和治疗决策。

（三）生殖实验室管理系统

生殖实验室管理系统是对生殖中心病人进入辅助生殖周期后的全流程管理系统。临床部分功能主要包括：基本信息、周期方案、女方病历、男方病历、排卵监测、病程记录、病历讨论、卵巢手

术、OHHS 监测、冻融移植、取精手术、取卵手术、人工授精、胚胎移植、随访记录、周期总结、减胎记录、减胎随访及手术通知等功能。实验室部分功能主要包括：精液处理、取卵记录、胚胎培养、囊胚培养、PGD 记录、胚胎移植、精液冷冻、胚胎冷冻、胚胎图像采集、胚胎解冻与培养、实验室小结等功能。在信息获取方面除病人基础信息与医疗信息以外，还将对病人的指纹信息、人脸信息以及结婚证件信息等进行获取。生殖实验室管理系统以病人进入辅助生殖周期后的医疗流程为主线，以管理为目标，全面、完整地记录医疗信息，使信息系统成为管理的载体和工具，提升生殖中心现代化管理水平，实现决策支持。

二、生殖中心信息系统的特点

生殖中心信息系统建设在坚持标准化、智能化、自动化、安全性和拓展性的基础上，还应注重以病人诊疗流程管理、质量控制与个性化服务为原则，遵守相关法律法规，结合生殖中心的特点进行整合，致力于管理和处理生殖中心的日常运营和业务流程，为生殖中心医疗特殊性提供全面的支持和便利。

（一）数据获取的底层建设

生殖中心医疗信息系统应基于标准化数据要求，充分考虑全流程数据获取的框架搭建，同时要注重专业学科的基础数据特殊性。在记录病人一般信息以外，生殖中心对特殊信息的获取与病人基本信息的获取同样重要。例如对病人的职业、文化程度、生活地区、性传播疾病史、避孕史、个人史、婚育史、配偶信息、配偶医学信息等信息的收集，以及在整个辅助生殖过程中的对专业诊疗数据的收集，都对临床决策起着至关重要的作用。搭建全面、完整的数据框架，实现信息的共享和协同工作，将使生殖中心的工作效率得到显著提高。

（二）特殊数据的记录与处理

生殖中心信息系统应注重对病人特殊诊疗数据在整个医疗周期中的连贯性与动态变化的记录与支持。辅助生殖技术是一项对病人身体指标、治疗信息以及进入试管周期等全流程指标监测要求较高的医疗技术。例如，激素水平、精子质量、卵泡监测、子宫内膜厚度等医疗信息。信息系统能为医生提供连续、直观的数据动态变化展示与分析，将帮助临床医生客观、精准的判断病人诊疗情况以做出更科学和个性的诊疗计划。

（三）系统的可拓展性

生殖中心信息系统的可拓展性主要表现在软件之间丰富的数据接口，以及软件与硬件之间的兼容性与适配度方面。兼容性是指硬件或软件能够在某个操作系统或系统环境下运行的能力。适配度则是指硬件或软件能够适应不同的操作系统或系统环境的程度。一方面要支撑数据准确、完整、流畅的在全流程各系统之间获取与交互。另一方面要满足新设备、新工具的系统接入，实现在不同环节、不同操作环境以及不同系统中的使用。实现灵活拓展以满足不断变化的医疗服务需求。

（四）数据安全与隐私保护

生殖中心的数据安全和患者隐私保护非常重要。虽然随着社会的发展，人们对于生育的观念也发生了转变，但大部分不孕不育患者群体仍对此方面隐私信息格外重视，所以必须采取一系列措施来保障。在数据安全方面，一方面要采取必要安全措施，确保病人隐私和医院的数据得到有效的保护，其中包括数据加密、权限管理、备份和恢复等功能，有效地防止敏感信息被非法获取或篡改的风险。另一方面要提供严格的信息核对功能，夫妻身份信息的核对、诊疗信息与患者信息的核对，患者与卵子、精子、胚胎的信息核对。隐私保护方面，强化制度和医护人员隐私保护培训的基础上，系统应提供完善的预约功能，减少患者之间的交叉就诊情况发生。同时，应在各环节支持患者就诊匿名化功能，提供便捷、流畅的诊疗服务的基础上，信息发布功能应去实名化处理。

三、医疗信息系统对临床决策的支持

（一）个性化诊断支持

个性化诊断支持系统可根据个人信息和健康状况，依托各辅助系统的数据集成与共享，整合包括基础信息、病历信息、检验信息、影像信息等多系统来源信息，为医生提供强大的信息整合与分析功能，全面展示病人生理和病理状态。通过数据挖掘和算法进行分析和挖掘，提取特征信息，并结合医学专业知识类似病例经验，提供相应的鉴别诊断和辅助决策依据。

（二）治疗方案推荐

临床决策支持系统基于临床数据和决策模型，通过解析电子病历中的大量数据，帮助医生做出更好的诊断和治疗方案，降低医疗错误率，提高医疗服务质量。如根据当前疾病流行病学信息提醒医生进行相关检查或手术，或者根据患者病情等相关因素，为医生提供病情排序的建议等。其次，在诊疗过程中，合理用药检测系统提供了药物相互作用和药物过敏反应的提示，帮助医生避免治疗

过程中的潜在风险。在治疗方案推荐方面，系统可以整合循证医学的治疗指南和临床实践指南，系统根据患者的诊断，为医生推荐最新、最适合的治疗方案。

（三）利用数据模型对生殖中心临床决策的帮助

生殖中心数据模型要基于本中心的实际情况和特点建立，这样可以更好地应用于本中心的预测和分析。但是，对于病例熟虑较少的中心仅仅基于本中心数据可能存在数据量不足、数据质量不高等问题，因此需要结合其他相关研究和文献，以获取更多的数据支持和理论指导。生殖中心临床决策中常用的数据模型主要有回归模型、决策树模型、神经网络模型、随机森林模型、支持向量机模型等。在建立数据模型之前，需要进行数据清洗和预处理，保证数据的准确性和完整性。同时，应该选择合适的模型进行建模，并进行参数估计和模型评估，以确保模型的准确性和可靠性。

生殖中心利用数据模型，可以为生殖中心的病人和医生提供丰富的决策依据。对于患者来说，通过分析夫妻双方的生殖健康状况、生活习惯、饮食结构等信息，预测夫妻不孕的原因以及提高生育能力的措施。在完成诊断之后，可通过夫妻双方自然情况、健康状况、相关病史、女性激素水平、男性精液质量、子宫内膜厚度及形态结合相关检查结果，对试管成功率做出预测。可以使夫妻双方更好的了解相关情况与病情，结合医生意见做出更准确的选择。对于医生来说，通过患者个体情况、检验检查结果以及相关的信息收集，可以预测不同诊疗方案应用在该患者治疗上所得到的结局，帮助医生选择最优诊疗方提供辅助依据。

四、结语

医疗信息技术对临床决策支持的帮助，是现代医疗技术中非常重要的环节，也是现代医学的重要组成部分。各类信息系统的综合使用是生殖中心现代化管理和运营的基础，而系统的数据收集、整理、分析等功能则是提高医疗质量和效率的重要工具。未来，随着医疗技术的不断发展和进步，信息系统对临床决策的支持将发挥更加重要的作用，为医疗服务质量的提升和人民健康水平的提高做出更大的贡献。

PU-053

DIP 背景下基于改良波士顿矩阵模型大型公立医院专科发展对策研究

罗慧中 程楠
河南省人民医院

1 研究背景

《中共中央国务院关于深化医疗保障制度改革的意见》中指出，为充分发挥医保在医药卫生改革中的战略性购买作用，推进医疗保障和医药服务高质量协同发展，促进健康中国战略实施，明确提出了大力推进大数据应用，推行以按病种付费为主的多元复合式医保支付方式改革。医保 DIP 付款方式全称为 Big Data Diagnosis-Intervention Packet，是一套现实病例为依据，以穷举法匹配主要疾病诊断和主要治疗方式，再以本类病例的医疗费用支出作为演算基础，对医疗机构进行结算的一种方式。

2018 年国家医疗保障局成立以后，大力推进 DRG/DIP 付费改革试点，2021 年 11 月，国家医保局发布《DRG/DIP 支付方式改革三年行动计划》，明确到 2022 年不少于 40%的地区开展 DRG/DIP 改革，到 2024 年底全国所有统筹地区全部开展 DRG/DIP 付费方式改革工作，到 2025 年底，DRG/DIP 支付方式覆盖所有符合条件的开展住院服务的医疗机构，基本实现病种、医保基金全覆盖。

2 存在的问题

DIP 支付方式是对以往按照项目结算的支付方式的一种颠覆，倒逼医疗机构对自身的医疗活动开展范围、秩序和能力进行全方位的自查自纠，真正实现运营方式精细化。然而，改革的双刃剑必定也会带来一些问题：

2.1 DIP 收付费制度下，超过的费用由医院自掏腰包，自行垫付。由于不同的诊断和治疗手段对分值影响较大，医疗机构为了避免亏损可能会进行“高编”歪曲真实的病案信息。

2.2 医院可能会挑选患者和病例，减少容易超出定额的病例，增加可以结余的病例，调整专科床位数量来进行成本节约或收入增加，直接损害了患者生命健康权益。

2.3 急危重症患者病在按照 DIP 结算的时候大概率是亏损的，ICU 科室作为医院的高水平诊疗团队，在 DIP 背景下成为了医院的纯成本，阻碍了医院向上发展的动力；

2.4 DIP 支付改革对医院的运营水平提出了新的更高的要求，需要医院能够从全局视角出发，既能控制无序的医疗资源浪费，又能维持住医院的诊疗水平而目前各大医院信息化水平参差不齐，信息系统缺乏标准化和系统化，数据信息的交流及再处理能力差，无法支撑 DIP 的顺利运行。

3 研究理论

波士顿矩阵模型（Boston Consulting Group Matrix, BCG Matrix）是一种常用的战略分析工具，该模型将企业产品或业务划分为四个象限：明星型业务、问题型业务、现金牛业务和瘦狗型业务。每个象限代表了产品或业务在市场中的地位和发展前景。改良后的波士顿将全院所有的专科或病种标注在一个二维矩阵图上，横轴代表人均盈亏金额，纵轴代表病种分值，将病种分为高分值高风险、高分值低风险、低分值高风险、低分值低风险四个类型。

4 研究内容

本文采用改良后的波士顿矩阵模型对 DIP 付费背景下的某公立医院（以下简称 A 院）的 2022 年的市医保结算数据进行分析。以往的波士顿模型多用于市场营销方面的研究，本次研究将商业化模型移植到医院领域，并且对波士顿模型的本身的短板进行了补充。以专科年度出院人次为横轴，病组的盈亏金额作为纵轴，气泡大小表示专科本年度的出院患者的总费用，对全院的整体运营情况进行画像，结果以改良后的波士顿矩阵模型进行呈现（如图 1 所示），同时结合统合综效理论对矩阵模型进行定制化特色化的补充，对专科内资源进行最大化利用和最合理配置，辅助学科发展良性化。

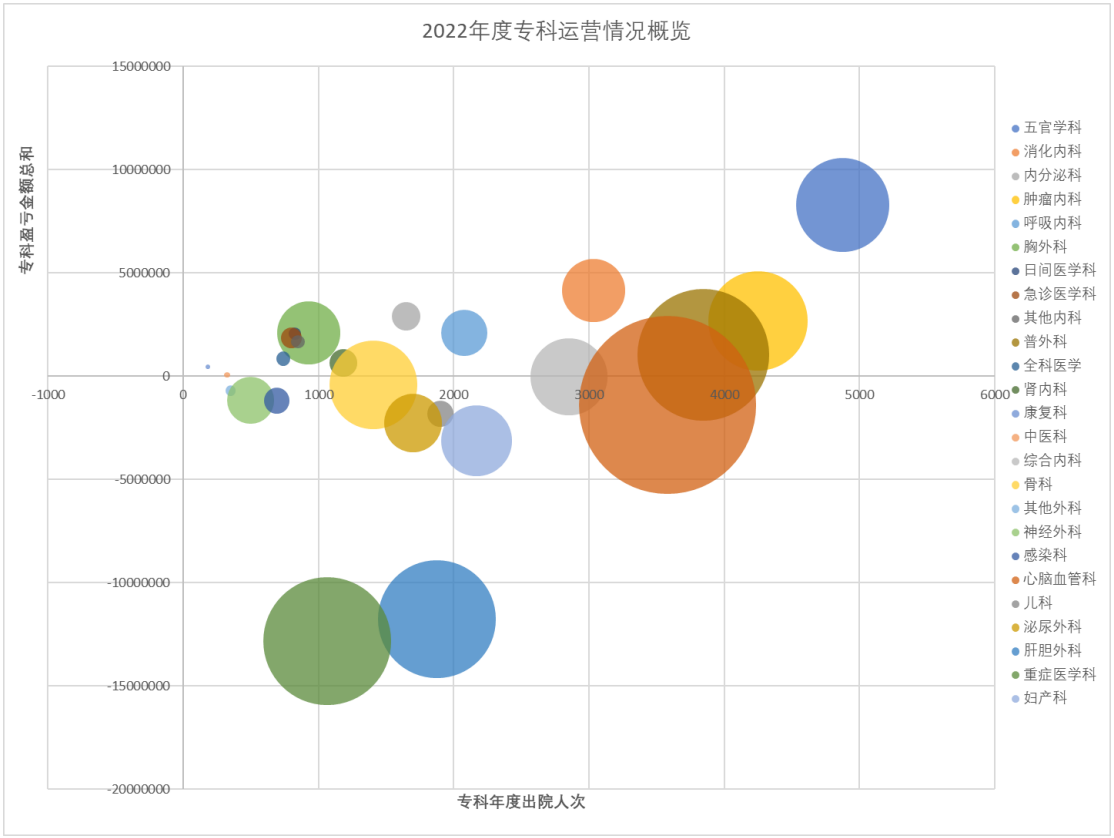


图 1 全院专科运营情况

5 研究结果

全院盈亏的专科数量分布较为平均，盈利专科 13 个，亏损专科 14 个，基本面呈现亏损。按照出院人数和盈亏金额两两组合，以 2500 人次的出院人数为中轴线，将全院专科划分为四个大类，分别为成长型、成熟型、问题型和风险型（如图 2）。

成长型对应出院人数少，略有盈余或者亏损的专科，此类专科调整速度较快，难度小；成熟型对应出院人数多，治疗方案合理的专科，维持现状即可；问题型对应亏损巨大，但是出院人数较少的专科，此类专科往往更多的存在管理问题，调整难度大，需要行政科室的积极参与；风险型对应高出院人数，有亏损有结余，此类科室做好专科内病组管理即可实现提升，但是由于出院人数较大，对于医院的波动性影响也很大。

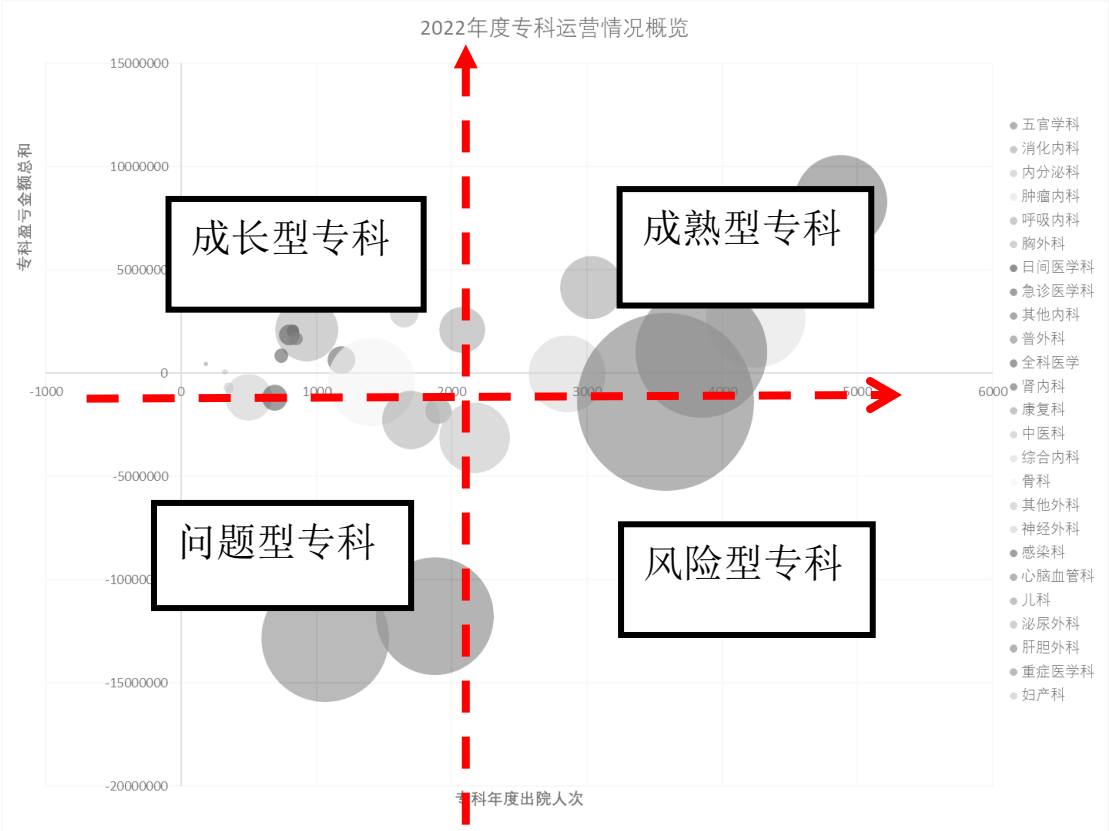


图 2 全院专科象限划分结果

亏损 TOP3 的是重症医学科、肝胆外科和血液内科（表 1），盈余的 TOP3 是五官学科、消化内科和内分泌学科（表 2）。

名次	专科名称	出院人次	盈亏总额（万元）	医疗总费用（万元）
1	重症医学科	1062	-1283.3	7524.4
2	肝胆外科	1876	-1177.5	6915.2
3	血液内科	1212	-366.4	3287

表 1 全院亏损前 3 位的病种概况

名次	专科名称	出院人次	盈亏总额（万元）	医疗总费用（万元）
1	五官学科	4876	828.8	5508.6
2	消化内科	3034	414.6	3739.8
3	内分泌科	1648	290.7	1662

表 2 全院盈余前 3 的病种概况

重症医学科和血液内科都属于问题型科室，但是不同于肝胆外科的情况，这两类专科是受限于 DIP 目录库无法体现疾病的危重程度和复杂程度，具有天然的亏钱属性，短时间内不会有大的改变，作为收治急危重症患者的主力科室，是医院高水平诊疗的体现，也是医院必须要承担的成本支出。

在对肝胆外科进一步的分析中发现，肝胆外科收治病人量全院第 10，外科大类中收治病人量全院第 2，病种分布存在极不均衡，低分值（小于 200）的病例数占据专科全部病例数的 62%，300 分以下的病例数占比 78%。低分值反映出了肝胆外科治疗方式存在偏保守的情况，收治患者不符合专科定位，属于战略目标错位、多劳多损类型的专科。需要对专科内人员进行专业方向细分，定量考核手术台次、四级手术率、DIP 分值。

问题型专科的共同特点是病组患者数量和分值均低于平均水平，在管控上，应根据专科特点和上一年度的盈亏情况分类处理，但是更新迭代治疗技术，增加疑难危重病例的收治是共同关注点。

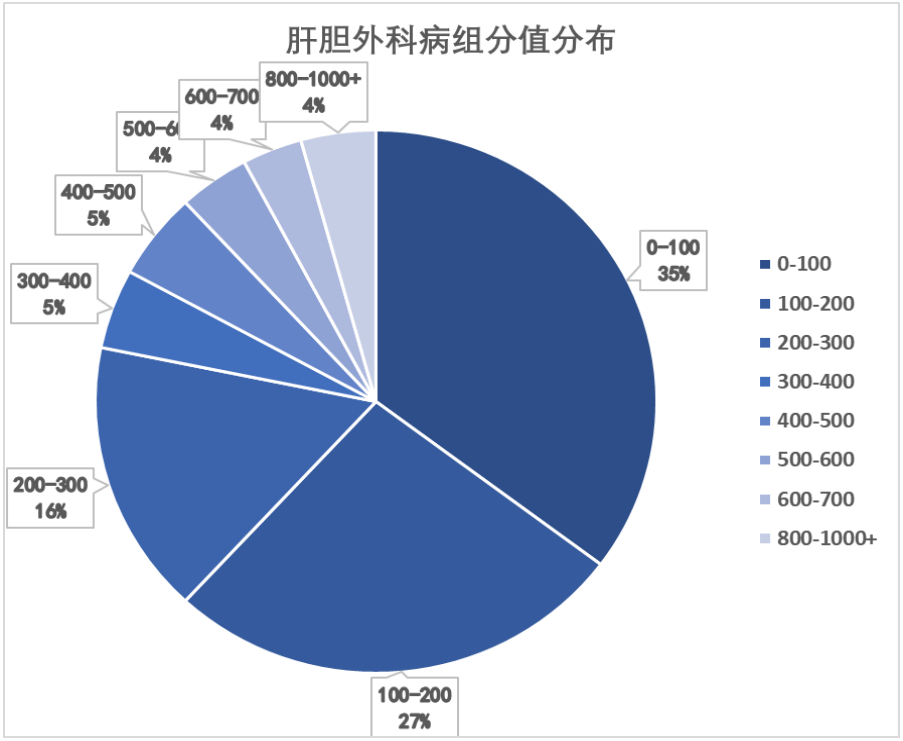


图 3 肝胆外科病种结构

类比于胸外科（成长型科室）收治病人量全院第 18，外科大类中收治病人量全院第 4，病组种类精简均衡，低分值（小于 200）的病例数占据专科全部病例数的 50%，中分值、高分值各占将近 30%。

成长型专科的病组疑难程度较高，但收治患者数量较少，服务效率和品牌吸引力也有待提高，可以加大资源投入，提升专业技术水平和服务质量，成就为院内口碑型专科，确保患者应收尽收和应治尽治”。

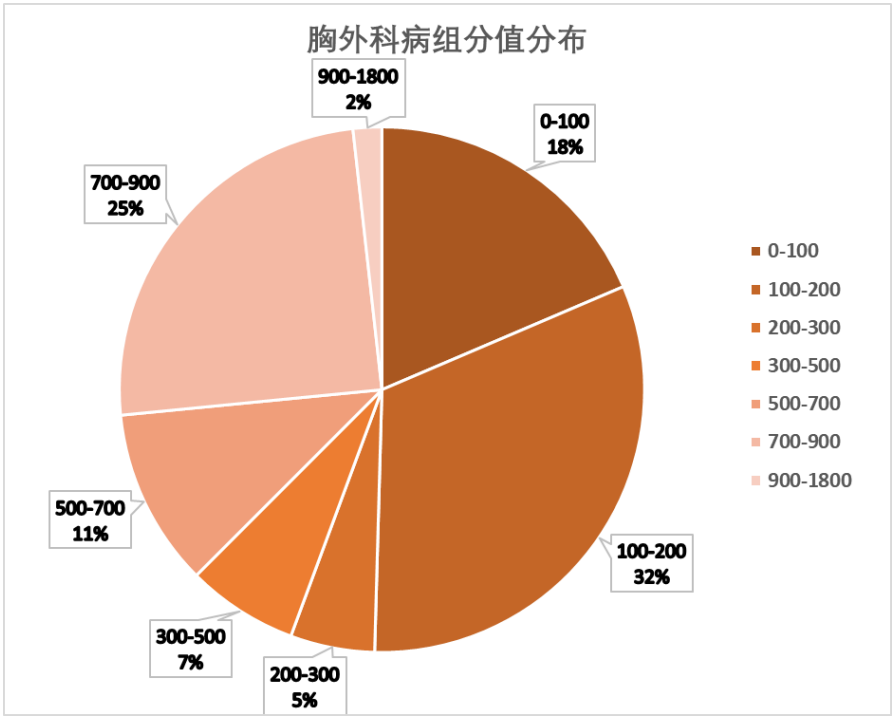


图 4 胸外科病组结构

医院需要重点关注风险型科室，如心脑血管专科、综合内科和神经内科。此类专科具有大规模的患者基数和高治疗成本，对医院的医疗资源消耗大，病种结构健康度欠缺，有很大的提升空间。聚焦于心脑血管专科来看，出院人数的病组分值分布较为均匀，在高分值区间存在断档断层（如图 3），与药占比、材占比、平均住院天数呈现明显的相关关系（如表 4），医务管理部门应该加强对此类专科的病组进行细分管理，运营管理部门应制定明确的成本目标并且严格跟踪，以降低风险型专科的次均费用。

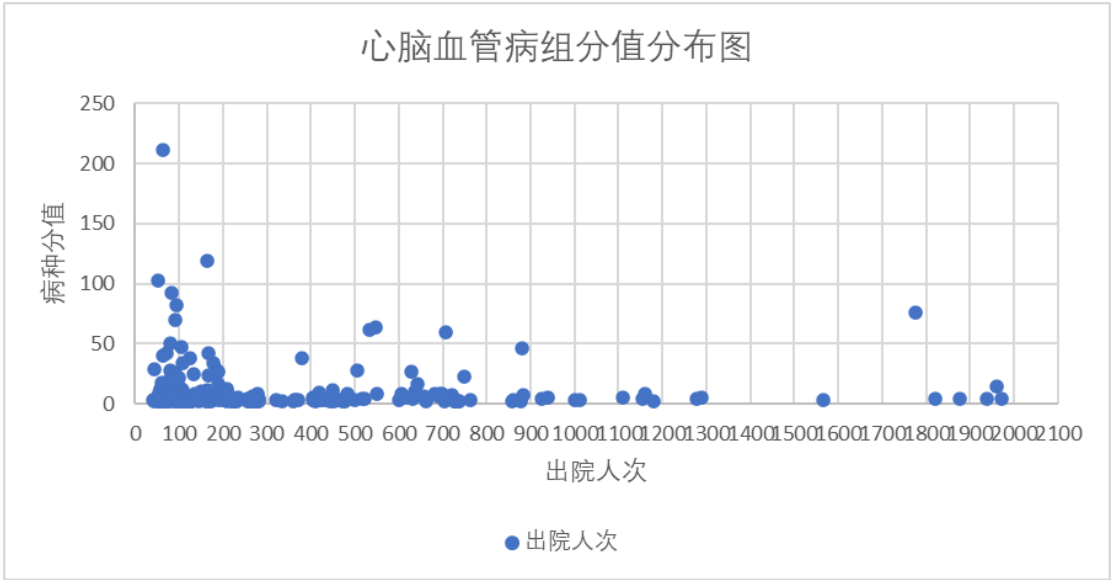


图 3 心脑血管专科病种概况

	药占比	材占比	平均住院天数	盈亏金额

病种分值	-28.82%	68.17%	47.07%	2.73%

表 4 病种分值与药占比、材占比、平均住院天数、盈亏的相关性结果

6 结论

总体而言，为了保证医保保障水平，各地持续强化医保支付管理将是必然措施、同时持续降低个人卫生费用占比也是国家社会发展的重要目标，任何医疗改革都是一种挑战，但是“以病人为核心”的目标始终不变，为患者提供更优质的服务目标始终不变。积极正面地面对医保管控收紧，优化医院自身的管理与医疗行为，以 DIP 作为管理工具，真正实现精细化管理，让优质的医疗资源得到充分利用、努力实现医、患、保三方的利益平衡、建立起适应新环境的服务方式才是可行的医院发展路径。

PU-054

医院信息系统与临床决策辅助 CDSS 应用的管理规范和应用效果

王秉慧
山西省忻州市人民医院 信息科

引言：

临床决策支持 CDSS 作为医院信息系统的一部分，能够为医院信息化建设提供重要的支持辅助。它可以与医院其他信息系统进行集成，实现临床数据的共享与传递，提高临床工作效率。CDSS 应用还可以提供丰富的临床知识库和决策支持工具，为医务人员提供有效的决策辅助，提高医疗质量和安全性。

临床决策辅助 CDSS 应用在提高临床决策准确性、工作效率和医疗质量安全方面具有重要的帮助和意义。然而，应用 CDSS 也面临一些挑战，包括系统搭建和维护的资源投入、准确性和可靠性的验证和优化、医务人员的培训和理解等。为了进一步提高 CDSS 应用的效果，需要加强管理规范和进行应用效果分析。

医院应制定相应的管理规范，确保 CDSS 系统的正常运行和有效使用。管理规范应包括 CDSS 系统的数据管理、知识库更新、决策算法验证等方面，以保证其符合临床实践的要求。此外，还需要加强对 CDSS 应用效果的监测与评估，及时发现问题并进行改进。

方法：

1. CDSS 应用管理规范的符合性

在保证 CDSS 应用有效性和可靠性的前提下，制定和遵守管理规范是至关重要的。以下是几个方面的管理规范：

1.1 数据管理规范

确保 CDSS 系统中所使用的临床数据的准确性、完整性和保密性。包括规定数据采集的方式和标准，保障数据的及时更新和备份，以及设立权限控制机制，限制数据的访问范围。

1.2 知识库更新规范

针对临床知识库的更新，确保其中的医学知识、最新研究和指南等信息时刻保持最新和可靠。定期审核和更新知识库内容，并与专业机构、学术团体等合作，确保知识库的权威性和科学性。

1.3 决策算法验证规范

CDSS 系统中的决策算法应定期进行验证和校准，以保证其准确性和可靠性。通过与实际临床案例的比对和评估，验证决策算法的正确性和适用性，并对不符合要求的算法进行修订和替换。

1.4 培训和教育规范

对医务人员进行相关培训，使其了解 CDSS 系统的使用方法和决策支持原理，掌握正确的使用技巧，并能够理解 CDSS 系统提供的决策结果和建议的依据。此外，还要提供定期的继续教育和培训，以使医务人员不断更新和提升相关知识和技能。

2. CDSS 应用效果分析

评估 CDSS 应用的效果是不可或缺的，以下是几个方面的应用效果分析：

2.1 临床决策准确性

通过与医生的决策结果进行对比，评估 CDSS 系统在提供决策支持方面的准确性。通过分析准确率、误诊率等指标，评估系统对临床决策的帮助程度。

2.2 临床工作效率

通过与传统方式进行对比，评估 CDSS 应用在临床工作中的效率提升。包括减少医生的决策时间、提高临床流程的顺畅度、减少重复工作等方面。

2.3 医疗质量安全

评估 CDSS 应用对医疗质量和安全的影响。包括减少不必要的检查和治疗、降低不良事件的发生、提高患者满意度等方面。

2.4 用户满意度

通过医务人员的反馈和调查，评估 CDSS 系统的用户满意度。了解医务人员对系统的易用性、决策结果的合理性、界面的友好性等方面的感受。

通过以上的分析，能够对 CDSS 应用的符合性和应用效果进行全面评估，为医院信息系统与临床决策辅助 CDSS 应用的管理与改进提供参考依据。这些成果对于推动医院信息化建设和提高临床决策质量具有重要意义。

3. 持续改进与质量管理

在管理 CDSS 应用的过程中，持续改进和质量管理是不可或缺的环节。以下是几个关键措施：

3.1 用户反馈和需求反馈

医务人员的反馈和需求是改进 CDSS 应用的重要依据。建立反馈渠道，定期收集用户对系统的意见和建议，并及时进行分析和处理。通过用户参与的方式，不断优化 CDSS 应用的功能和用户体验。

3.2 风险管理与安全措施

CDSS 应用涉及医疗决策和患者数据，必须重视风险管理和安全措施。建立完善的安全机制，加密敏感数据，监控系统的访问和操作，及时发现和应对潜在的风险和威胁。

3.3 与临床实践的整合

CDSS 应用需要与临床实践紧密结合，以确保提供的决策支持符合实际需求。与临床团队、专业学会和临床指南的合作，整合最新的临床知识和指南，不断更新 CDSS 应用的决策算法和知识库。

3.4 性能监测与评估

建立性能监测和评估机制，对 CDSS 应用的性能进行定期监测和评估。包括系统的响应时间、数据准确性、决策结果的准确性等方面，以确保系统的稳定性和可靠性。

4. 贡献和结论

本文在医院信息系统与临床决策辅助 CDSS 应用的管理规范符合性和应用效果方面进行了深入研究和分析。通过对全医院信息化建设的支持辅助和临床大夫的帮助及意义的探讨，阐述了 CDSS 应用对医院信息化建设和临床医生的重要作用。通过对 CDSS 应用的应用效果分析，评估了其在临床决策准确性、工作效率和医疗质量安全等方面的影响。

同时，本文还对 CDSS 应用的管理规范进行了探索和总结，包括数据管理、知识库更新、决策算法验证和培训教育等方面。并提出持续改进与质量管理的重要性，强调用户反馈、风险管理、临床实践的整合和性能监测与评估等关键措施。

综上所述，本文对医院信息系统与临床决策辅助 CDSS 应用的管理规范符合性和应用效果进行了全面研究和分析，并提出了相关的建议和措施。这些研究成果对于医院信息化建设和提高临床决策质量具有重要指导意义，为 CDSS 应用的推广和应用提供了有益的参考。

本研究采用文献综述和分析的方法，对医院信息系统与临床决策辅助 CDSS 应用的管理规范和应用效果进行研究和分析。通过收集相关文献和案例，对 CDSS 应用的管理规范和应用效果进行整理和总结，并提出相关的建议和措施。

总之，医院信息系统与临床决策辅助 CDSS 应用的管理规范符合性和应用效果对于提高医院信息化建设和临床决策质量至关重要。通过合理的管理和不断优化，CDSS 应用将为医院信息化建设和临床医生的工作带来更大的效益和意义。

结果与讨论：

本文通过对医院信息系统与临床决策辅助 CDSS 应用的管理规范符合性及应用效果进行研究，得出了以下结论：CDSS 应用对医院信息化建设提供了重要的支持辅助；它对临床医生具有重要的帮助和意义；CDSS 应用在临床决策中能够提高准确性、工作效率和医疗质量安全。然而，应用 CDSS 也存在一些挑战。首先，CDSS 系统的搭建和维护需要投入大量的人力和物力资源。其次，CDSS 系统的准确性和可靠性需要不断验证和优化，以确保提供准确的临床决策支持。此外，医生和其他临床人员需要接受相关培训，了解如何使用 CDSS 系统，并理解其决策依据和局限性。

PU-055

基于 CDR 的信息闭环管理助力医院高质量发展

朱骄锋
国药同煤总医院

1 引言

作为公立医院改革试点工作的重要任务之一，建立和完善以电子病历为核心的医院信息系统，是实现现代化医院管理目标的重要措施[1]。随着我国医药卫生体制改革的日益深入，医院面临越来越大的挑战。怎样营造一个经营求特、服务求优、医术求精、收费求廉的就医环境，使医院在当今日益激烈的竞争环境中生存、发展，是我们面临的主要问题。医院流程再造是一个可选择的正确方向，其中电子病历的应用又起到了非常重要的作用[2]。

2011 年国家颁布了《电子病历系统功能应用水平分级评价方法及标准》，此标准既是国家对医院信息化建设水平的评测工具，也是国家对医院信息化的建设指南[3]。同煤集团总医院至 2003 年开始逐步完善信息化建设，按照电子病历六级评价标准，利用信息化建设优势再造诊疗流程，通过各诊疗流程的信息化闭环建设，建设闭环监控平台，找出流程中的缺失环节、隐性不安全因素等，明确各个环节责任人, 使得业务过程的每个重要环节实现了数据管理，做到全员追踪、全程追溯，从而规范了员工行为，保证了患者诊疗安全。下面以药品、检验、输血闭环管理为例，介绍一下闭环管理体系建设的方法与应用。

2 闭环管理建设方法与应用

2.1 药品闭环管理

医嘱是医师在医疗活动中下达的医学指令，药品医嘱是最重要的医学指令，几乎所有疾病的治疗都离不开药品。依照药品的用法分为输液、口服、注射等。

2.1.1 药品医嘱的整体业务流程

以住院静配药品为例，医生开立药品医嘱后，自动审方系统自动审核医嘱，系统报警的医嘱由药师进行人工审核，审核通过后由护士签收医嘱，配液中心按医嘱进行备药、冲配、出仓打包，交由外勤人员配送到病区，护士接收药品并床旁执行。药品执行后，护士定时巡视，出现药品不良反应立即与主管医生反馈并按不良事件处理。医生可以明确开立的每一条医嘱的每一个执行状态，方便制定下一步的诊疗计划。

2.1.2 关键应用点

①在医生下达药品医嘱、护士核对药品医嘱、药师审核药品医嘱时，能够自动检查药品的合理性，系统给予明显标志，以保证用药合理性与患者用药安全。②医生可以快捷方便的查询清楚，哪个药品何时发放、使用到哪个患者身上、患者使用后有无不良反应，以及整个闭环环节中涉及相关人员的完整记录。

2.2 检验闭环管理

根据卫计委 《医疗机构临床实验室管理办法》、《临床检验专业医疗质量控制指标（2015 年版）》等，为加强对医疗机构临床实验室的管理，提高临床检验水平，保证医疗质量和医疗安全，对检验各环节进行了闭环管理，分为门诊检验闭环、住院检验闭环、不合格标本的闭环、门诊危急值闭环、住院危急值闭环以及三级微生物监测的闭环管理。以危急值管理为例。

2.2.1 检验危急值管理流程

住院患者危急值报告流程：当检验出现“危急值”时，检验者首先要确认检验仪器和检验过程是否正常，在确认各环节无异常的情况下，才可以将检验结果利用系统向病区护士、医生工作站报

告，如果 5 分钟内病区护士不点击接收，检验科电话告知病区主班护士签收。主班护士应立即向主管医生报告，医生需立即（10 分钟内）采取相应诊治措施，如找不到主管医生，需及时报告上级医生或科主任，由他们负责采取相应措施。主管医生需 6 小时内在结构化病程中记录接收到的“危急值”检验报告结果和诊治措施。

门诊患者危急值报告流程：当检验出现并确认“危急值”时，检验科利用系统向门诊医生工作站、门诊导医台系统发送危急值报告，由导医台护士签收并与门诊医生确认，门诊医生联系患者后给予相应措施。

2.2.2 关键应用点

①LIS 系统一旦出现危急值，结果显示黄色（参考值范围内黑色，高于参考值红色，低于参考值蓝色）。 ②报告一旦确认发送，如为病区标本，护士工作站立即出现报警，除非手动解除，否则报警不会停止。③医务科每天对自动统计的危急值项目抽查临床处理情况。在关键节点进行系统控制，保证了检验结果及时、安全送达，确保患者的诊疗安全。

2.3 输血闭环管理

根据卫生部《临床输血技术规范》相关规定，对患者输血进行全过程管理与控制，是医生开立输血医嘱、输血科配血、外勤人员接送血液、患者用血等各个环节之间形成输血的闭环管理模式。

2.3.1 输血流程

①用血申请：医生下达配血医嘱，系统自动弹出输血申请单，填写完毕后提交申请单。②标本采集：护士根据医嘱进行血样采集，对采血管粘贴条码。③标本运送：外勤人员按照配送提示去相应科室接收血样。④标本检验、配血：输血科通过系统接收输血申请，根据接收血样进行配血，根据 LIS 系统中血型、病毒检测结果等信息，对输血适应证、血型等进行质控。⑤取血：配血完成后，输血科利用外勤管理系统通知外勤人员取血，后勤人员到血库利用 PDA 扫描血袋上条码取血，做扫码交接。⑥输血：外勤人员与临床护士扫码交接血制品，核对患者信息无误后由病房护士或手术室护士输血前双人核对并为患者进行输血操作。⑦巡视：在输血过程中，必须到患者床旁使用 PDA 扫描血袋上的标签巡视输血反应，如发生不良反应必须立即通知医生处理。⑧输血结束：护士填写电子输血护理单，输血袋由外勤人员在 2 小时内送返血库回收，空血袋在输血科至少存放 24 小时进行相应处理。

2.3.2 关键应用点

①适应症控制：医生开立医嘱时知识库提醒，输血科配血系统会对不合理用血进行提示。②血型控制：医生开立配血申请时，输血科配血时系统可根据历史血型进行提醒。③输血核对：通过 PDA 扫描进行血制品与患者核对。系统上线后无一例输血差错。④输血巡视：时间控制在 30 分钟内。随着 PDA 巡视系统的上线使用，患者输血时，护士每次查房的时候通过扫描血袋、患者腕带、护士工牌等，可以详细记录每次巡视的情况，如果发现不良反应，能够得到及时的处理，并将相关的记录和反馈信息及时进行通知传达。

3 闭环监控平台的建设与管理

医院创新性将各大闭环的执行时间与执行情况作为医院质量管理的重要抓手，建设了基于 CDR 的闭环实时监控平台，对各大闭环每个时间节点和行为操作进行实时监控，对执行不好的项目和个人直接通过系统进行提醒，医务科、护理部等职能科室进行相应问题的考核和督导，以不断通过系统督促加强医疗行为的规范性，不断提升医疗质量。

4 总结

基于信息化辅助手段，实现闭环管理模式，借助二维码技术，实现了各个诊疗环节的全流程信息化监控，提高了准确性，减少了护理医疗事故的发生率[4]。因此，通过闭环管理拓展移动医疗系统建设，既是新一代临床信息管理基础，也是数字化医院建设的重要内容，对医院的信息化发展具有重要意义[5]。闭环管理体系已渗透到医院诊疗的各个环节，使医院医疗质量管理更加精细化、科学化，医疗安全得到进一步的保障。

PU-056

医疗信息系统临床评价提升问题初探

孙勇 温俊
南京医科大学第四附属医院

本文的医疗信息系统指的是在医院中为临床业务的运转所建设的信息软件系统，主要包含有医院信息系统（HIS）、实验室信息系统（LIS）、影像归档与通信系统（PACS）、电子病历系统（EMR）等基础功能模块，也包括智慧护理、合理用药（PASS）以至临床决策支持系统（CDSS）等以知识库为核心的的辅助决策支持系统。近年来，医疗信息系统获得了越来越多的关注，且应用的水平和范围也不断得到提升和扩张。但一个普遍存在的现象是：在大多数医疗机构中信息系统建设都呈现出临床评价不高甚至负面性评价的特征。这种现象通常被认为是一个发展中的问题而被忽略。现实情况却是，这个问题已经对医疗信息系统在医疗机构中充分发挥作用并持续更新发展形成了阻碍，需正视并探求其原因加以改善解决。

一、信息系统视域中的研究启示

医疗信息系统临床评价提升这一问题其内涵是医疗信息系统使用者对医疗信息系统的评价不高甚至负面性评价，需要加以分析提升。当前，学界对这一问题的研究还较少，但如果把医疗信息系统看作是信息系统的一种，则在信息系统视域中所作的此类研究则呈现出增多的趋势，而这与信息系统自身的不断升级发展也是相辅相成的。

孙效华等学者提出，信息系统已经逐步从人与数字系统的交互过渡到人与智能系统的交互阶段[1]，这是当前信息系统包括医疗信息系统发展的一个重要趋向和特征。董文莉也指出，信息系统使用者与信息系统的关系逐步从主要控制者的角色转变为与信息系统共享控制的合作者[2]。这个判断隐含地表明了处在这一关系中的信息系统使用者在地位上所面对的结构性变化。同时，正因为系统智能性的不断提升，信息系统与信息系统使用者之间的知识能级落差在加大，系统相对于使用者来说可解释性趋于减弱，较之前更接近黑箱，特别是深度学习等技术手段的加入使其算法透明度降低，愈加影响信息系统使用者的感受[3]从而也影响了其对于信息系统的评价。

除了结构性研究这一本源研究之外，更多的研究开始专注于信息系统使用者在体验认知层面的问题。随着以情境感知、自适应学习、自主决策及主动交互与协同为关键特征的智能系统的应用，系统的自完善性和主动性不断增强，但同时带来的是对作为系统使用者的人的能动性的一种挑战[4]，对人的自主性的压抑带来的负面情绪还将反作用于信息系统，减弱信息系统使用者的配合意识，并影响系统控制行为的成熟度。除了使用者的能动性之外，信息系统与信息系统使用者之间信任关系的构建也是使用者对信息系统认知和评价的关键。当信息系统协助信息系统使用者于不确定性的事件做出决策时，对可能产生的效果或结果的预期就反映了信息系统使用者对于信息系统的信任程度[5]。而真实发生的效果或结果又会反馈影响信任关系的强弱[6]，若由于信息系统不成熟，判断归因考虑不完整、不充分使得协助效果不佳，从而导致信息系统使用者不能与信息系统构建可靠的信任关系，则会持续影响信息系统使用者对于信息系统的认知和评价。

通过信息系统视域中系统使用者对信息系统认知和评价的分析，可以发现，信息系统使用者对于信息系统做出不高的评价有其相互关系的结构性演变这一根本原因，同时由于信息系统使用者是人，也不能忽视其中人本的主观认知和情感倾向的因素。这启示我们，对于医疗信息系统的临床评价不高的问题，也需要从客观和主观两个方面进行综合性的分析和研究。

二、医疗信息系统临床评价的分解探析

把目光从信息系统聚焦到医疗信息系统，在信息系统中存在的系统使用者对系统评价不高甚至负面评价的问题同样存在于医疗信息系统。同时，为解决医疗信息系统的具体问题，也需要进一步深入探究医疗信息系统自身的特殊性所带来的个性问题。

医疗信息系统同样处于数字系统向智能系统的转变之中，在这一过程中，作为医疗信息系统使用者的临床工作人员也在逐步从单方面的使用者角色向智能化系统的合作者角色转变，而在这两个转变的过程中也必然会伴随着临床工作人员对医疗信息系统认知的模糊和评价不高等问题。同时，大多数的医疗信息系统在其设计和建设之初都少有考虑到设置发挥临床工作人员主观能动性的空间

及在系统智能发挥过程中与之构建信任关系的问题，这同样是医疗信息系统难以得到临床正面评价的重要原因。

除了上述导致评价不高甚至负面性评价的根本性及普遍性原因之外，医疗信息系统还存在一些基于其自身特殊性所带来的评价问题的自因。首先，医疗信息系统是一个专业性、流程性极强的系统，由于医疗过程的极度严谨和可追溯要求，其确认点、控制点、记录点非常多，进一步加剧了系统的繁琐程度暨操作难度和学习成本。

其次，随着 CDSS（辅助决策支持）、AI 智能等知识库和算法类型的智能应用成为医疗信息系统的重要补充甚至快速成为整个医疗信息系统的必要构成的情况下，临床实践中的诸多决策不再成为临床工作者的独有权力。一定程度上医疗信息系统的内置知识库和算法开始与临床构成竞争关系，这一点在消灭影像和皮肤科医生的呼声中展现的最为明显，同样也加剧了临床在危机意识主导下对尚不完善的相关系统给出较低的甚至是负面性的评价。

再其次，由于医疗保险、外联平台和医疗监管信息化[7]等第三方系统及要求的增多，医疗信息系统的独立性逐步丧失，合规性要求日益增加，这对于医疗信息系统本身是极大挑战，在合规和系统操作自主性、流畅性之间缺乏平衡极易为医疗信息系统本身带来负面性评价。

最后，在医疗信息系统临床操作者的主观视野中，存在着社会公共系统平台及商业软件系统的比较性评价，不论是微信、支付宝这些国民应用平台还是 photoshop、office 这些专业工具软件系统，都有强大的开发团队和海量的用户，他们可以根据海量用户的实际需求进行敏捷的迭代开发，同时在市场竞争的压力下其软件易用性和人性化均在快速进化。对于使用习惯了这些系统平台和软件的医疗信息系统使用者来说，版本相对固定，需求开发相对迟缓的医疗信息系统将更加固化其负面印象，且在日常非自觉、非对称的比较中不断强化。

另外，医疗信息系统在其设计和使用过程中还存在专业性、弹性、稳定性不足等多种问题，给临床使用者带来困惑和不便。而相应的培训也流于形式和表面化，不能真正达到入脑、入心的良好效果，进一步加剧了医疗信息系统的临床评价不高甚至负面性评价的问题。

三、主客融合提升医疗信息系统评价

通过分析，医疗信息系统在临床存在评价不高甚至负面性评价问题的原因可以归结为客观和主观两方面。客观方面在于医疗信息系统与其临床使用者在角色地位上的变化消减，主观方面则由于在系统设计、运行及维护上没有充分考虑到使用者的主体性地位，导致临床使用者对医疗信息系统存在认识上的片面和错位。但要对其加以解决，不能再犯将主观与客观加以割裂的错误，应遵循主观见之于客观，主客融合的思路进行。

第一是易用性，医疗信息系统在其设计之初就应充分考虑到未来临床使用者的功能预期和使用感受，对系统的界面、流程加以简化、优化。减少临床工作者在系统操作上所要付出的学习成本和时间成本，提升临床使用者的主观感受。

第二是专业性，医疗信息系统流程设计应充分理解和尊重医疗行为背后的特有逻辑，真正把医疗理念融入到系统设计中去。具体则要引导临床工作人员在系统流程设计的初期就发挥其主观能动性，将医疗专业的思维与系统流程设计思路相碰撞，得到符合临床实际的系统流程，只有这样临床使用者才会在使用系统的过程中感到得心应手，有助于给出较高的评价，这在治疗、输血等系统流程设计中体现的最为明显。

第三是一致性，医疗信息系统应通过系统的集成化、平台化和统一数据存储打通各子系统、功能模块间的信息墙。统一数据格式，统一数据调取，减少因各子系统数据的不一致而导致在临床展现时发生信息错误、缺漏的问题。一致性的保障可以有效提升临床的使用效率和感受，有助于建立临床工作人员与医疗信息系统间的信任关系。

第四是互动性，随着人工智能的发展以及社交软件的高度普及，医疗信息系统应对其理念加以利用和借鉴。如临床辅助决策支持系统核心在于知识库，而构建一个成功的临床知识库的关键在于发动每一个临床工作者参与到其增加、审核、修正的过程之中，这样才能形成最符合特定医院当时、当科、当人实际的知识库。同时应为临床工作者个人及之间甚至包括患者提供展示状态、交流互动、投票表达的空间和渠道，努力把临床信息系统打造为以工作为核心的聚合平台。这将大大提升医疗信息系统对临床使用者的吸引力从而提升美誉度。

第五是弹性，医疗信息系统的弹性指的是其功能扩展的难易程度，如果增加一个科室或是增加一个诊断都很困难，或是新增一个检查流程都难以办到，那临床使用者在束手无策的同时必然要对

系统提出抱怨。增加医疗信息系统的弹性需要在设计之初就在数据库中保留表及字段的扩展余量，同时在系统设计中贯彻配置化管理的思想，并在必要的位置增加自定义和个性化设置的功能，照顾到使用者操作及界面上的个性化需要，提升自由度和参与感。

第六是稳定性，系统的不稳定会极大地影响到医疗信息系统的临床评价，因为临床始终在面对患者，这种即时压力是系统设计者应感同身受的。而要提升系统的稳定性包括软件和硬件两个层面，软件层面应强调代码审查和上线前测试的可靠，尽可能采用模块化编程方式并在前台设置完整可读的错误提醒。硬件层面一方面在服务端采用双机、超融合等高可用架构，尽可能减少因故障停机时间，另一方面 PC 端应标准化配置，采用准入控制和终端杀毒等手段保证其稳定可靠。

第七是培训，培训是否有效很大程度上决定了临床使用者对医疗信息系统的先发印象和上手难度，首先应在医疗信息系统上线前进行普遍、逐一的培训，上线后如有重要模块或功能加入，需对涉及到的人员进行专项培训。在日常系统维护过程中，也应以发布图示、公告的方式进行问题引导下的强化培训。另外，还应注重临床使用人员的自我培训，培养一些信息专员成为科室的内培员，提升培训质量和问题沟通处理的及时性，提升临床满意度。

综上所述，虽然当下的医疗信息系统事实上已大幅提升了临床工作的效率和准确性，但却没有得到与之相称的评价。这是不能被忽略的问题，更不能忽视由此引发的对医疗信息系统的厌恶和抵制情绪，如任由此情形发展，将大大降低医疗信息系统的使用效能，使系统非强制的一些统计性、智能性、质控性和安全性功能被闲置，这不仅导致医疗信息系统的进一步建设发展停步不前，更会对医院的管理体系建设特别是医疗安全体系建设带来不利影响。

PU-057

医学信息系统助力临床决策

智鹏
解放军医学院

目的 探讨医学信息系统在临床决策中的作用，以评估其对卫生保健提供者的支持和患者护理质量的影响。

方法 进行多个医疗机构的案例研究，分析医学信息系统的实际应用情况以及与临床决策相关的关键指标，如诊断准确性、治疗效果和患者满意度。此外，开展一项横断面调查，收集了医护人员和临床决策者的观点，以了解他们对医学信息系统的看法和体验。

结果 医学信息系统在临床决策中发挥了重要作用。这些系统提供了即时的、可靠的患者数据和医学知识，有助于医疗专业人员做出更准确的诊断和制定更个性化的治疗计划。案例研究表明，医学信息系统在减少医疗错误、改善治疗效果和提高医疗效率方面取得了显著的成果。横断面调查结果显示，医护人员和临床决策者普遍认为医学信息系统提高了工作效率，减少了文书工作，同时提高了患者安全性和满意度。然而，一些挑战仍然存在，包括系统互操作性问题、学习曲线陡峭以及信息安全和隐私问题。

结论 医学信息系统在临床决策中的应用对医疗卫生领域产生了积极的影响。它们为医疗专业人员提供了更多的数据支持和决策工具，有助于提高诊断和治疗的准确性，同时改善了患者的医疗体验。然而，为了充分利用这些系统，必须解决技术和安全性方面的挑战，并提供培训和支持，以确保医疗专业人员能够有效地使用这些工具。综上所述，医学信息系统在提高临床决策质量和医疗效率方面具有潜力，但需要继续不断改进和优化。

PU-058

数智化协同模式下的全医嘱质效双控决策支持平台构建

徐蕾
无锡市人民医院

一、引言

随着我国公立医院高质量发展的推进，公立医院发展方式由规模扩张转向提质增效，运行模式由粗放管理向精细化管理过渡，需要通过持续改进医疗质量管理体系，提高医疗服务同质化水平[1]。以电子病历、智慧服务、智慧管理“三位一体”的智慧医院建设[2]对智能化的管理决策提出了新的要求，医院需要建立可持续改进的管理决策支持体系与工具，来提升精细化管理水平。本文通过数智化协同模式下的全医嘱质效双控决策支持平台构建，分享存在问题分析、建设思路与方法以及应用效果，望能为智慧医院建设提供一定的建设经验。

一、存在问题分析

经过调研发现，我院在医疗质控管理上存在以下两个问题：

1、医生对政策不了解，管控节点滞后：对于医生而言，在诊疗过程中对政策不了解，不足以直接做出判断和干预，只能等管理部门发现问题后再进行处理，滞后性严重。

2、管理效率低，管控面局限：对医院相关管理部门而言，面对日益精细化、个性化、过程化的监管政策和绩效考核指标，人为管理存在易出错、管不全、数据提取不方便、管控面局限的问题，常规的管理手段已经难以快速响应政策的变化。

当前的临床决策支持系统主要分为基于知识库和基于机器学习或深度学习两种类型形式[3]。Coiera[4]将 CDSS 的功能概括为 8 个方面，根据这 8 个功能的作用可归纳为 4 个类别：判别诊断类、疾病预测类、流程支持类和信息检索类。

包含合理用药系统、病历质控系统、诊疗知识库系统等。这些系统存在管控节点分散、管控功能单一；缺乏自定义维护功能；缺乏数智化协同管理决策能力、且无法从医嘱源头前置管控等问题。

综合以上问题，本文以数智化协同和前置管控为出发点，提出构建“全医嘱质效双控决策支持平台”，以“一体化全流程质效双控”为建设目标，以全医嘱业务流程为管理主线，以医疗质控管理中存在的痛点难点为抓手，积极探索并构建具有质效双控能力的智能化精细化管理决策平台。

二、全医嘱质效双控平台建设思路与方法

围绕质量安全、协同效能、医保控费开展调研及问题梳理，结合国家相关政策规范及制度，制订精细化的管理决策规则，建立基于临床大数据的管控决策知识库；通过智能知识引擎，将临床业务对象实体化，在临床知识与临床对象建立属性关联，搭建全医嘱质效双控决策支持平台。利用大数据 ETL 技术来实现临床数据的采集处理，选用开源 NoSQL 类型数据库 MongoDB 作为临床主题数据中心的核心数据库系统。利用领域特定语言（DSL）思想，结合 MongoDB 动态化数据结构的存储特点，通过决策树、表达式及关系二元组的有机整合，实现任意知识（规则）的自由表达和高性能解析及计算。利用 SpringCloud 框架实现全部业务组件逻辑微服务化，并采用 Actor 框架实现容器内各业务组件的高性能线程复用并发模式。主要功能包括：

1、全医嘱精细化管控：当医生提交医嘱时，系统将根据底层的管控决策知识库与患者的 360 诊疗信息相结合，进行实时化分析判断，给出警示。全医嘱精细化管控模型体系从业务上分为：用药管控、检验管控、检查管控、手术管控、护理管控、诊疗项目管控、病历辅助管控。

2、质控指标主动预警：系统可在临床诊疗过程中实时监测各类质控相关的预警指标和监管指标（如：平均住院床日、抗菌药物使用强度、平均住院费用）以及针对患者个体治疗相关指标（如：肝肾功能指标、血压、感染指标等），当触发警示信号后及时向对应的管理部门发送主动预警信息，提醒相关医务人员进行重点关注和相应的处置干预。实现医疗过程事中的过程化动态化的主动监测预警。

3、医保控费决策支持：以医保三目录、医保报销支付政策及医院相关管理制度为标准，在医生开具各类处方医嘱时，进行医保合理性合规性前置审核及警示干预，从源头上降低医院的医保控费风险和医保飞检压力。

4、临床决策辅助参考：查阅本系统内包括括疾病、检验、检查、手术、药品等在内的各类权威准确的临床参考信息。

5、决策知识库管理：知识库管理模块的操作界面和数据呈现方式直观可视，配置界面简单易操作。

三、应用效果

汇同医务处、药学部、医保办、财务科、公卫处，结合各职能监管工作中存在的实际问题整理制定 7 大类 20 多套共计 1020 个精细化管控规则并全部维护到全医嘱质效双控平台。涉及疾病诊断、临床用药、检查检验、手术麻醉、医保控费、护理和治疗。下面介绍六个典型应用案例及应用效果。

4.1 氧气吸入项目控费

在院内的医保基金违规检查中，发现项目收费存在多收、重复、互斥等不合理不规范问题，导致我院医保费用消耗增加，不仅直接影响医保合规性检查，还影响我院国考成绩。针对这些问题在质效双控平台上进行规则编排，如不能同时开具有互斥关系的医嘱、限制医嘱频次、限制加收项目的医嘱和限制同时开具部位重叠的检查项目。比如：医生开过“氧气吸入（小时）”后，不能再开“氧气吸入(日)”项目。管控规则在 2021 年 10 月上线，截止到年底 12 月经数据分析发现，各科室临床医生开具氧气吸入医嘱的合理率从 43%逐步提升至 75%，不合理的氧气吸入医疗行为明显减少，效果如图 1 和图 2 所示。

图 1 氧气吸入管控提示框

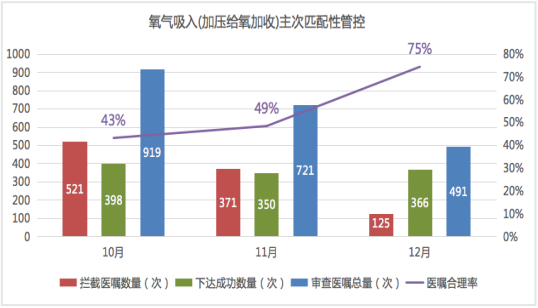


图 2 氧气吸入管控效果

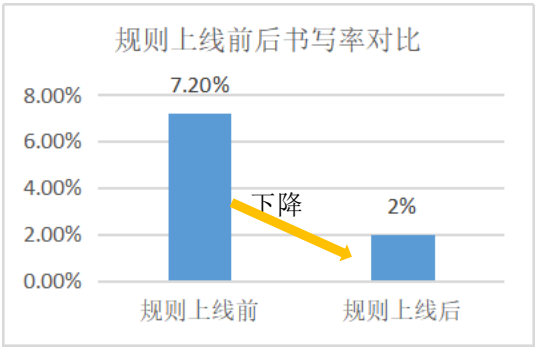
4.2 PPIs 的临床预防性用药控制

针对院内临床用药检查过程在中发现除了应用于消化系统的治疗性使用 PPIs，我院 PPIs 的临床预防性用药比例一直不断增加，尤其是在外科领域，部分医生为消化道损伤高风险的普通手术患者常规使用 PPIs；利用平台自定义能力，配置相关 PPIs 管控规则。提示效果如图 3 所示。PPIs 预防用药精细化管控自上线以来，运行效果良好。PPIs 各项运行指标提升明显（在使用 PPIs 的病历中随机抽取管控前后各 100 份进行点评分析，合理使用 PPIs 的医嘱比例由 46%上升至 88%，医生用药合理率大幅提升。PPIs 管控上线后，我院 PPIs 使用量明显减少。PPIs 使用金额从 2019 年上半年 1892 万降到 2020 年上半年 1650 万，同比下降 12.8%。），之前 PPIs 不合理使用问题均明显改善。在提高临床合理用药水平的同时也进一步降低了患者的经济支出。

图 3 PPIs 预防用药精细化管控提示框

4.3 电子病历时效质控

以电子病历文书书写时效性控制为业务抓手，结合患者病情及就诊方式，强制要求医生在不同的诊疗阶段及时书写相应的病历文书，否则系统将干预长期医嘱的正常开具。 管控前，统计的 21 年 4 月到 21 年 10 月，期间入院人次 38562，超 24 小时未书写入院录的病历人次 2773，书写超时率 7.2%；管控后，统计的 22 年 3 月至 22 年 9 月，期间入院人次 38401，超 24 小时未书写入院录的病历人次 780，书写超时率 2%。降幅达 5.2%。应用效果如图 4 所示。



4.4 深静脉血栓

针对院内 VTE 防控，质控办采取的方式主要是让护士（尤其是心内、心外、神内、骨科等重点科室）填写评估表，但是临床医生存在查看不及时且出现漏开处理医嘱等问题，导致院内深静脉血栓管控流程不规范，增加院内医疗纠纷风险。就质控办所提出的管理问题，从护理系统取数据，在全医嘱质效双控平台上配置 VTE 相关评估量表判断规则。当 Caprini>=3 或 Pauda>=4 时触发规则，提醒医生开 D-二聚体检测和血管超声检查，否则禁止医嘱下达。提示效果如图 5 所示。

图 5 VTE 防控管控效果

4.5 院感重点药物管控

根据《国家医疗质量安全改进目标的通知》，我院院感处牵头落实专项改进行动，协调相关部门，利用信息化手段，持续开展重点药物治疗前病原学送检相关监测。根据院感处对重点药物监测管理规则，利用平台自定义能力，自由配置相关管控规则。例如开具第二类重点药物注射用盐酸万古霉素，需要开具相关病原学检验项目医院，否则禁止下达医嘱。提示效果如图 6 所示。

图 6 重点药物监测管控效果

4.6 疫情期间公卫管控

在 2022 年我市卫健委颁布关于新冠肺炎疫情监测预警相关通知。我院利用平台规则自定义能力快速配置出院患者必须填写 SARI 患者表的管控规则，在开具明日出院医嘱时，自动判断 SARI 患者表的填写状态，起到提醒的作用，及时有效的保证了疫情期间诊疗流程规范。提示效果如图 7 所示。

图 7 SARI 患者表填写管控效果

四、结语

通过对数智化协同模式下的全医嘱质效双控平台建设，将我院各职能部门的管理质控需求融合于一体化平台，有效推动医院管理提质增效，助力医院一体化精细化管理。满足在事中管控，在医嘱源头进行管控，且要多系统联动判断的需求。为患者提供更加安全、高效的医疗服务。且该平台能快速响应政策要求，从业务科室接到需求开始到科室通过智能规则引擎自行编排新功能上线平均仅需要 1.5 天，提高了各职能部门的管理效能。但当前的质量管控主要集中在医生开医嘱阶段，围绕护士开展的管理质控工作做的还不够，需要加强推进研究。

PU-059

基于 PDCA 循环法的临床决策支持系统应用进展

解亚楼* 黄心莹
阜外华中心血管病医院

目的 临床决策支持系统是智慧医疗建设重点，通过大量临床数据，分析医疗行为，减少医疗差错。阜外华中心血管病医院响应国家智慧医院建设的号召，组织各部门,按照 PDCA 循环的工作模式，丰富临床决策支持系统的知识库，推进决策支持系统应用。
临床决策支持系统的根本目的是为了提升医疗质量。系统主要关注和医生的互动，辅助医疗行为，提供建议供临床医生参考。
临床决策支持系统基于它被使用的节点有不同的作用。诊断前，可完成对疾病的初步诊断。在诊断中，可以根据检查检验结果，修正初步诊断。诊断后，可以用于挖掘患者既往医疗信息、临床研究之间联系，以便于预测其将来的健康问题。
方法 1.基础工作
为保证按时完成对接,医院成立由医务、护理、药学部等多部门组成的联合工作小组,明确知识库和系统策略维护的内部分工和操作流程。
2.PDCA 循环方法

根据计划、实施、检查和处理四个步骤,对药品、检查、检验、护理知识库数据质量优化。
2.1 计划阶段
第一,对 2023 年一季度 10 万余条医嘱进行质控，筛选问题医嘱。
第二,按医生级别收集知识，包括但不限于抗菌药物使用、配伍禁忌、检查冲突、病历诊断和过敏史冲突等类型知识。
第三,原因分析:按照 80/20 原则选出 4 项原因: ①病历书写不规范。②信息系统不完善。③诊疗行为不规范。④知识欠缺。
第四,设定目标:根据影响因素,以优先提高诊疗质量为原则,分别设定相关指标的目标值:例如病案首页主要诊断、主要手术操作完整率的目标是 100%;医嘱数据质量合格率目标值第一轮定为 98%。
第五,制定计划:拟定系统改善流程,包括组织培训、数据质控等。
2.2 实施阶段
第一,组织培训学习:组织人员参加培训,规范诊疗行为,学习开展人机交互工作。
第二,优化信息系统:电子病历系统需要嵌入的业务流程,由各业务科室确认。建立知识和策略对应关系。设置判断规则,自动查找错误。
第三,加强知识库数据质控:完善院、科两级质控体系,加强环节质控,提高诊疗质量。
第四,建立沟通机制:工作小组建立工作群,及时沟通。
第五,制定奖惩措施:根据不同的专科特点和知识库贡献率,制定奖惩措施。
2.3 检查阶段
第一,病历数据质量检查:按照病历书写规范实施细则要求,在系统中检查病历数据的完整性和合理性。
第二,医嘱相关指标检查:医务部和药学部在系统中检查医嘱,违规药品、耗材、非适应症医嘱视为不合格。
2.4 处理阶段
每月汇总各指标检查结果,反馈存在问题,提出整改措施,科室及时调整数据;每季度向主管院领导汇报,落实整改措施,确定改进目标,进入下一轮 PDCA 循环。
结果 临床决策支持系统采用应用层，服务层和数据层三层架构体系。
3.1 数据层
数据层包含各种医疗知识，将政策法规、权威书刊，通过整理和审核，导入知识库。使用时，及时的将文献、规定等内容展示。
3.2 服务层
服务层包含搜索引擎和质控策略。
通过关键词来查找各种文献和案例，将需要的内容一一展示。类似于检索器的服务层，简化了临床决策支持系统的操作，方便了临床医生的查询。
管理部门设置了各种策略，如下达医嘱时，能够参考药品、检验结果、过敏史、诊断等至少 4 个维度进行检查并给出提示。
3.3 应用层
应用层利用系统分析电子病历，找到最适合的决策方案，在一定程度上减少了人为影响带来的失误。方式简单，处理迅速。
3.4 临床决策支持系统功能和演示
临床决策支持系统具有主动和被动的两种方式。
主动方式是临床决策支持系统主动的给出提示，避免医生可能采取错误的诊疗方案。如：医生开立处方时，系统采取主动的方式，根据知识库内容，判断处方、检查、检验的合理性。
被动的方式则是依靠医生的查询来发起，医生打开系统查询。此方式相当于一个数据库查询，而主动方式，则更像一名助手，协助医生共同治疗病人。
结论 临床决策支持系统，短期目标是提升病历书写质量,减少不良事件的产生。中期目标是优化电子病历,结合政策法规,对更好地服务临床。长期目标是深挖数据价值,规范医疗流程，提升医疗质量。
综上所述,应用临床决策支持系统能有效提升医疗质量, PDCA 循环能更加有效提高知识库的可用性,在提高医疗质量和改进管理模式的同时,提升临床决策支持系统的科学性和医疗费用管理的精准性。

PU-060

问题驱动的静脉血栓栓塞症智能临床决策支持体系构建

任晓霞[1]，黄桃[2]，陈虹[3]

1.重庆医科大学附属第一医院信息中心

2.重庆医科大学附属第一医院重症医学科

3.重庆医科大学附属第一医院呼吸与危重症医学科

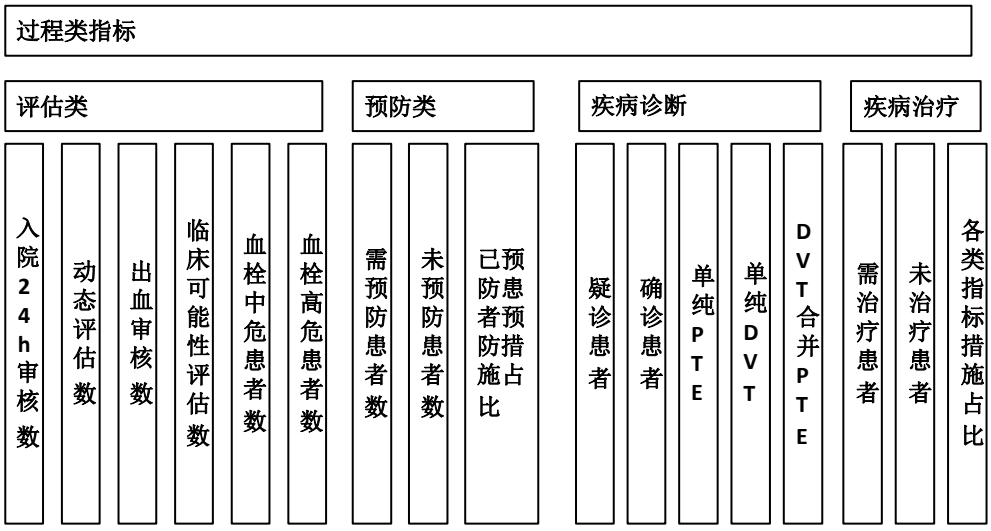
0 前言

为了规范静脉血栓栓塞症（VTE）的规范诊疗，自 2008 年开始，国家卫健委陆续颁布更新了《三级医院中心 VTE 防治体系建设标准》、《全国肺栓塞和深静脉血栓形成防治能力建设项目信息化应用和质控管理建议》等规范，要求医疗机构持续提高院内 VTE 防治能力[1]。至今全国二级及以上公立医疗机构 VTE 防治意识已大幅度提升，并利用信息系统初步实现了院内 VTE 防治流程的信息化[2-7]。随着 VTE 质量管理的逐步深入，多节点评估工作量大和不同医生诊疗经验的差异，导致 VTE 关键节点的评估质量和防治质量仍参差不齐。

临床决策支持系统利用人工智能技术，可实现标准诊疗规范和实际诊疗过程相融合，实现连接临床观察与临床知识，影响临床决策，改善临床结果的作用，并在很多的医疗领域取得了良好的效果[8-11]。针对 VTE 防治的临床辅助决策研究，之前多集中在辅助评估和简单的预防措施推荐，未形成针对全流程的诊疗决策体系[12]。本文参照国家医疗质量安全改进目标中要求各医疗机构建立本单位的质控监测指标类型要求，首先构建院内标准质控指标集，以此指标集为基础，分析防治和管理流程中存在的问题，综合诊疗规范、质控管理、合理用药、合理检验检查等多学科知识，建立构建 5 大临床诊疗知识库（CDSS），覆盖诊疗全过程实现 AI 赋能，数据驱动的 VTE 智能管理体系。

1 质控指标集构建

以国家质控指标为导向，结合院内病种精细管理和未来发展规划要求，建立覆盖 VTE 诊疗全流程的指标体系。主要分为过程指标和结局指标两大类。过程指标主要用于临床诊疗过程的监控，跟踪患者预防和治疗措施实施情况及结果；结局指标主要用于科室和全院的质量分析，发掘管控弱点制定改善方案，推进落实后检查实施效果。具体指标参见下图 1.



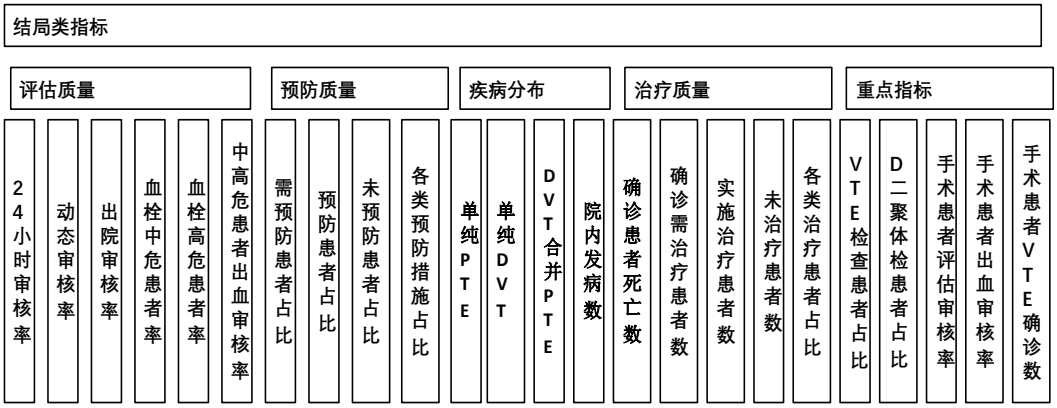


图 1 VTE 质控质控指标集

2 基于质控指标集的现状分析

2.1 过程类指标分析

存在入院 24 小时节点评估不及时；评估不准确，漏项情况；中高危患者出血风险率有待提高，动态评估率低。规范预防率不同科室差异较大，部分重点科室预防率仍偏低。经分析，评估表单的复杂性是导致的评估不准的主要原因，而不同医生之间 VTE 防治意识差异，是造成的关键节点评估不及时、漏评和未及时落实预防措施的症结所在，提醒级规则无法有效规范医生的行为；VTE 护理措施也存在未落实执行的情况。

2.3 结局类指标分析

质控人员对结局性指标分析困难，无法及时准确获取各类质控数据，在分析临床处置合理性时需查看各类资料，耗时耗力，导致质控效率降低，影响了对不合理诊疗行为的提前发现和预警。

3 数据驱动的 VTE 综合辅助决策质控体系构建

以指标分析发现的问题为导向，重点解决 VTE 各诊疗阶段评估效率、准确性、诊疗过程难规范和数据分析困难问题。以临床数据中心的患者多模态数据为基础，运用自然语言处理技术，将患者数据拆分为患者、疾病、检验、检查和治疗五大类因素，结合评估细项解读、诊疗指南和各学科诊疗经验，形成预警规则库 185 项，覆盖 9 个评估量表；诊疗决策规则库 75 项，聚焦预防、诊断和治疗 3 大阶段；文献库 37 项；事件控制规则 54 条，分三个层级进行关键节点的质量管控[10-12]。整体架构如图 2 所示。

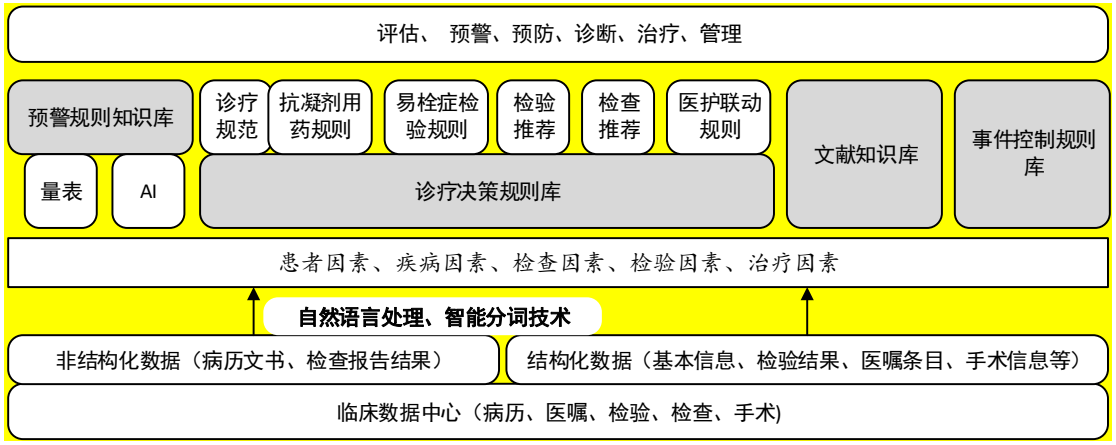


图 2 VTE 综合诊疗知识库

3.1 预警规则知识库：评估阶段自动评估风险等级，提高效率和评估准确性

根据评估标准解读，将每项评估项关联患者五大类数据，形成 185 项预警规则，针对所有在院患者，通过预警规则实时匹配患者数据，系统自动自动风险项评估，形成风险等级预警提示，将评估结果推送给医生审核，并对中高危患者进行重点预警。医生直接根据预警信息进行审核确认。

3.2 诊疗决策规则库：预防、诊断、治疗阶段推荐诊疗方案，提高诊疗规范性

将 VTE 防治指南和循证医学建议[13-16]，抽象成具体 75 条诊疗决策规则，嵌入医护工作流程。根据病人的个体信息、疾病信息等具体情况，结合风险评估、出血评估的结果，给出个性化的诊疗方案；结合临床可能性评估结果，给出 D-二聚体、肺动脉 CTA、血管超声等进一步检验、检查项目推荐。医生的评估结果同步给护理系统，并根据医生开具诊疗医嘱自动生成对应的护理措施，形

成医护一体化的诊疗流程，解决措施执行不到位问题。临床药师总结的预防抗凝药使用规则具体逻辑如图 3 所示。

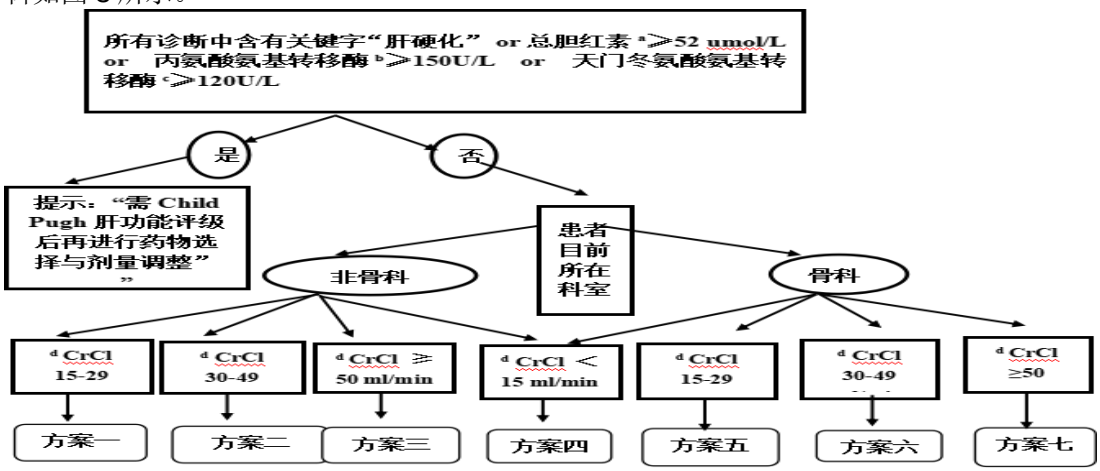


图 3 住院患者预防抗凝剂选择规则

3.3 文献知识库

综合 VTE 诊治相关的文献、指南形成文献库，辅助医生进行循证医学证据检索。

3.4 事件控制规则库：实现全流程的精细化质控管理

将患者分为普通、手术、肿瘤、重症等 8 大类；控制分为预警、提示、控制 3 个等级，根据患者类型、控制节点和控制等级 3 个属性，将需要管控的流程定义为事件，共构建 54 个事件规则。例如，事件 1：除入院确诊为 VTE 的患者在医嘱开立之前，强制需先完成风险评估,实现入院节点风险评估质控；事件 2：手术患者，开立手术申请之前强制需先完成风险评估，实现手术患者术前评估节点质控；事件 3 药品禁忌患者，仍开立抗凝医嘱，弹出风险提示等。

3.5 综合质控监管平台：整合 VTE 相关数据；提高分析效率

以质控指标集为基础构建指标数据平台，实时获取评估和诊疗数据。过程类指标实现医疗组长、科室质控员和医务管理者对所管辖范围内患者的 VTE 相关疾病发生情况、防治情况、确诊病例及不同类型血栓分布情况等关键指标的实时把控；结局性指标用于质量屏蔽和问题分析；将 VTE 相关诊疗事件整合集中展示，包含诊断信息、评估信息、手术信息、VTE 相关检验检查信息、病历记录信息等，结合指标值，辅助质控人员更高效得监督判断该患者诊疗过程的合理性。

4 结果分析

以质控指标集为标准，计算整改前后 3 个月的的关键指标值见表 1。通过对比发现，临床决策支持库的建立，对标 VTE 诊疗的各个关键环节，实现了关键节点的及时提醒与控制，院内 VTE 防治质量整体提升明显。

表 1 关键质控指标对比表

关键指标值	24h 评估审核率	出院前评估审核率	中高危患者出血审核率	预防措施实施率	院内新发 PE 患者数	院内新发 DVT 患者数	VTE 确诊患者数
整改前	13.7%	19.3%	17.1%	38.91%	43	201	220
整改后	100%	100%	100%	60.49%	29	132	144

5 讨论

医院标准 VTE 病种质控指标集的建立，为医院针对 VTE 病种的精细化管理提供了有力抓手，以质控数据为导向的综合质控体系构建，为医院进一步的诊疗流程优化和质量提升提供了更有针对性的切入点和目标评价办法。

基于 VTE 诊疗数据集、AI 技术、诊疗规则和流程质控相融合的一体化质控体系，为未来实施更多个性化的专病辅助诊疗提供了一个可扩展性的系统架构，在此基础上，可进一步扩展更加个性

化的预防、治疗方案，通过结果数据实现对不同方案的效果评估，找出最优方案，达到共识后，进一步完善疾病防治指南；另外，在质控指标集和辅助诊疗规则运行成熟后，可考虑在区域内推广，形成 VTE 区域质控中心的质控和管理标准。

PU-061

基于物联网技术的手术医疗行为管控系统设计与思考

刘阳
中南大学湘雅医学院附属海口医院（海口市人民医院）信息管理处

0 引言

随着互联网+医疗健康的不断成熟和普及，物联网、云计算、大数据等技术的应用已经成为智慧医院建设不可或缺的载体，物联网技术在医疗行业中的应用更是成为了大势所趋。医院手术室是外科手术和急危重症抢救治疗的基地,也是集多个学科、多个专业的人员于同一平面工作的复杂场所,手术室的运行质量直接影响医院运行效率。因此,基于物联网技术的手术医疗行为科学化管理势在必行。

1 存在的主要问题

医院手术室的人员及洗手衣鞋的管理还存在薄弱环节。进出手术室人员的准入控制难度大、洗手衣回收不到位、洗手衣鞋发放效率低以及统计难度大、手术衣管理不规范导致交叉感染风险、手术不准时造成手术间利用率低等问题一直以来困扰着手术室的管理。

1.1 手术人员准入控制难度大

进入手术中心的人员种类繁多，各类人员的权限难以把控，人员核查流程繁琐，引发现场人员混乱；人员饱和程度难以及时反馈，前台无法控制手术间可进入人员数量。

1.2 洗手衣回收不到位

手术衣物借用后不按规定归还，衣物丢失情况严重，造成物资浪费；非手术室人员随意穿着手术衣，影响手术室现场管理；后勤布草无法清晰核对送洗衣物数量，丢失或损坏的衣物无法追溯。

1.3 洗手衣鞋发放效率低以及统计难度大

传统的存衣柜、存鞋柜钥匙发放回收流程繁琐低效，导致长期占用的情况多发；存衣柜、存鞋柜的使用数据（如高频及低频使用时段、使用率、平均使用时长、最大使用人数等）难以统计，存衣柜、存鞋柜的管理配置优化寸步难行。

1.4 手术衣管理不规范导致交叉感染风险、手术不准时造成手术间利用率低

使用后/污染后的手术服乱扔乱放，不按着装规范穿着手术服，不按规范穿外出衣出手术室，加大交叉感染风险。手术医生迟到导致手术不准时，造成手术间利用率降低。

2 方法

2.1 总体架构

自助衣物/鞋柜管理
安全准入管理
人员行为管理
系统后台管理

2.2 平台内容

2.2.1 手术区域医疗行为基础管理系统

通过在手术衣中安装 RFID 智能芯片，可按尺码大小管理手术衣，能根据手术室的需要预先配置好手术衣类型，可维护更新手术衣的信息。

通过在拖鞋中安装 RFID 智能芯片，可智能化管理手术拖鞋，结合自动换鞋柜，系统可以自动绑定医护人员与其领用的拖鞋，能分组管理拖鞋，可维护更新拖鞋的信息。

2.2.2 手术区域医疗行为门禁控制系统

系统需根据严格的准入控制管理机制，针对医护人员设置不同的准入条件，如各个门禁点的设置控制，准予通过的条件设置，医生出入各个流程过程中信息的记录，整个流程信息的追踪管理。

系统能根据医院手术室实际管理流程设定手术室出入、换衣、换鞋等流程，能够借助门禁、自动发衣机、自动换衣柜等自动化设备，自动通过设备屏幕进行视窗提示或语音提示违反相关流程的医护人员。

2.2.3 手术区域医疗行为设备控制系统

自动发衣机控制。医护人员在任何一台自动发衣机、换鞋柜或换衣柜刷 NFC 卡，系统能够自动与中心自动货控制服务器比对识别人员身份，并作出发衣、开门、提示等，并自动发放相应尺寸的手术衣（如：大、中、小号）。系统具有群组管理能力，手术室内所有的自动发衣机都在一个平台上进行控制及管理，系统具有衣物数量告警功能，当自动发衣机中某种型号的衣服低于设定的数量，系统会自动进行提示，提醒工作人员及时补充衣服。

自动收衣机控制。医护人员领取衣服时，将衣服与医护人员绑定，医护人员在出手术室之前，将所领的衣服放到自动收衣机的回收托板上，刷 NFC 卡后，收衣机能通过红外物品检测装置或 RFID 检测装置对衣服进行识别并自动进行回收，并且在后台解除绑定，以便医护人员具有权限开启手术室大门或下次领衣权限。支持还衣过程拍照，便于可后期图片查看。

自动收鞋机控制。医护人员领取拖鞋时，将拖鞋与医护人员绑定，医护人员在出手术室之前，将所领的鞋放到自动收鞋机的回收托板上，刷 NFC 卡后，收鞋机会通过红外物品检测装置或 RFID 检测装置对拖鞋进行识别自动回收，并且在后台解除绑定。支持还鞋过程拍照，便于可后期图片查看。

自动换鞋柜控制。手术室配置有多组自动换鞋柜，所有自动换鞋柜都安装有智能识别及控制器，并连接 IP 网络，由后台统一管理，协调工作，医护人员只要在换鞋柜的刷卡区域刷 NFC 卡，系统自动将所检测到的医护人员相关信息与中心服务器数据进行交互通信，根据持卡人的身份权限就近开启自动换鞋柜的柜门或进行相应的提示，并自动记录柜门开启时间。

自动换衣柜控制。手术室配置有多组自动换衣柜，所有自动换衣柜都安装有智能识别及控制器，并连接 IP 网络，由后台统一管理，协调工作，医护人员只要在换衣柜的刷卡区域刷 NFC 卡，系统自动将所检测到的医护人员相关信息与中心服务器数据进行交互通信，根据持卡人的身份权限就近开启自动换衣柜的柜门或进行相应的提示，并自动记录柜门开启时间。

2.2.4 手术区域患者通道智能管理系统

医患身份核对管理。医护人员信息系统与患者腕带扫描系统融合，患者信息与医生信息绑定，进门刷卡即可核实患者信息。

3 结果

3.1 运用物联网 RFID 标签感应技术，进行衣物自动发放和批量扫描

运用物联网技术，在手术衣物中植入 RFID 电子标签，将标签的唯一编码与物品绑定，可实现对物品的信息化管理。

识别人员身份后，自动发放医护人员相应尺码的手术衣物，规
避衣物乱发滥用情况。可选择使用批量扫描仪，批量感应衣物，快速核对衣物数量，且衣物状态全追溯。

3.2 智能柜自动分配和后台统计数据

存衣柜、存鞋柜智能化，人员登记后可配置固定柜、流动柜规则。柜号即时显示，可刷卡等方式随时查询。手术衣归还后可延时自动释放人员对应的存衣柜与鞋柜。

存衣柜、鞋柜的所有使用数据在系统中保存，可查询生成数据报表，协助优化资源配置、管理方式。

3.3 手术室排班和流程管理优化

3.3.1 与手术麻醉信息管理系统全面集成，根据手术排班信息自动配置人员进入手术室的权限。

3.3.2 人员识别放行

门禁多方式识别人员身份，访问系统权限情况予以放行。

3.3.3 与管理台可视通话

支持来访人员与管理台进行可视通话申请进入手术室权限。

3.4 手术时间和回收闭环管理

3.4.1RFID 反查询：运用物联网技术，反向查询乱放衣物责任人，进行相应管理。

3.4.2 回收闭环管理机制：系统可设置为未正确归还衣物的医护人员无法再次使用系统内的任何设备，直到正确处理异常。

3.4.3 手术时间管理功能：手术前多方式提醒医生，后台记录所有数据形成报表，针对性规范行为，提高手术开台准时率。

3.4.4 行为管理大屏：显示数据，显示不规范人员行为名单，一屏监控不规范行为。

4 讨论与思考

4.1 建设成效

4.2 问题讨论

5.2.1 业务实现复杂度高

5.2.2 个人隐私保护等安全问题不容忽视

5.2.3 手术室管理需要形成有效地管理机制

5 结论与展望

随着海口市人民医院基于物联网技术的手术行为管理系统成功搭建，让物物相连的手术室的管理水平能够再上一个台阶。

PU-062

放射报告分配流程优化

黄娟，屈健宁
重庆医科大学附属第二医院

放射信息系统（Radiology Information System, RIS）的基本功能包括检查登记、分诊管理、数据采集、报告书写、报告审核、查询统计等，其中报告书写、审核流程面临一个难点问题，就是报告如何相对公平的分配给当日值班医生，缩短报告的完成时间、提高诊断医生的工作效率[1]。重庆医科大学附属第二医院放射科根据多年的科室管理经验，联合厂商对RIS放射的报告分配流程进行优化，探索出了一条适合本科室的报告分配模式，较好的满足了科室的业务需求，并取得满意的效果。

1 现状

我院放射RIS系统已运行多年，根据科室需求，利用PDCA循环对各流程进行持续改造[2]，其中报告分配的改造始终不能较好的满足科室的业务需求。初写报告由实习生、进修生、研究生等完成，他们的工作量不纳入绩效考核，所以对报告分配没有要求，问题主要集中在审核报告如何分配，经过两次较为典型的改造后仍存在一定的问题，具体为：（1）首先科室根据诊断医生专长进行分组，分为四个专业组，包括腹部组、神经头颅肌骨组、心胸乳腺组、普放组，在对患者信息登记时根据检查项目进行自动“检查分组”，分组后的医生只能书写该检查分组的报告，但会造成一个问题，医生会尽量选择简单、权重系数低的报告先书写，越复杂、难度系数越高的报告迟迟没人书写，极大地延误的此类报告的完成时间，甚至延长了病人的住院时间。（2）以报告的影像号尾号作为分发规则，按照人工约定的方式随机分配给当日的诊断医生。比如腹部组今日有三个诊断医生A、B、C，那就约定影像号尾号为1、3、5由A来完成，2、4、6由B来完成，7、8、9由C来完成，那影像号尾号为0的由谁来完成呢？该分配模式在分配的均衡性方面又存在一定的问题[3]。

2 报告分配流程改进

针对系统及科室存在的问题，放射科根据自身需求出发，联合RIS厂商对系统流程进行改造。急诊病人不纳入报告分配规则中，门诊、住院、体检患者纳入报告分配队列中，其中体检病人有独立的报告分配流程。主要有两种规则：（1）多劳多得。对于当日进行报告分配的诊断医生，会设置已分配待审报告的最大值，只要用户的实际已分配未审的数小于该最大值，就可以给继续该诊断医生分配报告。该模式下只要诊断医生审核的报告越多，分配给诊断医生的报告数越多，这样可以极大提高诊断医生的工作积极性，鼓励医生多劳多得。（2）公平原则。报告分配时不考虑每位诊断医生的报告完成情况，报告分配尽量向公平、公正的原则。每份报告按部位数计算，或者按部位数*权重值计算，将报告相对公平的分配给当日值班的诊断医生[4]。

放射科目前根据科室发展需求，在报告分配的规则上选择了相对公平的第二种规则，且报告只分配给审核医生，具体的流程如下：

2.1 配置收费项目的部位数

放射科完成收费项目的部位数设置，不同的收费项目部位数不同，为后期分配报告时计算部位数做好准备。

2.2 医生排班

放射科管理员每周根据分组后的诊断医生进行排班，排班机制灵活，可以为诊断医生单独设置上午班、下午班、长白班三种排班，排班时可以为每个诊断医生设置分配系数、最大份数、最大分值、是否含VIP报告权限等，见图1，这些参数为报告自动分配时计算数值提供了依据[5]。具体的代表含义为：

分配系数：为该组里某个医生单独设置分配系数，比如给A设置分配系数为1.5，其余人都为1。当总共350份报告需分配给3人，A的工作量为150份报告，B和C各自为100份。

最大份数：由于某些特殊原因，A今天只能完成50份报告后将不再为其分配报告。

最大分值：可以为某诊断医生单独设置报告的分值数；

VIP报告权限：该组中的VIP患者只由某位诊断医生单独书写，其他人无权限。

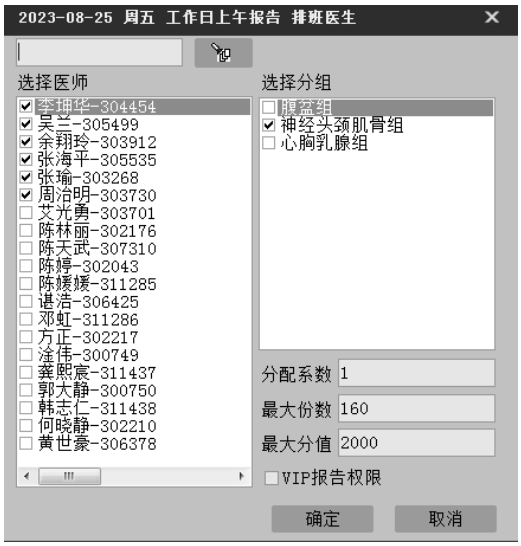


图1 医生排班

系统默认上午班时间为5:00-13:00，下午班开始时间可以灵活调整，截止时间为19:00，长白班时间为5:00-19:00。如某天腹部组检查患者特别多，腹部组的医生分为上、下午班，当上午班的每个诊断医生的最大份数为160份时，可以提前修改下午班的时间，让上午的报告提前分配给下午的医生，防止出现上午诊断医生工作量特别饱和，而下午诊断医生工作量不足的情况，系统设置如图2：

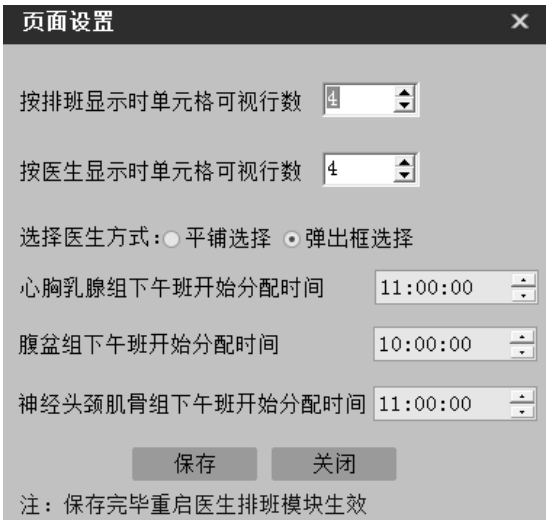


图2 下午班时间调整界面

2.3 制定报告分配策略

放射科报告分配策略是：

- 首先统计该组里所有诊断医生已分配报告的收费项目；
- 根据收费项目的部位数对照关系，计算出相应的部位数；
- 找出符合分配且已分配部位数最少的诊断医生；
- 分配给该诊断医生时间最早或优先级最高的报告。

报告分配的机制定时循环执行，每一分钟执行一次。图3表示根据医生的排班自动分配报告的情况，具体的数值分别表示“已分配待处理报告数/已分配报告数/已分配部位数”。

图3 某天报告分配图

3 效果

报告分配功能上线后，分配合理，使用便捷，得到了放射科医生的一致好评，同时缩短了报告的书写完成时间[6]。自2023年7月报告分配功能上线后，对比放射科2023年4月和2023年7月病人检查完成后，CT和MR报告平均诊断时间，统计发现，报告平均书写诊断时间缩短，医生诊断效率提高，数据如表1所示

表1 报告分配功能使用后时间对比表（分）

	2023年4月	2023年7月	缩短百分比
CT	954	899	5.77%
MR	531	483	8.26%

4 讨论

报告分配是医学影像信息系统充分利用信息技术替代人工操作的技术改造，科室精细化管理的有效手段，严谨的规则设计充分取代了人工分配的随意性，避免了科室员工挑选权重系数低的报告，缩短了报告的完成时间，减少病人等待时间，提高放射医生与临床医生的工作效率，为科室的绩效改革奠定了基础[7]。

在系统试运行阶段，放射科本着报告分配时相对公平的原则，报告分配时只考虑了部位数的数量问题，没有计算该部位数的权重值，没有体现出报告的难易程度，并不能从根本上解决放射科的问题。本着PDCA循环持续改进的原则，系统已根据科室需求设置收费项目的权重值，后期在报告分配时如何考虑权重值的问题值得讨论，但我们相信，只有规范、合理的规则设计才能提高放射科医生的工作效率和积极性。

PU-063

2014-2018 年河南省某三甲医院超长住院日患者临床特征分析

隋美丽

郑州大学第一附属医院病案管理科

平均住院日（average length of stay，ALOS）是指一定时期内每一出院患者平均住院时间的长短，即出院患者占用总床日数/同期出院患者，是反映医疗服务效率、评估有效利用医疗卫生资源情况和医院管理水平的一项重要指标[1]。超长住院日严重影响医院整体平均住院日，不仅加重了患者的经济负担，而且浪费了大量宝贵的医疗资源本研究[2]。在确保医疗服务质量的前提下，有效缩短平均住院日，充分利用现有卫生资源，提高医院整体运行效率，是医院发展的大势所趋[3]。《三级综合医院评审标准（2011版）》明确要求对住院时间超过30天的患者进行管理，并纳入医疗质量安全管理考核[4]。本研究通过对 2014年1月1日至 2018年12月31日河南省某三甲医院超长住院日患者临床特征的分析，结合实际情况提出对策性建议，为缩短患者平均住院日，合理利用医疗资源，提高医院质量管理水平提供科学依据。

1 资料与方法

1.1 资料来源 通过医院信息系统提取河南省某三甲综合医院 2014 年 1 月 1 日至 2018 年 12 月 31 日收治的超长住院日[5-6]（即住院日≥30d）患者的病案首页信息。排除标准：资料不完整者，包括就诊时间、性别、 年龄和/或诊断缺失；剔除诊断不明确、无法归类者。最终采纳资料完整者 93070 人次。疾病诊断按照国际疾病分类 ICD-10 编码归类，以病案首页主要诊断作为判定标准。

1.2 研究方法：查阅病历资料，提取的信息包括住院患者的姓名、病案号、住院流水号、出生日期、性别、年龄、婚姻状态、民族、职业、户口所在地、工作单位等一般信息和患者住院时间、出院诊断、病理分型、肿瘤分期等疾病信息。

1.3 统计学分析：使用 WPS Office 2022 和 SPSS 22.0 软件进行数据预处理和处理，分类变量采用频数或频率 n（%）进行描述，构成比、病死率等指标的组间比较采用χ^2 检验，检验水准 α=0.05。

2 结果

2.1 基本情况

在 2014-2018 年间，超长住院日患者达 93070 人次，2014-2016 年逐年下降，2017 年有所上升，2018 年下降；共死亡 628 例，病死率 0.67%，不同年度病死率差异无统计学意义（χ^2=6.30，P>0.05）。见图 1。

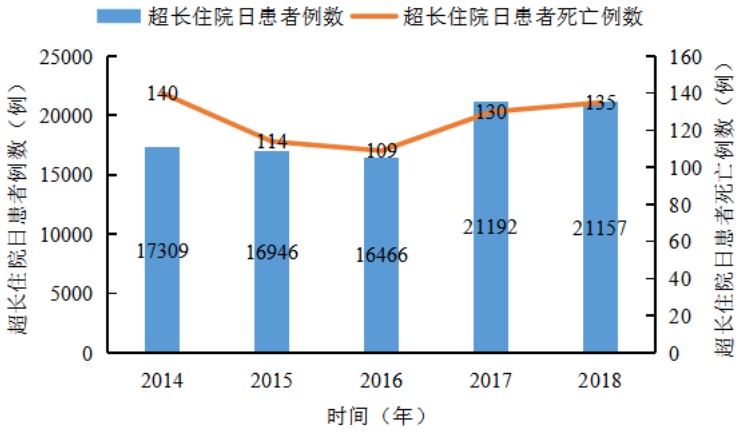


图 1 2014-2018 年河南省某三甲医院超长住院日患者例数及死亡例数

2.2 性别、年龄分布：男性 53029 例（56.98%），女性 40041 例（43.02%），男女比 1.32:1；男、女性患者年龄分别以≥60 岁-组和 45 岁-组为主，分别占 31.68%和 35.32%，不同性别的年龄构成差异有统计学意义（χ^2=287.40，P<0.001）。见图 2。

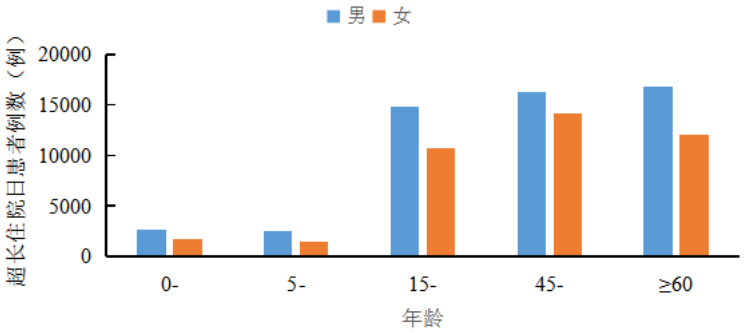


图 2 2014-2018 年河南省某三甲医院超长住院日患者的年龄分布

2.3 超长住院日患者疾病构成：在 93070 例超长住院日患者中，疾病分布前 5 位依次为 C00-D48、Z00-Z99、I00-I99、K00-K93、N00-N99，共 63377 例，占 68.10%。不同性别住院患者疾病构成差异有统计学意义（χ^2=2270.73，P<0.001）。见表 1。

表 1 2014-2018 年河南省某三甲医院超长住院日患者疾病分布构成及顺位 [n(%)]

疾病分布	男性		女性		合计	
	例数及构成比	顺位	例数及构成比	顺位	例数及构成比	顺位
A00-B99	2366(4.46)	8	1414(3.53)	9	3780(4.06)	8
C00-D48	12560(23.69)	1	10173(25.41)	1	22733(24.43)	1
D50-D89	621(1.17)	15	497(1.24)	14	1118(1.20)	14
E00-E90	780(1.47)	13	667(1.67)	12	1447(1.55)	12
F00-F99	463(0.87)	17	359(0.90)	16	822(0.88)	17
G00-G99	1885(3.55)	9	1256(3.14)	10	3141(3.37)	9

H00-H59	256(0.48)	18	174(0.43)	19	430(0.46)	18
H60-H95	67(0.13)	19	50(0.12)	20	117(0.13)	20
I00-I99	6391(12.05)	3	3994(9.97)	3	10385(11.16)	3
J00-J99	3200(6.03)	7	1665(4.16)	7	4865(5.23)	7
K00-K93	4155(7.84)	4	2371(5.92)	5	6526(7.01)	4
L00-L99	642(1.21)	14	357(0.89)	17	999(1.07)	15
M00-M99	1398(2.64)	10	1683(4.20)	6	3081(3.31)	10
N00-N99	3814(7.19)	5	2390(5.97)	4	6204(6.67)	5
O00-O99	0	20	209(0.52)	18	209(0.22)	19
P00-P96	807(1.52)	12	582(1.45)	13	1389(1.49)	13
Q00-Q99	520(0.98)	16	459(1.15)	15	979(1.05)	16
R00-R99	1254(2.36)	11	886(2.21)	11	2140(2.30)	11
S00-T98	3727(7.03)	6	1449(3.62)	8	5176(5.56)	6
Z00-Z99	8123(15.32)	2	9406(23.49)	2	17529(18.83)	2
合计	53029(100.00)		40041(100.00)		93070(100.00)	

3 讨论

平均住院日是一个评价医院医疗效益和效率的综合指标，平均住院日的长短涉及医院工作的各个方面，主要与医疗技术水平、护理水平、收治病人类型、手术构成、医疗设备等因素有关[7-8]。本研究从统计情况看，该院在 2014-2018 年间超长住院日患者病死率较稳定。

从患者性别分布可知：男性超长住院日患者例数是女性的 1.32 倍，这可能与男性患者的社会属性有关，男性一般都承担着更大的社会责任、更多的工作生活压力，并且危险性较高的特殊职业大多数亦为男性，其发生伤害等的风险明显高于女性，增加了超长住院日发生的概率[9-10]。

从患者年龄分布可知：不同性别的年龄构成差异有统计学意义,男、女性患者年龄分别以≥60 岁-组和 45 岁-组为主，关于年龄分布差异可能的原因，既往文献鲜有提及[7-10]，考虑可能与此年龄段女性人群为急慢性疾病高发期有关，提示该类人群健康状况值得重点关注。

从疾病构成来看：在 93070 例超长住院日患者中，疾病分布前 5 位依次为 C00-D48、Z00-Z99、I00-I99、K00-K93、N00-N99，男、女性患者前三位排序一致，第四位和第五位病种相同但排序有所不同。关于 C00-D48 肿瘤相关疾病，大多需要手术且术后恢复时间较长，从而增加了超长住院日发生的概率[10]；此外，许多慢性疾病需要定期治疗，治疗期间因同时合并其他合并症或并发症从而增加了平均住院日[9]。

总之，超长住院日的存在占用了大量的医疗资源，造成医疗资源分配不均，医院管理部门应根据超声住院日患者的疾病分布等情况合理分配资源，将超长住院日的天数和患者数量控制在合理范围。

PU-064

基于淋巴结阳性率等多因素预测完全切除的 III-N2 期非小细胞肺癌术后放疗的生存期的机器模型的开发和验证

郭强，魏益平

南昌大学第二附属医院

目的 多模式治疗是局部晚期(LA)非小细胞肺癌(NSCLC) (IIIA-C 期)患者的标准治疗方法。术后放疗(PORT)常被推荐用于 III 期 N2 非小细胞肺癌患者，但其作用一直存在争议。淋巴结比率(LNR)，定义为阳性 LNs 的比率除以总 LNs 数，并已被证明在几种实体癌类型中是一个强大的独立预后指标。也有研究表明 LNR 也是 N2 病患者的独立预后因素。更高的 LNR 与更低的生存率相关。然而，关于 LNR 与 PORT 之间的相互联系的研究却很少。机器学习能够从越来越多输入算法的数据中学习,并且可以给出数据驱动的概率预测。这种在当今大数据应用中快速有效地利用和应用高度复杂算法的能力是一种相对较新的发展。与传统的线性模型(最常见的是线性、logistic 和 Cox 回归)相比，ML 模型方法在解释非线性关系和相互作用方面提供了优越的能力，并可能超越这些技术的预测能力。因此，我们对 SEER 中已完全切除的 pN2 非小细胞肺癌患者进行了回顾性分析，利用基于机器学习的技术，纳入 LNR 等因素，开发并验证一种可对 III-N2 期非小细胞肺癌术后放疗的患者进行生存期的预测。

方法 回顾性分析 SEER 数据库（2010-2019）的 1647 例接受了完全切除手术的 III-N2 期非小细胞肺癌术后放疗的患者。使用逻辑回归（LR）、决策树（DT）、随机森林（RF）和支持向量机（SVM）来构建用于分类的机器学习模型。使用以下指标评价模型：曲线下面积（AUC）、校准曲线（CAL）、决策曲线分析（DCA）、一致性指数（C-指数）和 Brier 评分。

结果 LNR、年龄、性别、种族、原发部位、分级、病理、T 分期、手术方式、放疗等 10 个因素被选择纳入模型，并使用 deeolearning，drf，gbm 和 glm 建立模型。在训练队列中，AUC 为 0.8084,0.9223,0.7425,0.6354（图 1）;在验证队列中，AUC 为 0.6304、0.5919、0.6513 和 0.6290（图 2）。显示了令人满意的预测准确性; DCA 表明 RF 模型具有较好的临床应用价值（图 3）。

结论 基于淋巴结阳性率等多因素建立的预测模型可预测完全切除的 III-N2 期非小细胞肺癌术后放疗的预后状态。

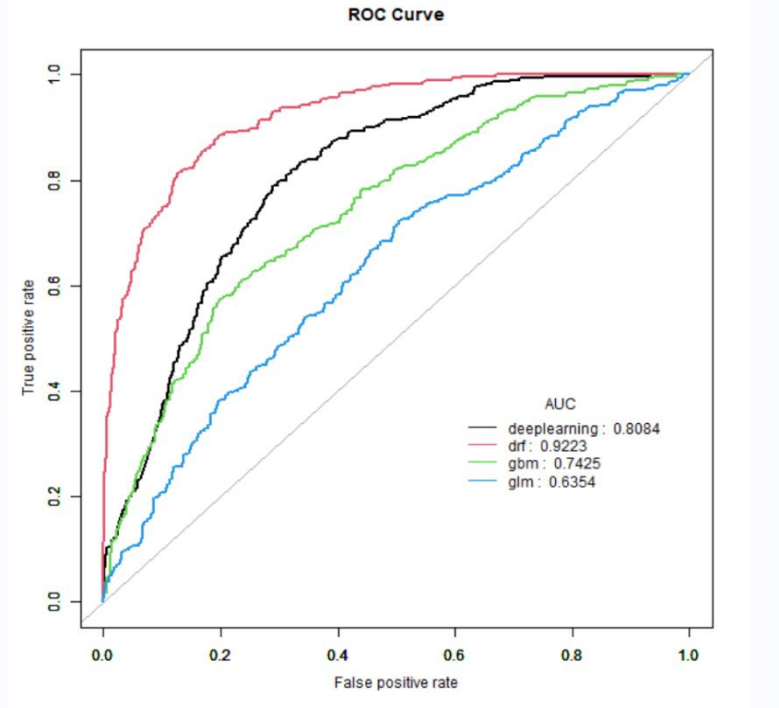


图 1 训练队列的 ROC 曲线

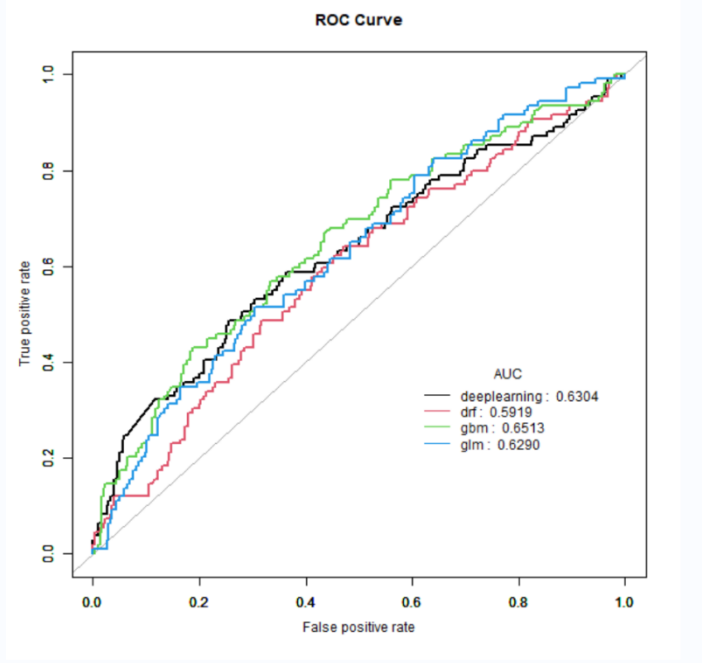


图 2 图 1 验证队列的 ROC 曲线

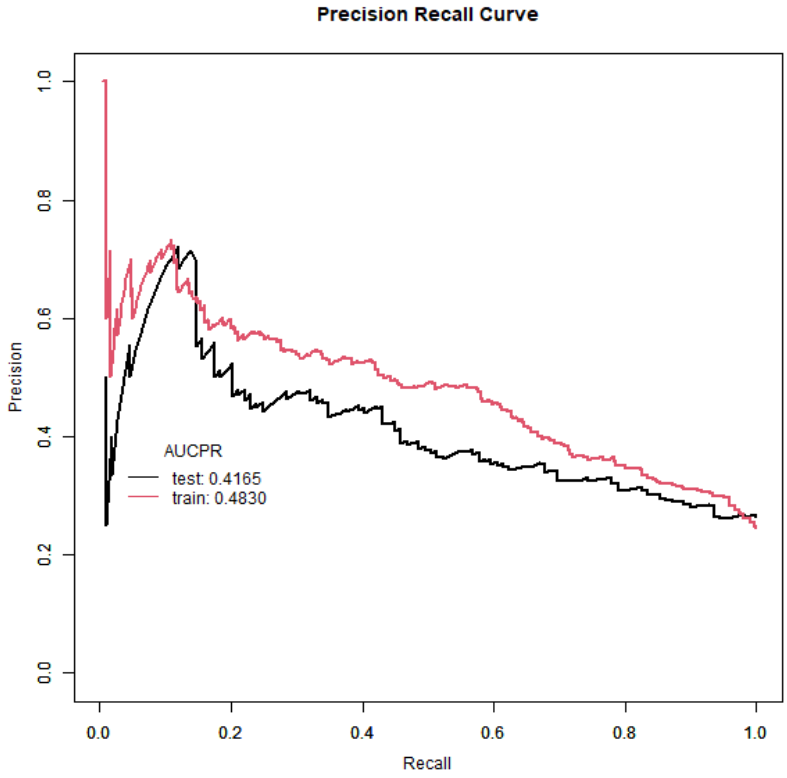


图 3 DCA 曲线

PU-065

人事报工管理系统在医院精益化管理中的应用

朱骄锋*
国药同煤总医院

1 背景

目前国内大部分医院临床医技科室的信息化发展较早，但是职能科室发展相对滞后，信息化技术可以促进医院各项业务推进和发展，加快职能科室信息化建设是促进医院发展的必然要求。医院员工的绩效是人力资源管理非常重要的环节，人为管控与计算必然会造成很多问题，之前我院人事报工流程为考核办管理员根据其他科室提交上来的数据，计算出本月每个科室的总绩效效益考核表，然后我院考核办人员将该表打印出来按科室拆分，逐个科室进行下发说明到科室报工员，科室报工员再根据该表计算出每个人的绩效工资明细，形成表格，将此表最终打印出来提交到人事科。人事科需要把每个科室计算的数进行统计，如果遇到多科室报工还需要汇总，最终形成员工绩效工资，而且科室进度也无从知道，此流程工作量大、跨度时间长，同时也无法实现精准化管控。

2 解决方案

针对我院报工流程存在的问题，我院进行人事报工系统开发，优化工作流程，提高工作效率。

具体从以下几个方面着手解决：

1、增加考核管理：通过实现考核办报表电子化功能，考核办人员制作完成考核表之后，只需要统一发送，每个科室报工员则可以看到各自的考核办报表，无需考核办人员逐一下发该表。并且在考核管理中增加关于绩效的附件查看功能。报工员可以查看具体绩效核算标准。彻底解决考核办考核管理员手工下发考核表和说明的问题，实现无纸化管理。

2、增加报工管理：报工员根据考核办发布的考核表进行拆分计算，将本科室每个员工的绩效进行计算，系统增加单科室与多科室逻辑规则，报工表的每项会与考核表每项进行校验，不通过的无法进行上报，校验通过直接提交，人事科即可查看到精确地科室绩效明细，真正做到了精准化管控报工。

3、绩效发布：人事科可以通过报表实时查看报工员报工进度，并在最终所有科室全部报工完毕情况下进行绩效统一发布，系统会自动进行逻辑计算，将所有员工的工资统一汇总，形成员工绩效工资列表。

3 具体功能应用

3.1 考核办管理：

对医院手工录入的每月考核办的【效益考核数据】进行全程电子化管控，实现考核办绩效表的信息化，包含考核数据添加、删除、修改、附件管理、导入、导出、发布功能。不用繁琐打印以及人工下科室处理。

附件管理：将每个科室的效益考核数据进行附件电子化处理。

导入：将【效益考核数据】表格，进行自动获取匹配科室然后自动存入系统记录中。

发布：统一发布，根据每条科室数据以及后台维护的当前科室的考核接收员进行自动发布。

3.2：报工管理：

对医院手工录入的每月报工进行全程电子化管控，实现报工管理的信息化，包含报工数据人员花名册、添加、删除、修改、导入、导出、提交、上传、发布、驳回、考核校验功能。省去繁琐的人力过程以及计算与打印，提高工作人员的效率，实现无纸化管理。

人员花名册：将本科室本月参与报工的员工进行管理，包含添加、删除、维护

导入：将报工表格的所有项内容，根据人员编号获取数据然后自动存入系统记录中。

提交：统一提交，根据每条人员数据以及后台维护的当前科室的报工审核员进行自动提交并带有消息提醒。

上传：统一上传，根据每条人员数据以及后台维护的当前科室的报工管理员进行自动上传并带有消息提醒。

驳回：当管理员发现数据问题，可以进行驳回，并可以加原因，录入人可以直观看到消息进行修改。

考核校验：系统有大量逻辑规则，包含出勤天数与病假、事假、旷工的逻辑判断，也包含自动将所有报工项与考核表的多项目进行智能校验，同时实现多科室与单科室的考勤校验，彻底实现精准化管控。并且提交的时候会进行自动考核校验，有一项不通过则无法提交本月报工数据。

3.3：科室报工情况：

可以对全院所有部门报工情况进行实时查看（包含未录入、未发送主任、主任未审核、已完成）。并可以查看历史报工数据。

3.4：本院绩效发布：

将报工表的项进行整合合并，然后按员工类别一键生成全院所有员工的工资绩效列表，如遇到多科室报工自动合并汇总。

3.5：邮箱功能：

进行一般电子邮箱的收发功能，也可以添加图片、文档。加强员工的互相沟通，为办公自动化提供便利。

3.6：用户管理：

按科室类型划分全院人员进行统一管理，有员工基本信息，以及报工办或考核办相关权限的录入功能。包含姓名、性别、工号、部门、出生日期、职称、电话、工作单位、用户级别、人员级别、人员类型、人员状态、主管属性。

1：员工档案管理：对单位每个员工进行统一管理，将每位员工信息保存到员工档案记录中。

2：新聘用员工进行增加记录到档案记录中，对于解聘员工从档案记录中进行类别更改。

3：当员工信息发生变动，可以修改员工档案记录中的相应的属性。

4：报工员与考核员权限的录入。

5：多科室权限的录入功能。

3.7:部门管理：

将全院科室按临床科室、管理部门、医技科室、医辅部门等进行分类管理

1：科室档案管理：对全院每个科室进行统一管理。每个科室保存到档案记录中。

2：对于科室的增加修改可以进行维护。

3.8：级别管理

将员工级别进行分类维护：其中有报工用户、考核用户、报工考核用户、系统用户。

1：报工用户有操作报工管理权限。

2：考核用户有操作考核办管理权限。

3：报工和考核用户有操作报工考核双权限。

4：系统用户操作所有功能权限。

3.9：密码保护：

由于涉及到工资与绩效，该软件的用户密码进行重置的时候会进行 6-20 位数字与字母保护设置，提高密码强度，防止被别人轻易破解。

4 总结

我院进行人事报工系统建设，把本单位人员从输入、计算、统计等繁重的工作中解脱出来，极大提升了医院所有部门绩效工资工作的效率，同时利用信息化的精准管控，加强了医院的绩效和考勤管理。相信今后利用信息化可以让医院的发展、管理得到进一步提升。

PU-066

日间手术模式下的信息化建设方向探究

曹靖霄
山西省眼科医院

1. 概述

1.1 日间手术定义

2003 年，国际日间手术学会（IAAS）对日间手术给出如下定义：病人在入院、手术、出院在 1 个工作日内完成的手术，不包括在医师诊所或医院开展的门诊手术。对需要过夜观察的患者，则建议称为“日间手术—延期恢复病人”，其定义是，在日间手术中心/单元（独立的或在医院内

的）治疗的病人，需要延期过夜恢复，次日出院。

2015 年中国第三届日间手术年会上，中国日间手术合作联盟（CASA）结合国内日间手术的实际情况，正式发布对日间手术的定义：病人在 1 日内（24 小时）完成入、出院的手术或操作，需要说明的是，日间手术是对病人有计划进行的手术和操作，不包括门诊手术。另外对于特殊病情需要延期住院的病人，住院时间最长不超过 48 小时。

1.2 日间手术发展情况

近年来，日间手术在国内受到了医疗机构极大的关注。自 2005 年开始，上海、天津、武汉、成都等地的医院陆续开展日间手术，据国家卫生和计划生育委员会 2014 年调查，全国已有 100 多家医院开展日间手术。在学术组织方面，国家卫计委卫生发展研究中心于 2013 年 3 月成立“中国日间手术合作联盟”，推动国内日间手术学科发展。我国在同年 5 月被国际日间手术学会（IAAS）正式接纳为成员。目前，一些医院日间手术占择期手术的比例已超过 20%。总的来看，我国日间手术数量虽然难以与发达国家相比，但日间手术已经逐渐被国内医疗机构和医生所认可，上升势头明显，发展前景广阔。

截至 2017 年 12 月，全国有一半以上的三级医院开展了日间手术，占择期手术比例提高至 12.8%，其中 639 家医疗机构设置日间手术中心，同比增加 243 家；1340 家医疗机构开展日间病房服务，提供化疗、新生儿蓝光照射、康复理疗等服务，7825 家医疗机构实施临床路径管理，86%的医院实现同级检查检验结果互认。

1.3 日间手术管理模式

医院内的日间手术通常有以下布局管理模式：

集中独立的日间手术中心：

这种布局形式是将与日间手术治疗相关的功能整合在一起，独立于医院的门急诊、医技和住院病房三大核心功能区，形成相对独立的日间手术医疗功能区，包括日间手术室、日间病房及与之配套的综合服务功能区，

这种模式能有效地提升医院的管理效率，但也会适当增加运行成本、同时对医院的管理要求也相对较高。这种模式中的日间手术能得到全面的支撑服务，病人能在医院门诊做术前检查评估，并往往能尽快实施手术，极为方便病人。

图 1 独立集中式的日间手术中心

分散布局的日间手术中心：

这种布局模式，是将日间手术治疗所需的全部或部分功能空间按其原有的功能分散地布置在医院的门诊、医技和住院病房三大传统功能区中。其特点是资源相对分散，管理不够方便，进而导致日间手术医疗模式的优势得不到充分发挥。

需要说明的是，若将日间病房布置在主要病房区域，会因为日间病房的人员流动性更大，出入更加频繁，给病区的环境和安全带来不利影响，也增加了病房的交通压力；若将日间病房布置在门诊区域，这样会因日间病房昼夜开放的特点而使得门诊区域无法在夜间进行关闭。此外，若将日间手术室与医院住院手术部合并设置，不仅会降低日间手术的效率，而且也会使得对日间手术病人的服务质量有所降低，因而被公认为是最不理想的日间手术中心布局形式，国际日间手术协会也不主张采用这种方式，除非在医院手术部留出足够数量的专用日间手术室。

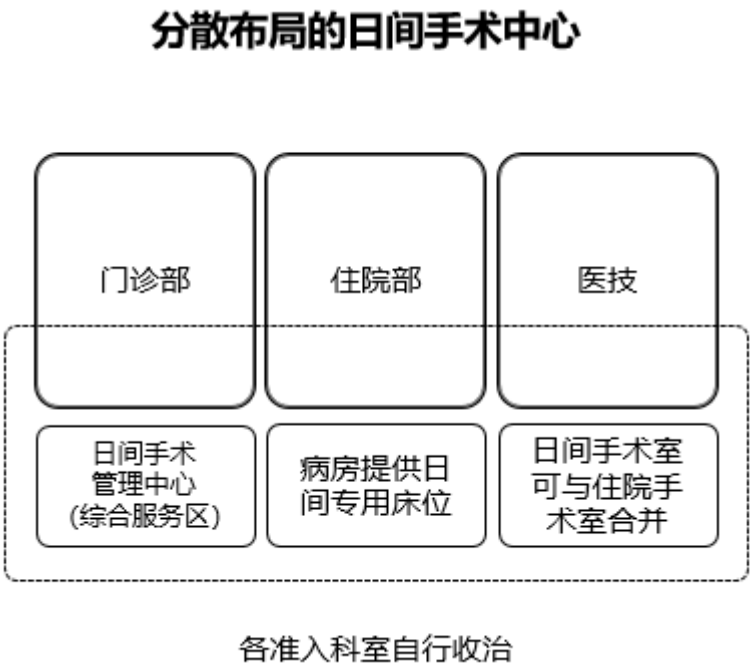


图 2 分散布局的日间手术中心

2 现状

2.1 我院日间管理模式

目前我院日间手术采用的是独立集中式和分散布局式结合的混合式收治方式，结合眼科疾病特点，院内设置了独立的日间医疗中心，负责所有日间手术的预约、随访和信息管理，部分科室日间患者被分配至日间医疗中心的病房，其余科室日间患者则被分配至相应的专科护理单元，此种模式结合了集中运营和分散运营的特点，是最为灵活的一种运营模式，可根据医院不同时期的床位实时运转利用情况进行符合当时医院状况的较为适宜的安排，但同时也对医院管理提出了更高的要求一一医院需要整合全院资源，包括床位信息、医护人员、设备占用情况等，并进行及时的更新与处理。

2.2 我院日间手术发展历程

2012 年 10 月，我院开始开展白内障日间手术，2014 年 3 月，我院获批开展 10 个病种的日间手术，2018 年 5 月，我院获批开展眼科 45 个病种日间手术，2019 年，我院获批开展 132 种眼科术式日间手术。

2015 年我院正式成为中国日间手术联盟第 32 家成员单位，2017 年成立了日间医疗中心，是国家首批 129 家日间手术试点医院之一，2021 年成立了独立的玻璃体腔注药中心。

目前我院日间手术量约为 2000 例/月，日间准入术者 73 位，2021 年全院日间手术占择期手术比例为 48.53%，患者平均住院等待时间由 1 个月余降低至目前的 1 周左右，平均住院周期由 2017 年的 5.89 天降低至 2021 年的 3.65 天。

2.3 我院日间手术业务流程

在日常收治流程中，日间医疗中心作为日间手术流程的枢纽，承担大部分工作，并协调门诊、病房、手术室、医技科室等各方，为患者带来一站式服务体验。

图 3 日间医疗中心为枢纽，协调各方

典型的日间手术流程：患者经门诊医生诊断，符合日间准入要求，并由内科医生诊断满足手术条件的，门诊医生为病人预约日间手术日期，开具术前检验检查项目，包括血常规、尿常规、血糖、A 超、B 超、OCT、角膜内皮镜检查、角膜地形图检查、胸部 DR 等，患者在指定日期内完成上述检验检查项目后，持所有检验检查结果，于手术前一日至专科病房或日间医疗中心办理入院手续，由病区护士完成入院宣教、入院评估等，并向手术室发送手术通知，手术当日患者提前到专科病房或日间医疗中心报道，护士完成术前准备，护理员引导患者至手术室完成手术，术后留观完成即可办理出院手续，出院后由护士登记完成三次术后随访。

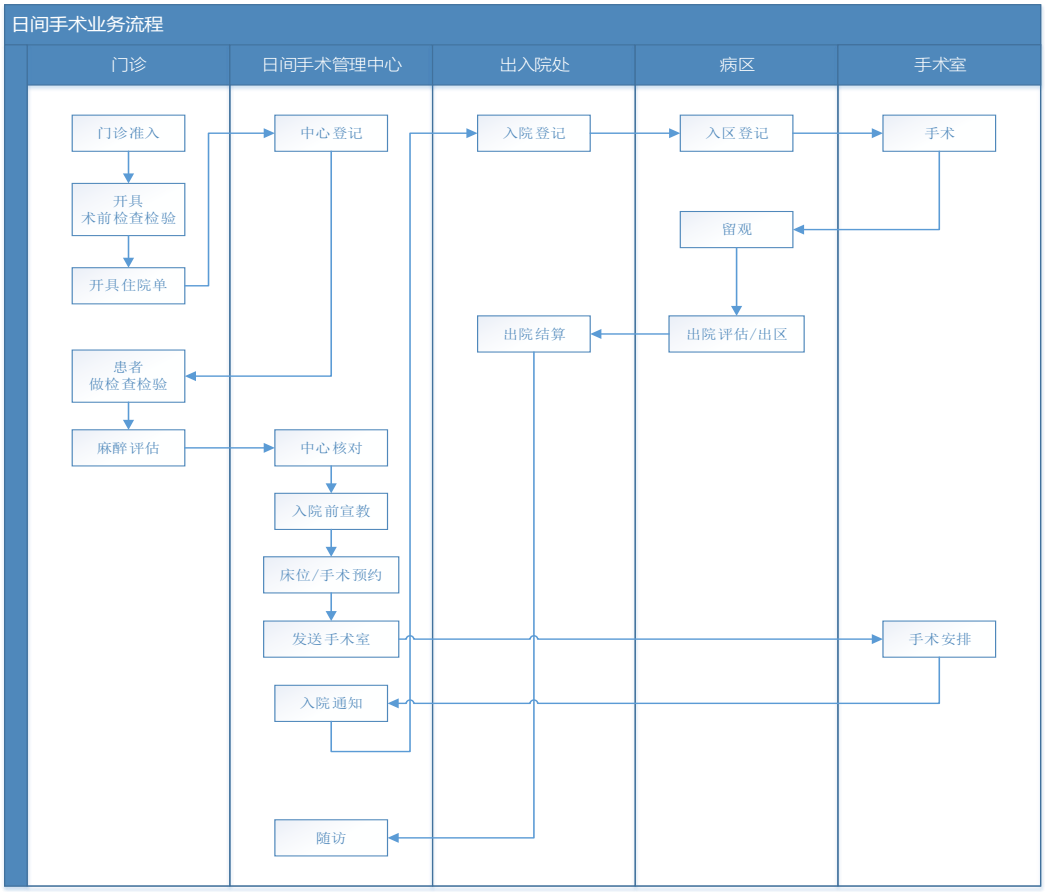


图 4 日间手术典型流程

3 日间手术信息化

3.1 日间手术管理对信息系统的需求

根据日间手术患者与院内管理特点，其对信息化的需求主要为以下方面：

日间手术的业务特点	对信息化的需求
满足独立、分散或混合模式管理	根据医院内部管理特点，应满足多种管理模式的要求。
患者术前术后不在院	医、护、患之间需要沟通畅通：包括围术期的宣教、患者的咨询、手术的通知、术后的随访。
手术计划贯穿门诊、病房、手术室	手术预约与计划的及时互通：包括流程化管理，以及手术申请信息、预约信息、报告结果、患者进程信息的及时互通等。
快速、准确和安全	质控与监管：对医生、手术的准入与授权；患者的全程示踪；对于日间手术患者的爽约率、手术的取消率、患者术前平均等待时间等数据的监控。
医保费用	一般认为日间手术有以下几种支付方式： 一是将日间手术视为门诊手术，在门诊按项目收费。二是将日间手术视为住院手术，仅将手术当天费用列入住院费用进行报销，前期在门诊进行的检查按门诊费用支付。三是将手术日之前在门诊进行的检查一并列入住院费用，按住院费用比例进行报销[1]。四是采用单病种、DRG 等支付方式。

表 1 日间手术信息化需求

3.2 我院现行日间手术信息化手段

我院目前在日间手术业务中，以日间手术管理系统为核心，对接门诊医生站、住院医生站、住院护士站、手麻系统等，基本覆盖了业务全流程。

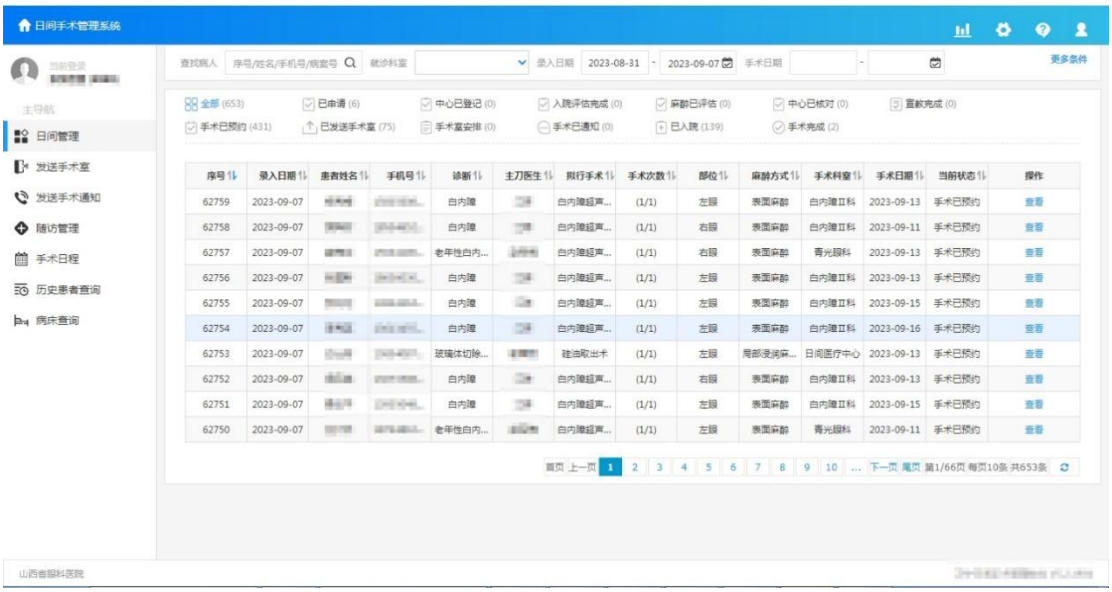


图 5 日间手术管理系统

日间手术管理系统：

患者信息检索，日间患者手术申请，入院评估，手术通知单发送手术室，预约、手术信息短信推

送，出院评估，患者随访，历史患者查询，工作量统计，节假日设置，患者、主刀大夫、术式准入等。

门诊医生站：

门诊接诊患者，门诊病历书写，开立日间检验检查项目等。

住院医生站：

日间患者住院病历书写，医嘱、处方开立等。

住院护士站：

患者床位管理，医嘱执行，医嘱计费，检验检查费用划转等。

手麻系统：

手术排程，围手术期信息收集、统计、查询，术后恢复等。

针对我院混合式日间管理模式，在现行的日间医疗系统中，可以对每个开展日间手术的临床科室单独设置采用集中式或分散式模式，日间手术量较多的科室如白内障科、青光眼科、视网膜病科等，设置为分散式模式，患者由专科病房自行收治，使用病房床位；日间手术量较少的科室如眼底病科、眼肿瘤眼眶科、角膜病科等设置为集中式模式，患者由日间医疗中心集中收治管理，使用日间医疗中心床位。

医保患者费用报销方面，我院于 2012 年上线了诊疗卡一卡通系统，门诊患者持诊疗卡可以完成充值、检验检查处方缴费操作。在此基础上，日间患者持诊疗卡完成日间检验检查项目，费用全部从诊疗卡上扣除，在办理入院后，护士在住院护士站系统中可完成诊疗卡门诊费用退费、划转住院费操作，将检验检查费用按住院费用比例报销。

3.3 我院日间手术信息化优化目标

（1） 随访管理

在现行的日间手术管理系统中，患者随访模块包含了随访短信发送、随访信息记录两个功能，在护士日常随访中存在信息反馈不及时、电话拒接、短信漏看、随访记录浮于形式等问题，对工作效率提高帮助有限。

针对上述问题，我们考虑将随访功能接入受众更加广泛微信、支付宝平台，通过微信公众号或支付宝生活号向患者定期发送随访信息推送，患者在平台上可以直接填写问卷，及时反馈自身情况。此项举措将有效提升随访效率，减轻护士工作量，大幅提高患者满意度。

（2） 电子诊疗卡

我院目前门诊及日间手术流程中大量使用诊疗卡，患者需在就医前办理，在门诊检验检查及处方扣费、日间检验检查项目扣费、日间费用划转住院费等场景中均需使用诊疗卡，有诊疗卡遗失、扣费异常、降低患者就医体验等风险。

在我院诊疗卡系统已运行十余年，各项业务流程已成熟的基础上，我们考虑将诊疗卡转至线上，对接省挂号预约平台，为每位患者开通个人诊疗卡账户，并适配全院诊疗信息系统。此项举措将有效简化就医流程，减轻医护人员工作量，降低信息系统维护成本、人力成本、购买诊疗卡及配套设备的物料成本，一举多得。

（3） 智能导引系统

基于我院日间手术业务现状，相比正常住院流程，日间手术流程更快，相应的对患者要求较高，需要患者根据单据上的导引位置、提示信息自行前往各检验、检查科室及收费处、入院处、病房等，患者易产生负面情绪，无法及时到达指定地点。

在我院智慧医院建设规划中，设计了院内智能导引系统，对接省预约挂号平台、门诊医生站、住院医生站等系统，从患者挂号开始到门诊、检验检查、住院、手术、出院，把院内涉及到的所有业务流程位置导引集成到手机 APP 或微信公众号、支付宝生活号上，患者可以根据当前就医流程，实时查看下一步位置及相关注意事项等。此项举措可以显著提升患者就医体验，提高满意度，同时彰显我院智慧医院整体形象。

4 日间手术信息化探索方向

4.1 基于可穿戴设备和人工智能的日间手术随访管理模式

随访是日间手术管理的重要环节，目前主流的短信、电话、微信支付宝推送等方式中，均存在受患者主观意识影响大，配合程度低等现象，由此易进一步造成医院随访工作完成度低，医护人员工作难度大等现象。

通过手机 APP、可穿戴设备及物联网技术，能够实现对院后患者相关指标数据的主动采集和监测，是推动术后患者管理的良好抓手，在提高日间手术的效率和安全性方面能发挥良好的作用。

此外，也有许多的研究表示将探索基于物联网技术的日间手术管理方式，有医院将物联网技术应用到了日间手术业务流程的优化管理中，实现了重要医疗设备追踪、药品追踪、医护人员及患者管理[2]。

同时，可以推进人工智能 AI 随访助手上线。人工智能 AI 随访助手是个机器人，可以根据规定问题模板模拟“医生”打电话给病人，按照不同术种患者的出院记录制订个性化的随访计划。AI 的上线不仅可以大大提高随访效率，还能确保随访信息采集的全覆盖及准确性，更有效地减轻医护人员工作压力。此外，AI 随访助手还会对采集到的患者随访海量信息进行统计分析，为临床和科研工作提供有效的数据支持，还可进行患者的满意度测评，用于日间手术医疗服务质量的持续科学改进。

4.2 结合床旁结算的日间手术管理模式

日间手术业务流程主打快速、便捷，尽可能地优化每个就医环节，但对于入院登记及出院结算的优化有限，相比传统住院流程，简化不明显。

通过上线床旁结算系统，在专科病房及日间医疗中心部署结算自助机，大部分患者可以省去住院处、医保科窗口排队办理的步骤，自行办理入出院登记、医保登记及结算业务，既能进一步简化日间手术流程，提高床位流转率，又可以减轻患者负担，提升满意度。

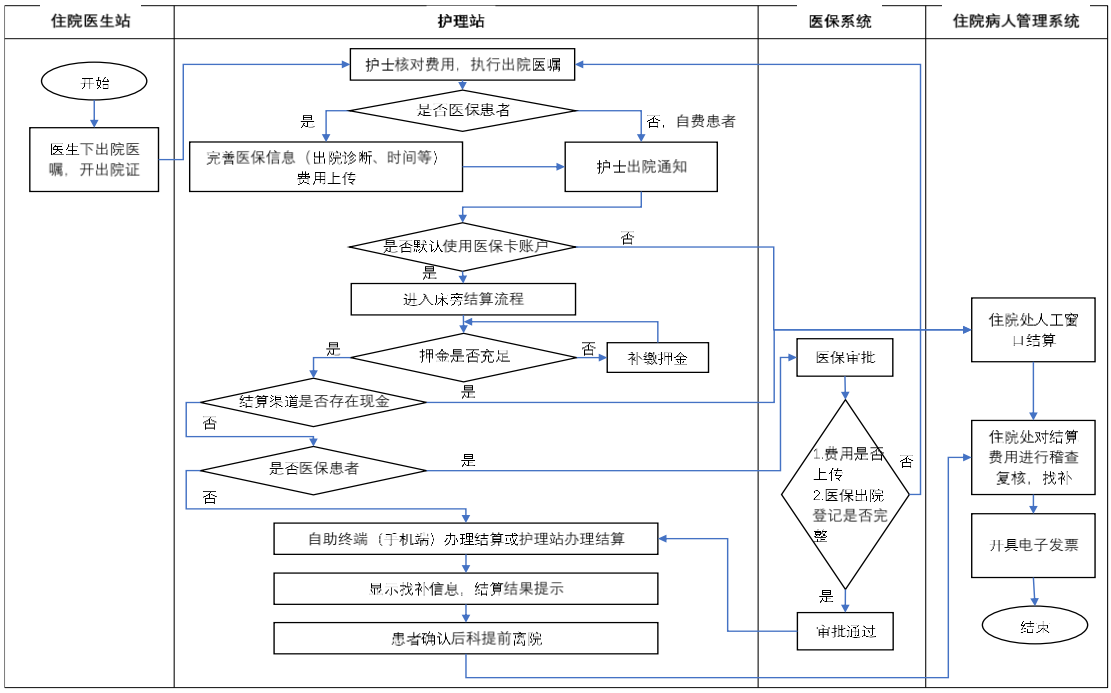


图 6 床旁结算业务流程

5 结语

虽然日间手术在我国起步较晚，但近年发展势头非常迅猛，种种数据资料也证明日间手术将成为一种重要甚至主要的医疗模式。国家大力推广日间手术，制定出台了一系列相关政策和制度，旨在进一步优化患者就医流程，提高医疗机构运转效率，推动发展全民健康医疗，看好病，花少钱。在此过程中，合理运用信息化技术是推行日间手术的重要手段，也是日间手术乃至医疗行业大力发展的必经之路。

由于日间手术具有医疗护理过程紧凑、医务人员工作密度大、牵涉科室多等特性，因此如何如何准确、快速处理日间手术过程中的信息成为开展日间手术的重要前提条件之一，而信息化的进一步推进则是信息管理的基础。对于专业的日间手术中心而言，信息化的建立是中心的硬件条件之一，而对于开展日间手术的综合性医院而言，信息平台的建立更是不能或缺，它不仅能将原本独立的科室紧密地联系起来，从而有效管理和开展日间手术，更能避免各种医疗事故的发生。

医院日间手术建设引入信息化手段诠释了精益医疗的内涵，信息化建设带给患者安全快捷的医疗服务体验。日间手术开展中应用信息化管理，更符合医院实施信息化发展方向。

PU-067

HIS 存储设备升级及性能优化

韦吉程[1] 陈汉云[2] 林方茂[3] 王欣[4] 文钊[5]
三亚市人民医院

1 引言

我院 HIS 系统于 2010 年 10 月份上线，至今已 12 年，上线之初从单一的院内业务、模块之间的集成不断地扩展，到现在系统之间的集成、数据集成、线上业务等功能的不断扩展，HIS 系统已经延伸到了医院应用的各个方面。HIS 系统的网络架构采用 FCSAN，双机集群，采用 ORACLE RAC，双链路单存储。目前的业务应用除了 HIS，还有对接了 20 多套接口，接口方式各不相同。2021 年 8 月份开始逐渐出现系统运行卡顿，尤其体现在门诊挂号收费、门诊医生站、微信公众号等业务，尤其是到了年底，核酸预约增多，高峰期卡顿现象更严重，有时候一个收费业务需要等五分钟左右，造成工作人员、患者等投诉十分严重。 根据院领导的指示，需要尽快解决问题，信息科对问题现象进行分析，初步判断硬件设备性能不能满足需求，紧急更换设备，保证了业务是顺畅运行，也为疫情防控起到了关键作用。

①韦吉程，海南 三亚 (572000)，研究方向：医学信息技术，312593961@qq.com
信息科，海南 三亚 (572000)，研究方向：医学信息技术，312593961@qq.com
②陈汉云，女，工程师，三亚市人民医院|四川大学华西三亚医院，信息科，572000，83986755@qq.com
③王欣，女，工程师，三亚市人民医院|四川大学华西三亚医院，信息科，572000，116562188@qq.com
④林方茂，男，工程师，三亚市人民医院|四川大学华西三亚医院，信息科，572000，
[illegible]@qq.com
⑤文钊，男，工程师，三亚市人民医院|四川大学华西三亚医院，信息科，572000，

2 HIS 系统及业务延伸应用的意义

目前医院的 HIS 系统已不同于 2010 年刚上线的时候，最初的系统只是内部业务模块之间的交互，居于一体化的设计，随着医院信息化的发展，业务已经延伸到线上互联网医院、医院与平台之间、医院与医保等外部的交互，有实时和非实时的，HIS 系统已俨然平台，不但自身业务，还要提供集成服务，服务器及数据库同时提供读写、高并发的服务。所以，保证 HIS 系统业务的高可用、高并发已十分迫切，同时兼顾数据安全。

3 方法和措施

3.1 背景 目前我院 HIS 系统网络架构搭建于 2010 年，采用 FCSAN 架构，双机集群，ORACLE RAC 数据库高可用，双链路单存储；服务器采用两台 IBM X3850M5，光交换机采用两台 IBM B24，存储采用 IBM DS5020；数据块采用 RMAN 定时备份，于每天凌晨自动备份，另外还有 ORACLE DATAGRUAD 备份机制，但由于服务器故障已失效；HIS 目前集合了 30 个业务模块，包括收费、门诊、住院、医保、药房药库、医技等；HIS 开发语言为 PowerBuilder+ORACLE，C/S 二层架构，操作系统 LINUX RADHAT6.5；客户端直连服务器数据库；外部系统有对接 20 多套业务接口，包含公众号、健康三亚、扫码支付、PACS、医技系统、体检、医保、区域平台、三医联动等；采用点对点方式，接口方式有 DBLINK、WEBSERVICE、HTTP+JSON、动态库等方式。

3.2 问题分析 2021 年中伊始，HIS 开始不间断的出现卡顿现象，主要体现在门诊收费工作站、门诊医生工作站，住院医师护士工作站也会偶尔出现，收费工作站出现缴费、退费点击确认时卡顿，一两分钟，长则五分钟都有，门诊医生工作站也是出现登录、开处方、保存数据时非常卡顿，还有通过医院公众号、卫健委健康三亚平台预约、缴费时也卡顿，严重影响了门急诊业务，8 月份到年底更严重，在网络预约，尤其核酸预约突增时，卡顿严重，患者投诉十分频繁。

出现问题后，信息科组织相关技术人员排查。首先排查网络硬件问题，因为硬件架构从 2010 年就部署上线，我们首先考虑问题可能发生在服务器、存储、光交换机设备是否老化导致性能下降的问题，从存储 IOPS、吞吐量，到服务器 CPU、内存、RAID 缓存、光交换机实施速率等进行分析。

首先分析存储，我们编写 SQL 脚本在服务器上运行，对存储设备进行读写，测试 IOPS（8KB 数据块）、最大吞吐量、延迟三个参数，发现最大 IOPS：2300 次；最大吞吐量：274MB/S；延迟：19S。这几个测试参数在正常负载下相比新购时已下降。

其次检查服务器及光纤交换机，服务器 CPU 占用率 10%以下，内存也不高，数据如下：

Cpu(s): 11.1%us, 1.3%sy, 0.0%ni, 87.3%id, 0.1%wa, 0.0%hi, 0.2%si, 0.0%st

相关的网络端口及状态等均正常，磁盘、仲裁盘等正常；光纤交换机光模块 8Gbps，状态正

常。

硬件网络问题初步判断主要在存储性能问题上，在大并发和吞吐量高时，卡顿严重，且目前数据库 RMAN 备份也在本机上，存储风险较高。

软件问题分析，我们在卡顿时对 HIS 数据库进行分析，发现锁表现象，主要在门诊信息表（MS_mzxx）、处方表（CF01）、YJ01（医技表），当使用 kill session 解锁时，卡顿现象消失，查看锁进程，主要集中在健康三亚公众号、医院公众号、扫码支付程序，锁终端主要也是对应着健康三亚接口服务器、扫码支付接口服务器。我们进一步根据锁进程发生的时间、终端查找 HIS 日志，对应健康三亚程序、患者、时间等，查找到相应的操作问题主要发生在健康三亚使用患者在操作退费时，发生锁表，相关表有 MS_MZXX、MS_SFMX、MS_FKXX，进而导致线下缴费、医生工作站卡顿，线下缴费业务不能 UPDATE 业务表。进一步分析，在健康三亚点击作废、退费业务时，会检索前面相关几个表，发现健康三亚可以检索到医院公众号及扫码支付业务无效的条目，且程序检索完后继续 UPDATE 相关表，设置相关字段，发现表中至少存在 40 多万条记录，卡顿的时间就发生在 UPDATE 锁表的时候。

至此，经过分析，软件问题主要存在于没有做相关的业务隔离及 IF 判断条件的退出。

3.3 解决方法及实施

3.1 解决思路 针对以上问题及考虑到整体信息化情况，我们总结如下，首先更换主存储设备，替换为全闪存盘，考虑到目前数据量共 2TB 及每月增长率医技我们总体信息化推进情况，我们采购容量 6TB 满足 2-3 年即可；软件方面，我们考虑到程序需要更新代码及数据备份的情况，另外购买一台数据服务器作为 DG 备份，替换出来的旧存储作为 RMAN 备份，可以达到数据两份物理备份，提高可靠性。

3.2 实施步骤 采购完成后进入实施阶段，我们在 2022 年 4 月底实施更换。首先把新存储接入光交，双链路，划分 RAID5，划分后的可用容量为 6TB，在光交划分 ZONE 后，服务器可识别存储，接下来由 HIS DBA 对新存储分区，划分数据区及日志区块，将区块挂载到相应文件系统及存储加入 ORACLE ASM 管理，接着对数据实施 AMS 同步，同步完成后，在凌晨 1：00 点停业务，将路径指向新存储并把就旧存储重新挂载。在替换完成后发现存储设备有告警提示，提示多路径问题，经分析，主要是 LINUX 版本的多路径 multipath 版本安装问题，重新安装存储厂家提供的多路径棒棒后重启 multipath 服务，告警消失。至此，设备替换完成。

新增加的服务器作为软件代码改进的临时测试环境及切换完成后的 DG 备份服务器。作为 DG 备份，实时根据主数据库的 redolog 来更新备库，同步效果在秒级内。同时利用旧存储作为 RMAN 备份，凌晨定时备份数据，备份时间在 5 个小时左右。

软件方面，对代码层面进行修改，做业务隔离，让健康三亚退费程序与其他线上业务隔离。

4 结果和作用

设备更换完成且运行一段时间后，业务基本上不再有卡顿现象，门诊医生站、收费处的卡顿已经没有，基本上解决了本文开头提到的问题。观察存储的 IOPS、吞吐量等均能满足。但是观察 ORACLE 的 redolog 发现 log file syn 切换频繁，6 组日志在短时间内切换，我们通过增加日志文件空间来改进。

5 总结及改进

5.1 总结 本次更换设备基本上解决了我们医院 HIS 系统目前存在的问题，且最重要是解决了最棘手的门诊业务的卡顿现象，为医院的业务正常流转起到了保驾护航的作用，同时对数据备份增加了一份双保险。

5.2 改进 本次升级及性能优化完成后，经过稳定运行一段时间，我们进一步分析后续还有一些可以改进的地方：

①作为医院的主业务系统，HIS 系统承载着同时读写和大量并发访问的业务，我们可以根据系统外部接口情况，将一些读操作转移到 DG 库，如 HIS 视图的访问等，减少主库的压力。

②本次更换由于资金及时间的压力，只是更换了存储设备，但由于服务器、光纤交换机的年限也老旧，性能下降，下一步将作为主要问题来解决。

PU-068

智能化自助入出院便民服务的创新研究与应用

王际强、黄盛明、骆振龙、李建标、廖国兴、梁胜焜、张耀龙
广西医科大学第二附属医院

一 、研究背景

2023 年，国家医政司发布《全面提升医疗质量行动计划（2023-2025 年）》，计划中提到“以病人为中心，以提高医疗服务质量为主题”，改善医疗服务行动计划，不断增强就医群众获得感、幸福感、安全感。

笔者所在的医院为一省级三甲医院，近年来，随着医院的发展，住院患者日渐增多，住院结算中心经常会出现患者扎堆，排长队办理入出院手续的情况，相信这也是国内各省市级医院经常会面临的情况。人员扎堆容易导致安全问题，也十分影响患者的就医体验感。通常医院为解决这种情况，基本只能通过新增办理窗口和工作人员来解决，但这种方式一方面会加重医院成本负担，另一方面有限的新增窗口在很多情况下仍旧无法满足数量庞大的病人需要。 在这种背景下，本着以患者为中心的指导思想，解决住院结算中心病人聚集扎堆办理出院手续的问题，笔者所在的信息技术团队多次进行现场调研，了解流程堵点和排队原因，收集各方需求意见后，最终提出三个创新功能来解决当前难题:

（1）一站式入院登记：简化入院登记流程，分流入院患者到临床护士站办理入院登记，让每个临床护士站都可以变成临时的“入院登记处”，帮助患者直接在护士站完成入院入科，无需挤在全院入院处排队等候。

（2）医保数据即时自动上传：对医保患者，护士办理出院操作后，系统后台即时上传医保结算明细到医保局结算平台，患者到达住院收费窗口可直接结算，不需要再等待医保结算明细上传。

（3）出院排队手机取号：出院患者通过手机微信端小程序，在线获取出院排队号，并且可以实时刷新获得最新排队信息，合理安排出院手续办理。

二 、研究与应用过程

2.1 研究方法

遵循软件工程设计思想，进行系统规划、系统分析、系统设计、系统实施、系统运行和维护，确立项目的研究方法如下：

（1）分析统计法：对我院入出院人员情况进行分析统计，分析历年入出院人次，逐步细化到季度、月、周、日、时，再对相关节假日数据抽取出来细化分析，形成入出院人员情况分析资料。

（2）文献研究法:通过学术文献检索，国内医疗信息化相关网站，查找入出院流程相关功能优化方案资料，结合医院的情况，选取有价值的部分进行参考。

（3）个案研究法：从向同行单位获取到的先进经验，结合医院自身情况，总结对比同行方案中的成功之处与不足之处，组织开展针对性的学习与探讨研究，探讨适合本院的优化方案。

（4）座谈会法：通过组织院内护理、财务、门诊、信息等部门相关人员讨论会议，集思广益、相互启发及彼此印证，通过头脑风暴的讨论方法，获取会议有价值的内容，最终形成入出院流程优化方案。

2.2 技术实现

（1）HIS 相关部分患者基本信息、住院登记操作、医保结算明细数据后台上传功能设计，使用 M 语言开发，采用 B/S 模式，通过 CSP 实现功能，数据库采用 Cache，插件采用 MedTrak。护士入出院办理逻辑及页面即通过本框架进行程序开发实现。

图 1.护士办理入院登记

（2）互联网医院小程序患者完善基本信息、住院押金补交操作，使用 JAVA 语言开发，采用 SpringBoot、SpringCloud 架构，数据库采用 MYSQL。患者端登记逻辑流程及页面实现通过小程序端开发实现，患者及住院信息通过 webservice 接口互联统一。

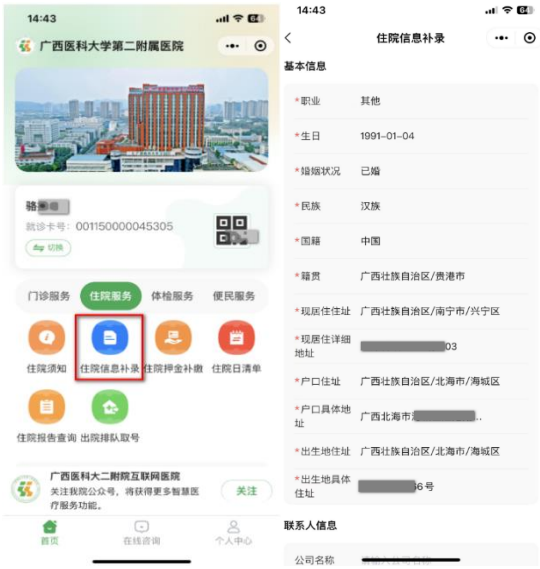


图 2.患者自助入院登记

（3）新住院大楼使用了全新的楼宇建设一体化，出院办理自助取号采用了新的排队呼叫系统，队列使用了全新的工作量负载算法，根据业务窗口数量和排队人数自动按窗口当前工作量进行智能分配。

（4）呼叫系统出院排队取号功能的算法设计依据预约出院人数、出院结算窗口数量、患者到达及结算状态、排队规则、系统排队最大值、历史排队数据这五个基础属性进行判断，提供更合理的预约取号服务给患者。出院排队取号功能可显示患者排队号、排队队列、当前呼叫号、当前等待人数（截止取号时的等待人数）、预计可办理结算时间。

- 预约出院人数：截止取号时，已预约办理出院的人次但未结算的人次。
- 出院结算窗口数量：指正在开放使用的出院结算窗口数量，影响结算速度的因子。

- 患者到达及结算时间：这里分两个指标，分别是患者预约取号后并到达住院结算中心开始结算的时间以及患者从开始结算到结算完成的时间，通过取号后等待时间及结算操作消耗的时间因子，预判取号后需等待时间。
- 排队规则：先取号先结算，同步现场取号机数据源。
- 系统排队最大值：设置时段排队人数最大上限，用于提醒可能等待的时间会稍长。
- 历史排队数据：历史数据对比，通过同一周期内同一时间点的预约排队情况，给出合理的提示。

（5）互联网医院小程序出院排队取号功能，通过获取呼叫系统排队序号，合理安排前往住院结算中心办理出院结算，患者也可通过互联网小程序实时查看预约取号后最新的排队情况。

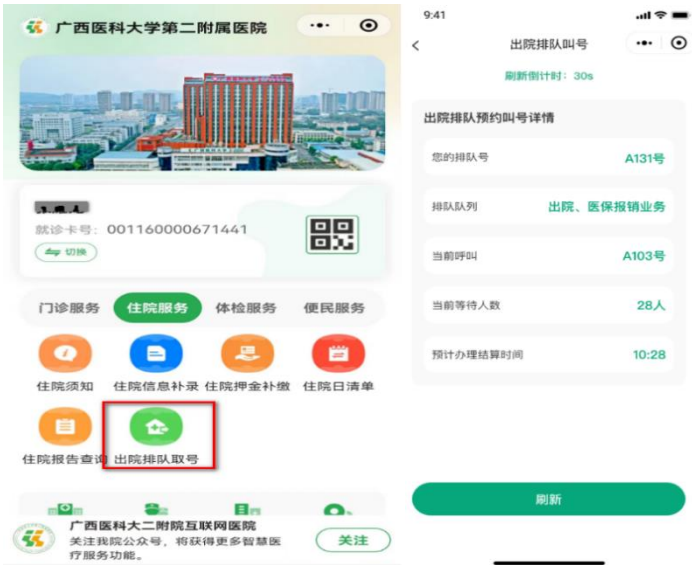


图 3.手机出院排队取号

2.3 推广过程痛点难点及解决

该研究项目技术上代码实现并不复杂，主要难点在推广使用、沟通协调、数据保障等问题上：

（1）护士工作量增加，协调护士配合是关键。患者通过互联网医院小程序自助补录个人住院信息，并支付住院押金即可完成入院入科。对于不熟悉手机操作或没有智能手机的患者，护士需要指导操作或在系统上为其办理入院登记。这增加了护士的工作量，影响护士对新流程的配合和认可。

（2）患者对一站式入院办理流程不熟悉，部分患者不习惯通过新流程办理入院。患者对新流程的认可及配合至关重要，需要引导患者从被动排队等待转变成主动自助办理业务。

（3）功能调整涉及不同功能模块的数据交互，要确保数据准确安全。在技术层面上，一是涉及院内 HIS 系统功能的调整，将住院结算中心的住院登记模块调整并开放给护士；二是通过手机互联网医院，实现患者补录住院基本信息，并同步数据到院内 HIS 中；三是医保患者结算明细数据自动上传到医保中心；四是患者通过手机互联网医院小程序办理出院结算取号。数据传输过程中，在技术上，既要保障医院数据的安全，又要保证数据传输过程中的完整。

（4）出院排队手机取号，智能化算法计算预计办理结算时间有难度。通过智能化算法，结合开放窗口数、出院人数、历史数据、排队规则等这些因素生成出院患者排队顺序号，对项目实施也是很大的考验。

面对上述难点，技术团队根据确定好的流程，积极主动与护理部、财务科、医务部沟通，及时调整优化系统。通过组织培训的方式，培训护理人员熟悉系统功能；通过短视频宣传，引导患者通过手机自助办理入院登记及出院手机自助预约取号；通过协调住院结算中心调整开窗的时间和数量，提升办理出院手续的效率，缩短窗口等待时间。通过多部门努力，项目得到了医护的大力支持，取得了不错的效果。

三 、成果及创新

3.1 应用成果

项目实施以来，实现了以病人为中心，提高医疗服务质量；实现信息多跑路，患者少跑腿的信息化服务；入、出院及结算效率显著提高等目标，体现在以下四个方面：

（1） 入院流程快速便捷，超过 60%的入院患者通过护士站一站式入院办理入院，初步实

现了让信息多跑步，患者少走路的目标。

（2） 医保患者出院结算，不需要等待结算中心上传医保结算数据，缩减医保结算患者办理出院结算的时间，提高患者办理业务满意度。

（3） 出院患者提前通过手机互联网小程序获取出院排队号，合理安排出院手续办理。

（4） 住院结算中心，不再存在患者扎堆排队办理入出院手续的现象，有效消除人员扎堆的安全隐患。

3.2 创新点

（1） 入院流程业务创新，患者到达临床护士站，线上自助完善基本信息并缴纳住院押金，全程只需要一部手机即可完成入院流程。

（2） 医保患者结算业务流程创新，护士操作患者出院后，后台即时上传医保结算明细数据，患者到达住院收费窗口直接结算，不需要等待医保数据上传。

（3） 出院排队流程创新，出院患者提前通过手机互联网小程序取出院排队号，并能实时获取排队情况，并可预判办理出院的时间，合理安排出院，不需要到现场排队等待。

（4） 出院流程技术创新，通过出院排队取号功能，结合现场实际情况与历史数据，智能化算法实时统计更新，给患者提供合理的引导服务。

3.3 推广普及意义

本项目在解决住院结算中心聚集办理业务方面有极大的推广意义：可以有效解决医院住院结算中心排队问题；系统的升级改造是对医院业务流程的创新再造，技术难度适中；项目投入产出比高，不仅就医群众得到实惠，同时也提升了医院在信息化新形式下的服务质量。

3.4 下一步研究方向

当前，笔者医院的患者通过护士站一站式办理入院的占比已经超过 60%，医保患者结算速度明显提升，住院结算中心现场扎堆排队情况已经不存在，现场取号或通过手机预约取号排队井然有序，智能化自助入出院便民服务平台上线实施初有成效。下一步，我们将继续研究优化并推广项目的实施，主要的研究方向有以下三点：

（1）继续努力提高护士站一站式办理入院的占比，方便更多病患；

（2）继续优化出院预约取号功能，结合节假日等因素，获取更合理准确的排队时间；

（3）按照国家医保推出的移动医保支付，继续优化医保患者出院手续，探索床旁实现医保结算的功能。

PU-069

医院信息化项目建设与管理

尹彭 刘寅*
岳阳市中心医院

一、引言

医院信息化是指利用互联网、大数据、人工智能、网络通信技术等现代化手段，对医院及其所属部门进行综合管理，以提高医院医疗质量、服务水平和运行效率的过程。医院信息化项目建设与管理是医院发展的重要支撑和引擎，对于提高医院核心竞争力、实现高质量可持续发展具有重要意义。本文将从医院信息化项目建设的现状、基本原则、主要内容和未来发展趋势等方面进行探讨[1]。

二、医院信息化项目建设与管理的现状

随着信息技术的不断发展，我国医院信息化项目建设已经取得了一定的成果。大部分医院都已经建立了比较完善的信息系统，实现了一定程度上的医院资源的整合和业务流程的优化。例如，很多医院实现了电子病历、医生工作站、护士工作站、实验室信息系统等临床信息系统的应用，提高了医疗服务的效率和质量。

然而，在实践中，我们也面临着一些问题。首先，信息化基础设施建设尚不完善，存在信息孤岛现象。这主要是由于不同时期不同厂商开发的系统缺乏兼容性所致。其次，医疗数据标准化程度较低，影响了信息共享和利用。尽管国内一些地区已经建立了医疗数据标准，但实际应用中仍然存在一些困难。最后，缺乏高水平的医疗信息化建设人才，制约了医院信息化项目的发展[2]。

三、医院信息化项目建设与管理的基本原则

在进行医院信息化项目建设与管理时，应遵循以下基本原则：

1、统筹规划，分步实施

医院信息化项目建设是一项庞大而又繁杂的系统工程，需要从整体上进行规划，分阶段逐步推进。在制定信息化战略时，应根据医院的实际情况，明确目标、任务和时间表，确保项目的可操作性。

2、数据标准化，信息共享

医院信息化项目的建设应注重数据的标准化和信息共享。采用统一的医学术语和数据标准，有利于规范医疗行为，提高医疗数据的准确性和可靠性，实现数据的互通互换。同时，通过信息共享，可以降低医疗资源的浪费，提高医疗服务的效率和质量。

3、以患者为中心，服务导向[5]

医院信息化项目的建设应以患者为中心，服务导向[3]。利用信息技术手段，提高患者就医的便捷性和满意度，实现医疗服务的个性化和智能化。同时，应关注医务人员的需求和使用体验，提供方便、高效的工作支持。

4、保障安全，注重隐私

医院信息化项目的建设应注重信息安全和隐私保护。采取有效的安全措施和技术手段，确保医疗数据的安全性、可靠性和完整性，防止信息泄露和非法访问。同时，应建立完善的网络与信息数据安全管理制度，明确责任分工和操作规范，确保系统的稳定运行。

四、医院信息化项目建设与管理的实施步骤

在进行医院信息化项目建设与管理时，一般可以分为以下几个步骤：

1、需求分析

在进行信息化建设前，需要对医院的需求进行深入分析。这包括对医院的业务流程、管理模式、服务理念等方面进行全面梳理，明确医院信息化的需求和目标[1]。

2、系统规划与设计

根据需求分析的结果，制定系统规划与设计方案。这包括系统架构设计、系统功能设计、数据结构设计等[6]。同时，需要考虑系统的可扩展性和可升级性，以满足未来医院业务和管理的需求。

3、系统开发与实施[4]

根据系统规划与设计方案进行系统开发与实施。在这个过程中需要注重系统开发的质量和进度管理；同时需要与医院的业务部门和管理者密切合作，确保系统的功能和性能符合需求。

4、系统测试与验收[2]

系统开发完成后需要进行系统测试和验收。通过系统测试可以发现系统存在的问题和缺陷；通过验收可以确保系统的质量和性能符合要求。

5、系统培训与推广

系统上线后需要进行系统培训和推广。通过培训可以提高医务人员对系统的熟悉度和操作技能；通过推广可以增强医务人员对系统的信任度和支持度。

6、系统维护与优化

系统上线后需要进行系统维护和优化。通过系统维护可以确保系统的稳定性和安全性；通过优化可以提高系统的性能和效率.

五、医院信息化项目建设与管理的主要内容

医院信息化项目建设与管理的主要内容包括以下几个方面：

1、硬件设施建设：医院信息化项目的基础设施建设包括服务器[4]、存储设备、交换机、路由器、终端设备等基础设施的采购与安装调试，以及与外部网络的连接。

2、系统软件建设：医院信息化项目的系统软件建设包括操作系统、数据库、中间件等软件的选型、安装调试与优化，以及与硬件设备的兼容性测试。

3、网络建设：医院信息化项目的网络建设包括内部网络的拓扑结构设计和优化，以及外部网络的接入和安全保障。

4、应用程序开发：医院信息化项目应用程序开发包括业务应用软件的开发和测试，以及用户界面设计、功能模块划分等[3]。

5、数据中心建设：医院信息化项目数据中心建设包括数据存储、备份与恢复、数据全生命周期安全保障等方面的工作，以及与外部数据中心的连接。

6、终端设备及应用程序支持：医院信息化项目终端设备及应用程序支持包括为终端设备提供相应的应用程序，以及对设备进行定期维护和升级。

7、安全管理：医院信息化项目安全管理包括网络安全管理[2]、数据安全管理、用户权限管理等，以确保系统安全可靠。
7、运维管理：医院信息化项目运维管理包括系统运行监控、故障处理、计划性维护等，以确保系统的稳定性和持续性。

六、医院信息化项目建设与管理的未来发展趋势

医院信息化项目建设与管理的未来发展趋势主要表现在以下几个方面：
1、一站式医疗信息化平台[3]：“一站式医疗信息化平台”是一种集成的医疗信息化解决方案，它将医院的所有医疗活动和行政后勤支持服务都整合到一个平台上。这个平台可以提供包括临床、科研、教学、办公、质控、物资、设备、财务等在内的全方位的信息化服务。
2、数据驱动的医疗决策：通过数据挖掘和分析，能够挖掘在大量医疗数据中的有用信息，从而为医院的决策提供科学依据。未来，医疗决策将更加依赖于数据，而不再仅仅是基于经验和直觉。
3、人工智能的应用[3]：人工智能技术已经在医疗领域得到广泛应用，未来这一趋势将继续发展。人工智能可以在医学影像分析、疾病诊断和治疗方案制定等环节发挥作用，提高医生的诊断和治疗效率。
4、医疗服务向云端转移：随着云计算技术的不断发展，未来的医疗服务将更多地转向云端。云服务平台可以提供高效、安全、灵活的计算和存储服务，使得医疗服务更加便捷和高效。
5、医疗信息安全和隐私保护：随着医疗信息化的深入发展，医疗信息安全和患者隐私保护问题将更加突出。未来，医院需要更加重视医疗信息的安全性和隐私保护，建立健全相关的制度和措施。
6、患者全流程管理[2]：未来的医疗服务将更加注重患者全流程的管理，包括患者预约、挂号、诊疗、支付、随访等方面。通过信息化手段，能够提高患者就医体验和服务质量。
7、跨区域、跨科室的医疗协作：未来的医疗服务将更加注重跨区域、跨科室的医疗协作。通过信息化平台，不同地区、不同科室的医生可以更加方便地进行交流和合作，实现资源共享和优势互补。

总之，未来医院信息化项目建设与管理将更加注重集成化、智能化和人性化，以进一步提高医疗服务质量、提升患者满意度和降低医疗成本。

七、医院信息化项目建设与管理的注意事项：

具体包括以下内容：
1、重视应急预案：对于信息系统，应该将应用软件安装包、配置文件、数据库、人员信息等都列入应急方案[4]，确保在系统崩溃时能够快速响应并保证医院信息系统的稳定运作。
2、注重机房容灾中心管理：数据中心机房是医院信息科维护工作的重中之重，必须加强对机房容灾中心的管理。
3、加强信息科培训：随着信息技术的不断更新，信息科的培训计划也必须跟上时代的步伐。未来信息科专业人才的培养将更注重信息系统研究，优化医院信息流程，充分发挥信息化赋能。
4、完善管理制度：好的管理离不开制度，完善制度是做好科室工作的首要任务。完善的制度，包括信息系统标准、网络设备管理、资产管理、人员管理等等。
5、重视资产管理：对于信息科日常维护工作，很多工作都是重复的，因此需要进行资产管理。

此外，在开展医院信息化的项目建设与管理时，还应注意以下几点：
1、整体规划、分步实施、效益驱动：制定完整的信息化规划是前提，必须具有完整的信息化规划[6]。
2、医院主导，需求为先：强调医院在信息化建设中的主体意识，信息系统必须适合医院的特点和满足医院的需求。
3、考虑预算：医院能够接受的预算是关键，不同的预算等级会对信息系统选型的范围和方向造成重要影响。
4、充分的市场调查：信息系统选型中大量的工作是对市场上相关产品及其供应商进行调查和评估。
5、技术先进还必须可用：需要在技术先进性和技术可用性之间找到平衡点。
6、强调信息系统的“三分技术、七分管理、十二分数据”：管理要到位，数据的作用要强调。
7、强调业务流程改进：管理的规范比先进的信息系统更重要。

PU-070

新时代党建引领下医院防统方系统的运用与管理

唐鑫洁

川北医学院附属医院

1 引言

医疗行业的有序运行涉及全体人民的身心健康，事关中华民族伟大复兴的重要任务，是建设社会主义现代化强国的关键领域。然而，医疗行业已不知不觉沦为腐败的重灾区，正面临一场“全链条、全领域、全覆盖”的系统整治。今年以来，全国公开通报被查的医院负责人已达到 200 多名，实在令人痛心疾首。

非法统方是医疗腐败的重要环节，泄露患者个人信息，扰乱正常市场秩序，抹黑医院形象，危害信息安全，降低人民满意度，属于严重的非法行为。不少医院也部署了防统方系统，却没有得到较好的运用与管理，单纯依靠信息化技术的传统手段收效甚微。康德有云，敬畏的事物唯有“头上的星空和内心的道德律”，而今天我们有了比之更上一层楼的精神准则，那就是中国特色社会主义精神内涵。比起依靠“技防”，“人防”的作用更加关键。因此，医院需要研究出一条在党建引领下，一体化统筹推进防统方工作的新路径。

2 现有防统方信息系统存在的局限性

2.1 硬件系统陈旧，落后于当前信息化速度

不少医院的防统方信息系统上线较早，上线时间在数年前甚至十多年前，硬件系统性能是在满足当时的情况下部署的，系统架构、覆盖范围、处理机制上已经远远跟不上当前不断增长的信息系统总量。受限于能处理的最大流量，数据中心等医院各业务系统的镜像数据库无法纳入监管，是最大的安全隐患。另一方面，即使检测到统方行为，也只能予以示警和追溯，无法有效阻断，难以避免统方事故的发生。

2.2 需大量人工干预，而纪检监察人员不具备计算机专业知识

以往防统方系统实施时，采用关键表、关键字段的传统 SQL 语句检测技术，手段单一，精度不足，会将各业务科室的正常统计需求也检测出来并示警。在示警信息层出不穷的情况下，需要人工进行鉴别筛查，设置安全过滤句柄。而纪检监察人员普遍无法阅读语句，在判断是否确切属于统方上无能为力，且此工作量及其庞大，很难由人员稀少的纪检监察部门独立完成。

2.3 统方途径多样，超出系统能追溯的范围

传统防统方系统普遍采用旁路监听的方式部署，对数据库审计功能和医院信息科的网络管理依赖严重。倘若数据库审计功能或者网络管理不完善，即使能检测到统方行为发生，也无法有效阻断和追溯。例如自助机、电子显示屏、内网 AP 等设备，在医院各个地方暴露出大量的网络接口，很容易被手持移动电脑的别有用心的专业技术人员利用。在事故发生后，嫌疑人早已逃之夭夭。另一方面，现在业务服务器广泛采用虚拟化技术，传统的物理交换机流量监听方式已无法适用，成为监控的盲点。

3 现有防统方制度安排存在的局限性

3.1 无法有效监管公司技术人员

随着医院信息化发展，大小业务模块引入了各自的信息系统，它们往往来自不同的厂家，公司驻场技术人员随之不断增多。在项目上线及运维过程中，他们有大量的机会接触医院核心数据，并以测试的名义旁建测试库。倘若其有心进行非法统方，在项目验收完成之后即行离开，会给医院带来难以估量的损失。而当前并没有出台专门针对公司技术人员的制度安排，亦没有建立公司人员行为管理的全流程监控。

3.2 人员权限管理制度混乱

权限管理与系统增多、更新迭代息息相关，许多上线较早的系统采用同一超级管理员账号，信息科人员和厂商人员也使用该账号，带来巨大的安全隐患。各种各样的子系统都有自己的可能涉统方的统计模块，并随着工作需要不断更新，逐渐超出固化的防统方检测语句范围。医院在信息安全制度执行上往往有所不足，保密协议签署不完善，相关人员的查询权限没有按照对应岗位职责精确划分，且发生工作内容和岗位调整的时候，往往权限并没有同步调整。

3.3 思想工作及安全培训开展不足

受限于医院繁忙的日常事务，在未雨绸缪展开防统方安全意识宣传培训这块力度不够，没有形成长效机制，没有将其纳入完备的绩效考核改革体制之中。安全意识的缺乏在临床医技科室体现得尤为明显，同事之间、师生之间知晓系统账号密码的情况比比皆是，因事离开电脑前时，很多人没有关闭程序、打开桌面密码的习惯，误用他人账号的情况时有发生。

4 党建引领防统方工作的必要性

4.1 党的领导是各项事业成功的根本保证

党的百年光辉实践充分证明，办好中国的事，关键在党。为人民谋幸福、为中华民族谋复兴，党的坚强领导是我们区别于世界上其他任何国家和政党的根本。因此我们才能求同存异，画出最大同心圆，集中力量将各项事业不断推向前进。医院防统方工作也是如此，靠单一部门力不从心，需要在党组织的统筹引领下，一盘棋规划，各部门通力合作。

4.2 党的领导是新时代医院高质量发展的必然要求

当前我们面临的最大实际，就是世界正处于经历百年未有之大变局的风云变幻之中。立足于新时代新形势，如何把握高质量发展的方向，是医院无法回避的问题。我们要毫不动摇地发挥党委的锚定航向领导作用，勇于改革创新，祛除制约高质量发展的弊病、痼疾，确保各项事业行稳致远。信息安全是发展的大前提，防统方工作需要在医院党委的领导之下，为服务高质量发展保驾护航。

4.3 党建引领是推进医院业务融合发展的必经之路

医院业务广泛、部门众多，涉及到各方面的利益千头万绪，交织在一起，往往会产生大大小小的冲突。“隔科如隔山”的情况长期存在，试图单单依靠运管部门来规划出符合医院整体利益的长远发展目标根本无能为力。而新形势下，医院想要在激烈的竞争中脱颖而出，必须打破制约高质量发展的藩篱，走党建引领的业务融合发展之路是必然要求。防统方并非独立于临床医技后勤业务之外，它本就应当是开展业务工作的准绳之一、核心组件之一。

5 党建引领防统方工作的实施路径

5.1 更新防统方全流程，打牢确保党委决策执行效果的基础

医院党委制定的政策制度想要不折不扣地投入到实践中去，就必须全面落实党委领导下的院长负责制。重大事项由院长办公会部署到具体责任部门和责任人，或召开专题会。这样党政分工合作，各部门齐抓共管，就建立起了从上至下、层层压实的责任体系。“磨刀不误砍柴工”，防统方系统的更新换代虽势在必行，然其是一个长期过程，需要持之以恒，查漏补缺。不仅要花费一笔不菲的资金，且工作维度广泛，单由信息科或纪检监察室都难以推动。有了这样一套分工合作机制体系，就打牢了确保党委决策执行到位的基础，然后就是循序渐进地解决防统方硬件系统的缺陷、网络安全的漏洞，完善数据库审计机制，完善权限管理制度，强化员工的信息安全意识。

5.2 完善党委班子联系机制，营造从善如流的防统方氛围

医院党委是医院各项工作的领导核心，也是医院高质量发展的指挥棒。唯有充分发挥党委的举旗定向作用，坚持办人民满意的社会主义医院，强化顶层设计，明晰各部门通力合作机制，才能保证全院上下心往一处想、劲往一处使。这就要求必须落实党委班子联系机制，每一位党委领导负责联系自己分管的部门，确保党委的决策准确无误地传达到位、执行到位。“其身正，不令而行”，防统方工作需要党委领导以身作则，同时常到分管部门宣讲，常抓常新，其作用是无可代替的。

5.3 创新党建工作举措，发挥基层党组织防统方战斗堡垒作用

严密高效、充满活力与战斗力的基层党组织是我们党的鲜明特色，也是党的优势所在、力量所在。医院需要在现有组织机构之下，结合自身实际，组织好各基层党支部，选强配齐专职或兼职党务干部，保障党委工作的高效运行。基层党支部是联系医院党委与每一个人的纽带，防统方工作借助基层党组织，创新党建工作举措，方能够常抓不懈，凝聚共识、团结一心，取得更好的效果。具体而言，在个人自学之外，可以召开多种多样的支部活动，如专题研讨会、读书报告会、观看主题教育影片等等。让防统方工作不再是冷冰冰的戒律，而是真正入脑入心，是大家携手共进的光荣事业。

5.4 鼓励主动作为，将防统方融入业务工作绩效考核机制

“路虽远，行则至”，充分发挥人的主观能动性是我们认识世界、改造世界无往而不利的法宝。党风廉政建设尤其是如此，唯有“不想腐”，才是廉洁自律的最高标准和最终目标，才能最大程度凝聚干事创业的同心力。要鼓励全体工作人员主动思考、主动作为，在自己的日常工作中不断全方位思考防统方的长远意义，将其作为完成工作的标准、改进工作的准则，乃至站在医院整体高质量发展的远景目标上思考问题。如此，若在工作中取得了实效，则应从绩效考核上予以侧重、鼓

励，并积极宣传先进模范作用，还有利于选拔出一批有想法、敢作为、能成事的干部，实现医院发展的正循环。

6 结语

党的二十大指出，构建新发展格局，推动高质量发展，党的领导至关重要。风清气正、廉洁高效是公立医院高质量发展的应有之意，是医学回归人文，打造以人民为中心的新时代医院的必要条件。做好防统方工作是医院纪检监察的关键环节，是从严治党的内在要求，需要的不仅是信息系统保障的“术”，更重要的是党建引领的“道”。唯有加强党建引领，才能为防统方工作注入源源不断的生机活力，保证正常医疗秩序，不断提高人民群众的医疗获得感。

PU-071

基于文献计量学的医院人力资源管理与信息化建设

林静，徐敏*，张仲男
甘肃省肿瘤医院，甘肃省医学科学研究院

人力资源部门是医院不可或缺的核心管理部门，负责医院的人事信息统计、人事档案管理、工资福利、职称晋升与评聘、人才招聘与培训、职工考勤、社保缴纳等多项业务[1]。近年来，人们生活水平和收入水平逐步提高，对于医疗服务的需求也越来越多，这就要求医院不断扩大规模，招募更多的专业技术人才，人力资源部门要处理的人才引进、人事信息统计、人事档案管理等业务也更加繁多。人力资源工作关系到员工的切身利益，一旦出现差错，后果极其严重。传统的人工管理模式已不能满足现实发展的需要，为解决人工管理的弊端，必须进行人力资源数字化建设，降低人工成本，减少人工差错，提高管理质量和工作效率，推动人力资源管理数字化、现代化、精细化[2]。

1 对象与方法

1.1 数据来源

分别以“甘肃、医院、人力资源、人事管理、数字化”为自由词，检索中国期刊全文数据库(CNKI)中、万方知识服务技术平台，检索时间为 2001-2020 年，有关医院人力资源管理及数字化建设相关文献，共检索出文献 432 篇。

1.2 文献处理方法

检索筛选文献并下载题录和保存，使用 ENDNOTE20.0 对文献进行查重，剔除重复文献 82 篇[3]。最终 350 篇文献纳入研究。使用 EXCEL 软件的高级筛选功能以及数据透视表对文献信息进行提取。

1.3 基于统计学方法的处理

通过 Excel 对纳入研究的文献建立收录文献数据库。数据库包含以下关键字段：题名、作者、关键词、单位、文献来源等 5 大类。

2 结果

2.1 发文年代

统计某一研究领域文献的发表年代分布情况，可以清晰的表明该领域的发展里程，帮助其他研究者准确掌握研究方向，预测发展趋势[4]。统计结果表明，2001-2020 年 20 年间共计发文 350 篇，2001-2017 年间发文量为 271 篇，年平均发文 15.9 篇，呈明显上升趋势，表明随着网络技术的发展创新，数字化建设对医院人事管理的重要性日渐凸显。部分年代发文较少，相关研究较少，具体情况如图 1 所示。

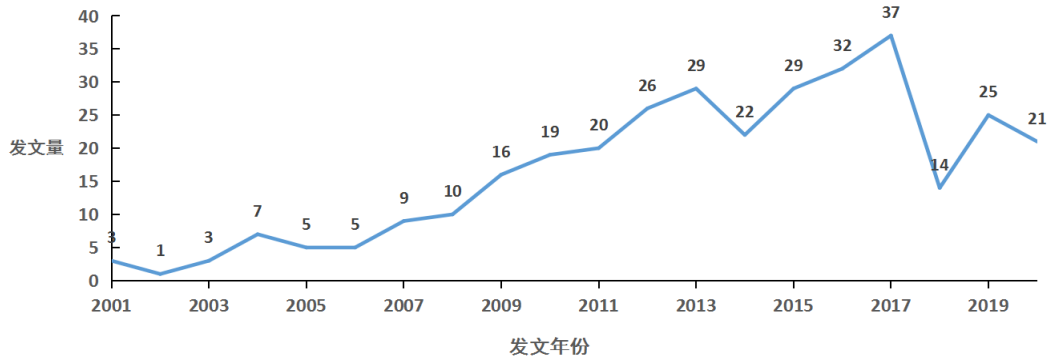


图1

图 1 发文年代与发文量

2.2 期刊来源

350 篇文献中，196 篇为期刊论文，刊载于 121 种期刊，其中 15 种期刊为国家级期刊（为 CSCD 期刊/统计源期刊/北大核心期刊），载文 28 篇，占比 14.29%，省级期刊 35 种，载文 102 篇，占比 52.40%，发文情况见表 1。总体来看国家级期刊发表文献较少，但论文质量及研究价值相对较高。

期刊名称	期刊级别	发文量	占论文总数比例（%）
中国中医药信息杂志	CSCD/北大核心	1	0.51
中国循证医学杂志	CSCD/北大核心	1	0.51
中国康复医学杂志	CSCD/北大核心	1	0.51
中国康复理论与实践	CSCD/北大核心	1	0.51
中华护理杂志	CSCD	1	0.51
中国医院管理	CSTPCD	5	2.60
西部中医药	CSTPCD	7	3.57
中国医药导报	CSTPCD	2	1.02
中国数字医学	CSTPCD	2	1.02
现代预防医学	CSTPCD	1	0.51
护理学杂志	CSTPCD	2	1.02
中国卫生经济	CSTPCD	1	0.51
中国实用护理杂志	CSTPCD	1	0.51
中国医院	CSTPCD	1	0.51
重庆医学	CSTPCD	1	0.51
疾病预防控制通报	省级	1	0.51
护理学报	省级	3	1.53
卫生软科学	省级	1	0.51
中国现代临床医学	省级	1	0.51
护理管理杂志	省级	1	0.51
中国初级卫生保健	省级	2	1.02
管理观察	省级	3	1.53
西北民族大学学报（自然科学版）	省级	2	1.02
职业与健康	省级	1	0.51
西北农林科技大学学报（社会科学版）	省级	1	0.51
中医药管理杂志	省级	1	0.51
甘肃医药	省级	9	4.60
农业科技与信息	省级	1	0.51

现代经济信息	省级	3	1.53
卫生职业教育	省级	30	15.3
齐鲁护理杂志	省级	2	1.02
现代医院	省级	1	0.51
电脑知识与技术	省级	1	0.51
护士进修杂志	省级	1	0.51
护理研究	省级	2	1.02
西北国防医学杂志	省级	5	2.60
人力资源	省级	13	6.63
全科护理	省级	1	0.51
临床医药实践	省级	1	0.51
现代医院管理	省级	1	0.51
护理实践与研究	省级	1	0.51
延边医学	省级	1	0.51
海南医学	省级	1	0.51
中国医学理论与实践	省级	1	0.51
中国卫生监督杂志	省级	1	0.51
现代职业教育	省级	1	0.51
中国农村卫生	省级	2	1.02
中国保健营养	省级	3	1.53

表 1 期刊来源与发文量

2.3 核心作者

使用普赖斯（Price）计算公式对纳入文献开展统计分析工作。核心作者发文数（Mo）与发文数量最多作者的发文数（Nomax）之间的关系：$Mo=0.749*\sqrt{No\max}$。上述检索作者中发文最多的是 12 篇，计算 Mo=3.46，即发表论文 3 篇及以上的作者可视为核心作者候选人[7]。根据普莱斯定律（Price law），计算核心作者候选人发表论文累积最低被引量Mc：$Mc=0.749*\sqrt{Nc\max}$。对上述检索中作者发文依次进行计算，得到被引频次累积最高的为 148 次，计算出累积最低被引量Mc＝9.11，即累积最低被引频次 9 次及以上者被视为核心作者候选人。对收录文献数据库中核心作者背景信息进行分析比对，对重名作者（核心作者）分别记录发文量，再剔除不符合条件的。最终得到结果，甘肃省人民医院韩琳发文最多，共计 12 篇，总被引次数为 148 次，占所有发文量的 3.41%，其次为西北民族大学马玉霞，第三为甘肃省人民医院王晨霞，且三者存在一定数量的合作，说明其以形成明显的合作团队。如表 2 所示。

作者姓名	发文篇数	被引频次	作者姓名	发文篇数	被引频次	作者姓名	发文篇数	被引频次
韩琳	12	148	张丽	3	5	任晓燕	3	1
马玉霞	11	63	张彩云	3	2	彭素梅	3	1
王晨霞	6	57	杨琴	3	4	裴娟	3	1
郑慧	5	41	杨灵歌	3	2	李林贵	3	1
郑小玲	4	11	严慧萍	3	2	李娟	3	1
张菊霞	4	7	唐兰兰	3	1	丁国武	3	0
苏萍	4	11	魏雪峰	3	1	白飞	3	0

表 2

3 问题

当今社会，人才是非常重要的资源，大到国家、民族，小到企业、公司凡是拥有较高水平的人才，均能得到快速的发展。我国在当前环境的影响下，投入大量资源高度重视人才发展、人才管理，建设有助于人才管理和控制的机制，以促进人力资源开发。目前已有多地多家医院立足于现代医院数字化转型的长期发展战略，通过人力资源信息化平台搭建，实现医院信息资源的合理配置与数字化集成，有效提高人力资源管理效率，为医院科学决策和人才管理发展提供数据支持。与此同时，实施人才战略，不断更新管理理念，持续完善人才选、用、育、留四位一体的管理体系，健全人才管理制度，保证人才在当下的医院管理体系下得以发展，人才能力能够发挥极致，以便为患者提供更优质的服务。但是在当今大数据时代下，仍存在很多问题，由其在西北经济落后地区，尤为严重。

3.1 甘肃省医院人力资源管理数字化仍处于非常落后水平

2001-2020 年，我省共发表相关领域文献 350 篇，年发表量 17.5 篇，虽然一直呈现上身趋势，但发文量仍然较低，且核心期刊的发布水平和数量还有待于提高，同时根据普赖斯定律，还不能形成核心团队。

3.2 对医院人力管理信息化建设认识不足

医疗机构管理者往往更加关注的是医院发展、医院文化、医院管理以及医疗服务质量等问题，对医院人力管理信息化建设的重视、支持和参与不够。这便往往就是制约医院人力资源信息化发展的重要的原因。而医院分管领导对许多问题难以平衡和解决，数字化管理系统购买不及时，这就导致医院人力管理信息化建设流于形式。同时，也反映出医院人事部门工作人员对于医院的数字化建设的认识不够，信息化管理能力欠缺，积极性、主动性低，管理系统利用率低。

3.3 缺乏相应的规范管理及业务流程

引进任何一种先进的技术，需要有制度规范和管理，但很多医院在人力资源信息化建设中并没有充分遵循这个原则，没有及时地制定医院人力信息化工作规范，明确工作内容、工作职责、工作程序、工作要求以及制定相应的考核办法等，导致整个医院人力信息系统没有充分利用起来，造成资源浪费。与此同时，对医院的管理流程进行分析，没有重视与医院的具体业务和人才发展战略相结合，照搬别人的系统来设计管理系统，造成了信息系统与医院人事工作的业务流程不匹配，没有发挥所期望的作用，从而使设计的系统使用不便或很快过时，降低了信息系统的效率。

3.4现有的基础设施匮乏

对于公立医院人力资源建设而言，基础设施有着两方面的内涵，分别是医疗结构的分布以及医疗人员的专业程度。首先，我国是世界上的人口大国，现有的医疗机构、医护人员数量相对比较少，而出现上述问题的原因在于行业发展不平衡和医护人员培训周期过长两个方面。常年来，我国医疗体系中的供需不平衡使得众多的医护人员每天都面临超负荷的工作。

与此同时，我国各公立医院之间服务水平参差不齐。在传统的医疗活动中，人们只注重最低的就医需求，能够有好的医生、先进的技术给自己提供一定的医疗服务就已经得到满足。但是随着我国经济社会的不断发展，人们生活水平不断提高，在物质生活获得极大的丰富和满足之后，人们开始注重精神品质的提高，在生病住院的过程中，除了希望医院能够提供高

水平的医疗技术之外，对医疗服务也提出了较高的要求，希望在看病的过程中能够实现更高质量的服务。其次，员工缺乏职业培训。在各行各业的发展过程中，人才都是企业获得长足发展和进步的不竭动力和源泉，医疗行业也不例外。通过组织医院的医护人员进行职业培训，才能够使得医护人员的医疗技能得到充分的发展。但是目

前我国公立医院对于医护人员的培训比较缺乏，一方面是组织的培训较少，另一方面是培训缺乏系统性和针对性，往往起不到很好的培训效果。

4.对策

4.1 医院人力资源管理系统的构建

目前数字化人事管理以广泛应用于医院人力资源管理中，其不仅帮助人事管理人员有效的做好医院人力档案管理，同时也有助于医院管理者掌握当下医院人力资源，促进医院未来发展[5][6]。近年来，关于医院管理数字化建设的政策及制度已相继出台，

我省需要立足于医院发展实际，重视推进医院数字化人事管理进程，构建安全完善的医院人事管理系统，将人事化管理理论与实际相结合，以期进一步提高医院数字化人事管理建设水平[7]。

4.2 医院人力资源管理人员的数字化培训

当前信息技术已被各行各业应用广泛，其中也包括医院人事管理工作。想要推进医院人事管理的数字化建设，就需要组建一支具备较强综合能力的人才队伍[8]。在人才培养计划中，不仅要让其了解医疗机构和相关领域的重要性，还要提升其计算机应用及数据维护能力，此外采取定期举办培训班、微信推送数字化人事管理小知识、现场观摩的等形式，加强医院人事管理演员的数字化素养，为医院人事档案数字化管理建设相关工作的顺利开展奠定坚实的基础[9]。

4.3 医院人力资源数字化管理水平的提升

人力资源数字化是指基于互联网的、高度自动化的人力资源管理工作,囊括了最核心的人力资源工作流程,如招聘、薪酬管理、培训等。医院人力信息资源则是从医院所需要的医疗卫生领域人才着手，包含了医、药、护、技，其相关执业信息、教学科研信息以及职称信息等数据，相对复杂。因此，在医院人力资源管理数字化建设中，需要多部门分工机制。人事科主要负责职工人事变动、工资绩效变动、年度考核以及职称晋升等材料的填写；科教科整合职工的各项科研成果、论文发表、进修学习等工作材料的汇总；医务科负责职业医师执业、护理职业资格、药师职业资格等注册及变更等材料；党办负责定期交送入党材料、政审和干部任免材料；工会和纪检监察部门负责职工奖惩材料。相关科室应充分发挥各个职能科室的作用，严格遵循相应的规章制度，对人力资源档案材料的形成过程进行全方位的审核和把关，以此来确保最终医院人力资源信息具备较高的可信度，提高利用率，从而有效推动医院人力资源数字化管理的高速发展[10]。

PU-072

微服务架构在医院信息系统建设中的应用

赵霞 姚佳璇 曾艺 李小华*

南部战区总医院

微服务是基于 SOA 架构发展起来的系统架构。微服务具备优良的敏捷性、扩展性、灵活性、可控性、容错性和维护性，在互联网、金融、政务等领域得到越来越多的应用。在医疗卫生领域，微服务也开始展现出良好的应用趋势，许多医院信息系统架构，包括主流的医院信息平台产品，都开始采用微服务技术。

本文通过介绍微服务的技术原理，展现微服务在医院信息平台建设中的应用。

一、微服务概念

1.微服务定义

微服务的定义有多种表述：

一般性定义：微服务是一种架构，是将单个的整体应用程序分割成更小的项目关联的独立的服务。一个服务通常实现一组独立的特性或功能，包含自己的业务逻辑和适配器。各个微服务之间的关联通过暴露 API 实现。这些独立的微服务不需要部署在同一个虚拟机、同一个系统和同一个应用服务器中。微服务是 SOA 的一种变体。

Thomas Eri 定义[1]：一种非不可知服务，通常包括特定处理和实现需求的逻辑小功能范围。

Gartner 定义[2]：微服务是一个面向服务的应用程序组件，具有严格的范围、强封装、松散耦合、可独立部署和可独立扩展。

Microsoft 定义[3]：微服务是一种用于生成应用程序的体系结构方法，应用程序中的每个核心功能（或服务）都单独进行生成和部署。微服务体系结构是松散耦合的分布式体系结构，因此一个组件的故障不会中断整个应用。各个独立组件协同工作，并通过定义明确的 API 协定进行通信。生成微服务应用程序，以满足快速变化的业务需求。

AWS 定义[4]：微服务是一种开发软件的架构和组织方法，其中软件由通过明确定义的 API 进行通信的小型独立服务组成。

百度定义[5]：面向服务的体系结构（SOA）架构样式的一种变体，提倡将单一应用程序划分成一组小的服务，服务之间互相协调、互相配合，为用户提供最终价值。每个服务运行在其独立的进程中，服务与服务间采用轻量级的通信机制互相沟通（通常是基于 HTTP 的 RESTful API）。每个服务都围绕着具体业务进行构建，并且能够独立地部署到生产环境、类生产环境等。另外，应尽量避免统一的、集中式的服务管理机制，对具体的一个服务而言，应根据上下文，选择合适的语言、工具对其进行构建。

2.微服务特征

在 SOA 架构中，为了面向服务，需要将业务流程在功能上分解为一组细粒度操作，这使我们能够识别可能成为服务和服务能力的潜在功能上下文和边界。在这个初始分解阶段，可以将业务流程操作组织为两个主要类别：不可知上下文与非不可知上下文的。

在每个不可知上下文中，解决方案逻辑将进一步分解组织成一组不可知的服务能力，即解决独立关注点的服务能力。因为它们是不可知的，所以能力是多用途的，可以重复使用以解决多个问题，即重用。

在不可知服务逻辑分离之后剩余的是针对业务流程的逻辑，这个逻辑在本质上具有单一目的的，归类为非不可知逻辑。非不可知逻辑的子集（或“微任务”），往往具有特殊性能或可靠性需求，当这类逻辑被抽象单独的服务层时，可以通过微服务加以实现。微服务实现通常是高度自治的，以便最小化对它们功能边界之外资源的依赖性。微服务通常被放在容器内。容器内部除了微服务，还含有微服务所需的公共服务和系统组件等。

SOA 服务可以通过对企业原有的非 SOA 服务，经过 SOA 封装后形成，对服务的颗粒度或者说业务功能范围并没有限制。但微服务应该具有小的颗粒度，或者说是最基本的、最小的业务功能逻辑。可以理解为微服务具有原生服务的属性，即微服务没有超类，是不能再抽象的服务[1，6]。

二、微服务原理

1.微服务的基本架构

微服务的基本架构如图 1 所示[7]。

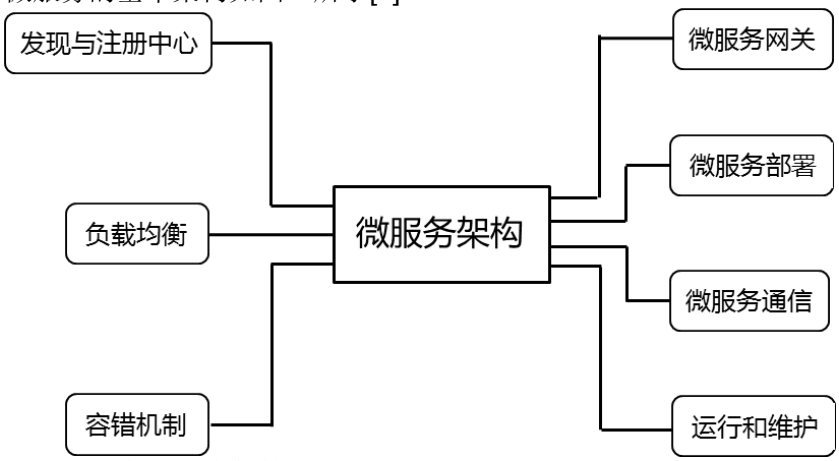


图 1 微服务的基本架构如

微服务的基本架构主要包括以下部分：

（1）微服务发现机制和注册中心

作为一种轻量化的应用，微服务在运行过程中时刻都在动态地变化，服务随时可能被复制、删除或重定位，因此需要创建一种微服务的发现机制，用来感知服务是否正常。一个微服务部署启动后会在服务注册中心进行注册，注册的信息包括微服务本身的网络、实例地址和路由信息等。

（2）微服务网关

微服务网关主要作为微服务架构中的所有微服务和客户端的出入口，接收所有的 Web 请求，实现包括路由转发、身份权限认证、网关安全防护、协议适配和流量控制等功能。

（3）负载均衡

在微服务架构中，需要部署多个服务实例即采用集群化、主备、双活等部署方式，来实现微服务的高可用性，同时要求支持主流的负载均衡算法。

（4）微服务部署和通信

微服务架构中，微服务独立进行部署，各微服务之间不存在部署上的强相关性。微服务通信是指微服务之间在网络传输过程中是如何高效准确地实现信息交互和消息传递，一般分为同步和异步通信两种。

（5）容错机制

在微服务体系中，通常都会有多种微服务之间的互相调用和前后依赖关系，某个微服务出现故障都会导致多个相关联的微服务不可用，从而造成业务流程中断。比较常见的有超时重试、限流和熔断等容错机制。

2.容器化技术和 Docker

容器化技术是指独立的应用进程以容器的形式部署运行在操作系统之上，容器化技术中应用最多的就是 Docker。应用和服务经过打包移植到容器中，在不同平台上进行发布和虚拟化的实现。Docker 是开源的， 基于 Go 语言进行开发，遵循标准化的 Apache2.0 协议，采用 C/S 架构，容器创建和管理等通过 API 实现。

镜像（image）、容器（Container）和镜像库（Docker Repository）是 Docker 技术中比较重要的几个概念。镜像就是模板，用于创建容器，容器是基于镜像创建的运行实例，镜像库则是用来管理镜像。Docker 技术的基础架构如图 2 所示，主要由客户端（Client）、守护进程（Daemon）、镜像（image）、容器（Container） 和镜像仓库（Docker Repository）等几部分组成。

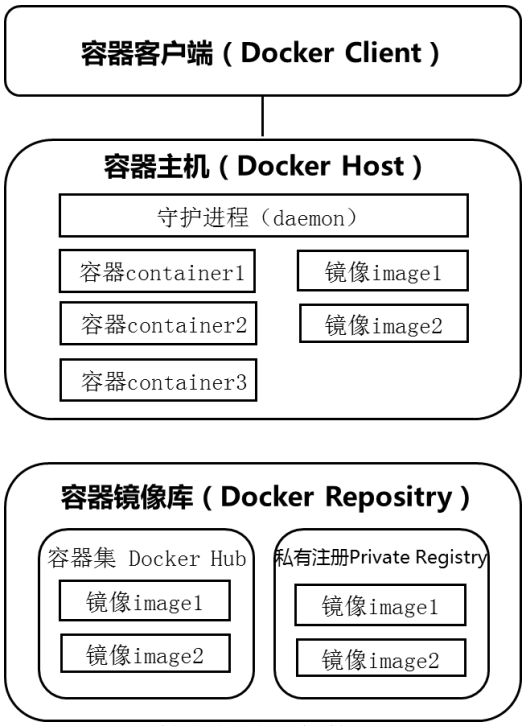


图 2Docker 技术的基础架构

Docker Client 客户端会向服务端发送包括容器创建、启动等请求，服务端的守护进程 Dockerdaemon 负责接收处理这些请求。客户端和服务端可以部署在一台机器上，也可以通过远程

API 互相通信来完成容器的相关功能。通常容器和镜像都会部署在 Docker host 上， 镜像库（DockerRepository） 既可以提供 Docker 官方的镜像库（Docker Hub），也可以将私有镜像进行注册（Private Registry）。

3.容器集群管理 **Kubernetes**

Docker 技术已经成为主流的容器虚拟化技术，为了方便对 Docker 容器进行规模化和集群化管理，谷歌公司推出了 Kubernetes 的容器集群管理系统。Kubernetes 包括 Master、Node、Cluster（集群）、Service（服务）和 Pod 几个重要概念[8]。

（1）Cluster（集群）表示一个由 Master 和 Node 组成的 Kubernetes 集群。

（2）Master 是指集群的控制节点，在每个 Kubernetes 集群中都需要有一个或一组被称为 Master 的节点，来负责整个集群的管理和控制。Master 通常占据一个独立的服务器，是整个集群的核心。

（3）Kubernetes 集群中除 Master 外的其他服务器被称为 Node。与 Master 一样，Node 可以是一台物理主机，也可以是一台虚拟机。Node 是 Kubernetes 集群中的工作负载节点，每个 Node 都会被 Master 分配到一些工作负载（Docker 容器）。

（4）Service 通常由多个程序副本提供服务。Kubernetes 里的 Service 具有一个全局唯一的虚拟 Cluster IP 地址，一旦 Service 被创建，在 Service 的整个生命周期中，它的 Cluster IP 地址都不会改变。客户端可以通过这个 Cluster IP 访问该服务。Service 可以理解为微服务架构中的一个微服务。Service 定义了服务的访问入口，实际上代理了由一组 Pod 构成的应用实例。

（5）Pod 是 Kubernetes 的最小单元，容器包括在 Pod 中，一个 Pod 由一个称为“根容器”的 Pause 容器和一个或多个紧密相关的用户业务容器组成。Pod 相当于逻辑主机，每个 Pod 都有自己的 IP 地址，称为 Pod IP。

图 3 是 Kubernetes 集群整体结构图。

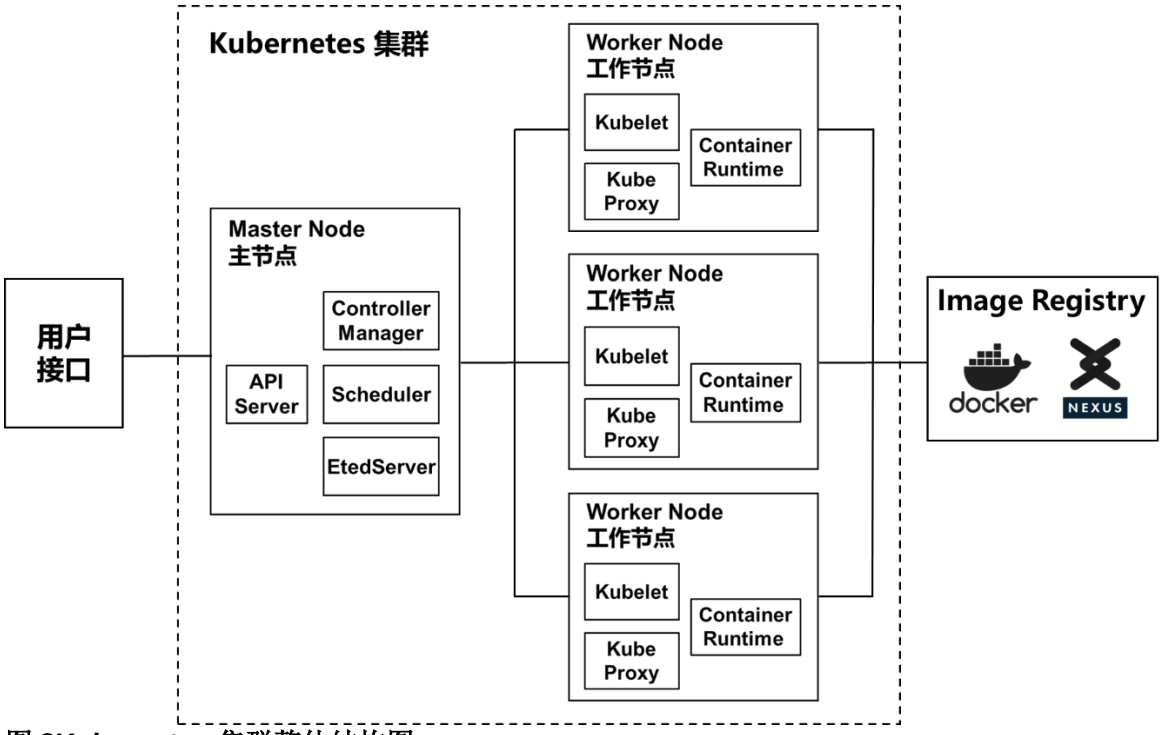


图 3Kubernetes 集群整体结构图

图 3 展示了共四个节点的 Kubernetes 集群的结构图。其中包含一个 Master 节点和三个 Worker 节点。Master 节点上运行着 API Server，Controller Manager，Scheduler 和 Etcd Server 四个关键进程。其中，API Server 进程提供 HTTPRest 服务，负责资源的增删改查，是用户访问集群的入口。Controller Manager 进程负责执行集群层面的操作，包括生命周期方法以及 API 业务逻辑。Scheduler 进程负责调度容器，会根据容器所请求资源的可用性、服务质量要求、亲和力约束等调度容器至最合适的节点上。Etcd Server 进程负责管理所有资源对象的数据，因为 Kubernetes 底层存储是基于 Etcd 键值数据库实现的。

Worker 节点上运行着 Kubelet．ContainerRuntime 和 KubeProxy 组件。Kubelet 是 Kubernetes 集群中最重要的控制器，负责 Pod 对应的容器的创建、停止等工作，并和 Master 节点进行协作，完成各种管理工作。ContainerRuntime 组件通过容器引擎接口调度底层的容器引擎。管理员可以自由的选择合适的容器管理引擎，例如 Docker 作为容器运行管理引擎。Kube Proxy 组件是实现 Kubernetes 中 Service 通信与负载均衡的重要组件，创建虚拟 IP，使得客户端可以直接访问该虚拟 IP．并透明代理服务中的 Pod。

三、微服务在医院信息平台的应用

基于微服务优良的敏捷性、扩展性、灵活性、可控性、容错性和维护性等特征，微服务架构很快受到医院信息化软件开发厂商的青睐，纵观市场上的医院信息平台产品，基本上都提到采用微服务技术。尽管微服务应用起源于互联网应用，但随着医院互联网业务的增加，特别是微服务在医院信息系统建设的成功应用，微服务得到越来越多的关注，并出现许多很好的研究和实施案例[9-10]。

基于微服务的特性和优势，微服务技术用于医院信息系统时，需要注意把握以下要点：

1.适用范围

微服务作为信息系统体系架构，属于平台层面应用，用于实现信息系统之间的信息交互，与目前常用的集成平台、交换引擎相似。不同的是，微服务细分了传统 SOA 服务的颗粒度，并将这些细分的微服务放进一个容器中，以简化服务之间的接口，即简化软件开发、又实现系统高效运行。因此，微服务的应用应该是在平台层面上，或者说是新平台建设抑或现有平台升级。

2.业务系统改造

微服务的特征在于“微”，即微服务是包括特定处理和实现需求的逻辑小功能范围（一种非不可知服务）。一个微服务是一个小的程序单元，这就要求在应用微服务架构时，需要将现有 SOA 服务或业务系统程序进行微服务化，将原有服务细分为系列的、可重用的微服务，这是应用微服务技术的关键点和难点。只是个别或部分的业务系统微服务化是不够的，起码是在一个学科或者领域的范围。

3. 架构过渡

除非新建医院和采用整套微服务架构业务系统，否则难以在短时间内完成微服务架构体系的升级。在微服务架构升级中，医院会面临不同系统体系架构、不同架构业务系统的混合运行情况。不同平台架构之间的中间件、数据交互接口，业务系统之间的衔接等等，都是需要认真考虑和解决的问题。

4.技术准备

最后，微服务作为医院信息化应用新技术，医院信息部分的工程技术人员应该掌握微服务架构原理、微服务分析与建模、微服务 API，以及微服务的支撑技术（Docker、Kubernetes）等知识。

PU-073

多院区医院互联网出口整合改造与安全建设探索

宫晨曦 郝晓伟 许锋波 史乐萌 刘轲

河南科技大学第一附属医院

1 引言

多院区医院（也称一院多区医院）是指以资本或长期的经营管理权等为纽带建立的拥有两个及以上院区的医院[1]。目前，河南科技大学第一附属医院拥有景华院区、开元院区及医教院区三个院区，每个院区均拥有独立的互联网出口。外网服务器区业务分散在各院区，不利于进行整体、系统的安全防护，分散了网络及安全资源，业务可靠性差。同时，使用多个互联网出口也不便于对上网用户进行认证、审计。现考虑将两个院区的多个互联网出口统一为一个出口，集中资源，做到设备及网络架构的全冗余设计，以便于进行集中管控，统一互联网出口和外网服务区安全策略的配置，

提高对应区域业务的可靠性和安全防护级别。[2]对广大职工及病患提供便捷的无线上网服务，提升患者就医体验。同时加强对互联网访问尤其是病患上网行为的管控和审计力度，满足上级监管机关的要求。

2 现状及需求分析

2.1 医院网络发展现状

2016 年以前，医院的主要业务全部在内部局域网（不可访问互联网的局域网）网络运行，对互联网的需求极少，仅办公业务需要访问互联网。从 2017 年开始，医院的互联网业务飞速发展，先后开通了线上支付、线上预约挂号、互联网医院等业务。尤其是新冠疫情出现后，患者对互联网医疗的需求爆发式增长，对医院的互联网出口提出了更高的要求。

经过多年发展，医院多个院区互联网局域网（以下称外网）的管理出现了诸多问题：

一、医院拥有多个互联网出口，外网服务器区业务分散在多个院区，每个院区的外网均为独立运行的局域网，内部资源无法共享。

二、进行网络安全建设需对每个院区分别设计建设，不仅需投入更多的资金，且管理难度高。

三、使用多个互联网出口不便于对上网用户进行认证、审计，难以满足公安机关的监管要求。

四、院内各类管理数据的采集合并工作难度大、效率低、人力成本高。

景华院区改造前拓扑图，见图 1：

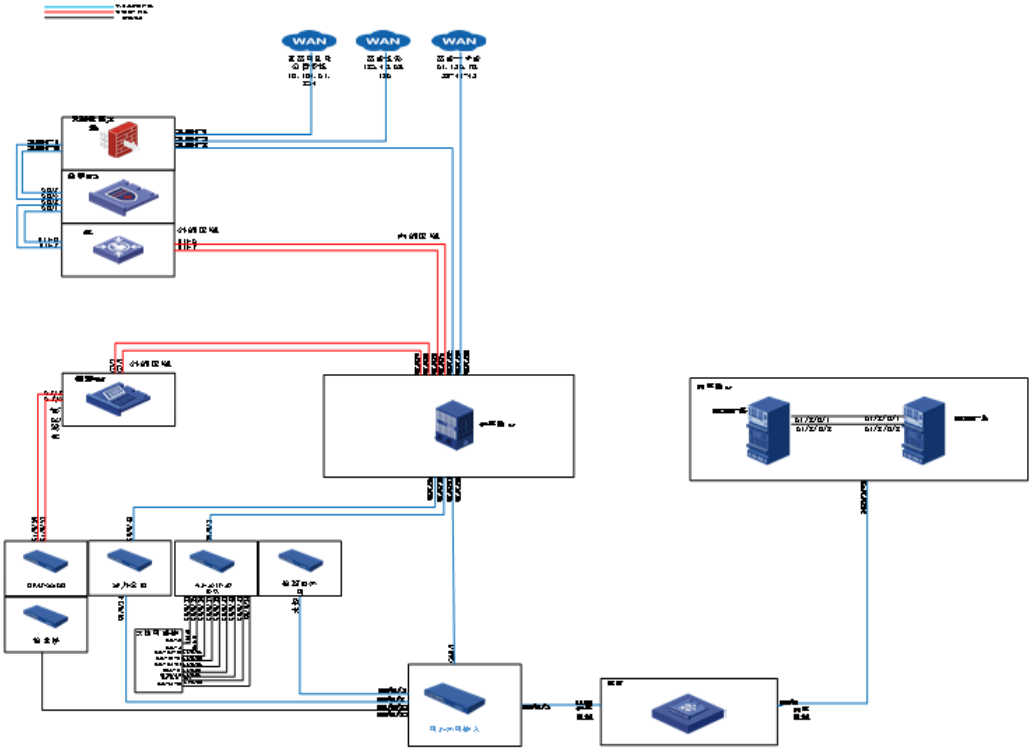


图 1 景华院区改造前拓扑图

开元院区拓扑图，见图 2：

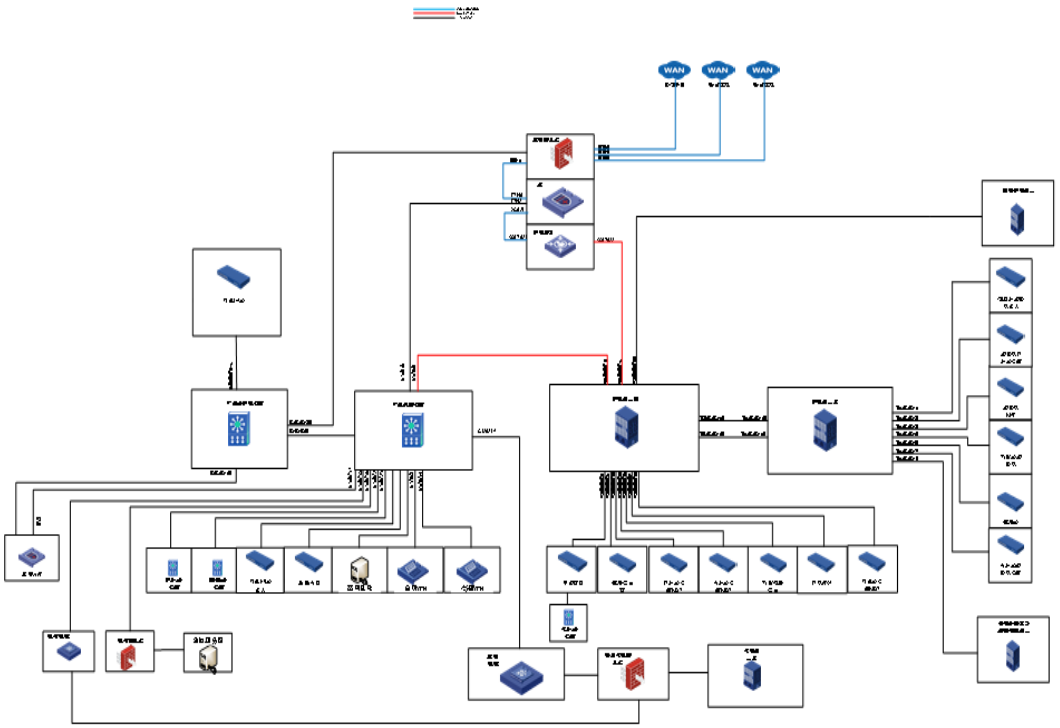


图 2 开元院区改造前拓扑图

2.2 医院外网安全需求分析

2.2.1 有效对抗网络攻击

目前互联网出现的网络威胁种类越来越多，不仅有非法入侵、网络渗透，还有网络欺骗、DOS/DDOS 攻击、恶意软件、垃圾邮件等。最严重的攻击是无法终止的攻击，如 DOS/DDOS 攻击，它还呈现出以下特点：攻击突发性强、攻击随机性强、攻击的方向不确定、攻击复杂度在提高。[3]

2.2.2 实现应用层安全防护

自 2003 年以来，蠕虫攻击、入侵技术日益革新并演变到应用层面（L7），这些有害代码会伪装成正常业务进行传播。传统的网络安全设备工作在传输层、网络层及数据链路层（L2-L4），不具有对数据流进行综合分析、深度监测的能力，无法有效识别伪装成正常业务的非法流量。因此，特洛伊木马、垃圾邮件、恶意代码和蠕虫等非法流量能轻而易举地通过业务系统开放的端口进出网络，从而达到攻击的目的。[4]

2.2.3 互联网数据的高速转发

互联网出口作为公网地址与内部地址 NAT（Network Address Translation，是指网络地址转换）转换的核心区，针对出口部署的设备需要提供高速 NAT 转发性能。其中，衡量设备性能的主要指标包括用户并发会话数与每秒新建连接数。用户并发会话数决定互联网到互联网地址翻译的吞吐量；每秒新建连接数的数量决定互联网用户突发同时访问互联网的情况下用户网络速度的快慢。

2.2.4 互联网多出口链路负载

医院接入多条互联网出口以保障出口链路的冗余性。采用多个出口容易造成其中一条或多条链路长期拥塞，而其它出口宽带闲置的问题。只有合理的均衡多个出口进行负载，才能在改善用户体验的同时提高网络带宽资源利用率，节约带宽成本。

2.2.5 互联网多出口智能选路

跨运营商访问的问题很明确，由于多运营商互联互通的问题，导致跨运营商访问存在相对高的延时，甚至有些访问会不通。因此，需要出口设备能够根据目的 IP 地址所属的运营商，智能地为用户选择最快最优的访问线路。

2.2.6 互联网出口高可靠性保障

互联网出口主要承担整个医院的数据转发、安全防护、NAT 转换、策略控制等业务功能。一旦出现故障，医院的对内和对外业务将全部中断。需要做到设备及网络架构的全冗余设计，避免单点故障。

2.2.7 互联网接入身份认证

医院患者、陪护流动性大，若采用非认证方式上网容易被犯罪分子利用互联网的“匿名”性，进行各种违法犯罪，导致医院因无法配合网监等部门完成网络违规犯罪行为的调查被问责。因而需要对上网用户尤其是病患的互联网访问进行认证，以便将 IP 对应到个人或者手机号，同时审计相关人员的访问日志，进行上网行为的分析及访问控制，如限制非法网站的访问甚至恶意发帖，维护公共网络安全。

2.2.8 Web 安全防护

Web 业务系统被互联网访问，访问对象不可控，需要注重其安全性及抗攻击性。在安全性方面，需要考虑 Web 系统的运行安全和内容安全两个方面，运行安全首先要保障系统可以在一个稳定的环境下对外提供稳定的 Web 服务，保障该服务可以被用户访问；内容安全，要保障 Web 对外提供的内容是我们希望对外提供的内容，要防止网站内容被恶意篡改，导致发布错误信息，影响医院形象。在抗攻击性方面，结合 Web 系统的应用特点，要从网络层和应用层两个方面去考量，首先 Web 系统对访问的流量要进行访问控制、过滤清洗，将非法流量合理引导，将正确流量准入引入。其次，要对访问的请求进行进一步的安全检测，保障访问数据的规格，请求目的都是在可允许范围之内。在提升链路可用性的同时实现深度防御，保障服务器区业务的安全性。

3 建设成效

根据现状分析，笔者对网络安全架构进行了优化设计。两院区外网核心交换机之间进行互联互通，通过路由变更将多院区变成平行的结构。访问互联网的数据统一通过景华院区的外网核心交换机，透过上网行为管理、IPS、防火墙等设备组成的安全防御层，交由出口负载均衡设备。通过出口负载均衡设备进行策略匹配、地址转换、寻址转发后，数据包再经过抗 DDOS 设备清洗流量，转发至合适的运营商完成数据通信。

外部访问的流量同样需先经过抗 DDOS 设备的流量清洗，再通过出口负载均衡设备的策略匹配、地址转换、寻址转发，后经过边界区安全防御层的多台设备，由核心交换机转发至 DMZ 服务器区，透过 DMZ 防火墙及 Web 防火墙组成的 DMZ 安全防御层交由服务器区负载均衡设备，通过负载设备的代理功能，访问到其所需业务系统。

本次项目改造采购的设备均支持 IPv6 业务的改造。随着医院院区的不断增加，智慧医院改造的兴起，新业务系统、服务器、终端电脑、物联网设备成几何指数增加，IPv6 的改造需求越来越迫切。IPv6 报文具有拓展报头，能提供路径选择，支持网络安全联动的整体化部署。IPv6 改造是未来几年医疗信息化发展的方向。因此，本次改造也为医院后续 IPv6 建设奠定了基础。[5]

改造后的拓扑图如下，见图 3：

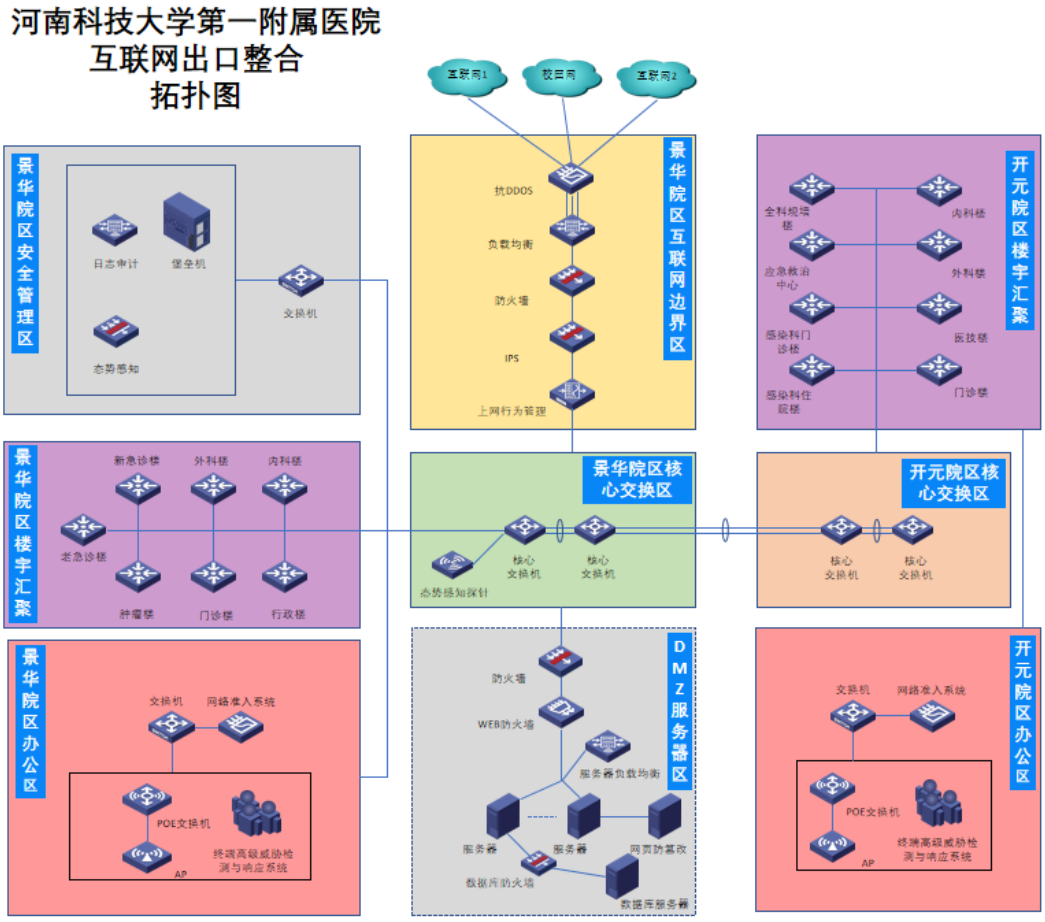


图 3 改造后的拓扑示意图

根据互联网出口整合改造前后对比，改造后有以下几点优势：

一、多院区外网局域网互通，有利于原有多院区独立局域网内的业务进行对接，避免占用出口带宽；

二、多院区原有的多个互联网出口迁移至单一院区，便于集中管理，集中进行安全防护；

三、避免信息化建设资金重复投入。医院原本需在各个院区分别投入建设，改造后，显著减少了医院的后期建设投入。目前，经过统计预期结余了 200 余万元的后期投入；

四、优化院内访问互联网的出口资源。互联网出口整合后，出口带宽大幅提升，跨运营商访问延时改善明显，提升了病患在医院的线上就医体验；

五、减少医院人力资源投入，减轻运维人员的压力；

六、规范医院对互联网的管理，提高医院管理效率。融合后互联网相关业务流程统一、报表统一，提高了多院区的整体性和协同性，构建了高效运行的信息管理体系。

七、进一步提升了医院网络安全防御体系的防护能力。改造完成后的近一年内，医院的网络安全事件大幅减少，同比上一年事件减少了 80%。

4 讨论

本次改造项目为探索性改造工程，已完成互联网出口整合建设，将多院区的出口进行了合并，以便于采用统一安全策略进行集中管控，在提升网络安全性、稳定性的

同时，降低了运营成本，减少了网络故障的发生，提高了安全响应效率，使医院获得了良好的安全收益和经济收益。然而本次改造仅完成了单链路组网建设，关键节点不具备冗余性，一旦该节点或设备出现故障，整个网络就会受到影响，甚至可能导致整个网络系统的崩溃。[6]为了保障出口的高可靠性、高可用性，下一步计划在互联网边界区进行冗余改造，避免单点故障。二期改造计划拓扑如下，见图 4：

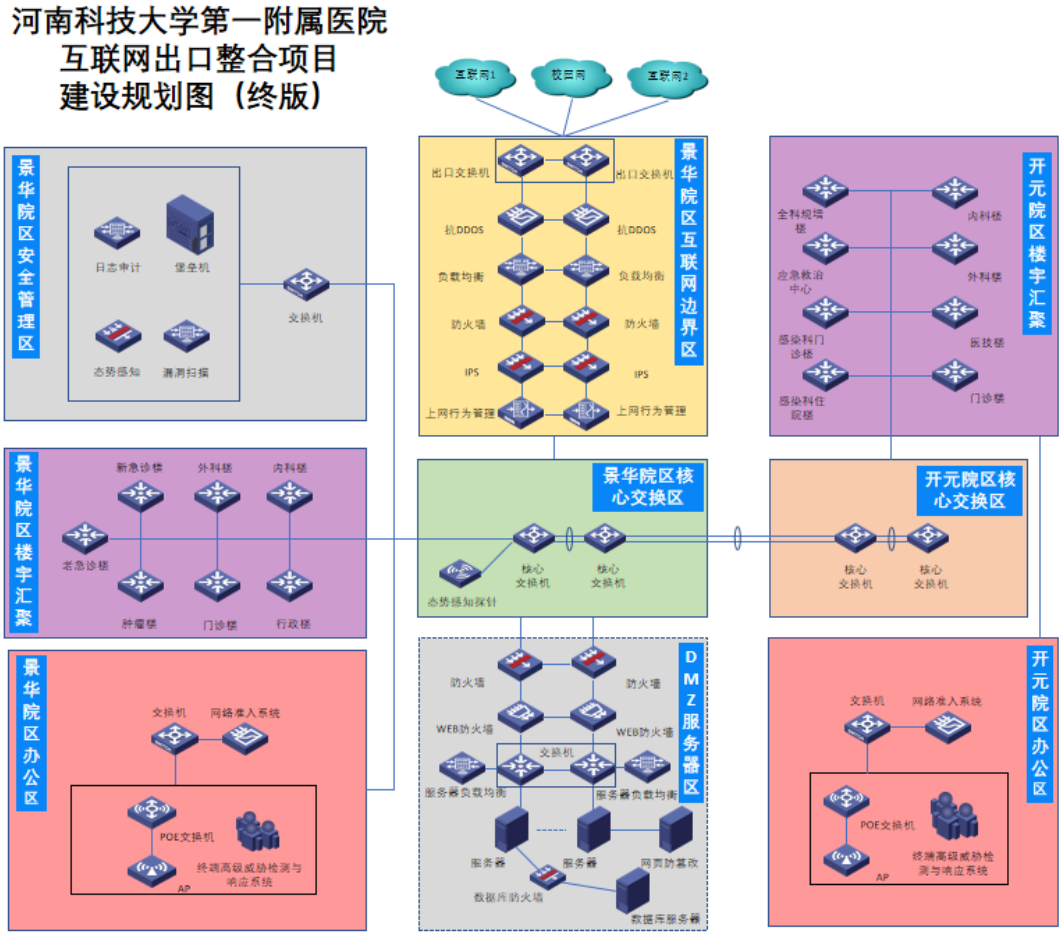


图 4 二期改造规划拓扑图

5 结论

多院区互联网出口改造项目建设实施近一年来，系统总体表现平稳高效，网络稳定可靠，医院职工对改造后的网速满意度较高，实现了出口资源的整合，降低了医院安全设备的投入，加强了医院互联网区域的安全防护能力，保障了互联网医院信息系统的正常运行，持续地为患者提供安全、方便、快捷、优质的医疗服务。网络安全防范体系的建设不是一个一劳永逸的工作，而是一个长期不断完善的过程，技术规划方案要能够随着安全技术的发展、外部环境的变化、安全目标的调整而不断完善。本次建设是在对医院网络系统可能面临的安全威胁进行风险分析的基础上，结合安全区域的划分，从计算环境安全、区域边界安全、通信网络安全和安全管理中心四个主要安全层面进行了整体的安全设计，在完成针对关键节点冗余部署的基础上，综合多种有效的安全防护措施，进行多层和多重防御部署，实现纵深防御，协同防护，为医院信息化建设保驾护航。

PU-074

以互联网“云平台”构建基层高效便捷超声质控体系

杨希[1] 李雪[1] 刘潇[1] 杨宇宏[1] 赖利[1] 张群霞[2] 冉海涛[2*]

重庆大学附属沙坪坝医院（重庆市沙坪坝区人民医院） 重庆市沙坪坝区超声质量控制中心[1]
重庆医科大学附属第二医院超声科 重庆市医学影像（超声诊断）质量控制中心[2]

1 引言

2021年，沙坪坝区人民医院超声医学科走访调研发现，沙坪坝区内21家社区医疗机构中超声检查已成为社区居民就诊、中老年体检、两癌筛查、慢病管理等不可缺少的重要辅助检查手段；其次发现基层超声科设置不规范、超声医师身兼数职、医技共职等现象较为普遍，超声诊断水平参差不齐、操作不规范、质控标准不统一、超声专业描述及诊断不准确，从而导致超声检查及诊断的漏诊及误诊现象常有发生[1]；然而基层医师规范化学习机会少，对超声质量控制工作陌生，无超声质量控制体系，无法落实部分医疗核心制度等情况，存在着极大的医疗安全及纠纷隐患[2]。

2 现状分析

针对基层医疗机构存在的以上诸多亟待解决的现实问题，结合全国和重庆市关于深入贯彻落实党中央、国务院关于实施健康中国战略和深化医药卫生体制改革的决策部署，加快推进区域医疗中心建设，完善医联体管理，发展“互联网+医疗健康”，优化资源布局，促进分级诊疗，解决基层群众看病难和看病贵问题[3]。因此依托现代5G通信技术和超声医学技术的发展，结合“互联网+医疗”和“医共体建设”以及“分级诊疗”的探索创新[4]，建设一套高效便捷的基层互联网超声医疗质控体系显得尤为重要。

3 质控体系架构设计

由重庆市沙坪坝区卫健委及区人民医院牵头，率先以区人民医院超声医学科为试点单位，以互联网医疗上接重庆市医学影像（超声诊断）质量控制中心，下联沙坪坝区基层21家社区卫生服务中心，开创互联网“远程超声诊断云平台”的建设应用。

3.1 创建覆盖沙坪坝区基层的互联网远程超声诊断云平台

2021年05月获批建设沙坪坝区远程超声诊断中心云平台。

项目平台系统群由三层级构成：

以重庆市医学影像（超声诊断）质量控制中心为专家指导，负责疑难病例三级会诊及质控把关，形成金字塔第一层级；

以沙坪坝区人民医院超声医学科专家、设备、质控体系、科研学术、教学等优质医疗资源为依托，构成项目平台系统第二层级，承担项目平台系统的分布式云中心及其相关核心专家组评审和数据资源云分享的功能；

以沙坪坝区21家基层社区卫生服务中心，构成项目平台系统第三层级。其分布于沙坪坝区各街镇，承担项目平台系统的基层超声医疗保障、质控服务和学习培训、大数据重要来源等工作。

2021年10月，远程超声诊断云平台正式运行。平台通过双向、即时、同步的音视频现场链接，实现群众在社区医疗机构就近首诊，即可获得与上级医院同质化的超声诊疗服务，节省就诊时间、降低看病费用、降低出行成本，解决基层病患对优质诊疗质量的需求。同时在面对突发紧急重大疾病、突发公共事件，提高诊疗效率。

3.2 实现沙坪坝全区基层医疗机构同质化的超声医疗质控

通过互联网远程超声诊断云平台，以现场网络会诊、报告审核的方式，实现全区21家社区卫生服务中心超声诊疗同质化质控，大幅降低基层医疗纠纷风险，提升基层医疗水平和口碑，提高患者就诊满意度及诊疗效果。

3.3 适应发展要求，推动基层超声人才规范化培养

对基层超声医师队伍进行规范化、系统性、长期持续的培训学习，既是远程超声诊断中心云平台建设的需要，更是保障平台日常高效、高质、高量运行的需要。为此，云平台在创建初期及运行后对基层超声医师进行多轮次规范化培训学习，推动基层超声人才规范化培养。

3.4 探索基层超声质控体系建设，力争打造全国基层超声质控样板工程

基层医疗机构的超声质控工作开展较困难，超声从业人员本身对质控工作陌生、不理解；缺乏专业的监督管理人员对定期其进行完成质控指导等。通过云平台的实施，向上连接重庆市医学影像（超声诊断）质量控制中心，向下覆盖21家基层社区卫生服务中心，架构起了全区统一的基层超声质控体系，力争将其打造称为具有自身特色的全国基层超声质控样板工程。

3.5 整合沙坪坝区超声诊疗数据资源，创建和开发医疗专项大数据库

通过云平台覆盖沙坪坝全区基层医疗机构的网络，可以对区内超声诊疗病例资源的分布、构成等大数据[5]。通过云平台进行图像采集、入库、建档、分析、统计并对其潜在科研价值进行开发研究，不仅有利于指导全区医疗工作规划的科学性、前瞻性，更有利于建立微区域的流行病调查、地方病监测、人口健康情况跟踪预测等功能[6]。

4.基层质控体系建立的创新成效

4.1 远程超声诊断云平台创建近2年，功能应用覆盖全区21家社区医疗机构，实现日常稳定的常态化运行，有效达成基本目标。

自2021年10月投入运行以来，截止2023年08月已完成接通21家社区卫生医疗机构、累计远程诊断超声病例19万余例、开展实时视频会诊191余次，为基层人员培训19轮次，不仅受到基层普遍欢迎，而且取得了强化基层超声质控、提升基层服务水平、增值社会效益等基本目标。

4.2 通过云平台建设，大幅规范和提升了全区21家社区医疗机构超声诊疗流程、标准化测量培训。

2021年进行8场规范化培训，从超声仪器的使用及调节、到腹部、心脏、血管、妇产各系统检查流程、存图规范，标准化测量，近400人次参与培训，解决了基层医疗机构“基本规范”问题；在平台投入运行后培训学习实现定期常态化，针对工作中发现的问题，不断持续改进，2022年进行12次专项规范化培训，主要针对心血管系统、妇产系统，从而做到“做好规范”的要求。通过远程超声诊断云平台使之成为基层医师与上级医院优质医疗资源的双向交互学习培训，从而大大提升了沙坪坝区基层医师专业技能水平。

4.3 实现21家社区医疗机构远程超声诊断实时审核。

基层社区医疗机构的超声医师，可随时通过互联网呼叫和连接云平台，将所检查的每位患者超声图片、视频、超声报告迅速上传至平台，平台值班专家立刻对所上传图像、报告进行审核，对不足之处进行修正完善超声报告，并及时反馈回传至社区医疗机构。2021年10月运行至今，已上传实时审核192713例超声病例，发现各类缺陷病例8231例，避免了严重医疗纠纷隐患62例。切实将优质医疗资源通过互联网技术下沉至基层。

4.4 云端实时超声视频会诊，与基层超声医师共同面对疑难病例

对于较复杂的疑难病例，通过申请预约，基层检查医生通过5G互联网技术将现场上机医生超声操作手法和超声影像实时双屏传输至云平台，平台值班专家通过网路

与基层医生进行交流，指导其实时操作手法与技巧，共同完成对疑难复杂的超声检查。目前与辖区内基层机构完成191余次实时视频会诊。

4.5 完善事前、事中、事后质控环节，实现PDCA质控闭环循环，初步构建基层超声全面质量控制体系

统一全区基层超声各系统标准化流程、统一更新全区超声危急值项目，制作《超声报告质量控制表》、《缺陷病例差错分析》《阳性病例统计登记》、《疑难病例讨论》、《超声诊断与临床符合率》《业务学习记录本》等记录表，完善全区基层机构超声科停电、网瘫、消防、患者跌伤、休克、猝死等应急预案，科室新进员工培训流程，拟定重庆市沙坪坝区基层超声质量控制方案，形成全区统一的事前质控环节。对运行超声诊疗过程进行监督及干预，加强超声诊疗过程中的事中环节质控。每月对平台全区的误诊、漏诊超声病例进行总结、分析，并向上至卫健委汇报，向下至基层社区医疗机构反馈实现PDCA质控闭环循环。超声检查的规范化、制度化、标准化是最终目的，也是提升医疗服务水平的重要突破口[7]。远程超声平台的实施，初步构建基层超声全面质量控制体系。帮助实现了分级诊疗中的“同质化”最困难的问题。

4.6 以为超声诊疗为突破口，创新推进分级诊疗

2022年，通过远程超声诊断中心云平台发现复杂疑难病例671例，其中约403例患者转诊到我院进行进一步超声、增强CT、磁共振及胃肠镜检查，约311例患者入院进行治疗。基层社区门诊转诊79例患者到我院行介入超声穿刺活检，主要为颈部包块、体表包块、腮腺包块、肾囊肿、肝囊肿、甲状腺结节、乳腺肿瘤等，这些患者分别入住我院耳鼻喉科、普外科、肾内科、消化科、肿瘤科行相关手术或内科治疗。危急值对于病人的治疗和康复非常关键，因为它能够提醒医生或病人本人及时采取措施，以避免可能的重大风险和后果[8]。为了更好完成2022年国家医疗质量安全改进目标-提高超声危急值10分钟内通报完成率，2022年09月我们对全区社区卫生服务中心超声危急值进行台账管理，因重庆疫情影响，全年仅上报12例超声危急值，但东部地区7例患者转入我院进一步治疗。云平台的运行，不仅解决了基层超声发展最难的同质化问题，同时也为基层医疗分级诊疗提供了新思路。

5 基层超声质控体系未来展望

5.1 以区域超声质控中心为依托，通过2年远程超声诊断中心云平台的运行，进一步推进分级诊疗体系建设，补齐沙坪坝区基层医疗机构超声诊疗发展的短板，真正让优质医疗资源通过互联网技术便捷下沉，为基层百姓办实事、办好事。

5.2 基于5G核心技术的远程超声会诊平台，不仅突破了医疗空间的限制和医师经验的局限[9]，而且通过对“云平台”运行管理的探索，希望能探索出一条可在重庆其他区域乃至全国可复制、可推广的互联网+超声医疗创新模式和有效途径。

5.3 值得探讨，在不远的将来，如果全国各区县都能建立起像“沙坪坝远程超声诊断中心云平台”这样从专科起步、融合“互联网+医疗、医共体、分级诊疗、智慧医院”“云思维”“人工智能”等探索范例并实现互联互通、资源共享[10]，将形成基于“区块链”模式下的全国超声诊断大数据资源，该数据库的潜在科研开发价值不可估量。

PU-075

加速推进医联体双向转诊建设的研究与探讨

曾婷
长沙市中心医院

1 引言

2023 年，我国卫生健康信息化发展总指数为 73.60，指数表明互联互通的信息化平台基础日益牢固、医疗卫生机构之间协同不断强化、卫生健康信息化组织保障体系逐步完善、便民服务的应用成效日益凸显、群众幸福感、满足感进一步增强。目前我国正面临着城镇化进程加速、人口老龄化亟待缓解、疾病谱日益复杂等挑战，与此同时，基层医疗卫生机构的资源配置水平不高、服务能力存在不足。如何疏通摆在基层医疗卫生机构面前的难点和堵点是国家卫生健康委一直探讨的核心问题。

分级诊疗制度是中国特色基本医疗卫生制度的重要组成部分，而医联体信息化建设是推动分级诊疗制度建设的重要抓手。信息化平台基于互联网技术提供诸如健康档案互查、双向转诊、远程会诊和监护、远程教学等功能，从很大程度上为优化医疗服务格局赋能，实现优质医疗资源下沉基层、区域内院间资源共享、信息互通，是医疗卫生服务体系的一次自我整合、自我提升。

2 研究现状与需求分析

目前，国内多个医院积极响应国家号召,对信息化平台建设进行了不同程度的探索。例如，上海市儿童医院先后于2017年和2021年应用了基于云平台的区域医联体信息系统[1]和基于“互联网 +”设计了跨区域医疗信息共享与服务协同平台[2]；2018年,靖江市卫生和计划生育委员会致力打造了区域卫生信息平台[3]；2021年，中山大学附属第一医院设计了医联体医疗信息平台[4]；2023年，广东省第二人民医院阐述了基于信息化的互联网医疗服务一体化医联体运行模式[5]。经过近些年的发展，部分地区医联体医疗信息平台建设取得了一定成效，涌现除了一批典型的经验。浙江湖州、山东日照、广州花都、辽宁大连推进城市医联体网格化布局管理，实现了资源下沉和资源共享。浙江德清、福建尤溪推进紧密型县域医共体建设，有效提升了县域医疗服务能力[6]。

“十四五”全民健康信息化规划发展目标指出，到2025年，要初步建设形成统一权威、互联互通的全民信息平台职称保障体系，基本上要实现公立医疗卫生机构与全民健康信息平台全覆盖，二级以上医院基本实现院内医疗服务信息互通共享，三级医院实现核心信息全国互通共享[7]。因此，制定切实可行的“双向转诊”配套支持政策，引导患者小病、慢病去基层医疗卫生机构，大病、危重病和疑难病到上级医院；制定临床路径，统一诊疗规程，实现单病种规范化治疗，为基层医疗卫生机构提供同质化、一体化服务是必要的。

目前，长沙市医院之间主要的转诊方式为人工转诊、转检，而人工转诊、转检都要去窗口挂号、缴费，时间成本较高、便捷性不强且无法按照按要求自动采集相关数据。针对本市缺少基层医疗机构与上级医院互联互通转诊系统的实际情况，若要实现双向转诊、转检，则双方的信息系统必须进行深度对接，从而实现多个医疗卫生机构之间的信息共享。加速推进医联体建设是搭建上下联动的新型诊疗服务模式、破除资源共享局限、畅通数据共享渠道、联动医疗服务能力、提升病患转诊转检体验感的重要方式，也是践行健康中国战略的有效路径之一。

3 系统架构与功能模块

3.1 系统架构

信息化系统设计应当遵循先进性、实用性、安全性、开放性与标准化等原则，在保证系统建设、使用、维护低成本、高可靠性与易用性和易于维护性的同时还应当具有良好的可扩展性，以保证医院在不断发展壮大形势下的管理需要。客户端软件系统要体现易于理解掌握、操作简单、提示清晰、逻辑性强、直观简洁等特点，并能针对医院输入项目的特点对输入顺序专门定制，保证操作人员以最快速度和最少的击键次数完成工作。医联体信息化平台中的双向转诊应能够实现病人预约就诊、上转下转、检查预约、检验预约、床位预约等业务，上、下级医院可以相互查看对方产生的检验结果、病历记录等信息，同时系统应为业务部门提供其所需的统计分析报表。大致架构设计可参考图1：

图 2 架构设计示意图

3.2 功能模块

双向会诊主要涉及到基层医疗卫生机构将病人上传到上级医院以及上级医院将病人下转到基层医疗卫生机构，其流程可以参考图 2。

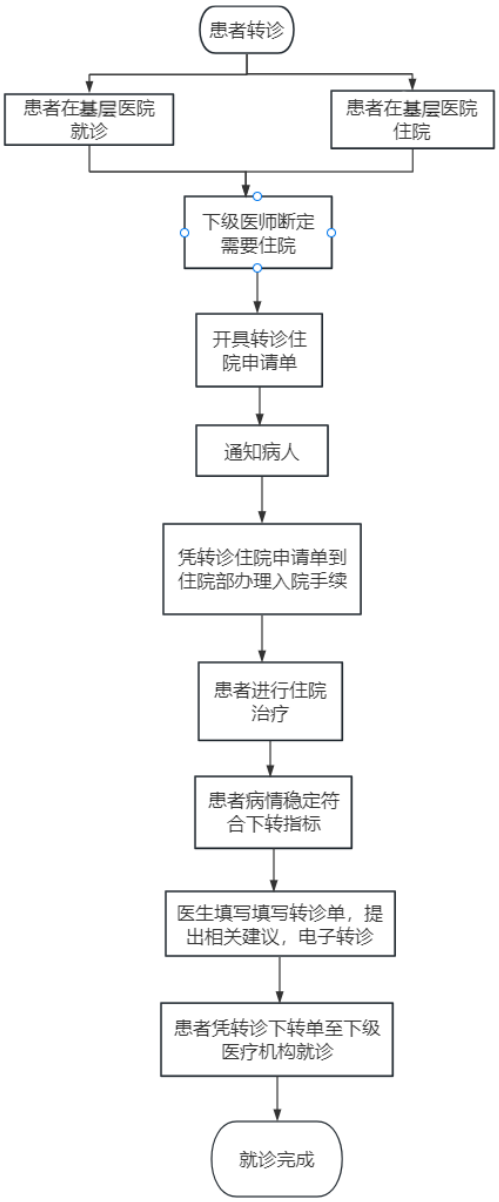


图 2 转诊流程示意图

（1）上传管理

上传管理即为基层医疗卫生机构将病人向上级医院进行转诊。一般需要包括上传申请、上传审批、转诊病人管理、查看各类检查结果等功能。其中，上传申请部分由基层医疗机构根据上级医院要求提供上传所需信息并打印出上转申请单，病人凭上传申请单到上级医院办理住院或免费挂号，上转审批部分由上级医院审批基层医院的转诊申请单，并根据情况办理入院或挂号信息。转诊病人管理部分针对转诊的病人进行统计，基层医疗机构与上级医院可以通过此目录能查询到转诊病人的相关信息。查看各类检查结果部分向上级医院提供查看病人在基层卫生机构的检查结果的功能。

（2）下转管理

下转管理即为上级医院将病情稳定的病人向基层医疗卫生机构进行转诊。一般需要包括下转申请、下转接收、查看病历、查看各类检查结果等功能。其中，下转申请部分由上级医院根据医院要求提供上传所需信息并打印出下转申请单，病人凭上传申请单到对应基层医院办理住院或免费挂号。下转接收部分由基层卫生机构审批上级医院的转诊申请单，并根据情况办理入院或挂号信

息。基层卫生机构可以通过查看病历部分来查看病人在上级医院产生的病历信息，通过查看各类检查结果部分查看病人在上级医院的检查结果。

（3） 预约管理

预约管理可以根据检查的类别进行设计，例如检查预约、病理检查、超声检查预约、放射检查预约、内窥镜预约、治疗预约、床位预约等功能。其中，检查预约部分由基层卫生机构向上级医院预约细菌检验、基础检验、免疫检验、生化检验等。病理预约部分由基层卫生机构向上级医院预约组织病理学诊断、细胞病理学诊断等检查。超声检查预约部分由基层卫生机构向上级医院预约腔内彩超、产科彩超、血管彩超、腹部彩超等检查。放射检查预约部分由基层卫生机构向上级医院预约 CT 检查、放射检查、磁共振等检查。内窥镜预约部分由基层卫生机构向上级医院预约电子胃镜、电子肠镜、电子直肠镜、电子乙状结肠镜等等内镜检查。治疗预约部分由基层卫生机构向上级医院预约放疗定位等等治疗。床位预约部分由基层卫生机构向上级医院预约床位。图 3 为检查预约流程示意图。

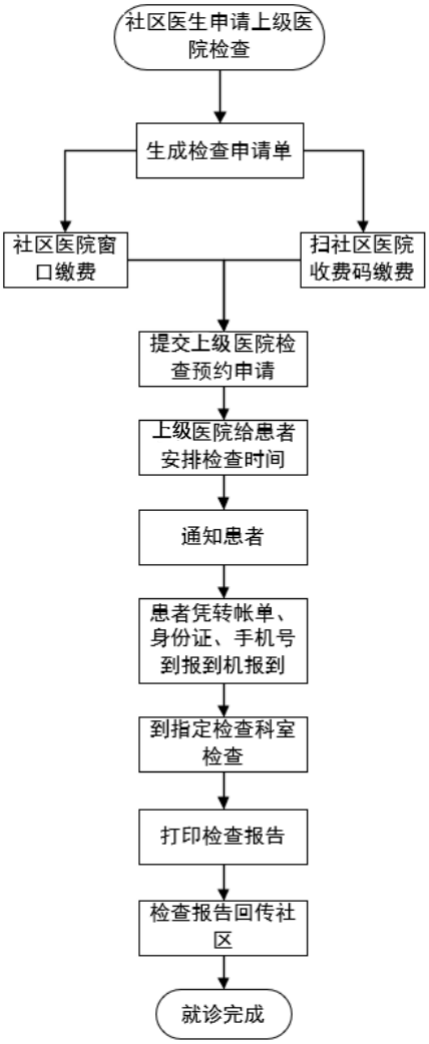


图 3 检查预约流程示意图

（4） 财务管理

财务管理是规范医院财务行为，提高资金使用效益的必要模块。根据实际需求可以设计收费明细、收费汇总等功能，其中收费明细部分向基层卫生机构与上级医院提供按条件统计在上级做的各类收费明细信息。收费汇总部分向基层卫生机构与上级医院提供按条件形成的总览报表情况。

（5） 三目对照

三目对照主要是为了畅通上下级医疗机构的交流，保障项目的同步性与一致性，包括医嘱对照、收费项目对照、诊断对照。由于双向转诊是为了病人更好享受到优质医疗队伍的服务，因此进行医嘱对照、收费项目对照时，基层卫生机构可以采用上级医院的医嘱和收费项目，诊断对照部分由基层卫生机构将本级诊断信息与上级医院的诊断情况进行对照，从最大程度上进行资源整合并保证医院的营收。

（6） 远程会诊

远程会诊一般采用音视频交互式，因此对远程硬件设备、网络以及服务器有较高的要求。一般可以购买第三方公司的相关设备和服务。在基层卫生机构以及上级医院各自配备一套远程设备，并按照实际需求建设两三家基层医疗卫生机构和上级医院的远程会诊服务以实现云会诊，同时平台需要自动上传病历资料、报告等信息。

4 结语

加速推进医联体建设是打通城乡医疗资源共享最后一公里的必要措施。目前，医联体信息化建设取得了长足的进展，但是也存在一些问题。例如，针对新型业态的管理体系尚不完善、基层医疗卫生机构设备配置存在缺位、偏远基层医疗队伍互联网意识相对欠缺、网络环境质量较差等，从一定程度上降低了医患体验感。针对上述问题，可以从以下两个方面入手：一方面，充分发挥卫生强省带头作用，抓典型，取经验，辐射带动周围省市开展实践。制定考核与奖励办法，定期做好健康宣教与指导，提高基层医护人员对双向转诊、远程医疗的认知度与参与度。另一方面，加大基层基础信息设施建设，实现网络全覆盖，发挥大数据、云计算技术优势，推进多方医疗数据共享与医疗资源联动。

PU-076

经济不发达地区三甲医院信息化建设与项目管理

孙盛轶、姜永峰
佳木斯市中心医院

1.引言

1.1 研究背景

随着卫生行业的快速发展和信息化技术的广泛应用，信息化建设已成为提高医疗服务质量、优化资源配置、提升管理效率的重要手段。经济不发达地区三甲医院作为城市最重要的医疗卫生机构之一，具有很大的影响力，其信息化建设程度直接关系到所在地区的发展水平。当前三甲医院的信息化建设仍面临着一系列挑战，如信息孤岛、系统老化、安全隐患等，这些问题在一定程度上制约了医院的发展。因此，开展信息化建设与项目管理研究具有重要的现实意义。

1.2 研究目的

本研究旨在探讨经济部发达地区三甲医院信息化建设的现状、问题及原因，提出加强信息化建设的对策建议，同时为类似医院的信息化建设提供参考和借鉴。通过本研究，希望能够提高三甲医院的信息技术水平，优化医院的管理流程，提升医院的综合竞争力，为市民提供更加优质、高效的医疗服务。

1.3 研究方法

本研究将采用问卷调查法、访谈法等研究方法，以确保研究的科学性和可靠性。

问卷调查法：针对医院的相关医务人员进行问卷调查，了解他们对医院信息化建设的认识、需求和建议，以便更好地把握现实情况。

访谈法：对医院的领导和主要负责人进行访谈，深入了解医院信息化建设的具体实施过程、取得的成效以及遇到的问题，为研究提供详实的第一手资料。

2.经济不发达地区三甲医院的信息化建设现状

2.1 经济不发达地区三甲医院信息化建设滞后的原因

资金限制：经济不发达地区的医院通常面临资金紧张的问题，信息化建设需要购买硬件设备、软件系统，并进行培训和维护等，这些都需要相当的资金支持。

技术水平不足：信息化建设需要医院具备一定的技术实力和人才储备，但经济不发达地区的医院可能缺乏专业的信息技术人员，无法有效地开展信息化建设工作。

人员培训和意识问题：信息化建设需要医院内部的医务人员接受培训，学习使用新的系统和工具。然而，经济不发达地区的医院可能缺乏培训资源和意识，使得医务人员对信息化建设的重要性和使用方法缺乏了解。

管理体制和文化差异：信息化建设需要医院内部的管理体制和文化支持，但经济不发达地区的医院可能存在管理体制不完善、决策效率低下、对新技术和变革的抵触心理等问题，阻碍了信息化建设的推进。

2.2 经济不发达地区三甲医院信息化建设的必要性

提高医疗服务质量：信息化建设可以提高医院的管理效率和服务质量，通过电子病历、电子医嘱等系统的应用，可以减少人工操作和纸质文件的使用，提高医疗信息的准确性和及时性，从而提高医疗服务的质量和安全性。

提升医疗资源利用效率：信息化建设可以实现医疗资源的优化配置和利用，通过电子排班、预约挂号等系统的应用，可以减少患者等待时间，提高医疗资源的利用效率，缓解医疗资源紧张的问题。

促进医疗协作与共享：信息化建设可以实现医院之间的协作与共享，通过电子病历、远程会诊等系统的应用，可以方便医生之间的交流与合作，提高医疗诊断和治疗的水平，特别是在经济不发达地区医疗资源匮乏的情况下，可以借助信息化手段实现资源共享，提高医疗水平。

便利患者就医体验：信息化建设可以提供更加便捷的医疗服务，通过在线挂号、在线支付等系统的应用，可以减少患者的排队等待时间，提高患者的就医体验。

3.医院信息化项目管理的挑战

3.1 项目需求分析与规划

在信息化项目的初期阶段，需要进行项目规划和需求分析，明确项目的目标和范围，确定项目的关键节点和交付物，以及明确用户的需求和期望。项目计划是项目管理的基础，应包括项目的时间表、资源分配、风险管理等内容。通过制定详细的项目计划，可以有效地管理项目进展，并及时应对风险和问题。

3.2 项目实施过程管理

在项目实施过程中，需要进行项目的进度控制和质量管理，确保项目按时按质完成。同时，需要与相关部门和人员进行沟通和协调，解决项目实施过程中的问题和风险。实施过程中应建立跨部门的项目管理团队，信息化项目需要各个部门的协作和配合，因此建立跨部门的项目管理团队非常重要。该团队应由具有项目管理经验和医疗行业知识的人员组成，以确保项目的有效实施。

3.3 项目验收与评估

项目管理团队应定期监控和评估项目的进展情况，及时发现和解决问题，确保项目按计划进行。在项目完成后，需要进行项目的验收和运维，确保项目交付的成果能够正常运行和使用。此外，还需要进行项目的评估和改进，不断提升信息化建设的效果和水平。

4.案例分析

4.1 佳木斯市中心医院信息化建设案例

我院位于黑龙江省佳木斯市，是一所市属三级甲等综合医院，医院信息化始建于 2000 年，中间经历过几次升级改造，随着医院不断发展，原有的信息系统已不能满足国家对三甲医院信息化建设要求和医院自身信息化方面的需求。为满足医院在电子病历应用水平评审、医院互联互通等级评测、智慧医院等级评审，实现医院人、财、物全要素协同，持续提高就医体验、诊疗效率和质量、运营管理水平及创新能力，我院于 2022 年 1 月上线了新系统。

信息系统主要功能模块、硬件基础架构、实施方案选择：

主要模块功能：HIS、EMR、LIS、PACS、“掌上医院”、手麻重症系统、透析系统、输血系统等所有业务系统，均使用国内某知名品牌产品，实现了医院信息的一体化管理，大幅降低了资金支出，也避免了产生大量接口和由此带来的工期延误和扯皮推诿。

硬件基础架构：我们采用服务器虚拟化技术，这样大大减少实体服务器数量，虚拟服务器可以动态调整个应用服务器配置资源，按照不同模块应用服务器需求进行量身配置，为经济不发达地区的我院大幅节省硬件资金投入，存储采用分布式存储，具有高可靠性、高扩展性、易于维护、投入资金少的特点。

实施方案选择：结合我院实际情况采用先分后总的实施方案，即先在规模较小的分院区进行试点实施上线，稳定后再在总院区进行全面铺开上线。这种实施方案的选择极大减低了项目实施的风险，避免了因信息系统上线造成医院正常运营混乱的现象，同时系统切换方案采用新老系统部分时间并行的方式，即老系统中住院患者在新系统上线后仍在老系统中进行操作，新系统只接收新入院和门诊系统，待老系统中所有住院患者全部出院后，老系统彻底下线不再使用，此种上线切换方案更加稳妥，避免新老系统间因数据流交叉造成财务、医疗等方面的混乱，使新系统上线更加平稳的同时也可以在新系统上线出现问题时快速切换为老系统，不影响医院正常运营。经过分院区新系统试运行半个月，新系统运行稳定，使用效果良好，得到医院领导、医护、患者良好的反馈，总院区于半月后正式上线，有分院区上线经验作为借鉴，总院区上线平稳，取得的良好的实施效果。

4.2 案例分析结果

提高了医疗服务质量。我院顺利通过省卫生健康委电子病历评级，评定级别为四级，同时通过电子病历系统等技术的应用，医生可以更加方便地获取病人的诊断和治疗信息，提高医疗服务的准确性和效率。

提高了医院管理效率。通过医院管理系统的应用，该医院实现了各部门之间的信息共享和协同工作，减少了重复劳动和错误决策的发生。

促进了医院大数据的互联互通，让数据更精准，让管理和决策更有依据，进一步规范了医疗、护理和物价等工作流程，减少了临床医护工作差错的发生；极大地方便了患者就医，各种自助设备、手机端医院，让患者少排队、不排队，真正做到了让信息多跑路，让患者少跑路，通过智慧医院建设，优化诊疗流程，优化医院资源配置，提高患者就医体验，一部手机轻松解决医院看病问题。

4.3 经验总结

我院在信息化建设方面取得了较大的成功，以下经验值得借鉴：

项目管理团队的设立至关重要。在项目实施过程中，项目管理团队始终保持与各方的良好沟通与协调，确保项目的顺利推进。

基础数据字典完整性和准确性，对于信息化项目建设，项目是否顺利部署、测试、培训、上线、交付等，很重要的内容是数据的基础字典的质量，这些信息在前期准备阶段至关重要，需要保证每一条数据都准确无误，一步错步步错，将严重影响上线的质量，甚至导致上线的失败。

新系统的培训与操作熟练度，综合项目建设实施和管理经验，特别是一些大型的项目或者一个新的项目，信息系统的培训也是项目正常上线的至关重要因素，培训直接影响了正式使用的系统操作熟练程度，往往一些新的系统的上线由于功能、界面、操作、流程等变化，操作人员都需要提前熟悉操作，提高新系统的操作速度。

第三方厂商标准化程度及配合度，医院信息化普遍具有供应商多、系统多、标准不统一、数据质量差异等问题，往往一个新项目建设，不管接口方式是点对点还是集成平台，都一定会涉及到和第三方系统对接，如何保障第三方系统积极配合接口开发是此项目建设的痛点、难点，因此需要和第三方厂商积极沟通协商，争取最大的配合。

5. 结论

信息化建设是三甲医院提升医疗服务质量和管理水平的重要手段。然而，信息化项目管理面临着一系列的挑战。通过明确项目目标和范围、建立跨部门的项目管理团队、制定详细的项目计划以及持续监控和评估项目进展，三甲医院可以有效地应对这些挑战，推动信息化建设的顺利进行。只有在良好的项目管理下，信息化才能真正发挥其在医疗服务中的潜力，为患者提供更好的医疗体验。

PU-077

医疗高值耗材自动计费的应用研究

周逸帆 周敏* 倪剑
浙江大学医学院附属第一医院

前言

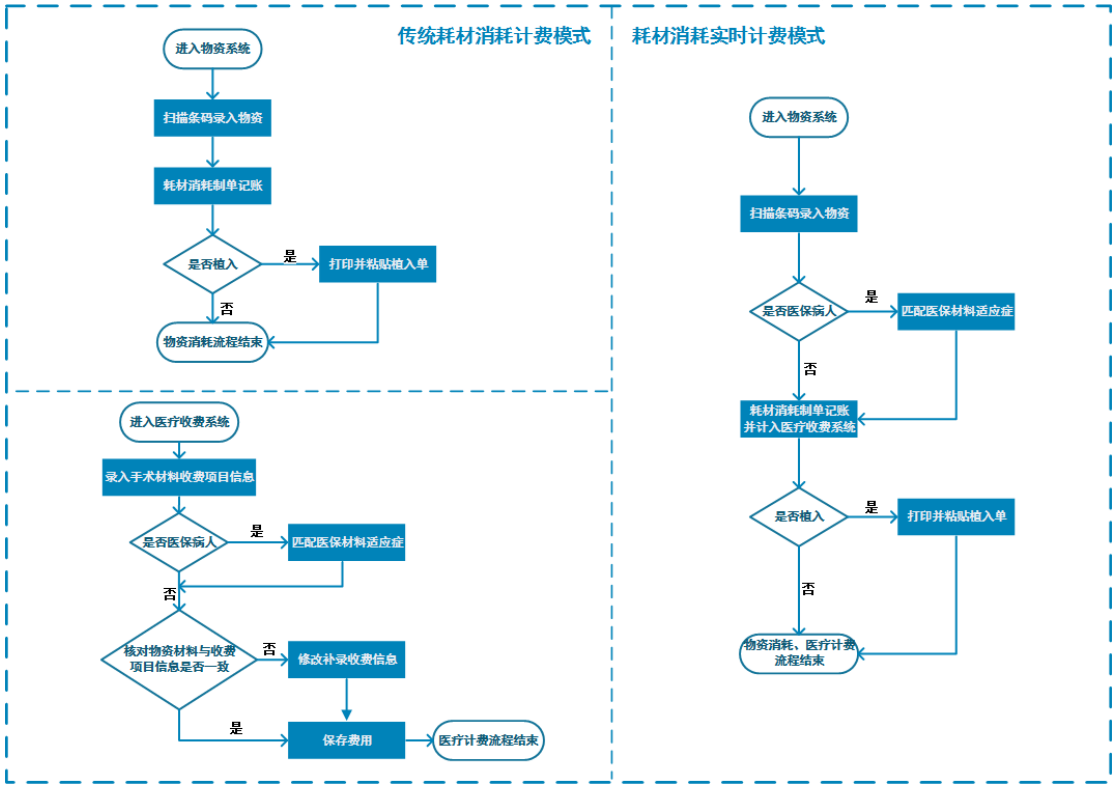
医疗高值耗材主要指对人体使用，关系人体生命安全的，且在生产、运输、配货、消费、使用过程中须严格控制，限定于特定专科专用，价格相对较高的消耗性医疗器械。目前常见高值耗材分为心脏介入类、外周血管介入类、人造关节类、人工晶体类、消化、麻醉材料类和其他类。随着人民生活水平的提高，人民对医疗救治水平，医用耗材的使用产生巨大的需求，大量新型设备、新种类高值耗材逐步应用在医院手术室/介入室等重点部门[1]。医疗高值耗材已成为疾病医疗过程中不可或缺的重要组成部分，由于医疗高值耗材使用的特殊性，如何进行医疗高值耗材使用管理，成为各大医院耗材科学化、精细化管理的重难点[2]。

自动医疗耗材计费与传统医疗耗材计费应用模式比较

近年来医疗信息系统飞速发展，医疗收费、控费系统、物资管理系统也迎来了多次迭代更替，虽然新的信息系统在各自业务范围内为业务的流程梳理、固化带来诸多好处，但也随之也带来了业务数据流与财务数据流的数据流割裂，医用耗材使用数据与 HIS 计费数据无法协同问题，难以实现数据资源的高效交流与共享[3]。

传统医疗耗材计费模式涉及两套系统间重复切换录入，容易产生漏费、错费、溢库等情况。术后还需逐一耗材比对，确认耗材使用登记与录入的耗材收费项目价格、数量、型号规格是否一致[4]。

自动医疗耗材计费模式将两套系统间计费简化为一套，手术室耗材计费系统与医疗耗材使用登记系统自动一对一关联，实现一物一码[5]，使用登记、手术计费、物资记账、术后追溯的全过程记录，减少多头操作，避免形成信息“孤岛”，实现各部门、各环节间精确、有序的数据流动，提高医疗计费管理效率[6]。



图一、传统耗材消耗计费与实时计费模式比较

进行自动医疗耗材计费遇到的困难

进行自动耗材计费改造时遇到多方面、多部门改造的困难，进行自动耗材计费面临改造的系统较多，涉及多种软件开发商间协调问题，具体问题如下：

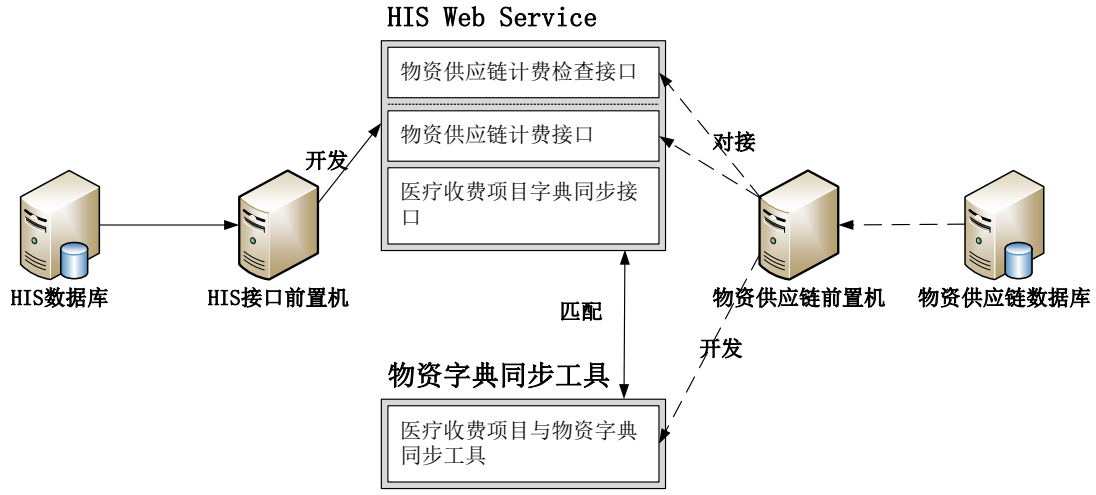
1) **医用耗材与医疗收费项目对照匹配难。**医用耗材与医疗收费项目将面临对照匹配，需要医疗物资管理人员、临床一线进行高值耗材使用的护士与医保控费管理部门三方面都需要进行物资与医疗收费字典的匹配，任何一个项目的匹配出错将面临错收费风险。

2) **医保交易材料适应症关联难。**多种医保计费交易模式不同，所需上传的字段也不同，涉及不同医保代码、医保材料适应症与医疗耗材关联的问题。

3) **医保耗材改价交易退费难。**收费项目与材料对照面临耗材采购价格变更时，需要更新收费项目价格的问题。如何统一变更两个系统价格成为计费与医保结算问题。例如：手术室护士在 A 材料价格变更前给某患者自动计费使用，高值耗材价格为 B 元，并且已通过 HIS 医保交易接口上传，医保处记录为“消耗 A 材料，金额 B 元，医保代码 C”。且此时 A 材料库房内已完全消耗完毕，更换为同类型、不同厂家、同名、同医保代码的新产品 A’，价格为 B’元，出院时病人或因退保、停保等个人原因无法使用医保交易，需要调用医保退费，自费出院。当前高值耗材 A 对应的医疗收费项目价格已经在系统上更改为新产品 A’的 B’元。调用医保退费接口时，医保处记录为“消耗 A 材料，金额 B 元，医保代码 C”，而退费上传是记录为“退除，消耗 A’材料，金额 B’元，医保代码 C”。医保处会报错相同医保代码上传医保金额与退费医保金额不一致，交易失败。

总体设计解决方案

1) 由于物资系统与 HIS 系统间涉及交换数据较多，接口数目较多，为降低开发难度，减少业务风险，采用医院信息科与物资系统共同开发接口的方式来实现，HIS 端 WebService 接口由医院信息科自行开发，降低项目接口沟通协调难度。



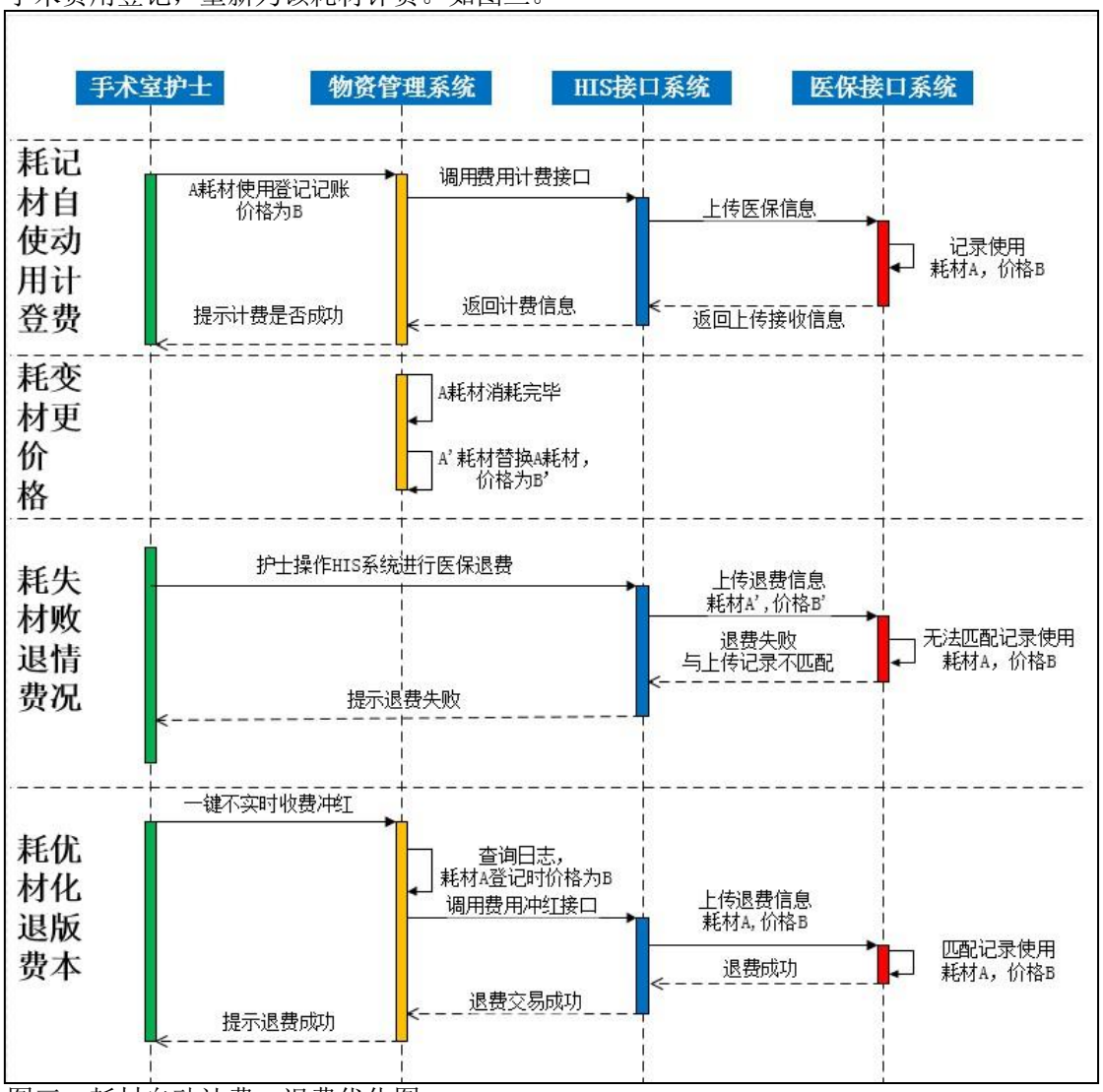
图二、HIS 系统与物资系统接口交互分工与调用过程

2) 设计多维度、多索引的收费项目与医用耗材结构化对照工具，配置产生以医疗费用或医用耗材为索引方式的多对多字典关联工具。

3) 在 HIS 接口系统方面，将所有种类的医保及其涉及的字段进行了梳理。护士在医疗物资管理系统使用登记时，录入使用患者，系统会主动查询患者医保归属地类型，并关联对应医保所匹配的材料适应症，在护士选择时材料适应症时，默认选中同种诊断、同种医保、且被选次数最多的医疗材料适应症代码，提高护士操作便利性。

4) 关于医保交易过程中材料价格变更问题，探访多家医院各家都有不同解决方案，甚至有医保办不管理材料收费项目，全局由物资管理部门维护的情况。鉴于此情况的偶发性、复杂性及多部门合作退费的不便利性，系统设计了一键不实时收费冲红功能。利用此功能冲红医

疗周期内价格变更的耗材，当自动收费的医疗耗材于医疗物资管理系统冲红时，医疗物资管理系统后台会主动根据记账时日志，形成一笔记账时价格的 HIS 退费交易，并调用 HIS 退费接口主动向医保进行退费交易（此时退费记录为“退除，消耗 A 材料，金额 B 元，医保代码 C”）。当后台医保退费交易完成后，该耗材冲红才算完成，然后护士需按原有操作流程的在 HIS 的手术医疗计费系统，录入使用的耗材对应的物价收费项目和医保材料适应症代码，保存手术费用登记，重新为该耗材计费。如图三。



图三、耗材自动计费、退费优化图

总结

上线实时收费功能以来，医院手术室、心导管室、放射导管室、肝胆胰介入中心等重要植入性材料、高值耗材消耗部门，均未出现高值耗材费用错漏情况，用户使用反馈良好，提升医护协同效率。通过医疗高值耗材自动计费模块的研究应用，将物资流与财务流相互关联、穿插合并、业务协同，极大地提高了手术室护士操作的便利性，通过实现计费关联从财务层面也可反向追溯患者材料使用情况，并从源头上避免手术室高值耗材使用的漏费、错费、溢库等问题[7]。对医院耗材使用走向精细化、规范化起重要的推动作用。

PU-078

基于 PMC 指数模型的卫生健康信息化建设政策评价分析

鲁志威，王剑莉，贺莲，王宁宁，李彩文，徐沛颖，高俊，谢银梅
广东省卫生健康宣传教育中心

1 引言

“数字中国”建设是推进中国式现代化建设的重要引擎，中共中央 国务院将卫生健康信息化建设列为保障医药卫生体系有效规范运转的重要支柱之一，卫生健康信息化建设的重要性得到进一步凸显。近来年，卫生健康信息化建设广泛应用于“互联网+医疗”[1]、基本医疗卫生[2, 3]、健康教育[4]等领域，有效推动了卫生健康事业的高速发展。卫生健康信息化建设的发展离不开相关政策的大力支持，当前关于卫生健康信息化建设的政策文本评估主要集中在政策工具[5]和词频分析[6]，尚缺少系统评估。政策建模一致性（Policy Modeling Consistency，以下简称 PMC）指数模型是衡量政策内部一致性和政策内容完整性和优劣性的重要计量工具，广泛应用于卫生健康领域的政策量化评估，包括公立医院改革[7]、公共卫生[8, 9]、中医[10]等领域。本文基于 PMC 指数模型，对国家层面的卫生健康信息化建设政策进行系统分析，为进一步提升卫生健康信息化建设政策质量提供建议。

2.资料与方法

2.1.政策来源

2.1.1 政策文本的检索范围及规则。 本文系统整理从 2010 年以来的国家层面出台的卫生健康信息化政策，时间范围为 2010 年 1 月 1 日-2022 年 12 月 31 日。政策来源主要包括两个渠道：一是北大法宝数据库（https://www.pkulaw.com/），北大法宝数据库是目前比较完善的政策文本数据库之一；二是国务院及国家卫生健康委、工业和信息化部、国家中医药局等官方网站公开发布的文件。用“信息化”“智慧”等关键词在上述两个渠道进行全文检索。经检索，2010 年 1 月 1 日-2022 年 12 月 31 日期间，国家层面累计发布卫生健康信息化政策 33 条。

2.1.2 政策文本的纳入标准。①由国家层面发布的政策，包括中共中央、国务院及国家各部委；②政策发布的时间为 2010 年 1 月 1 日-2022 年 12 月 31 日；③政策应为卫生健康信息化建设的综合性政策，结果通报、信息发布等政策文件不纳入其中。经过筛选，从 33 条政策中剔除 9 条非综合性政策，最终将 24 条政策文本纳入分析（详见表 1）。

表 1 纳入分析的卫生健康信息化建设政策汇总

编号	政策名称	发文部门	发文时间
P1	关于推进以电子病历为核心医院信息化建设试点工作的通知	卫生部办公厅	2011 年
P2	关于做好 2012 年卫生统计和信息化工作的通知	卫生部办公厅	2012 年
P3	关于加强卫生统计与信息化人才队伍建设的意见	卫生部办公厅	2012 年
P4	关于加强卫生信息化建设的指导意见	卫生部、国家中医药管理局	2012 年
P5	关于印发卫生监督信息化建设指导意见（2012 版）的通知	卫生部办公厅	2012 年
P6	关于加快推进人口健康信息化建设的指导意见	国家卫生和计划生育委员会、国家中医药管理局	2013 年
P7	关于加强疾病应急救助工作信息化管理的通知	国家卫生计生委办公厅	2017 年
P8	关于开展智慧健康养老应用试点示范的通知	工业和信息化部办公厅、民政部办公厅、国家卫生计生委办公厅	2017 年
P9	关于印发全国医院信息化建设标准与规范(试行)的通知	国家卫生健康委员会办公厅	2018 年

P10	关于进一步推进以电子病历为核心的医疗机构信息化建设工作的通知	国家卫生健康委员会	2018 年
P11	关于印发全国基层医疗卫生机构信息化建设标准与规范(试行)的通知	国家卫生健康委、国家中医药管理局	2019 年
P12	关于做好疫苗信息化追溯体系建设工作的通知	国家药监局综合司、国家卫生健康委办公厅	2019 年
P13	关于加强信息化支撑新型冠状病毒感染的肺炎疫情防控工作的通知	国家卫生健康委办公厅	2020 年
P14	关于印发《新冠肺炎疫情社区防控工作信息化建设和应用指引》的通知	民政部办公厅等 4 部门	2020 年
P15	关于组织申报《智慧健康养老产品及服务推广目录（2020 年版）》的通知	工业和信息化部办公厅、民政部办公厅、国家卫生健康委办公厅	2020 年
P16	关于做好信息化支撑常态化疫情防控工作的通知	国家卫生健康委办公厅	2020 年
P17	关于进一步加强远程医疗网络能力建设的通知	工业和信息化部办公厅、国家卫生健康委办公厅	2020 年
P18	关于组织开展 5G＋医疗健康应用试点项目申报工作的通知	工业和信息化部办公厅、国家卫生健康委员会办公厅	2020 年
P19	关于印发全国公共卫生信息化建设标准与规范(试行)的通知	国家卫生健康委办公厅、国家中医药局办公室	2020 年
P20	关于印发《智慧健康养老产业发展行动计划(2021—2025 年)》	工业和信息化部、民政部、国家卫生健康委	2021 年
P21	关于开展 2021 年智慧健康养老应用试点示范遴选工作的通知	工业和信息化部办公厅、民政部办公厅、国家卫生健康委员会办公厅	2021 年
P22	关于印发公立医院运营管理信息化功能指引的通知	国家卫生健康委办公厅、国家中医药局办公室	2022 年
P23	关于组织开展 2022 年智慧健康养老产品及服务推广目录申报工作的通知	工业和信息化部办公厅、民政部办公厅、国家卫生健康委办公厅	2022 年
P24	关于印发“十四五”全民健康信息化规划的通知	国家卫生健康委员会、国家中医药管理局、国家疾病预防控制局	2022 年

2.2 卫生健康信息化建设的政策文本分析

本文通过 Nvivo12plus 软件对上述 24 条政策进行文本分析，剔除“提升”“加强”等没有具体指向性的高频词汇，从高频词汇中获取政策文本涉及的主要内容，结合《关于印发全国基层医疗卫生机构信息化建设标准与规范(试行)的通知》关于信息化建设的工作分类，将政策涉及的内容整合成四类。一是服务业务类，涉及“档案”“健康教育”“预防接种”“养老”“诊断”“保健”等高频词；二是管理业务类；涉及“随访”“指导”“个案”“培训”等高频词；三是平台服务类，涉及“采集”“统计”“审核”“查询”“流程”等高频词；四是信息安全类，涉及“安全”“风险”“监控”“权限”等高频词。

表 2 卫生健康信息化建设政策本文主要词频分布

主题词	词频（个）	主题词	词频（个）	主题词	词频（个）	主题词	词频（个）
信息	2886	知识	295	智能	164	协同	123
数据	2755	平台	289	终端	163	老年人	123
卫生	2068	校验	287	人才	162	行为	122

健康	1722	逻辑	262	治疗	160	药物	118
医院	1429	资源	255	监控	159	医学	117
疾病	1245	质量	247	临床	154	报表	117
预防	1178	居民	237	药品	153	规划	115
医疗	909	培训	234	随访	150	中医药	114
统计	792	规范	233	防治	149	权限	114
采集	654	报告	222	调查	148	环境	113
信息化	650	病历	217	应急	143	科研	113
安全	563	评估	216	引证	143	统一	112
监测	535	养老	215	数据库	142	认证	112
系统	520	档案	212	指导	137	检验	111
查询	464	审计	204	互联网	136	支撑	110
监督	456	健康教育	201	流程	135	个人	109
网络	437	产品	199	预算	134	成本	109
审核	421	智慧	184	预警	134	机制	109
设备	377	知识库	176	存储	131	运营	109
患者	375	识别	170	远程	129	预约	108
电子	345	体系	168	处理	128	实验室	105
个案	338	诊断	168	辖区	128	预防接种	105
记录	333	资产	168	病例	127	职业病	103
检测	298	风险	168	企业	125	疫情	101
检查	297	保健	167	诊疗	124	软件	100

图 1 卫生健康信息化建设政策词语云

2.3 PMC 指数模型构建

PMC 指数模型的构建，需综合考虑卫生健康信息化建设政策文本内容及异质性影响因素。本文参考多位学者关于 PMC 指数模型的修订以及卫生健康信息化建设政策的相关研究，结合政策文本，构建我国卫生健康信息化建设政策的 PMC 指数模型体系。卫生健康信息化建设政策 PMC 指数模型选取 9 个一级变量，37 个二级变量。规定一级指标均服从二项分布 N[0,1]，若政策文本满足评判标准，对应的二级指标赋值为 1，否则赋值为 0。因 X8 和 X9 的二级指标在分类上具有互斥性，对其参数赋值作出单独规定，满足条件时赋予对应的值。赋值时由两位研究人员分别进行赋值，并对赋值情况进行对比，赋值情况不一致的经讨论后确定最终赋值。具体评价指标体系详见表 3。

表 3 卫生健康信息化建设政策 PMC 指标体系设置及有关说明

一级变量	编号	二级变量	二级变量评价标准	评价参数	来源或依据
政策结构 X1	X1:1	依据充分	是否有充分的政策制定依据	[0,1]	参考翟运开[11]等文章修改
	X1:2	目标明确	是否有明确的政策目标指向	[0,1]	
	X1:3	方案科学	是否有科学的政策方案设计	[0,1]	
	X1:4	规划详实	是否有详实的政策规划安排	[0,1]	
政策工具*X2	X2:1	需求型	是否使用需求型工具	[0,1]	参考杨玉红[12]等文章修改
	X2:2	供给型	是否使用供给型工具	[0,1]	
	X2:3	环境型	是否使用环境型工具	[0,1]	
政策性质 X3	X3:1	强制	是否具有强制性政策倾向	[0,1]	杨赐然[7]、张志国[13]等文章修改
	X3:2	建议	是否具有建议性政策倾向	[0,1]	
	X3:3	描述	是否具有描述性政策倾向	[0,1]	
	X3:4	预测	是否具有预测性政策倾向	[0,1]	
	X3:5	监管	是否具有监管性政策倾向	[0,1]	

政策时效 X4	X4:1	长期	＞5 年	[0, 1]	根据李祥飞等文章修改[14]
	X4:2	中期	3~5 年	[0, 1]	
	X4:3	短期	1~3 年	[0, 1]	
	X4:4	年度内	1 年以内	[0, 1]	
政策内容 X5	X5:1	服务业务	是否涉及服务业务内容	[0, 1]	根据卫生健康信息化建设政策文本有关内容确定
	X5:2	管理业务	是否涉及管理业务内容	[0, 1]	
	X5:3	平台服务	是否涉及平台服务内容	[0, 1]	
	X5:4	信息安全	是否涉及信息安全内容	[0, 1]	
政策实施 X6	X6:1	强化组织领导	是否具备组织领导保障措施	[0, 1]	根据郭强[15]等文章修改
	X6:2	宣传推广动员	是否具备宣传推广动员措施	[0, 1]	
	X6:3	监督考核评估	是否具备监管考核评估措施	[0, 1]	
政策受体 X7	X7:1	行政部门	是否涉及行政部门	[0, 1]	根据卫生健康信息化建设政策涉及的对象及政策文本有关内容确定
	X7:2	医疗机构	是否涉及医疗机构	[0, 1]	
	X7:3	第三方机构	是否涉及第三方机构	[0, 1]	
	X7:4	普通群众	是否涉及普通群众	[0, 1]	
政策级别 X8	X8:1	规划类	法律、规划、纲要、规定、规范、战略	1	根据成全[16]等文章修改
	X8:2	方案类	方案、细则、指引、指南	0.8	
	X8:3	办法类	办法、措施、条例	0.6	
	X8:4	意见类	意见、决定	0.4	
	X8:5	通知类	通知、公告	0.2	
发布机构 X9	X9:1	中共中央、国务院	中共中央 国务院及其办公厅	1	根据成全[16]等文章修改
	X9:2	多个部委	多个部委联合发文	0.8	
	X9:3	单一部委	单一部委单行发文	0.6	
	X9:4	直属机构	副部级及以下单位行文	0.4	

注：★政策工具使用“Rothwell＆Zegveld”政策工具。

2.4 指数测算公示

按照赋值原则，对各项政策的二级指标的进行赋分。依据公式（1）计算各个一级指标（X1-X7）的平均水平，X8 和 X9 直接赋值；然后依据公式（2）和公式（3）计算 PMC 指数得分与绘制 PMC 曲面图。

（1） $X_i = \sum_{j=1}^{n_j} X_{ij} / n_j$ i=1,2，...7

（2） $PMC = \sum_{j=1}^{4} \frac{X_{1j}}{4} + \sum_{j=1}^{3} \frac{X_{2j}}{3} + \sum_{j=1}^{4} \frac{X_{3j}}{4} + \sum_{j=1}^{4} \frac{X_{4j}}{4} + \sum_{j=1}^{7} \frac{X_{5j}}{7} + \sum_{j=1}^{3} \frac{X_{6j}}{3} + \sum_{j=1}^{4} \frac{X_{7j}}{4} + X8 + X9$

（3） $PMC\_Surface = \begin{pmatrix} X1 & X2 & X3 \\ X4 & X5 & X6 \\ X7 & X8 & X9 \end{pmatrix}$

2.5 PMC 指数质量等级划分

依据 PMC 指数等级划分依据[13]，对所计算的 PMC 指数得分进行质量等级划分。当得分为 8~9 时，该政策的质量等级为完美；在得分为 6~7.99 时，该政策的质量等级为优秀；在得分为 4~5.99 时，该政策的质量等级为良好；当得分低于 4 时，该政策的质量等级为不良。详见 4。

表 4 PMC 指数等级划分

编号	PMC 指数得分	质量等级
1	8~9	完美
2	6~7.99	优秀
3	4~5.99	良好
4	0~3.99	不良

3.结果

从整体来看，24 份卫生健康信息化建设政策的 PMC 指数平均值为 6.41，政策质量总体上处于优秀水平。PMC 曲线图的平缓程度与政策质量呈现正相关，曲面越平缓，则政策质量得分越高，24 项政策平均水平的 PMC 曲线图（图 2）显示，政策级别和政策实施倾斜度较大，说明在政策级别和政策实施方面存在薄弱环节，其余部分较为平稳。

从各政策维度的值来看，政策时效（X4）、政策结构（X1）的平均得分大于 0.8，比较理想。政策级别（X8）、实施主体（X6）的均值比较低，建议改进顺序为：X8-X6-X7-X9-X3-X5-X2-X1-X4。其中得分最高的是政策实效（X4），为 0.85，表明大部分政策覆盖了短期、中期和长期，具有较长的政策效力。政策结构（X1）的得分为 0.84，说明大部分政策文件有充分的依据、明确的目标、政策方案设计科学，有详实的政策规划安排。得分最低的是政策级别（X8），主要表现为通知类和意见类等政策级别低的文件较多；另外，实施主体（X6）的得分也较低，说明卫生健康信息化建设在组织领导、宣传推广、考核评估等方面还需要进一步加强。

从政策评价等级来看，处于完美等级的政策有 1 项，为国家卫生健康委员会、国家中医药管理局、国家疾病预防控制局《关于印发“十四五”全民健康信息化规划的通知》；75.0%的政策处于优秀等级，且在评价过程中没有发现处于不良等级的政策，说明卫生健康信息化建设政策总体分布趋向于完美和优秀。

表 5 卫生健康信息化建设政策 PMC 指数计算情况

	编	X	X	X	X	X	X	X	X	X	P	质
	P	0	0	1	1	0	0	0	0	0	6	优
	P	1	0	1	0	0	1	0	0	0	5	良
	P	1	0	0	1	0	0	0	0	0	5	良
	P	1	0	0	1	1	0	1	0	0	7	优
	P	1	0	0	1	0	1	0	0	0	6	优
	P	1	0	1	1	1	0	0	0	0	6	优
	P	0	0	0	1	0	0	0	0	0	5	良
	P	0	1	0	1	0	0	1	0	0	6	优
	P	1	0	0	1	1	0	0	1	0	6	优
	P	0	0	1	1	1	0	0	0	0	6	优

	P	1	1	0	1	1	0	0	1	0	7	优
	P	1	1	1	1	1	0	0	0	0	7	优
	P	0	0	0	1	0	0	0	0	0	5	良
	P	1	0	0	1	0	0	1	0	0	6	优
	P	0	1	0	0	0	0	0	0	0	5	优
	P	0	0	0	1	0	1	0	0	0	6	优
	P	0	1	0	1	0	0	0	0	0	6	优
	P	0	1	0	1	0	0	0	0	0	6	优
	P	1	0	0	1	1	0	0	1	0	6	优
	P	1	0	0	0	1	1	1	1	0	8	优
	P	0	1	0	0	0	0	0	0	0	5	良
	P	1	1	0	1	1	0	0	0	0	7	优
	P	0	1	0	0	0	0	0	0	0	5	良
	P	1	1	1	0	1	1	1	1	0	8	完
	合	0	0	0	0	0	0	0	0	0	6	优

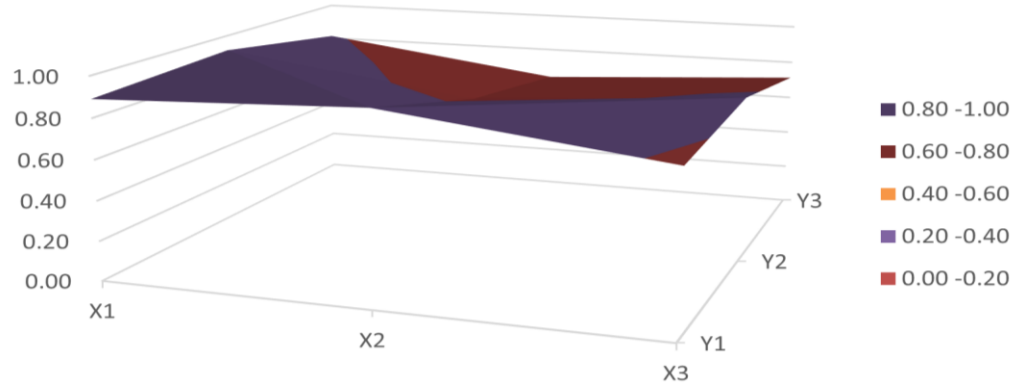


图 2 卫生健康信息化建设政策整体的 PMC 指数曲线图

4.讨论

4.1 卫生健康信息化建设政策整体质量优秀

分析结果显示，国家层面出台的卫生健康信息化建设综合政策总体得分为 6.41，处于优秀水平，但低于医保[14]、公立医院[7]的政策评价得分。质量等级为优秀及以上的政策文件占比 70.8%，且未出现质量等级为不良等级的政策文件。国家层面出台的卫生健康信息化建设政策的文本总体上具有依据充分、目标明确、设计科学、规划详实的特点，具有完整的结构；另外，政策时效长，政策工具应用全面，能够指导当前卫生健康信息化建设工作。

4.2 政策推动力度不足

分析结果显示，政策实施的总体得分为 0.53，具有宣传推广措施的政策占比 37.5%，具有监督考核评估措施的政策占比 54.2%，具有强化组织领导措施的政策占比 66.7%；政策等级得分为 0.45 分，健康卫生信息化建设政策的政策级别主要为效力较低的通知类，规划类、纲要类等政策级别较高的政策文件占比较低。从发文主体来看，主要为工信部、国家卫生健康委牵头发文，缺少从中共中央 国务院层面下发的文件。从政策实施、政策等级和发文主体来看，当前卫生健康信息化建设的推动力度不足，与陈慧等的观点一致[5]。下一步应在加强顶层设计，强化宣传推广，深化监督考核等方面加大工作力度，进一步推动卫生健康信息化建设各项工作落地落实。

4.3 政策受体的分布不均衡

分析结果显示，全部政策均涉及行政部门及医疗机构，50%的政策涉及第三方机构，仅有 25%的政策涉及普通群众。当前智慧医疗、数字健康等卫生信息化成果未能广泛应用，居民共建共享程度低[6]。卫生健康信息化建设最终的目的是惠及全体人民群众，政策制定时，要充分考虑人民群众的需求，引导人民群众参与其中，维护人民群众的健康权益。可积极引导第三方机构积极参与

卫生健康信息化建设，构建政府-医院-第三方-普通群众四位一体的卫生健康信息化体系，更好服务于卫生健康工作大局。

5.结语

政策评价需要考虑的维度很多，本文研究基于国家层面出台的卫生健康信息化建设相关政策的文本分析，缺乏对政策实际执行过程及执行效果的评价。下一步将在政策文本研究的基础上，收集政策执行相关数据，对政策执行、政策效果开展更加全面、系统和深入的分析，以获得更加合理的结论和结果。

PU-079

国家医疗保障局公开通报欺诈骗保典型案件分析及防范对策探讨

王宇彤[1] 谢忱[1] 朱卫国[1,2] 孙学勤[1] 董蕊[1] 韩丁[3]
[1]北京协和医院医疗保险管理处
[2]中国医学科学院北京协和医院全科医学科（普通内科）
[3]中国医学科学院北京协和医院

1 引言

国家医保局曝光台[1]由国家医保局设立，具有公开性、及时性、真实性、公共利益性和影响力等属性，公众可以通过电话、邮件、网站等渠道向国家医保局曝光台举报问题，国家医保局联合相关部门进行调查、惩处并曝光。国家医保局在“十四五”时期就如何加强医保基金监管中着重提出“健全完善要情报告制度，用好基金监管曝光台”[2]。曝光台的设立旨在达到四个目的：一是惩处违规者；二是维护医保制度的公正和透明；三是号召公众参与，激发社会关注；四是保护患者权益，使患者能够获得公正、安全、高质量的医疗服务。本文通过整理并分析曝光台最近 2 年公布的典型案件，归纳违规违法行为的主要特点，从预防类似案件再次发生的角度探究解决问题的途径，增强医保基金监管力度，维护医保基金安全。

2 数据来源与方法

2.1 资料来源

收集 2021 年 7 月 1 日至 2023 年 6 月 30 日期间，国家医保局曝光台曝光的 10 批典型案件，共 112 例。涉及全国 31 个省级行政区。

2.2 研究方法

对 112 例案件进行整理，提取案件一般特征，包括案件发生时间、省份、责任主体、违法违规具体行为和涉案金额等内容。归类总结违法违规行为的责任主体、问题类型、发生环节，以频数和构成比描述案件的基本情况，探讨分析欺诈骗保防范对策。

3 结果

3.1 案件基本情况描述

2021 年 7 月 1 日至 2023 年 6 月 30 日，国家医保局曝光台共曝光典型案件 112 例。从地域分布看，江苏 8 例，浙江、福建、江西各 6 例，吉林、四川、贵州、西藏各 5 例，天津、内蒙、黑龙江、安徽、河南、湖南、广东、广西、重庆、云南、新疆各 4 例，山东 3 例，北京、河北、山西、辽宁、宁夏、海南各 2 例，湖北、陕西、甘肃各 1 例。案件发生时间范围为 2018 年至 2022 年，主要集中在 2021 年（64 例，占 57.14%）。从责任主体看，可分为三类，一是医疗机构（45 例，占 40.18%），二是参保人或参保单位（37 例，占 33.04%），三是定点零售药店（30 例，占 26.79%）。涉案医保基金共计 117,569,727.91 元，按责任主体分类统计，医疗机构涉案金额最高（101,002,756.02 元，占 85.91%），参保人/单位和定点零售药店涉案金额分别为 9,116,173.41 元和 7,450,798.48 元，合计占比 14.09%，见图 1。

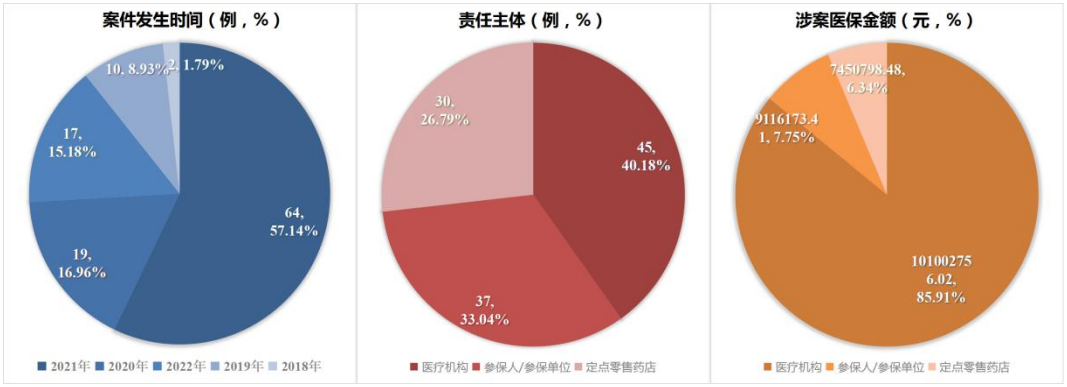


图 2 案件基本情况描述

3.2 违法违规行为方式统计

在 112 例案件中，各责任主体的违法违规行为方式具有相似性，可总结归纳为八大类，包括以串换药品或诊疗项目、重复收费为主的违规行为，统称违规医疗或销售行为；伪造用于医保报销的各种相关材料；滞留参保人医保/社保卡套现；冒用他人医保卡就医购药；药耗进销存不符；虚构参保人劳动关系；参保人就医购药流程中的监督管理问题；诱导参保人消费或治疗等。某一个案件中存在的问题可能多于一种，例如一典型案件曝光了某医院医保目录内药品维护不及时和将医保目录外药品串换成目录内药品等问题，则该医院的违规行为涉及“监管不到位”和“违规医疗和销售行为”两类。按照此方式进行统计，112 例案件中违法违规行为合计 132 频次。其中，以违规医疗或销售行为占比最高（48 次，36.36%），其次为伪造相关材料（31 次，23.48%）和滞留参保人医保/社保卡套现（17 次，12.88%），见表 1。各责任主体违法问题构成情况见图 2。

对于参保人/参保单位来说，违法违规行为以三类为主，一是伪造病历、住院、票据等虚假材料（20 次，54.05%），主要发生在报销环节，伪造的材料常用于办理异地就医、门诊慢特病报销；另有职工身份和城乡居民身份重复报销，以及发生交通事故谎称骑车摔伤等情况发生。二是冒用他人医保卡挂号并接受治疗开药等（9 次，24.32%），主要发生在就医购药等环节，常见方式为冒用他人医保卡开药、挂号、住院等；另有持卡人死亡后医保卡仍发生费用的情况产生。三是虚构劳动关系，主要发生在生育险中，包括参保单位虚构职工劳动关系和参保人伪造出生证明两种情况（7 次，18.92%）。

对于医疗机构来说，主要以《医疗保障基金使用监督管理条例》第三十八条中提及的串换收费、重复收费、超标准收费等违规医疗行为为主（35 次，66.04%），同时存在伪造或销毁医学文书、虚构医药服务项目、诱导患者消费或治疗等违法行为。此外，因医疗机构日常监督管理不当，或对患者审核不严格而造成冒用医保卡就医、未按照规定保管治疗检查记录药品和医用耗材出入库资料、检验报告单雷同等行为，除了会对参保人进行处罚外，医院也会承担相应的管理责任。

对于定点零售药店来说，主要发生在药品销售环节，违法违规行为以三类为主，一是医保/社保卡套现（15 次，35.71%），包括利用参保人空刷医保卡套取医保基金、留置参保人医保/社保卡等问题，同时伴随伪造虚假购药单据等行为。二是以串换药品为主的违规医疗或销售行为（13 次，30.95%），包括为非定点药店提供医保费用结算，协助非特门定点药店实行特门补偿。三是药品进销存不符（7 次，16.67%）。此外还有诱导消费，以及医保目录内药品维护不及时等监管问题的发生。

表 1 违规违法行为方式情况

违规违法行为方式	频次	构成比（%）
违规医疗或销售行为	48	36.36
伪造相关材料	31	23.48
滞留医保/社保卡套现	17	12.88
冒用他人医保/社保卡就医购药	11	8.33
进销存不符	10	7.58
虚构劳动关系	7	5.30
监管问题	4	3.03
诱导消费或治疗	4	3.03

合计	132	100.00

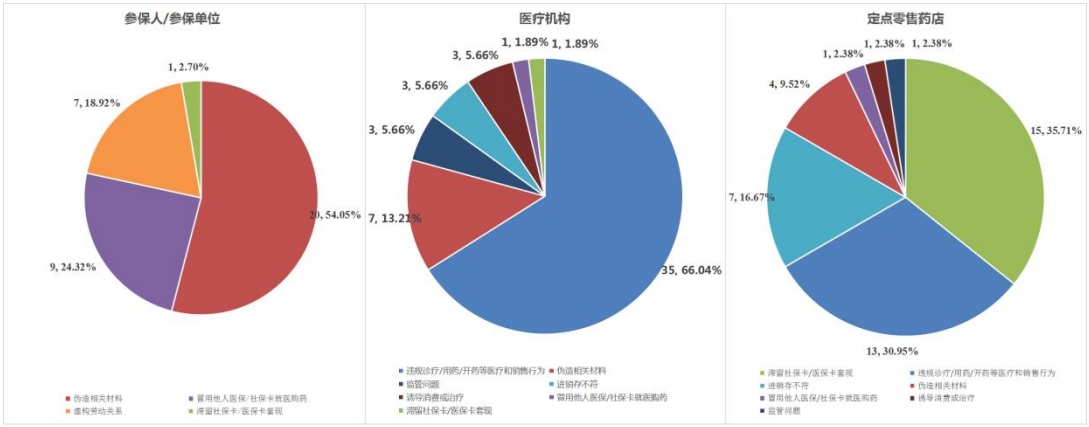


图 3 各责任主体违规违法行为构成情况

4 欺诈骗保防范对策探讨

通过案件梳理，理清当下欺诈骗保突出问题，例如“冒用他人医保/社保卡”、“伪造相关材料”、“药耗进销存不符”等，从基金监管的角度，分析探索避免此类违规违法行为发生的方法：推进医保卡实名使用制度，推进异地医保直结信息建设，推进违法违规案件源头治理。

4.1 推进医保卡实名使用制度

避免他人冒用医保卡是一个长期性的工作，需要各级医疗卫生机构、医保主管部门、公安机关等执法部门多方合作。可从五个方面建立完善相关制度，同时借助流程优化和技术支撑，真正逐步实现“医保实名制”：①医保卡身份验证核查制度，医院可以通过读卡器等设备，对患者的医保卡进行验证和身份核查，确保医保卡的使用者和持卡人相符[3]；②医保卡使用登记制度，对患者的医保卡使用情况进行记录和管理，发现异常情况及时采取相应措施。③医保卡使用培训制度，加强对医师医保卡合理合规使用的培训，提高医护人员的职业道德和法律意识。④欺诈骗保举报制度，鼓励患者和医护人员积极举报医保卡的冒用等违规行为，加强对医保卡的监督和管理。⑤医保电子凭证推广使用制度，医院应当设立专门的技术支持团队，建立完善的技术支撑体系，及时解决患者和医务人员在使用电子凭证过程中遇到的问题；提供便利的使用和患者服务流程，例如通过医院APP、微信公众号或者支付宝等平台进行电子凭证的扫码验证，简化患者的操作流程。

4.2 推进异地医保直结信息建设

通过总结归纳，异地就医报销是典型薄弱环节，为很多不法分子欺诈骗保提供可乘之机，推动异地就医购药、结算报销等方面的信息化建设，不仅是便民惠民的有力举措，更是从信息的采集、标准、传输等角度建立数据安全保护机制，加强对数据的监督和管理。全国跨省异地就医直接结算规模由 2018 年的 131.8 万人次增长到 2022 年的 3812.35 万人次，增长 28 倍[4]，在住院、门诊及门诊慢特病、备案服务等方面持续打通异地就医瓶颈堵点。但仍存在政策不统一、信息沟通渠道不通畅等困境[5]，同时成为欺诈骗保的风险点。思考切实可行的举措，例如推进异地门诊慢特病直接结算信息建设、医保移动支付信息建设和医保内药品决策支持系统建设等，从源头环节杜绝因参保人身份认定不准确、伪造报销相关材料、非医保内药品医保基金结算等问题的发生。

4.3 推进违法违规案件源头治理

本研究分析的 112 例案件中，医疗机构平均涉案金额为 2,244,505.69 元，显著高于参保人/参保单位（246,383.07 元）以及定点零售药店（248,359.95 元）。合理追回违规金额不仅是执法部门按照规章制度执行的法律手段，院方作为医生医保基金使用行为的直接管理者，更应承担起判定医院内部责任主体的职责。医保违规问题可由多种因素导致，包括因政策研读不到位、故意骗保等主观因素和政策性因素等非主观因素，若采用统一的处理方式，例如全部由医院承担或全部由科室承担则有失偏颇，因此亟需提出一种相对来说较为公平公正且普适的机制，让每笔违规金额有源可溯，有理可依，合理追回。追回违规金额是一主要目标，但单纯从金额维度进行处罚，可能会造成小金额违规事件、人数较多科室违规事件处罚力度不强等多种问题，不能起到警示科室或个人的作用。因此亟需对违规问题进行责任评定，评议造成违规问题的因素构成，明确责任比例，进而采取通报、警告、扣罚相应绩效等措施。从医院管理的角度，院方可以两个维度预防违规违法行为的发

生。以本研究中归纳的问题为基础，自查医院内部是否有同样风险点和类似隐患，实现违规违法行为的就事论事对症处理，以及举一反三全面整改。在此基础上，从思想认识、政策理解、制度制定、流程设计、信息支撑、字典维护、人员培训、监管落实和绩效配套 9 个环节深入分析违法违规行为产生的根本原因，以案明纪，以案促改。

5 结论

综合分析归纳近两年案件，违规违法行为方式较为集中。医保信息化、数字化建设以及数据治理是遏制医保基金违法违规使用，打击欺诈骗保行为的重要手段[6]。医疗机构应提高站位，承担起建章立制和源头治理的主要责任，贯彻落实全国“两会”精神和《政府工作报告》的部署要求，全领域、全链条、全覆盖的开展医药购销领域专项整治工作，提升公立医院高质量发展新效能，助力健康中国、数字中国建设。

PU-080

基于 SWOT 方法分析中医药在防治传染病中的发展契机——以新冠疫情防治为例

薛丹阳[1] 于琦[1] 岳玉婷[2]
1. 山西医科大学
2. 英国利兹大学

1 引 言

1.1 研究背景与研究意义

自古以来，传染病的防治都是政府的难题，因为它不仅会对公众健康造成极大威胁，而且给现有的卫生政策和服务体系带来了极大的挑战。尤其是在 21 世纪，传染病的肆虐挑战现代医学技术，严重影响人类的生命生活质量。因此，如何有效防治传染病成为当今公共卫生领域研究的重点。中国历史上曾多次发生过重大传染病疫情，如：明末时期的鼠疫、东汉时期的天花疾病等，中医在疫情防治工作中充分发挥辨证施治的科学性优势、攻补并施的兼容性优势。面对多种多样传染病，如果单靠西医治疗效果甚微，而中医由于其灵活度高、副作用小的特点越来越受到传染病防治领域专家的重视，比如我国伟大的科学家屠呦呦的青蒿素治疗疟疾。近些年来，中医药在防治新发传染病方面快速高效的优势逐渐凸显。

新型冠状病毒感染肺炎疫情的大爆发，引起国家高度重视，疫情不仅给人民生命生活质量带来极大危害，同时也改变了传统医护工作的方式，而中医作为中华文化的瑰宝在抗击疫情的过程中扮演着至关重要的角色，这一点得到了广大医务工作者和众多患者的认证，比如：经国家中医药重点研究室认证，中成药连花清瘟对于新冠的治疗有显著作用，可改善普通型患者发热、咳嗽、乏力、气促等症状。

故为了探究中医药在防治传染病中的作用，本文以中医药参与新冠肺炎疫情防控救治为例，梳理不同时期中央出台的中医药抗疫相关政策文本，使用 SWOT 分析方法和政策文本分析方法，探讨中医药在防治传染病中的作用、优势劣势、发展机遇和威胁。同时也为今后能更好地开展中医药抗疫工作提出建议和展望。

由于中医药在新冠疫情防治中表现出无可替代的特色优势，以及国际上认可度的提高，呼吁我们在传染病防治过程中将眼光放到中医药上来。尤其是在面对不断涌现的突发公共卫生事件时，我们必须发挥中医药在疫情防治中的优势特长，积极推进中医药的传承与创新发展，以应对新的挑战和机遇。中医药防治手段在此新冠疫情中彰显了无可替代的优越性，迎来了高光时刻，在国际上的认可度也逐渐提高，但是在新的发展机遇下，中医药将面临更多挑战。在面对不断涌现的突发公共卫生事件时，我们必须持续发挥中医药在应急救治管理中的预防控制作用，积极推进中医药的传承创新发展，以应对新的挑战和机遇。

1.2 国内外相关研究综述

我国对于中医药防治传染病的研究：传染病是属于中医“瘟病”的范畴，中医认为疾病的发生发展是正邪斗争的结果，战胜邪气扶持正气是战胜传染病的关键因素，如果正气不敌邪气，就可能导致病毒入侵机体，造成传染病的爆发与流行，因此降低邪气扶持正气则是防止传染病的关键。中医采取“辩证”施治，充分发挥“湿毒疫邪”“治未病”思想，在抗击新冠肺炎疫情中总结出的“三药三方”（清肺排毒汤、化湿败毒方、宣肺败毒方和金花清感颗粒、连花清瘟胶囊、血必净注射液）对于改善患者临床症状，促进患者恢复等方面具有明确疗效，充分彰显了中医药在防治传染病中的独特优势。

国外对于中医药防治传染病的研究：近年来，随着新冠疫情在全球肆虐，对各国社会经济发展带来巨大打击，中国积极向世界各国分享救治经验，西方国家的一些医药机构已经开始对中医药进行研究，医药企业购买中医药材进行成分检测，连花清瘟胶囊已经在多个国家和地区获得上市许可，例如：加拿大，巴西等国，四季茶药方具有预防流感、咽炎等功效，也获得国际的广泛认可，经过美方医疗专家机构检测鉴定，宣肺败毒颗粒获得已经获得美国的第二期临床批件，津巴布韦将中医药运用于临床实践，高度赞赏中医药在抗击新冠肺炎疫情的独特作用。

1.3 研究方法

1.3.1 SWOT 分析方法

SWOT 分析方法常用于战略规划，用于评估一个组织、项目或者政策等。SWOT 分别是优势（strength）、劣势（weakness)、机会(opportunity)、外部的威胁与挑战(threats)的首字母缩写。运用 SWOT 方法分析梳理中医药在防治传染病工作中的优势、不足、机遇、外部的威胁与挑战，继续发扬中医药防疫优势，改善不足，以期谋求更好的发展，实现中医药自身传承创新发展的同时为人类健康福祉作出贡献。

1.3.2 文献政策分析方法

政策分析方法一般包含两个方面，一方面是指政策制定分析过程中每个步骤（问题界定—目标确立—方案搜寻—结果预测—方案比较—效果评估）中可以具体应用的方法；另一方面则是指具有理论指导性且较为抽象化的方法，主要有系统分析方法、计量经济学方法、创造性思维的方法等。

2 结 果

2.1 优势分析 (Strength)

中华文明源远流长，中医作为其杰出的传统文化结晶，已有两千载光辉历程。中医在预防和治疗传染病方面积累了丰富的经验和独特的治疗方法，经过长期的实践和积累，已成为一种不可或缺的治疗手段。尤其是随着科学技术水平的不断提高和人类疾病谱的变化，对传染病发病规律、流行趋势及其预防措施进行系统研究成为当务之急。在公元前四世纪的《黄帝内经》中，就有对于预防和治疗传染病的探讨。自古以来，中医药一直被视为中华民族应对传染病威胁的一项至关重要的手段，为保障人民身体健康发挥着不可或缺的作用。自秦汉时期起，众多医学专家便已经开始致力于预防和治疗传染病，为保障人类的健康和福祉作出了巨大贡献。随着社会经济发展和科学技术进步，人类疾病谱也随之变化，由单一病原体导致疾病向多病原体混合感染转变。据不完全统计西汉以来我国共发生大型传染病疫情至少 300 余次，中医先贤总结了许多治疗疫病经典理论，浓缩了中华民族抗击传染病疫情六经辨证、卫气营血辨证和三焦辨证思想的经验和智慧，为后世医务工作者留下不可估量的参考经验，对传染病的防治和人民健康的保障做出了巨大的贡献。

进入现代以来，众多科研机构及专家学者对中医药进行了多方位试验研究，发现大部分中药药物所含活性成分皆具有抗病毒、消炎和增强免疫力的功效，可协助治疗和缓解患者的不适症状，且相关临床试验表明，相较于西医用药，中医药物的治疗毒副作用较小、对药物依赖性小；中医非药物疗法不仅包括传统药物治疗，还涵盖了导引按跷、针、灸、刮痧、按摩、点穴等多种特色疗法，

这些疗法原理深入，对于一些疑难杂症具有显著的治疗效果，同时也可用于养生和保健，对于疾病的预防和治疗也有着显著的作用。

健康是全球人民的共同诉求，随着全球经济一体化进程加快，面对重大传染病疫情任何国家和人民都不能独善其身。中医药注重整体观念，把人体看成是不可分割的整体,同时注重内外协调平衡，以期实现人体与自然的共生共存，中方积极为海外抗疫贡献中国力量，构建人类命运共同体。

中医理论认为，新冠病毒属于一种寒湿邪气，容易消耗人体邪气，导致身体虚弱无力，因此常常需要运用针灸、推拿等手段来消除湿气。连花清瘟注射液主要成分为连花和连翘，具有清热解毒、化湿化痰、疏散外邪等功效。连花清瘟中还使用了麻杏石甘汤，可以应用于清瘟解毒、宣肺泄热等治疗领域，因此有不少专家认为其对新冠肺炎的治疗效果显著，同时，莲花清瘟在治疗新型冠状病毒方面也表现出良好的疗效，这一点经过临床实践得到了证实。近年来研究发现连花清瘟具有抗炎作用和免疫调节作用，且无明显不良反应，可用于预防及治疗新冠病毒感染引起的呼吸系统疾病。因此，该药品已被《新型冠状病毒肺炎诊疗方案》（试行第四/五/六/七/八版）以及 20 余个省市的新冠肺炎诊疗方案所推荐，成为备受推崇的治疗方案之一。

早在新冠疫情爆发初期，张伯礼院士和其他中医药专家就将他们使用中医药防治新冠疫情的临床经验与《新型冠状病毒感染的肺炎诊疗方案（试行）》第三版至第七版相融合，并在全国范围内推广实施。此后，该系列方案相继通过多个国家级文件予以推广实施。国家卫生健康委和国家中医药管理局于 2020 年 2 月 6 日联合下发了《关于推荐在中西医结合救治新型冠状病毒感染的肺炎中使用“清肺排毒汤”》的通知。国家卫生健康委与国家中医药管理局于 2020 年 2 月 12 日共同发布《关于在新型冠状病毒肺炎等传染病防治中建立健全中西医协作机制的通知》。这一系列政策的出台充分表明了国家对中医药参与防治新冠疫情的高度重视，同时这也是带有中国特色的传染病控制处理方案。

2.2 劣势分析（Weakness）

马克思辩证主义思想告诉我们看待事物要用一分为二的观点，中医诊法的确具有不可替代的优越性，这是毋庸置疑的；但是中医在临床实践中的弊端也是不可否认的，所以我们必须辩证地看待中医药对于防止传染病的的作用。

首先，中医是古代传统观念和实践经验总结的产物，中医疗法和药物的疗效缺乏科学证据的支持，这使得它们的可靠性和有效性受到了质疑，大多都没有通过现代科学实验方法、临床试验检测进行评估；其次，中医文献所记载的内容存在一定的局限性：主要基于回顾性研究，可能涉及到回忆和选择的偏倚；各项研究使用的中医药组方、剂量等没有固定标准，可能存在信息偏倚；纳入研究的患者病程、体质、年龄、基础疾病史等情况不同，样本很难统一，控制变量存在难度，可能会造成研究结果与真实情况存在差异。

中医以“望、闻、问、切”为主要诊疗手段，因此，与西医相比，中药的诊断具有主观性，对于新型冠状病毒感染后的基本临床表征，由于个人体质因素的差异，对于用药的剂量和中草药类型的严格把控显得尤为重要，因为疾病诊断结果可能存在主观性，仅凭医生的主观经验判断可能会导致误诊。此外，中药的复方制剂中含有多种中草药成分，其中一些草药需要严格遵守剂量控制标准，而一些具有潜在毒副作用的药物可能会与现代药物发生不良反应。因此，在疾病诊疗过程中，必须遵循辨证施治原则，根据患者具体情况采取个体化方案，中西医结合用药也许科学考证。虽然中医治疗已被实践证明具有显著疗效，但对于新发传染病的中医药物选择和用量，仍存在一定的挑战，使用中医药时可能存在潜在的安全风险，因此需要进行科学的论证和临床实践。

此外，由于中医在汤药熬制过程中所需的繁琐时间和携带不便，导致许多传统中成药的疗效不及西医，逐渐失去了其“便捷、经济、高效”的优势。中医药企业的药材来源大多数以野生药材为主，受环境气候等因素影响，中药材质量、产量很难保证。随着现代科学技术的飞速发展和互联网信息技术的普及，传统中医药已经开始与现代科学相结合，并不断地得到继承和发扬。由于中医的治疗过程具有私密性和主观性，传统的传承方式是以师为师，因此中医诊断技术的传承和发展面临着巨大的挑战，很难成为公共学习资源，也难以培养大量的中医人才。此外，越来越多的年轻人不愿意花费时间重新学习中医知识，甚至有些宝贵的临床实践经验也因此消失殆尽。

2.4 机遇分析 (Opportunity)

2.4.1 国家对中医药的政策支持（以本次新冠肺炎防治为例）

随着中医药在现代疾病治疗过程中越来越显著的优势，中医药事业的发展工作受到党和国家高度重视，因此出台了一系列支持中医药发展的政策，如：《健康中国 2030 规划纲要》《国民经济和社会发展第十四个五年规划纲要》、《关于扶持和促进中医药事业发展的若干意见》、《中医药创新发展纲要(2006—2020)》、《中华人民共和国中医药法》的颁布、中共中央和国务院发布的第一个中医药文件《关于促进中医药传承创新发展的意见》，使得我国中医药事业的平稳发展得到了有力的保障，同时也为我国中医药事业发展带来全新机遇。

在新型冠状病毒肺炎疫情爆发后，中医药医护人员的参与程度和规模达到了前所未有的高度，他们毫不犹豫地投身于救援行动之中。在这场没有硝烟的战斗中，广大医务工作者充分发挥自身专业优势和特长，发挥了重要作用。在抗击疫情的过程中，中西医学紧密合作，积极深入参与，为国际抗疫事业提供有力支持，同时也提高人民对中医药的认可度。新冠肺炎患者多是老年人和慢性病人，他们往往体质虚弱、抵抗力差，需要更长时间才能恢复健康。急症以西医为主，慢性病则以中医为主。新冠肺炎疫情期间，全国各地都涌现出一批优秀中医药工作者。中医药曾被贴上“缓慢”的标签，但在汹涌澎湃的疫情面前，它正在从幕后走向前台，从“慢郎中”蜕变“前线战士”。在 2020 年 1 月 8 日，北京地坛医院接收了一位患者，该患者表现出发热、头痛、乏力等症状。经过疾控中心的检测和专家讨论，患者被诊断为新型冠状病毒感染肺炎。随后，2020 年 1 月 13 日，该患者被转移至北京地坛医院接受中医药治疗。经临床观察，该例患者病情发展迅速，高热不退，咳嗽咳痰，胸闷气短，全身不适，肺部听诊呈实变浊音。在北京新型冠状病毒感染肺炎的治疗中，中医专家采用了辩证清肺宣热疗法，经过中医诊治，症状逐渐缓解，体温迅速恢复正常，核酸检测结果为阴性，符合国家规定的出院标准。这是首例主要采用中医药治疗的患者，中医药的辨证施治的优势能够最大限度缩短病程、减少并发症，同时还能提高机体免疫力，促进康复。随着新型冠状病毒感染的肺炎疫情的蔓延，国家中医药管理局迅速拟定了工作方案，并成立了肺炎疫情防控领导小组和专家组，以“湿、热、毒、淤”为病因特征，拟订了中医治疗方案。

该方案于年份 2020 年 1 月 22 日正式纳入国家卫生健康委员会和国家中医药管理局联合发布的《新型冠状病毒感染的肺炎诊疗方案（试行第三版）》，将该疾病分为湿邪郁肺、邪热壅肺、邪毒闭肺、内闭外脱等 4 个中医证型，并为每个证型提供了推荐的方药。目前对于新冠肺炎的研究已经取得一定成果，但仍有很多问题有待解决。根据国家卫生健康委员会发布的《新型冠状病毒感染的肺炎诊疗方案（试行第五版）》方案，该疾病属于中医疫病范畴，其病因在于感受疫戾之气，病位在肺，其基本病机特点为湿、热、毒、瘀”；临床主要表现为卫气营血辨证及三焦辨证两种分型。根据不同地区的病情、气候特征和体质等因素，可以采用推荐的方案进行辨证施治。

全国各地因地制宜制定中医药诊疗方案，很多地区着重对高危人群进行中药预防，可有效降低感染率，实现一级预防，结合当地实际情况纷纷推出抗新冠肺炎处方，且不断传来佳讯。山西省卫生健康委严格贯彻落实国家和省委、省政府决策部署，先后组织印发《山西省新型冠状病毒肺炎中医药防治方案》等，指导和督促各级、各中医医疗机构全面做好医疗资源储备，提升救治能力，作为新冠肺炎中医药专项应急救治三个试点省份之一，山西成立了省、市中医专家组，致力于挖掘名老中医的潜力，建立了中医药应对机制，充分发挥中医药的优势，并将中西医紧密结合，共同发挥协同作用，以应对疫情，前期研制并获批的益气除瘟颗粒、除湿清肺颗粒、解毒护肺颗粒、葶苈泻肺颗粒、补肺健脾颗粒五种制剂在临床应用效果良好,有效降低了发病率和病亡率。此外还有甘肃的岐黄避瘟颗粒、河南的金蒡清疫颗粒、广东的透解祛瘟颗粒等，这些中药和方剂“简、便、验、廉”，构筑起一道道防疫屏障。实践证明，中医药历久弥新，“好使、管用”是百姓对中医药最朴素的评价，中医药参与疫情防控已经取得了阶段性进展。在总结中医诊治新冠肺炎康复经验基础上，国务院应对新型冠状病毒肺炎疫情联防联控机制医疗救治组组织专家制定了《新型冠状病毒肺炎恢复期中医康复指导建议(试行)》。

中医药防治手段在此次新冠疫情重大公共卫生事件中的实践证明，中医药在防治传染病方面是有极大优势的，中医药在防治常见病、多发病、慢性病、重大疾病和新发传染病方面是有独特优势和价值作用的。

2.4.2 人民对于中医疗法的认知度提高

由于中医药在疫情防控中的优异表现，极大改变了大众对于传统中医药的刻板印象。自西方现代医学传入我国之后，中医受到了前所未有的威胁，疗效慢、治疗时间长等成为了中医疗法的标签，长期处于西医的辅助地位，然而中医在新型冠状病毒肺炎中发挥的不可磨灭的作用，让人们重新认识到了中医药，提高了国民认可度。

人类命运共同体，疾病面前任何国家和人民都不能独善其身。跨越国界，跨越种族的疫情肆虐，唯有人类共同奋斗，才能在这场人与病毒的“战争”中获得胜利。中国从来秉持人类命运共同体观念，无私的向世界贡献中医药力量，已同数十个国家进行中西医结合治疗抗疫经验分享，为全球新冠疫情的防治交上了“中国答卷”，有效提升中医在全球范围的知名度与认可度，吸引大批国外学者对中医药在防治传染疾病的兴趣。

2.5 威胁分析（Threats）

党的十八大以来，以习近平同志为核心的党中央高度重视中医药的发展，中共中央、国务院出台了《关于促进中医药传承创新发展的意见》，中医药在新冠疫情防控中发挥了至关重要的作用，其传承创新和发展呈现出蓬勃生机，迎来了具有历史性意义的发展战略机遇期。尽管中医药在突发急性传染病中的防治意义重大，但其仍未被纳入国家传染病应急体系，并且现有中医药在应对突发传染病方面的能力和科研水平仍有待提高。目前，我国对重大突发公共卫生事件采取了“预防为主”的方针，在疾病预防控制中发挥着重要作用。而中医药在突发传染病防治中的参与不仅受到科学上的质疑，而且在政策的具体实施和执行过程中也遭遇了各种阻碍和困难。

当前，中医药发展仍然存在一些亟需解决的突出问题：

一是中医参与传染病临床诊疗没有规范性机制；在公共卫生应急管理中对于中医药主体的相关规定还很缺乏。在抗击新冠肺炎疫情保卫战中，国家层面而言，《中华人民共和国传染病防治法》《中华人民共和国突发事件应对法》等法律和《突发公共卫生事件应急条例》《国家突发公共卫生事件应急预案》等制度中未明确提到中西医结合相关的具体内容，各级政府在应对突发公共卫生事件中也没有关于中西医结合治疗的相关政策文件。在当前西医主导的情境下，中医药及时介入控制传染病的时间和参与度都受到影响，比如，武汉等省份出现中医药介入防止传染病不及时的现象。事实上，“中医不是慢郎中。防治急性传染病，可以有效缓解西医药物的激素作用，中医药具有一定疗效，应当参与治疗越早越好。

二是中医药在预防和治疗突发传染病服务体系尚未健全，应对传染病和突发公共卫生事件应急服务水平还需提高，防疫临床网络不够全面，临床试验基地，数据网络平台以及中医药的学科建设和资金投入不足，因此，中医药的防疫能力和水平还难以达到及时应对突发公共卫生事件传染病当中。

三是对于出台的中医抗疫的相关政策性文件，应当更加灵活，因地制宜，因人而异，不能一概而论。国家卫生健康委员会办公厅、国家中医药管理局办公室相继发布《关于进一步做好新型冠状病毒感染的肺炎中医药防控工作的通知》《关于进一步做好新型冠状病毒感染的肺炎中西医结合救治工作的通知》《关于推荐在中西医结合救治通知》《关于推荐在中西医结合救治新型冠状病毒感染的肺炎中使用“清肺排毒汤”的通知》《关于加强信息化支撑新型冠状病毒肺炎疫情中医药防控工作的通知》《关于在新型冠状病毒肺炎等传染病防治工作中建立健全中西医协作机制的通知》，这些政策的颁布，为预防和控制传染病提供了核心的政策保障，但需要注意的是，政策性文件在专家论证研讨和公众参与方面相对宽松，与法律法规相比存在差异，因此，在政策制定可能缺乏一些合理性，刚性过强。例如《关于在新型冠状病毒感染的肺炎中医药治疗中推荐使用中药协定方的通知》该政策文件中指出确保所有患者服用中药，但是未考虑到患者的个体体质差异性，灵活性和可及性有待提高。

3 结论与建议

3.1 突发公共卫生事件处置机制中应当更加重视中医作用

中医在疫情防控救治工作中处于边缘化地位，无法在第一时间介入究其原因是由于我国传染病应急响应机制还不健全，还未充分考虑中医对于传染病应急救治工作的突出优势。应当加强中西医优势互补，健全中西医联防联控处置机制，加强中西医信息资源共享，高效整合疫情信息资源，发挥中西医各自职能优势，提高突发传染病救治效率，保证中医全程参与救治，使得中医医师可以在传染病发现时第一时间对病人“望闻问切”，摸清疾病的发病机制并制定诊疗方案，搭建中医药物资源统一调度平台。明确各级中医药部门在传染病防控工作中的权责义务，对于相关职能部门参与传染病救治设置奖惩制度，统筹协调各级部门的资源调配、财款拨放、防疫研究等职能。

3.2 应做好中医药防控传染疾病的宣传工作

大力宣传中医药文化，提高民众中医知识普及率。国家中医药管理部门可通过社交网络平台宣传中医保健知识，例如：发布保健操八段锦教学视频、养生饮食党参鹌鹑汤、四神汤等食疗补品的制作方法，以视频教学等通俗易懂、简便易学的方式，在宣传中医文化的同时，也可以帮助民众改善自身免疫力，提高疾病抵御能力。相应完善中医药救治防控网络数据平台，坚持预防为主，中医药可与流行病学、卫生统计学等学科开展合作，构建传染病预警大数据平台，加强各卫生主管部门之间的信息沟通，提高传染病预防、控制的运作效率。

3.3 加强中医院校应急防治的学科体系建设并与中医院开展合作演练

加强中医应急防治传染病的专业人才培养。鼓励中医药临床实践技术的传承发展与学习名老中医防治传染病的思想精髓，对中医药防治传染病的古籍文献进行研究梳理。建立中医防治传染病的科研基地，由政府主导，提供充足物资、设备、技术、专业人才，完善中医药防疫团队建设。此外，中医院、中医职业技术院校应开设中医传染病应急救治课程，培养中医防治传染病专业人才，整合各地防治传染病特色资源，促进经验共享、交流学习，促进中医学科长效发展。保障中医药资源应急储备和生产，用以调度应急资源，保障中医药可以在第一时间参与到传染病防控工作中。

3.4 加强中医药的国际交流与合作

病毒不分国界，疫情不分种族，全人类只有共同努力，才能战而胜之。构建人类命运共同体，做好与全球各个国家传统医药的交流与合作工作，开展相关的学术交流讲座论坛，打造中医特色品牌企业，规范中医药企业运行机制，提高中药药材的产品质量。推进中医药走向世界。可以借助多种形式宣传中医药文化，借助海外孔子学院、扎住海外的医疗机构、个人中医工作室等平台，在海外宣传中医特色产品，以及通过提供中医义诊服务，比如：拔罐、刮痧和正骨等，让更多的海外民众可以切实体验到中医的特色疗法，领略中华深厚的历史文化底蕴。向世界无私分享以中医药疗治为代表的中国智慧、中国方案，彰显中国推动构建人类命运共同体的真诚愿望。

PU-081

基于集聚度和 GIS 技术的湖北省卫生资源配置公平性评价

高梦婷* 汤梓菲 袁磊
武汉大学人民医院信息中心

“没有全民健康，就没有全面小康”的重要论断说明医疗卫生服务的平等性在建设实现全面小康社会的目标时举足轻重。长期以来，由于经济社会发展的地区差异，医疗卫生资源分布不均衡是我国卫生领域的突出问题[1]，也是造成跨省就医、增加群众就医负担的重要原因。近年来，我国自上而下推出各种形式的区域医疗中心建设试点，旨在形成地理上的医疗卫生集群，促进卫生资源利用。《湖北省医疗卫生服务体系发展规划（2015—2020 年）》提出打造六个省级区域医疗中心[5]。湖北省既存在武汉这种国家级中心城市，又有恩施、十堰等经济发展较为落后的山区，其卫生资源配置的公平性尚无研究。既往对卫生资源配置的公平性评估多是基于洛伦兹曲线、基尼系数、泰尔指数等方法，较少考虑地理面积和空间位置的影响[2-4]。本研究基于 Lorenz 曲线和基尼系数分析 2011-2019 年湖北省卫生资源配置公平性的同时，利用集聚度和 GIS 技术对 2019 年湖北省卫生资源集聚特征进行空间分析，为卫生资源合理配置提供数据支持。

1 资料与方法

1.1资源来源

分别从 2012-2020 年湖北省各市州统计年鉴中提取卫生资源指标[6]，本研究纳入的指标有卫生机构数、床位数和卫生人员数，卫生人员数包括卫生技术人员数、执业/助理医师数、注册护士数。同时提取各行政区域常住人口数作为评价参考。

1.2评价方法

1.2.1 Lorenz 曲线和基尼系数

Lorenz 曲线和基尼系数是反映资源分配公平程度最常用的经济指标。先将各地市卫生资源指标分别按照升序排列，计算出人口累计百分比和卫生资源累计百分比，然后以人口累计百分比为横轴、以资源累计百分比为纵轴，绘制 Lorenz 曲线。曲线越弯曲，表明资源分配差异愈明显。在 Lorenz 曲线的基础上计算基尼系数，具体计算方法为：Lorenz 曲线与绝对公平线（即 45 度线）之间的面积为 A，Lorenz 曲线下面积为 B，基尼系数 G=A/(A+B)。

1.2.2 集聚度

考虑到地理面积对卫生资源配置的影响，卫生资源集聚度（HARD）将地理面积作为基准，HARD 的计算公式为：$HARD_i=(HR_i/HR_n)*100\%/(A_i/A_n)*100\%$。式中， HR_i 和 HR_n 分别为 i 地区卫生资源数量和全省卫生资源总量。A_i 和 A_n 分别为 i 地区地理面积和全省地理面积。

1.2.3 局部空间相关性分析

采用 Local Moran’s I 指数评价区域卫生资源配置状况，并绘制局部空间相关性地图。Local Moran’s I 的计算公式为：

$$I=\frac{n^2}{\sum_i\sum_j W_{ij}}\times\frac{(x_i-\bar{x})\sum_j W_{ij}(x_j-\bar{x})}{\sum_j(x_i-\bar{x})^2}$$

式中，n 为地区数量，W_{ij} 为权重矩阵，x_i、x_j为各空间单位的卫生资源水平。I>0 表示区域存在“高高”聚集或“低低”聚集，I<0 表示区域存在“低高”聚集或“高低”聚集[7]。

2 结果

2.1 2011-2019 年湖北省卫生资源配置总体水平

2011-2019 年，湖北省卫生资源量总体呈增加趋势，除卫生机构数之外，每千人口卫生资源数均逐年增加。每千人口床位数自 2011 年的 3.890 增至 2019 年 6.858，每千人口卫生技术人员、执业/助理医师、注册护士分别由 2011 年的 4.656、1.773、1.773 增至 2019 年的 7.078、 2.616、3.304。

表 1 2011-2019 年湖北省每千人口卫生资源数

年份	卫生机构数(个)	床位数（张）	卫生技术人员（人）	执业/助理医师（人）	注册护士（人）
2011	0.619	3.890	4.656	1.773	1.773
2012	0.610	4.378	4.996	1.888	2.002

13	20	0.614	0	4.97	5.334	2.021	2.206
14	20	0.620	3	5.47	5.769	2.168	2.476
15	20	0.618	5	5.87	6.282	2.324	2.821
16	20	0.616	4	6.14	6.549	2.420	2.977
17	20	0.615	9	6.35	6.772	2.501	3.123
18	20	0.615	1	6.65	6.943	2.567	3.225
19	20	0.599	8	6.85	7.078	2.616	3.304

2.2 Lorenz 曲线和基尼系数

2011 年、2019 年湖北省卫生资源分布的 Lorenz 曲线见图 1，两年的结果均显示出卫生机构数的 Lorenz 最为弯曲，2011 和 2019 年卫生机构数的基尼系数分别为 0.29、0.21。卫技人员数和床位数的 Lorenz 曲线较接近，基尼系数在 0.13-0.18 左右。

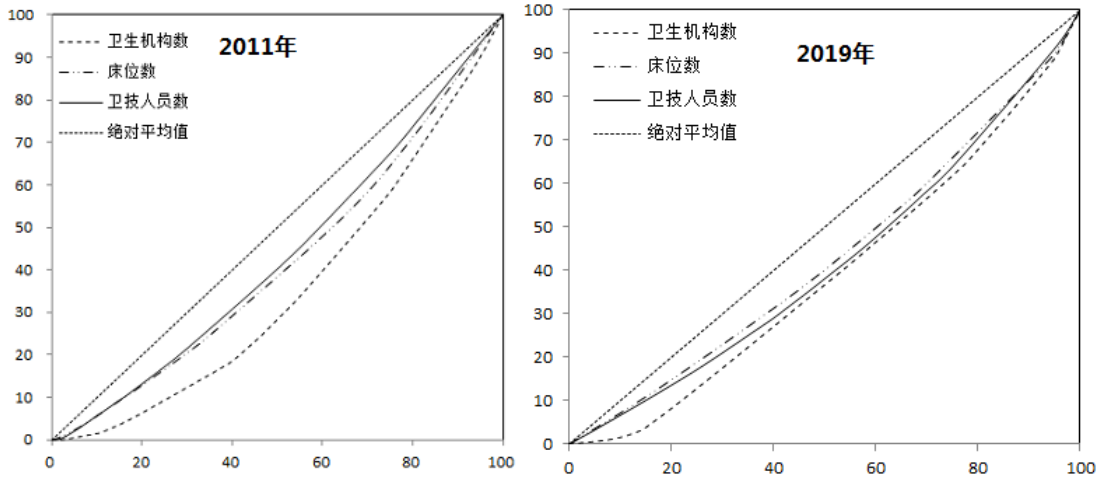


图 1 2011 年和 2019 年湖北省卫生资源分布 Lorenz 曲线

2.3卫生资源集聚度

图 3 为 ArcGis 绘制的湖北省卫生资源集聚水平分布图，颜色越深表明卫生资源集聚水平越高。结果显示，2019 年湖北省卫生机构数、床位数、执业/助理医师数和注册护士数四项指标均存在明显的地区差异，鄂东地区卫生资源集聚度明显高于鄂西，武汉市医疗卫生资源集聚度最高。

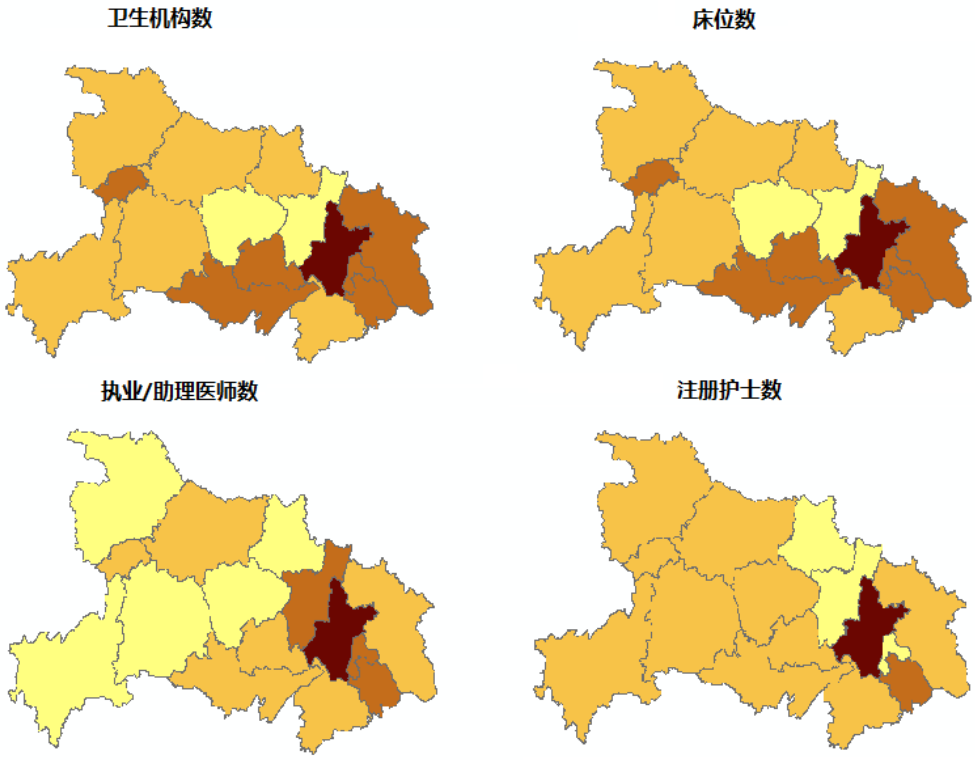
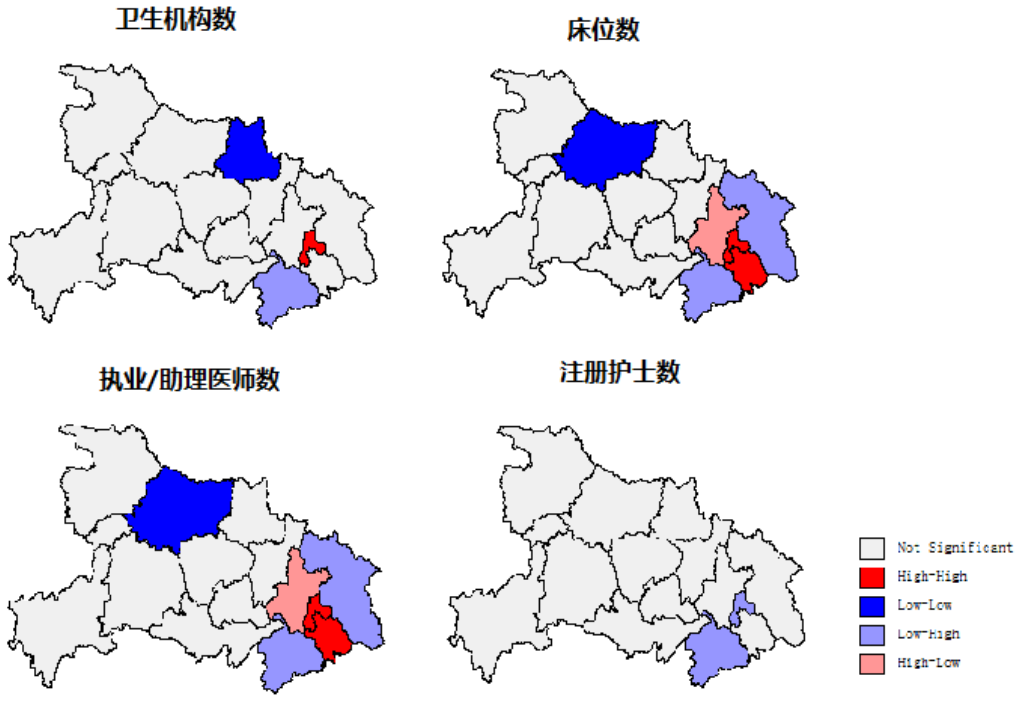


图 3 湖北省卫生资源集聚水平分布图

2.4卫生资源集聚度的空间相关性分析

基于上述卫生资源集聚度进行空间相关性分析，卫生机构数的“高高”聚集区为鄂州市，“低低”聚集区为随州市。床位数、执业/助理医师数的空间相关性分析结果与卫生机构数相似，“高高”聚集区为鄂州市和黄石市，“低低”聚集区为襄阳市。注册护士数未发现显著的“高高”聚集区或“低低”聚集区。



3 讨论

3.1 湖北省卫生资源总量逐年增加，按人口配置的卫生资源分布总体向好，卫生技术人员和执业/助理医师数仍有一定缺口

2011-2019 年湖北省每千人口床位数、卫生人员数、执业/助理医师数、注册护士数均逐年上升，说明卫生资源总量呈增加趋势。基尼系数结果显示 2019 年卫生机构数、床位数的基尼系数较 2011 年有所降低，公平性提高，按人口配置的湖北省卫生资源分布总体向好。值得注意的是，2019 年湖北省每千人口床位数和注册护士数高于全国平均水平（6.30、3.18），但是每千人口卫生技术人员数和执业/助理医师数仍低于全国平均水平（7.26、2.77），相比北上广、江浙仍有较大差距[8]。

3.2 湖北省内卫生资源分布不均衡，东、西差异明显

每千人口卫生资源数和基尼系数仅考虑到人口对卫生资源配置的影响，未考虑到地理位置和面积的作用[9]，本研究进一步使用集聚度和地理信息系统来评价湖北省卫生资源配置的公平性。集聚度结果显示湖北省存在区域间卫生资源分布不均衡，卫生机构数、床位数、执业/助理医师数、注册护士数均集中于鄂东部，尤其是武汉“1+8”城市圈内[10]：卫生机构集聚度超过 1 的 9 个地市中，7 个属于武汉“1+8”城市圈；床位数集聚度超过 1 的 8 个地市中，7 个属于武汉“1+8”城市圈。卫生资源集聚度的空间相关性分析同样表明卫生资源“高高”聚集区均位于鄂东，“低低”聚集区均位于鄂西。“高高”聚集表明区域卫生资源充沛，“低高”聚集、“高低”聚集区表明区域周边存在卫生资源聚集，形成辐射效应，可通过区域间交流提升低水平地市卫生资源水平或共享医疗卫生资源，“低低”聚集区域卫生资源欠缺，应为关注重点。

3.3 鄂西部未形成卫生资源聚集，应进一步加强区域统筹规划

湖北省在制定经济和社会发展规划时，已经开始注重鄂西部重点城市建设。2017 年提出的“一主两副多极”发展战略便以武汉市为中心，以襄阳、宜昌为两个省级副中心[11]。在医疗卫生服务发展方向，湖北省 2016 年便提出在鄂西部分别以十堰、襄阳、荆州和宜昌为中心打造三个区域医疗中心。本研究结果显示 2019 年武汉市卫生资源集聚度在 4.14-6.83 之间，龙头作用凸显，在武汉周边形成卫生资源热点。但鄂西部的三个医疗卫生中心均未发展成熟，尚未形成卫生资源聚集，尤其是鄂西北地区，卫生资源水平整体较低。

综上所述，2011-2019 年湖北省卫生资源总量总体向好，但相比全国平均水平仍有一定差距。集聚度和 GIS 技术的卫生资源空间分布图显示，各地市间卫生资源配置不均衡，卫生资源集中在武汉及周边城市，鄂西部较为欠缺。应进一步加强区域卫生规划，统筹卫生资源，改善鄂西北卫生资源欠缺现状，满足人民群众就医需求。

PU-082

数字健康应用现状与挑战

肖月①，邱英鹏①，游茂①，赵佳钰②，秦佳璇③

1.国家卫生健康委卫生发展研究中心暨国家药品与卫生技术综合评估中心

2.中国药科大学国际医药商学院

3.华中科技大学同济医学院医药卫生管理学院

一、 数字健康应用现状

1. 数字健康政策发展

数字健康是指将数字技术与医疗健康相结合的领域，旨在通过数字技术和医疗健康的结合，促进个人健康的监测、治疗和管理。数字技术在医疗卫生领域的应用具有重要的战略价值，其有助于提升医疗健康效能、增进公众福祉、推动公共卫生治理现代化、促进大健康产业的高质量发展。

在 2010 年之前，我国数字健康发展处于萌芽期，医疗信息化建设随着门户网站、搜索引擎等业态的发展开始受到重视；2010 年之后，远程医疗开始快速发展，医药电商管理逐渐规范化、互联网医疗问诊模式开始推广和落地；随着数字技术不断发展，2015 到 2019 年间人工智能、大数据、云计算等技术兴起并广泛应用，“互联网+医疗健康”服务新模式不断涌现，互联网医院快速发展，政府部门高度重视，出台多项利好政策。2016 年发布的《“健康中国 2030”规划纲要》，更是将互联网医疗提升到国家战略层面。

2020 年至今，疫情蔓延、技术创新等加速了数字医疗的快速发展，国家政策支持力度持续加强。《关于推进新冠肺炎疫情防控期间开展“互联网+”医保服务的指导意见》《关于深入推进“互联

网+医疗健康”“五个一”服务行动的通知》（国卫规划发〔2020〕22 号）、《关于印发互联网诊疗监管细则（试行）的通知》（国卫办医发〔2022〕2 号）等文件陆续出台，有效支撑数字健康发展；2021 年 12 月，中央网络安全和信息化委员会印发《“十四五”国家信息化规划》，推动普惠数字医疗提供及建设公共卫生应急数字化行动。2023 年全国两会上，“数字医疗”再度成为热议话题，我国正推进“以治病为中心”向“以健康为中心”转变，在相关政策的指导下，实施国家数字卫生战略，积极探索数字健康新范式。

2. 数字医疗发展规模

数字技术快速发展的当下，医疗数字化转型成为大趋势，全球数字医疗市场规模不断加大。据 Grand View Research 数据，2022 年全球数字医疗市场规模为 2110 亿美元，2023—2030 年将以 18.6%的年均复合增长率增至 8092 亿美元[1]。近年我国数字医疗规模不断增长，据统计，中国数字医疗产业技术市场规模 2030 年有望达到 788 亿美元左右，但从我国数字健康在医疗健康支出的占比来看，2018 年以前渗透率不足 3%，2021 年仅增长至 5%左右。此外，远程问诊、线上购药、疾病管理等场景渗透率均处于较低水平。互联网医院成熟度整体不高，慢病管理、远程门诊、“互联网+”家庭医生签约服务等核心业务渗透率不足 60%[2]。

互联网医院集中建设于沿海及经济领先省市。例如北京市在规模以上数字医疗企业数量和代表企业估值上均处于领先水平；广东东莞滨海湾新区利用其优势，打造粤港澳大湾区信息技术与生命健康（IT+BT）融合发展高地。但总体来看，我国数字医疗仍处于发展初期阶段，应用场景有待深化，数字健康建设地区差异仍然存在。

3. 基础设施建设

目前国家级全民健康信息平台基本建成，健康医疗大数据全国总体战略布局逐步落实，远程医疗服务体系逐步健全。公立医院已逐步接入区域全民健康信息平台，并依托平台实现不同医疗机构之间诊疗信息共享。根据健康医疗大数据应用发展的“1+5+X”总体规划，确定江苏省（南京、常州）、福建省（福州、厦门）、贵州（贵阳）、山东（济南）、安徽（合肥）为东南西北中五大医疗健康大数据中心试点省市，分别承担国家健康医疗大数据中心、区域中心、应用发展中心和产业园建设等国家试点工程任务。截至 2020 年底，全国 29 个省份建立省级远程医疗平台，全国远程医疗服务县（区、市）覆盖率达到 88.5%，建成面向边远贫困地区的远程医疗协作网 4075 个，实现 8 个省级行政区 832 个脱贫县的远程医疗全覆盖。

4. 数字医疗的实际应用

数字医疗的应用领域十分广泛，涵盖了例如健康监测及管理、远程医疗、电子病历和医院信息系统、健康教育、医学和科学研究、药物管理和用药指导等多个方面，随着技术的不断发展和创新，数字医疗应用的领域还在不断扩展和深化。

人工智能辅助诊断。人工智能在医疗领域的应用日益广泛，其中最为突出的就是辅助诊断。通过机器学习和深度学习等技术，人工智能可以从大量的医疗数据中学习和模拟医生的诊断思维，从而辅助医生进行疾病诊断。例如处方智慧审核系统，能够应用大数据、云计算、人工智能等信息技术，辅助药师进行处方审核，改善传统用药审核系统误报率较高的缺点，提高医务人员工作效率及诊疗质量；冠状动脉 CT 血管成像人工智能辅助诊断技术能够在患者完成 CT 扫描得到原始图像后，承接手动图像后处理工作和低年资医师的初步人眼判读图像工作，缩短处理时间，此外还能对冠脉管腔狭窄、斑块情况的判读，能够对不同资历的医师起到提示作用。

数字化慢病管理。数字化精准慢病管理服务能为慢病患者和医生赋能，提升慢病管理效率，缓解公共医疗系统的负担。对患者，能够更为方便获取处方及药物，提高治疗依从性；对医生，能够更高效管理慢病患者，获取更为全面的医药产品组合；对医院，能够分流基础复诊患者，缓解公共医疗系统负担，提升医院管理和治疗效率。

数字技术适老化。数字技术可以帮助老年人预防或尽量减少因身体、认知变化而产生的诸多风险，协助慢性病的自我管理，满足日常生活和保健需求。基础电信企业推出视频客服、远程柜台等适老化服务举措，已惠及超 7000 万老年人用户；终端设备企业对智能终端产品进行适老化改造，积极推进智慧康养解决方案；互联网企业运用数字技术赋能智慧养老产品服务，百度与清华大学等联合推出的“百度五福养老智慧平台”，已为北京多个社区家庭生活的老年人提供语音交互的智能服务。

数字健康技术在疾病预测与治疗、远程医疗、数据安全等方面也取得重要进展。例如痴呆风险预测模型 UKB-DRP 能够前瞻性智能预测阿尔茨海默病的发病风险[3]；基于 Hyperledger Fabric 的医疗数据共享解决方案[4]、互联网医疗服务监管平台存储系统[5]等能够促进医疗信息共享、防止

医疗数据被篡改或泄露，保护患者隐私；手术机器人能够进行微创或远程精细操作，促进医疗手段进步，提高患者生活质量。

在新冠疫情及信息技术发展等多重因素的影响下，数字技术与医疗行业深度融合，数字治理不断完善，数字健康逐渐成为各国重点关注领域，“大健康”应用场景将拉动数字健康持续发展，人工智能、大数据等作为医疗保健领域的创新能力加速器，也将继续驱动数字健康升级[6]。

二、 我国数字健康发展存在的主要问题

1. 医疗健康数据共享困难

健康医疗数据共享是发挥数据价值的重要环节，但我国健康医疗数据共享尚处于初级阶段，面临着数据主体“不会”、“不愿”和“不能”共享的问题。“不会”共享方面，虽然全民健康信息平台标准规范和医院信息化建设标准规范已经基本建立，但未形成国家层面统一的标准规范体系，地方与国家级标准管理工作缺乏衔接，部分标准尚未真正落地应用。在线诊疗技术标准、医疗软件产品开发使用标准等重要行业标准尚未建立，医疗机构在线服务、企业产品开发无标准可依。“不愿”共享方面，数据控制者间权利、义务不明确，医疗机构通常仅在监管部门要求下才会对外提供数据。企业数据多表现出资产属性，在非政务场景下通常仅支持内部联通。医疗机构间、各类健康信息系统间共享机制尚未完全建立，参与各方的需求和价值期望未充分满足。同时，隐私保护制度的不完善也导致数据主体间不愿共享。“不能”共享方面，主要考虑到法律责任和安全问题，国家间数据共享更涉及主权问题，以及是否能统一信任第三方国际组织。

2. 数字健康核心技术创新不足

在多种应用驱动的场景中，我国数据健康技术创新能力已大幅提升，在人工智能算法研究方面已取得较大进展，但技术创新和应用能力仍需要进一步加强。我国数字健康前沿技术的理论基础薄弱且实践经验不足，创新大多是商业模式和应用创新，基础研究突破和核心技术创新较少。大数据、人工智能、区块链等技术在药物研发、检查检验、病患监护等医学领域核心应用有待进一步融合发展，临床医学领域与数字健康领域之间的协同、创新和融合度较低。以人工智能产品为例，多项人工智能医疗技术都停留在辅助诊疗和实验研发阶段，现有 AI 技术可解释性差，需结合医学文献、电子病历、医学影像等多源异构数据，构建可解释、适应性更强的人工智能模型和方法。

3. 落地应用场景有待深化

数字健康服务涉及政府部门、医疗机构、供应商等利益相关者，后续推广与应用存在诸多问题。多元参与主体协同困难，极易产生利益分配和责任分担不清等问题。数字健康“三医”联动机制有待健全。如，人工智能辅助诊断系统如何定价，由谁付费等问题尚未解决。尚未充分发挥数字健康在常见病诊疗、慢性病全生命周期管理方面的优势和作用。同时，我国部分年长者、受教育水平较低的群体、基层居民等信息素养较低人群缺乏将所获得的健康数据信息用于处理或解决与健康相关问题的能力。

4. 医疗健康数据安全面临挑战

医疗健康大数据具有隐私性强、传播链条长、易于复制的特点，在收集、存储、应用等环节都可能出现数据安全和隐私泄露问题。数字健康在商业化过程中也有可能引发隐私泄漏风险,部分机构为了商业利益将消费者医疗数据出售给第三方，导致消费者与数据使用主体之间信息危机。此外，医疗健康大数据在知情同意、信任和公平等方面也存在问题，需要在保护隐私以及数据权限的透明、合理、利益共享方面作出努力。最后，医疗健康数据跨境流动其中所有权、使用权、管理权涉及多个部门，目前不同国家和地区的数据保护法规对数据跨境流动的要求存在差异，如何处理数据跨境安全合规与数字健康发展的矛盾也是亟待解决的难题。

5. 基础设施建设和资金有待加大投入

当前我国卫生健康数字化专网尚未建立，基层医疗卫生机构基础设施建设力度不够。经济发达地区数字基础设施配备相对完善，落后地区产业结构、科技水平、人才布局等与发达地区差距较大，导致数字基础设施薄弱。数字健康领域准入投入较大，目前补贴政策及比例较少，部分处在研发初期的中小型企业因不符合固定资产投入要求，即使重点项目也很难申请补贴。银行信贷对于软件型企业，因无法提供质押物，贷款授信额度很少。支持科研类政策，如加计扣除政策，对研发费用的认定较为严格，软件升级、代码完善、运维等后端服务都不计入研发费用，导致可申请加计扣除政策费用少于投入。虽然目前为鼓励科创企业推行了各类税收优惠措施，但由于我国实施垂直税收管理体制，导致目前以税收政策为手段促进企业科技创新政策的效力不足，加上地方税务部门与科技部门政策协调不力，不同部门对政策认定存在差异，使得优惠政策实施存在困难。

6. 国家层面数字健康体系尚未形成

政策宏观布局方面，尚无数字健康战略整体安排，管理层面政策和相关法律的建设有欠缺，政府对数字健康的管理、数字健康的建设和规划投入不足。政府层面缺少对数字健康体系监管责任、监管方法、监管框架的明确，缺少对数字健康相关技术和健康相关产品管理、管控、认证方面的政策。尚未形成协同的健康数据治理模式，缺乏强制性政策要求和有操作性的政策措施，未建立统一的规范标准和术语体系，对数字健康发展的推动效果不明显，措施落实不彻底。缺乏同质化数字健康建设质量评价体系，部分医疗机构或单位在开展健康管理和数字管理的时候出现偏差，导致一些隐患或法律问题的产生。

三、 我国数字健康发展建议

1. 建设数字健康标准体系助力医疗健康数据共享

未来需深入医疗数据标准的研究，自上而下推动建立统一的数据标准。推进中国医学名词体系建设，促进医疗机构采用标准化的数据格式和编码，便于数据交换和共享。同时，不同医疗机构使用的信息系统各异，需要克服技术兼容性问题，进一步加强技术标准的推广力度，深化信息交换平台，促进数字健康各业态联动融合，形成数字健康完整产业链和生态系统。提升医疗行业各参与方对健康医疗数据背后蕴含的宏观价值认知，如健康医疗数据赋能医疗机构资源管理、智慧便民医疗服务、辅助临床诊断、患者院外健康管理等场景仍需社会各方持续深入探索，将健康医疗数据和行业实践相结合，充分发挥数据价值。规范健康医疗数据生态的利益分配机制，包括数据生态各参与方如何对数据主张合法权利、主张何种权利。加快建设指导行业利益分配机制的制度——数据产权制度、数据要素流通规则和交易制度。

2. 研发具备自主知识产权的数字健康前沿技术

进一步夯实我国数字健康前沿技术的理论基础，研发具备自主知识产权的数字健康关键技术。深化医护人员与技术人员之间有机交叉的合作，促进技术的融合创新。在重点研发计划、重点实验室、创新团队和支撑平台上向数字健康领域交叉项目倾斜。加大硬件和设备方面的支持力度，为医疗领域提供大规模的云服务资源以及更多算力和存储支持。加强技术效果和技术卫生经济学的评估，如综合考虑医保支出情况等。

3. 强化数字健康技术的落地应用

加强医学人工智能和区块链等前沿技术的试点应用。规范引导各级医疗机构在医学影像辅助诊断、临床辅助决策支持、公共卫生智能管理等领域开展试点应用。依托健康医疗大数据和基于机器学习的人工智能技术转化应用，组织研发部署服务于医疗机构、医务人员、患者的智能临床决策支持系统，助力提升基层医疗服务效率和水平。围绕数据互联互通、有序溯源、安全管理，研究创新区块链在医疗健康领域的应用场景，鼓励医疗卫生机构探索区块链技术在医疗联合体、电子处方、身份识别、电子健康档案、药品追溯、医疗保障等方面的应用。

4. 重视健康医疗数据安全和隐私保护

国家需进一步完善相关法律法规，设立专门的监管机构制定进行安全审计和监测，重视科学数据安全团体标准并推动其研究，涵盖数据加密、访问控制、身份验证、漏洞管理等方面。利用联邦学习、隐私查询、区块链追溯等技术[7]对数据采集、传输、存储、处理、交换、销毁等数据全生命周期各环节进行管理，建立跨部门跨领域密切配合、统一归口的数据共享机制。同时，加强国际合作与交流，分享数据安全的最佳实践和经验，共同应对全球范围内的网络安全威胁。

5. 加强数字健康产业资本供给

资本的有序介入和长久供给对于推动健康医疗数据产业稳定发展和突破创新非常重要。应持续加大对新兴数字技术融合创新、标准化体系建设和安全保障体系等薄弱环节的资金投入。创新财政资金支持方式，加大国家科技计划统筹力度，支持数字诊疗、智能装备、健康和老龄化数字化等领域创新研发。进一步完善针对科技企业的注册制改革中的税收、支持系统。金融机构可对不同类型的中小科技企业提供有针对性、差异化的服务，提供更丰富的金融产品。关注健康医疗数据背后对于社会各方带来的巨大收益和价值，促进社会资本投融资合作对接，激发各类社会资本活力。

6. 夯实新基建赋能数字健康

持续加大数字健康基础设施投入，强化全民健康信息一体化建设，提高国家、省、市、县四级人口健康信息平台互联互通水平，缩小东中西部、城乡地区医疗信息化差距。大力建设覆盖医联体、医共体域内的信息基础设施，提高诊疗设备物联接入能力和覆盖水平，推进数字孪生应用。布局数字健康基础设施和运行空间，搭建起联通国家医学中心、国家区域医疗中心、国家临床医学研究中心，以及国家卫生健康委、药监局、医保局等信息系统的网络枢纽，构建多类型、多途径、数据采集利用的运行机制，形成卫生健康领域的人工智能算力枢纽和生物医学计算高地，通过数字

化、智能化提升医学科技创新体系核心基地的支撑能力。完善公共卫生服务信息化体系，提升慢性病、传染病信息化管理水平，构建突发公共卫生事件信息联动机制。

7. 统筹制定数字健康战略发展规划

开展我国数字健康战略实施的顶层设计，完善数字健康服务供给策略。规划数字健康产业布局，制定健康服务与产业的监管治理机制针对数字健康产业各环节开展系统的监管与治理，在数据生成、数据共享、数据质量和准确性、政策优先顺序、人工智能技术、数据引用伦理原则等方面制定全面的监管制度，实现规范化发展。加强数字健康不同主体之间的协同治理能力。联合发改委、卫健委、医保局、药监局等部门，共同鼓励医疗机构、科研院所、企业等协同创新，加强产业链上下游资源的组织协调，共同打造数字健康良好发展的生态体系。

PU-083

我国基层数字健康发展现况及未来展望

肖月，邱英鹏，游茂
国家卫生健康委卫生发展研究中心暨国家药品与卫生技术综合评估中心

一、基层数字健康发展

1.概述

根据世界卫生组织最新定义[1]，数字健康指与开发使用数字技术用以改善健康相关的知识和实践，不仅包括各类智能设备及其消费者和用户，还包括应用物联网、人工智能、大数据和机器人技术的卫生健康服务。我国将人民健康作为民族昌盛和国家强盛的重要标志，同时近年来高度重视数字经济发展，制定实施“健康中国”和“数字中国”战略，以实现国富民强和助推高质量发展目的。因此，数字健康在我国被赋予特殊的意义，是健康中国和数字中国目标交汇的稳压器，也是健康数字化发展和健康经济创新的助推器。

数字健康在我国及全球新冠疫情大流行应对中均获得了加速度发展的机遇，显示出对卫生服务质量、效率、可及性、可负担性等改善的潜在价值。目前全球超过 120 个国家建立了全国性数字健康政策或策略，加强了卫生信息化建设，多国近期通过立法和治理机制推动健康数据的共享利用，2023 年第 76 届世界卫生大会和亚洲经济合作组织（APEC）2023 年会均强调数字健康对实现可持续发展目标及全民健康覆盖的促进作用。数字健康的潜在优势包括支持个体主动管理健康和福祉，助力医务人员遵循临床指南并提供高质量服务，并可通过优化人力资源配置使用和诊疗技术服务供应系统加强卫生服务体系。

2.我国数字健康发展

十八大以来我国数字经济取得了长足发展，数字技术研发和产业布局加速，同时健康中国战略制定实施，为“数字+健康”交融汇聚提供了重大机遇；二十大进一步要求推动战略性新兴产业融合集群发展，构建新一代信息技术、人工智能、生物技术等一批新的增长引擎，促进数字经济和实体经济深度融合，数字健康不仅仅是“数字+健康“，还融入了对经济发展的牵引驱动，成为促进卫生高质量发展、助推经济社会高质量发展的新引擎。

我国卫生健康领域在 2000 年后加快信息化发展步伐，加大信息基础设施建设，推动信息数据标准化；2014 年《关于推进医疗机构远程医疗服务的意见》促进了远程医疗及相关产业的迅速发展，2015 年乌镇互联网医院建立标志着互联网+医疗的实体化和产业化模式，起人工智能、大数据、云计算等新技术应用广泛，医疗信息化建设不断增强，互联网医院逐步发展，线上线下医疗健康服务闭环开始打通；2020 年起新冠疫情加速数字健康应用，《关于深入推进“互联网+医疗健康”“五个一”服务行动的通知》《远程医疗服务管理规范（试行）》《国家卫生健康委办公厅关于做好信息化支撑常态化疫情防控工作的通知》《互联网诊疗管理办法（试行）》《互联网医院管理办法（试行）》《互联网诊疗监管细则（试行）》等文件为推进引导数字健康的良性发展提供政策支撑。随着卫生健康数字化转型发展，初步显示出以下三方面成效：

一是促进卫生质量、可及性和效率改善。数字化技术的应用使互联网诊疗、远程医疗和医联体医共体高效协作成为可能，一方面可降本增效、减轻医护人员工作负担，改善卫生体系整体资源投入模式、提升服务效率，另一方面通过提升诊疗安全有效性，改善患者顺畅便捷就诊体验，提高卫生健康服务在社区和家庭的可及性。

二是加速创新诊治技术扩大应用。人工智能、大数据等技术的应用显著加速了药械研发应用进程，同时也促进了创新数字疗法的研制应用，为建立精准诊疗方案提供先进技术手段。云胶片、云药房、个体健康信息汇集预警等新干预措施支持主动健康等新理念、新模式、新业态的发展。

三是驱动治理模式创新发展。医药、医疗、医保不断深化信息化标准化工作，通过提升数据治理水平，促进便民利民政策措施的制定实施和贯彻落地，增强医药产品供应能力，促进诊疗服务提质增效，提升医疗保障水平。

3.基层数字化进展

老龄化社会加速到来、慢性病高发、新发重大传染给卫生服务体系带来持续压力，“十四五”时期，我国 60 岁及以上人口占总人口比例将超过 20%，78%以上的老年人至少患有一种以上慢性病，失能失智问题突出，国家大力发展养老产业背景下，健康老龄化要求推动卫生服务模式向预防保健和居家管理方式转变，基层数字化有助于推动主动健康理念贯彻，利用数字技术支持社区和居家健康管理和疾病监测，以可负担方式提供经济、方便、高效的整合健康照护服务。

随着数字健康体系不断完善，为确保上述健康需求，基层数字化加速发展。截至 2022 年 9 月，全国所有省份、85%的地市、69%的区县已建立区域互联互通的全民健康信息平台；建立完善省市县乡村五级远程医疗服务网络，已覆盖所有的地市和 90%以上的区县，通过基层数字健共体、云上医共体等建设推动优质资源要素流动和服务模式创新；建成应用全国统一的医疗保障信息平台，实现医保政务服务事项“跨省通办”，2022 年底前每县至少 1 家定点医疗机构实现医疗费用跨省直接结算，农村居民可异地便捷办理医保业务，实现无卡结算、全国通用；82 类已在我国获批上市的人工智能医疗技术产品逐步在基层落地应用，支持心脑血管疾病、肺癌、眼病等重点疾病的基层诊疗能力水平的提升。

4.典型经验与作法

全国多个地区涌现出基层健康数字化典型案例和经验，包括运用大数据、5G、智能应用程序、人工智能等数字技术支持慢性病基层管理、优化患者就医流程和改善诊疗体验、加强诊疗行为管理和提升标准化规范性。

案例一：天津市基层数字健共体支持慢性病连续管理

天津市卫生和医保部门联合指导下，委托第三方互联网医院平台牵头，协同全市 266 家基层医疗卫生机构组建了紧密型基层数字健共体，搭建数据管理和运营平台，初步实现医疗、医药、医保、基本公卫、健康管理等数据聚集和联通，为慢性病管理模式数字化转型提供基础。71 家基层医院建立标准化慢病管理中心，负责共同管理逾 90 万名糖尿病、高血压等慢性病患者，基于数据平台建立家庭医生为主体的人头付费和区域总额打包付费机制，促进慢性病连续性管理，糖尿病按人头付费实施后节省医保资金近 1289.09 万元，显示出数字赋能基层健康管理的效益。

案例二：安徽省六安市智慧审方 AI 支持处方行为规范管理

2019 年，安徽省人民政府发布《关于 2019 年实施 33 项民生工程的通知》、六安市卫生健康委依据《2019 年安徽省智医助理建设实施方案》和《2019 年安徽省“智医助理”技术方案》，在舒城县、金寨县推进合理用药 AI 前置审方试点应用。由政府、医疗机构、企业三方参与试点工作并签订协议，运行后取得减少不合理处方数量、降低不合理药品费用的应用效果。以舒城为例，智能审方终端覆盖 15 个镇，6 个乡，共计 21 个卫生院、286 个卫生室，纳入 5 万余品类药品，干预前后不合理处方张数减少 14.23%，不合理用药费用降低 12.23%。

二、问题及挑战

尽管卫生健康数字化转型发展势头迅猛，但是基层健康数字化仍面临一些主要的问题与挑战。

1.健康相关数据应用和治理问题

数据互联互通和基层可获得性问题。医联体、医共体及各医疗机构之间各业务系统建设条块分割，系统之间信息互通不畅，“信息烟囱”和“信息孤岛”现象突出。在实践中，大多医联体、医共体内部各医疗机构间医疗信息数据尚未实现整合，基层医院开展上下级转诊面临与多家系统对接的困境。另外，基层医院与不同医疗机构间，甚至基层医院内部都存在“数据孤岛”，一些基层医院尝试通过 HIS 系统整合，但整合效果不佳。

数据流动和应用问题。仍需深入研究健康数据产权制度、数据要素流通规则和交易制度，探索如何结合区域医疗资源结构布局和群众健康需求，推动医联体率先探索跨院、跨地域的数据应用模式。健康医疗数据经过标准化治理后需在各级医疗机构之间流动，进一步打破优质医疗服务资源在空间和地域中的限制、实现医疗集团内部的统筹管理，加速深化分级诊疗制度改革。

2.基层数字化转型仍需构建系统保障体系

基础设施建设及配套监管措施不足。基层医疗卫生机构信息化基础设施建设力度仍有待加强。互联网医院准入“门槛”高、标准严、诊疗限制多，基层信息化很难满足国家信息安全等级保护认证等标准。互联网诊疗数据的存储、应用、共享等都可能存在隐私保护与安全风险隐患问题，需要相对成熟、完备的监管手段，基层医疗机构在这方面总体相对较弱。

适宜的数字技术应用有限。当前各类数字技术的规范化、标准化程度低，对实际临床环境适应性不足，应用场景不明确，基层医务人员及群众对数字产品认可接受度低导致应用落地难，实际应用不足情况下识别遴选高效干预技术措施面临困境，加之创新数字技术投入和付费机制未建立，导致基层医疗卫生机构对于以人为本、基于循证医学的数字技术手段应用不足。

三、未来展望

在健康中国和数字中国双轮驱动下，随着健康数据标准化和规范化共享工作推进，基层健康数字化发展潜力巨大，数字技术赋能基层的项目措施已遍地开花，给基层卫生发展带来机遇，也给数据治理、数据信息共享应用、基础设施建设、保障配套措施等基础工作提出了更多的挑战。为适应数字健康在我国发展新形势，未来基层卫生的数字化转型仍需要关注以下重点：

一是围绕重大疾病需求、重点人群健康需要，依法依规定义标准化的人口健康核心数据集，向具备信息基础、在区域诊疗服务体系发挥网底作用的基层医疗卫生机构赋予数据积聚、分析、应用的核心职责，通过以基层为中心探索推动数字健康治理和数据共享利用，为基层卫生数字化转型提供基础条件。

二是在信息基础设施建设过程中，向基层倾斜，特别需要加强数据接口、信息管理平台、知识管理工具、创新人工智能器械的投入建设，支持基层运用适宜的数字技术以实现服务质量效率提升和效果效益优化。

三是在基层卫生改革中深入贯彻主动健康理念、加快构建主动健康体系、积极提升主动健康意识，推动数字技术在建立以社区家庭为核心的健康维护、促进和保障模式中创新应用，普及基本数字技术应用，提升人群数字素养和运用能力，挖掘高价值数字技术和有潜力应用场景，发挥线上咨询问诊、人工智能辅助疾病筛检、数字化慢性病管理等干预服务在基层应用效果效益。

PU-084

一体化电子病历系统在临床应用的研究

高卫光
济源市人民医院

引言

电子病历是利用计算机技术以及互联网技术，实现病人病历的电子化管理。 一方面，电子病历可以将病人所有问诊记录进行汇总和梳理，为医生提供全面的病历数据，方便医生进行科学诊治；另一方面，电子病历为病患解决病历丢失等相关问题，利用电子病历可以实现病历管理的高效化和集中化。不仅如此，随着现代电子病历的发展和研究，针对电子病历的科学研究，可以实现病人的周期化管理，以数据的集中梳理进一步实现病历内容科学整合。例如，电子病历的数据分享，可以实现病人社区医院与城市医院的信息互补，同时还能够按照计算机的设定形式，为病患提供对应的治理方案，从而减少医生的诊治压力。电子病历的意义和价值，还需要进一步发现和挖掘。

一体化电子病历系统概念

电子病历系统（Electronic Medical Record， EMR）是医疗信息技术的一种重要应用，是指经过计算机和其他信息技术（例如网络和智能手机）处理的患者诊疗过程中发生的一系列关于临床诊疗的记录，以最大化患者和医疗机构之间信息交换和共享的效果。

一体化电子病历系统整合全院信息，建立以病人为中心全面优化诊疗流程，实现患者院内院外诊疗衔接，简化就诊既往繁杂程序，缩短患者就医过程，实现患者诊疗流程及数据的全程共享，提升医院整体运营服务能力和患者就医体验。在实现电子病历全覆盖、临床信息资源全面整合、临床

信息流程全面优化的基础上，拓展临床辅助决策应用，全面提升医护人员的工作效率，保障医疗质量，并为科教研和协同医疗提供基础保障。

一体化电子病历系统建设必要性

应用电子病历系统能够提高医疗质量和安全性，提高工作效率和便捷性，促进医疗信息共享和协同工作，便于数据分析和研究，同时节约资源和减少存储空间，是医疗行业信息化发展的必要举措。

提高医疗质量和安全性：电子病历系统可以实现医疗信息的标准化和规范化，减少信息传递和记录过程中的错误和遗漏。通过电子化的方式记录患者信息，可以提高医疗质量和安全性，减少医疗事故的发生。

提高工作效率和便捷性：电子病历系统可以将病历记录、查询、管理等功能集成在一个系统中，实现信息的快速获取和处理。医生和护士可以更快速地查找患者信息、制定治疗方案、进行药物管理等，提高工作效率和便捷性。

促进医疗信息共享和协同工作：电子病历系统可以实现跨部门、跨机构之间的医疗信息共享和协同工作。医生和其他医务人员可以实时获取患者的病历信息，进行多学科的讨论和决策，提供更综合、准确的医疗服务。

便于数据分析和研究：电子病历系统可以提供大量的医疗数据，为医学研究和临床决策提供支持。通过对大数据的分析，可以发现潜在的疾病风险、药物不良反应、治疗方案效果等信息，促进医学科学的进步。

节约纸质资源和减少存储空间：传统的纸质病历需要大量的纸张和存储空间，不仅浪费资源，还不易管理和保护。而电子病历系统可以实现无纸化记录和存储，节约资源，并通过数据备份和恢复功能，保障病历信息的安全性和可靠性。

应用一体化电子病历系统需考虑的问题

一体化电子病历系统在临床应用的研究涉及到将传统的纸质病历转化为电子形式，并以数字化的方式存储、管理和共享患者的健康信息。这种系统的应用主要关注以下几个方面：

流程优化：如何通过电子病历系统改进医疗流程，提高临床工作效率和患者照料质量。例如，电子病历可以减少繁琐的纸质文档处理和信息查找时间。

数据安全：确保电子病历系统中患者隐私和敏感信息的安全。如何进行合理的数据加密、访问控制和权限管理，以防止未经授权的访问和数据泄露。

数据标准化和互操作性：制定统一的数据标准和交互接口，以实现不同医疗机构之间电子病历数据的互操作性。这可以促进医生和医疗团队之间的协作，提供全面的患者健康信息。

数据分析和决策支持：如何利用电子病历系统中积累的大量数据进行分析，从而提供更好的临床决策支持。例如，通过数据挖掘和机器学习技术，可以发现患者特征、治疗模式和疾病趋势，从而改进临床实践。

用户界面设计：如何设计用户友好的界面，使医生和其他医疗专业人员能够轻松使用电子病历系统。良好的用户界面设计可以提高工作效率并减少错误。

临床决策支持：通过整合临床指南、最新研究和患者数据，为医生提供个性化的治疗建议和决策支持。

移动应用和远程访问：随着移动技术的发展，如何将电子病历系统扩展到移动设备上，并实现远程访问。医生可以在任何时间和地点查看和更新患者的健康信息。

数据迁移与整合问题：将现有的纸质病历或其他数字化系统中的数据迁移到新的电子病历系统中，并进行数据整合，确保历史病历和现有数据的一致性和完整性。

用户培训和接受度：如何培训医生和其他医疗专业人员使用电子病历系统，并提高他们的接受度。培训和接受度的提高可以确保系统的有效使用和最大化的效益。

一体化电子病历系统在临床的应用

一体化电子病历系统功能

病历管理

诊疗计划

以患者住院时间为横轴，显示为住院日、手术日和元年时间，以患者在院期间全部临床诊疗事件为纵轴，包括患者的体征、病历、病程、护理、处置、手术、用药、检验检查等。

同时该集成界面也可以根据设定好的诊疗计划，提示用户每日需要完成的临床工作，并对未完成的工作进行提醒。通过该集成界面，用户可以完全掌握当前病人的情况，并对需要开展的工作有了直观的了解。

病历录入

可以实现医生、护士等医务人员对患者的病历进行录入和管理，提供多种方式的结构化病历书写模式，支持知识库、医学术语等内容自定义动态插入病历文书中，提高医生的编写效率和质量。包括基本信息、病史、疾病诊断、检查结果、用药情况等，方便查阅和修改。

病历打印

支持整体打印、连续打印和套打功能，并且能够按照卫生管理部门和医院对病历格式的要求进行适当调整。

系统的打印输出功能支持整体打印、连续打印和套打功能。

病历查询

可查询门诊资料和历史病历。医生可以在医生工作站上查看护士填写的各类护理资料内容。

模版管理

医院管理人员可以对医院的病历模板进行维护，增加、删除、修改等操作权限，病历元素，即病历中的项目，通过病历元素实现病历数据的一元化管理。按照纸质病历的样式，支持病历模板的自定义和可视化维护。将病历模板分配到指定的科室，该科室下的医护人员就能够使用该模板进行病历的书写。

诊断辅助功能

可以根据患者的病情和病历数据，提供一些辅助诊断的功能，比如自动推荐可能的诊断结果、提供相关的医学知识和参考资料等。

医嘱管理

医嘱录入

可以记录医生对患者的诊疗医嘱，包括药物处方、治疗方案、手术安排等，方便医务人员随时查看和调整。

检查和检验结果管理

可以记录和管理患者的各种检查和检验结果，比如血液检查、影像学检查、实验室检验等，方便医务人员进行综合分析和评估。

临床路径管理

系统可以直接与临床路径系统整合，根据临床路径的诊疗内容，进行医嘱及诊疗项目的关联，并能够根据临床路径管理的要求，对医嘱的下达进行提醒。

病历质量管理

可以统计和分析患者的病历数据，评估医疗质量和效果，提供医院管理决策的依据，促进医疗质量的提高。

三级质控管理

完善病历闭环质控体系，整体内容涵盖病历评分、缺陷提取、反馈平台、临床整改；系统中的质控管理功能分为个人、科室和医院三个级别，分别满足不同管理的要求。

流程管理

电子病历具有流程控制功能，每个医院都可以按照自己的流程在电子病历系统中进行预先定义，医生在编辑相关病历文档时，即可按照规则进行检查，对不符合的操作予以提醒，避免了病人在出院的时候，医生才去补写一些相关的医疗文书。

事先提醒

系统能够根据医院的病历书写要求，对临床医护人员进行各种预先的提醒功能。

事中监督

由系统根据预先设置好的病历书写要求，对所有住院病人的病历书写情况进行统一监督。通过电子病历的使用，完全改变了以前病历质控人员要下到病房，去抽查病历的情况。

指标维护

提供医疗质控指标基础知识库，指标覆盖多级部门、各类人员、所有业务流程，自定制指控标准，质控标准为总分、甲级最低分、乙级最低分等设置。

临床路径

搭建临床路径管理，在医院的内部形成信息的整合和统一管理。临床路径的应用贯穿于整个医疗过程中，路径系统以临床路径管理功能为主体，以标准路径为引擎，并结合临床知识库，进而实现路径制定、路径执行、变异分析、统计评估、路径改进的管理功能，形成临床路径的闭环管理流程。

多科室协同

可以实现不同科室之间的信息共享和协同工作，比如医生之间的会诊、药师的调配工作等。

数据安全和隐私保护

可以对病历数据进行权限管理和数据加密，保障病人隐私安全。

数据共享

可以与其他医疗机构、健康管理平台等进行数据共享，方便患者在不同医院就医时的病历传递和信息查询。

一体化电子病历系统建设成效

一体化电子病历系统是一种能够改善医疗流程和提升医疗质量的工具。它不仅可以改善患者和医务人员之间的信息交流，还能更有效地管理患者的临床信息和医疗病例记录。通过提供准确、及时、安全、有效的医疗信息记录，一体化电子病历系统为医疗活动提供了重要的支撑，最大限度地提升了医疗质量和效果。

具体来说，一体化电子病历系统能够改进病历记录的完整性和准确性，保障医疗记录的安全性。它能够降低行政成本和医疗安全风险，减少患者在接受检查和治疗流程中的时间和经济消耗。同时，一体化电子病历系统还能改善医护人员的诊疗流程，简化病案的维护和管理流程，提高医疗组织的效益和预算支出。

此外，一体化电子病历系统还可以利用智能技术，提供更多的医疗历史记录和实时体征监测，帮助临床工作者进行更好的诊断和治疗。它能够便捷地获取患者的历史数据，并通过自动化和统计分析等功能，加快诊疗流程。医院可以准确提取患者的医疗信息，并定期更新和修改，以确保医疗信息的准确性和及时性。这样，医院能够更清楚地了解患者的情况，并为患者提供更好的医疗服务。

结论

如今，医疗信息化建设已经进入了一个具有挑战的阶段。然而，无论我们如何发展和引入新的技术，我们始终要回归医疗的本质，即医疗安全和质量的保证。电子病历的发展为医疗机构的信息化管理进程提供了强有力的支持和帮助。为此，引入一体化电子病历系统，构建以电子病历为核心的医院临床信息平台，能够有效地保障医疗的安全性和质量，这对于提升医院的综合实力，打造数字化医院具有重要的保障作用。

PU-085

手术操作编码标准体系的国际经验与启示

刘立煌* 黄锋* 张霞* 秦秋彦*
北京大学人民医院

1、引言

手术操作作为一类医疗技术是医疗服务过程的重要组成部分，因此手术操作数据是卫生服务利用的一个侧面，是卫生资源规划配置的一项重要参考数据，自 2019 年国家公立医院绩效考核（后简称国考）开展以来，手术操作数据的相关统计指标愈加得到重视。然而当前我国使用的编码标准为基于线分类的 ICD-9-CM3，由于其扩展性较差难以适配当前高速发展的医学和满足高质量发展的统计需求。本文对美、英、澳、加四国的手术操作编码标准及统计应用情况进行了系统比较，为完善现有的编码标准和相关分析指标框架，充分挖掘手术操作数据价值提供参考。

2、四国手术操作编码标准比较

澳大利亚和英国的手术操作编码多年沿用 ACHI 和 OPCS-4，这两种都是以线分类为底层模型的知识分类模型，其优点是分类条目为先验术语，更加贴近现实，且层级清晰，但不足是扩展空间非常有限，架构相对固化，更新成本较高，特别是对于积累了多年数据的国家。美国和加拿大最新的手术操作分类都是基于面分类体系的编码架构，与世界卫生组织开发的 ICHI 的基础架构相似，其最鲜明的特点是结构化模型，结构化的维度一般包括部位/术式/入路/设备/药物/组织/手术状态/解剖/范围等，利于计算机处理，且分类条目界定清晰，扩展性强，代表着未来医学编码的发展趋势。

表 1 美、英、澳、加四国手术操作编码模型比较

国家	手术操作编码体系	分类方法与编码方法	分类轴心	特点
澳大利亚	ACHI	线分类 一术式多编码	部位/术式/入路	层级清晰
英国	OPCS-4	线分类，一术式多编码	主要章节：部位/术式/入路 附属章节：解剖/范围	层级清晰；统计以主要手术操作为主；通过附属章节标注细节
美国	ICD-10-PCS（FY2015~）	面分类&线分类（部位），一术式一编码	部位/术式/入路/设备/范围	分类方法先进
加拿大	CCI	面分类&线分类，一术式一编码	部位/术式/入路/设备/药物/组织/手术状态/解剖/范围	分类方法先进，扩展维度最全面

3、四国手术操作编码数据统计应用现状

2.1 美国

美国的住院患者手术操作由美国疾病预防控制中心（centers for disease control and prevention, CDC）下属的国家卫生统计中心（national center for health statistics，NCHS）统计，数据来源分两个时期，1965 年至 2010 年是全国出院调查（national hospital discharge survey，NHDS），每年多阶段分层抽样调查非联邦短期（平均住院天数＜30 天）社区医院；2011 年后国家医院服务调查（national hospital care survey，NHCS）取代了 NHDS，并加入了门诊手术调查，2007 年符合条件的样本医院为 477 家医院数。2015 年启用 ICD-10-PCS 用于手术操作编码，每个病例收集最多 8 个手术或非手术室操作。

美国自 1965 年起持续至今，每年在国家卫生统计年度报告（National Health Statistics Reports）报告手术操作统计结果，通过样本对全国手术量和每万人手术操作数进行推断估计，并在 ICD-CM3 的章节系统和特定术种、性别、年龄等维度进行交叉描述性分析，并对部分术种数量进行时间趋势分析。最新的年度报告显示男性最多的手术操作为心血管系统（超过 1/5）、消化系统、骨骼肌肉系统，而女性则为产科相关、消化系统、心血管系统的手术操作（排除诊断性和治疗性操作，章节 87–99、00）。

2.2 英国

英国国民医疗服务体系（National Health Service, NHS）管理部门常规收集住院患者手术操作信息，患者级别的手术操作信息被纳入医院诊疗统计数据库（Hospital Episode Statistics，HES）进行上报，手术操作分类标准使用的是 OPCS-4，每个病例最多可以上传 24 个手术操作，上报时需要指定主要手术操作（main procedure），其定义是本次住院资源消耗最大的手术操作。OPCS-4 保留了残余分类.8 和.9，但只能用于主要手术，被设计主要用于流行病学统计、医疗机构支付、卫生资源规划与配置、绩效评价等。

基于 HES 数据英国 NHS 每年常规报告主要手术操作和全部手术操作统计结果，分别在 OPCS-4 的章、节、三位码、四位码四个颗粒度统计，主要手术操作的统计分析维度包括性别、入院途径（急诊或者门诊）、是否日间手术、平均/中位等待时间、平均/中位住院天数、年龄等，全部手术操作则对是否主要手术操作、性别、年龄、是否日间手术进行分布描述。

3.3 加拿大

加拿大除魁北克省以外所有地区的所有急症医疗机构都需向地区卫生行政部门上报手术操作及其他出院摘要信息，形成出院摘要数据库（Discharge Abstract Database, DAD)，魁北克省是直报给加拿大健康信息研究所（Canadian Institute for Health Information，CIHI），DAD 和魁北克省的直报信息整合而成医院发病数据仓库（Hospital Morbidity Database,HMDB）。2001 年之前加拿大的手术操作标准主要使用 ICD-9-CM3，之后主要使用加拿大医疗服务干预分类（Canadian classification of interventions, CCI），由 CIHI 开发用于获取医疗机构的手术操作信息，共 17417 个编码。

急症医疗机构的手术操作编码经过全国分组系统（National Grouping System，NGS)形成大手术术种（不含诊断性和治疗性操作）进行全国/地区内比较，其定义方式是枚举 CCI 编码，例如胆囊切除术：1.OD.89.^^，扁桃体切除术：1.FR.78.^^, 1.FR.89.^^，复杂的比如骨折手术包括 136 个 CCI 编码。CIHI 每年发布全国和各地区前十的手术统计结果，并对不同地区前十术种的性别分布、平均住院日进行了描述性分析。

4.4 澳大利亚

澳大利亚的住院手术操作上报包括两个部分，一个是所有公立医院需要上报择期手术相关信息形成的择期手术等待时间数据库（NESWTDC），包括择期手术登记患者数据和移除患者数据，主要用于统计择期手术等待时间，该数据库监测 155 个目标手术，比如白内障摘除、膀胱镜检查等，被分成了心胸外科、耳鼻喉科、普通外科等 12 个大学科，手术术种采用了自定义分类，同时在指导手册里做了 ACHI 的等价编码说明，分析结果作为专题报告发布。此外澳大利亚对目标择期手术的数量和等待时间进行了学科、地区和医院的分布描述。

另一个是出院统计，手术操作包含在国家最小数据集（national hospital morbidity database）中的住院患者信息部分，所有公立和私立的急症、精神疾病、日间医疗机构、酒精与药物依赖治疗中心都需上报，形成国家公立医院发病数据库（National Hospital Morbidity Database），手术操作采用是 ACHI（Australia Classification Of Health Intervention），AIHW 基于国家公立医院发病数据库（National Hospital Morbidity Database）每年发布手术统计的结果汇总表，分析维度包括手术的专业、术种、年龄、性别、是否日间的分布情况，比如 2020 年度澳大利亚最常见的操作是消化系统、泌尿系统、骨骼肌肉系统。

3、讨论

四国的经验充分体现了手术操作在卫生管理中的分析价值，而目前我国使用的线分类 ICD-9-CM3 已难以满足当前管理需求，特别是国家公立医院绩效考核评价，例如微创手术，由于 ICD-9-CM3 诞生之初腹腔镜、胸腔镜等技术尚未发展成熟，故 ICD-9-CM3 中大部分的分类条目缺乏手术入路的分类轴心，或者缺乏明确的腔镜入路，造成其在手术统计方面的局限，特别是容易出现细目与扩展条目微创、四级手术属性不同的现象，而面分类体系的入路维度恰好能够解决这个问题，此

外其还有扩展性强、便于统计等多个优点，美国的 PCS、加拿大 CCI 均是采用面分类体系架构，我国应积极借鉴向面分类模型发展，以满足临床和管理需求。

此外从四国对手术操作数据的统计应用分析领域看，美国更侧重于质量安全、手术谱等方向，而英、澳、加三国更加侧重于手术操作的可及性评价。从卫生体制的角度分析，美国是作为典型的市场主导卫生体系，效率问题相对不突出，手术的质量安全是其主要的关注点，因此其侧重于在质量评价的手术统计应用，其他三个国家都是政府主导型卫生体系的国家，整个体系的问题与美国不同，因此统计分析侧重点也有所不同。英国和澳大利亚都非常关注手术等待时间，英国是一个国民卫生服务体制，其卫生系统的效率问题最为突出，澳大利亚也是全民覆盖的国家卫生保险系统，效率问题也是首当其冲，两国的手术等待时间均较长。加拿大是联邦、地区、省的混合筹资国民卫生体制，具有完备的绩效评价模型，如 CIHI 的数据分析框架是健康状态、患者安全、医疗质量、医疗可及性、非医疗性卫生决定因素、初级医疗机构的用药六个维度，手术操作统计的统计应用主要在中间三个维度。因此，结合未来对整合型卫生服务体系的发展需要，我国也可以挖掘手术操作编码在高质量发展评价其他维度的指标，例如在连续协同方面，对于紧密型医联体可以采用转诊手术比例，在人民群众健康水平方面可以采用术后生存时间（OS/PFS/DFS），以及术后生存质量评分等，充分挖掘全国手术数据集对高质量发展评价的供给能力，充分发挥数据要素的应用价值

PU-086

贵阳市健康医疗大数据标准制定研究

邵建华[1] 李 挺[2] 田智华[3]
1.贵阳市第一人民医院信息科
2.贵阳市卫生健康信息中心
3.贵阳市医学会

引言

目前我国健康医疗大数据标准国际化水平低，特别是随着大数据蓬勃发展，当前的健康医疗大数据标准对健康医疗领域支撑力很不够，健康医疗标准研制与医疗产业的蓬勃发展的应用还有相当大的差距。国外在很早以前就已经形成了健康医疗大数据标准化的机制[1]，在基础标准建设和技术体系架构上制定出足够支撑国家建立医疗大产业发展的数据标准，而我国目前健康医疗大数据标准还处在刚刚起步阶段，标准的研制进程满足不了日益蓬勃发展的医疗信息化建设的需要。

为促进和规范健康医疗大数据应用发展，推动“健康中国”建设，国务院先后发布多项政策性文件，指引健康医疗大数据标准建设方向[3]，贵阳市人大在 2018 年 10 月发布《贵阳市健康医疗大数据应用发展条例》对数据采集的标准和要求，数据采集汇聚的对象范围，数据汇聚存储的平台载体进行规范，建立医疗检验、检查结果互认机制。作为全国的第一部地方性法规，该条例的出台对于贵阳市健康医疗大数据的采集、医疗大数据的互通共享和应用发展具有相当积极的作用，也为全国健康医疗大数据标准建设提供了生动实践。为贯彻落实条例的要求，推进贵阳市大数据标准化的落地，结合我市实际的应用和发展，组织开展针对我市实际情况的健康医疗大数据标准研究工作。

1、贵阳市健康医疗大数据发展现状

贵阳市于 2018 年启用了贵阳市人口健康云平台（以下简称健康云）。健康云平台是按照国家卫健委信息化建设的“46312”体系框架的要求进行的区域医疗大数据平台。平台从建设标准体系，IT 基础设施，信息安全体系等角度进行构建。从横向、纵向多维度覆盖医疗健康多个环节的功能应用。实现了全市公立医疗机构和基层社区卫生服务中心（乡镇卫生院）等三百家医疗机构的信息化覆盖。打通了市县乡三级公立医疗机构诊疗数据通路，支撑起市县乡村四级基层公共卫生体系，整合了全市妇幼保健数据，儿童预防接种数据等。

2019 年底，贵阳市卫生健康局在“健康云”平台基础上，继续实施了“贵阳市健康医疗大数据治理项目”。项目建成后将能够更加深度和全面的整合全市医疗机构临床诊疗数据，对贵阳市医疗大数据进行收集、清洗、分析、应用、呈现，支撑相关监管部门决策及医疗大数据的综合应用。并形成优良的健康医疗数据为更多需求场景提供坚实的基础。

伴随着这些各类医疗业务平台和大数据平台的实施，在相关系统的整合中，我们发现数据缺陷也逐渐凸显出来。各家方案提供商按照自己的产品体系和建设目标，容易形成各自为政，产生新的系统和数据壁垒，从而大大降低数据互联互通性的要求。因而我们必须从顶层设计层面提出一套完

整的健康医疗大数据标准体系来规范各类医疗系统和各家医疗机构的数据标准，从而提高数据质量，明确数据归口。

2、贵阳市健康医疗大数据标准建设的必要性

国发 22 号文要求到 2020 年二级及以上医疗机构数据要互联互通、检查检验结果要实现互认，这就对各医疗机构信息系统提出了更高的要求，从以前要求院内信息系统实现互联互通到要求各医疗机构之间信息系统互联互通。随着贵阳市人口健康云平台、贵阳市智慧妇幼平台、贵阳市健康医疗大数据治理平台、贵阳市互联网健康服务平台等各大平台的建设和实施，对数据质量的要求越来越高，数据元格式统一的问题、数据分类分级、格式转换等一系列问题呼之欲出，这就要求有一套完善的数据标准体系来规范贵阳市健康医疗大数据的建设。“数据整合，标准先行”，健康医疗大数据标准的制定是贵阳市实施健康医疗大数据的前提，为贵阳市各类医疗信息化平台和医疗机构信息化建设提供参考规范。

3、健康医疗数据的特点分析

3.1 健康医疗数据分类

健康医疗数据可分为医疗服务类、公共卫生类、健康管理类、基因类、管理运行类、安全与隐私类以及一些规则和知识库类。

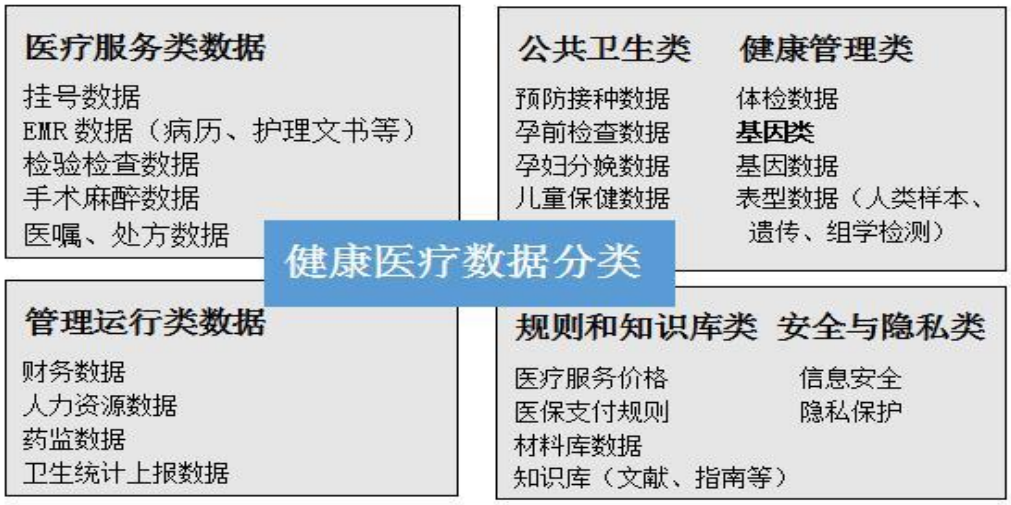


图 1 健康医疗数据分类

与其他行业数据相比较，健康医疗行业数据不但重要而且特殊，不仅仅是与人类的生命、健康、疾病紧密相关，而且具有更多样性和复杂性，以及需要我们探讨研究更多的未知性。健康医疗数据不同于一般的数据，不仅具有大数据 5V 的特点，还具备主观性、时效性和冗余性的特点。

3.2 主数据和参照数据分析

所谓的主数据主要是用来描述医疗业务实体的数据，比如处方，患者基本信息，病历，医嘱等。参照数据是用于将其他数据进行分类或目录整编的数据，是规定数据元的值域范围。参照数据一般是有国标可以参照的，固定不变的，如果没有国标的参照数据，我们也可以根据标准制定规范

自己定义。主数据和参照数据的标准化是我们制定健康医疗大数据标准的核心。

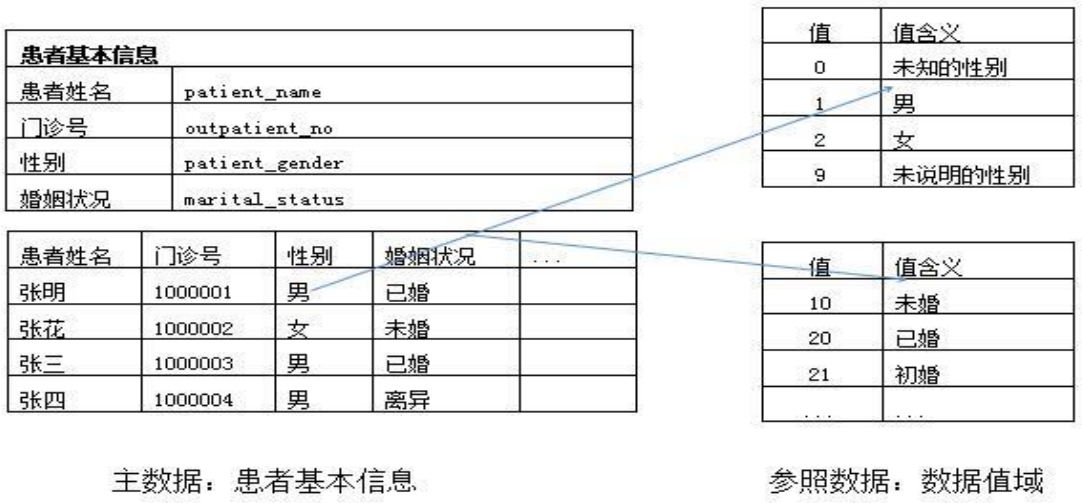


表 1 主数据与参照数据

4、贵阳市健康医疗大数据标准制定

4.1 制定原则

数据标准应在《贵阳市健康医疗大数据应用发展条例》的政策指导下，在原贵阳市人口健康云数据接口标准的基础之上，结合贵州省全民健康信息平台数据共享规范，并参照国际、国家标准以及行业标准，根据贵阳市医疗信息系统建设的实际情况制定，以继承、发展、提高、创新为出发点，全面梳理国际标准、国家标准、行业标准及地方标准，结合大数据技术及产业发展现状与趋势分析[4]，建立适应贵阳市健康医疗大数据产业发展需求的标准体系，该标准应函盖区域卫生，智慧医疗、各医疗机构、健康医疗服务企业、公共卫生、妇幼保健、乡镇卫生院（社区卫生服务中心）、中医馆、村卫生室等所有数据，以面向政府卫生管理需求、面向医疗机构需求、面向公共卫生需求、面向社区系统等方面需求为导向对贵阳市健康医疗大数据标准进行编写，遵循科学性、系统性、协调性、先进性、可扩展性原则。资源整合、统筹规划，以培育并形成完善的健康医疗大数据服务体系为总体目标[5]，明确标准化工作思路、内容及具体的推进措施，整合数据资源，统筹医疗产业规划，破除数据孤岛，强化应用服务，保障贵阳市健康医疗大数据标准服务体系目标清晰、技术可行、结果可见。

4.2 制定研究方法

对现有的国际、国内健康医疗数据标准进行总结研究、对比研究，同时针对贵阳市各级医疗机构实地调查、问卷调查、专家论证等方法统筹规划，制定贵阳市健康医疗大数据标准。

1.资料收集：对目前国外、国家、省级层面的健康医疗数据标准进行收集，采用文献分析结合贵阳市医疗健康信息发展现状进行对比研究、政策分析，进行了吸收、扩展、融合。

2.现场调查：对贵阳市目前 28 家二级以上医院进行现场调查。同时对贵阳市范围内的社区医院、口腔诊所、中医馆、村卫生室发放调查问卷，抽样实地调研。详细了解其信息化建设程度、规划、投入等。

3.健康医疗大数据标准制定：收集各调查问卷和现场调查情况后，结合目前国内外医疗大数据标准进行统计分析因地制宜制定贵阳市健康医疗大数据标准。

4.健康医疗大数据标准制定后进行专家论证，专家评审，评审通过后形成标准，经相关主管部门发布。

4.2 制定技术路线

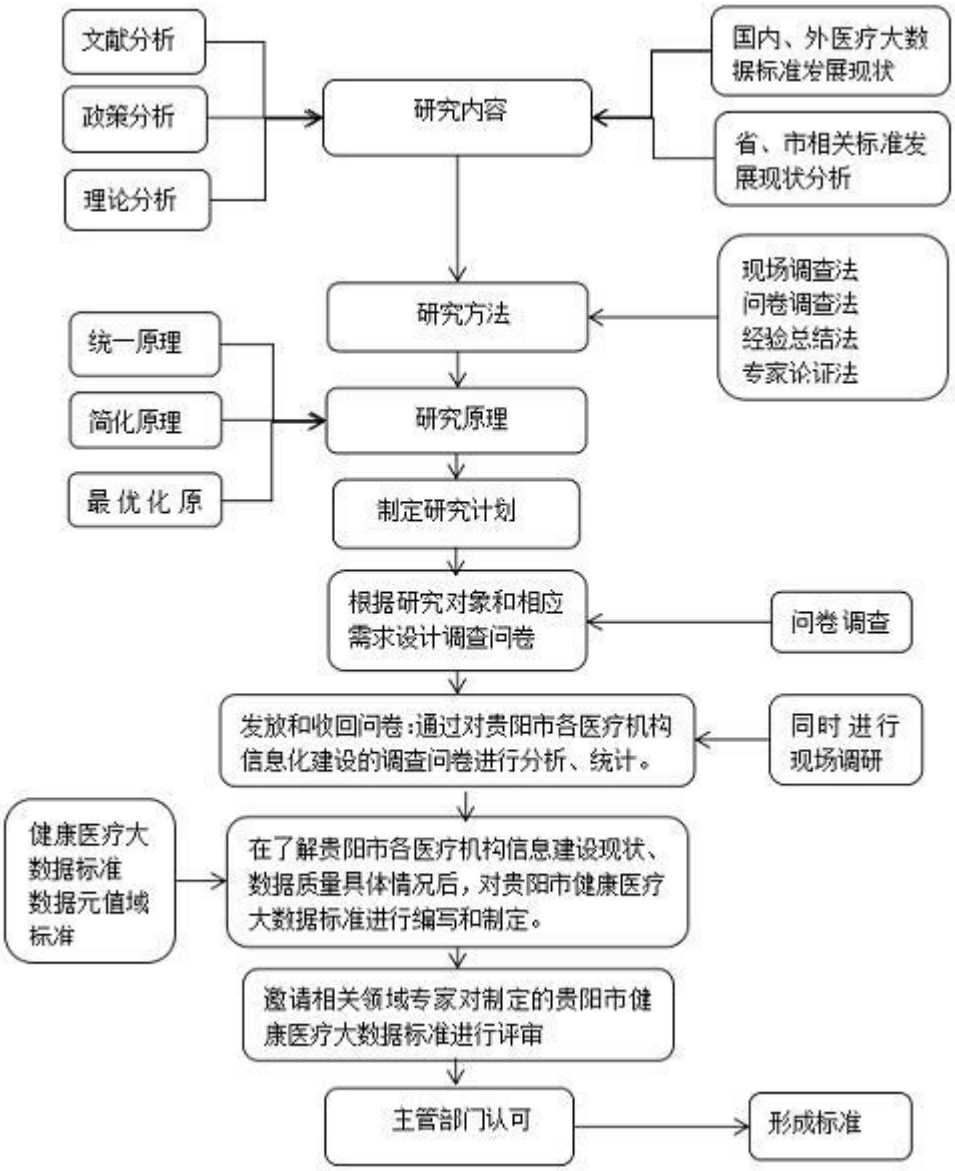


图 2 技术路线图

5、贵阳市健康医疗大数据标准框架

通过对贵阳市健康医疗大数据建设现状和各级政府对健康医疗大数据要求进行分析，贵阳市健康医疗大数据标准应该包括公共卫生、妇幼保健、医疗服务、基因数据等。

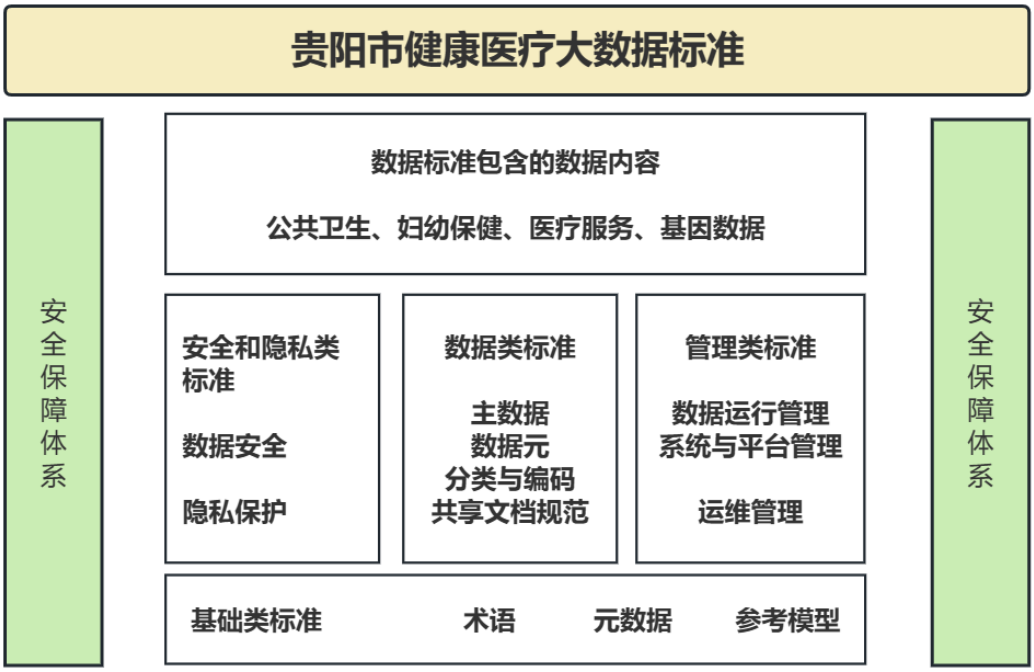


图 3 贵阳健康医疗大数据标准框架

6.1 公共卫生

国发 22 号文要求结合区域全民健康信息平台[6]，实现现有公共卫生信息系统与居民电子健康档案的联通整合，在此次标准制定我们充分考虑了公共卫生数据标准和医疗服务标准的深度融合，主要包括居民健康档案管理服务、重点人群（老年人、糖尿病、高血压、精神病、结核病、孕产妇、儿童等）、签约服务、家庭信息、健康教育服务等。

6.2 妇幼保健

妇幼保健卫生信息化的管理与建设是切实落实国家孕产妇风险防控和分级管理、生育全程基本医疗保健服务等工作要求。它要依托已有的各类医疗卫生信息资源，通过外部数据转换接口在子系统之间、上下级系统之间实现数据传递的互联互通与共享。数据互联互通离不开始数据标准，妇幼保健服务数据标准主要包括妇幼档案基础数据、婚孕检记录、妇女疾病筛查、两癌筛查、分娩信息、出生证、托儿机构及学校等数据的标准规范，为各妇幼保健信息系统数据共享提供参考，同时在制定该标准规范时，我们充分考虑和医疗服务、公共卫生、基因组等数据标准的深度融合。

6.3 医疗服务

医疗服务数据标准是此次数据标准制定的重点部分，由于医疗服务数据涉及机构多，且各医疗机构信息化水平参差不齐，我们根据卫生部门对监管数据的侧重情况对数据采集分为三个等级。一级：属于卫生管理部门要求的核心数据，必须采集；二 级：属于卫生管理部门要求的重点数据，积极采集；三级：属于卫生管理部门要求的扩展数据，根据实际情况采集。医疗服务数据标准包括：门诊、住院、EMR、检验、检查、手麻以及运行管理等相关数据标准。

6.4 基因数据

精准医疗是当前医学界讨论和发展的热点，精准医疗是以个体化医疗为基础，随着基因组测序技术快速进步以及生物信息与大数据科学的交叉应用而发展起来的新型医学概念与医学模式[6]。因此本次数据标准把部分医疗应用层面的人体基因信息与医疗相关的数据纳入此次数据标准统一进行规范和管理，人体基因数据标准包括基因数据（测序信息、生物信息分析、数据质控信息、组学通用信息、数据应用信息）、表型数据（人类样本资源信息、孕产妇样本资源信息、家族史信息、症状信息、生活方式及习惯、组学检查相关信息）等。

7、结语

本文通过分析国内外健康医疗大数据发展现状，结合贵阳市健康医疗大数据建设的实际情况，通过采用现场调研法、问卷调查法、经验总结法、专家论证法等方法对贵阳市健康医疗大数据标准制定进行研究，并制定出符合贵阳市范围内的健康医疗大数据标准，为贵阳市范围内的各类区域医疗平台、医疗机构信息化建设提供统一的数据规范。同时，由于此次标准涵盖的范围较广且贵阳市

各级医疗机构信息化水平参差不齐，尤其是一些社区卫生院、民营诊所等，信息化水平几乎还处于起步阶段，所以该标准在实施上就具有一定难度，需要加大政策引导、和资金的支持等；由于大部分二级及以上医疗机构信息化已初步建成院内数据的互联互通，在标准制定后，各级医疗机构在现有的信息系统建设上是否会升级改造和调整，还是只有新建信息系统参照此标准执行，需要各级医疗机构根据实际情况来规范执行；同时，随着“健康中国 2030”深入开展，大健康范围逐步扩大，相关数据的产生也需要逐步进行规范，这些因素产生的问题需要我们在以后的研究中进行分析和解决。

PU-087

高通量测序在结直肠癌的研究进展

赵思琪[1,2,3]，渠鸿竹[1,2,3]，张昭军[1,2,3]，方向东[1,2,3]
1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室
2. 中国科学院大学
3. 基因组与精准医学检测技术北京市重点实验室

1 引言

CRC 是一种常见的恶性肿瘤，其特征是结肠或者直肠器官中的细胞异常增殖，大多数 CRC 起源于腺瘤，当早期的良性腺瘤无控制的分裂则会进展为高危腺瘤或 CRC[1]。尽管大多数 CRC 起源于腺瘤，但只有部分腺瘤（3%-5%）会发展为癌症，这小部分腺瘤累积在 5-10 年后部分进展为肿瘤[2]。从近五年的研究数据显示，CRC 每年约有 180 万新发病例和 90 万死亡病例[3]。CRC 无论是国内还是国外的新增发病率和死亡率都排名前三，CRC 是全世界最常见的癌症类型之一，直肠癌是 CRC 的一种亚型，占 CRC 的 1/3[4]。结直肠癌的高发病率和高死亡率使得其诊断和治疗仍然存在较大的挑战。

2 高通量测序在 CRC 中的研究概况

2.1 全基因组测序在结直肠癌的研究进展

CRC 是一种异质性基因组疾病，其中各种改变，例如点突变、基因组重排、基因融合或染色体拷贝数改变，都可能导致疾病的发展[5]。*APC* 是一种肿瘤抑制基因，参与 DNA 复制和修复过程，它在 CRC 病例中的突变率为 57%，，已被证明会强烈影响化疗反应[6]。在 2%-20%的 CRC 病例中观察到 *SMAD4* 基因突变，与西妥昔单抗治疗反应不佳相关[7]。某些 *RAF* 突变与诱导基因组不稳定有关，从而促进癌细胞增殖。在近 40%的 CRC 患者中已发现异质 *KRAS* 突变[8]，最常见的是位于 G12C 位置的替代，与抗表皮生长因子受体治疗耐药性相关[9]。

WGS 在 CRC 的早期筛查和诊断方面具有重要作用。通过患者基因组的测序，可以检测与 CRC 相关的遗传突变、基因变异和致病基因。某些基因的突变会增加个体患 CRC 的风险，帮助早期发现潜在的 CRC 患者[10]。例如，几种基因（如 DNA 错配修复基因 *MUTYH* 和 *APC*）以突变形式遗传，会使携带者患 CRC 的风险很高[11]。此外，通过 WGS 检测 CRC 相关的遗传突变，我们可以区分 CRC 与其他疾病或良性肿瘤之间的差异。这对于选择最合适的治疗方案和制定个体化医疗计划非常重要[12]。因此，通过检测与 CRC 相关的基因变异和致病基因，可以确定患者的遗传易感性和风险，并为早期筛查和预防提供依据，从而提高 CRC 的早期检测率，并优化诊断与治疗方案的制定。

2.2 转录组测序在结直肠癌的研究进展

转录组测序捕获细胞或者组织中的 RNA 分子表达信息[13]，在 CRC 防治中，具有广泛的应用[14]。转录组测序技术能高通量测量每个基因的表达水平，帮助确定 CRC 中异常表达的基因，如驱动基因、抑癌基因、转录因子和信号通路相关基因等。通过功能注释和调控网络分析，可以全面了解这些基因在 CRC 发展中的作用方式[15]。基于基因表达谱信息，利用聚类分析、机器学习等方法，可 将患者划分为不同的亚型或子组，进一步研究其特征，为治疗和个体化护理提供指导

[16]。此外，转录组测序可以鉴定 CRC 生物标志物，这些生物标志物可用于早期诊断、疾病分型、疗效监测和预后评估等临床应用。转录组测序还有助于预测患者对特定药物的响应性和抗性[17]。

单细胞转录组测序（scRNA-seq）是转录组测序的一种高精度类型，可在单个细胞内测量基因的表达水平，被广泛用于分析肿瘤的异质性和肿瘤微环境。肿瘤微环境对于肿瘤的发生、进展、侵袭、转移和耐药性发挥着至关重要的作用。肿瘤异质性的分子基础表现为不同的恶性亚群，恶性亚群基于结构变异、染色体重排事件、表观遗传修饰和基因表达特征等因素而存在[18, 19]。scRNA-seq 检查单个细胞的多组学特征，从而在各种肿瘤之间绘制 TME，为准确的分子亚型和精确治疗提供了支持证据[20]，单细胞分析可以用于揭示肿瘤细胞对药物反应的潜在变异性[21]。随着单细胞多组学数据整合能力的不断提高，我们可以更深入地了解各种癌症亚型和细胞亚分类下的多种分子疾病。例如，利用单个细胞的细分群对慢性淋巴细胞白血病进行亚分类，可以为患者提供更加个体化的治疗方案[22]。

2.3 表观组测序在结直肠癌的研究进展

过去十年中，人们在对 CRC 的表观基因组理解方面取得了很大进展。现在认为，在每个肿瘤中都存在数千种表观遗传改变，其中一小部分可能是 CRC 的驱动事件。DNA 甲基化谱是研究 CRC 最广泛的表观遗传特征之一，DNA 甲基化谱包括具有不同分子特征的患者群，被归类为 CpG 岛甲基化表型[23]。表观遗传改变，尤其是 DNA 甲基化，在肿瘤发展的早期阶段以及从正常粘膜到腺瘤性息肉的过渡阶段就会发生。最近进行的全基因组甲基化分析显示，DNA 甲基化可能成为诊断、预后和预测 CRC 治疗反应的有用工具[24]。

结直肠癌患者体内的 DNA 甲基化、组蛋白修饰和非编码 RNA 等表观遗传变化，为个体化治疗、早期诊断以及预后评估提供了新的依据[25]。DNA 甲基化的异常参与了肿瘤抑制基因失活和癌基因过度活化；组蛋白修饰变化也与肿瘤进展和预后相关；非编码 RNA 在结直肠癌的调控中发挥重要作用，可以作为潜在的生物标志物和治疗靶点。综上所述，表观遗传学研究在结直肠癌诊疗中为更好地理解疾病、发现治疗靶点和标志物、实现个体化治疗提供了助益[26]。

2.4 微生物组测序在结直肠癌的研究进展

健康的人类结肠中有数百至数千种不同的微生物物种，它们组成了约 38 万亿个微生物，主要存在于覆盖隐窝的外粘液层内[27]。这些微生物的功能包括消化食物、合成维生素和其他营养物质、解毒代谢物和分泌抗菌产物，并且对肠道稳态和人体健康起着重要的影响[28]。最近的研究表明，微生物组的组成和多样性与肠道疾病，如肠癌的发生有关[29]。低膳食纤维、肥胖、缺乏身体活动以及摄入过多的红肉和加工肉类，可能会改变肠道微生物组的组成，从而增加患结直肠癌的风险[29]。微生物组还通过分泌代谢物和化合物来影响人体健康[30]。最近一项针对小鼠的研究发现，产生乳酸的细菌与增强的肠道干细胞增殖之间存在机制联系[31]。虽然肠道干细胞直接使用乳酸来促进增殖，但微生物分泌的乳酸激活 Paneth 细胞和基质细胞表面的 *GPR3*，从而增加 WNT 分泌，从而促进肠道干细胞干性和增殖。

在人类微生物组中，一些常见微生物可能促进或者预防结直肠癌的发生。例如，带有基因毒性 *pks* 的大肠杆菌、核梭杆菌和产肠毒素的脆弱拟杆菌等微生物与结直肠癌相关[32]。一项研究发现，带有基因毒性 *pks* 的大肠杆菌在结直肠癌患者的结肠中的检出率为 55%，而对照组仅为 19%[33]。微生物能够产生毒素，引起 DNA 损伤来促进结直肠癌的发生[34]。另一些微生物可以防止结直肠癌的发展，如肠道微生物的缺乏与结直肠癌肿瘤发生有关。而许多微生物物种可以通过代谢色氨酸来增强 DNA 修复，抑制细胞增殖，从而减缓结直肠癌的发展[30]。还有一种来自罗伊氏乳杆菌的微生物代谢物，通过不同的机制预防结直肠癌的发生[35]。人类微生物组中的微生物对结直肠癌的发生具有重要影响，一些微生物可能增加患结直肠癌的风险，而其他微生物则可能提供保护作用。通过进一步研究微生物组及其代谢产物的作用机制，为结直肠癌的预防和治疗提供新的方法。

3 讨论

结直肠癌是一种常见的恶性肿瘤，其诊断和治疗方法在不断更新。高通量测序技术在结直肠癌的研究中发挥着重要的作用，在基因组学研究中，高通量测序技术可识别致病基因、突变和遗传变异。这有助于了解结直肠癌的发病机制，发现新的治疗靶点，并为个体化治疗提供依据。转录组学研究通过高通量测序技术揭示癌基因的表达情况、信号通路的异常及其调控机制，也可发现新的生

物标志物，用于早期诊断和预后评估。表观遗传学研究利用高通量测序揭示结直肠癌中的 DNA 甲基化和组蛋白修饰等表观遗传变化，解析肿瘤的表观遗传调控模式，识别肿瘤抑制基因和癌基因的异常表观遗传改变，并为治疗选择和预后评估提供参考。微生物组学研究利用高通量测序技术分析结直肠癌组织中的微生物组成，功能及其与肿瘤之间的相互作用，为肿瘤预防和治疗提供新思路。

综上所述，高通量测序技术在结直肠癌的诊疗研究中取得了重要的进展，为深入了解结直肠癌的发病机制、寻找新的治疗靶点以及实现个体化治疗提供了新的途径，有望为结直肠癌患者带来更好的临床疗效。

PU-088

“双一流”背景下临床医学 ESI 学科预测模型构建及提升路径研究

李日[1]，李文臣[2]，徐行[1]，曲珈锐[1]，伦志军[1*]
1 吉林大学第一医院图书档案馆
2 吉林大学第一医院神经创伤外科

1. 引言：

学科是大学的基本组织单位，是人才培养、科学研究、社会服务的载体，为高等学校实现内涵式发展提供根本支撑。2015 年，我国发布《统筹推进世界一流大学和一流学科建设实施办法》，提出以学科建设为基础，统筹推进世界一流大学和一流学科(简称为“双一流”)建设的要求[1]。目前国际影响力较大的 US NEWS、THE、ARWU、QS 等大学排名，都在学科评价指标体系中强调科研产出对学科排名的重要贡献[2]。国内的学科评价积极对接国际标准，不论是“双一流”建设的遴选入围，还是已开展的五轮教育部学科评估和软科世界一流学科排名，都将科研成果作为学科竞争力的重要标志[3-4]。

Essential Science Indicators(简称 ESI) 是当今全世界范围内关注度、科学性和影响力最大的科研成果评价和分析工具之一，通过 ESI 学科排名，可以实现全球机构在同一标准下的学科评比，是衡量学科世界水平的重要参考指标[5-6]。许多评价体系已将学科是否入围 ESI 排名世界前 1%作为一项重要的指标[7]。如何提高学科的 ESI 排名，进入 ESI 世界前 1%，已经成为各大高校重点关注的内容。目前关于 ESI 排名的研究主要集中于分析某一省份或者某一大学的学科排名以及学科入围的预测[8-11]，对于全球范围内影响临床医学 ESI 学科排名的综合因素尚未有研究。本研究以 ESI 学科中的临床医学学科为例，分析学科中已经进入 ESI 世界前 1%和未进入 ESI 世界前 1%的中外高校的科研竞争力水平，选取 Incites 数据库中的论文发表数量、第一作者发文数量、通讯作者发文数量、篇均被引频次、高被引论文数量、热点论文数量、Q1 区论文数量、Q2 区论文数量、国际合作情况和国内合作情况等指标，研究影响学科进入 ESI 世界前 1%的综合因素，构建学科未来进入 ESI 排名的预测模型，展示一个学科的全景图，为我国对标世界一流大学，寻找 ESI 临床医学学科建设途径，提供参考和借鉴。

2. 研究方法和数据来源

2.1 研究方法

本研究在科研成果评价指标选择方面参照中国科学评价研究中心提出的 RCCSE 评价指标体系及 ESI 学科评价方法，共设立了科研生产力、科研影响力、科研创新力、科研发展力和科研合作力共五个一级指标测算科研产出情况[12]，运用文献计量法和德尔菲专家咨询方法确定二级指标，形成科研竞争力评价指标体系。见表 1。

表 1 ESI 学科科研竞争力评价指标体系

一级指标	二级指标
科研生产力	ESI 收录论文数
	第一作者发文量
	通讯作者发文量
科研影响力	ESI 收录论文篇均被引次数
	高被引论文数量
科研创新力	热点论文数量

科研发展力	**被引排名前 10%的论文数量**
	发表在 Q1 区论文数量
	发表在 Q2 区论文数量
科研合作力	**国际合作论文数量**
	国内合作论文数量

备注：Q1 区和 Q2 区论文为 Web of Science 分区

2.2 数据来源

数据来源于 Incites 数据库，收集时间为 2013 年-2022 年， 机构类型: “Academic”； 学科分类体系: “Essential Science Indicators ”；文献类型: “Article”和“Review ”，研究方向: “Clinical Medicine”。以当期临床医学学科 ESI 机构被引阈值 4181 为标准，以负百分数表示与 ESI 机构被引阈值的差距，选择被引频次与机构被引阈值分别相差 90%、80%、70%、60%、50%、40%、30%、20%和 10%作为分层标准，选取中外高校作为未进入 ESI 世界前 1%的高校样本; 选择 ESI 世界前 1%的中外高校作为临床医学学科世界高水平高校样本。

2.3 统计学方法

本研究的统计学处理采用 SPSS21.0 软件包完成。计量资料进行正态性分布检验，符合正态性分布的资料采用均数±标准差(X±S)表示，不符合正态性分布检验的资料采用中位数（四分位间距）表示，计数资料采用构成比表示。将评价体系中筛选的指标作为自变量，是否进入 ESI 前 1%作为因变量（$Y_{未进入}=0$，$Y_{进入}=1$），构建二分类多因素 Logistic 回归模型。通过回归模型计算的 Logistic（P）得出模型的 ROC 曲线下面积，评价模型构建的预测效果，研究采用双尾检验，检验水平设置为 α=0.05。

3. 结果

3.1 **国内外高校基本情况**

在 Incites 数据库中共检索出 6464 个临床医学高校或机构，本研究共纳入 1696 个高校样本，其中国内高校 149 所，占比 8.8%，国外高校 1547 所，占比 91.2%；进入 ESI 世界前 1%的高校有 649 所，其中国内高校 46 所，占比 7.1%，国外高校 603 所，占比 92.9%， 不同 ESI 水平的样本高校数量见图 1。对比 ESI 临床医学世界排名前三和中国排名前三的高校发文情况，哈佛大学的论文发文量及论文被引频次远远高于其他高校，世界排名前三的高校的篇均被引频次为我国排名前三高校的 2 倍左右，见表 2。

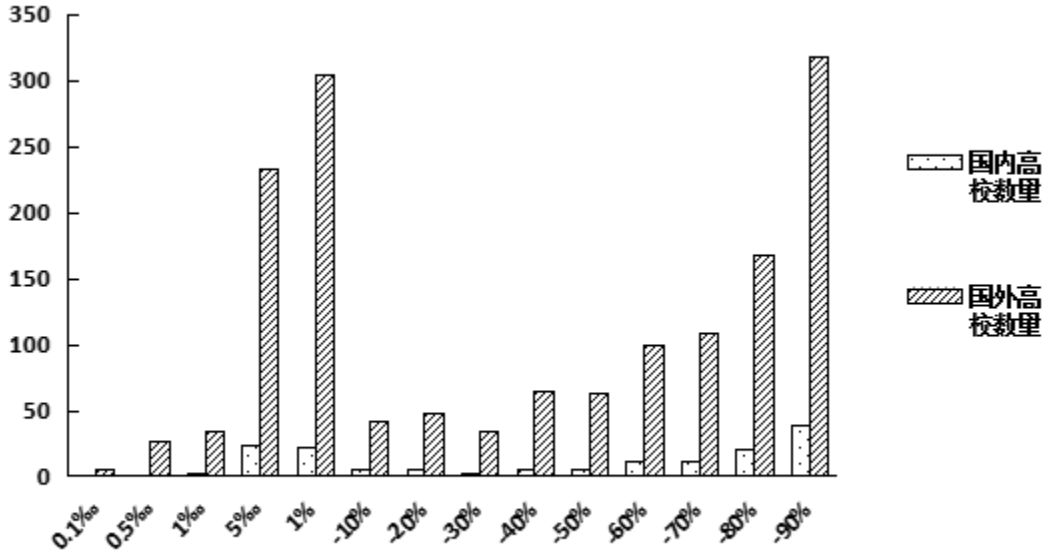


图1 不同ESI水平的样本高校数量情况

表 2 ESI 临床医学世界排名前三和中国排名前三的高校发文情况

排名	中文名称	国家或地区	论文数	论文被引频次	篇均被引次数
1	哈佛大学	美国	1050	3593357	34

				23		
				5275		
	2	多伦多大学	加拿大	6	1603007	30
				4528		
	3	约翰斯·霍普金斯大学	美国	2	1517846	34
	5			3016		
7		上海交通大学	中国	4	519707	17
	7			2473		
3		复旦大学	中国	7	443015	18
	8			2444		
4		中山大学	中国	3	423226	17

3.2 国内外高校“临床医学”学科科研竞争力评价

科研生产力方面，无论学科是否进入 ESI 世界前 1%，国内高校的发文量、第一作者论文占比和通讯作者的论文占比均高于国外高校，差异具有统计学意义（*P*<0.001）。科研影响力方面，无论学科是否进入 ESI 世界前 1%，国内高校的论文篇均被引次数和高被引论文占比均低于国外高校，差异具有统计学意义（*P*<0.001），见图 2。进入 ESI 世界前 1%的高校，国内高校热点论文占比低于国外高校；对未进入 ESI 世界前 1%的高校分析发现，国内高校热点论文占比高于国外，差异具有统计学意义（*P*<0.05）。科研发展力方面，无论学科是否进入 ESI 世界前 1%，国内高校的被引排名前 10%的论文占比和发表在 Q1 区论文占比均低于国外高校，而发表在 Q2 区论文占比国内高校高于国外高校，差异具有统计学意义（*P*<0.05）。科研合作力方面，国外高校的国际合作论文占比高于国内高校，而国内合作论文占比低于国内高校，差异具有统计学意义（*P*<0.001），详见表 3。

表 3 国内外高校“临床医学”学科科研竞争力情况

变量	进入 ESI 前 1%			未进入 ESI 前 1%		
	国内（n=46）	国外（n=603）	*P*	国内（n=103）	国外（n=944）	*P*
科研生产力						
ESI 论文总数	8750（6044, 14233）	4906（2769, 9552）	<0.001	907（546, 1591）	617（310, 1115）	<0.001
第一作者论文占比（%）	0.68（0.62, 0.70）	0.43（0.39, 0.47）	<0.001	0.57（0.49, 0.63）	0.42（0.35, 0.49）	<0.001
通讯作者论文占比（%）	0.67（0.60, 0.70）	0.38（0.34, 0.43）	<0.001	0.51（0.47, 0.57）	0.37（0.29, 0.44）	<0.001
科研影响力						
ESI 收录论文篇均被引次数	14.60（12.50, 17.20）	26.8（21.6, 32.7）	<0.001	11.7（10.0, 14.4）	18.5（13.4, 29.1）	<0.001
高被引论文占比（%）	1.20（0.90, 1.50）	3.00（2.10, 3.80）	<0.001	0.81（0.53, 1.17）	1.77（0.92, 3.35）	<0.001
科研创新力						
热点论文占比（%）	0.05（0.03, 0.08）	0.12（0.07, 0.19）	<0.001	0.00（0.00, 0.06）	0.00（0.00, 0.02）	0.001
科研发展力						
被引排名前 10%的论文占比（%）	12.03（9.60, 13.05）	16.86（13.04, 19.51）	<0.001	9.93（8.09, 13.93）	11.57（8.46, 15.96）	0.017
发表在 Q1 区论文占比（%）	27.94（23.43, 32.69）	47.29（37.41, 52.31）	<0.001	21.98（17.33, 33.90）	36.58（26.32, 46.21）	<0.001
发表在 Q2 区论文占比（%）	30.22（29.09, 31.44）	26.51（24.07, 29.34）	<0.001	30.60（28.20, 32.63）	28.28（23.40, 31.92）	<0.001
科研合作力						
国际合作论文占比（%）	18.70（15.90, 21.53）	43.50（31.73, 53.23）	<0.001	15.74（12.36, 25.49）	40.64（26.77, 56.52）	<0.001
国内合作论文占比（%）	52.53（46.06, 57.31）	45.08（34.82, 53.77）	<0.001	62.08（58.00, 67.54）	47.34（32.23, 58.76）	<0.001

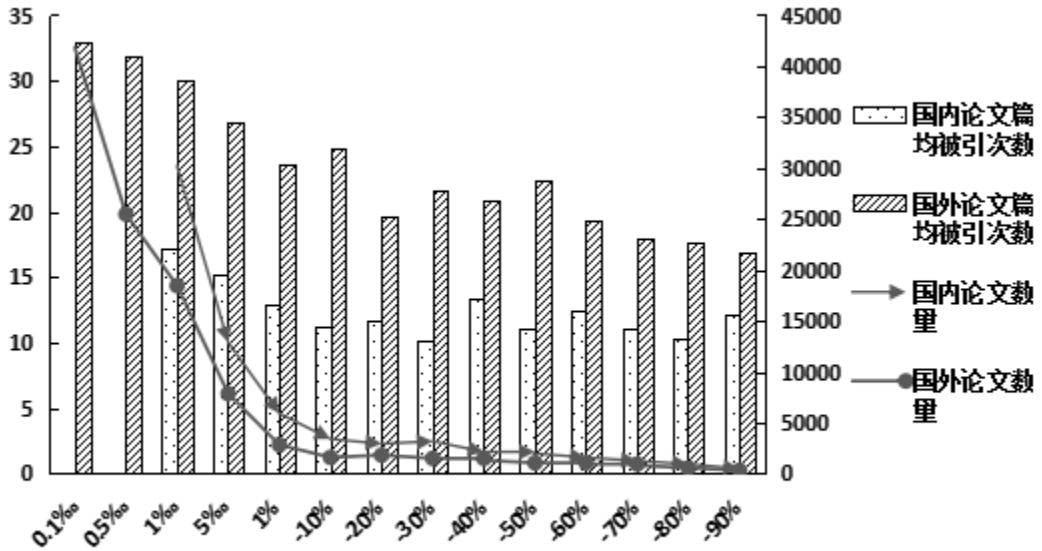


图2 国内外高校发文数量与篇均被引频次

3.3 临床医学 ESI 学科预测模型的建立

为了解影响高校学科排名能否进入 ESI 世界前 1%的综合影响因素，对纳入的变量进行 Logistic 回归分析，结果显示，论文的篇均被引频次、高被引论文数量、热点论文数量、被引排名前 10%的论文数量、发表自 Q1 区论文数量和发表在 Q2 区论文数量是影响高校进入 ESI 世界前 1%的影响因素，见表 3。根据回归模型分别计算出学科进入 ESI 世界前 1%的回归方程：

$$Logistic（P）_{ESI\ 世界前\ 1\%}=-0.21X_{论文总数}+0.007X_{篇均被引频次}+0.286X_{高被引论文数量}-0.388X_{热点论文数量}-1.98X_{被引排名前\ 10\%论文数数量}+0.5X_{Q1\ 论文数量}+0.67X_{Q2\ 论文数量}-11.115$$

表 3 全球范围内 ESI “临床医学”Logistic 回归模型

变量	*B*	Wald	*P*	OR	95%CI	
					下限	上限
论文总数/10	-0.021	1.317	0.251	0.979	0.944	1.015
第一作者论文/10	-0.001	0.001	0.976	0.999	0.956	1.045
通讯作者论文/10	0.027	1.232	0.267	1.027	0.98	1.077
篇均被引频次	0.007	18.183	<0.001	1.007	1.004	1.01
高被引论文	0.286	107.797	<0.001	1.331	1.261	1.404
热门论文	-0.388	10.708	0.001	0.678	0.538	0.856
被引排名前 10%论文/10	-0.198	18.206	<0.001	0.821	0.749	0.899
Q1 区论文/10	0.050	8.374	0.004	1.051	1.016	1.087
Q2 区论文/10	0.067	6.120	0.013	1.069	1.014	1.127
国际合作论文/10	0.009	0.260	0.610	1.009	0.974	1.046
国内合作论文/10	0.007	0.163	0.686	1.007	0.973	1.043
常量	-11.115	131.305	<0.001	0.000		

3.4 临床医学 ESI 学科预测模型的验证

根据 Logistic 回归方程计算每个样本高校学科进入 ESI 世界前 1%的预测概率 Logistic（P）。根据预测值及是否进入的真实值绘制 ROC 曲线，计算得出 AUC $_{学科排名}$=0.996（95%CI:0.995,0.998），预测模型的预测面积均大于 0.5，说明本研究构建的模型预测价值较好，见图 3。

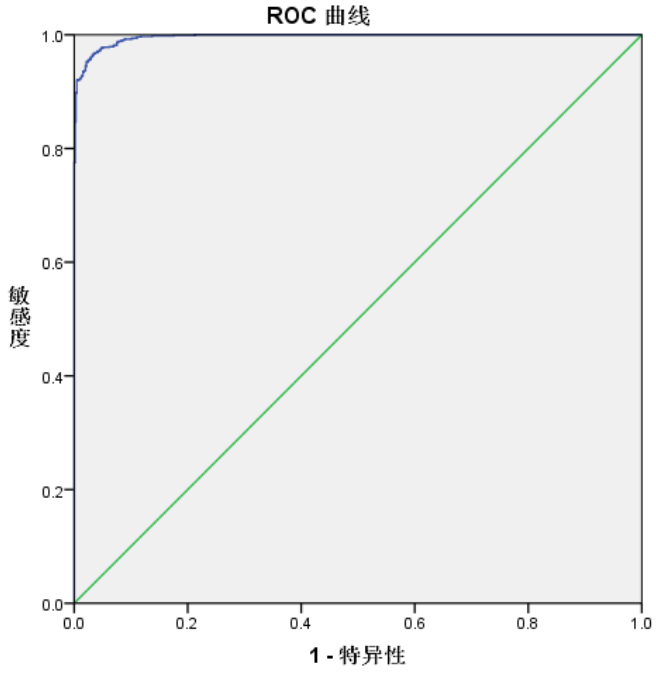


图 3 预测模型的 ROC 曲线下面积

4. 结论与讨论

本研究以 ESI 学科中的临床医学为例，在 Incites 数据库检索共检索到全球 6464 个高校或机构，筛选出 1696 个高校样本，进行 ESI 学科科研竞争力分析，构建高校学科未来进入 ESI 世界前 1%的预测模型，分析国内高校提升临床医学 ESI 学科排名的路径，得出以下结论和建议：

4.1 科学定位学科发展，强化政策支持

ESI 排名前 1‰即为世界一流学科，前 0.1‰即为世界顶尖学科。本研究发现，我国没有高校进入临床医学世界前 0.5‰和 0.1‰，进入世界顶尖学科的高校数量明显不足。排名数据的背后需要高校的政策支撑、资金的支持和高素质人才的保障。“拼指标”背后“拼”的是学校的人才队伍和管理水平。学科的高质量发展离不开政策的保障，近年来，国家和一些省市都出台相应的人才队伍建设计划[13-14]，从政策上给科研人员“松绑”，并给予充足的资金支持，让科研人员能全身心投入科研工作，激发科研工作者的创造热情。高素质人才是学科高质量发展的不竭动力源泉，通过与国内外顶尖高校在人才培养领域合作，开展专项计划培养人才、吸引人才、留住人才，打造出学科在全球领域的“研究高地”。

4.2 提高科研水平，产出高质量成果

本研究发现，无论学科是否进入 ESI 世界前 1%，国内高校的论文发文量高于国外高校，而论文的篇均被引频次却低于国外高校，说明相当一部分高校的临床医学学科存在着“以数量换质量”的畸形发展模式。高校在学科建设中不应简单关注 ESI 名次上的变化，避免学科通过不断发表“低质量”论文带来被引频次的“量变”，而应该通过一定数量“高质量”的被引论文带来被引频次的“质变”[15]。

4.3 聚焦学科研究前沿，提高科研创新力

热点论文是指某学科领域发表在最近 2 年内的论文在最近 2 个月内被引次数排在 0.1%以内的论文。热点论文数是衡量科研机构和学者科研创新力的重要依据[16]。本研究发现，ESI 排名进入前 1%的高校，国内高校热点论文占比低于国外高校。高校应结合自身发展特色与优势，聚焦于当前研究前沿，与国内外高水平机构开展合作研究，在合作化发展中寻求差异，在差异中寻找自己的科研特色，促进临床医学学科原创成果的产出，提高科研创新力。

4.4 加强国际合作，扩大学术影响力

国际合作论文的质和量是一个高校国际影响力的重要体现，国际合作论文占比是衡量高校科研持续发展能力的重要指标[17]。2021 年，爱思唯尔出版集团发布《中国临床医学学术产出及国际合作解读》，数据提示国际合作发表的论文具有更高的引文影响力。本研究发现，国外高校的国际合作论文占比远高于国内高校。国际合作交流在推动学科建设中发挥不可或缺的作用[18]，ESI 排名不区分发表论文的单位顺序，每位作者的贡献度一致，因此，国内高校应加大与国际相关机构的合作，提高合作论文的发文数量，提升合作论文比例，不要过分强调第一作者和第一单位的论

文数量。

5. 结语：

本研究以 ESI 学科中的临床医学学科为例，分析学科中已经进入 ESI 世界前 1%和未进入 ESI 世界前 1%的中外高校的科研竞争力水平，从科研生产力、科研影响力、科研创新力、科研发展力和科研合作力等五个维度研究影响学科进入 ESI 世界前 1%的综合因素，构建学科未来进入 ESI 排名的预测模型。在“双一流”建设背景下，我国高校应科学定位“临床医学”学科发展方向、强化政策支持、提高科研产出水平和科研创新能力、加强国际合作交流，建设我国“临床医学”的世界一流学科和顶尖学科。

PU-089

气管镜介入技术的研究趋势及热点的 VOSviewer 聚类分析

徐行，李日，伦志军（通讯作者）
吉林大学第一医院

气管镜介入技术是一种针对呼吸系统疾病的侵入性诊断和治疗操作，可用于治疗由良、恶性病变引起的复杂气道病变，也可用于诊疗胸膜疾病和肺血管病变[1]。自 20 世纪 90 年代起，气管镜介入技术开始兴起，各种借助气管镜完成诊治技术在临床实践中得到广泛应用，扩大了呼吸道疾病的诊断和治疗范围[2]。近 30 年来，虽然国内外的专家和学者已经对气管镜介入技术的应用进行了大量研究，但对该领域的研究进展仍缺乏讨论。文献计量学通过数学和统计的方法对文献本身包含的引用数据进行分析，所得结果在一定程度上能够反应某一领域的研究趋势[3]。可视化研究是文献计量学的方法之一，其中借助 VOSviewer 软件绘制的可视化图谱具有信息量大、视觉效果好的特点[4]。因此，本研究选用 VOSviewer 软件对 Web of Science 核心合集数据库于 1992—2022 年间收录的气管镜介入技术的相关文献进行计量分析和聚类分析，探索气管镜介入技术领域的研究趋势和热点问题，并明确气管镜介入技术的未来研究方向。

1. 资料与方法

1.1 数据来源与检索策略

本文选取 Web of Science 核心合集数据库作为文献来源，以“Interventional Bronchoscopy”、“Bronchoscopic surgical procedure”、“Bronchoscopic Surgery”为主题词进行检索。文献发表时间限制为出版时间在“1992 年 01 月 01 日至 2023 年 12 月 31 日”，文献类型限定为 Article or Review。检索后，获取相关文献 2021 篇。相关文献的引用数据以“纯文本格式”导入 VOSviewer（1.6.19）软件进行分析。

1.2 数据分析

利用 Endnote 中使用“Subject Bibliography”功能对相关文献的年份、作者、机构、国家/地区、文献来源期刊及被引情况进行分析。利用 VOSviewer 软件对作者、国家/地区、机构及合作情况，文献关键词进行分析，并生成可视化图谱。文献检索和信息处理由 2 名研究者完成，以确保研究结果的准确性。

2. 结果

2.1 年发文量分析及高被引文

气管镜介入技术领域的论文刊发量在近 30 年间呈三种主要趋势。图 1 为 1992 年—2022 年间该领域发文量的变化情况。1995 年前，该领域的年发文量尚不足 10 篇。1996 年—2017 年间，该领域论文数量整体呈增长趋势。自 2018 年起，该领域内论文发刊量呈爆发式增长，2022 年的发文量虽有所下降，但仍高于 2018 年前的发文量。在近 30 年内，2021 年的发文量最多（184 篇）。气管镜介入技术领域内的十大高被引文见表 1。

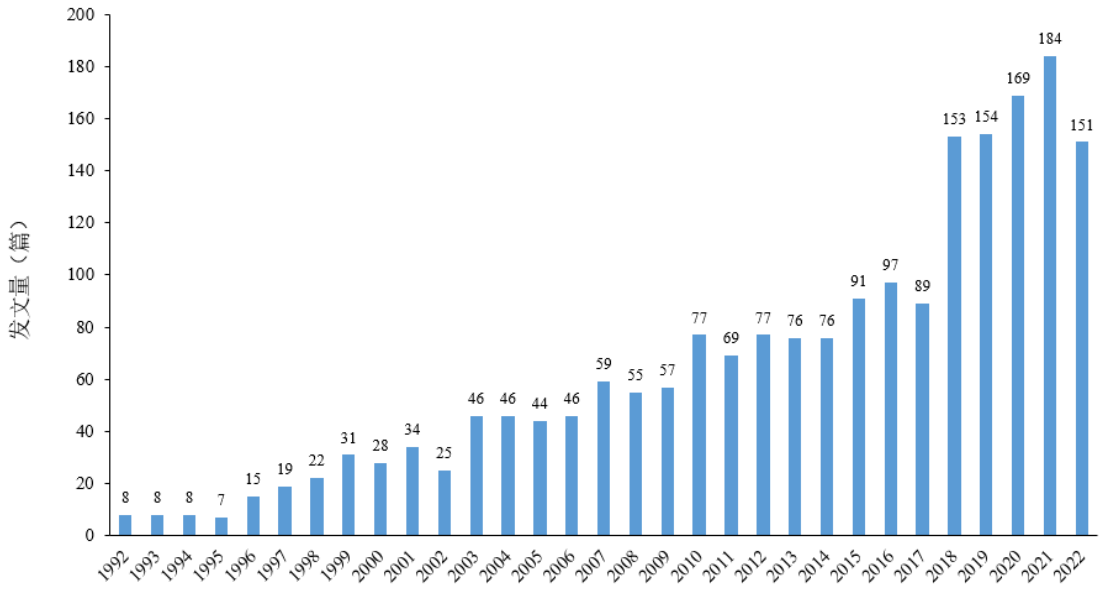


图 1 年发文量统计

Fig.1 Statistics of annual published articles volume

表 1 十大高被引文章

Table 1 Top 10 most-cited articles

作者	标题	期刊	发表时间	被引频率
Castro,等	Effectiveness and Safety of Bronchial Thermoplasty in the Treatment of Severe Asthma A Multicenter, Randomized, Double-Blind, Sham-Controlled Clinical Trial	American Journal of Respiratory and Critical Care Medicine	2010	525
Yasufuku,等	Real-time endobronchial ultrasound-guided transbronchial needle aspiration of mediastinal and hilar lymph nodes	Chest	2004	510
Ernst,等	Interventional pulmonary procedures: Guidelines from the American College of Chest Physicians	Chest	2003	429
Eberhardt,等	Multimodality bronchoscopic diagnosis of peripheral lung lesions - A randomized controlled trial	American Journal of Respiratory and Critical Care Medicine	2007	380
Yoon,等	Bronchial and nonbronchial systemic artery embolization for life-threatening hemoptysis: A comprehensive review	Radiographics	2002	335
Tanaka,等	Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients	Journal of Trauma-Injury, Infection, and Critical Care	2002	328
Freeman,等	A meta-analysis of prospective trials comparing percutaneous and surgical tracheostomy in critically ill patients	Chest	2000	278
Shirato,等	Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy	International Journal of Radiation Oncology Biology Physics	2003	271
Babiak,等	Transbronchial Cryobiopsy: A New Tool for Lung Biopsies	Respiration	2009	268
Bolliger,等	Therapeutic bronchoscopy with immediate effect: laser, electrocautery, argon plasma coagulation and stents	European Respiratory Journal	2006	253

2.2 文献来源期刊分析

为进一步了解发文质量，按发文量高低对文献来源期刊进行排序。表 2 为 1992—2022 年间在气管镜介入技术领域发文量排名前 10 位的期刊。排名前 3 的期刊为“Chest”（5.6%）、“Journal of Thoracic Disease”（4.8%）、“Respiration”（4.4%篇）。

表 2 主要期刊文献数量

Table 2 Number of papers in major journal

来源期刊	发文量（篇）
Chest	113

Journal of Thoracic Disease	97
Respiration	88
Respirology	76
Annals of Thoracic Surgery	46
European Journal of Cardio-Thoracic Surgery	40
BMC Pulmonary Medicine	31
Journal of Bronchology & Interventional Pulmonology	29
Journal of Thoracic and Cardiovascular Surgery	29
Medicine	28

2.3 高发文国家/地区分析

1992 年—2017 年，在气管镜介入技术领域中发文量最大国家为美国（579 篇），德国（227 篇）和中国（196 篇）位于第二位和第三位（图 2）。利用 VOSviewer 对发文量≥20 篇的国家进行合作分析，结果显示中国同多个国家和地区在气管镜介入技术领域开展了合作研究，其中与美国合作最多，其次为德国、澳大利亚（图 3）。密度视图结果显示美、中、德、日四国在在气管镜介入技术领域研究较为集中（图 4）。

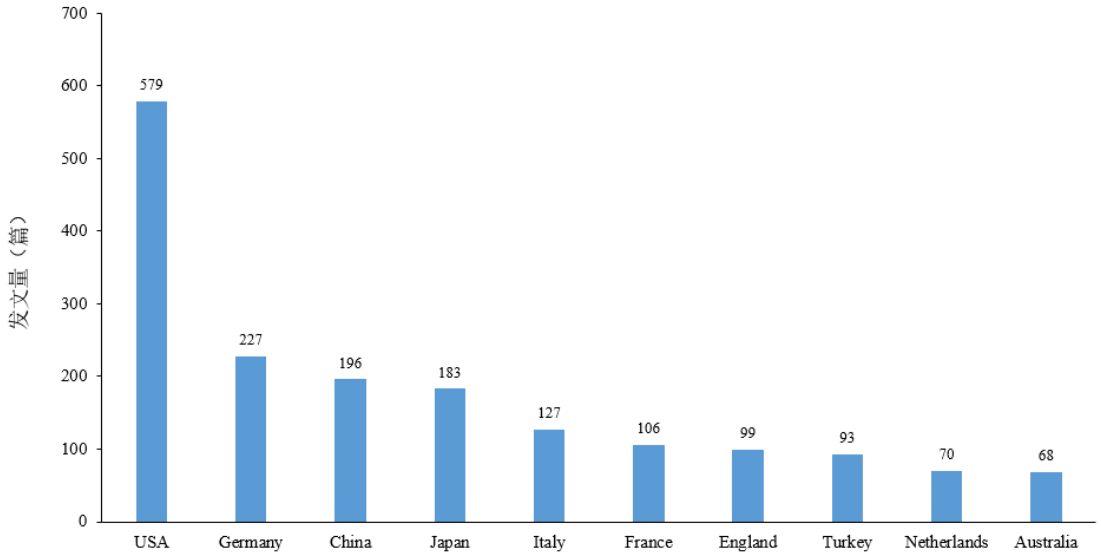


图 2 发文量前十名国家/地区

Fig.2 Top 10 countries/regions with the largest amount of published research

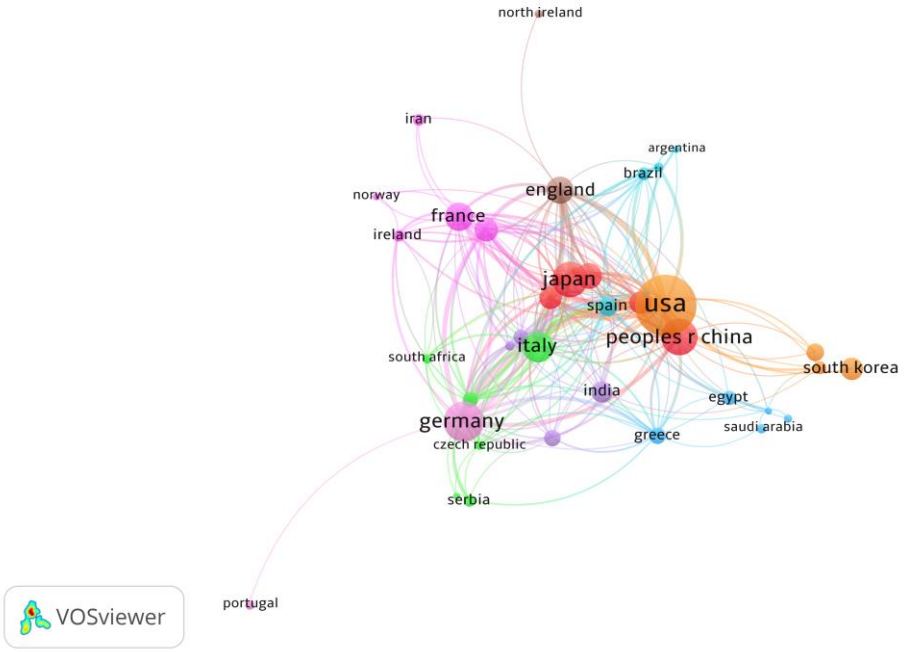


图 3 国家或地区合作的可视化图谱

Fig.3 Visual map of countries or regions cooperation

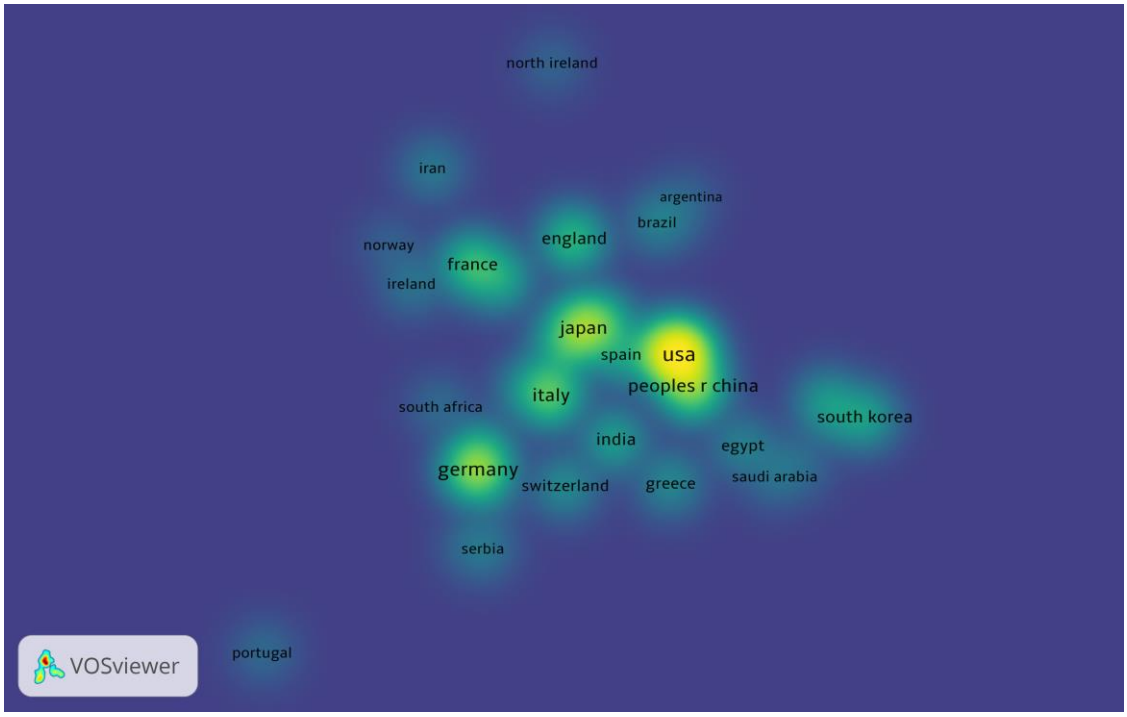


图 4 国家或地区合作密度的可视化图谱

Fig.4 Visual map of countries or regions cooperative density

2.4 论文作者、机构及合作网络分析

在 2021 篇文献中共有 8659 名作者，发文量≥20 篇的作者 166 名。发文量排名前三的作者是 Herth, Felix J. F.（58 篇）、Eberhardt, Ralf（32 篇）和 Slebos, Dirk-Jan（30 篇）。发表量前十

的其他作者见表 3。利用 VOSviewer 对发文量≥20 篇的作者进行合作分析，可视化图谱显示在支气管技术领域以 Herth, Felix J. F.为中心、以 Slebos, Dirk-Jan 为中心、以 Yarmus, Lonny B.为中心、以 Ernst, Armin 为中心的团队内部合作关系比较密切，但各关键作者之间的合作关系强度不一（图 5）。

表 3 发文量前十的作者

Table 3 Top 10 authors with the largest amount of published research

作者	发文量（篇）
Herth, Felix J. F.	58
Eberhardt, Ralf	32
Slebos, Dirk-jan	30
Yarmus, Lonny B.	27
Ernst, Armin	24
Lee, Hans J.	22
Wahidi, Momen M.	22
Shah, Pallav L.	20
Dutau, Herve	17
Majid, Adnan	17

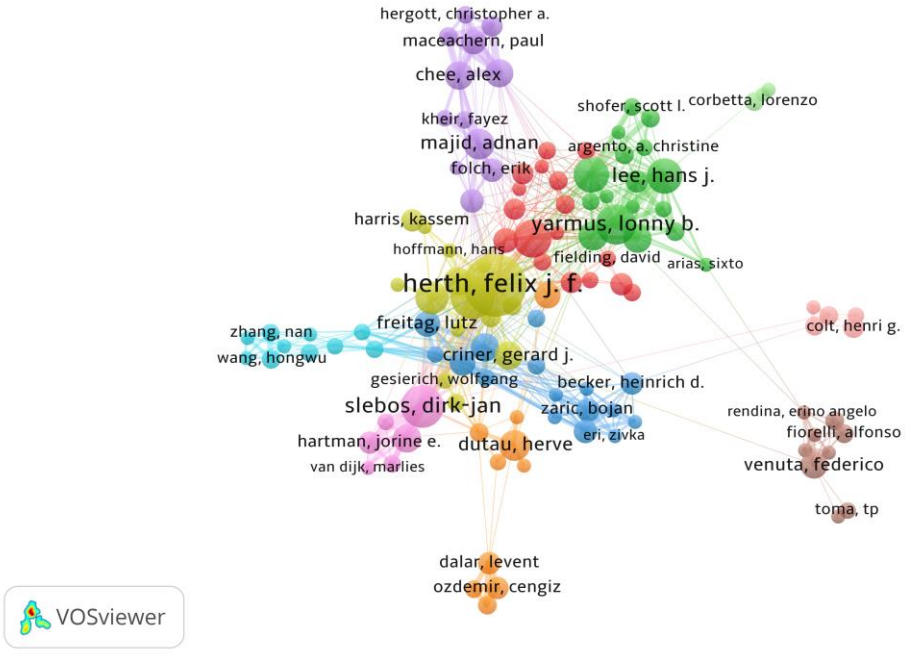


图 5 作者合作的可视化图谱

Fig.5 Visual map of authors cooperation

在 2021 篇文献中共涉及 2245 个研究机构，发文量≥20 篇的研究机构有 62 个，德国海德堡大学（56 篇）、美国约翰·霍普金斯大学（42 篇）、荷兰格罗宁根大学（33 篇）位于前三位（表 4）。利用 VOSviewer 对发文量≥20 篇的研究机构进行合作分析，可视化图谱显示在支气管技术领

域内研究机构间合作以区域内合作为主，隶属中国的研究机构与其他地域的研究机构开展的合作相对较少（图 6）。

表 4 发文量前十的研究机构

Table 4 Top 10 research institutions with the largest amount of published research

机构	发文量（篇）
Heideberg University	56
Johns Hopkins University	42
University of Groningen	33
Duke University	29
Harvard University	27
Mayo Medical School/Cleveland Clinic	25
Harvard med sch	24
Vanderblit University	20
University of Texas MD Anderson Cancer Center/Sungkyunkwan University	19

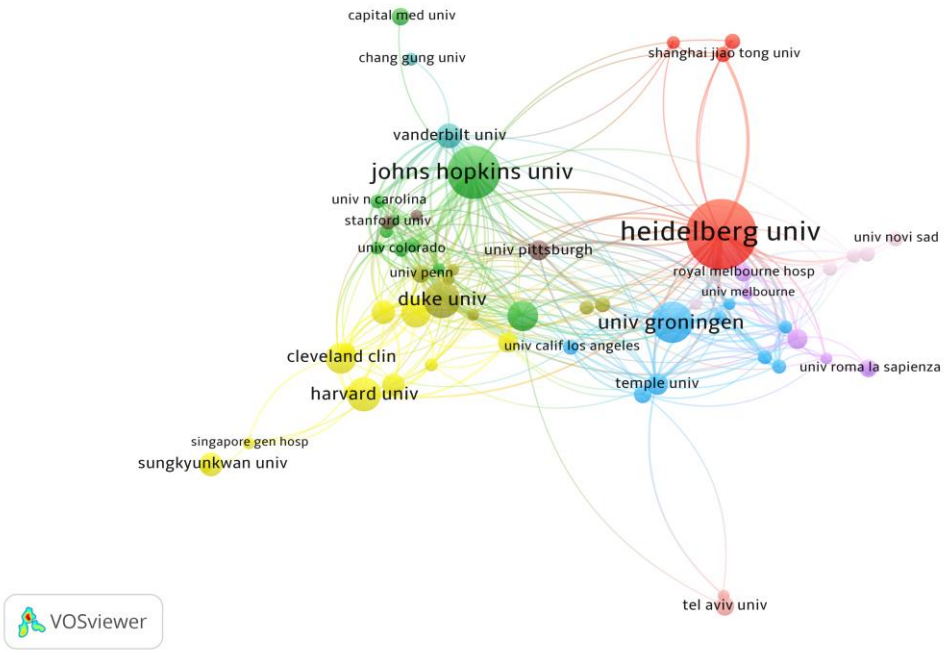


图 6 研究机构合作的可视化图谱

Fig.6 Visual map of research institutions cooperation

2.5 关键词聚类及热点主题分析

为使分析结果严谨，先对文献中的关键词进行整理，将关键词的大小写、全称简写、单复数进行统一，如“tumour”“tumor”统一为“tumour”，“bronchoscopic biopsy”“transbronchial biopsy”统一为“biopsy”。利用 VOSviewer 软件对整理后出现频次≥20 次的关键词进行共现分析，绘制可视化图谱。图 7 和图 8 分别为关键词共现网络图谱和时间叠加图。根据关键词共现聚类结果可将目前气管镜介入技术领域的研究方向分为四类，红色区域：经支气管镜活检对肺癌、结核的诊断；绿色区

域：经支气管镜介入手术的麻醉及通气；紫色区域：经支气管镜介入治疗恶性肿瘤所致气道狭窄；蓝色区域：经支气管镜肺减容术治疗肺气肿及预后。根据关键词时间叠加共现聚类图谱，介入性肺病学、支气管内膜炎、肺过度通气是气管镜介入技术领域最新研究热点。

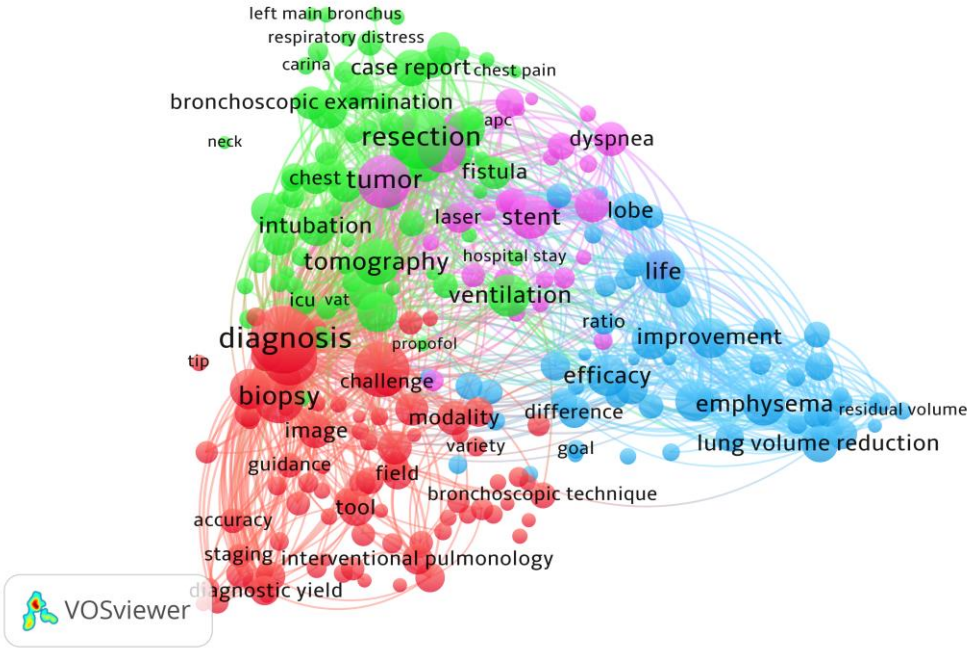


图 7 关键词共现聚类图谱

Fig.7 The co-occurrence network of key words

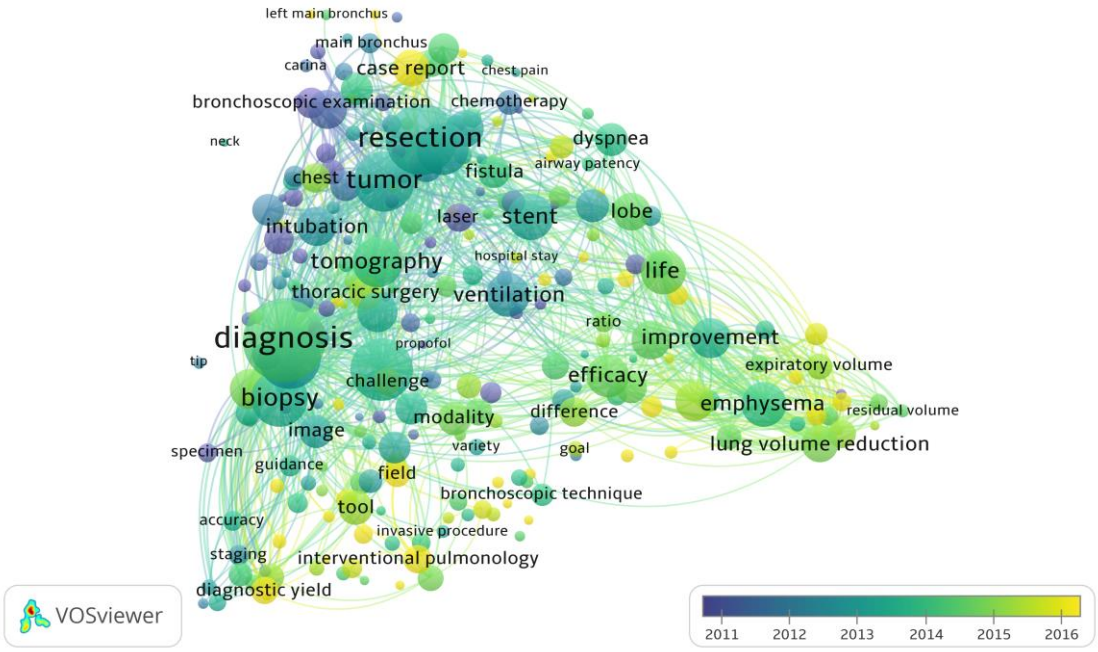


图 8 关键词时间叠加共现聚类图谱

Fig.8 Co-occurrence analysis of key words and time superposition

3. 讨论

本研究对 1992 年至 2022 年间气管镜介入技术领域的论文进行了文献计量学分析，并探讨了该领域的国际研究趋势和热点研究问题。发文量的增减变化能够反映出某一领域内知识量的变化情况和研究者的研究进度，是衡量科学研究发展的一项重要指标[5]。1995 年前，气管镜介入技术领域的论文年发量不足 10 篇，这可能是当时该领域缺乏关注所致。1996 年—2017 年，进入了该领域研究的发展期，年均发文量基本持稳定增长状态，这是由于 1995 年国际上正式对“介入性肺病学”提出了明确定义，即应用内窥镜在胸腔和气道内进行较普通内窥镜诊治更深入的诊疗过程[6]。介入性肺病学包含诸多内容，如经气道壁针吸活检技术、气道内微小球囊置入治疗大咯血、气道内化疗技术等，此类治疗往往需借助气管镜完成。随着对介入性肺病学研究的深入，全球气管镜介入技术领域的论文也随之开始增多。2018 年起出现了爆炸式增长，年发文量达到以往的两倍左右，表明现已有大量专家学者参与了这一领域的研究。

中国在气管镜介入技术领域尽管已成为发文量排名前三的国家之一，但高被引文章前十名中并无中国学者发表的论文，这说明中国学者发表的论文质量仍需提高。此外，高发文量前十的机构和文献来源期刊也均不隶属于中国。国内大陆地区高发文量≥20 篇的研究机构中仅有上海交通大学、复旦大学和首都医科大学与国外的研究机构建立了跨地域的科研合作，但研究合作仅局限于德国海德堡大学、美国约翰霍普金斯大学和美国范德堡大学，这说明我国研究机构同外部的科研合作较少，同时国内研究机构也未建立紧密的合作关系，这可能与我国在介入性肺病学领域开展研究的起步时间较晚有关。未来我国研究者应考虑在气管镜介入技术领域开展更多的多中心研究，同时加强与领域内优势研究机构的合作，从而提高我国气管镜介入技术领域的学术国际影响力。

应当注意的是，本研究仅统计分析了 WOS 核心合集数据库收录的气管镜介入技术领域的研究文献，这可能会遗漏报告在会议论文集、书籍和灰色文献中的研究最新进展。同时，本研究未分析出现频次低的关键词，这类关键词中仍可能包含未来的潜在研究热点。

综上所述，本文利用文献计量学法，通过 VOSviewer 软件探究了气管镜介入技术领域的研究趋势和热点。结果显示近 5 年内气管镜介入技术领域的发文量呈爆发式增长，研究范畴涉及经支气管镜活检对肺癌、结核的诊断、经支气管镜介入手术的麻醉及通气、经支气管镜介入治疗恶性肿瘤所致气道狭窄、经支气管镜肺减容术治疗肺气肿及预后。中国相比于欧美国家在气管镜介入技术领域内的跨地域合作相对薄弱，研究团队建构较为单一，研究成果质量不高，应当从多方面予以加强。

表 1 十大高被引文章
Table 1 Top 10 most-cited articles

作者	标题	期刊	发表时间	被引频率
Castro,等	Effectiveness and Safety of Bronchial Thermoplasty in the Treatment of Severe Asthma A Multicenter, Randomized, Double-Blind, Sham-Controlled Clinical Trial	American Journal of Respiratory and Critical Care Medicine	2010	525
Yasufuku,等	Real-time endobronchial ultrasound-guided transbronchial needle aspiration of mediastinal and hilar lymph nodes	Chest	2004	510
Ernst,等	Interventional pulmonary procedures: Guidelines from the American College of Chest Physicians	Chest	2003	429
Eberhardt,等	Multimodality bronchoscopic diagnosis of peripheral lung lesions - A randomized controlled trial	American Journal of Respiratory and Critical Care Medicine	2007	380
Yoon,等	Bronchial and nonbronchial systemic artery embolization for life-threatening hemoptysis: A comprehensive review	Radiographics	2002	335
Tanaka,等	Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients	Journal of Trauma-Injury, Infection, and Critical Care	2002	328
Freeman,等	A meta-analysis of prospective trials comparing percutaneous and surgical tracheostomy in critically ill patients	Chest	2000	278
Shirato,等	Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy	International Journal of Radiation Oncology Biology Physics	2003	271
Babiak,等	Transbronchial Cryobiopsy: A New Tool for Lung Biopsies	Respiration	2009	268
Bolliger,等	Therapeutic bronchoscopy with immediate effect: laser, electrocautery, argon plasma coagulation and stents	European Respiratory Journal	2006	253

PU-090

“互联网+教育”视域下知识图谱在医学教育的研究进展

徐瑾[1,2] 王亚军[1] 卢洁[1]

1.首都医科大学宣武医院 教育处

2.首都医科大学宣武医院 图书馆

近年来，习总书记曾多次指示深入研究“互联网+教育”的本质和内涵[1]。利用信息化手段扩大优质教育资源的覆盖面，支撑教育高质量发展成为当今医学教育的工作重点。“互联网+”正在触发医学教育教学模式变革，知识图谱成为“互联网+教育”视域下医学教育研究的新思路。

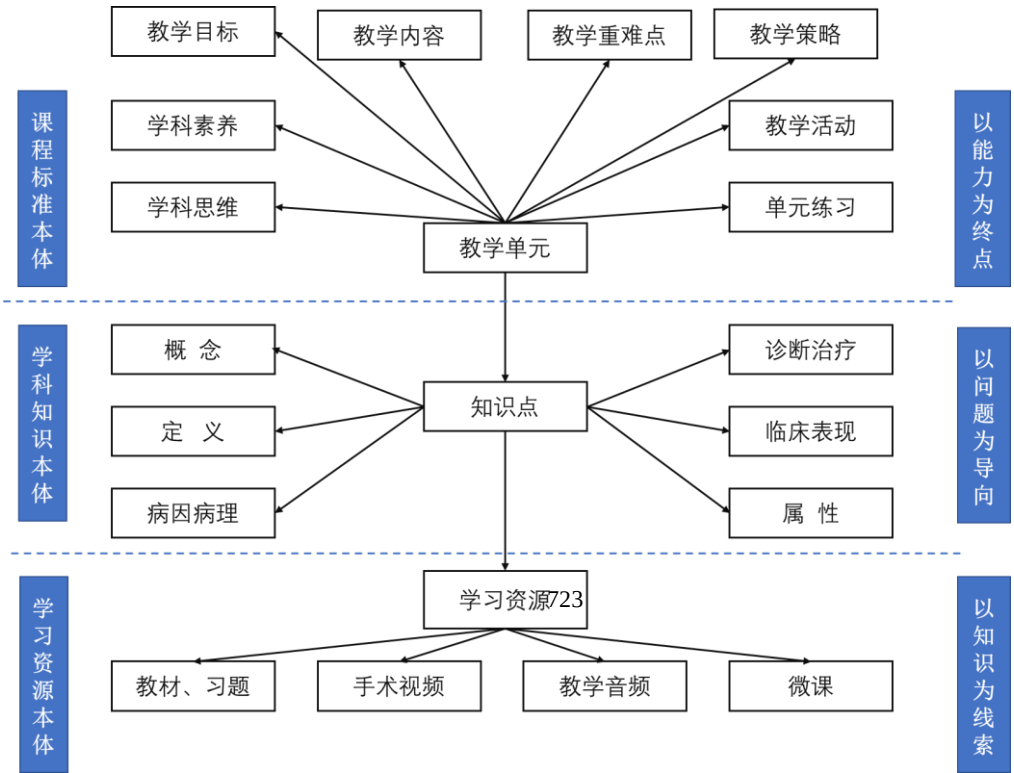
一、教育知识图谱的内涵及外延

知识图谱使用实体、实体属性和实体关系的表述方式描述客观世界的各个概念实体及其之间的联系[2]，它作为一种知识工程领域的人工智能技术，可有效地处理和分析数据，重构知识之间的联系，为数据决策提供支持。教育知识图谱在行业内尚未形成统一定义[3-5]：从狭义上讲，以学科资源作为研究对象，教育知识图谱可以准确地表征各类知识点及其关联的教学资源之间的联系，起到对学科资源进行建模与管理的作用。学生作为研究对象主体，教育知识图谱可建构学生的学情信息，准确表征学生认知状态。通过将学习资源与学情信息的有机融合，教育知识图谱通过生成多种智能教育应用提升教育教学质量[6]。从广义上讲，教育知识图谱既能传授传统意义上的学科知识，同时可对教学场景中产生的各类教学资源以及教学数据信息进行加工管理。该技术打破传统学科知识点与学科资源之间的联系的静态知识图谱模式；还可以建设侧重于学科知识点与教学过程的活动、教学事件之间联系的动态知识图谱。

二、医学教育知识图谱构建技术的框架及构建技术

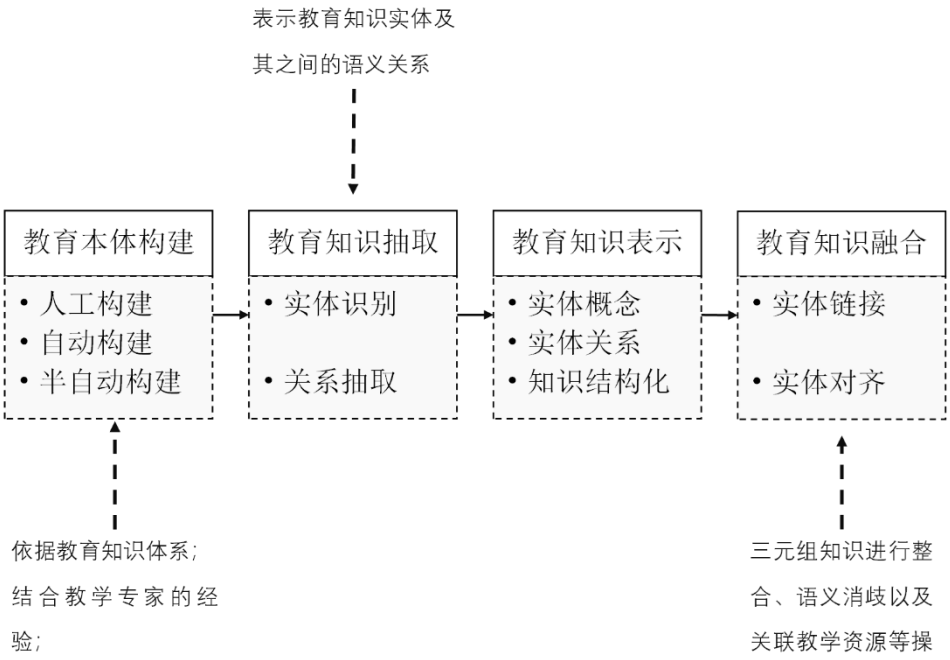
教育知识图谱的构建设计应从教育教学需求和目的出发，包括“自下而上”和“自上而下”两种构建模式。在初次尝试构建或构建多方面知识融合的通用知识图谱时，对知识图谱的整体框架尚不明确，通常采用“自下而上”的模式；即从具体的知识实例入手，通过收集知识实例并进行分析和提炼，从而逐渐明确整体框架和知识模式。而对于整体框架、知识内容明确且具备一定知识深度的领域知识图谱构建，一般采用“自上而下”的构建模式，即构建架构后填充知识实例。对于高等医学院校构建虚拟教研室，打造学科间的知识图谱则应考虑“自上而下”构建方式[21]。

医学教育数据的组成较为复杂，除传统书本教材之外，还包含图片、视频等由多媒体组成的非



结构化数据，专业网站上医学资源组成的半结构化数据，以及由医学数据库组成的结构化数据[22-24]。构建医学教育知识图谱，需要通过知识图谱构建模型训练得到包含不同专业或不同教研室间关系的三元组形式。其具体构建流程见图 1。

图 1 教育知识图谱构建流程

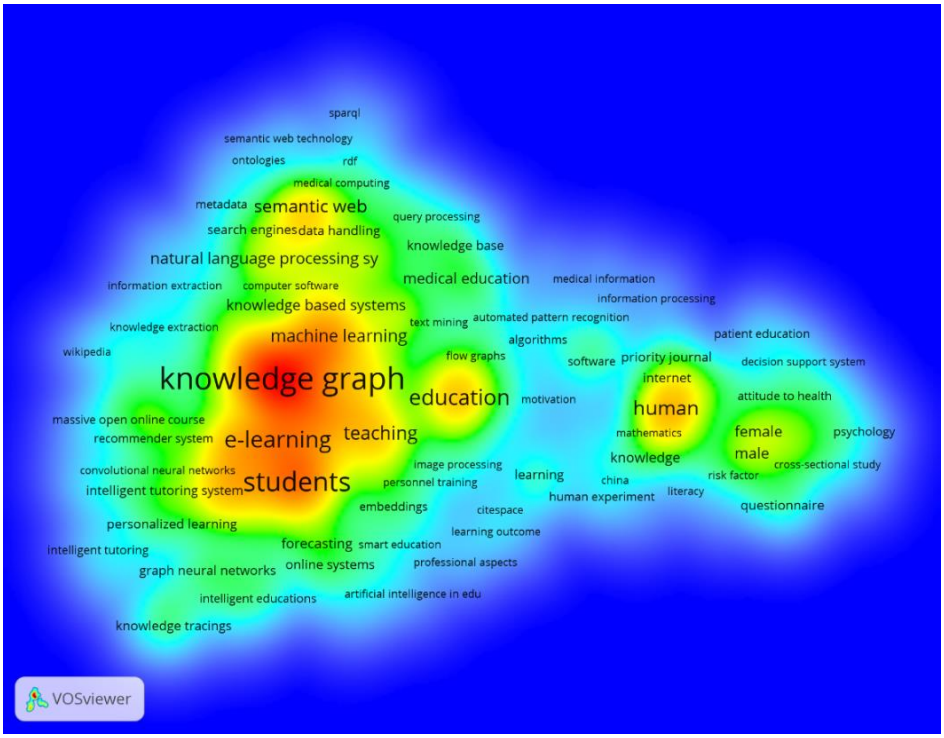


教育知识图谱的构建多以文本数据为主，多媒体等数据为辅。对实体信息的准确表征通常通过对文本数据进行自然语言任务处理，多媒体数据进行辅助特征。在实际医学教学过程中，为满足“智慧教学”应用的多元化需求，扩展教育知识图谱的表现形式，多媒体数据通常会被连接到文本知识实体中，对知识进行补充说明。本研究通过对教学过程中的多媒体等资源的处理，绘制教育知识图谱的构建技术流程图（图 2）。

图 2 医学教育知识图谱以课程为单元的构建流程图

二、知识图谱在“互联网+教育”中相关实践应用

笔者通过对知识图谱在“互联网+教育”中应用在 Scopus 数据库中进行检索，检索策略如下：



(TITLE-ABS-KEY(education OR edu*) AND TITLE-ABS-KEY(Knowledge Mapping OR Knowledge Graph*) AND ALL(web OR internet OR online))。由于知识图谱概念从 2012 年提出，因此年份限制在 2012 年至今（检索日期为：2022 年 11 月 28 日），对检索到的 896 篇文章进行筛选，最终 826 篇纳入本文分析对象，使用 Citespace、VOSviewer 软件进行可视化分析（图 3）。

图 3 知识图谱在“互联网+教育”中的应用关键词共现图

由关键词共现图（图 3）可知，随着人工智能以及互联网的发展，越来越多的研究转向了机器学习以及各种教学平台的研发，通过对学科类别进行分析，多集中在数学、地理学、工程学、软件、信息抽取等领域。目前知识图谱已经开始在医学教育教学中崭露头角。通过对词云的梳理，中国已经成为世界上知识图谱技术应用于教育中的先驱者。使用 Python 对关键词的频次进行统计，当关键词出现频率大于等于 5 次计数 1，共得出 316 个高频关键词，取出现前 20 的关键词使用 SPSS 进行分析，绘制出现前 20 关键词的层级聚类图（图 4）。可知在聚类关系上：知识图谱与学生在线课程的关系更为紧密，对于知识图谱的研究已经从理论的研究转向至教育教学的应用。在技术层面上多使用机器学习和文本挖掘技术。

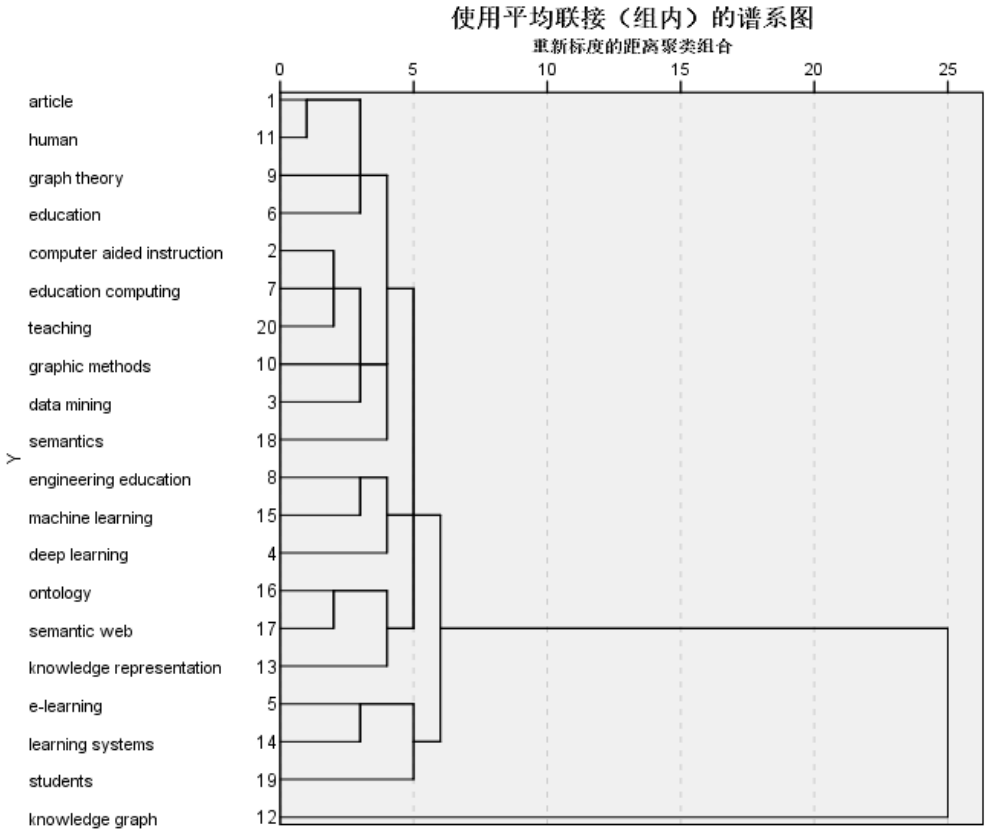


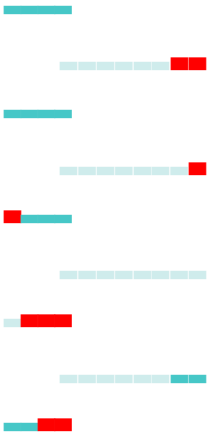
图 4 知识图谱在“互联网+教育”中高频关键词层级聚类图

对 2012 年至今的关键词进行突现度（表 1）统计。在知识图谱在教育领域发展的十年中，知识图谱技术的发展依附于信息及互联网技术的发展；随着计算机可视化技术的更新迭代，从原始的技术领域也逐步应用于计量学研究以及制作概念图。随着大数据的发展，人工智能、机器学习广泛深入地应用于教学之中。全球范围内正共同致力于将高精尖的信息技术投入于教育领域，以期获得在教育教学领域的长足进步，进而使得学生能够更好更方便地获取知识。

表 1 知识图谱在“互联网+教育”中的应用十大关键词突现表

Keywords	Year	Strength	Begin	End	2012 - 2023
信息 information	2013	4.43	2013	2016	
互联网 internet	2013	3.43	2013	2019	
知识建构 knowledge construction	2014	2.46	2014	2017	
卫生保健 health care	2015	2.48	2015	2019	
科学教育 science education	2016	3.86	2016	2018	
概念图 concept map	2012	2.98	2017	2019	

教育技术 educational technology	2018	6	2.8	2018	2019
合作学习 collaborative learning	2019	3	2.3	2019	2020
人工智能 artificial intelligence	2021	7	2.5	2021	2023
知识图谱 knowledge graph	2018	1	2.5	2022	2023



通过对知识图谱的计量学分析以及聚类分析，笔者将其归纳为以知识为线索、以问题为导向、以能力为终点的知识图谱的构建（表 2），不同的应用场景及应用目的选择不同类型的构建技术,对知识图谱的构建方法、特性利用程度、实验设计、系统评价效果等方面来对知识图谱在医学及教育中的应用进行综述。

表 2 知识图谱在医学及教育领域中的应用分析

构建需求	医学/教育类型	知识图谱构建技术及特性
以知识为线索	教育资源推荐	知识存储、知识匹配、可视化交互
	知识检索/获取	实体识别、关联分析、链接预测
	知识点问答	本体构建、关联分析、路径分析
以问题为向导	学习路径分析	本体构建、知识链接
	考试/学习效果预测	知识点图谱构建、认知诊断
以能力为终点	课程设计	关联分析、链路预测
	教案评估	知识匹配、关联分析、路径分析

2.1 教育资源推荐

如何在海量课程资源中为学生提供个性化内容成为当今“互联网+教育”时代无法忽视的问题。学生缺乏对整体知识结构的深入了解，互联网学习资源良莠不齐，学生课程选择困难，信息过载、学习不成体系从而导致课程通过率低[7]。近年来国内外学者通过学习新闻头条、网络视频、“自媒体”短视频等多种平台的个性化推荐技术，有效解决了各领域信息过载问题，从而实现教育资源精准推荐[8]。Xia [9]提出一种基于深度学习技术构建知识图谱的学习资源精准推荐模型，这一模型的学习资源类型不受限制，通过学生的基本信息、学习资源信息的加工可以为目标学生生成推荐列表。Xu[10]通过使用知识图谱将项的语义信息嵌入到低维语义空间中，计算推荐项之间的语义相似度，然后将该项语义信息融合到协同过滤推荐算法中，可提升语义级别的推荐性能。结果表明，与传统推荐算法相比，该算法能够有效地为学生推荐课程，在精度、召回率和 F1 方面具有更高的数值。

2.2 知识检索、获取以及知识点问答

学生在线课程的学习过程中只是单纯被动接收零散、冗余的信息知识,缺乏直观系统的知识梳理，为学生记忆与理解增加了难度。零散的知识点使学生通常不能建立清晰的知识框架，学生没有在有效期内进行复习、整合、梳理建立知识结构，就难以达到学习目标和预期的学习效果。Yang [11]基于 ERNIE 的水利学科实体关系联合提取模型 WDERJE。WDERJE 中的联合提取用于提取输入语句中存在的实体和实体间关系，利用知识图谱有序、完整地涵盖了水利学科各个层次的知识，以目标-课程-知识单元的结构化层次结构，通过深入挖掘课程之间的实体间关系建立课程链接，从

而帮助水利学生根据自己的目标制定个性化的课程计划。有组织的知识单元提高了知识的可读性和利用率，方便学生高效完成课前预习、课上学习、课后复习的学习任务。

2.3 学习路径分析

教育学家奎克曾指出：教育应该按照时间和顺序，遵循自然方法，从易到难进行[12]。连接主义理论曾指出：如果把知识化为节点，则学习是一个不断连接知识节点或资源的过程[13]。知识图谱可结合学生的领域知识结构与认知结构，通过虚拟网络形成符合学生的学习需求的学习路径。Shi [14]设计一个多维度知识图谱框架，提出了基于多维度知识图谱框架的学习路径推荐模型。该模型以加权系数评分的学习路径选择方法生成明确的学习路径，为学生合理安排学习对象的顺序，帮助学生高效、高质地完成学习目标。

2.4 考核及学习效果预测

学习效果的评价通常与考核测评以及就业息息相关。通过对学习习惯、知识的检索和获取习惯、学生成绩及学习效果的抽取和清洗,可找到成绩数据中潜在的知识和信息,对提高个性化教学质量甚至对学生就业有着积极的指导意义。Lu [15]提出一种学习框架，该框架基于人工智能的自适应辅导系统，以解决自适应辅导系统的问题；这种知识图谱基于深度神经网络[16]应用于学习资源建议知识模块，专为个性化设计的高效算法建议和学习路径，分析学生弱点、掌握状况、学习习惯。行为检测模块[17]可以有效分析学生的状态并提出注意事项，在未来可以针对学生薄弱点进行进一步训练。FETTACH [18]提出了通过使用知识图谱对招聘广告以及学生能力进行关系提取，通过对广告等招聘信息的数据帮助学生确定就业市场需求，企业通过对学生能力数据的提取，建立更有效的人才评估方法，以此打破教育和就业市场之间的技能不匹配，有助于缓解社会的就业压力。

2.5 课程设计以及教案评估

随着互联网课程的普及，生成出海量学生的学习习惯数据，利用这些真实世界的数据支持教师课程、学位规划、大学战略层面的行政管理。管理海量数据，通常来自多个异构来源，其属性有时反映语义不一致，造成运算复杂和处理数据的低效。Manuel [19]提出的 e-LION（电子学习集成本体）语义模型，通过对数据进行高级语义查询，依据学生的最终成绩和互动时产生的数据进行预测建模和预测时间序列，生成学生行为分类的 SWRL 推理规则，以此设计出更有针对性的课程设计。胡辉[20]在个性化习题推荐研究中引入了此类知识图谱的分析方法，根据图网络特性简化分析过程，提升分析效率。该研究通过使用教案映射知识图谱的分布规律,引入路径矩阵和中心性算法,提出基于路径矩阵和知识点序列的教案特征类型的判定方法。

四、知识图谱在医学教育中前景及展望

“互联网+教育”正在触发教育教学模式变革，医学人才培养目标的转变亟需课程教学内容体系重构、人才培养培养模式的转变、教育服务及管理方式的创新。“互联网+教育”视域下的知识图谱应用于医学教育可以从三方面进行深化改革创新：一、构建“问题为导向”的知识图谱医学人才培养新模式。探索基于知识图谱的新型教学模式，实施因材施教、个性化学习的新型教学组织方式。重塑教学评价和教学管理方式，变结果导向的单一评价为综合性、过程性的多维度评价，由仅注重知识传授向更加注重能力素质培养转变。二、发展知识图谱为蓝图的医学教育服务供给新模式。通过知识树或思维导图的形式构建数字医学教育资源体系，缩小医学教育数字鸿沟，为打造医学教育服务供给模式提供政策的扶持和丰富的教育资源“沃土”。三、探索知识图谱资源下的医学教育管理新模式。利用知识图谱采集到的学生学习行为；缜密布局构建医学临床教育教学为蓝本的教育管理信息系统，教与学全周期内参与者进行纳入的教育管理与监测体系。通过对各层级各个学科的合纵连横设置虚拟教研室，以大数据为抓手对教育管理方式进行深化变革，最终形成规范统一、互联互通、安全可控的医学教育数据开放体系。

“互联网+教育”视域下的知识图谱在医学教育中的研究仍处在建设和应用阶段，存在应用不全面欠深入、医学课程实操创新性不足等问题。比如：对基于知识图谱教育教学新模式的探索，知识图谱学习资源的开发、课程体系建设和专业的交叉知识融合。相信随着人工智能机器学习技术日新月异，虚拟现实技术必然与大数据、人工智能等现代信息技术融合，构成“智慧”医学教育的重要技术支撑。

PU-091

基于生物信息学筛选 IgA 肾病 Hub 基因

陈凯

开封市人民医院 科研教学部

IgA 肾病（IgA nephropathy，IgAN）是我国最常见的肾小球疾病，也是我国终末期肾病的重要病因[1]。IgAN 最常见的临床表现为无症状性血尿，部分患者可出现蛋白尿、高血压或表现为肾病综合征，病理上以 IgA 型免疫复合物在肾小球系膜区过多沉积为主要特征，同时可伴有不同程度的细胞外基质增生、系膜细胞过度增殖等组织学改变。IgAN 的具体发病机制尚不明确，目前认为半乳糖缺陷的免疫球蛋白（immunoglobulin，Ig）A 可能在疾病的发生中起重要作用[2]，研究表明在大多数 IgAN 患者及其直系亲属体内铰链区半乳糖缺陷的 IgA 分子水平升高，针对这种半乳糖缺陷的 IgA 分子机体可产生自身抗体 IgA、IgG。此外，感染、氧化应激所造成的“二次打击”也可促进自身抗体的生成，由此形成的抗原抗体免疫复合物沉积在肾小球，通过激活补体系统介导炎症反应、肾小球硬化、间质纤维化[3-4-5]，导致疾病的发生发展。因此深入了解 IgAN 的发病机制，对于疾病诊断、治疗、预后判断具有重要意义。

近年来，随着测序技术的发展，应用基因表达谱技术可筛选出与疾病发病机制相关的差异表达基因，并从这些差异表达基因中分析鉴定不同信号通路、生物学过程及分子机制的各种相关基因。研究表明，IgAN 的发生发展受多基因共同调控，因此仅研究单个基因具有明显的局限性。通过加权基因共表达网络分析（Weighted correlation network analysis，WGCNA）可以鉴别基因表达与临床表型间复杂的相关性，根据样本基因表达谱对相关基因进行聚类，构建共表达模块，将模块与临床表型进行关联，从而筛选出与临床表型相关性较高的模块，并从这些模块中筛选出在疾病发展中起重要作用的关键基因[6-7-8]。

本研究通过 WGCNA 方法鉴定并识别与 IgAN 患者发病相关的共表达模块及中枢（Hub）基因，通过对筛选出的 Hub 基因进行 GO 功能富集以判断这些基因在 IgAN 发病中的作用，以此为深入了解 IgAN 的发病机理、疾病诊断、治疗以及预后判断提供理论依据。

1 材料和方法

1.1 数据来源及处理

从 GEO 数据库（https://www.ncbi.nlm.nih.gov/geo/，于 2023 年 8 月 22 日访问）中获取 IgAN 患者及正常对照人群表达谱及临床信息。研究中包含的 GEO 数据集是基于 GPL14663[HGU133A_CDF_ENTREZG_10]阵列的微阵列，数据集 GSE35488 中的样本来源于 25 名 IgA 肾病患者及 6 名正常对照人群的肾组织活检。

1.2 差异表达基因

使用 R 软件包 limma(version 3.40.6)进行差异分析，以获得不同比较组与对照组间的差异基因。对表达谱数据集 GSE35488 利用 lmFit 函数进行多重线性回归，进一步使用 eBays 函数，对差异表达进行经验贝叶斯调制，计算调制后的 t 统计量、F 统计量和对数几率，最终获得每个基因

的差异显著性。使用 *P*<0.05 和 logFC>0.1 作为 DEG 的筛选标准。logFC＜0 的 DEG 被认为是下调的基因，而 logFC>0 的 DEG 被认为是上调的基因。使用火山图绘制工具绘制相关图片。

1.3 WGCNA 鉴定基因集与表型之间相关性及 Hub 基因筛选

利用差异基因表达谱，分别计算了每个基因的 MAD，剔除了 MAD 最小的前 50%的基因,利用 R 软件包 WGCNA 的 goodSamplesGenes 方法去除了离群的基因和样本，进一步的使用 WGCNA 构建共表达网络。先对所有成对基因进行 Pearson 相关矩阵和平均连锁法，然后用幂函数 a_mn=|C_mn|^β 构造加权邻接矩阵。β 是一个软阈值参数，可以强调基因之间的强相关性。在选择 22 的幂后，将邻接性转换为拓扑重叠矩阵（TOM），该矩阵可以测量一个基因的网络连通性，该基因定义为其与所有其他基因的邻接性之和，用于网络基因比率，并计算相应的相异性（1-TOM）。为了将具有相似表达谱的基因分类为基因模块，根据基于 TOM 的相异性测量进行平均连锁层次聚类，基因树状图的最小值为 30。设置敏感度为 3，为了进一步分析模块，计算了模块特征基因的相异性，选择了模块树状图的切割线，并合并了一些模块。通过合并距离小于 0.25 的模块,最终获得了 2 个共表达模块。计算了模块特征向量与基因的表达相关性以获得 MM 阈值，设置 MM 阈值为 0.8，GS 阈值为 0.6，权重阈值为 0.1，筛选出 15 个 Hub 基因。

1.4 Hub 基因的 GO 功能富集分析

对筛选出的 Hub 基因使用 R 软件包 org.Hs.eg.db(version 3.1.0)中的基因的 GO 注释，以此作为背景,将基因映射到背景集合中，使用 R 软件包 clusterProfiler(version 3.14.3)进行富集分析，以获得基因集富集的结果。设定最小基因集为 5，最大基因集为 5000，*P*< 0.05(根据需要自己修改) and FDR<0.1 认为差异具有统计学意义。

2 结果

2.1 IgAN 组与 Normal 组差异基因筛选

通过对数据集 GSE35488 进行 Limma 差异分析，按照差异倍数 1.5、显著阈值 0.05，结果显示 IgAN 组与 Normal 组间共存在 188 个差异基因，其中较 Normal 组表达上调的基因 58 个，表达下调的基因 130 个。（图 1）

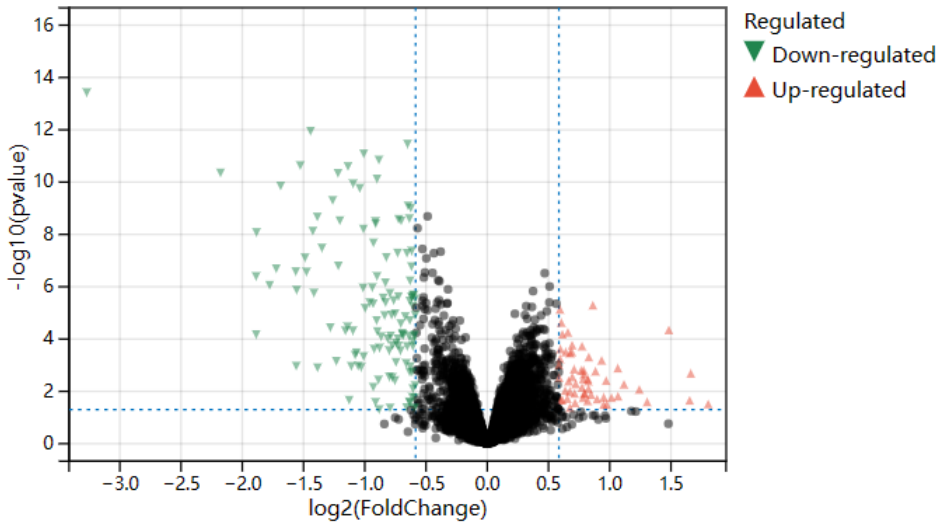


图 1. IgAN 组与 Normal 组间差异基因

2.2 Hub 基因筛选

通过 WGCNA 进行基因和表型间关系分析，绘制标度独立性，软阈值为 0.86（图 2a），绘制平均连通性，阈值为 0.32（图 2b）。对数据进行模块特征向量聚类，得到 turquoise 模块与 grey 模块（图 2c），通过绘制模块表型相关性热图，结果显示 grey 模块与 IgAN 组呈负相关（r=-

0.89，*P*＜0.05），与 Normal 组呈正相关（r=0.89，*P*＜0.05）（图 2d），选择 MM 阈值为 0.8，GS 阈值为 0.6，权重阈值为 0.1，筛选出 Hub 基因 15 个，基因分别为 CEBPD、TRIB1、GEM、CYR61、RGS2、DUSP6、FILIP1L、CHSY1、FAM134B、CTGF、CCL2、IER3、KLF6、GADD45B、GDF15。对上述 15 个 Hub 基因与上调差异基因及下调差异基因取交集，绘制韦恩图，结果显示 15 个 Hub 基因均为下调的差异基因（图 2e）。

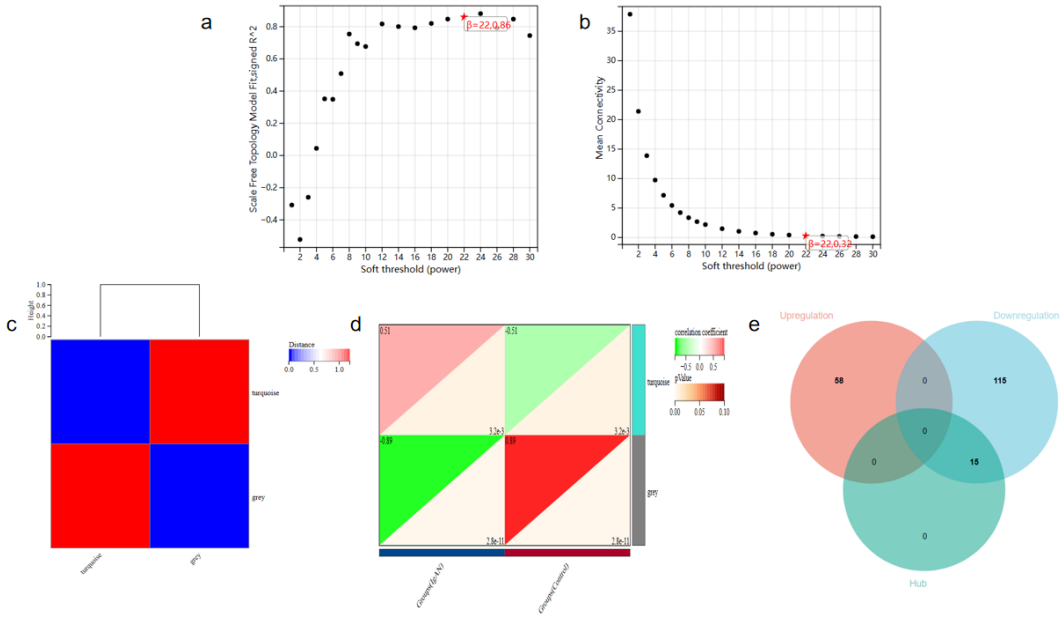


图 2a. 标度独立性；2b. 平均连通性；2c. 模块特征向量聚类图；2d. 模块表型相关性热图；2e. 差异表达基因与 Hub 基因韦恩图

2.3 GO 富集分析

对 Hub 基因进行 GO 富集分析，选取 P＜0.05、最小基因集 5 作为检验标准，选择分子功能、细胞成分、生物学过程三种方面功能富集。在生物学过程方面，基因被富集在蛋白质磷酸化信号转导、MAPK 级联反应、激酶活性调节、蛋白激酶活性调节、MAPK 级联反应调节、丝苏氨酸激酶活性调节、MAPK 的调节、激酶活性负调节、蛋白激酶活性的负调节（图 3a）；在细胞成分富集方面，基因被富集在高尔基体、有丝分裂纺锤体、核仁、质膜胞质侧、细胞膜胞质侧（图 3b）；在分子功能方面，基因被富集在 MAP 激酶酪氨酸/丝氨酸/苏氨酸磷酸酶活性、MAP 激酶磷酸酶活性、丝裂原活化蛋白激酶结合、泛素蛋白转移酶调节活性、G 蛋白 α 亚基结合、钙调素结合、GTP 酶活性（图 3c）。

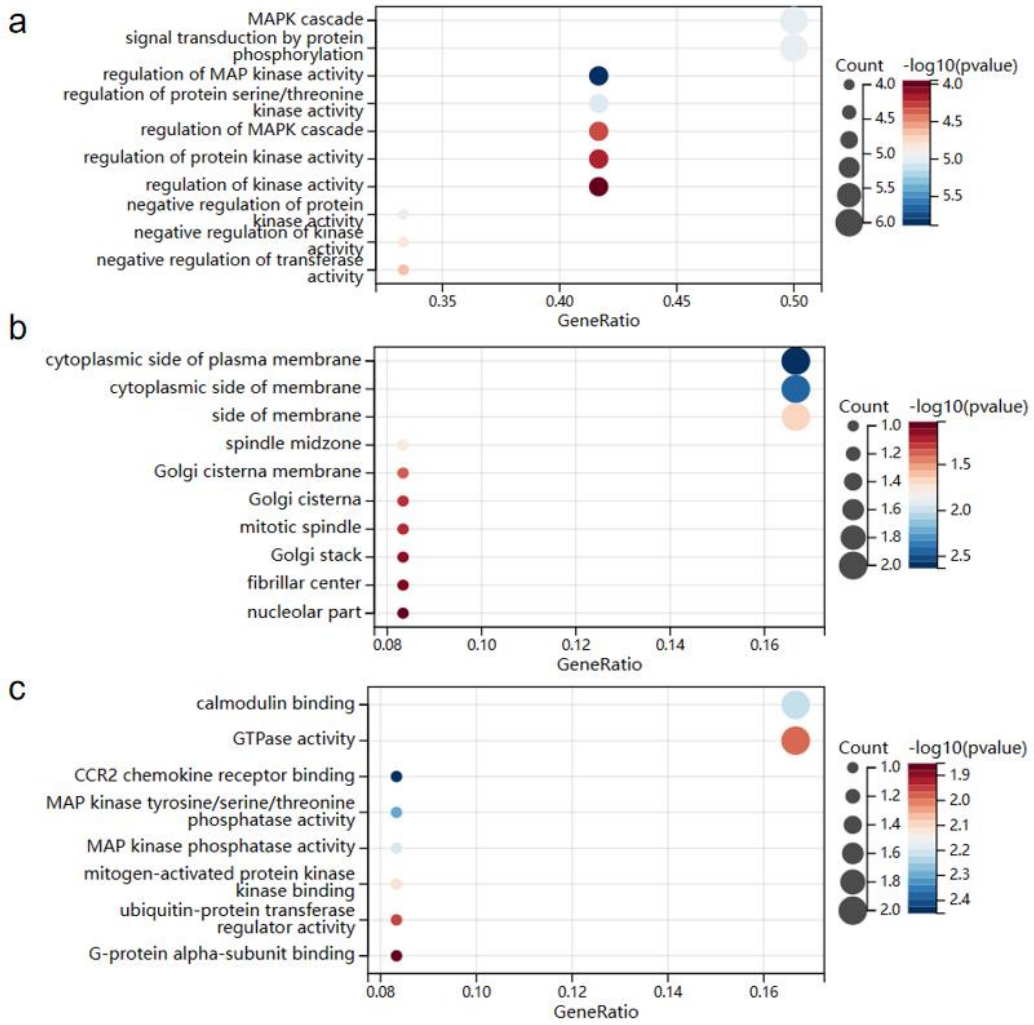


图 3a. 生物学过程 GO 功能富集；3b. 细胞成分 GO 功能富集；3c. 分子功能 GO 功能富集

3 讨论

IgAN 约占我国原发性肾小球疾病的 30%至 40%[9],是我国导致终末期肾病的重要病因之一。研究表明，IgAN 的发生可能与遗传倾向、免疫调节异常有关[10-11]，但由于其具体发病机制尚不明晰，关于 IgAN 的最佳治疗方案也尚不明确，当前仍缺乏特异性治疗手段。目前 IgAN 的治疗目标仍是控制血尿、控制血压、减少蛋白尿、延缓肾脏向终末期肾病进展[12-13]。相较于治疗，在肾脏发生不可逆损伤前早诊断、早治疗仍是最有效的干预手段。因此，深入了解 IgAN 的发病机制，才能更好地为诊断、治疗提供有利的理论支撑。

本研究通过 Limma 差异分析、WGCNA 等生物信息学方法筛选出 CEBPD、TRIB1、GEM、CYR61、RGS2、DUSP6、FILIP1L、CHSY1、FAM134B、CTGF、CCL2、IER3、KLF6、GADD45B、GDF151 共 15 个 Hub 基因，通过 GO 功能富集分析显示，在细胞成分富集方面，基因被富集在高尔基体、有丝分裂纺锤体、核仁、质膜胞质侧、细胞膜胞质侧；在分子功能方面，基因被富集在有丝分裂原活性蛋白激酶（mitogen-activated protein kinase，MAPK）酪氨酸/丝氨酸/苏氨酸磷酸酶活性、MAP 激酶磷酸酶活性、丝裂原活化蛋白激酶结合、泛素蛋白转移酶调节活性；在生物学过程方面，基因被富集在蛋白质磷酸化信号转导、MAPK 级联反应、激酶活性调节、蛋白激酶活性调节、MAPK 级联反应调节、丝苏氨酸激酶活性调节、MAPK 的调节、激酶活性负调节、蛋白激酶活性的负调节。上述结果提示 15 个 Hub 基因涉及蛋白质合成过程，并且主要参与激酶调节尤其是 MAPK 信号通路的调节过程。

激酶是一类能够催化生物学过程中磷酸化的酶，通过将磷酸基团添加到目标蛋白的特定位点上发挥作用。这些磷酸化事件对细胞信号传导、细胞增值、细胞凋亡、细胞分化等多种生物学途径至关重要[14-15]，且激酶功能异常可参与多种肾脏疾病的发生发展[16]。AMP 活化蛋白激酶（amp-activated protein kinase，AMPK）与人糖皮质激素调节激酶(serum and glucocorticoid inducible kinase，SGK)是一类在细胞能量代谢和离子平衡调节中发挥重要作用的激酶[17-18]，当二者功能出现失调，可影响肾小管细胞的离子通道、药物转运和水平衡，从而影响尿液的生成和组成，引发肾小管疾病；MAPK、细胞因子受体相关激酶-信号转导和转录激活因子（Janus kinase/signal transducer and activator of transcription，JAK-STAT）、磷脂酰肌醇激酶-Akt-靶向雷帕霉素（Phosphatidylinositol 3-kinase/Akt/mechanistic target of rapamycin，PI3K-Akt-mTOR）等激酶通路的异常活性可导致炎症反应、细胞凋亡、肾小球损伤，从而介导肾小球肾炎的发生[19-20-21]。已有研究表明 MAPK 信号通路在 IgAN 中发挥重要作用[22-23]，MAPK 包括胞外信号调节激酶（Extracellular signal-regulated kinase，ERK）、c-Jun N 末端激酶（c-Jun N-terminal kinase，JNK）、p38 有丝分裂原活化蛋白激酶（p38 mitogen-activated protein kinase，p38）。在 IgAN 患者中 MAPK 信号通路尤其是 JNK 和 p38 的活性增加，从而导致肿瘤坏死因子-α（Tumor necrosis factor-α，TNF-α）、白细胞介素（Interleukin，IL）-1β、IL-6 等促炎细胞因子生成增多、炎性损伤增强。此外，JNK、p38 通路还可以介导细胞凋亡信号的传递而促进肾小球和肾小管细胞的凋亡[24-25]。也有报道显示，活化的 MAPK 信号通路可调控胶原合成、基质金属蛋白酶活性以及细胞外基质的重塑，导致肾小球硬化和间质纤维化等病理改变。

本研究中，CYR61、RGS2、DUSP6、FILIP1L 等 15 个 Hub 基因均与 IgAN 组呈负相关，与 Normal 组呈正相关，并且 GO 功能富集分析显示这些基因与激酶、尤其是 MAPK 的调节有关，提示这些 Hub 在 IgAN 患者中表达降低可能解除其对 MAPK 信号通路的抑制作用，而促进 IgAN 的发生。已有研究显示，CYR61、RGS2、DUSP6、FILIP1L 等 Hub 可通过作用于 MAPK 的不同环节而抑制 MAPK 信号通路的活性。例如，TRIB1 可通过抑制 ERK 和 JNK 的活性来调控 MAPK 通路活性[26-27]；RGS2 则可以通过抑制 p38 和 ERK 以抑制 MAPK 信号通路的转导[28]；DUSP6 的编码产物 MKP-3 可通过去磷酸化效应抑制 ERK 的活性[29]；GDF15 是一种生长分化因子，属于 BMP 家族成员，其通过与 BMP 受体结合出发 Smad1/5/8 的磷酸化。进而形成 Smad1/5/8 与 Smad4 的复合物，此复合物可进入细胞核调节靶基因的表达而抑制 MAPK 信号通路活性[30]。

综上所述，IgAN 的发生是一个多基因、多途径的过程。激酶调节基因，尤其是 MAPK 调节基因在 IgAN 的发生发展中可能起到重要作用。通过本研究可以为日后续深层次揭示 IgAN 的发病机制，更好地诊断、治疗、判断患者预后提供理论支撑。

PU-092

基于定量转录组与免疫浸润分析探究线虫感染对衰老的影响机制

张力中[1]，张钧栋[2]，陈浩然[3]，卢学春[2,4*]，于琦[3*]

[1]山西医科大学基础医学院

[2]解放军总医院第二医学中心血液病科

[3]山西医科大学管理学院

[4]国家老年疾病临床医学研究中心

目的 通过联合线虫（Nematoda）感染人单个核细胞样本的定量转录组分析与对线虫感染小鼠肠上皮细胞的免疫浸润分析，明晰线虫感染对衰老的影响机制与免疫浸润特征。

材料与方法 在 GEO（Gene Expression Omnibus）数据库中以“Whipworm OR Hookworm OR Filaria/Trichinella spiralis”为检索式，以样本量大于等于 3，具有完整注释文件为标准，筛选线虫纲动物感染人或小鼠的基因表达谱数据，结合 HAGR 数据库中人衰老相关基因，提取丝虫感染人后衰老相关基因表达矩阵。采用 Limma 软件包对衰老相关基因表达谱进行差异表达分析获得差异基因（DEGs），应用 clusterProfile 软件包对 DEGs 进行 KEGG 功能富集分析；利用 CIBERSORT 软件包以小鼠免疫细胞特征矩阵为参考对旋毛虫感染小鼠的数据集进行免疫浸润分析。

结果 最终选取编号为 GSE2135 的丝虫感染人数据集与编号为 GSE67365 的旋毛虫感染小鼠数据集。联合 HAGR 数据库分析得到丝虫感染人后衰老相关差异基因 31 个。KEGG 富集分析显示上述衰老相关差异基因主要与细胞衰老 、FoxO 信号通路、HIF-1 信号通路等相关，GO 富集主要与衰老过程、受损 DNA 结合功能相关。将上述衰老相关差异基因导入 String 数据库绘制 PPI 网络，并使用 cytoHubba 插件中 MCC 算法得出 10 个 Hub gene，分别为 VEGFA、MYC、PIK3R1、ATM、CAT、SERPINE1、PIK3CB、NFE2L2、PTGS2、APOE。对旋毛虫感染小鼠的免疫浸润分析发现，正常组与感染组的小肠上皮免疫浸润的情况无显著性差异。

结论 本研究分析了线虫感染对生物体衰老的影响，得到了线虫感染对人衰老相关基因的调控情况，其中线虫对 FoxO 信号通路的激活或对衰老产生一定影响。而对小鼠免疫浸润的结果说明线虫感染后肠上皮组织并未有免疫细胞募集，表明线虫感染对免疫具有一定抑制作用，证明寄生虫可能对衰老相关的慢性炎症也有一定的预防作用。

PU-093

国内外知情同意能力研究热点的共词聚类分析

朱彦霞[1]，吴玉彬*[1]，华南[2,3]

1.河南省职工医院 科研教学科

2.河南中广智媒科技有限公司 产品创新支撑部

3.中国广电河南网络有限公司 未来电视业务部

0 前言

随着现代医学技术的不断进步和人们对医疗权益的日益关注，知情同意能力越来越受到人们的关注和重视。评估知情同意能力和保护患者知情权益是医学研究和临床实践中不可或缺的组成部分，同时也是医学伦理学研究和实践的重要内容之一。美国最早将知情同意权立法，是第一个通过判例形式确定知情同意权的国家[1]，其知情能力评估也是邀请会诊的重要原因之一[2]；英国已为患者的知情同意能力立法[3]。我国政府及相关领域的医学、法学、伦理学专家也在知情同意及能力评估方面给与了高度关注。例如，2016 年 10 月，国家卫健委颁发第 11 号令《涉及人的生物医学研究伦理审查颁发》（2016 版本），明确了在开展医学研究的伦理审查过程中关于知情同意的审查原则；杨民来[4]、王欢[5]则分别探讨了医疗大数据过程中知情同意原则并提出了保障医疗数据利用的知情同意的实现路径；庞敏[6]研究了我国无民事或限制民事行为能力人、特殊情况下由代理人行使知情同意的实用范围、实用条件、行使主题及行使顺序等。然而，中外文献在研究范围、重点与方法上的差异尚待弥合；知情同意评估标准的个性化程度和评估工具的科学性待优化；特殊人群下的知情同伦理问题还需深入研讨。

因此，本文旨在通过 Bicomb、CiteSpace、VOSviewer 等软件对国内外近 10 年的知情同意能力相关研究进行共词聚类分析，及时了解国内外相关研究热点，为我国学者在该领域的研究提供参

考与借鉴。通过对相关文献进行数据挖掘和分析，可以深入了解知情同意能力的研究热点、研究趋势等方面的信息，推动该领域的发展。

1 资料和方法

1.1 文献检索策略及数据筛选

以“主题=知情同意 AND 主题=能力 OR 主题=知情同意能力”为检索条件，在 2013 年 1 月 1 日至 2022 年 12 月 31 日的时间范围内，对中国知网数据库（CNKI）进行中文文献检索。同时，以“Informed consent capacity”为主题词，在 PubMed 中进行检索，也在相同的时间范围内检索到 3031 篇文献（其中：中文文献 706 篇，英文文献 2325 篇）。通过排除无关研究、无关键词、重复文献等筛选过程，最终得到了 2965 篇有效期刊文献（其中：中文文献 706 篇，英文文献 2259 篇）。

1.2 研究方法

采用文献计量学的方法,对筛选出的文献进行文献发文量统计、关键词、共词聚类分析及共词网络、密度图等可视化展示。具体步骤如下:

（1）使用中文共现分析系统 Bicomb2.0 对文献进行自动分词,提取关键词,并统计关键词频次,获得高频关键词, 从而确定研究热点和关注重点。

（2）使用 CiteSpace 软件构建关键词的共现网络图,可呈现关键词间的相关强弱和研究领域的发展脉络。此外, 采用 CiteSpace 软件进行关键词时区图和突变图的绘制：关键词时区图可以展示关键词在不同时间段的变化趋势，以观察知情同意能力研究的演化进程和发展动态；突变图则可以揭示某一关键词在某个时间突然增加或减少的情况，从而发现可能的研究突破点和转折点。

（3）使用 VOSviewer 软件分析各个研究领域的知识结构,构建密度图：密度图可以通过节点的大小和颜色来表示学科的重要性和关联程度，从而帮助我们发现知情同意能力研究中的学科交叉和合作领域。

（4）综合运用多种文献计量分析方法,系统梳理不同时期“知情同意能力”研究的主流方向和热点焦点,为后续研究提供参考。

2 结果

2.1 各年份文献发文量对比分析

通过对筛选出的文献进行发文量统计分析，发现国外对知情同意能力的研究发文量相对较高，在 2016-2018 年期间发文呈现迅速递增趋势，在 2018 年发文量 600 余篇；国内对知情同意能力的研究基本稳定在每年 100 篇左右，2020 年发文量较高（达到 100 篇）。见图 1。

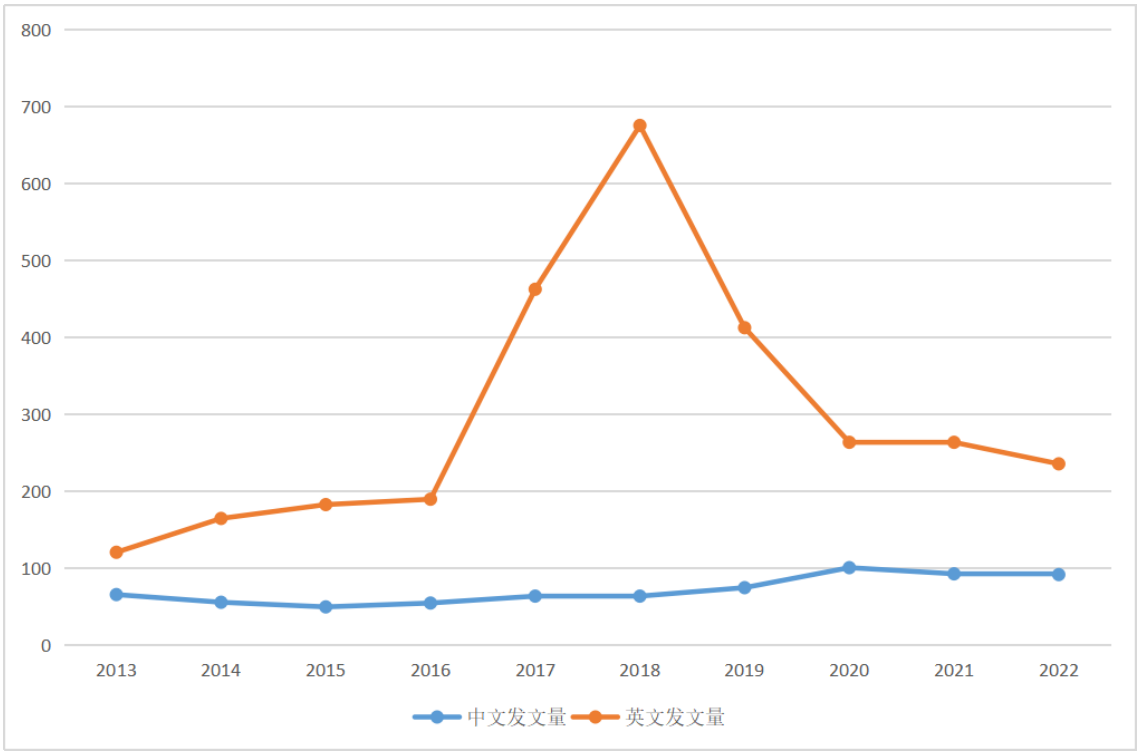


图 1“知情同意能力”中英文文献发表年代趋势图

2.2 高频关键词分析

对某一时期文献中出现频率较高的关键词进行词频分析，可发现该领域在该时期的研究热点[7]. 本研究基于 Bicomb2.0 分别对“知情同意能力”相关中英文文献进行关键词统计分析，考虑统计结果的代表性，分别截取国内外研究中频次>3 的前 32 个中英文关键词，具体见表 1。

表 1 基于 Bicomb2.0 的国内外知情同意能力研究关键词分析

名	中文关键词	次	百分比 1(%)	累计百分比 1（%）	英文关键词	次	百分比 2（%）	累计百分比 2(%）
	知情同意	10	3.5324	3.5324	Informed Consent（知情同意）	88	4.8253	4.8253
	知情同意权	1	2.28	5.8125	Mental Competency（心智能力）	14	2.6614	7.486
	医患关系	9	0.9313	6.7437	Decision Making（决策制定）	98	2.4624	9.949
	告知义务	2	0.7065	7.4502	Biomedical Research（生物医学研究）	8	0.97	10.919
	伦理审查	1	0.6744	8.1246	Mental Disorders（精神障碍）	9	0.7337	11.6528
	患者知情同意权	8	0.578	8.7026	Pulmonary Disease，Chronic Obstructive（肺疾病，慢性阻塞性）	8	0.7213	12.3741
	脑卒中	7	0.5459	9.2486	Lung（肺）	7	0.7089	13.0829
	影响因素	6	0.5138	9.7624	Dementia（痴呆）	2	0.6467	13.7296

9	临床试验	5	0.4817	10.2441	Exercise（运动）	5	0.5596	93	14.28
10	生前预嘱	5	0.4817	10.7258	Research Subjects（研究课题）	2	0.5223	16	14.81
11	伦理问题	4	0.4496	11.1753	Personal Autonomy（个人自主）	2	0.5223	39	15.33
12	伦理	4	0.4496	11.6249	Quality of Life（生活质量）	0	0.4975	14	15.83
13	个人信息	3	0.4175	12.0424	Patient Participation（患者参与）	0	0.4975	88	16.32
14	受试者	2	0.3854	12.4277	Health Policy（健康卫生政策）	6	0.4477	65	16.77
15	患者	2	0.3854	12.8131	Cognitive Dysfunction（认知功能障碍）	5	0.4353	18	17.21
16	临床研究	1	0.3532	13.1663	Schizophrenia（精神分裂症）	3	0.4104	22	17.62
17	生活质量	1	0.3532	13.5196	HIV Infections（艾滋病毒感染）	2	0.398	01	18.02
18	未成年人	1	0.3532	13.8728	Primary Health Care（初级卫生保健）	1	0.3855	57	18.40
19	医患沟通	1	0.3532	14.2261	Neoplasms（肿瘤）	1	0.3855	12	18.79
20	伦理委员会	0	0.3211	14.5472	Randomized Controlled Trials as Topic（随机对照实验）	0	0.3731	43	19.16
21	医疗行为	0	0.3211	14.8683	Exercise Therapy（运动疗法）	0	0.3731	74	19.53
22	侵权责任	0	0.3211	15.1895	Clinical Trils as Topic（临床实验）	9	0.3607	8	19.89
23	最佳利益	0	0.3211	15.5106	Treatment Refusal（拒绝治疗）	7	0.3358	38	20.23
24	精神障碍	9	0.289	15.7996	Health Personnel（个人保健员）	7	0.3358	96	20.56
25	质性研究	9	0.289	16.0886	Mesenchymal Stem Cells（间充质干细胞）	7	0.3358	54	20.90
26	老年人	9	0.289	16.3776	Stroke（中风）	7	0.3358	11	21.24
27	紧急救治	8	0.2569	16.6346	Physicians（医师）	6	0.3233	45	21.56
28	知情同意能力	8	0.2569	16.8915	Alzheimer Disease（阿尔茨海默病）	6	0.3233	78	21.88
29	药物临床试验	8	0.2569	17.1484	Delivery of Health Care（照料服务提供）	6	0.3233	12	22.21
30	同意能力	8	0.2569	17.4053	Aging（老龄化）	5	0.3109	21	22.52
31	自主权	8	0.2569	17.6622	Patient Selection（患者选择）	5	0.3109	3	22.83
32	医疗自主权	8	0.2569	17.9191	Third-Party Consent（第三方同意）	4	0.2985	15	23.13

从表 1 可以看出，中文文献中，知情同意相关的研究主要涵盖了医患关系、告知义务、伦理审查、脑卒中、临床试验等方面。研究重点主要关注患者知情同意权的保障和限制，以及伦理问题的探讨和解决。此外，也探讨了个人信息保护、未成年人知情同意、医疗行为侵权责任等问题;而英文文献中，知情同意所涉及的领域更加广泛，包括生物医学研究、精神障碍、肺疾病、痴呆、运动、肿瘤、中风、阿尔茨海默病、老龄化等。研究热点主要涵盖了个人自主权、患者参与、生活质量、健康卫生政策等方面。同时，也探讨了临床实验、随机对照实验、拒绝治疗、第三方同意等问题。

总体来说，中外研究都强调了知情同意在医疗行为中的重要性和保障作用，但具体研究重点和方向各有不同。在未来的研究中，应加强对知情同意权的保障和限制、伦理问题的探讨和解决、个人信息保护等方面的研究。

2.3 共词网络图

运用 CiteSpace 对国内外文献分别进行关键词共现分析，如图 2 所示。在中文共现图谱中，关键词“知情同意”凸显，其他关键词如：知情同意医患关系、同意能力、告知义务、伦理审查、患者知情同意权、脑卒中、影响因素、临床试验、生前预嘱、伦理问题等为题围绕其展开。这些关键词反映了知情同意能力研究的多个方面，包括知情同意能力、伦理审查、患者知情同意权利、临床试验过程中涉及的相关研究等，这些研究对于促进知情同意实践、保障患者权利以及加强临床试验的伦理管理具有重要意义。

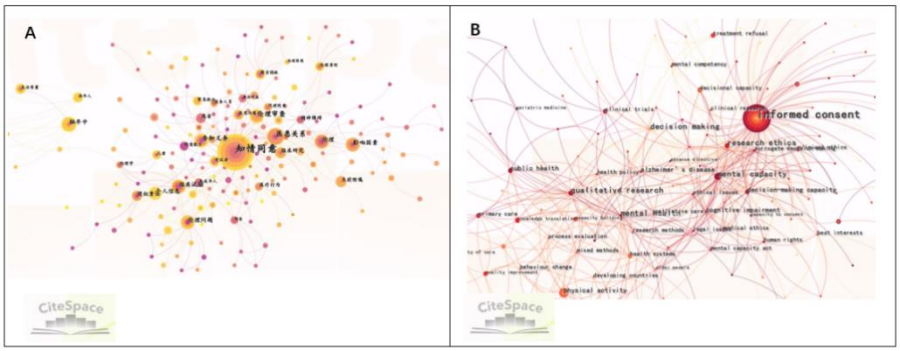


图 2 基于 CiteSpace 的国内外知情同意能力研究关键词共现图谱

在英文共现图中，“Informed Consent（知情同意）”关键词凸显，Mental Competency（心智能力）、Decision Making（决策制定）、Biomedical Research（生物医学研究）、Mental Disorders（精神障碍）、research ethics（伦理研究）等围绕其展开。这些关键词反映了英文文献中对知情同意能力相关研究的关注点和研究方向。其中，心理和认知因素、决策制定的影响因素、生物医学研究中的知情同意问题、与精神障碍相关的知情同意问题，以及伦理研究的角度分析知情同意问题等，都是研究中的重要方向。

与关键词词频分析结果基本保持一致，关键词的共现图谱进一步揭示了这些研究方向之间的联系和互动。这些研究方向都是在为了更好地理解和保障患者的知情同意权，探索如何在不同情境下评估和提高患者的知情能力，以及如何在实践中应用知情同意原则。在未来的研究中，可以进一步深入探讨这些问题，以推动知情同意相关研究的进一步发展和应用。

2.4 共词聚类分析

运用 Cite Space 进行共现词聚类分析（如图 3 所示），将相关关键词分为不同的聚类，如“知情同意”、“同意能力”、“伦理研究”、等，加深了该研究主题的进一步理解，同时发现国内外在“知情同意能力”研究中存在一些差异。

对国内文献关键词聚类分析，结果聚集在：知情同意、医患关系、伦理审查、临床试验、伦理原则、影响因素等群组。这些群组反映了国内研究者对知情同意相关问题的关注点和研究方向，包括知情同意的实践、伦理审查、医患关系、临床试验等；同时，影响因素的聚类也表明了国内研究者对于影响知情同意实践的因素进行了探讨，这对于制定相关政策和规范实践具有重要意义。

对国外文献关键词聚类分析，结果聚集在：informed consent(知情同意)、mental capacity(心智能力)、decision making(决策制定)、qualitative research(质性研究)、quality of life 等群组。这些群组反映了国外研究者对知情同意相关问题的关注点和研究方向，包括知情同意的心理和认知因素、决策制定的影响因素、质性研究、生命质量等。这些研究对于提高知情同意的质量、促进实践的发展以及保障患者权益具有重要意义。

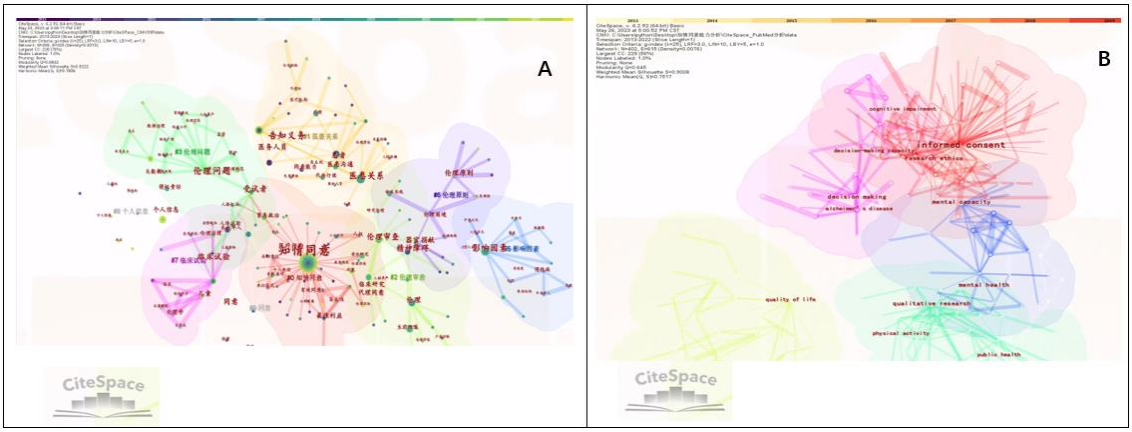


图 3 基于 CiteSpace 的国内外知情同意能力关键词聚类分析图谱

2.5 关键词突变热点分析

基于 CiteSpace 的“知情同意能力研究”关键词演化分析如图 4 所示，图中红色线段表示关键词突现的时间。在国内文献中，“同意能力”、“知情同意能力”、“同意”、“医疗自主权”等关键词的突现时间较长，自 2017 年以来，“代理同意”、“医学伦理学”、“医疗自主权”、“生前预嘱”等关键词逐渐成为国内研究的热点；而在国外文献中，“clinical research（临床研究）”、“decisional capacity（决策能力）”、“decision making（决策制定）”、“mental competency（精神能力）”等关键词的突现时间较早，自 2017 年以来，“primary health care（初级保健）”、“quality improvement（质量改进）”、“capacity building（能力增强）”、“knowledge translation（知识转化）”等关键词逐渐成为国外研究的热点。

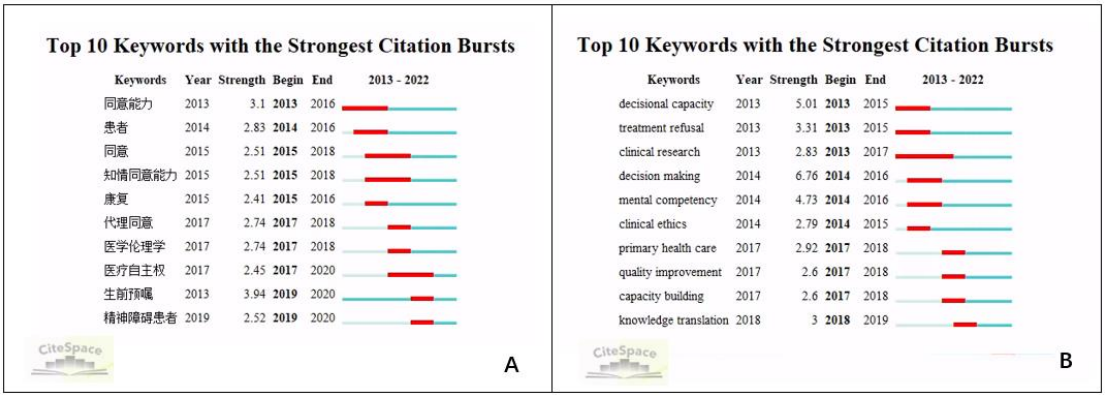


图 4 基于 CiteSpace 的知情同意能力研究关键词突变分析图谱

2.6 关键词密度分析

共现密度图是一种用于反映关键词在文献中出现频率、权重和相关性的可视化工具，以为研究人员提供更全面、更准确的研究视角和研究方向，在图中，与黄色接近程度越高，则表示该关键词在文献中的权重越大，热度越高。基于 VOSviewer 的“知情同意能力研究”关键词密度可视化图如图 5 所示，该图是对“知情同意能力”研究领域的可视化呈现。

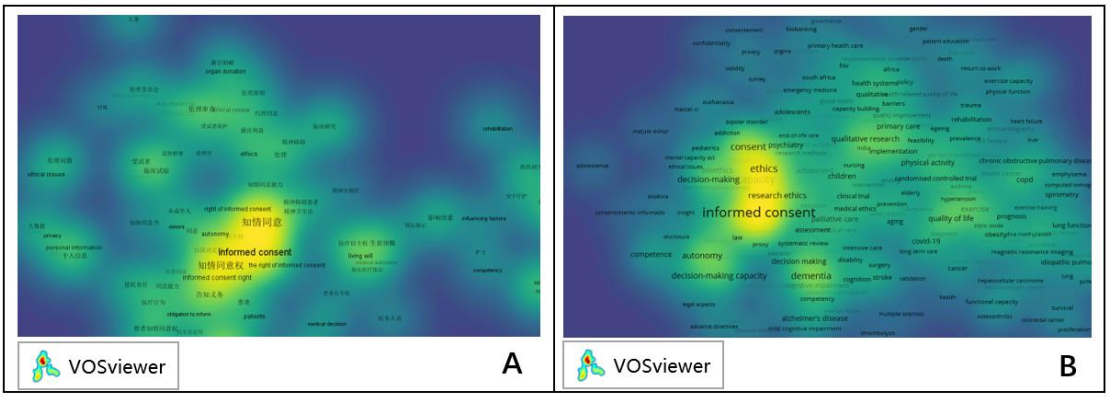


图 5 基于 VOSviewer 的知情同意能力研究关键词密度图

从图 5 中可以看出，国内和国外针对“知情同意能力”的研究热点存在一定的差异。在国内文献中，“知情同意”、“知情同意权”、“告知义务”、“医疗自主权”、“伦理审查”、“个人信息”等关键词是研究的热点，这些关键词主要涉及我国医疗伦理法规、患者知情同意、医患关系、个人隐私等方面的研究。而在国外文献中，“informed consent”、“research ethics”、“consent”、“ethics”、“dementia”、“qualitative research”、“quality of life”等关键词则是研究的热点，这些关键词主要涉及国外医疗伦理法规、老年痴呆、病人决策能力、生命质量等方面的研究。同时，图 5 也反映了国内外“知情同意能力”研究热点的主要格局。可以看出，国内和国外的研究热点存在一定的重叠，但也存在明显的差异。这些差异可能与不同国家或地区的医疗伦理法规、研究重点和文化背景有关。

3 讨论

3.1 研究主题和涉及领域的变化趋势

在 2013-2022 年的国内外知情同意能力研究中，研究主题和涉及领域呈现出多样性和复杂性。从研究主题来看，包括医学伦理、临床试验、生物医学研究、精神障碍、老龄化、运动、肺疾病、肿瘤、脑卒中、阿尔茨海默病、痴呆等多个方面。从涉及领域来看，涵盖了医学、生物医学、心理学、社会学、法学、伦理学等多个学科领域。

3.2 关键词共现分析和共词聚类分析的结果

通过对国内外知情同意能力研究相关的高频关键词进行共现分析和共词聚类分析，可以发现以下几个热点和趋势：

（1）伦理审查和知情同意过程的规范化

近年来，在国内外的研究中，规范伦理审查和知情同意过程的意识和需求日益增强，共现词网络图中“伦理审查”、“知情同意”、“决策制定”等词汇共现频率较高。越来越多的学者开始关注知情同意过程中可能存在的问题，如研究对象代表性不足问题、知情同意过程中的评估工具及方法等，并开始探讨如何通过工具和方法来规范伦理审查和知情同意过程。

黄晶晶[8]开展了关于精神障碍患者知情同意能力评估的研究，并构建了适用于临床诊疗过程中的精神障碍患者知情同意能力评定工具，为相关政策制定提供了学术参考；龚一玫[9]以“意定监护人医疗决定权”为主要研究内容，深入探讨了“医疗决定权”与“知情同意权”之间的关系，论述了我国目前相关立法现状并提出完善建议；PeterD.G[10]团队开发了一种知情同意能力评估工具，侧重于 AD 患者的主观信念及对临床试验事实的理解，提供了一种客观的临床知情同意能力衡量的标准参考；

ShepherdV[11]针对知情同意能力评估试验过程中经常存在的“缺乏同意的成年人”被排除在外而导致研究对象代表性不足的问题进行了深入研究并提出应对措施；RadenkovićM[12]肯定了决策能力工具在临床环境中获得知情同意的重要性，并对麦克阿瑟治疗能力评估工具（MacCAT-T），霍普金斯能力评估测试（HCAT）、能力/无能力评估测试的结构化面试等能力评估工具进行测试评估。

（2）医患关系和患者参与度的提高

近年来，随着医疗模式的转变和患者权益的提高，患者参与度得到了进一步提高，共现词网络图中“患者参与度”、“医患沟通”、“医疗自主权”等词汇共现频率较高，表明在国内外研究中，越来越多的学者开始关注患者参与度和医患沟通等问题，并开始探讨如何促进医患沟通、保障患者知情同意权等问题。

庞超[13]站在医患双方角度深入分析后发现：“过度医疗”、“畸形就医期待”、“医患信息不对称”等是造成医患知情同意保障困境的主要成因同时对破除困境提出了“规范医疗行为”、“加强医德医风”、“认可家庭参与知情同意实践模式的合理性”等建议；李爽[14]对医疗行为中的知情同意权开展研究，指出“医务人员告知义务”是患者知情同意的前提同时应进一步明确医务人员告知义务的范围及标准；和钰霖[15]深入研究了知情同意模式，指出知情同意原则对构建和谐医患关系具有积极意义；BeattieE[16]针对在涉及痴呆的研究中，默认选择知情同意代理而忽视痴呆患者自主权与代理权的问题进行研究并发现略超过 1/5 的痴呆症患者认为自己有能力签署知情同意书；MahafzahR[17]调查了 ICU 患者亲属对生物医学研究的态度和偏好，大多数 ICU 患者亲属对于代表患者提供知情同意持积极态度;PothA[18]试图通过开发增强程序（ECPs）改善痴呆症患者知情并自主决策的权力，试验结果虽未发现 ECPs 对同意能力的影像，但提高了痴呆患者对信息的理解。

（3）知情同意代理问题研究

随着社会和科技发展，“知情同意代理”的研究日益受到关注。在共词聚类分析中，“代理同意”，“生前预嘱”等词汇被聚集在一起，这表明国内外已有学者对知情同意代理问题展开广泛研究。屠亦伟[19]通过对患者知情同意权相关法规的研究，对保障紧急救治过程中行使知情同意代理应采取的措施；方志华[20]等精神卫生领域行使代理同意的伦理风险进行研究并针对性的提出应对措施;李彩虹[21]等研究了儿童非治疗性医学试验中“知情同意”的问题并构造了“知情同意”的条件及其判定准则;LaneT[22]等对急危重症知情同意代理流程进行研究，通过调查访谈在研究人员中达成共识：“信任”对增强危重症患知情同意代理流程的有效性具有重要意义；ShepherdV[23]等开发了一种新的测量项目——CONCORD 量表，经测试，该量表可用于评估缺乏同意能力的成年人的代理人所做决策的质量，并提供了一种评估干预措施以改善代理人决策质量的方法。

（4）伦理问题和个人信息保护的研究

由于知情同意过程中涉及到个人隐私和伦理问题，因此国外研究也逐渐关注到个人信息保护和伦理问题的研究。共现词网络图中，个人信息保护、伦理问题、伦理标准等词汇共现频率较高，说明国内外研究对于个人信息保护和伦理问题的研究也越来越重视。范贞从民法的民事行为能力的角度分析了未成年、精神神经疾病患者等特殊群体的知情同意能力[24]，提出完善《药物临床试验质量管理规范》有关“弱势受试者定义”的建议；孟祥菡[25]探讨了受试者知情同意权受法律保护的发展历程，对“知情”与“同意”群体、行使权力过程等多方面进行了分析，并指出随着近些年知情同意权受到侵犯的事件发生率逐年上升，得到了法学理论、词法事务界更多关注；满洪杰[26]认为“知情同意权”是受试者权力的核心，同时指出，知情同意能力受限群体是否参与试验应从有益性与风险性两方面进行评估，其权力应当受到特别保护；丁磊[27]深入研究了对侵害患者知情同意权的司法救济，提出了细化侵害患者知情同意权民事责任的具体建议；BirosM[28]探讨了弱势群体的知情同意能力，从临床研究人员的角度论述了目前知情同意流程的局限性并指出平衡科学进步的需要和保护人类受试者的需要对于研究的道德实践至关重要。

3.3 研究热点和重要发现的总结

综上所述，国内外知情同意能力研究主题和涉及领域呈现出多样性和复杂性。伦理审查和知情同意过程的规范化、医患关系和患者参与度的提高、知情同意代理问题研究、伦理问题和个人信息保护的研究等是国内外知情同意能力研究的主要热点和趋势。此外，国内外研究中也存在一些差异，可能与不同国家或地区的医疗伦理法规、研究重点和文化背景有关。

3.4 对未来研究的建议

未来的国内外知情同意能力研究需要关注和解决实践中的问题和挑战，建议深入对以下几方面进行深入研究与探讨：

（1）知情同意权的保障和限制

通过探讨如何加强知情同意权的教育和普及，提高患者对其知情权益的认识和保护意识来实现以避免知情同意权的被滥用或过度限制。

（2）知情同意过程伦理及有效性研究

通过提高知情同意过程的透明度和公正性，使得患者能够更加全面地了解研究相关信息，并自主做出知情同意决策，进一步解决知情同意过程中涉及到的伦理问题，平衡研究进展和患者权益，提高知情同意过程的有效性和质量。

（3）知情同意代理问题研究

改进现有的知情代理制度，通过加强对知情代理的培训和监管，提高其决策的质量和有效性并更好地保障患者的知情权益。

（4）关注和保护个人信息

制定相应的政策和法规，加强对个人信息的管理和保护，防止其被滥用或泄露来实现，以保障患者个人信息的安全和隐私。

（5）推进多学科合作和跨国合作

通过加强学科交叉和跨国合作，促进知情同意能力研究的跨界融合和创新发展，促进知情同意能力研究的发展和进步。

4 小结

知情同意能力是医学伦理学中一个重要的概念，涉及到患者的自主权和医生的职业道德。本文使用了高频关键词分析、共现词网络图和共词聚类分析等方法，对国内外知情同意能力研究热点进行了探讨。研究发现，知情同意能力研究是一个多学科交叉的领域，涉及到医学、法律、伦理学等多个学科。近年来，知情同意能力研究的热点和趋势多样，主要包括伦理审查和知情同意过程的规范化、医患关系和患者参与度的提高、知情同意代理问题研究、伦理问题和个人信息保护的研究等。这些热点和趋势反映了知情同意能力研究的复杂性和多样性。不同国家或地区的医疗伦理法规、研究重点和文化背景等也可能导致国内外研究存在一些差异。

未来的研究可以从以下几个方面展开：进一步探索知情同意能力的评估方法、加强对于认知功能障碍患者知情同意能力的研究和保护措施的制定、提高伦理标准和规范化程度、知情同意代理问题研究、个人信息保护的研究和实践以及加强知情同意能力研究成果在实践中的应用。这些研究对于促进知情同意实践、保障患者权利以及加强临床试验的伦理管理具有重要意义。总之，知情同意能力研究是一个复杂而重要的领域，需要不断地深入研究和探索，以推动伦理标准的提高和患者权利的保障。

PU-094

基于 GEO 和 TCGA 数据库探究肝癌早期诊断标志物

陈凯，王晴晴

开封市人民医院 科研教学部

随着大数据时代的到来，生物信息学受到广泛关注，并逐渐在生物医学研究和疾病机制探索中发挥重要作用。根据 2018 年全球癌症统计报告，2018 年肝癌成为世界上第六大最常见的癌症和第四大癌症死亡原因[1]。在所有肝癌病例中，90%以上是肝细胞癌 (HCC)，化疗和免疫疗法是最佳治疗选择。肝癌的主要危险因素是乙型肝炎病毒(HBV)[2-4] ，丙型肝炎病毒(HCV)[5-7] ，黄曲霉毒素污染的食物[8] ，大量饮酒[6,9,10] 。据统计，在世界上不同的国家和地区肝癌形成的危险因素不同。在中国、东非等大多数高危地区，慢性乙型肝炎病毒感染和接触黄曲霉毒素是肝癌的主要决定因素。相反，在日本和埃及，HCV 感染是肝癌的主要原因[11,12]。只有 5％至 15％的患者适合手术切除，这仅适合早期患者，并且由于肝脏再生能力减弱，通常没有肝硬化；与左肝切除术相比，右肝切除术术后并发症的风险更高[13]。大多数癌症是在晚期被诊断出来的，尤其是肝癌。现代医学研究表明，肝脏没有疼痛感，即使有任何肝脏疾病已经开始，身体也不能通过疼痛反馈机制感知或识别它。因此，肝脏疾病的临床表现非常轻，大多数肝癌患者由于缺乏及时的症状表现而在晚期被诊断[14-17]。早期肝癌的治愈率是非常乐观的，因此，如果在早期就能作出诊断，那么对癌肿的治疗将不会像最后阶段那样剧烈。

甲胎蛋白(AFP)是目前临床上唯一用于肝癌早期诊断的生物标志物，然而并不是非常准确的肝癌诊断生物标志物。32% 至 59% 的肝癌患者已被证明具有正常的 AFP 水平[18]。因此，寻找新的肝癌诊断标志物对于肝癌的准确诊断具有重要意义。对于癌症患者而言，癌症的预后和生存时间对于提高患者的生活质量以及采取何种诊断和治疗方案至关重要。本研究以 TGCA 和 GEO 数据库中的肝癌患者数据作为检索标准，通过数据挖掘来识别肝癌的诊断生物标志物。目的是提高肝癌早期诊断的准确性，实现早期发现、早期治疗，降低死亡率。

1 材料和方法

1.1.数据处理

肝癌基因表达数据的获取来源于 GEO 和 TCGA 数据库 。其中 TCGA 数据库包括 30 例正常肝组织标本和 263 例肝癌标本，并结合临床资料进行分析。GEO 数据库(GSE25097)，包括 212 个正常样本和 236 个肿瘤样本的信息。GSE25097 和 TCGA LIHC 数据集的原始微阵列数据分别用 Rlanguage 进行分析，以筛选差异表达基因(DEGs)。利用 DrawVenn Diagram 在线工具计算由两个不同数据集得到的两个差异 DEG 的交集，这两个差异 DEG 代表了常见的差异表达基因(Common DEGs)。

1.2.蛋白质-蛋白质相互作用网络

相互作用基因/蛋白质检索工具(STRING)是一个在线蛋白质相互作用工具，可以整合已知的蛋白质-蛋白质相关数据，建立蛋白质之间的上游或下游关系[18]。将常见的 DEGs 导入 STRING 软件

中，构建可视化蛋白质-蛋白质相互作用(PPI)网络。同时，利用 Cytoscape 软件(Cytos-cape _ v3.6.1) 筛选中心基因。以度值>＝5 的 top22 基因被选为中心基因。

1.3 绘制 Hub 基因的 ROC 曲线

采用 pROC 软件包对 22 个 hub 基因进行（ROC）曲线分析。以 AUC > 90%为确定 Hub 基因诊断意义的截止值。

2 结果

2.1 Hub 基因筛选

对 GSE25097 数据集进行 p 值< 0.05，且|logFC| > 2 的 R,DEGs 处理，筛选 570 个基因。对 TCGA LIHC 数据集进行分析，采用 DEG- seq2，以 p< 0.05 和|logFC| > 2 为截止标准，鉴定出 2032 个符合标准的基因。为确认基因序列在肝癌中的可靠性，我们获得了两个数据集的共同基因序列，其中包括 92 个基因。将 92 个 DEGs 输入 STRING，以建立 PPI 网络。将 PPI 网络图导出到 Cytoscape (3.2.1)软件计算度值。取度值 > = 5 的基因作为 Hub 基因，共获得 20 个 Hub 基因(表 1)。

Number	Name	Betweenness	BottleNeck	Closeness	Clustering Coefficient	Degree	DMNC	EcCentricity	EPC	MCC	MNC	Radiality	Stress
1	CYP1A2	577	5	29.39286	0.2	6	0.309	0.13562	19.57	7	3	7.7361	822
2	RRM2	0	1	25.28452	1	7	0.7683	0.11867	23.64	5040	7	7.0433	0
3	AURKA	967.42857	11	32.02619	0 4	11	0.6415	0.13562	24.93	5045	8	7.8002	1918
4	SOCS3	631.69841	6	31.05952	0.09524	7	0.3078	0.13562	25.03	7	2	7.8259	1378
5	FOS	617.74127	17	33.3	0.30556	9	0.3207	0.15823	27.17	29	8	8.0311	1918
6	PCK1	570	5	24.21786	0 1	5	0.3078	0.11867	10.7	5	2	7.0176	1074
7	FOSB	35.46667	1	27.55952	0 7	5	0.4538	0.13562	24.25	18	5	7.5821	166
8	PTTG1	0	1	25.28452	1	7	0.7683	0.11867	23.95	5040	7	7.0433	0
9	UBE2C	146.37619	1	27.25119	0.75	8	0.7683	0.11867	24.36	5041	7	7.2229	264
10	GPC3	44.47619	2	26.7	0 53333	6	0.3804	0.15823	24.62	20	6	7.364	138
11	DCN	303.6	3	28.55	0.28571	7	0 2853	0 15823	24.75	13	6	7.5308	782
12	EGR1	228.02381	3	30.21667	0.42857	7	0.4279	0.15823	25.63	25	6	7.8002	602
13	GMNN	0	1	25.28452	1	7	0.7683	0.11867	24.18	5040	7	7.0433	0
14	TOP2A	332.07302	3	30.27619	0.78571	8	0.6415	0.13562	24 54	.5042	8	7 7361	938
15	SPP1	1519.02063	30	35.65	0.22727	12	0.2545	0.18987	28.52	35	1	8.1723	3920
16	CXCL12	1415.00476	15	32.73333	0.11111	9	0.2842	0.15823	24.1	11	4	7.9926	2458
17	UBE2T	0	1	25.28452	1	7	0.7683	0.11867	24.05	5040	7	7.0433	0
18	AFP	281.74603	4	30 05	0.33333	7	0.3328	0.15823	26.15	17	6	7.7874	770
19	NUSAP1	146	2	25.78452	0.75	8	0.7683	0.11867	23.93	5041	7	7.0561	338
20	ESR1	2780.97619	58	38.78333	0.16667	13	0.2036	0.15823	29.27	28	1	8.416	5698

表 1 度值> = 5 的 Hub 基因

2.2 Hub 基因在肝癌患者中的表达

在分析了 20 个 Hub 基因在肝癌和正常肝组织中的表达后，发现 NUSAP1，TOP2A，SPP1，AURKA，UBE2C，AFP，CXCL12，GMNN.PTTG1， SPARCL1，dCN，SOCS3，RRM2 和 PCk1 等 13 个基因在肝癌组织中高度表达。

2.3 Hub 基因的 Roc 曲线分析

采用 pROC 软件包对 20 个 Hub 基因进行 ROC 曲线分析。以 AUC >为 90%为截断值，发现 20 个 AUC >为 90%的 Hub 基因中有 16 个(图 1)，分别包括 NUSAP1，TOP2A，SPP1，AURKA，UBE2C，AFP，CXCL12，GMNN、PTTG1， SPARCL1，FOS，DCN，SOCS3，RRM2，FOSB 和 PCk1。这些基因的表达水平在区分正常组织和肝癌组织方面具有很高的准确性，并可能成为一种潜在的“肿瘤生物标志物”。

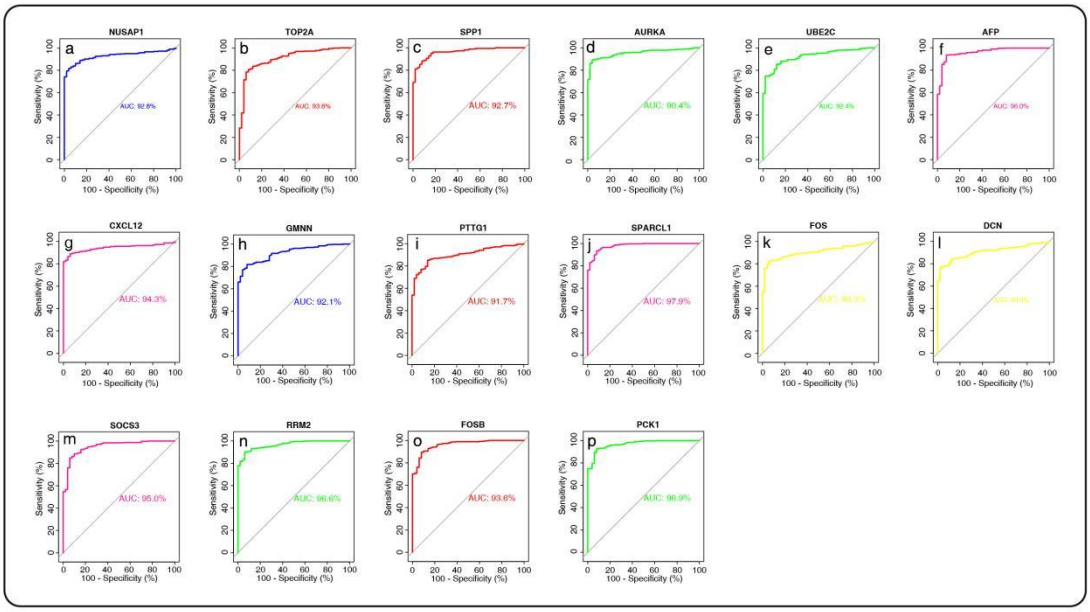


图 1 Hub 基因的 ROC 曲线

(a)NUSAP1、(b)TOP2A、(c) SPP1、(d)AURKA、(e) UBE2C、(f) AFP、(g)CXCL12、(h)GMNN、(i)PTTG1、(j) SPARCL1、(k)FOS、(l)DCN，(m)SOCS3、(n)RRM2、(o)FOSB、(p)PCk1

3 讨论

大多数肝癌患者在肝癌晚期出现症状后才求医，因此早期诊断肝癌对治疗具有重要意义。然而，由于肝癌的发病机制尚不明确，导致肝癌发病率高、侵袭性强、预后差。数据显示，约 70%的肝癌在切除或消融后 5 年内复发[18]。AFP 是 50 年前发现的肝癌诊断生物标志物，但有出现假阳性和假阴性的结果，存在诊断不准确的问题，造成了诊断的偏差。因此，有必要为肝癌的诊断寻找新的、更准确的生物标志物。为实现这一目标，在本研究中，为验证 Hub 基因在癌症和正常样本分类中的性能，应用了来自不同来源的两个数据集。通过使用不同的分类器进行验证，获得了较高

的分类精度。通过对比 20 个 hub 基因集，可以发现两个基因集的预测准确率相差不大，但由于基因数量减少，临床诊断的速度得以提高。

蛋白互作网络(PPI 网络)分析构建了 92 个重叠差异表达基因 ，并根据度值从 92 个重叠差异表达基因中筛选出 20 个 HubGene (表 1)。ROC 曲线是反映敏感性与特异性关系的曲线，对准确诊断疾病具有重要意义。采用 ROC 曲线分析 20 个 AUC 大于 90%为阈值的 Hub 基因，得到 16 个 Hub 基因。它们分别为 NUSAP1，TOP2A，SPP1，AURKA，UBE2C，AFP，CXCL12，GMNN、PTTG1， SPARCL1，FOS，DCN，SOCS3，RRM2，FOSB 和 PCK1。16 个 Hub 基因在肝癌中的表达水平可以准确区分正常肝组织和肝癌，因此这 16 个基因可以作为肝癌早期诊断的诊断生物标志物。而 ESR1 的 AUC 值为 68.7%，说明 ESR1 基因的准确诊断率较低，不适合作为诊断性生物标志物。这也证实了 SPP1 是一个高危基因。TOP2A 编码 DNA 拓扑异构酶 II α，参与染色单体分离、细胞周期进程和扭转应力缓解等过程。据报道，TOP2A 经常与 HER-2 共同扩增，从而降低膀胱癌和乳腺癌的临床结果[19]。RRM2 的高表达可能会导致非小细胞肺癌患者的预后不佳,因为 BCL2 是细胞凋亡的重要调节剂[20]。此外,研究显示,RRM2 可能是各种癌症的治疗靶点,如乳腺癌、非小细胞肺癌和膀胱癌[20]。同样，关于 SPP1，Lu C 等人发现，肝癌中 OPN 蛋白（由 SPP1 基因编码）的沉默降低了细胞克隆的数量和增殖率，并通过体内药效学观察到，小鼠的肿瘤体积降低了[17]。SPP1，AURKA，NUSAP1，TOP2A，UBE2C，AFP，gMNN，PTtg1，RRM2 和 SPARCL1 在肝癌发生过程中高表达，因此，这些基因可进行进一步研究，可能可以作为肝癌诊断的标志物。

综上所述,肝癌的发生、发展是一个多分子、多步骤、多因素的动态生物学过程，目前人们对其机制的认识还很有限。以往的研究主要集中在单个基因对肿瘤的影响，但细胞癌变的过程往往涉及多个基因的变化。本研究通过基因芯片筛选肝癌的差异表达基因,有助于进一步理解肝癌的发病机制。同时,通过 TCGA 数据库初步验证了核心基因的表达,为肝癌提供了新的研究方向,有待于进一步研究其具体的功能和作用。

PU-095

体力活动对炎症性肠病风险的影响和 BMI 的介导作用：孟德尔随机化研究

赵宇慧[2] 李红霞[2] 贾林霖[2] 郝佳炜[2] 陈淼然[2] 贺培凤[1] 于琦[3]

[1]山西医科大学临床决策大数据重点实验室

[2]山西医科大学细胞生理学重点实验室

[3]山西医科大学管理学院

1 引言

炎症性肠病（IBD），包括克罗恩病（CD）和溃疡性结肠炎（UC），是一种慢性并危及生命的炎症性疾病 [1]。CD 表现为透壁性炎症和纵向溃疡，影响整个胃肠道，特别是回肠末端和邻近结肠，而 UC 表现为粘膜炎症，仅限于结肠 [2]。近年来，我们对这些疾病发病机制的认识取得突破，使 IBD 有了更广泛的治疗选择；然而 CD 和 UC 是无法治愈的，患者往往需要终身使用药物治疗。为了提高 IBD 患者的生活质量，确定潜在的可改变的危险因素和治疗方法，以减缓病情的进展，甚至预防这些疾病的发生，是很重要的。

体育活动（PA）被定义为由骨骼肌动作产生的导致超过基础水平能量消耗的任何机械运动 [3]。早期的研究表明，与久坐不动的职业相比，涉及更多体力劳动的职业与更低的 IBD 风险相关 [4]。Chan 等人的一项研究报告了一个不同的结果 [5]，即 PA 与 UC 或 CD 无任何关联。许多观察性研究报告称，缺乏运动和久坐时间的增加与肥胖风险的增加有关 [6]。与此同时，大量和一致的流行病学数据表明，肥胖是发生 IBD 的高危因素之一 [7 8]。这些发现表明，在参与较低水平 PA 的个体中观察到的 CD 风险的增加可能部分是由体重增加介导的。然而，迄今为止，还没有研究证实这一假设。此外，传统的观察性研究可能会受到潜在的混杂因素、测量误差和反向因果关系的影响。孟德尔随机化分析 (MR) 有效地避免了潜在的混杂因素和反向因果关系。作为 MR 的最新发展，多变量 MR（MVMR）方法被用于调查中介。

在本研究中，我们旨在通过使用单变量双样本 MR (UVMR) 方法分析最新的 GWAS 数据，确定自我报告的闲暇时间中度至剧烈强度体力活动（MVPA）/闲暇屏幕时间 （LST）与 IBD（包括其两个亚型，即 CD 和 UC）风险之间的潜在因果关系。此外，利用 MVMR 方法，我们还旨在确定 BMI 是否介导 MVPA/LST 对 IBD（包括 CD 和 UC）的影响，以及 BMI 介导的影响比例。

2 方法

2.1 数据来源

2.1.1 MVPA 和 LST 的汇总统计数据

从最新发表的全基因组关联研究（GWASs）的荟萃分析中提取 MVPA 和 LST 的汇总统计量。该荟萃分析分析了 51 个 GWASs，共包括 703,901 名欧洲人（MVPA： n = 608,595；LST： n = 526,725）[9]。游泳和慢跑等活动被视为 MVPA，看电视、玩电子游戏、坐在电脑前等活动被视为 LST。GWAS 的结果根据主要协变量（包括评估中心、基因分型阵列、年龄和季节）进行了调整。

2.1.2 IBD 的汇总统计数据

IBD 的摘要统计数据是从一项 IBD GWAS 的荟萃分析中提取的 [10]，该荟萃分析共包括 25,042 例临床诊断的 IBD 病例（与 34,915 例对照）、12,194 例临床诊断的 CD 病例（与 28,072 例对照）和 12,366 例临床诊断的 UC 病例（与 33,609 例对照）（摘要见表 1）。对每个队列的前 10 个主成分进行了调整。

2.1.3 BMI 汇总水平数据

与 BMI 相关的全基因组重大（$P < 5 \times 10^{-8}$）遗传变异是从人体测量特征的遗传调查联盟（Genetic Investigation of Anthropometric Traits consortium）发表的一项全球基因组研究荟萃分析中提取的[11]，该分析共纳入了 681 275 名欧洲血统的个体（摘要见表 1）。对年龄、年龄平方和研究特异性协变量（包括用于调整人群分层的主成分）进行了调整。

2.2 IV 的选择

提取与 PA 相关的单核苷酸多态性（SNPs），MVPA 和 LST 的全基因组显著性阈值为 $P < 5\times10^{-8}$。满足相关性假设的 SNPs（$r^2= 0.01$，聚类距离= 10,000 kb）被认为与 MVPA 和 LST 相关，具有全基因组意义。此外，它们还被验证为是独立的，而不是处于连锁不平衡状态。此外，本研究只分析了有关欧洲血统个体的数据，并且人口分层在任何情况下都没有违反独立性假设（图 1）。

2.3 统计分析

2.3.1 总影响、直接影响和间接影响

首先，进行 UVMR 分析，以确定 MVPA 和 LST 对 IBD （包括 CD 和 UC) 的净效应，该效应被视为总效应。接着，再次进行 UVMR 分析，以估计 MVPA 和 LST 对 BMI 的影响。随后，在对 MVPA 和 LST 进行调整后，进行 MVMR 分析，以确定中介因子对 IBD （包括 CD 和 UC) 的影响。

暴露对结果的总效应可分为直接效应和间接效应。直接效应被定义为暴露对结果的影响，这种影响没有通过因果中间物进行中介，而间接效应被定义为暴露通过候选中介物对结果的影响（图 2)。在这里，间接效应采用系数乘积法计算，即在调整 MVPA 和 LST 后，将 MVPA 和 LST 对 BMI 影响的 MR 估计值与 BMI 对 IBD（包括 CD 和 UC）影响的 MR 估计值相乘。此外，通过间接效应除以总效应，计算出 BMI 所介导的效应比例。BMI 中介效应比例的置信区间（CIs)和标准误差采用 delta 法估算。

2.3.2 UVMR 和 MVMR 分析

为了评估 MVPA 和 LST 对 IBD（包括 CD 和 UC）发病风险的影响，我们使用了固定效应反方差加权法（IVW）作为 UVMR 和 MVMR 的主要方法。IVW 法用于估计遗传变异的几率比（OR），是 MR 研究中使用最广泛的方法。

3 结果

3.1 MVPA 和 LST 对 IBD、CD 和 UC 的总影响和直接影响

在 UVMR 中，IVW 分析表明 MVPA 与 IBD 以及 MVPA 与 CD 之间存在保护性因果关系 （表 2，图 3 A-B)。MVPA 每增加一个 S.D.，IBD 的相对几率降低 33% （OR = 0.67，95% CI = 0.47-0.95，P = 0.03)，CD 的相对几率降低 47% （OR = 0.53，95% CI = 0.30-0.94，P = 0.03)。IVW 分析表明，基因预测的 LST 越高，患 IBD 和 CD 的风险越高（表 3，图 4 A-B)。LST 每增加一个

S.D.，IBD 的相对几率（OR = 1.21，95% CI = 1.06-1.38，P = 0.004)增加 21%，CD 的相对几率（OR = 1.25，95% CI = 1.07-1.45，P = 0.005)增加 25%。相反，MVPA 与罹患 UC 的风险降低并无统计学意义（OR = 0.76，95% CI = 0.49-1.18，P = 0.23)（表 2，图 3 C)。此外，在 LST 和 UC 之间没有观察到有统计学意义的关系证据（OR =1.12，95% CI =0.95-1.33，P =0.18)（表 3，图 4 C)。使用其他 MR 方法，包括加权中位数法和 MR-Egger 法，对 IBD （包括 CD 和 UC) 也观察到了类似的因果关系估计值。

在 MVMR 分析中，对 BMI 进行调整后，观察到 MVPA 与 IBD 之间存在因果关系。MVPA 每增加一个 S.D.，IBD 的相对几率降低 32% （OR = 0.68，95% CI = 0.46-0.99，P = 0.042），CD 的相对几率降低 42% （OR =0.58，95% CI = 0.36-0.94，P = 0.026）（表 4）。相反，调整 BMI 后，LST 与 IBD（OR =1.07，95% CI =0.83-1.39，P =0.586）或 LST 与 CD（OR =1.11，95% CI =0.80-1.54，P =0.543）之间的因果关系没有统计学意义（表 4）。

3.2 MVPA 和 LST 对 BMI 的影响

为了探讨 MVPA 和 LST 通过 BMI 对 IBD（包括 CD 和 UC）的影响，我们进行了 UVMR 分析，以估计 MVPA 和 LST 对 BMI 的影响。结果显示，MVPA 水平越高，BMI 越低；LST 水平越高，BMI 越高。MVPA 每增加一个标准差，BMI 下降 0.15 个标准差单位（OR = 0.85, 95% CI = 0.74-0.97, P = 0.016）；LST 每增加一个标准差，BMI 增加 0.15 个标准差单位（OR = 1.15, 95% CI = 1.09-1.20, P = 1.48×10^{-8}）。

3.3 调整 MVPA 和 LST 后，BMI 对 IBD、CD 和 UC 的影响

在 MVPA-BMI-IBD 和 MVPA-BMI-CD 的 MVMR 分析中，MVPA 和 BMI 的条件 F 统计量均大于 10。在调整 MVPA 后，BMI 对 IBD 的直接影响被确定为 OR 1.02 （95% CI = 0.89-1.16，P = 0.769），BMI 的中介比例被确定为 0.8%。同样，在调整了 MVPA 后，BMI 对 CD 的直接影响被确定为 OR 1.15 （95% CI = 0.97-1.36，P = 0.104），BMI 介导的比例被确定为 3.7%。

在 LST-BMI-IBD 和 LST-BMI-CD 的 MVMR 分析中，LST 和 BMI 的条件 F 统计量均大于 10。在调整 LST 后，BMI 对 IBD 的直接影响被确定为 OR 1.05 （95% CI = 0.90-1.23，P = 0.528），BMI 介导的比例被确定为 3.5%。在调整 LST 后，BMI 对 CD 的直接效应为 OR 1.19（95% CI = 0.98-1.45，P = 0.081），BMI 的中介效应占总效应的 11% （P = 0.095）（表 4）。一系列敏感性分析，包括 MR-Egger 截距检验、Cochran Q 检验、留一分析和 MR-PRESSO 全局检验，证明了结果的稳健性（图 3 D-I,图 4 D-F）。

4 结语

我们的研究结果证实了增加 MVPA 和减少 LST 可以预防 CD，而且这种效应部分是由 BMI 介导的。因此，通过促进生活方式的管理，如减少休闲久坐活动和鼓励适当的运动，可以降低 CD 的发病率。此外，没有体力活动的人应该监测他们的 BMI，以防止 CD 的发展。这种做法可以减少 CD 发展的风险，以及在人类生产力方面对卫生保健和一般公共卫生回报的负担 [12].

表 1：纳入 MR 分析的 GWAS 的特征。

表型	GWAS 联盟	人种	样本量	PMID
MVPA	MTAG	欧洲	608,595	36071172
LST	MTAG	欧洲	526,725	36071172

IBD	IBD meta-GWAS	欧洲	25,042/34,915	28067908
CD	CD meta-GWAS	欧洲	12,194/28,072	28067908
UC	UC meta-GWAS	欧洲	12,366/33,609	28067908
BMI	GIANT	欧洲	681,275	30124842

表 2：MVPA 对 IBD、CD 和 UC 的因果效应。

暴露	结局	SNPs	方法	OR	95%CI	P 值	Q_p 值	截距 p 值
MVPA	IBD	9	IVW	0.665	0.465,0.952	0.026	0.315	0.538
			MR-Egger	0.364	0.056,2.353	0.324	0.267	
			WM	0.671	0.424,1.062	0.088		
	CD	9	IVW	0.533	0.302,0.942	0.030	0.079	0.654
			MR-Egger	1.076	0.054,21.636	0.963	0.057	
			WM	0.588	0.320,1.083	0.089		
	UC	9	IVW	0.761	0.489,1.183	0.225	0.368	0.093
			MR-Egger	0.096	0.011,0.809	0.068	0.668	
			WM	0.704	0.396,1.254	0.234		

Q_p 值采用 MR-Egger 和 IVW 方法检测异质性。截距 p 值来自 MR-Egger 回归的多效性检验。

暴露	结局	SNPs	方法	OR	95%CI	P 值	Q_p 值	截距 p 值
LST	IBD	85	IVW	1.213	1.063,1.384	0.004	0.027	0.698
			MR-Egger	1.079	0.591,1.973	0.804	0.024	
			WM	1.163	0.978,1.382	0.088		
	CD	86	IVW	1.245	1.070,1.449	0.005	0.351	0.313
			MR-Egger	1.765	0.884,3.526	0.111	0.353	
			WM	1.213	0.973,1.513	0.086		
	UC	87	IVW	1.122	0.950,1.325	0.176	0.026	0.371
			MR-Egger	0.800	0.376,1.702	0.564	0.026	
			WM	1.107	0.897,1.367	0.342		

表 3：LST 对 IBD、CD 和 UC 的因果效应。

Q_p 值采用 MR-Egger 和 IVW 方法检测异质性。截距 p 值来自 MR-Egger 回归的多效性检验。

表 4：我们研究中的中介分析。

暴露	中介	结局	a	se_a	b	se_b	c'	se_c'	a*b	se_a*b	c	中介比例
MVPA	MI	BBD	-0.167	0.069	0.020	0.067	-0.390	0.192	-0.003	0.011	-0.407	0.008
MVPA	MI	BCD	-0.167	0.069	0.140	0.086	-0.549	0.246	-0.023	0.017	-0.628	0.037
LST	MI	BBD	0.138	0.024	0.050	0.079	0.071	0.131	0.007	0.011	0.193	0.035
LST	MI	BCD	0.138	0.024	0.176	0.101	0.102	0.167	0.024	0.014	0.219	0.110

a 代表暴露与中介因子之间关联的估计值

b 代表调整暴露后，中介因子与结果之间关联的估计值

c' 代表调整中介因子后，暴露与结果之间直接影响的估计值

a*b 代表中介效应对结果的间接效应的估计值

c 代表 UVMR 结果中暴露与结果之间关联的估计值

中介比例按 a*b / c 计算

标准误差（se）采用 delta 法计算。

图 1：暴露和结果的孟德尔随机模型。该设计假设遗传工具变量与暴露相关，但与混杂因素无关，遗传工具变量仅通过暴露影响结果。遗传工具变量表示单核苷酸多态性。

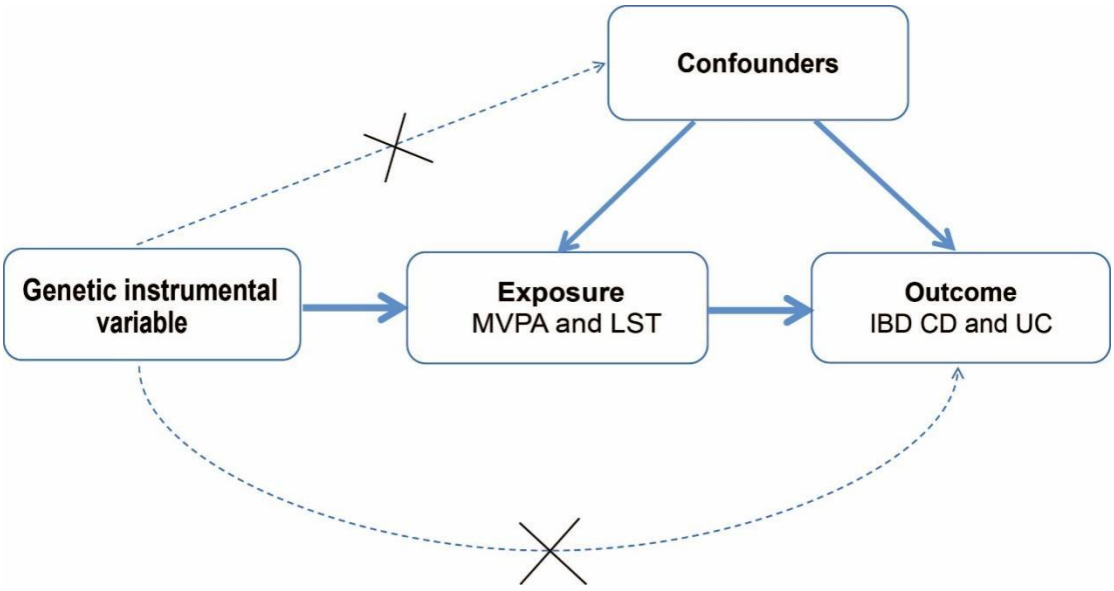


图 2：直接无环图说明 MVPA 和 LST 对 IBD、CD 和 UC 的总效应、直接效应和间接效应。直接无环图显示了 MVPA 和 LST 对 IBD、CD 和 UC 的总效应的假设方向 （A），以及 MVPA 和 LST 对 BMI 的效应的假设方向 （B），BMI 可能部分介导了 MVPA 和 LST 对 IBD、CD 和 UC 的效应。

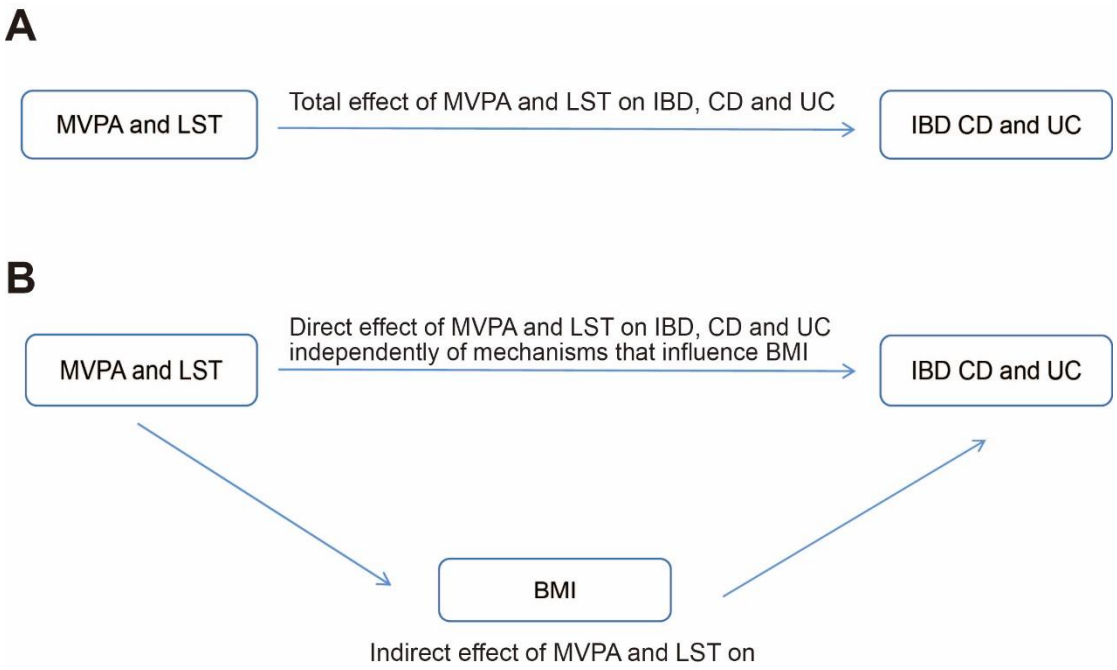


图 3：MVPA 与 IBD、CD 和 UC 相关性的散点图、漏斗图和缺失分析。（A-C） IVW、MR-Egger 和 WM 分析方法的散点图显示了 MVPA 对 IBD （A）、CD （B） 和 UC （C） 的影响。（D-F） MVPA 对 IBD （D）、CD （E） 和 UC （F） 的漏斗图，表明没有证据表明存在实质性异质性。（G-I） 留一分析，探讨 MVPA 与 IBD （G）、CD （H） 和 UC （I） 的因果关系是否由单个特定 SNP 驱动。

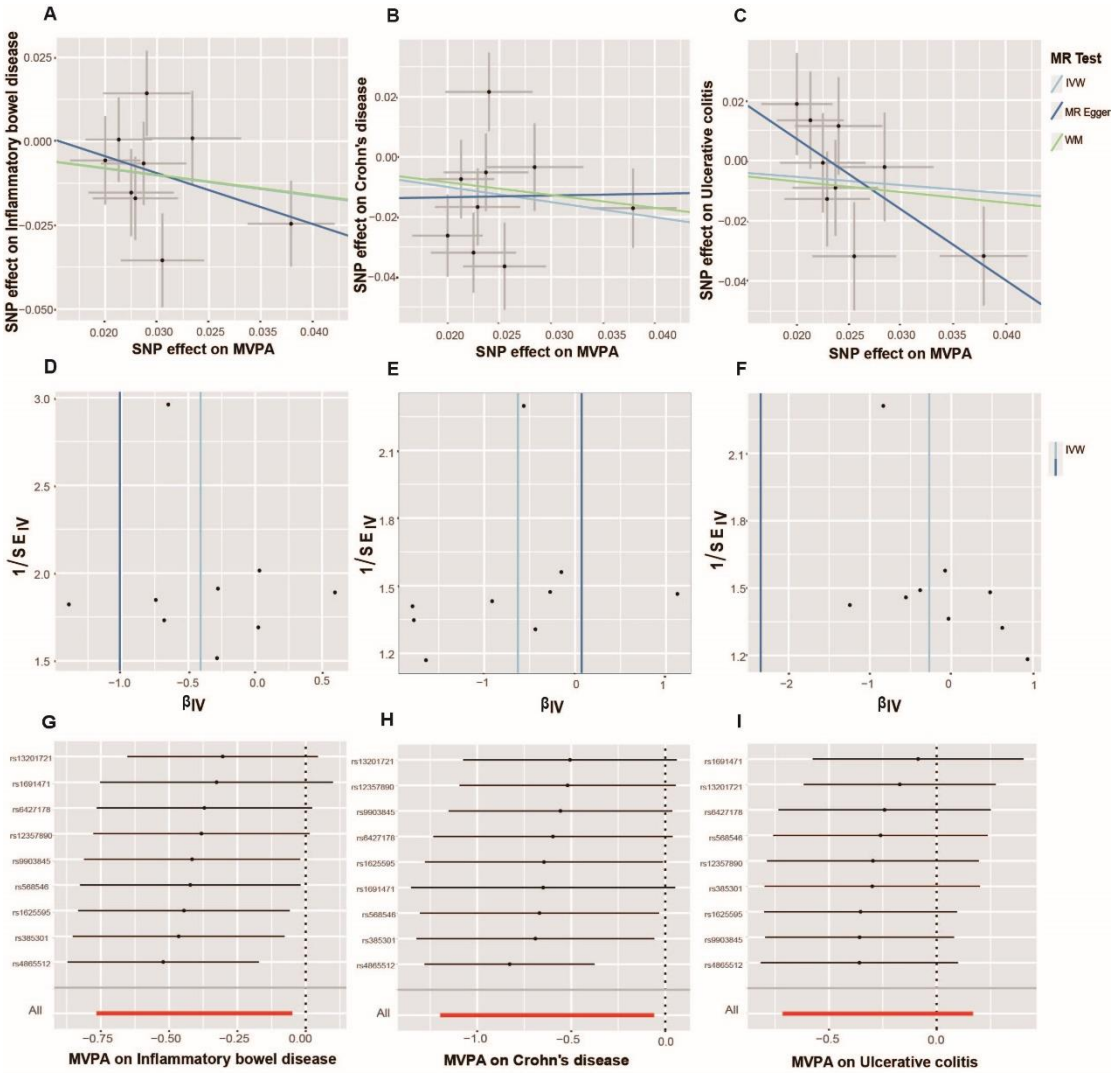
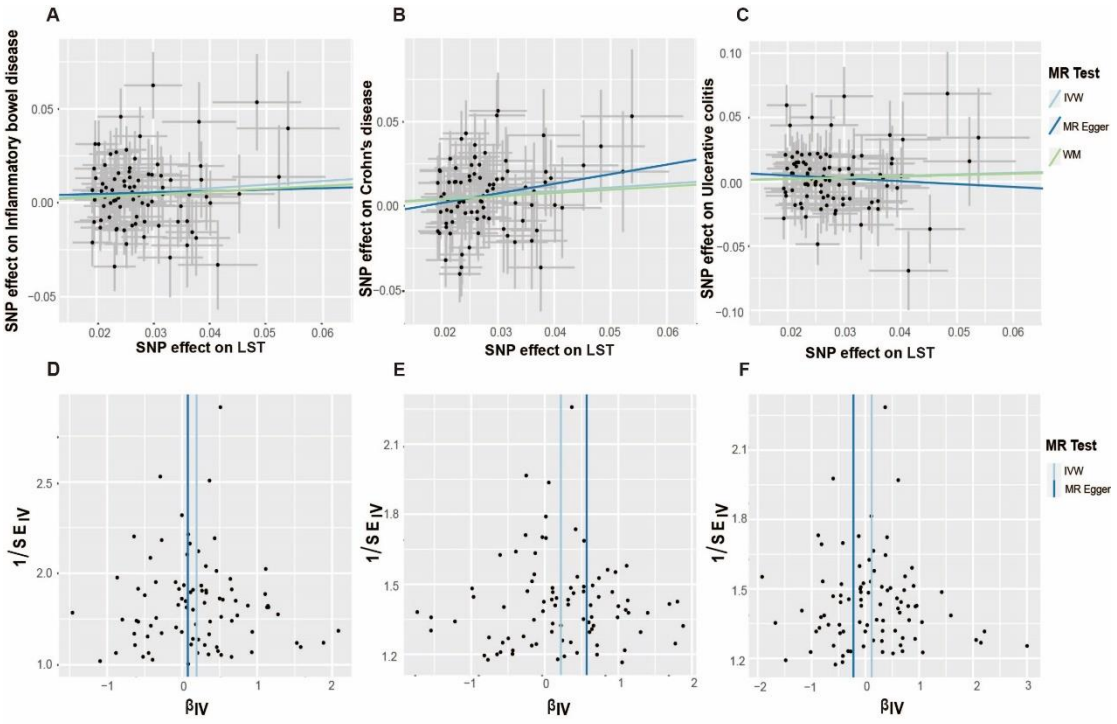


图 4：LST 与 IBD、CD 和 UC 相关性的散点图和漏斗图。（A-C） IVW、MR-Egger 和 WM 分析方法的散点图显示了 LST 对 IBD （A）、CD （B） 和 UC （C） 的影响。（D-F） LST 对 IBD （D）、CD （E） 和 UC （F） 的漏斗图，表明没有证据表明存在实质性异质性。


A
B
C
SNP effect on Inflammatory bowel disease
SNP effect on Crohn's disease
SNP effect on Ulcerative colitis
SNP effect on LST
MR Test
IVW
MR Egger
WM
D
E
F
1/SE IV
βIV

PU-096

指南制订对文献检索的要求与实践

丁文婧 仇晓春
上海交通大学医学院图书馆

临床实践指南 (下文简称“指南”)指基于系统评价的证据和平衡不同干预措施的利弊，形成的能为患者提供最佳保健服务的推荐意见[1]。高质量的指南能够规范诊疗行为，优化医疗资源配置，减少不同区域或不同机构临床实践的差异，提升医疗服务质量。近年来指南发布数量增长迅速，越来越多的学术机构和研究团队投入指南撰写。在指南制订过程中，文献检索是必要且非常关键的步骤。不同于一般的医学文献检索，制订指南需要检索的数据库和资源类型数量庞大，检索方式不尽相同，检索词纷繁复杂，检索式的撰写更是挑战专业检索能力，并且在指南制订的不同阶段对文献检索的内容和要求不同。全面规范的文献检索是完成高质量指南的前提。因此，本文拟采用文本分析法结合检索工作经验，就指南制订对文献检索的要求进行剖析，分享检索心得，以期为广大指南制订者提供借鉴和参考。

1 指南制订不同阶段对文献检索的要求

1.1 准备阶段

“工欲善其事必先利其器”。了解指南制订的原则方法、流程步骤和相关注意事项，才能让后续工作做到有的放矢，稳中求进。指南方法学的检索对象，一是指南制订手册、指南评价工具等指导文件，二是检索有关指南制订方法学研究的最新文献。

指南制订手册发布主体主要是国家权威指南制订机构或专业学会。例如，世界卫生组织（World Health Organization，WHO）指南评审委员会督促指南的制订过程并出版了《WHO 指南制订手册》[2]。Shabnam 等[3]比较了 19 部指南制订手册，对 27 个主要方面进行评估打分，得分较高的手册制订单位除了 WHO，还有英国国家卫生与临床优化研究所（National institute for health and clinical excellence，NICE）、瑞士国际健康中心（Swiss Centre for International Health，SCIH）和苏格兰地区学院之间指南网络（Scottish Intercollegiate Guidelines Network，SIGN）。其中，WHO、NICE、SCIH 指南手册对各项内容和流程的阐述最为详细。2017 年，由兰州大学陈耀龙教授发起，联合 12 个国家、7 个国际组织研发的卫生保健实践指南的报告条目（Reporting Items for Practice Guidelines in Healthcare, RIGHT）清单发布[4]，该报告规范已成为国内外指南规范化、科学化和透明化撰写的重要参考。国内多所高校学者联合编写的《中国制订/修订临床诊疗指南的指导原则（2022 版）》提供了符合国际标准、适应中国实践的循证指导和建议[5]。

指南评价工具尽管是用于评价指南质量，但其中罗列的评价条目可作为写作对照。如国际权威指南评价工具 AGREE Ⅱ（Appraisal of Guidelines Research and Evaluation II）、复旦王吉耀教授团队研发的中国临床实践指南评价体系（AGREE-China）等。2021 年，WHO 指南实施与知识转化合作中心、中华医学会杂志社指南与标准研究中心等联合成立了 STAR 工作组，发布了针对指南科学性、透明性和适用性的多维度评级工具 STAR[6]。在 STAR 评级网站上，不仅能查阅众多高质量最新指南而且还可浏览评级情况，所谓“他山之石，可以攻玉”，是指南撰写者学习借鉴的宝藏资源。

上述几个权威指南制订手册和评价工具通过网页检索或登录官方网站即可下载最新版本。另外，在文摘型数据库中检索有关方法学研究的最新文献也有助于了解指南制订的方法进展和学术动态，特别是像《BMJ》、《Annals of internal medicine》、《Journal of clinical epidemiology》等杂志经常刊载指南方法学、报告规范等文章，帮助指南制订者更新观念、开拓新思路。

1.2 注册与计划书撰写阶段

该阶段需要明确指南制订目的和应用对象，例如是制订新指南还是扩展现有指南。通常可先在指南数据库以及指南注册网站（如国际指南注册与透明化平台 Practice guideline Registration for transparency，PREPARE）检索是否已有相关主题指南发布或已注册开展，避免重复[7]。常用的指南检索网站有国际指南协作网（https://g-i-n.net/）、美国国家指南文库(https://ahrq.gov/gam/index.html)、英国国家卫生与临床优化研究所(https://nice.org.uk/guidance)、苏格兰校际指南网络（https://www.sign.ac.uk/our-guidelines/）、加拿大医学会临床实践指南信息库（https://joulecma.ca/cpg/）、澳大利亚临床实践指南平台（https://clinicalguidelines.gov.au）等[8]。若在文摘型数据库中检索指南文献，常用中文检索词：指南 or 临床指南 or 专家共识 or 标准 or 意见 or 推荐意见 or 诊疗规范。英文检索词：guideline or

guidance or clinical guideline or consensus or standards or opinion or recommendation or comments or clinical consensus statement。为确保在数据库中检索到的是最新版指南，有时要回到其学会或机构官方网站查看，同时留意是否有其他相关信息和补充修订材料发布，注意专业版指南和患者版指南、临床循证实践指南和专家共识、标准指南和快速指南等区别。

如果没有现成指南作内容参考，在采用 PICO（Population，Intervention，Comparison，and Outcomes）基本模式构建并优选临床问题的同时，重点检索当前可得的最佳研究证据。而最新高质量的系统评价能减少选择性引用的风险并提高决策的可靠性和精确性，是最佳证据的典范。系统评价是指运用减少偏倚的策略，全面检索、严格评价和综合针对某一具体问题的所有相关研究。根据 WHO 指南手册和 RIGHT 报告清单，须描述指南基于的系统评价是新制作还是使用已发表的，若用已发表的，应提供检索策略并采用国际公认的 GRADE 方法评价证据质量和推荐意见强度[9]。关于系统评价的检索，Cochrane 系统评价数据库(Cochrane Database of Systematic Reviews, CDSR)、澳大利亚 Joanna Briggs 循证卫生保健中心系统评价和证据应用报告数据库（The JBI database of systematic reviews and implementation reports）、Campbell 协作网（Campbell collaboration）是重要的检索资源。关于检索方法，《Cochrane 系统评价手册》[10]和 Campbell 协作网的系统评价检索指导[11]提供了详细攻略，不仅针对现有系统评价的全面检索，还涵盖了制作新系统评价的检索策略，有助于指南制订者开展快速系统评价。

1.3 正式撰写阶段

WHO 等多个机构的指南制订手册都强调应基于全面的系统检索和评价来撰写指南[12]。指南评价工具 AGREE Ⅱ也明确对指南各类证据的检索策略，包括检索数据库名称、检索词、检索时间、检索语言和详细检索式等都要进行报告(一般放在指南附录或补充材料里)。关于文献检索在绝大多数指南制订手册均有提及[13]，以下结合一些实践体会浅谈下检索数据库、检索词和检索式方面的应用思路供参考。

1.3.1 检索数据库

系统检索顺序可借鉴“6S 证据资源金字塔”自上而下逐级检索，从临床实践指南、系统评价、Meta 分析、临床随机对照试验、队列研究、回顾性研究、观察性研究、质性研究到专家意见、案例分析、个案报道等。此外，新西兰临床实践指南组指出还应依据不同类型的临床问题重点检索不同的研究设计文献。例如，治疗问题需要重点查找临床随机对照试验，诊断性问题需查横断面研究，预后研究要找队列研究等等。

临床指南和系统评价的检索数据库除了上文提及的，还可检索 BMJ Best Practice、Uptodate、Dynamed、ClinicalKey 等循证汇编数据库。这些数据库一般以具体的疾病、药物、诊疗措施等为主题，包含了临床试验、系统评价和临床指南等不同类型的循证医学证据且定期更新，不仅有参考文献出处，而且有 GRADE 质量评级及推荐意见。

临床随机对照试验、队列研究等其他类型的文献检索资源选择范围很广，需要特别注意各数据库收录的学科范围、地理区域和收录年限等特点，联合多个数据库检索可以减少漏检风险。WHO 和 SIGN 指南制订手册推荐 Medline 和 Embase 两个数据库作为必选数据库。SIGN 指南手册列出的数据库还有：针对 RCT 的 Cochrane 临床对照试验中心注册数据库（Cochrane central register of controlled trials, CENTRAL）和 WHO 临床试验注册平台（WHO international clinical trials registry platform），针对护理学和心理学等专业的 CINAHL、ERIC、PsycINFO 等。如制订中国指南，还需补充国内相关证据，需要检索中国知网、万方、维普等数据库。

1.3.2 检索词和检索式

指南制订会有好几个 PICO 构建的临床问题，每一个 PICO 问题往往由许多医学术语构成，同一个术语要尽可能多得考虑到同义词或近义词。Pubmed、Embase、PsycINFO、CINAHL、SINOMED 等数据库检索时应充分利用主题检索功能，浏览主题词注释信息和树形结构，了解上下位主题词和款目词，还可进行主题词加权以及扩展，选择合适的副主题词，使检索结果更符合检索需求。自由词检索途径通常选择在“Title/Abstract”或“Title/Abstract/Keywords”字段中进行限定检索。鉴于英文单词的单复数和变体形式，自由词检索可积极使用截词符、邻近运算符和布尔逻辑运算符。对于 PICO 的每一项常常采用主题词和自由词联合检索的方式。当然，一些临床问题并不包含所有 PICO 要素，比如没有对照组或结局变量不限定。当检索的文献数量过少，可以适当减少其中某一个 PICO 要素，扩大检索范围。总之，检索是一个随机应变、不断调整完善的动态过程。除了撰写检索式，追溯法也能帮助指南制订者检索文献。利用已经掌握的文献末尾所列的参考文献，

进行逐一追溯查找引文，还能从查到的引文中再追溯引文，如滚雪球一般扩大了检索文献数量。这种方法在刚开始检索所得文献量过少或者对检索词拿捏不准时可提供许多线索。

检索文献还需注重文献的时效性。医学领域文献的半衰期一般为 5 年[14]，引用过时的系统评价或证据文献可能会导致错误的判断和推荐意见。另外，由于大部分数据库更新速度很快，同一个检索式没过几天再次检索又会新增无数文献，因此除了每次检索需要记录检索行为发生的时间，还需要借助 Endnote、Note express 等文献管理软件对检索到的文献进行整理、合并和去重，方便后续根据制订的证据纳入和排除标准等进行筛选。

1.4 指南更新与修订阶段

各临床指南制订手册都指出应适时对临床指南进行评估和更新，才能保证其时效性以及在临床指导中的权威性。一般而言，指南更新的周期约为 2~5 年，当干预措施的利弊平衡发生改变，有争议的领域出现了新证据时则需要考虑更新[15]。指南更新和修订时，可以使用之前的检索策略系统检索并更新检索日期，其前提是检索词和检索式没有新的补充或者调整。如果遇到新药、新疗法或新发不良事件则需要重新选择检索词，制订检索式。另外当原指南 PICO 等关键元素的专业定义出现了更新、调整或者范围变化，例如疾病定义或诊断标准改变，也应及时重新斟酌检索词，不宜使用旧的检索式进行检索。

2 文献检索在指南制订中的实践思考

2.1 专业检索团队的参与

指南制订需要检索大量网站和数据库，而各数字资源的检索系统存在着很大的差异。检索时不仅要选择与检索目的对应的资源类型，还需要掌握不同的检索方法和规则。面对几十个网站和数据库的检索，成千上百个检索词和检索式的撰写，如果没有专业检索经验，很难适应纷繁复杂的检索任务。其次，检索是一个庞大的系统工程，尽管指南制订手册并未明确文献检索至少需要几个人完成，但参考《Cochrane 系统评价手册》对检索人员至少 3 人的要求，即 2 人独立检索遇到分歧请第 3 人判断，指南的检索应由多位指南研究领域专家、文献信息专家、专业检索员共同构成的检索团队来完成，这对提高文献检索的效率和检索质量至关重要。

2.2 自动化检索系统的发展

鉴于指南文献检索流程复杂且检索范围广，一些研究者尝试研发自动化工具加速检索过程，如同时整合多个数据库（Cochrane、JBI、DARE 等）并能使用多种语言检索的 Epistemoniko（https://www.epistemonikos.org）可以一站式检索；文献筛选工具 DistillerSR、EPPI-Reviewer 等能基于研究设计类型、标题和摘要进行筛选，有助于海量文献的管理[16]；利用机器学习、数据挖掘、文本分类等技术将大大提高检索效率，减轻检索负担[17]。在 2023 年“人工智能驱动的科学研究论坛”上发布了基于大语言模型+向量数据库的文献知识库“Science Navigator”，该数据库的性能优势体现在多模态、多模型、查询快、迭代快等特点，能更专业地理解检索目的，并自动完善检索输入[18]。虽然指南制订的自动化、智能化工具和平台已有较多的研究开展，但在实际应用中仍然存在信息源不稳定、算法黑箱、信息遗漏或数据丢失等情况，未来大数据与人工智能在辅助指南文献检索上仍需不断发展和优化。

2.3 文献检索质量标准的制订

为了处理不同检索方式之间的差异，增加文献检索的透明度和可重复性，会对检索过程进行评价，其中查全率和查准率是两个传统指标。查全率（recall ratio）计算公式的分母是所有应该检索到的文献；查准率（precision ratio）是指所检索到的文献中，被判断为相关文献的比值，与查全率成反比关系。而实际检索中，应有的总数情况往往无法预测，很难准确得到数值。有国内学者针对循证社会科学领域原始研究文献检索质量的评价，引入了信度评价指标[19]，探索了不同检索途径的复本信度与重测信度。进一步还提出了饱和度和冗余度等评价指标，能够更好衡量文献检索的质量和效率[20]。目前尚缺乏对指南文献检索的科学性评价标准，未来能否用信度、饱和度、冗余度、敏感度等标准来指导文献检索流程，使文献检索更加科学规范并且高效可靠，仍需要更多研究。

3 总结

从指南的筹划注册、计划书准备到正式撰写、更新修订，各个环节都需要进行文献检索，不同阶段对文献检索的要求不尽相同。指南制订中文献检索规模庞大且繁杂细琐，明确检索目的和掌握检索技巧至关重要，专业的检索团队也有助于高质量完成任务。未来，自动化、智能化等高科技检索工具和技术平台开发，以及文献检索质量标准的研究或将大大提升检索效率，推动指南文献检索更便捷、更规范。

PU-097

基于生物信息学分析方法探索霍奇金淋巴瘤一线治疗药物的优势与不足

张昊军
山西医科大学基础医学院

目的 本研究旨在利用生物信息学分析方法，深入探讨霍奇金淋巴瘤（Hodgkin's Lymphoma，HL）一线治疗药物的优势与不足，为 HL 的临床治疗提供全新的视角和可能性。
方法 我们首先综合检索了各个国家的 HL 临床诊疗指南，对 HL 的临床治疗药物进行了系统整理。随后，通过结合 DrugBank 数据库和 CTD 数据库，建立了药物与靶点之间的详尽映射关系。同时，我们运用 Biopython 包，深入检索 PubMed 数据库中的 HL 相关文献，筛选并提取摘要信息，以获取与 HL 相关的基因。接着，我们构建了药物-基因网络，并对其进行深入的关联分析。最后，我们运用基因本体学（GO）和京都基因与基因组百科全书（KEGG）分别对药物靶点和疾病基因进行功能富集分析，以揭示更多治疗机会和潜在路径。
结果 最终共得到 20 个 HL 的常规治疗药物，包括依托泊苷、甲泼尼龙、阿糖胞苷、多柔比星等，共得到 18 个药物的作用靶点与 93 条药物-基因关系对，绝大多数药物得作用靶点或者作用机制都与 MAPK1，CD40，NFKB1A 基因相关。文本挖掘共得到 338 个疾病相关基因。药物靶点与疾病相关基因取交集发现共有 8 个基因，包括 MAPK1，NFKB1A 等，且 MAPK1，NFKB1A 在网络分析的四个衡量指标上皆排名靠前，但仍有 330 个疾病基因没有参与药物作用途径。富集分析发现，NF-kappa B 信号通路和免疫应答等在药物靶点富集结果和疾病富集结果中出现，但仍有部分富集结果为疾病所独有，如 FoxO 信号通路等。
结论 目前 HL 的常规治疗方法是较为准确的，但仍有可完善部分，可针对其余疾病相关基因和信号通路进行药物筛选。

PU-098

基于知识发现工具 Bitola 探索自闭症关联药理物质

王康龙 钟小云 吕艳华*
山西医科大学

目的 通过知识发现工具 Bitola 探索自闭症的关联药理性物质，文章以黄满丽教授团队对络丝蛋白的研究成果为依据，选择络丝蛋白作为中间概念进行知识发现检索，提高研究的科学性和效率。
方法 基于非相关文献的知识发现开放式挖掘和闭合式验证两个过程。首先，我们选择了 Autistic Disorder（自闭症）作为起始概念 X，使用 Bitola 系统进行开放式知识检索，以找到与 ASD 相关的潜在候选蛋白质、基因。然后，我们将中间概念 Y 的语义类型设置为“Amino Acid, Peptide, or Protein”来筛选出最有价值的 Y 概念。接下来，我们使用 Pharmacologic Substance（药理物质）的语义类型，找到与 X 和 Y 之间存在新关系的药理物质 Z。通过这一过程，我们初步筛选出潜在的治疗 ASD 的药理物质。最后使用文献数据库对初筛后保留的药理物质进一步分析研究，发现 Recombinant Fibroblast Growth Factor-2（重组成纤维细胞生长因子 2）可能与自闭症相关。
结果 Bitola 系统的开放式挖掘阶段产生了多个与 ASD 相关的候选蛋白质和基因，其中包括了 FGF2。在进一步的闭合式验证和筛选过程中，我们发现 FGF2 与 ASD 之间存在潜在的联系。通过文献回顾和数据验证，我们进一步确认了 FGF2 在 ASD 研究中的重要性。
结论 实践证明，应用非相关文献知识发现方法挖掘出了治疗自闭症的潜在生物分子 FGF2。通过文献回顾和数据验证的方式，最终确认了这一关联，为自闭症治疗提供了新的方向。这也进一步突显了非相关文献知识发现方法在疾病研究中的潜力和重要性。

PU-099

关于清宫痹症医案用药规律的研究

吕艳华 李琦* 梁傲雪 赵宏霞
山西医科大学管理学院

1 引言

痹症泛指因风寒湿等邪侵袭或凝滞肌肤筋骨，经络痹阻而引起肢节疼痛、麻木、屈伸不利等为特征的一类疾病[1]。这是一种临床上的常见病及多发病,痹症的发病机制复杂且病因不太明确,因此治疗痹症目前无有效的治疗方法,西药治疗也不彻底,而且副作用较大,具有药物依赖性,严重损害了肝、肾等脏腑功能[2]。当前关于痹症治疗的研究主要有三类，一是外治法的探究，如吴中朝[3]进行了针灸疗法的研究；二是名家治疗经验的分析，如郭会卿等[4]总结了娄多峰教授治疗痹病的学术思想；三是对某方药效用的研究，如贺倩等[5]关于沙冬青治疗类风湿关节炎(痹症寒湿痹阻证)的临床研究。

与上述几类研究不同，清宫痹症医案拥有完整时期的治疗记录，包含多位御医的施诊档案，数量众多，用药手段丰富，构成了独特的宫廷医学特色。研究既可以挖掘其中核心药物及常用方剂，还能够挖掘出其中的关联药对及药组，由此把握出清宫医案治疗痹症的诊疗方向，为现代医学临床治疗痹症提供思路。

2 清宫痹症医案数据与处理

2.1 数据来源

研究数据来源于 2009 年北京科学出版社出版的陈可冀等主编的《清宫医案集成》[6]。

2.2 纳入标准与排除标准

纳入标准根据国家中医药管理局、国家卫生健康委员会印发的《中医临床诊疗术语第 1 部分：疾病》（国中医药医政发〔2020〕3 号）里关于痹症的规范描述来对清宫医案进行人工筛选，纳入范围以方剂、代茶饮、药丸为主，选择以痹症为主症的医案。排除记录不完整的医案（一例医案起码包含症状、方剂名/治法、药物组成、药引、患者、医家六大要素）。筛选出的痹症医案由中医专家审核后，最终得到 347 例清宫痹症医案。

2.3 药物数据规范化

一方面是药名规范化。以国家药监局、国家卫生健康委发布的 2020 年版《中华人民共和国药典》[7]为标准，对痹症医案所涉及的 203 种药物完成规范化。遇到《中国药典》没有记录的药物，则依次从《全国中草药汇编》[8]、《中药大辞典》[9]、《中华本草》[10]次序进行查询补充。像六一散、益元散等中成药保留原名，焦三仙等惯例缩写不进行拆分，保留原名。除生地黄、熟地黄、生甘草、炙甘草、生黄芪、炙黄芪、煅石膏、生石膏外，去除炮制规格。另一方面是药物类型及功效规范化。规范顺序依照药名规范顺序，首先参考《中国药典》。

2.4 构建清宫痹症医案数据集

通过设定的纳入标准和排除标准，筛选出合格的痹症医案，构建出清宫痹症医案数据集。根据症状、方剂、药物组成、药引、患者、医家这六要素，将医案信息转换成规范化数据，来建立数据集，以便接下来的数据挖掘分析。

3 清宫痹症医案药物挖掘研究

3.1 清宫痹症医案常用药物分析

将方剂数据输入至 Excel 表格，统计各药物的使用频数，203 种药物共计使用 3902 次，其中排名前十位的药物如表 1 所示，由 IBM SPSS Modeler 18.0 可视化得到药物使用情况。

前十位药物里，有一半具有止痛的效用，当归具有补血调经、活血止痛的效用，主要对应“肝郁血虚，不能荣经”、“气血两亏，不能荣养筋脉”等血虚类证候；白芍具有养血敛阴、柔肝止痛、平抑肝阳的效用，主要运用于“肝脾湿热，荣血不能下注”、“肝郁血虚，湿热下注”等涉及肝脏功能的证候；生甘草具有补脾益气、缓急止痛的效用，并且还有着清热解毒、调和诸药的重要功效，被南朝医学家陶弘景尊为“国老”[11]，故而生甘草在痹症方剂是为常用药物；川芎具备活血行气、祛风止痛的效用，主要针对“血虚有热，外受风凉”、“风湿未净”等受风邪影响的证候；

香附具有疏肝解郁、调经止痛、理气调中的功效，主应“肝郁血虚，不能荣经”、“气郁停饮”等相关证候。由此可见，缓解痹痛是治疗痹症的重点。

其余的五位药则是具备治疗痹症其它方面的功效。利水消肿的茯苓主要使用于消除“湿热下注”、“湿滞过盛”等证候的除湿方剂中；清热凉血的生地黄则在“养阴和肝汤”、“调荣润燥汤”等方剂中发挥主要作用；理气的陈皮有着“大橘皮汤”的方剂专门对于“气不化湿”等相关证候行使功效；而补气的白术对“气血未充”、“元气未复”、“荣卫不和”之类的证候发挥着健脾益气的功效，而“湿饮下注，脾胃虚弱”、“肝胃气道不调，化湿仍慢”等证候描述里，可以看出白术的功效最终目的也是为了除湿；清热燥湿的黄芩主治“湿热过盛”、“湿热下注”等证候。同时，白术和黄芩还具有安胎功效，成为常用药物或也与清宫医案内患者多为女性有关。从茯苓、陈皮、白术、黄芩可以分析得出，除湿是治疗痹症的另一重点。

表 1 清宫痹症医案前十位常用药物

序号	药名	频数	频率
1	当归	185	4.74%
2	茯苓	151	3.87%
3	白芍	126	3.23%
4	生甘草	109	2.79%
5	生地黄	106	2.72%
6	白术	106	2.72%
7	黄芩	85	2.18%
8	川芎	84	2.15%
9	香附	83	2.13%
10	陈皮	81	2.08%

3.2 清宫痹症医案常用药物关联规则分析

将药物数据整理为矩阵模式，导入 IBM SPSS Modeler 18.0 进行关联规则分析，以支持度和置信度作为依据，得到清宫痹症医案常用药对。其中，排名前三的二项关联药对及三项关联药组（根据置信度进行排名）（表 2、表 3）。

3.2.1 二项关联药对分析

清宫痹症医案共有 30 个方剂使用了党参-白术药对，而 30 个不同方剂具有共同的益气之效。这与党参和白术同为补气药不无关系，《素问·痹论》有载:“风寒湿三气杂至, 合而为痹也。其风气胜者为行痹, 寒气胜者为痛痹, 湿气胜者为著痹也。”[12]张双等论证乌头汤益气温阳蠲寒法也说明了痹症的发生多为人体正气不足[13]。故而益气可视为治疗痹症的一大手段，清宫痹症医案常以党参-白术达到益气目的。

当归-茯神药对则被 18 个方剂使用，主要方剂为保元养阴汤、益气养阴汤、保元益阴汤、益气养荣汤，综合所有方剂来看，可以归纳主治证候为“气血未充”、“心脾尚弱”、“元气未复”、“荣卫不和”等，与补血的当归和养心安神的茯神功效对应，除了再次强调了益气的重要性，还与臧堃堂教授在治疗痹症的“气血亏虚, 诸痹不已, 益气和营”经验符合[14]。气血并行，重视补血畅通也是治疗痹症的要点。清宫痹症医案则以当归-茯神作为常用药对调补气血。

泽泻、茯苓皆为利水消肿药，共同具有利水消肿及渗湿的功效，该药对被使用的方剂有养阴和肝汤、除湿拈痛汤、当归拈痛汤、除湿化滞汤，相关的证候为“湿滞受凉”、“湿热受风”等，验证了除湿的治疗重要性，也看出清宫痹症医案用于除湿的药对为泽泻-茯苓。

表 2 清宫痹症医案二项关联药对

后项	前项	支持度百分比	置信度百分比
白术	党参	17.87	95.16
当归	茯神	19.02	90.91
茯苓	泽泻	12.68	90.91

3.2.2 三项关联药组分析

茯神-生地黄-当归为二项关联药对茯神-当归的进一步延申，从包含此药对的 18 个方剂内缩小至三个方剂，分别为益气养营汤、清燥和胃饮、益气健脾汤。梁艳冰通过《临证指南医案》中地黄应用研究，挖掘得到茯神、当归皆为生地黄配伍的常用药物[15]。生地黄凉血补血，既与补血活血的

当归和治心虚血少的茯神相和，共同补养血气，还能够调和性温的当归。

同样的，党参-白芍-白术三项关联组是党参-白术药对的展开，共有 9 个方剂使用了该三项关联药组。白芍作为补血之品，与党参、白术同归类为补虚药，在益气的重心上添加了补血的功能。

党参-当归-白术三项关联组共有八个方剂，分别为益气养心汤、益气养荣汤、燥湿养荣汤、益气养营汤、益气理脾汤、益气养阴汤、益阴通络汤、益气健脾汤，根据方剂名可以看出，“益气”是主要治疗方式，还兼有“养心”、“养荣”、“养营”、“理脾/健脾”、“养阴”、“通络”的功效。党参、当归、白术同是补虚药，党参补脾肺气、补血、生津，当归补血调经、活血止痛，白术健脾益气、燥湿利尿，都具备完成方剂功效的性质。并且，这八个方剂内，益气健脾汤、益气理脾汤、益气养荣汤、益气养心汤、益阴通络汤、燥湿养荣汤包含了茯神-当归、党参-白术、党参-白芍-白术、党参-当归-白术；益气养阴汤包含了党参-白术、党参-白芍-白术、党参-当归-白术；益气养营汤包含了茯神-当归、党参-白术、党参-白芍-白术、党参-当归-白术、茯神-生地黄-当归。说明了从关联规则来看，清宫痹症医案主要围绕“益气”、“补血”、“补虚”进行治疗。

表 3 清宫痹症医案三项关联药组

后项	前项	支持度百分比	置信度百分比
当归	茯神 and 生地黄	13.54	100.00
白术	党参 and 白芍	12.97	97.78
白术	党参 and 当归	12.97	97.78

3.3 清宫痹症医案核心药物分析

常用药物并不能等同于核心药物。为了使研究更具可信度，在定义核心药物时，引入中介中心性指标作为参考依据，结合频数统计，两相综合来获取清宫痹症医案的核心药物。

中介中心性 （Between Centrality）源于分析社会网络中个体的重要性,以经过某个节点的最短路径数目来刻画节点的重要性指标[16]。研究通过 Gephi 软件完成中介中心性计算，得到了前十位药物（表 4），并由 IBM SPSS Modeler 18.0 得到了可视化的药物使用情况（图 1）。

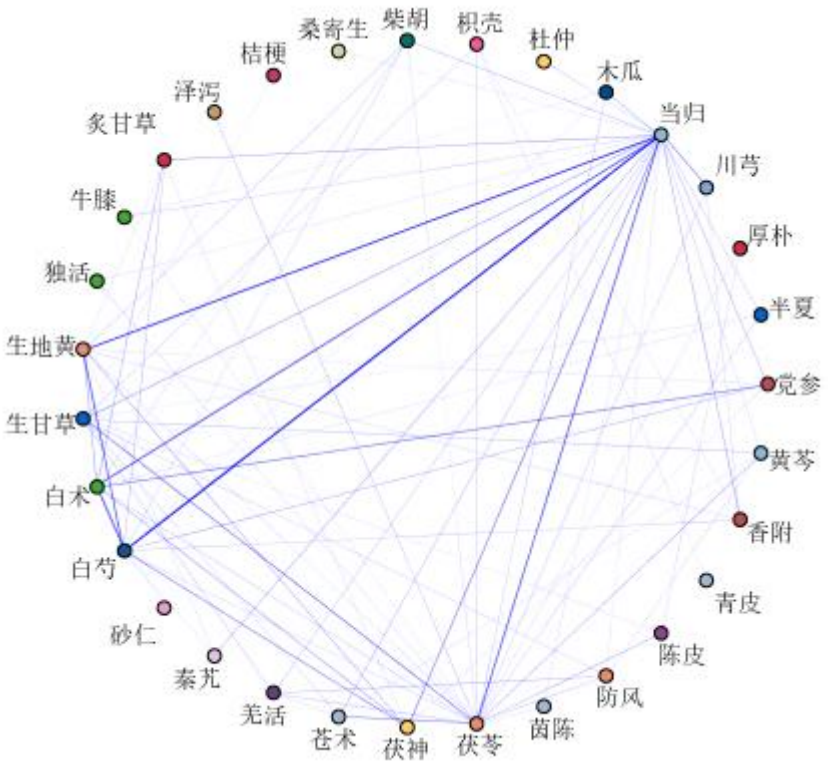
综合表 1 的排名，可以得出清宫痹症医案核心药物为当归、茯苓、川芎、白芍、白术、生甘草、香附、黄芩、陈皮。这九位核心药物中，补虚药占了四位（当归、白芍、白术、生甘草），理气药有两位（香附、陈皮），川芎为活血化瘀药，茯苓为利水渗湿药，黄芩为清热药。

木瓜从频次第 12 位跃至中介中心性排名第 6 位，证明它在该指标中具有一定地位。木瓜作为祛风湿药，具有舒筋活络、和胃化湿的功效，于痹症有着直接的对症功效。《中国药典》对其主治也有了说明：“用于湿痹拘挛，腰膝关节酸重疼痛，暑湿吐泻，转筋挛痛，脚气水肿。”木瓜对于痹症治疗经络痉挛疼痛具有作用。

表 4 中介中心性前十位药物

序号	药名	betweenesscentrality
1	当归	1053.65
2	茯苓	915.68
3	川芎	779.03
4	白芍	747.71
5	白术	630.45
6	木瓜	492.64
7	生甘草	481.91
8	香附	465.75
9	黄芩	465.25
10	陈皮	444.94

图 1 药物使用情况



4.结论

本文在综合分析了当前中医药领域内的数据挖掘进展后，结合多种数据挖掘方法，挖掘了清宫痹症医案的用药规律。通过常用药物分析和关联规则分析，得出清宫痹症医案药物主要采用补虚类药物，诊治方向主要为止痛、益气、补血、除湿四个方面，并利用频数统计和中介中心性两方面指标深层次地挖掘出九位清宫痹症医案核心药物，为现代医学临床治疗痹症提供了思路。

PU-100

近十年全球老年多重慢病研究趋势及热点演化分析

王洁 冯佳* 余丹 刘永恒 赵伟栋
吉林大学公共卫生学院医学信息学系 长春 130021

1.引言

随着全球人口老龄化进程的加快，老年人群的健康状况越来越受到医疗卫生机构及大众群体的重视。美国疾控中心的一项研究表明，老年人中患有慢性疾病的约有 80%，其中患有多种慢病的老年人占一半以上[1]。意大利的一项调查显示，老年人多重慢病患病率为 39.9%[2]。我国学者研究发现，中国大陆居民慢性病共病率为 36.3%，而 65 岁以上人群慢性病共病率则高达 66.1%[3]。目前，多种慢病共存已经成为老年慢病患者的常态化特征，为其个人及社会都带来了一定的负担。

2008 年，世界卫生组织对“Multimorbidity”做出了规范的界定：指同时患有两种及两种以上的慢性疾病[4]。近年来，国内外学者对老年多重慢病的相关研究在广度和深度上都有了进一步的扩展，研究成果日渐丰富，但对老年多重慢病领域的研究现状总结不足。2021 年，ZHOU 等人对 Web of Science 和 CNKI 数据库中老年人多重病症相关文献进行了文献计量分析[5]，2022 年，吕晓燕等利用 PubMed 和 Web of Science 数据库文献对全球范围内共病的研究现状和热点进行了分析[6]。郑晓等人[7]及卫安乐等人[8]利用 CiteSpace 工具分别对国内、国内外的老年多重慢病领域研究热点进行了可视化分析。现有研究多为利用文献计量法阐述了老年多重慢病领域的研究现状，但其并未对该领域的研究趋势及热点演变进行深入的分析与探讨。本研究全面收集了老年多重慢病相

关文献，利用文本挖掘技术，采用文献计量、聚类分析、三维战略坐标分析、热点演变桑基图等多种方法，深入探讨老年多重慢病领域的研究进展、研究趋势、研究热点演化过程等，充分发挥现有研究成果价值，为相关疾病的健康管理提供依据，同时也为该领域的深入研究与发展提供参考。

2.数据来源与方法

2.1 数据来源及预处理

老年多重慢病研究涉及了多学科多领域的交叉融合，由此出发，本研究选取了综合领域、医学领域、心理学领域相关的 Web of science 核心合集、Scopus 数据库、PubMed 数据库、APA-PsycINFO 美国心理学会数据库作为数据来源以提高检全率。以 elderly、multimorbidity、Multiple chronic diseases 等词及其同义词制订检索式（具体检索式见附件）。检索时间限制为 2010-2021 年。共检出 7996 篇文献，对文献进行去重及文献筛选，最终剩余 5237 篇文献作为本研究的数据。具体文献处理记录如图 1 所示。

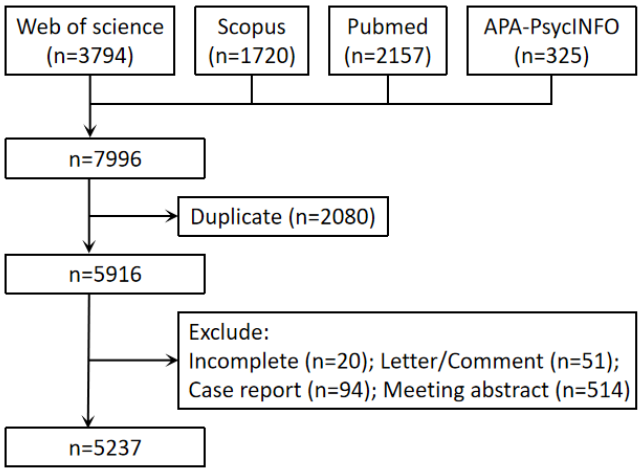


图 1 老年多重慢病相关文献来源及文献处理记录

2.2 研究方法

2.2.1 文献计量分析

使用 Excel2010 进行年发文量统计，使用 PASW18.0 对发文量趋势进行回归分析线性预测。使用 R4.2.2 软件的 Bibliometrix 包对文献来源期刊进行分析，并根据布拉德福定律[9]确定核心期刊。通过 CiteSpace6.2.2 对对文献作者、发文机构进行分析及可视化，并根据平方根法确定高产作者。使用 Bicomb2.04 对文献关键词进行提取，并生成词篇矩阵及共现矩阵。在关键词提取过程中，根据补充后的百度停用词表对关键词进行去停用词处理[10]，后合并关键词，合并原则为：含义相同或相近的词，如 old、aged；在 MeSH 词表有同一个入口词，如 Amentias、Dementia。

2.2.2 研究热点分析

根据普赖斯公式确定高频词阈值，使用 PASW18.0 对高频关键词进行聚类分析，探究老年多重慢病领域的研究热点。

2.2.3 主题演化分析

（1）战略坐标分析：x 轴代表向心度，主要用类团间的共词强度表示，y 轴代表密度，用来衡量主题的内在强度，密度的计算公式为：

$$d = \frac{\sum_m \theta_k^{(m)}}{M}$$

M 表示文档的总量，$\theta_k^{(m)}$ 表示主题 k 在文档 m 中的权重。根据各类包含关键词的时间分布确定主题新颖性，并对主题进行命名，绘制三维战略坐标图[11]，揭示热点主题的外部特征演化。

（2）桑基图：将文献数据根据时间划分成四个部分，分别提取每个时间段的高频关键词，计算相邻时间段的关键词相似性，绘制热点演变的桑基图，揭示热点主题的内容特征演化。

3.结果

3.1 文献计量

3.1.1 年发文量分布

自 2010 年至 2021 年，该领域每年的发文量整体呈上升趋势（图 2a），文献平均年增长率 15.84%。对数据进行曲线估计，其中指数模型的 r^2最大(r^2=0.979)，能够很好解释曲线的增长情况（图 2b）。老年多重慢病领域文献发文量呈指数增长，说明该领域正处于快速发展时期。

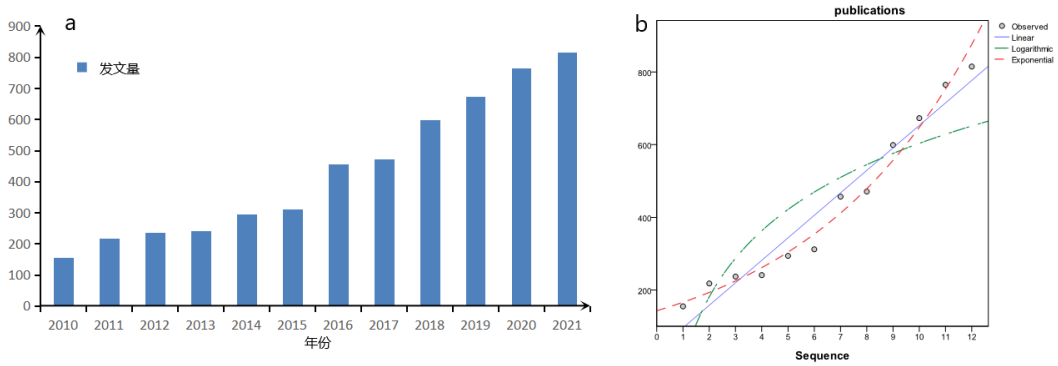


图 2 老年多重慢病领域年发文量随时间的分布趋势及曲线估计

3.1.2 来源期刊

由 Bibliometrix 统计结果，文献来源期刊共 1932 个。根据布拉德福定律计算得出核心期刊 72 个，占期刊总量的 3.7%。部分核心期刊及发文情况如表 1 所示。

表 1 老年多重慢病领域发文量最多的十大期刊

序号	期刊名	文章数
1	J AM GERIATR SOC	123
2	BMJ OPEN	122
3	PLOS ONE	122
4	BMC GERIATR	90
5	BMC PUBLIC HEALTH	54
6	J GERONTOL A-BIOL	49
7	J GERIATR ONCOL	44
8	ARCH GERONTOL GERIAT	43
9	AGE AGEING	37
10	AGING CLIN EXP RES	37

3.1.3 作者和机构

根据平方根法计算得出老年多重慢病领域高产作者 132 人，高产作者发文总量 1350 篇，占全部文献的 25.8%。作者合作网络知识图谱如图 3a 所示。

绘制机构合作关系图谱（图 3b），发文量最高的机构是瑞典的 Karolinska Inst，发表相关文献 77 篇，其次是 Univ Calif San Francisco 发表 53 篇，Johns Hopkins Univ 发表 51 篇。

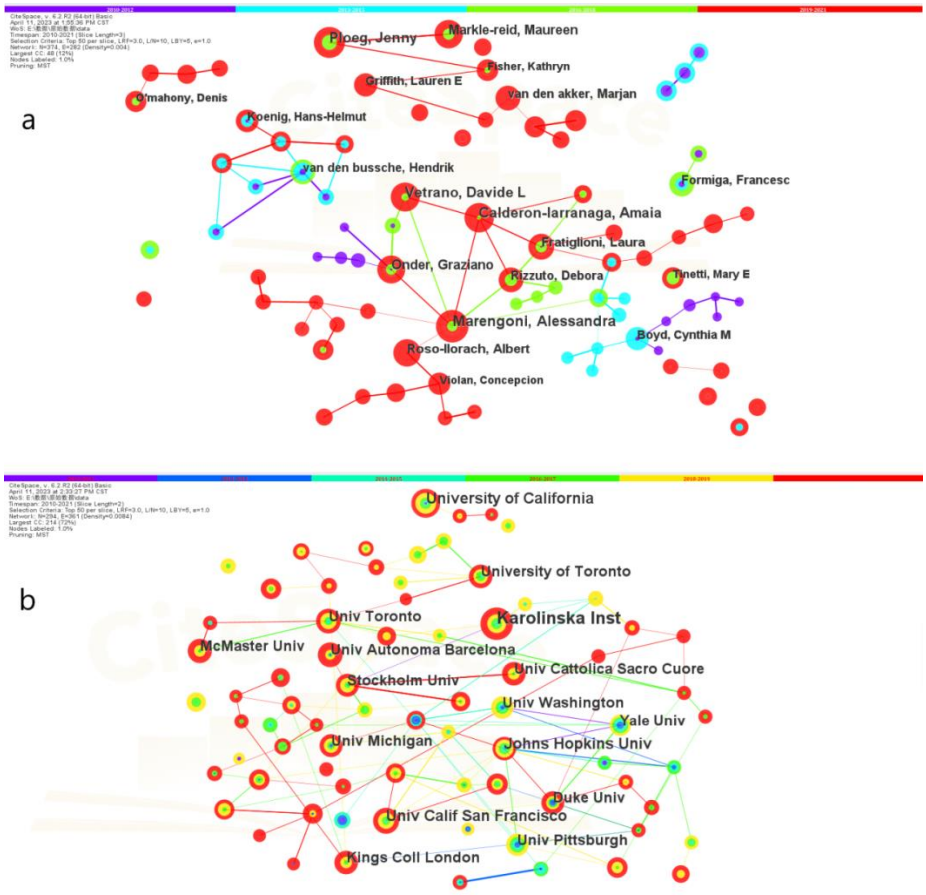


图 3 老年多重慢病领域高产作者和机构合作关系图谱

3.2 研究热点现状

关键词提取后首先对关键词进行去停用词处理，在百度停用词表的基础上结合关键词提取结果对停用词表进行补充，比如增加 analysis、sex、review 等停用词，并对关键词进行合并。根据普赖斯公式确定高频词阈值为 29，得到高频关键词 40 个，如表 2 所示。对纳入的高频关键词生成词篇矩阵和共现矩阵后进行聚类分析，可分成 8 类，如图 4 所示。

表 2 老年多重慢病高频关键词

序号	关键词	词频	序号	关键词	词频
1	comorbidity	1507	21	prognosis	58
2	aged	1398	22	risk factors	57
3	chronic disease	284	23	heart failure	52
4	polypharmacy	224	24	hypertension	51
5	mortality	217	25	hospitalization	49
6	frailty	189	26	cancer	43
7	depression	163	27	health-related quality of life	41

8	epidemiology	160	18	2	mental health	83	
9	primary care	160	19	2	cardiovascular disease	83	
10	quality of life	107	20	3	anxiety	73	
11	diabetes	87	21	3	self-management	73	
12	prevalence	84	22	3	obesity	72	
13	dementia	84	23	3	physical activity	63	
14	Charlson comorbidity index	83	24	3	chronic obstructive pulmonary disease	63	
15	survival	78	25	3	cognition	55	
16	disability	76	26	3	chemotherapy	54	
17	COVID-19	73	27	3	treatment	32	
18	geriatrics	73	28	3	surgery	32	
19	HIV	64	29	3	complications	31	
20	elderly patients	64	30	3	morbidity	30	

图 4 老年多重慢病领域热点主题聚类

3.3 热点主题演化

3.3.1 三维战略坐标图

以聚类结果为基础，结合时间维度绘制三维战略坐标图（图 5），并根据各类包含的关键词对各类进行命名。

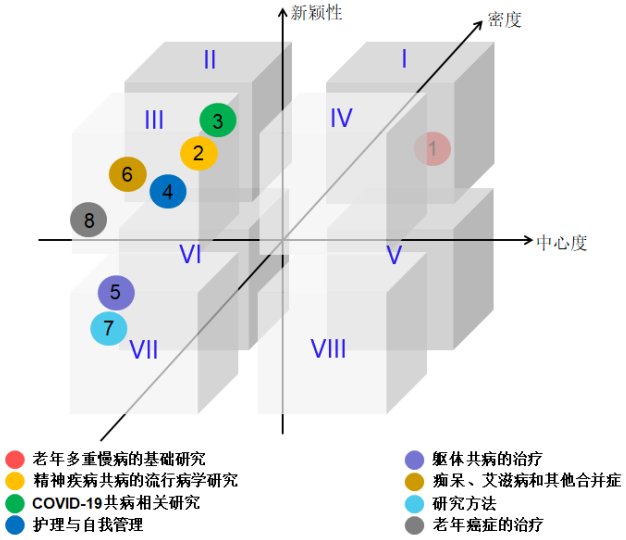


图 5 老年多重慢病领域热点主题分布三维战略坐标图

3.3.2 桑基图

将文献数据按时间划分成 4 个时间段，提取每个时间段的高频关键词，并计算相邻时间段关键词之间的相似性，绘制主题演变的桑基图如图 6 所示。

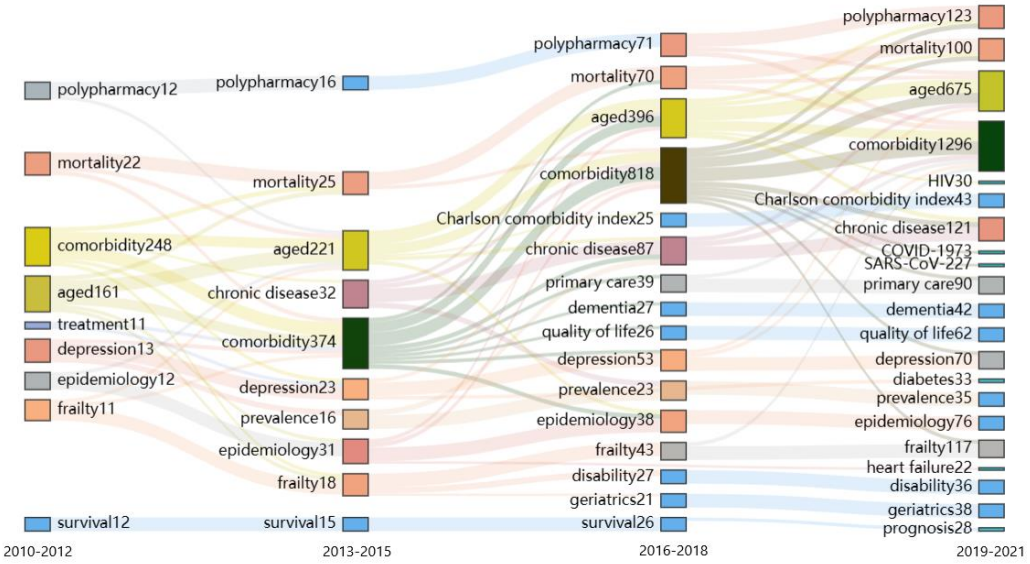


图 6 老年多重慢病领域热点主题演变桑基图

4. 讨论

4.1 文献计量分析

本文通过文本挖掘技术、文献计量学方法、三维战略坐标图等方法对对全球老年多重慢病相关的文献数据进行整理分析及可视化。我们发现老年多重慢病领域的研究正处于快速发展阶段，年发文量呈指数上升，我们预测该领域在未来几年将继续保持快速发展。对于核心期刊，J AM GERIATR SOC、BMJ OPEN、PLOS ONE 均发表一百篇以上文献，在所有的核心期刊中，英国

和美国的期刊较多。我们还发现老年多重慢病领域发文的学者们之间合作广泛，形成以 Marengoni Alessandra 等人为代表的多个合作团体，学者之间的广泛合作可以更好地促进学科的发展。高产出的机构普遍为各个高校，其中美国的大学占很大比例，且机构之间的合作网络相比与作者合作网络更加密集。

4.2 研究趋势分析

从高频关键词（表 2）可以看出目前老年多重慢病领域的研究热点，其所研究的疾病多为抑郁、痴呆等精神共病和糖尿病、心力衰竭、高血压等躯体共病。研究内容主要为多重慢病的流行病学研究[12]、护理保健[13]、自我管理[14]及治疗[15]等方面，且随着近年来新冠疫情的全球大爆发，新冠肺炎及其合并症相关研究[16]也快速增多。

从三维战略坐标图（图 5）结果来看，我们发现，老年多重慢病的基础性研究比如多重慢病的复方用药[17]研究已经相对成熟。抑郁症、痴呆等精神疾病共病相关研究成熟度和核心度欠佳但新颖性较高。有研究发现老年人的身体疾病负担沉重，会对心理健康产生不利影响，反之亦然[18]，抑郁症可导致其共病治疗的依从性降低，进而导致住院率、死亡率增加[19]。新冠肺炎相关研究是新兴的研究热点，近几年研究成果不断丰富。护理与自我管理相关研究的成熟度也有待提高。我们还发现了一些领域内相对边缘的热点主题，比如癌症与老年多重慢病患者的关系[20]，癌症仍然是人类尚未解决的医学难题，相关研究层出不穷，覆盖多个学科领域。综合第三象限来看，精神疾病共病、护理与自我管理等相关研究是目前老年多重慢病领域的研究热点，但其研究深度不足，尚未成熟，未来有待于进一步完善，可以成为未来研究的主要发展方向。我们还发现一些早年的研究热点主题，比如糖尿病、高血压等躯体共病的治疗以及老年多重慢病的相关研究方法。高血压和糖尿病是常见的合并症[21]，且糖尿病和高血压等慢性病在各国的中老年人群中患病率都很高[22-24]，对老年人身体健康及生活质量产生较大影响，被学者们广泛关注。研究方法主要涉及 Charlson comorbidity index，近年来也有学者在探索多重慢病评估的新方法[25]。

4.3 热点演变分析

从主题演变的桑基图中可以看出，随着研究的不断深入，老年多重慢病领域的研究成果逐渐丰富，热点主题逐渐增多。2010-2012 年主要是抑郁症及共病的基础研究[26]；2013 年开始流行病学相关研究[27]增多；2016-2018 年间热点主题显著增多，出现痴呆[28]、残疾[29]、老年多重慢病患者的生活质量及护理保健[30]等新的研究热点；近几年更多关注糖尿病[31]、心衰[32]等相关疾病合并症方面的研究。整体来看，抑郁症相关共病、老年多重慢病的流行病学、多重用药相关研究是一直以来持续关注的热点主题。

5 结论

老年多重慢病领域正处于发展阶段，未来一段时间内也将继续保持快速发展状态。结合热点主题的外部特征和内容特征我们发现，对于现有研究成果，老年多重慢病基础性研究已经相对成熟；抑郁症、痴呆等精神共病相关研究是学者们持续关注的热点，但成熟度有待提高，尚未形成完整的研究体系，未来可以在这些方面进行更加深入的研究，老年多重慢病患者的护理与自我管理是潜在的研究热点，未来研究可能更趋近于以人为本的健康促进。

PU-101
基于 Web of Science 可视化分析 ChatGPT 的应用研究热点

李娟，张仲男，韩斌孝，秦三利，丁颖，林静，张渊厚，石昕
甘肃省医学科学研究院 甘肃省医学情报研究所

chatGPT 是由 Open AI 公司开发的基于深度学习的自然语言处理模型[1-2]。自 2022 年 11 月底，chatGPT 一经推出，便在世界各地掀起一场人工智能的热潮，迅速成为人们讨论的热点，引发

市场极大的关注。chatGPT 在推出一个月末，其用户量已突破 1 亿，成为用户增长速度最快的应用程序。

chatGPT 作为一种语言处理工具，它能够生成自然流畅的语言文本，拥有强大的语言理

解能力，可以完成写脚本、文案、翻译、代码等多种任务[3-8]，是一个很有潜力的自然语言处理模型。chatGPT 强大的功能使其被广泛地应用于医疗、教育、信息、建筑、金融等各行各业。为了更好地了解 chatGPT 应用研究的现状和发展趋势，本文通过检索有关 chatGPT 的文献，绘制科学知识图谱，剖析 chatGPT 应用的研究热点和前沿，以期为国内的人工智能领域研究提供参考。

1 资料与方法

1.1 数据来源与检索策略

检索 Web of Science 核心合集数据库 2022 年 11 月~2023 年 9 月的关于 chatGPT 的相关文献。检索式为：主题[chatGPT] and 语种（English），文献类型主要为 article，共检索到 738 条有效记录。

1.2 分析方法

将纳入的文献以纯文本格式导出，导出的文献记录包含作者、研究机构、题名、关键词、摘要、发表年份、参考文献等信息。采用 VOSviewer 和 Citespace 软件对相关研究的发文国家、机构、作者、关键词、引用情况等进行可视化分析。最后，对结果进行解释（见图 1 ）。

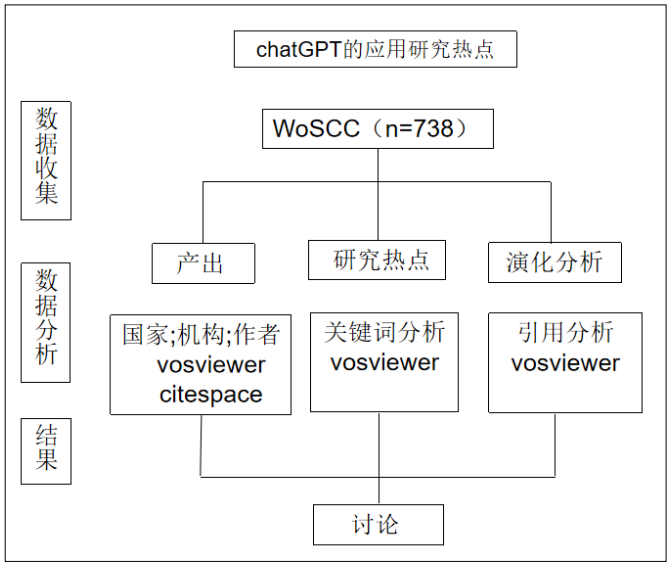


图 1

2 结果

2.1 发文国家分布

共 70 个国家发表了关于 chatGPT 的相关研究，美国发文量达 248 篇，位列第一，发文量占文献总数的 33.60%，遥遥领先于其他国家；中国继美国之后，位列第二，发文量为 83 篇，占总发文量的 11.25%；印度位列第三，发文量为 78 篇，占总发文量的 9.96%。具体情况见图 1、2。

图 2

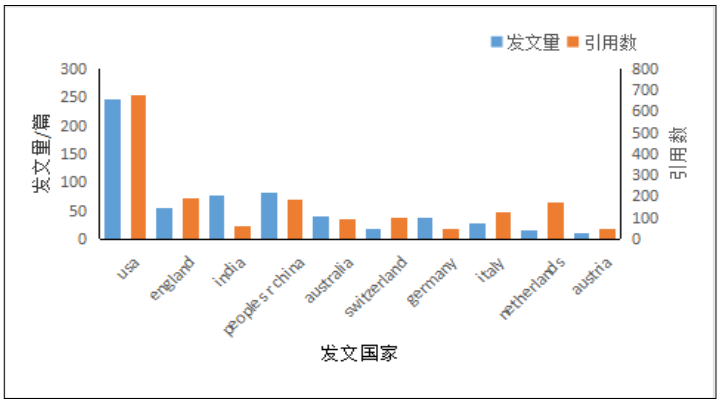


图 3

2.2 发文机构分布

经分析，共 1062 个机构从事 chatGPT 的相关研究，但发文量大于等于 10 的机构仅有 8 个。发文量排名前 3 的机构分别为印度昌迪加尔大学（Chandigarh University）、尼日利亚约瑟夫·阿约·巴巴洛拉大学（Joseph Ayo Babalola University）、中国科学研究院（chinese acad sci）。具体见图 3、4。

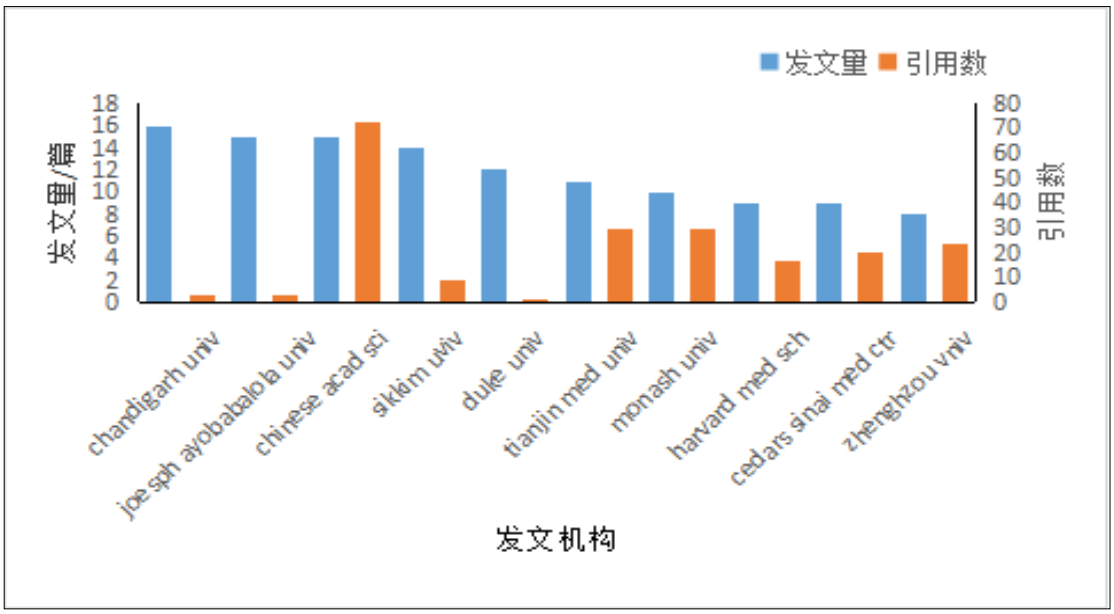


图 4

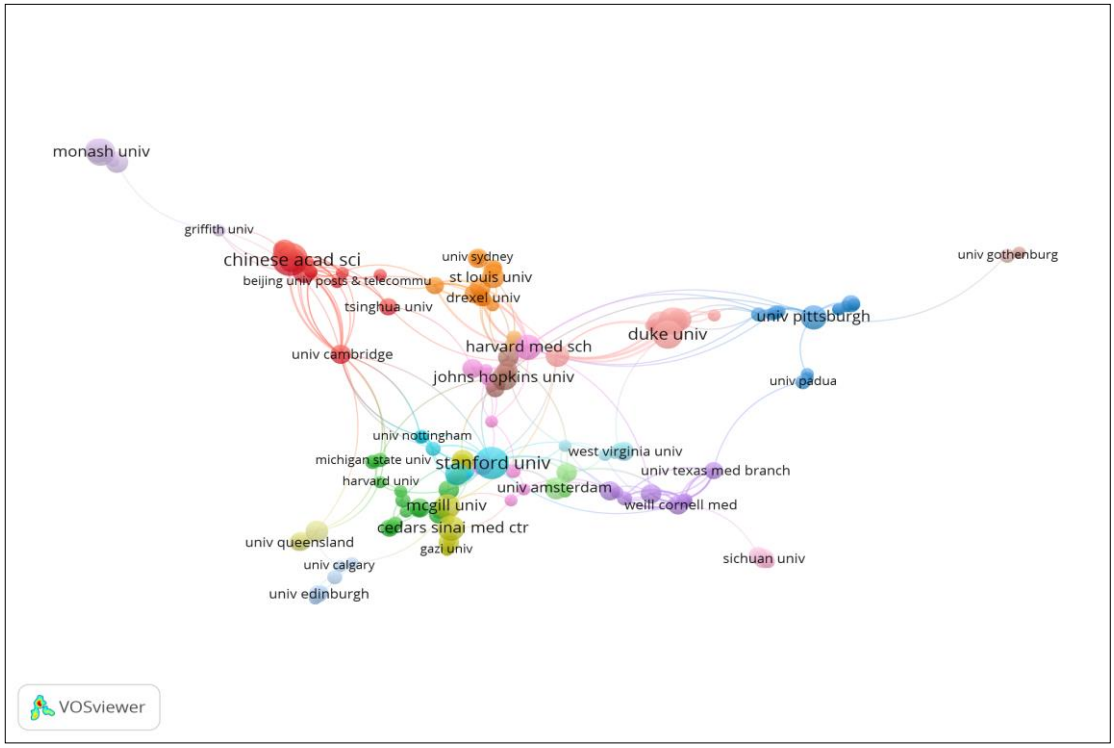
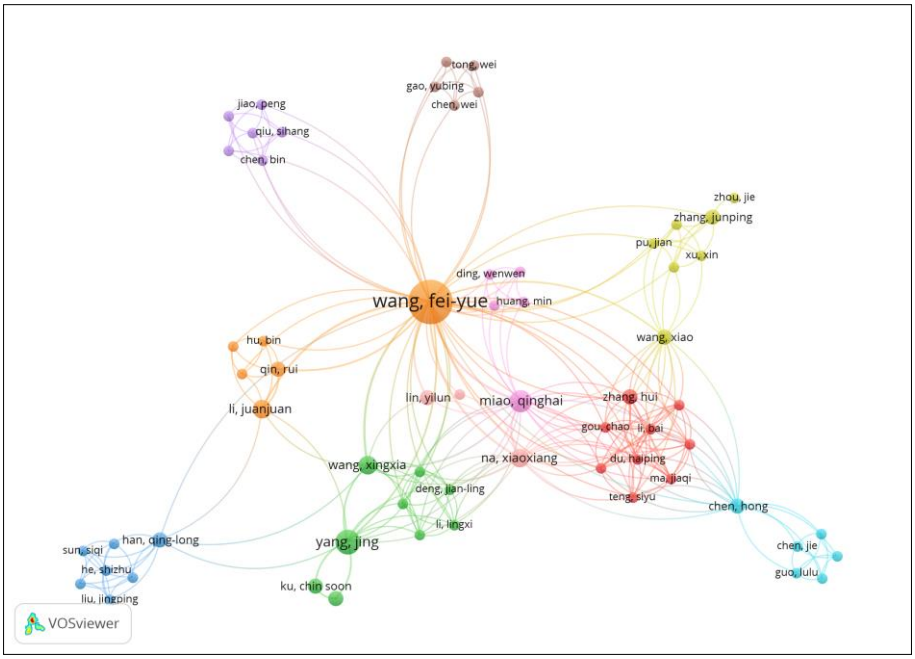


图 5

2.3 发文作者分布

共有 2115 位作者从事 chatGPT 的相关研究，但发文量大于 10 篇的作者仅有 6 位。发文量前 3 位的作者为 wiwanitkit, viroj（18 篇），wang, fei-yue（16 篇）及 kleebayoon, amnuay（15 篇）。具体见图 5、6。

图 6



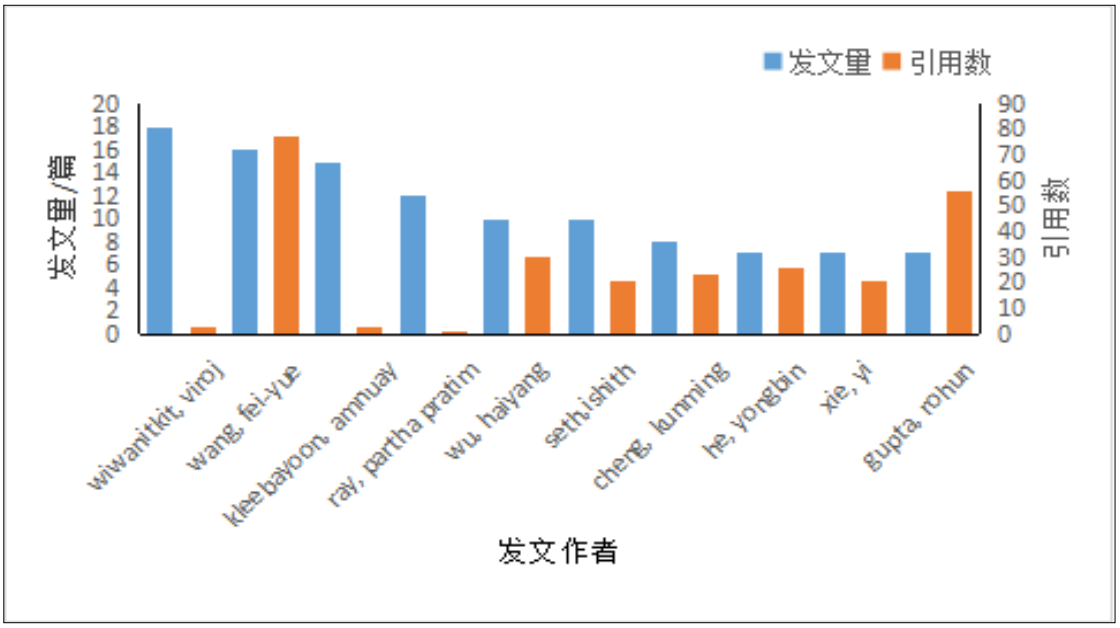


图 7

2.4 关键词可视化分析

关键词共现分析共发现 1318 个关键词，其中 16 个关键词共现次数大于 10 次；共现频次排名前 10 位的关键词见表 1。

表 1 chatGPT 领域前 10 位高频关键词

序号	关键词	频次
1	chatgpt	234
2	artificial intelligence	186
3	chatbot	37
4	large language models	32
5	machine learning	32
6	natural language processing	28
7	ethics	26
8	ai	24
9	medical education	19
10	deepling learning	14

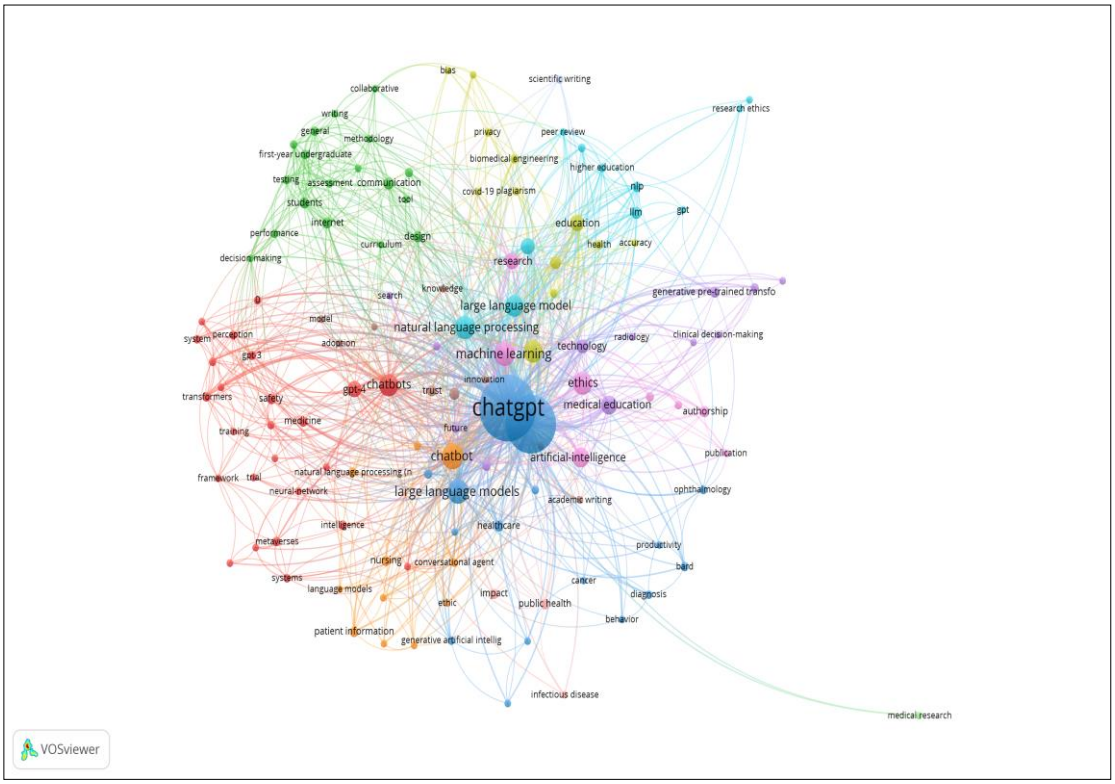


图 8

2.5 共被引文献分析

共发现 10882 篇文献被引用，其中 6 篇文献的被引次数超过 50 次。高被引论文的研究内容主要集中于探讨 chatGPT 的出现带来的效应以及 chatGPT 的实际应用。具体见图 9、图 10、表 2。

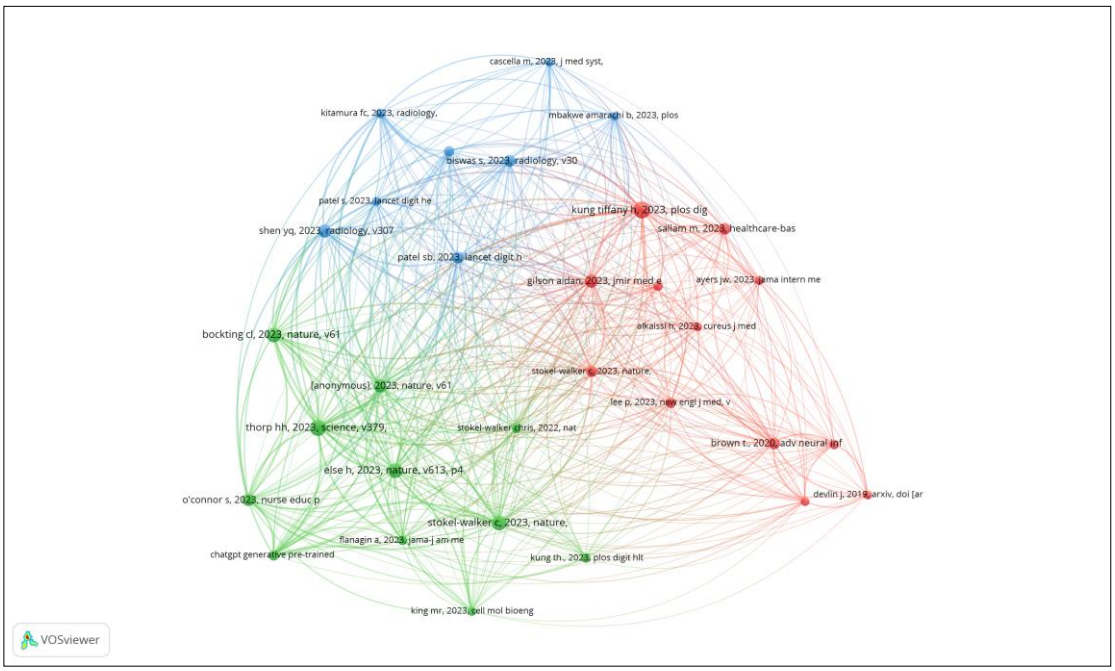

cascella m, 2023, j med syst,
kitamura fc, 2023, radiology,
mbakwe amarachi b, 2023, plos
biswas s, 2023, radiology, v30
patel s, 2023, lancet digit he
kung tiffany h, 2023, plos dig
shen yq, 2023, radiology, v307
sallam m, 2023, healthcare-bas
patel sb, 2023, lancet digit h
gilson aidan, 2023, jmir med e
ayers jw, 2023, jama intern me
alkaissi h, 2023, cureus j med
bockting cl, 2023, nature, v61
stokel-walker c, 2023, nature,
[anonymous], 2023, nature, v61
lee p, 2023, new engl j med, v
thorp hh, 2023, science, v379,
stokel-walker chris, 2022, nat
brown t., 2020, adv neural inf
else h, 2023, nature, v613, p4
devlin j, 2019, arxiv, doi [ar
o'connor s, 2023, nurse educ p
stokel-walker c, 2023, nature,
flanagin a, 2023, jama-j am me
chatgpt generative pre-trained
kung th., 2023, plos digit hlt
king mr, 2023, cell mol bioeng
VOSviewer

图 9

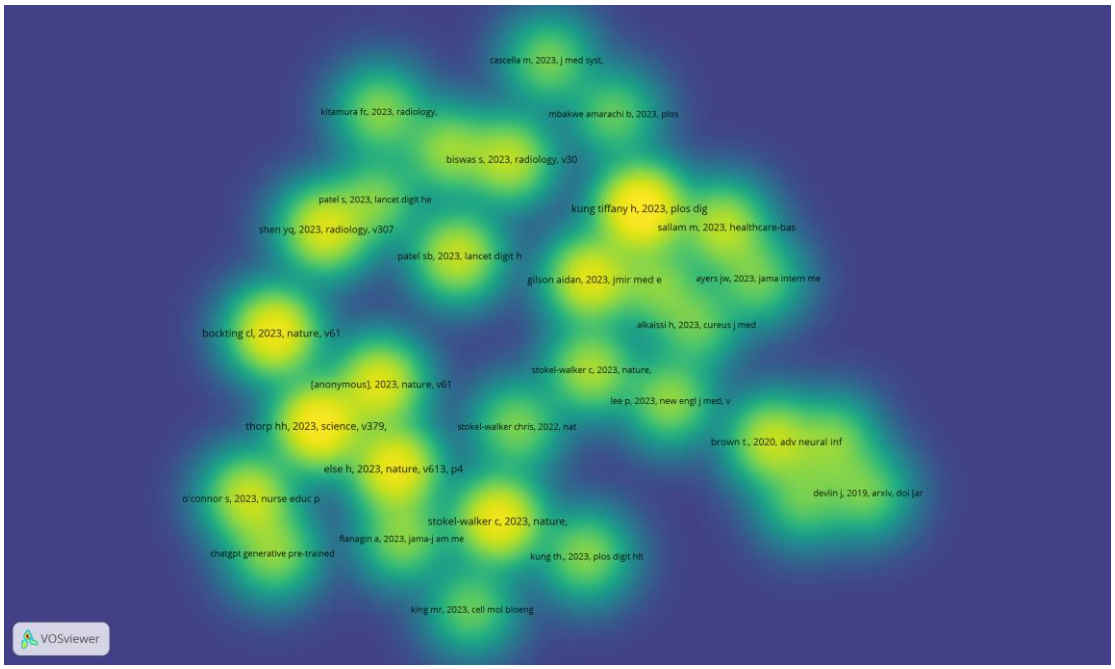

cascella m, 2023, j med syst,
kitamura fc, 2023, radiology,
mbakwe amarachi b, 2023, plos
biswas s, 2023, radiology, v30
patel s, 2023, lancet digit he
kung tiffany h, 2023, plos dig
shen yq, 2023, radiology, v307
sallam m, 2023, healthcare-bas
patel sb, 2023, lancet digit h
gilson aidan, 2023, jmir med e
ayers jw, 2023, jama intern me
alkaissi h, 2023, cureus j med
bockting cl, 2023, nature, v61
stokel-walker c, 2023, nature,
[anonymous], 2023, nature, v61
lee p, 2023, new engl j med, v
thorp hh, 2023, science, v379,
stokel-walker chris, 2022, nat
brown t., 2020, adv neural inf
else h, 2023, nature, v613, p4
devlin j, 2019, arxiv, doi [ar
o'connor s, 2023, nurse educ p
stokel-walker c, 2023, nature,
flanagin a, 2023, jama-j am me
chatgpt generative pre-trained
kung th., 2023, plos digit hlt
king mr, 2023, cell mol bioeng
VOSviewer

图 10

表 2 排名前 5 位的高被引论文

位次	题名	引用次数
1	Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models	75
2	ChatGPT is fun, but not an author	68
3	ChatGPT: five priorities for research	59
4	ChatGPT listed as author on research papers: many scientists disapprove	58
5	Abstracts written by ChatGPT fool scientists	57
6	How Does ChatGPT Perform on the United States Medical Licensing Examination?The Implications of Large Language Models for Medical Education and Knowledge Assessment	50

3. 讨论

chatGPT 自 2022 年底问世以来，人们对于它的讨论越来越多，2023 年这种讨论和关注更是达到一种高潮，这个词出现在各大媒体报道中，尤其以目前主流的自媒体平台为主[9-10]。大众和媒体对于 chatGPT 的讨论，除了其自身的的突破性研究进展外，人们也开始关注 chatGPT 在各行各业的应用以及其对人们生活、工作、理念的影响。chatGPT 的热度越来越强，意味着它的影响力也越来越高，结合时代的大背景，当下关注 chatGPT 的最新研究热点和研究动态就显得尤为必要。

随着 chatGPT 热度的上升，学术界对于 chatGPT 的探讨也越来越多，2023 年关于 chatGPT 的发文量相比 2022 年底开始激增，由此可见 chatGPT 的研究热度在不断增长，未来对于 chatGPT 的研究会越来越深入、系统和广泛。另外，关于 chatGPT 发文量的上升也说明 chatGPT 对人类的生活带来深远的影响和巨大改变，提示我们要对 chatGPT 带来的相关问题进行深入思考，并将其应用到对人类有利的地方，让其发挥更有意义的价值。

本次研究结果显示，有关 chatGPT 的国家发文量分布中，美国及中国的发文量排名位于前列，但美国的发文数量远比中国多，且美国的被引频率高于中国。这一结果表明，美国目前是研究 chatGPT 领域最活跃的国家，中国在该研究领域也具有突出的贡献。但我国人口基数大，人均发文量与美国相距甚远，因此国内研究学者应加大对该领域研究的投入，吸取国外优秀科研经验，加强与国外权威科研机构合作，提高我国的学术生产力，增强我国的学术影响力。目前，根据 chatGPT 的发展态势，可以看出 chatGPT 未来会在各个领域的应用中迎来大幅度增长，持续关注该领域的发展方向可为国内人工智能的研究带来新思路和新方向。

对关键词进行聚类分析发现，有关 chatGPT 的研究焦点主要集中在人工智能、聊天机器人、大语言模型等方面。这一研究结果从侧面说明 ChatGPT 已经成为目前人工智能生成内容的领跑者，这预示着人工智能内容生成的技术将在国内外迅速崛起并得到普及，人工智能内容生成将广泛进入人们的日常生活。ChatGPT 可以接受文本、图像、音频等多种输入，并生成相应的输出，不仅可以处理更复杂和丰富的信息，也可以提供更多样化和有用的服务。在未来的生活中，ChatGPT 有可能应用到各行各业，给人们的生活和工作结构带来巨大的影响。

综上所述，目前 chatGPT 的研究正处于逐渐深入和快速增长的阶段，国内研究可充分借鉴国外研究的热点与前沿，并结合自身特点发挥自身的优势，借助人工智能手段，加强交叉学科研究，有效促进 ChatGPT 技术在中国各行各业的应用于发展。

PU-102

基于 citespace 的复杂网络在中医药研究中应用分析

潘晔 赵移畛
河南省中西医结合医院

复杂网络 (complex network) 是构成复杂系统的基本框架，在自然界中存在的大量复杂系统都可以通过形形色色的网络加以描述。一个典型的网络是由许多节点与节点之间的连边组成，其中节点用来代表真实系统中不同的个体，而边则用来表示个体间的关系。复杂网络在中医药研究中有着较广泛应用，方药关联、证候关联都可以通过复杂网络得到清晰的呈现；中医药复方网络药理学也是其应用的主要方向之一。本研究通过文献计量学软件 citespace 梳理复杂网络在中医药研究中的应用现状，探究其创新发展方向。

1 文献计量学分析

1.1 数据来源与文献检索

在中国知网数据库中进行文献检索，设定检索关键词为 “复杂网络+中医药”为关键词，检索时间为 2013 年 1 月 1 日—2023 年 8 月 31 日。 共 检 索 到 667 篇文献,其中期刊论文有 316 篇，学位论文有 351 篇。

1.2 纳入与排除标准

纳入标准:①论文内容为通过复杂网络方法进行中医药研究;②发表的中文学术期刊及学位论文。排除标准:①报纸文献、新闻、会议论文、重复文献;②与研究主题不符的文献如复杂网络在相近行业的应用、综述、英文论文等。

1.3 文献计量学分析

将按照纳入排除标准检索得到的文献以通过两种格式导出，Refworks 格式文件用于 CiteSpace 软件分析， Excel 格式文件用于核对。 将 Refworks 文件通过 citespace 软件进行发文机构及作者分析、关键词共词分析、聚类时间线分析 3 种方式研究复杂网络在中医药研究中的应用。

1.4 数据处理

2 结果

2.1 文献机构及作者分析

从发文机构及作者的共现网络图（图 1）发现复杂网络是一种被广泛使用的中医药研究方法，各地区中医药大学均采用，未呈现明显的地域性。其中北京中医药大学及中国中医科学院、南京中医药大学的应用最多；胡孔法及周雪忠等领域专家发文数量较高，并有极强的辐射效应，各地区中医药大学间多有合作交流。

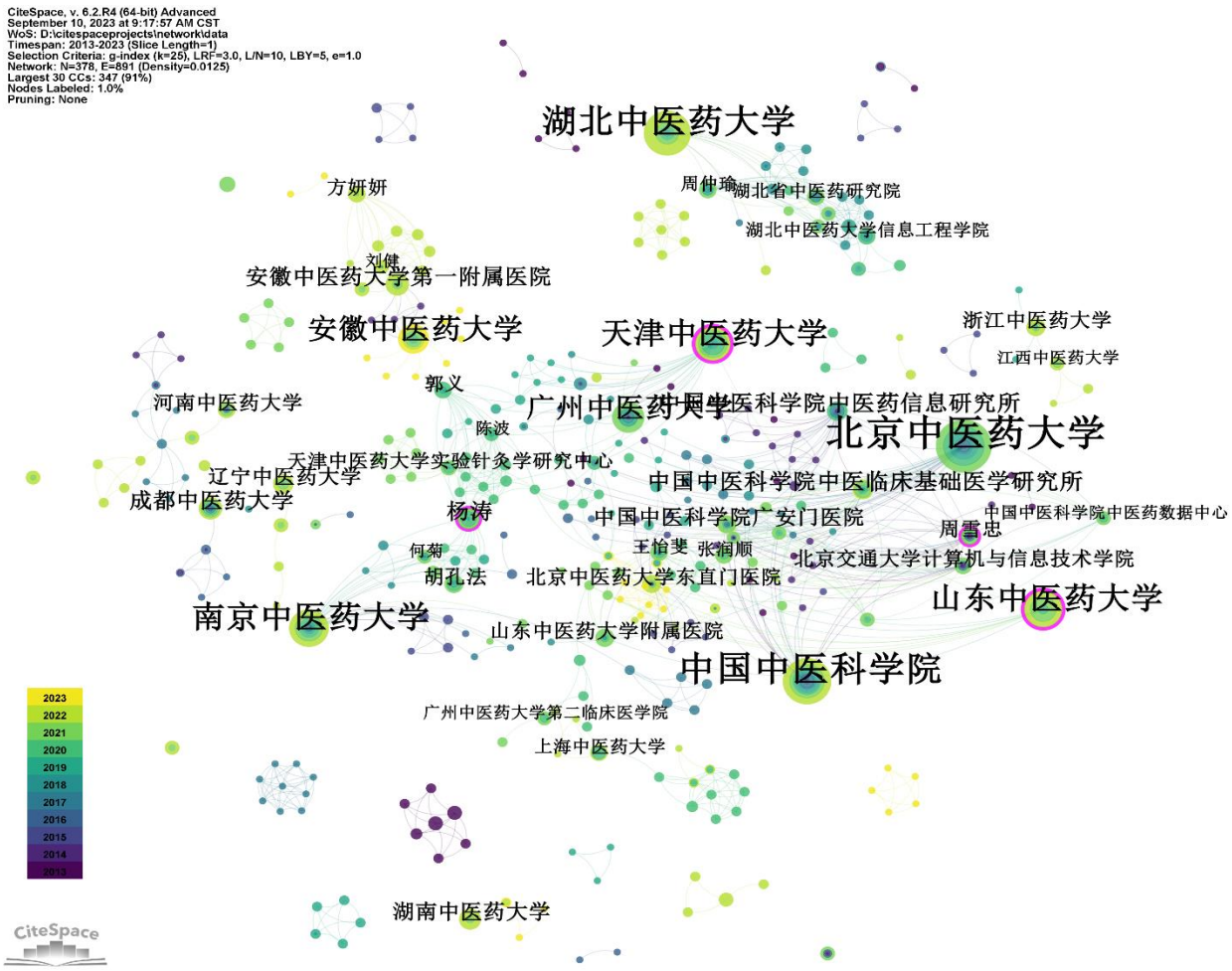


图 1 文献机构及作者共现网络图

2.2 关键词共词分析

从关键词共词分析（图 2）发现，较大的节点集中在数据挖掘、复杂网络、用药规律及中医药。在实际应用中，复杂网络常与数据挖掘共同使用，例如中医药数据挖掘的一种常用模式为利用关联规则得到药物间共现次数，再通过网络图进行图形展示。所以关联规则为用药规律及复杂网络之间的一个节点。复杂网络与数据挖掘的应用范围较为一致，多用于用药规律的发现，主要应用对象为名医经验、名老中医经验，发现药物间的配伍。

复杂网络常用于针灸穴位挖掘，穴位相对于中药饮片种类较少，通过复杂网络能比较清楚地展示多个穴位之间的关联。复杂网络常用于中医复方的机制研究，也就是中医复方网络药理学，通过检索相关数据库得到复方组成药物的成分及对应靶点，构建网络图，并对靶点的进行核心成分筛选最终进行机制的验证。关键词中还有分子对接，中药复方网络药理学逐渐与药物设计和分子生物学领域的技术结合，更深度地探索中医药复方治疗疾病的原理及作用机制。

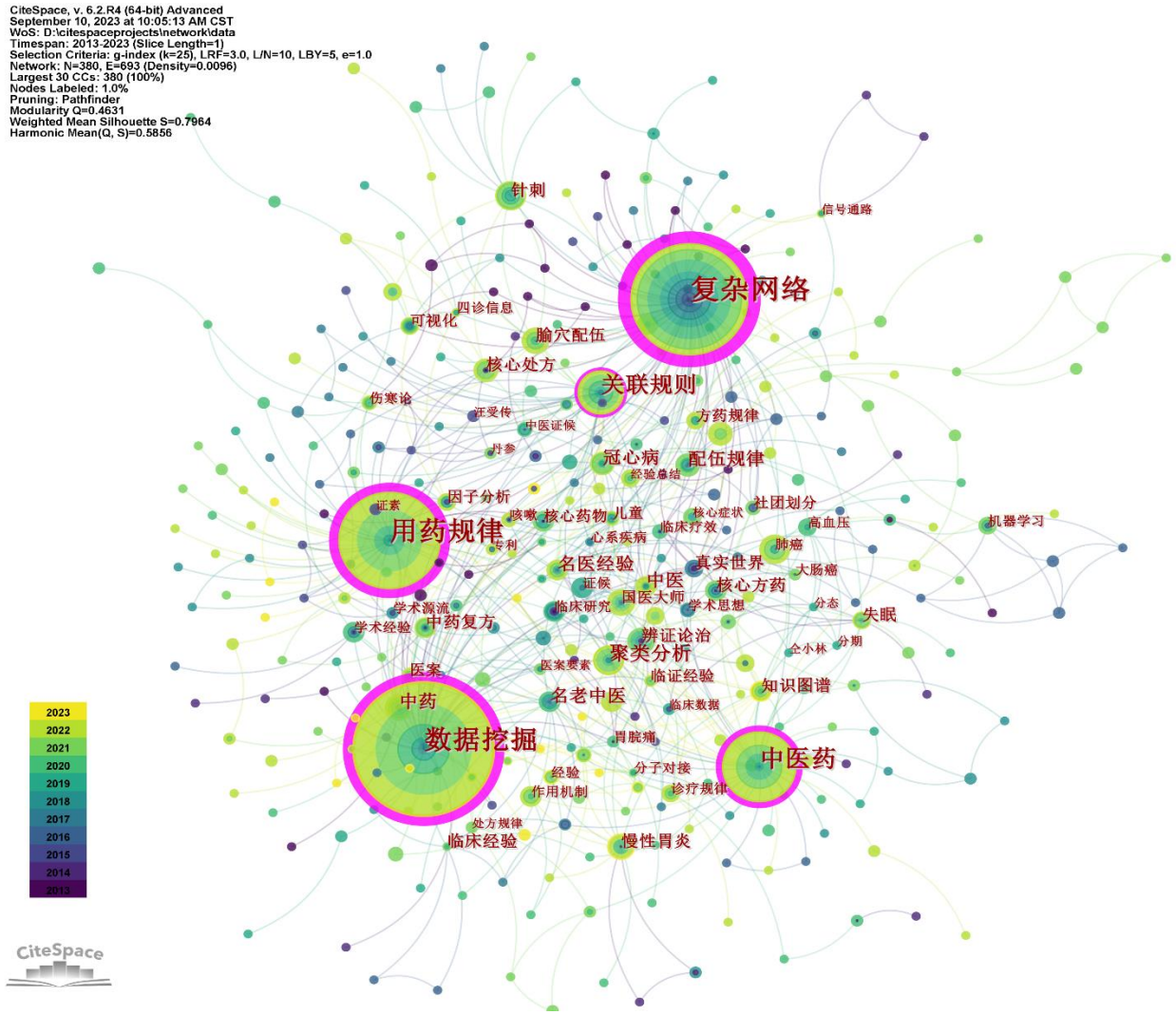


图 2：关键词共现网络图

2.3 复杂网络在中医药领域应用的时间线聚类分析

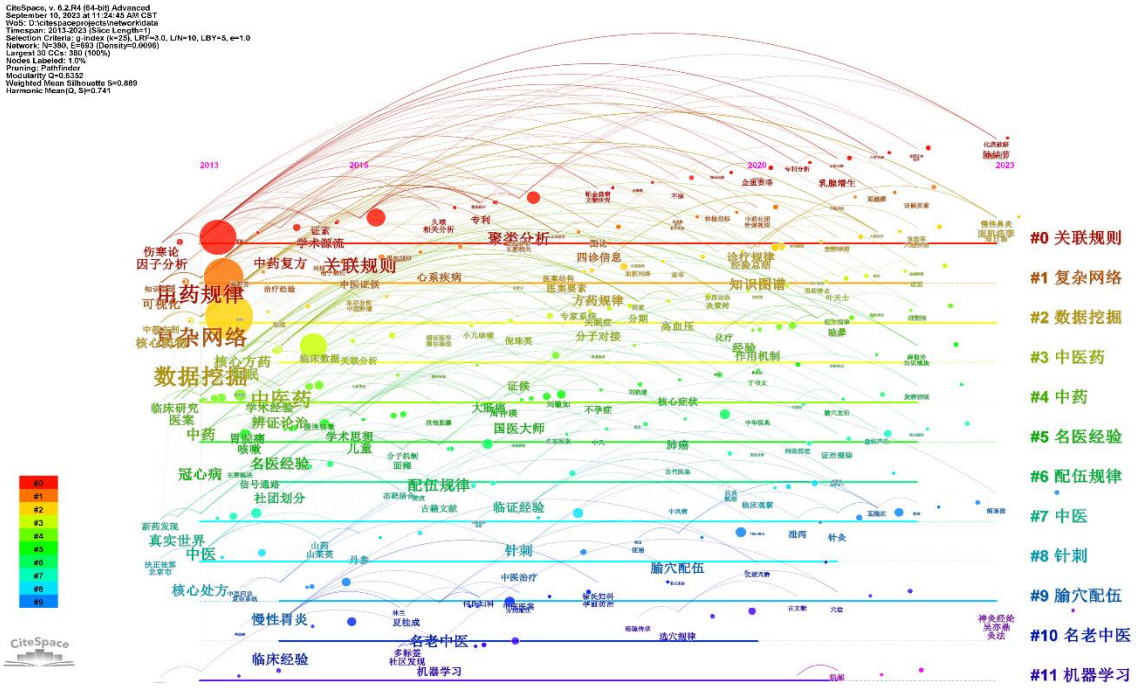


图 3：关键词聚类时间线图

从聚类时间线图，发现复杂网络应用始于 2013 年，稍晚于数据挖掘 [1]，两者有共同也有差异。两者均涉及了中医古籍、经验传承、医案挖掘等核心研究内容。均覆盖了中医辨证论治的主要过程，包含了“证型”、“症状”、“体征”、“治法”、“用药”等中医的重要元素。但数据挖掘往往分开探讨这些元素，复杂网络可以独立探讨构建，也能融合再一起。复杂网络在针灸玄学规律中应用更广泛，2015 年以来作用机制研究更多，也就是在网络药理学方面的应用更广泛。知识图谱这一基于复杂网络展示的知识表达方式，也在近年广泛开展。

3 分析

3.1 中医辨证论治与复杂网络

中医辨证论治有整体观及个性化两大特点，包含识别症状得到证型及通过证型进行用药两个过程。通过数据挖掘等方法进行研究时，为了能更深入，往往仅对其中一个过程进行分析。由于中医单味药种类丰富、组合变化丰富，所以用药规律一直是研究重点，很少有研究将这两个过程综合起来分析，而复杂网络则能满足这一需求。中国中医科学院中医药信息研究所通过构建知识图谱这一复杂网络方法，进行中医皮肤科流派用方经验的知识展示[2,3]，同时构建流派传承知识图谱，并且能融合在一起，实现更丰富的知识展现与应用。蔡云采用 Protege5.5 作为本体的建模工具,OWL 为本体描述语言,建立国医大师周仲瑛教授治疗肺癌领域本体类关系模型即概念树并形成可融合网络，对周仲瑛教授治疗肺癌经验知识进行全面梳理、建设数据库,并与后续数据挖掘成果相互印证[4]。

3.2 复杂网络与中医复方网络药理学

由于关键词的问题，中药复方网络药理学相关文献在此次检索中出现不多，但是近年来发展迅速。中药复方有多成分多靶点的特性，成分靶点间关系复杂，可以将每一味中药、蕴含靶点各看成一个小型复杂系统。所以不仅可以建立药物-成分网络、药物-疾病-靶点网络、还可以构建蛋白质互作网络（Protein-Protein Interaction Networks，PPI）、靶器官网络等等。随着中医复方网络药理学的蓬勃发展，网络分析法得到了充分的应用。但文超通过 TCMSP 数据库获取白术、枳实、荷叶的主要化学成分及其靶点,根据 ADME 筛选中药活性组分，通过 Gencards、OMIM、TTD、DRUGBANK 数据库获取血脂异常主要靶点，利用 String 平台进行蛋白质相互作用分析,构建 PPI 网络并挖掘网络中潜在的蛋白质功能模块，分析枳术丸治疗血脂异常的药理机制,为新药研发及经典方剂临床拓展运用提供参考[5]。石雅馨基于网络药理学探讨桂枝茯苓丸对任脉癥瘕类病"异病同治"分子机制研究，并采用灵活应用复杂网络方法构建关键靶标-组织分布网络图及关键靶标-亚细胞分布网络图，有所创新[6]。网络药理学由于过程严密、方法科学、形式新颖，引发了中医药研究者的极大热情。

3.3 复杂网络指标与应用

复杂网络的统计特征一般通过节点的度及度分布、平均距离、簇系数、介数、度相关性、社团结构等属性进行度量。其中节点的度表示与当前节点直接相连的边数，度分布描述节点度的分布情况。平均路径长度为网络中所有节点对的平均距离；节点簇系数为节点的所有邻居节点间实际存在的边数与最多可能存在边数的比值。在中医药研究中，这些统计指标的应用应用较少，仅仅是展示共现。而在复方网络药理学中，不仅仅通过网络图展示药物、成分、靶点、疾病之间的关联，还需应用这些统计指标进行关键成分、靶点的筛选。Cytoscape 软件在构建复杂网络中起到了关键作用，可进行各种个性化的网络构建、分割、融合。cytoHubba 插件可以通过对上述指标进行筛选从而到关键节点，生物信息学分析的多个插件有助于更好的进行机制验证。这些方法可以应用于中医药辨证论治规律的发现，但是尝试的较少。

3.4 复杂网络在中医药研究中应用的创新趋势

复杂网络将复杂系统内部元素抽象为节点，元素间关系抽象为边，从而构建网状结构。看似简单的结构，却具有强大的表现力。针对节点，可以通过大小、颜色、形状进行区分，针对边，可以通过颜色、粗细、透明度、方向区分，从而包含关系，展示知识。中医药研究者利用复杂网络的特性，结合团队特色进行了不同程度的创新。胡孔法等首先根据方剂指向药物的情况来衡量药物的重要性并对药物排序，使用 PRCN 算法来确定的核心药物，要比复杂网络中原有的仅仅根据药物的度来确定核心药物更加具有可信度，更深入[7]。随着通用人工智能技术的迅猛发展，标准化、智能化在中医药研究中更趋明显，复杂网络符合这一趋势。刘震团队自主开发了基于数字经络理论的智能针灸机器人，具备基于经穴主治与症状联系的小世界网络自动配穴这一核心功能，初步实现了基

于图论的针灸智能取穴功能，是复杂网络在选穴规律应用的深化[8]。将复杂网络与人工智能算法结合，将展现的知识充分利用，才是其发展的创新之处。

PU-103

高校图书馆中文图书纸电同步采访关键因素分析 ——以大连医科大学图书馆为例

邓岩彬* 刘强 刘玮
大连医科大学

1 背景及文献调研

馆藏中文电子书采访始于 20 世纪末，以超星、方正 Apabi、书生之家为代表的电子书馆配商将纸质图书电子化加工后以打包买断为主的方式提供服务。随着包括数字版权保护等技术的迅速发展和电子书相关政策法规的日益完善，纸电融合出版已经成为主流。中文电子书经过二十余年的发展，馆配商取得图书版权后将其电子化并销售的方式逐渐被以出版社为主导的定价和销售模式替代：馆配商提供平台，出版社自主上架电子图书，如北京人天集团的畅想之星平台、湖北三新有限公司的云田智慧图书馆云平台、中新金桥数字科技有限公司“可知”知识服务平台、科学出版社的科学文库等，这些都为高校图书馆的纸电同步采访提供了选择。

2022 年 6 月在 CNKI 平台检索“主题=中文电子书 OR 主题=纸电”，同义词扩展检索得到 390 篇文献，通过学科分布可视化分析可以看出文献主要集中在图书馆情报与数字图书馆（34.99%）、出版(25.92%)以及计算机软件及计算机应用(15.98%)三个学科领域。其中“主题=纸电”的 146 篇文献发表年度趋势显示 2016 年之前文献量极少，自 2016 年开始相关研究明显增加，并且每年呈显著上升趋势，近两年达到较高热度。电子书的发展和纸电同步采访取决于图书馆的需求、出版社的意愿以及计算机技术的支持。高校图书馆中文图书纸电同步采访工作已经由理论研究提上了实施日程。段晓林[1]对超星、中文在线、田田网和畅想之星四家平台的资源、功能、运营模式和理念进行了比较，分析其满足图书馆需求的程度。同时也指出面对各有优势的电子书平台，图书馆有必要根据自身情况制定相应的电子书平台采购策略。

2 研究思路和方法

2.1 研究对象

高校图书馆采访馆员能够熟知学校的专业特点和学科建设以及本馆的馆藏政策和采访原则，了解本校读者的专业需求，对纸书的采访有丰富的经验，可持续发展性强，其采访的纸质图书与读者需求契合度较高，因此该研究以本馆的年度入藏图书目录为对标样本。由于纸本图书从入馆到上架流通需要一定的加工时间，本研究容许电子图书时滞三个月到一年，平台电子书目的比对时间选在次年的三月份，即完成全年纸本入藏后三个月，使电子书目与对标样本书目，内容上相对同步。因为学校是以医学为主，理学、工学、管理学、艺术学、法学和哲学等多学科发展的医科大学，人民卫生出版社一直是本馆的首位核心出版社，通过与出版社沟通了解到“可知”知识服务平台与该出版社合作较密切，同时综合考虑电子书平台服务商的电子书规模、国内图书馆市场占有量、纸本图书服务等因素，选取畅想之星平台和“可知”知识服务平台（以下简称为“畅想之星”和“可知”）作为比较对象。本研究组成员自 2018 年起即筛选几个电子书平台与馆藏年度新书做比对，但 2018－2019 年各电子书平台还处于起步阶段，新书上架率低，图书种数匹配均不到 5%，研究意义不大。因此书目比对时间为 2020 年和 2021 年。

2.2 书目比对方法

导出本馆 2020 年度和 2021 年度纸本图书入藏清单，与“畅想之星”和“可知”分别在 2021 年 3 月份和 2022 年 3 月对相应年度入藏清单进行电子书匹配并返回结果。将数据清洗后使用 EXCEL 进行统计：筛选匹配的书目，合并相同出版机构的不同描述（如将中信出版集团、中信出版集团股份有限公司、中信出版集团有限公司、中信出版集团有限责任公司、中信出版社、中信出版社股份有限公司、中信出版社集团股份有限公司统一为“中信出版社），按出版社分类汇总。

2.3 确定本馆核心出版社的方法

通过纸书年度入藏书目按出版社汇总，人民卫生出版社、科学出版社、江苏凤凰文艺出版社、

中信出版社、北京联合出版公司、人民文学出版社、化学工业出版社、北京大学医学出版社等一直是本馆的重要核心出版社，年采购书目种数均达 150 种以上。可以这些出版社确定为本馆重要核心出版社。

根据二八法则，分别筛选 2020 和 2021 年纸书年度入藏种数前 20%的出版社，2020 年，410*20%=82 个出版社，查到第 82-85 位次出版社书目种数都是 25 册，最终确定该年度纸书的 85 家核心出版社，这 85 家出版社采购的纸书种数为 6826 种，占总数 8826 种的 78.03%。2021 年，386*20%=78 个出版社，查到第 78-80 位次出版社书目种数都是 25 册，最终确定纸书的 80 家核心出版社，这 80 家出版社采购的纸书种数为 6152 种，占总数 8186 种的 75.1%。两年的核心出版社重合达 75%，核心出版社筛选结果符合二八法则，核心出版社规律明显。

2.4 筛选纸电同步采访出版社

根据筛选的核心出版社确定两个平台上电子书匹配率大于等于 10%的出版社为纸电同步采访的主要出版社。对于非核心出版社中电子书匹配率较高（如高于 50%），或种数大于等于 10 种，匹配率 30%以上的出版社也可以做为关注的对象。

3 研究结果

3.1 匹配到的书目和出版社的数量为实施纸电同步采访提供依据

本馆 2020 年度和 2021 年度纸本图书入藏清单，与“畅想之星”和“可知”匹配汇统计结果见表 1：

表 1. 2020 和 2021 年度书目及出版社比对结果

年度	2020 年度			2021 年度		
对象	本馆	可知	畅想之星	本馆	可知	畅想之星
出版社数量	410	33	129	386	76	147
书的种数	8826	591	1086	8186	815	1334
出版社均图书种数	20.87	17.91	8.42	21.21	10.72	9.07

近两年电子书平台上架书目种类和出版社数量发展呈年均稳步增长。2020 年可知平台虽然匹配的出版社数量不多，但出版社平均匹配到的图书种数为 17.91 册，远高于畅想之星的 8.42 册，说明平台合作的出版社上架电子书积极性高。此外可知平台新加入的出版社两年增长一倍多（匹配率由 8%增到 18%），这也说明平台起步虽然相对晚，但发展较快。畅想之星平台上出版社和图书种类匹配率较高，两年入驻出版社满足本馆要求均不低于 30%。两个平台上无论是出版社数量还是图书种数都有明显增长，如图 1 所示。

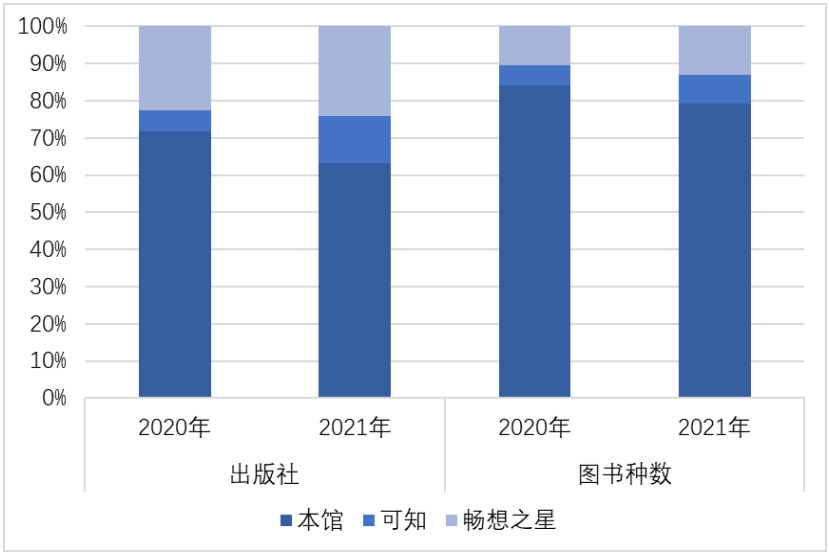


图 1 2020-2021 出版社和图书种数匹配占比情况

3.2 根据本馆确定的核心出版社筛选纸电同步采访出版社

将本馆的重要核心出版社作为纸电同步采访的重要目标，甚至可以据此去寻求合作的电子书平台。据统计数据可以看到可知平台上本馆重要核心出版社--人民卫生出版社图书匹配率达 19.8%和25.2%。

根据筛选的核心出版社确定两个平台上电子书匹配率大于等于 10%的出版社为纸电同步采访的主要出版社。2020 年和 2021 年畅想之星平台的核书出版社覆盖率分别是 51.8%（44/85）和50%（40/80），远高于可知的 15.3%（13/85）和 26.3%%（21/80），这也意味着该平台可以纸电同步采购的图书数量较丰富。

化学工业出版社电子书在这两个平台上的上架率都较高，与本馆年度入藏新书目匹配率达70%左右，是一个高度纸电融合出版的机构，可以作为图书馆纸电同步采访合作度非常高的出版机构。

表 2. 2020 和 2021 年度化学工业出版社电子书与本馆纸书比对情况

化学工业出版社	本馆纸书种数	可知	可知匹配率	畅想之星	畅想之星匹配率
2020 年	208	153	73.56%	142	68.27%
2021 年	106	74	69.81%	74	69.81%

从书目匹配率来看化学工业出版社、社会科学文献出版社、人民邮电出版社、电子工业出版社、东南大学出版社、中国人民大学出版社、浙江文艺出版社、北京出版社、北京大学出版社、北京时代华文书局、中国文史出版社、青岛出版社等与两个平台的合作关系相近。其它的出版社都各有疏近之区分。如果图书馆同时选择几个平台采购电子书，可以根据电子书匹配率进行分工分社重点采访。

图书馆在电子书出版社筛选时对于非核心出版社中电子书匹配率较高出版社也要加以关注，作为纸电同步采访的可选资源。

科学出版社是本馆的重要核心出版社之一，每年采购纸本图书量均排在前三位，但该出版社拥有自己的电子书平台“科学文库”，与其它电子书平台没有合作。这种提供电子书的服务方式也是几大外文出版集团电子书的主要发行服务方式。图书馆可以根据本馆需求选择与其单独合作。

4 存在的问题及解决方案

4.1 高校图书馆中文图书纸电同步采访面临的困境

4.1.1 纸电同步采访需求更加显著

今年 4 月 23 日第 27 个世界读书日到来之际，中国新闻出版研究院院长魏玉山发布第十九次全国国民阅读调查结果[2]中显示：成年国民数字化阅读倾向明显，中青年人成为数字化阅读的主体。高校读者以青年人为主体，2020 年以来疫情时代线上线下混合式教学越来越常态化，近年来通过访谈了解到新一代的大学生读者对电子图书的需求已经不仅仅是纸书的种类补充，而是更青睐于电子图书使用便捷的优势。在图书馆已有纸书的情况下还是希望能够使用电子图书，便于教师读者教材编写，方便学生做学习笔记等。包库的电子书采购方式满足率性价比均不高，高质量的纸电同步单本采购成为经济实用高效的解决办法。

4.1.2 高质量电子书种类不够丰富

本馆一直关注中文电子书市场的发展，2018 年和 2019 年对畅想之星平台、云田智慧图书馆云平台上架新书与馆藏年度新书做比对，匹配率均不到 5%。通过浏览目录也可看到平台电子图书数量不少但质量不高。辽宁师范大学图书馆李树凯[3]在纸电同步采购的实践中也感到可采选书目较少。出版机构的纸电出版率还是不高，特别是电子版学术性图书较少，可供采选的纸电同步书目较少。北京大学出版社张建民[4]也认为不管载体形式是纸质书还是电子书，高校图书馆在资源建设中看中的永远是图书内容本身，看资源是不是符合高校学术需求和专业需求，这才是根本。

4.1.3 电子书的著作权较纸书更为复杂

《著作权法》第十条第十二项规定：“信息网络传播权，即以有线或者无线方式向公众提供，使公众可以在其选定的时间和地点获得作品的权利”。中文电子书的信息网络传播权因作者、出版社等多方面因素授权时长不一致，存在购买后图书被平台下架的风险。尽管书商建议使用本地镜像的保存和访问方式，但严格意义上并未完全解决信息网络传播权问题。希望通过相关法规，完善电子书信息网络传播权领域的“避风港”原则。

4.2 高质量中文图书纸电同步采访解决方案

4.2.1 经费投入保障

张赟玥等[5]应用 ERS 评估外文电子图书绩效的实践探索中发现电子图书馆藏建设的连续投入较为欠缺，采购呈现明显的波动。对其建设定位，通常是作为经费充裕时的优先扩藏对象，与经费紧张时的停订首选对象。要发展高质量中文图书纸电同步采访就必须在经费投入上有保障。国内高校图书馆的纸本图书复本量大都不止一册，电子书替代一个复本是相对更合理的馆藏建设策略，纸电同步采访经费预算上可以与纸本图书协同，而不是纳入电子资源经费管理中。

4.2.2 纸电融合出版稳步发展

虽然高质量电子书种类不够丰富，但是近两年电子书平台上架书目种类和出版社数量发展稳步上涨。段双喜[6]使用书目比较法进行系统性的实证分析，其研究时期（2018 年以前）馆配商平台在纸本书出版一年内基本无法采购到相应的电子书，出版社第二至四年内才可能会渐次发售电子本。并提出高校馆可以率先采取“单纸本”策略，以消除出版方惜售电子本的心理。高质量纸电同步采访需要执行“单纸本”策略并打破包库购买的模式，与纸书一样采用单本采访的方式。这都将会促使高质量电子书上架，为图书馆常规化采购中文电子图书提供可选书目日趋丰富。

4.2.3 图书馆可直接与确定的目标出版社沟通谈判

图书馆可以直接与本馆的重点出版社、核心出版社进行沟通和谈判，表达对电子书采购的意向，促进出版社电子书上架决心，加快上架进程。期望出版社能顺应数字出版的发展，作出积极的回应，更好地服务于读者服务于社会。出版社应该意识到图书特别是学术研究类图书的电子化出版已势不可当，学术类图书电子出版或将成为主流，其相应纸质图书可能会发展为按需印刷。

4.2.4 电子书平台服务商与图书馆的关系变得更合作

电子书平台服务商成为图书馆和出版社之间的纽带和桥梁。有些出版社的电子书定价偏高，电子新书价格高达纸质新书的 1.5-2 倍。对于大宗长期有采购需求的出版社，图书馆可以充分与电子书平台服务商合作，与出版社进行价格谈判，达成良好合作获得较优的服务。

4.2.5 将中文图书纸电同步采访纳入馆藏策略中，并鼓励读者参与其中

高校图书馆在执行纸电同步采访时应本着循序渐进的原则，有计划开展该项工作。王子君[7]分析知识付费产业发展为出版业带来机遇，指出将会提升出版业内容输出的精准度。出版的本质是对知识内容的筛选和编辑，在知识付费时代，用户在出版环节的参与度和交互性发生了本质的改变。持续而稳定的购销关系以及借助高校读者决策采购（PDA）的采访模式将会反作用于出版社的出版策略。

5 结语

高校图书馆中文图书纸电同步采访关键因素即是图书本身，图书的纸电同步出版既包括内容的一致性，也体现在出版时间的同步性上。高质量发展纸电同步采访需要在包括计算机技术等各项技术支持下，图书馆、出版社以及电子书平台服务商为方便知识传播和再生产共同作出积极的努力。

PU-104

Venetoclax 耐药机制研究热点与趋势分析——基于 CiteSpace 文献计量视角

吴利俊[1]，王倩[2#]，应峻[1#]
1. 复旦大学图书馆
2. 复旦大学附属华山医院血液科

引言

维奈克拉（Venetoclax）作为全球首个获批的 B 细胞淋巴瘤/白血病-2（BCL-2）蛋白口服抑制剂，2018 年被美国 FDA 批准，与去甲基化药物或低剂量阿糖胞苷联合使用，适用于老年人或因条件不适合行诱导化疗的急性髓系白血病（Acute Myeloid Leukemia, AML）患者[1]。与其他 BCL-2 抑制剂相比，Venetoclax 引起血小板减少的不良反应小，对患者具有保护作用[2]。在年龄大于 75 岁的 AML 患者中分别给予阿扎胞苷单药或阿扎胞苷联合 Venetoclax 治疗，接受阿扎胞苷联合 Venetoclax 治疗患者总生存期较长，缓解率高于阿扎胞苷单药患者[3]。虽然 Venetoclax 单药或与

其他药物联合用药的疗效已得到认可，但在临床应用中发现患者存在耐药现象，目前关于 Venetoclax 耐药的机制仍不明确[4]。

目前研究主要围绕 Venetoclax 作用机制、在血液系统肿瘤中的临床应用以及包括耐药在内的一些相关问题进行探讨[5-8]。为了更全面了解 Venetoclax 的耐药机制研究现状，需要借助新的理论和研究方法。文献计量分析作为一种相对较新的文献研究方法，可以根据文献数据库提供的信息，通过定性和定量的方法来评估研究活动的发展趋势，掌握特定领域的发展情况，并为比较不同层次的贡献提供信息[9]。CiteSpace 是一种基于“共引分析理论”的文献计量工具，旨在揭示相关研究领域的新技术、热点和趋势[10]，也是科学知识图谱制作和分析的一项最常用工具[11-13]，可以对特定领域进行文献计量分析，绘制一系列图谱，探索科学研究领域演变的关键路径和前沿发展。

因此，本文基于文献计量学方法对 Venetoclax 耐药相关英文文献进行统计分析，并使用 CiteSpace 软件将统计结果进行可视化，希望可以帮助研究人员快速了解该领域目前的研究现状和热点以及未来趋势，从而为进一步的基础和临床研究提供新的思路和方向。

1.数据与方法

1.1 数据采集

本研究主要基于 Web of Science 核心合集数据库的科学引文索引扩展（SCIE）中检索了文献数据，检索字段设定在主题，使用的关键词主要包括：“venetoclax OR Venclexta OR ABT-199 OR GDC-0199 OR RG-7601 OR RG7601”和“resistance OR resistant OR resistent”，检索时间为 2023 年 1 月 18 日，限定文献类型为原始研究性论文、综述，文献语言为英文，排除会议、信件、编辑材料、书籍和摘要等文献类型，共获得 803 篇相关文献。

1.2 数据处理与分析

文献常用题录信息、引文和源期刊被导入 Microsoft Excel 2016 进行数据处理和可视化分析。以纯文本格式导出了检索到的文章，包含全记录和引用数据，名为“download_XXX.txt”，然后导入 CiteSpace 6.1.R2 进行进一步分析。在绘制知识图谱时，遵循 CiteSpace 软件主要步骤，设置了时间切片、阈值、建模、修剪、合并和映射[14]等主要参数，具体如下：从 2013 年 1 月到 2023 年 1 月，时间切片为 1 年。选择标准为 g 指数，比例因子为 κ=25。CiteSpace 的核心概念包括突发检测、中介中心性和异构网络，这有助于及时可视化研究现状、热点和前沿[15]。可视化图谱中的节点分别表示作者、机构、国家或关键字、参考文献等。节点的大小表示出现或引用的频率，节点的颜色表示出现或被引用的年份。

2.结果

2.1 论文年度发表和引用情况

通过对样本数据的分析，发现 2013 年至 2023 年间，该领域共有 606 篇研究论文和 197 篇综述论文获得发表。如图 1 所示，总引用次数为 21407 次，篇均被引频次为 26.74 次，年均引用频次为 1951.82 次。第一篇关注 Venetoclax 耐药性机制的文章发表于 2013 年。2013 年至 2022 年，论文发表数量逐年增加，呈现快速发展趋势，引用次数也迅速增加。2016 年和 2019 年出现了两个高峰，该领域的引用次数达到了峰值。从 2019 年到 2022 年，引用次数分别呈逐年下降趋势。

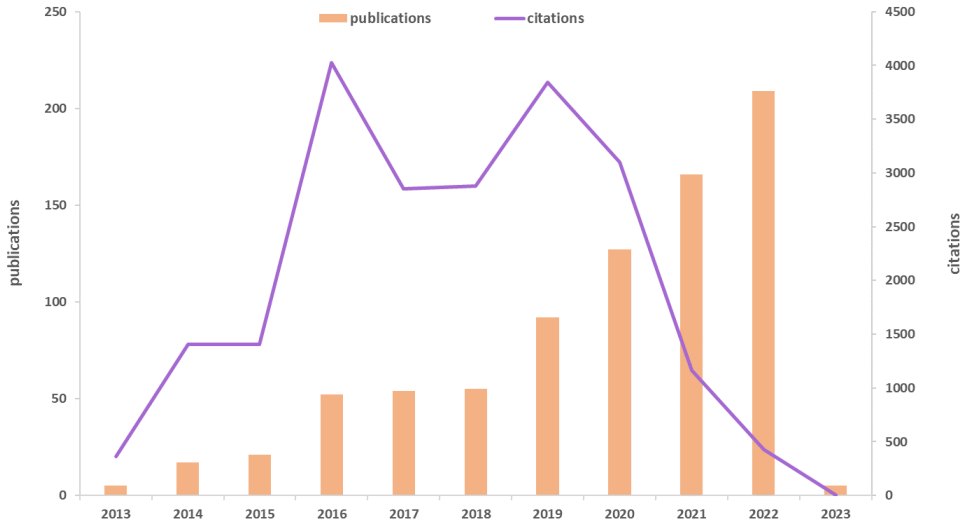


图 1. Venetoclax 耐药机制相关研究的年度论文发表和引用情况

2.2 期刊分析

通过分析样本数据发表所属期刊，可以发现该领域的论文发表在 256 种期刊。如表 1 所示，其中《Blood》是该领域发表最多的期刊，共有 42 篇论文，该期刊在过去 5 年中的平均影响因子 19.6，2022 年的影响因子为 22.54。排名第二的期刊是《Cancers》，共发表了 40 篇相关论文，该期刊 5 年平均影响因子为 5.6，2022 年的影响因子为 6.886。以上期刊均为 Q1 区期刊，表示在专业领域拥有较高水准。

表 1. 发文量排在前 10 位的期刊名称、影响因子以及 JCR 分区情况

Journals	Publications	Impact Factor	JCR Quartile
Blood	42	22.54	Q1
Cancers	40	6.886	Q1
Leukemia	32	12.256	Q1
Frontiers in Oncology	28	6.122	Q2
Blood Advances	25	7.545	Q1
Clinical Cancer Research	25	13.975	Q1
Cell Death & Disease	21	9.62	Q1
Oncotarget	18	5.312	Q2
Haematologica	17	10.29	Q1
Leuk. Lymphoma	15	3.011	Q3

2.3 作者、机构、国家共现分析

为了绘制可视化图谱，本研究选择了树环形状作为节点显示模式。节点之间的线条表示合作，线条的颜色表示第一次合作年份。该领域作者合作网络共有 384 个节点和 1252 个链接组成，如图 2 所示，针对具有最大轮廓值（加权平均值 0.9216）的 10 个聚类进行可视化展示。排在前 5 位的高产作者及其中心度分别是 Konopleva M（34.0.18）、Roberts AW（22.0.01）、Andreeff M（17.009）、Dinardo C（16.001）和 Wang Y（15.013）。尽管共有超过 5000 名研究人员发表了相关文章，但他们之间的合作相对很少，合作网络规模较小。此外，作者的中心度相对较低，这表明未来需要更多高质量和大规模的合作。以关键词作为聚类标签，可以发现知识图谱中显示的 10

个聚类的研究方向主要包括 AKT、急性髓系白血病、谱系、化学免疫疗法、滑膜肉瘤、治疗失败、mivebresib、病例报告等。

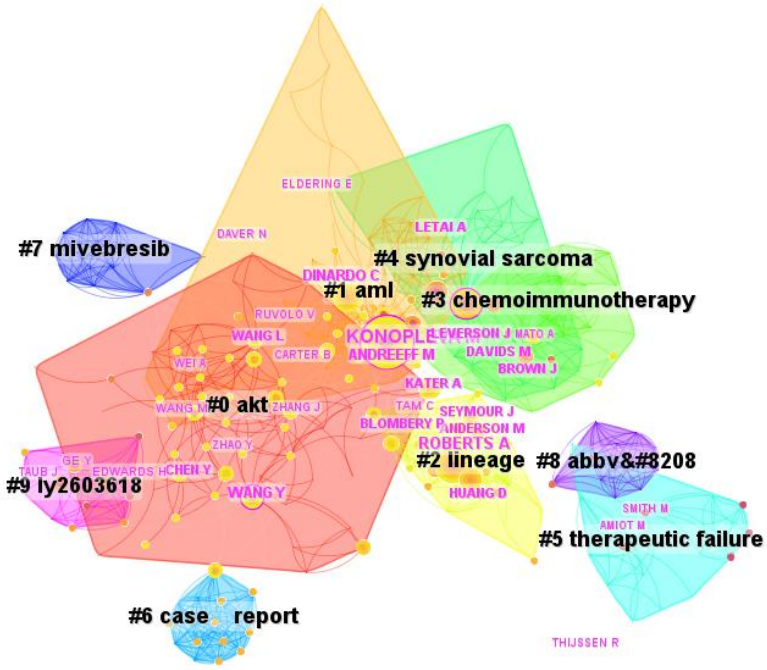


图 2. Venetoclax 耐药机制的作者合作网络和聚类分析

机构合作网络共包含 366 个节点和 1144 个链接（图 3A）。其中，排名前 5 位机构分别是美国德克萨斯大学 MD Anderson 癌症中心（92 篇）、Dana Farber 癌症研究所（47 篇），澳大利亚墨尔本大学（47 篇）、Walter 和 Eliza Hall 医学研究所（37 篇）和 Peter MacCallum 癌症中心（33 篇）。MD Anderson 癌症中心和 Dana Farber 癌症研究所的中介中心度均为 0.22，其他三个机构均低于 0.1，表明这两个机构在这一领域合作网络中有重要作用。如图 3B 所示，国家合作网络由 73 个节点和 287 条链接组成。美国是最领先的国家（402 篇），其次是中国（113 篇）、澳大利亚（68 篇）、德国（67 篇）和英国（52 篇）。美国和中国的中介中心度分别为 0.65 和 0.55，其他三个国家均低于 0.1，说明这两个国家在该研究领域具有重要位置。

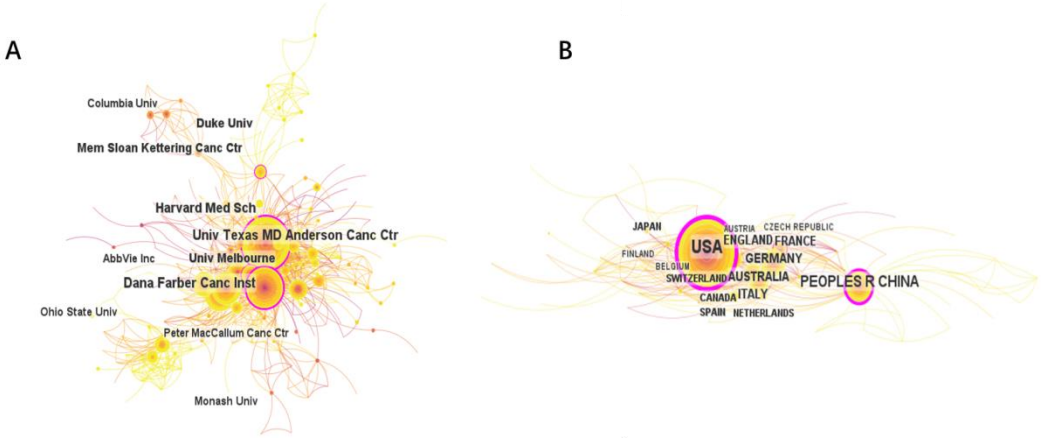


图 3. Venetoclax 耐药机制的机构（A）和国家（B）合作网络

2.4 关键词共现和聚类分析

关键词通常是一篇论文的高度提炼和概括性内容，通过从样本数据中提取关键词，基于关键词共现分析，可以找到那些出现频率高、中心性强的关键词，这往往代表了一个领域的研究热点和发展趋势。在本研究中，获得了由 382 个节点和 2734 个链接组成的关键词共现网络。表 2 列出了频率最高、中心性最高的前 10 个关键词。图 4 则是通过对数似然比（LLR）方法对关键词进行聚类分析，从纳入文献的关键词中提取了最大的 8 个聚类，以#作为标记。聚类标签主要包括急性髓系白血病、慢性髓系白血病、多发性骨髓瘤、伊布替尼、阿扎胞苷、表达、细胞色素 C 和家族蛋白等，这些聚类标签通常都是一些研究热点方向的表征。

表 2. 排在前 10 位的高频关键词

Rank	Frequency	Keywords	Rank	Frequency	Keywords
1	264	Venetoclax	6	129	expression
2	229	resistance	7	112	inhibitor
3	219	chronic lymphocytic leukemia	8	104	cancer
4	168	apoptosis	9	91	open label
5	143	acute myeloid leukemia	10	74	ibrutinib

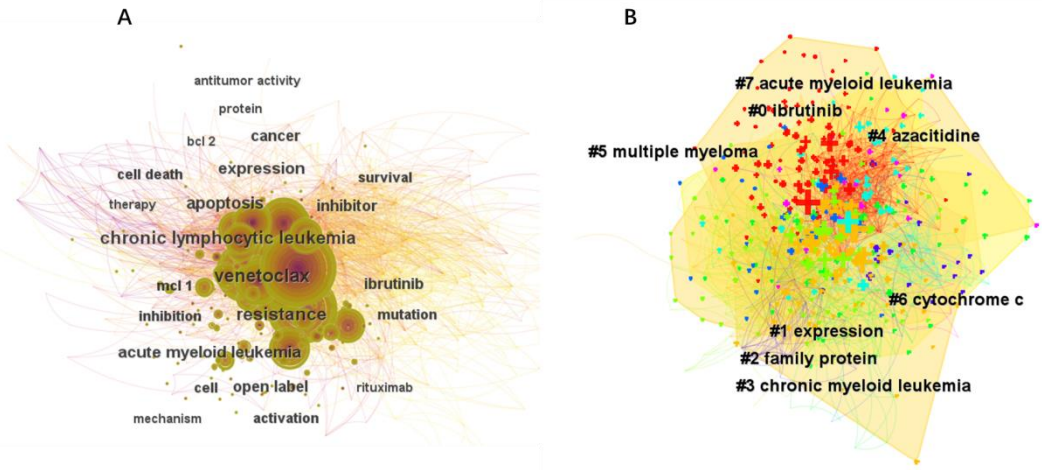


图 4. Venetoclax 耐药机制的关键词共现网络（A）和聚类分析（B）

2.5 引文共现分析

针对引用文献进行共引分析是 CiteSpace 最显著的特点。共被引网络聚类中包含的文献越多，通常表示该聚类字段的重要性就越大。本研究共提取了 699 篇共引文献作为节点绘制聚类图。图 5 显示了共被引文献的网络图（5A）以及聚类图（5B）。图谱中显示最大的 9 个标有#的聚类名称，这些通常代表的是 Venetoclax 耐药分子机制相关研究的核心热点，包括急性髓系白血病、多发性骨髓瘤、慢性淋巴细胞白血病、BTK 抑制剂、克隆动态、淋巴细胞白血病、套细胞淋巴瘤以及无进展生存期等。

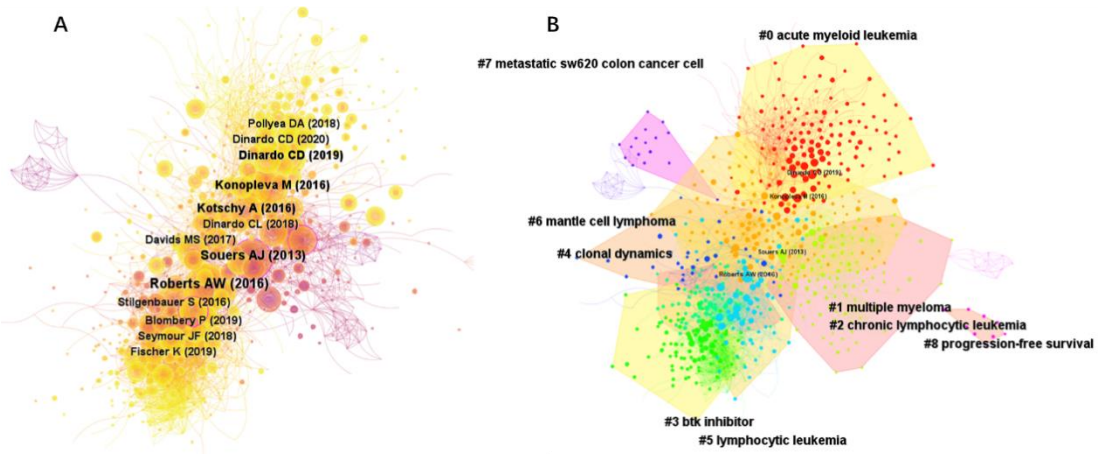


图 5. Venetoclax 耐药机制的共被引网络（A）和聚类分析（B）

表 3 则显示了共被引频次排在前 10 位的参考文献。共被引频次最高的是 Roberts AW 等发表在新英格兰医学杂志上的论文，探讨 Venetoclax 用于复发性慢性淋巴细胞白血病的疗效和机制[16]。Souers AJ 等则介绍了 ABT-99 作为 BCL-2 选择性抑制剂的药理学机制，这可能为治疗 BCL-2 依赖性血液学癌症提供前景[2]。Dinardo CD 等通过临床试验发现，将 venetoclax 与地西他滨或阿扎胞苷联合应用疗效明确，老年 AML 患者耐受性良好，不良事件较少[17]。Konopleva M 提供 AML 中的 IDH1/2 突变体显示 BCL-2 依赖性和线粒体功能障碍[6]。Kotschy 指出，靶向 MCL1 可以促进 BAX/bak 依赖性细胞凋亡，显示出 MCL1 抑制剂在抗肿瘤治疗中的价值[18]。Stilgenbauer S 评估了 Venetoclax 单药治疗复发或难治性 del（17p）慢性淋巴细胞白血病患者的疗效和安全性[19]。Dinardo CD 发现相比较于阿扎胞苷单独使用，与 Venetoclax 联合使用可提高反应率和总生存率，降低不良事件发生率[3]。Seymour JF 评估 Venetoclax 和利妥昔单抗联合使用，治疗复发或难治性慢性淋巴细胞白血病患者的疗效[20]。Blombery P 揭示了 BCL-2 中 Gly101Val 的获得性点突变导致耐药性产生的主要原因[21]。Dinardo CD 等对 Venetoclax 与低甲基化剂的组合进行了开放标签 1b 期研究[22]。这些高频次的共被引文献都是关于 Venetoclax 的一些重要的基础性和临床性研究，也是该领域的重点文献。

表 3. 共被引频次排名前 10 位的文献信息

Title	Author	Journal	Year	Citations
Targeting BCL2 with Venetoclax in relapsed chronic lymphocytic leukemia	Roberts AW	New Engl J Med	2016	192
ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets	Souers AJ	Nat Med	2013	148
Venetoclax combined with decitabine or azacitidine in treatment-naive， elderly patients with acute myeloid leukemia	Dinardo CD	Blood	2019	120
Efficacy and biological correlates of response in a phase ii study of Venetoclax monotherapy in patients with acute myelogenous leukemia	Konopleva M	Cancer Discov	2016	102
The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models	Kotschy A	Nature	2016	92
Venetoclax in relapsed or refractory chronic lymphocytic leukaemia with 17p deletion: a	Stilgenbauer S	Lancet Oncol	2016	89

multicentre, open-label, phase 2 study				
Azacitidine and Venetoclax in Previously Untreated Acute Myeloid Leukemia	Dinardo CD	New Engl J Med	2020	89
Venetoclax–Rituximab in Relapsed or Refractory Chronic Lymphocytic Leukemia	Seymour JF	New Engl J Med	2018	88
Acquisition of the Recurrent Gly101Val Mutation in BCL2 Confers Resistance to Venetoclax in Patients with Progressive Chronic Lymphocytic Leukemia	Blombery P	Cancer Discov	2019	85
Safety and preliminary efficacy of venetoclax with decitabine or azacitidine in elderly patients with previously untreated acute myeloid leukaemia: a non-randomised, open-label, phase 1b study	Dinardo CD	Lancet Oncol	2018	83

3.讨论

本研究基于 Venetoclax 耐药相关研究的 803 篇文献数据，以文献计量分析为主要研究方法，并通过 CiteSpace 绘制知识图谱以后，初步揭示了该领域研究现状、热点和趋势。我们发现，该领域研究论文数量从 2013 年开始就一直处于逐步递增状态，2016 年和 2019 年还出现引文突破形成高峰，说明自从该药被应用于临床以后，耐药问题就伴随而来。相关研究发现，以 Venetoclax 为代表的 BCL-2 抑制剂单药或联合去甲基化药物治疗具有较高的初反应率，但具体耐药机制尚不明确[1]。

过去十几年来，越来越多的研究人员开始研究 BCL-2 抑制剂的耐药相关机制，逐渐形成了一组以 Konopleva M、Roberts AW、Andreeff M、Dinardo C 和 Wang Y 等为代表的核心学者[23-25]。从研究者所属国家来看，欧美国家尤其是美国仍处于领先地位，例如，美国 MD Anderson 癌症中心和 Dana Farber 癌症研究所，仍是发文数量和影响力最大的两个研究机构。而中国虽然发文总量可以排在世界第二位，合作网络中心度也紧邻美国，但是无论从核心作者、核心机构，还是高影响力论文量角度来分析，与美国相比，仍有一定差距。

从研究热点来看，核心关键词包括 Venetoclax、耐药、急性髓系白血病、慢性淋巴细胞白血病、凋亡和伊布替尼等，而共被引频次最高的两篇论文分别是澳大利亚皇家墨尔本医院 Roberts AW 等所发表的探讨 Venetoclax 用于复发性慢性淋巴细胞白血病的疗效和机制的临床研究性论文[16]，以及美国艾伯维公司 Souers AJ 等最早于 2013 年报道的改良型 BCL-2 选择性抑制剂——ABT-199 用于 BCL-2 依赖性血液系统肿瘤的临床应用前景[2]。Venetoclax 作为选择性 BCL-2 抑制剂，临床治疗效果比较显著，但是作为抗癌药物，如何解决耐药性问题仍然是一个关键，将来可能需要更多的相关研究共同寻找作用机制以及破解方案。

4.结论

本研究发现美国在 Venetoclax 耐药研究领域仍处于主导地位，中国在发文数量上紧随其后，欠缺高质量论文，未来也需要与更多国家或机构开展更深入的合作。近年来，Venetoclax 针对白血病患者治疗作用，与伊布替尼、阿扎胞苷等药物联合应用，以及耐药分子机制探讨一直是研究的热点。本文从文献计量学的角度揭示了 Venetoclax 耐药领域的研究现状及趋势，以期对该领域相关研究机构和学者具有借鉴意义。

PU-105

基于本体的糖尿病知识图谱构建方法研究

满坚平，唐圣晟，杨瑞，周毅*
中山大学中山医学院

引言

糖尿病是一种常见的慢性代谢疾病，致残、致死率高，严重影响患者健康，造成沉重的社会经济负担。全球约有 5.37 亿糖尿病患者[错误!未找到引用源。]。我国是全球糖尿病患病率增长最快的国家之一，糖尿病患者超过 9700 万，糖尿病前期人群约 1.5 亿。目前，糖尿病的预防已经成为患者、临床医生、决策者的重要目标[错误!未找到引用源。]。

随着医疗信息化的发展，医院收集数据的能力大幅提升，医疗健康数据爆炸式增长。大量与糖尿病相关的知识呈碎片化分散在医学文献、百科类网站、研究报告等各类资源中。公众只能从不同的数据库进行信息检索和整合，对信息的挖掘和再利用带来极大不便。因此，迫切需要科学和有效的方法来组织和利用这些知识。

知识图谱（Knowledge Graph，KG）作为人工智能的重要组成部分，是信息处理和知识组织的重要手段[错误!未找到引用源。]。与传统的信息管理方法相比，知识图谱具有智能推理能力，能够可视化显示知识的推理过程，帮助我们更快速、更有效地获得知识之间的潜在关系。本文利用知识图谱，对糖尿病知识进行梳理和建模，可视化展示糖尿病的知识，促进公众对糖尿病知识的全面了解，增强自我主动健康意识，不断提高健康管理能力，推动全民健康。

一、医学知识图谱的概述

自 2012 年谷歌首次提出知识图谱，知识图谱一直是各行业研究的焦点[错误!未找到引用源。]。知识图谱是一种通过可视化图形来表示知识的结构化方法，用于描述现实世界中各种实体及其关系[错误!未找到引用源。]。其本质是一种语义网络的知识库，节点表示实体（Entity）或概念（Concept），边表示实体或概念之间的语义关系。知识图谱通常使用 RDF/RDFS/OWL 等网络知识表达语言来描述。本体是知识图谱的抽象化，知识图谱是本体的实例化。

医学知识图谱（Medical Knowledge Graph，MKG）结合知识图谱与医学知识，可以很好地表示现实中医学领域的实体（如疾病、症状、药物、基因等）及其关系和属性[错误!未找到引用源。]。其优势在于能够将大量的医学知识以结构化的方式表示，并提供交互式查询和推理功能。医学知识图谱能够帮助医生和研究人员更好地理解和利用医学知识，提高诊断准确性、加速药物研发过程，并为个性化医疗提供支持，也有助于促进医学领域的信息共享、协作和创新[错误!未找到引用源。]。

目前，国内外学者已经研究构建了针对不同疾病或医疗服务需求的医学知识图谱，如人类基因组知识图谱 GenomicKB [错误!未找到引用源。]、多模态精准医学知识图谱 PrimeKG [错误!未找到引用源。]、中文医学知识图谱 CMeKG [错误!未找到引用源。]、乳腺癌知识图谱 BCKG [错误!未找到引用源。]等，并在临床决策[错误!未找到引用源。]、智能问答[错误!未找到引用源。]、智能检索[错误!未找到引用源。]、文献可视化[错误!未找到引用源。]等方向取得了重大的进展。

二、医学知识图谱的构建方法

基于本体的知识图谱构建方法主要是自顶向下和自底向上。自上而下的方法首先定义本体（或数据模式），然后基于本体提取知识。自下而上的方法从数据中提取知识，并基于数据定义知识图谱的本体（如图 1 所示）。前者侧重于领域本体等知识模式，后者侧重于链接开放数据集等知识实例。通过知识建模、知识抽取、知识融合、知识存储、知识推理、质量评估等关键技术，实现医学知识图谱的构建。

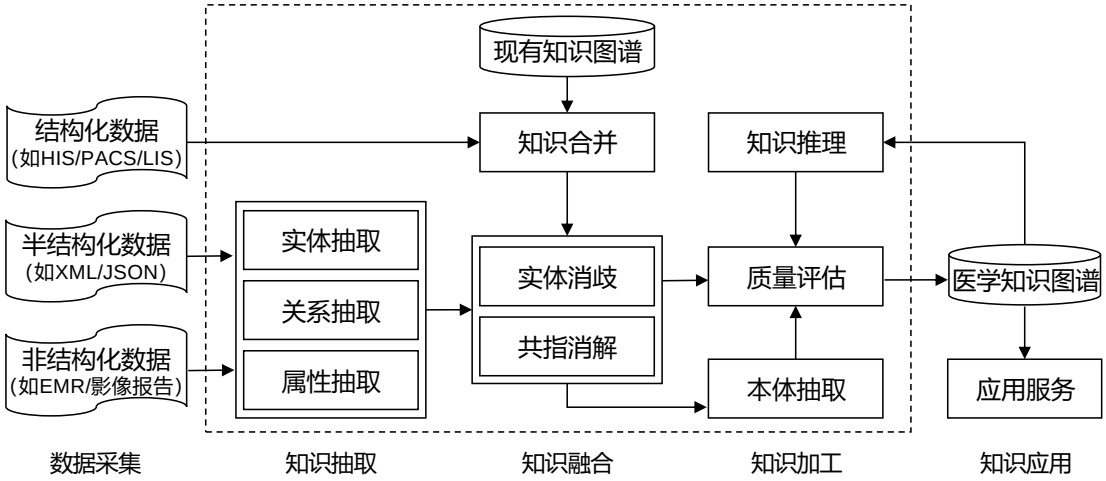


图 1 医学知识图谱的构建框架

（1）知识建模

知识建模是建立知识图谱概念模型的过程。为了合理地组织知识，更好地描述知识本身与知识之间的关系，需要根据实际应用的需求，进行知识图谱模式的定义[错误!未找到引用源。]。知识图谱包括模式层和数据层。模式层是构建知识映射的框架，涉及域中的概念、属性类型，以及各类之间的关系，多基于本体进行建模。数据层则是本体的实例。

（2）知识抽取

知识抽取是从非结构化或半结构化的异构医学资源中收集和提取结构化信息[错误!未找到引用源。]。知识抽取的数据源包括真实世界数据、医学文献、互联网资源、标准化术语和开放共享的医学知识库等，可以从中提取出医疗实体、属性以及实体间的相互关系，形成本体化的知识表达。其主要任务包括实体识别、关系抽取、属性补全、概念抽取、事件抽取等[错误!未找到引用源。]。

（3）知识融合

知识融合主要是将不同来源的关于同一实体或概念的描述信息连接起来，解决多源数据因命名不同、描述不清而无法关联分析的问题[错误!未找到引用源。]。在获得新知识之后，需要对其进行分析、验证、总结和集成，并消除矛盾和歧义，以扩大知识的存量。其主要任务包括实体对齐、实体链接、关系推演等。

（4）知识存储

知识存储是为已完成的知识图谱设计底层存储方式，实现各种类型知识的存储[错误!未找到引用源。]。知识图谱的存储方式主要包括基于关系型数据库的存储和基于图数据库的存储。常见的关系型数据库有 MySQL、Oracle 和 DB2 等，图数据库如 RDF4J、gStore、Stardog、Neo4j 等。考虑到实际应用中查询效率和响应效果，往往结合多种形式进行知识图谱的存储。

（5）知识推理

知识推理主要是基于知识图谱获取更多的隐性知识，同时也可以预测现有知识实体的相关性从而实现专业知识的深度挖掘和扩展[错误!未找到引用源。]。知识推理的模型主要包括基于规则的推理、基于分布式表示的推理和基于神经网络的推理。

（6）质量评估

知识图谱的质量评估不仅需要在知识图谱构建过程对知识进行评估，还需要在实际应用中根据特定场景的需要，对知识图谱的效用和性能进行度量，不断优化更新知识图谱[错误!未找到引用源。]。未来全面评估知识图谱的质量，可以从准确性、一致性、完整性、及时性、可信度和可用性六个维度开展研究。

三、糖尿病知识图谱的设计与构建

本文按照自顶向下的模式设计和构建糖尿病知识图谱，主要包括疾病实体，以及实体之间并发症、症状、诊断、治疗的关系。通过参考糖尿病相关文献、指南、词典和教材等，采用本体学习和手工编译的方法构建本体，从大量文本中学习词汇、概念和关系，并将其转化为本体形式，实现对糖尿病知识进行抽取和表示，形成结构化的糖尿病知识。

糖尿病知识图谱的构建包括对概念、实体、关系、属性的明确和定义。本文定义了疾病、症状、检查、药物、手术、部位等 6 个概念级别，并对已构建的糖尿病实体类进行数据属性定义。同时，本文还定义了 10 种关系类型，包括鉴别诊断、并发症、检查处理、疾病症状、用药、手术治疗、发病部位、拮抗作用、联合作用、手术部位（如表 1）。

表 1 糖尿病知识图谱的关系类型

关系类型	概念类	关系描述
鉴别诊断（diaDisease）	疾病-疾病	疾病诊断需要鉴别的其他疾病
并发症（hasComplication）	疾病-疾病	疾病的并发症
检查处理（needCheck）	疾病-检查	疾病诊断所需的检查检验项目
疾病症状（hasSymptom）	疾病-症状	疾病引发的症状
用药（needDrug）	疾病-药物	疾病治疗过程所使用的药物
手术治疗（needOperation）	疾病-手术	疾病治疗所需的手术
发病部位（causesOn）	疾病-部位	身体发生疾病的部位
拮抗作用（hasAntagonism）	药物-药物	药物间的拮抗作用
联合作用（hasCombination）	药物-药物	药物间的联合作用
手术部位（operatesOn）	手术-部位	手术过程所涉及的身体部位

四、糖尿病知识图谱的存储和可视化

Protégé软件是 Stanford 大学基于 Java 语言开发的本体编辑和知识获取软件，属于开放源代码软件，是语义网中本体构建的常用开发工具。本文使用 Protégé软件提供的图形界面创建和编辑概念和属性，构建了疾病、 症状、检查、药物、手术、部位六大概念类，并建立了概念间的层次结构（如图 2）。

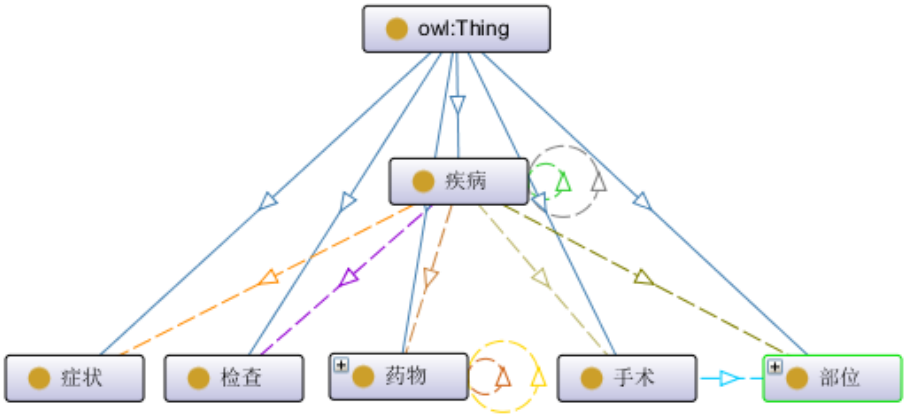


图 2 糖尿病知识图谱本体建模

构建概念是知识图谱建设的第一步，还需要添加属性、实例等来完善知识图谱的内容和功能。在 Protégé中构建了糖尿病的实体对象属性（Object property），为不同的概念之间创建关联关系，描述和表示概念之间的关联和连接（如图 3）。

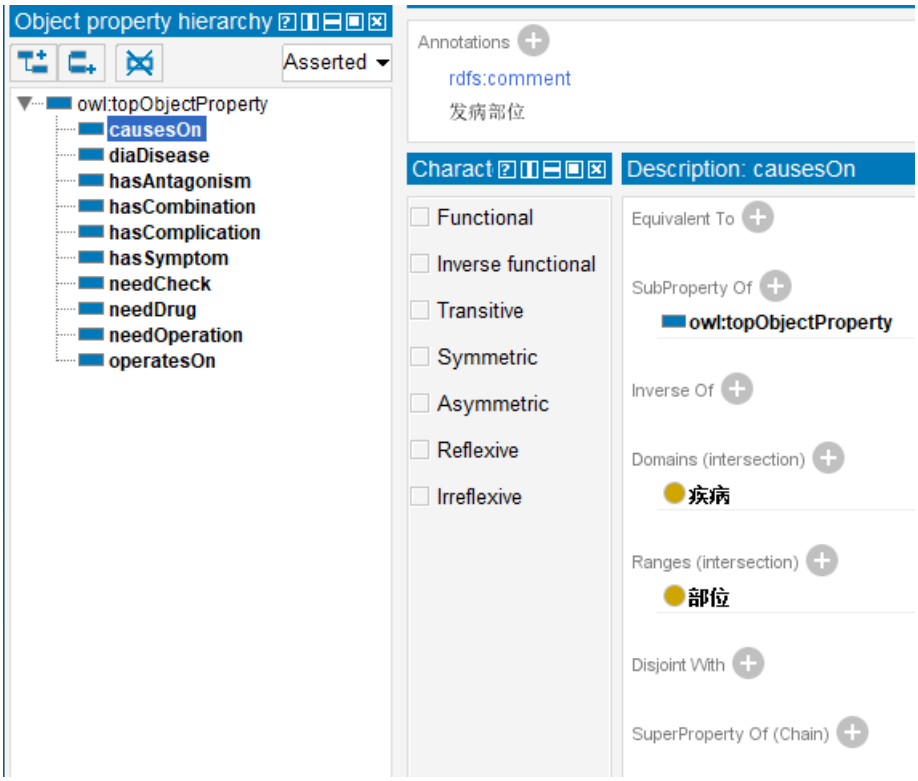


图 3 糖尿病知识图谱的实体对象属性建模

同时，构建和添加了糖尿病的实体数据属性（Data property），为不同的概念之间添加具体的属性值，描述和表示概念和实例的各种属性特征和值（如图 4）。

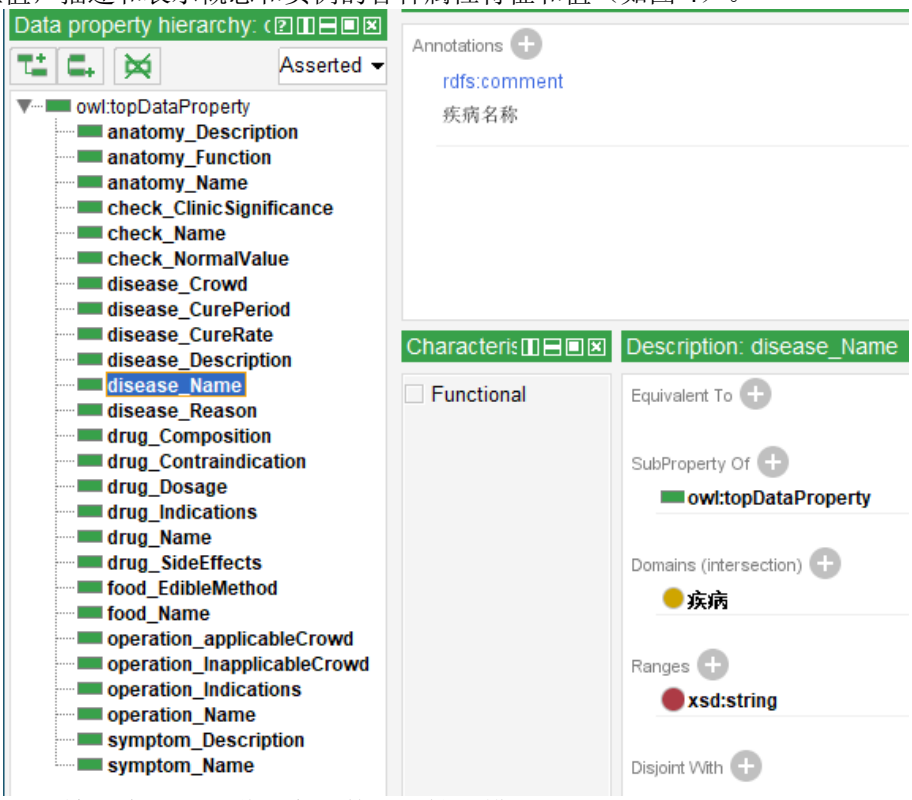


图 4 糖尿病知识图谱的实体数据属性建模

糖尿病知识图谱以图形化形式地展示糖尿病相关概念之间的关联（如图 5），并提供简要的概念信息。用户可以交互式浏览相关概念并选择其中一个概念，以进行进一步信息挖掘或检索。

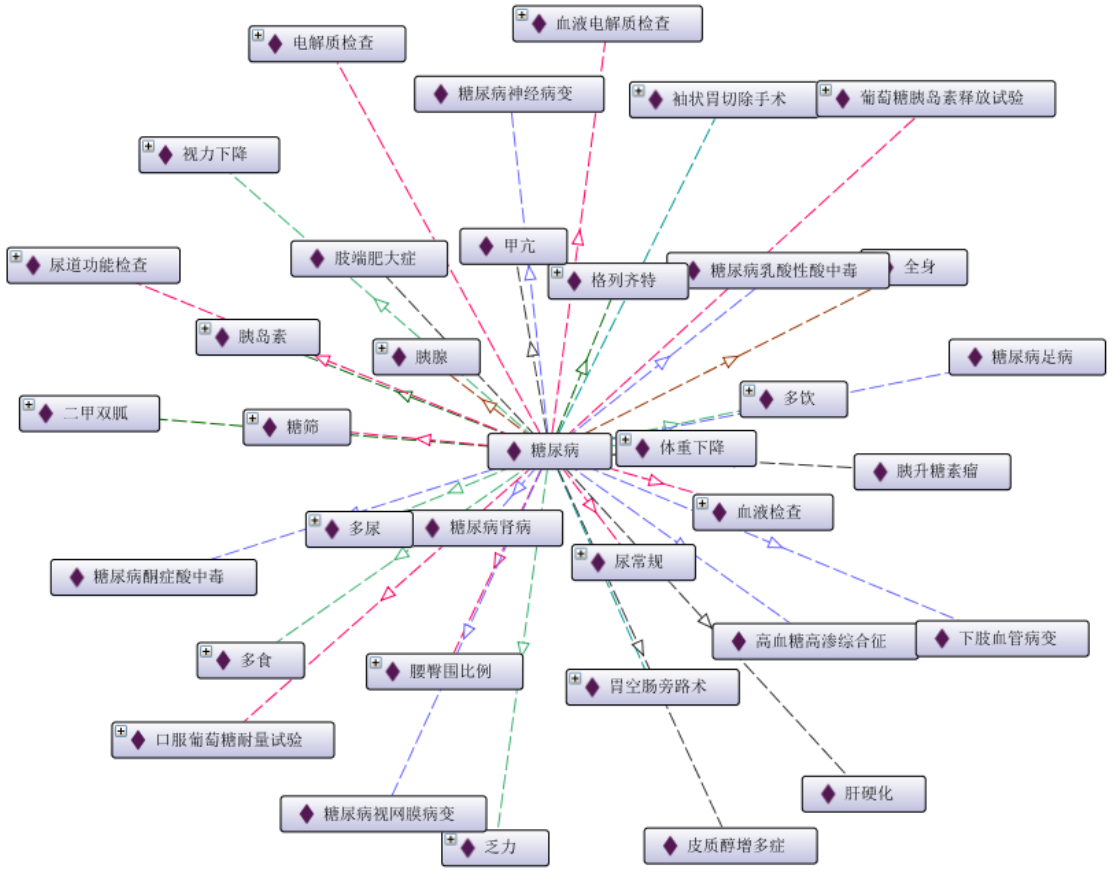


图 5 糖尿病知识图谱部分示例

五、总结与展望

本文通过对糖尿病知识的本体建模和知识抽取，利用 Protégé工具，构建了基于本体的糖尿病知识图谱，实现糖尿病知识的结构化表达和可视化展示。随着大语言模型的发展，医学知识图谱的研究将迎来新的发展动力。未来研究将利用大语言模型执行不同的知识图谱任务，丰富知识图谱的内容，增强对医学知识的语义理解和应用能力，提供更智能的推理和交互，为个性化临床决策和健康管理提供支持。

PU-106

基于疾病网络的血友病并发症挖掘与关联发现

邰杨芳，昝彭
山西医科大学

1. 引言

血友病是已知时间最长的疾病之一。尽管它的发病率相对较低，但仍然是一个严重的医疗和社会问题[1]。在临床医疗方面，血友病是一种遗传出血性疾病，是由凝血因子缺乏引起，血友病的主要特征为反复的自发性或外伤性出血，在没有预防措施的情况下，严重的血友病患者会出现反复的关节出血、最常见于踝关节、膝关节和肘关节[2]。由于血友病的特殊性质，使得血友病患者在临床治疗过程中极易引起其他疾病，或是血友病患者在进行其他疾病治疗过程中给患者带来更大就诊难度，如伴有特殊疾病的人工耳蜗植入术[3]、vWF-Thr789Ala 基因多态性与 2 型糖尿病的相关性[4]、血友病骨骼肌肉并发症功能评估及康复治疗[5]等，从上述研究中不难验证这一点。在社会生活方面，血友病患者从发病到后续治疗过程中，由于出血的长期反复性和由出血引起的身体损害会对患

者的心理造成严重影响，患者也会因为疾病给家人带来困扰而感到自责、惭愧，从而产生焦虑、抑郁的负面情绪[6]。

疾病的并发症是临床治疗时影响患者治疗康复的重要原因之一。临床并发症一般分为两类，一是一种疾病在发展过程中引起另一种疾病或症状的发生，后者即为前者的并发症；另一种并发症是在诊疗护理过程中，病人由患一种疾病合并发生了与这种疾病有关的另一种或几种疾病[7]。目前临床上对于并发症的预测多基于常规的疾病检查和医生的临床经验判断，多是在患者出现特殊症状后再进行诊断和治疗，这就导致在时效性上可能会有所延误，因此对疾病可能存在的潜在并发症的预测在临床诊疗中就显得尤为重要。

血友病作为一种遗传性疾病，患者常常需要长期接受治疗来控制出血和减轻疼痛。然而，在治疗的过程中，患者面临着并发症带来的沉重经济负担[8]。首先是血友病治疗中常用的凝血因子替代治疗剂价格昂贵，超出了大多数患者的经济承受力。其次，由于血友病患者具有易出血的特点，他们往往需要更频繁的医疗服务。此外，血友病治疗还需要一系列的康复措施，例如物理疗法和康复训练。这些措施旨在帮助患者恢复关节灵活性和肌肉力量，减轻疼痛和预防并发症。所以揭示血友病并发症间的关联规律有利于提前预防新的并发症的出现或及时采取措施，减轻患者痛苦，降低医疗费用负担。

2. 数据与方法

2.1. 数据获取与处理

（1）数据来源

在 PubMed 数据库中以"Hemophilia A"[Mesh] OR "Hemophilia B"[Mesh] OR "Factor XI Deficiency"[Mesh]构建主题检索式，并将 MeSH 词表中的入口词以“OR”的逻辑关系连接，以“[Title/Abstract]”为限制条件，构建高级检索表达式，将两个检索式以“OR”的逻辑关系连接，以保证检索文献的全面，并去掉其中的综述类文献，最后共检索到 16930 条文献，通过 NE 等软件导出文献的 PMID 号、题目及摘要以备后续研究。

（2）MetaMap 实体抽取

语义知识表达（Semantic Knowledge Representation，SKR）是基于一体化医学语言系统（Unified Medical Language System，UMLS）进行自然语言处理、提取文献中概念的系列在线工具[9]。其中 MetaMap 是 SKR 系统的核心组件，MetaMap 拥有 UMLS 的超级叙词表和专家词典的强大功能，因此抽取出的疾病实体拥有医学专业性，在 UMLS 的叙词表中对疾病这一类实体有专业的语义类型[dsyn](Disease or Syndrome)与之对应，因本研究主要围绕疾病的并发症展开，所以在通过 MetaMap 进行实体抽取的同时对实体类型以[dsyn]做为限定条件，最后得到结果如下图所示：

```
Processing '18652280.ti.1': [Secondary amyloidosis connected with neoplasma of urinary bladder and acquired
disorders of blood coagulation. Case report].

Phrase: [Secondary amyloidosis
>>>>> Phrase
secondary amyloidosis
<<<<< Phrase
>>>>> Mappings
Meta Mapping (1000):
  1000  AMYLOIDOSIS, SECONDARY (Reactive systemic amyloidosis) [Disease or Syndrome]
Meta Mapping (1000):
  1000  Secondary amyloidosis (AA amyloidosis) [Disease or Syndrome]
<<<<< Mappings
```

图 1MetaMap 疾病抽取结果示例（部分）

MetaMap 对原始文本实现逐句处理，其中“Processing "18652280.ti.1': [Secondary … Case report]”说明这些字段来源于编号为 18652280(PMID)的文献 ti(标题)字段的第 1 句内容中，每篇文献的语句经过 MetaMap 解析后所得的结果出现在 MetaMapping 后，例如“1000

AMYLOIDOSIS, SECONDARY (Reactive systemic amyloidosis) [Disease or Syndrome]”结果中数字“1000”代表 UMLS 中所使用的 Semantic Type 的编号。具体来说，“1000"对应的是"Disease or Syndrome”（疾病或综合征）这个语义类型，“AMYLOIDOSIS, SECONDARY”是罕见病次级淀粉样变性病在文献中的具体表达，“(Reactive systemic amyloidosis)”是对应疾病的规范化表述。

经上述步骤共抽取到 514 种疾病实体，其中出现频次前十的疾病如表 1 所示：

表 1 疾病命名实体抽取及频次结果（部分）

疾病名称	出现频次
Hemophilia A	1811
Hemophilia, NOS	1147
Blood Coagulation Disorders	305
Hepatitis C	277
Hemophilia B	239
HIV Infections	171
von Willebrand Disease, Type 1	146
Hereditary Factor XI Deficiency	137
Malnutrition	95
von Willebrand Disease, Type 3	84

（3）疾病网络构建

利用 python 程序对上述命名实体抽取得到的原始数据进行整理和结构化存储，一条记录代表一条文献，由文献 PMID 号和抽取出的疾病实体构成。并利用 COOC 数据分析软件生成共病矩阵，通过 UCINET、VOSviewer 等软件形成可视化网络图谱，根据可视化网络则可以探索疾病的并发症状况，具体构建流程如图 2 所示:

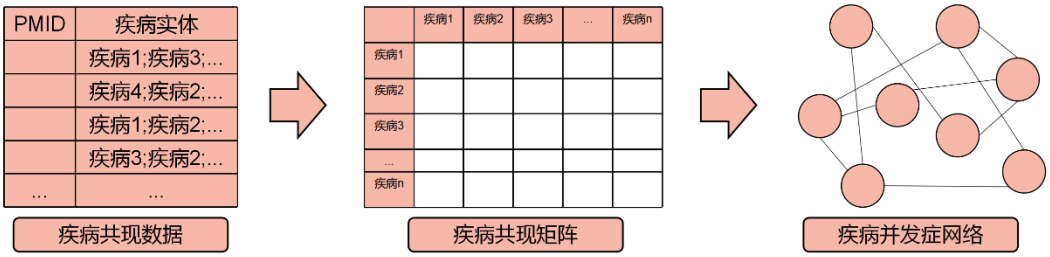


图 2 并发症疾病网络构建流程

2.2. 研究方法

（1）社会网络分析法

社会网络分析法是情报学研究的前沿手段，是对由多个实体之间相互联系所构成的社会关系结构及其属性进行分析的一套规范与方法的集合，是对行动者之间的关系以及由此形成的整个网络、结构进行量化研究。采用社会网络分析法能够从网络的个体属性、整体属性等方面进行分析，挖掘出更多深层次的信息。通过可视化技术直观展示关键词网络可增强研究问题的现实感，使人们更容易了解网络内部结构，揭示隐性知识。

（2）关联规则挖掘

关联规则挖掘是指在大规模数据集中寻找频繁项集并推导出项集之间的关联规则的一种方法。其特点是能够挖掘数据集中的隐藏关系，从而支持商业决策和市场分析。常用的关联规则分析算法包括 Apriori、FP-Growth 等。

Apriori 算法，作为关联规则分析的经典算法。关联规则挖掘作为数据挖掘的重要项目，受到了广泛的关注，并在体育、商务、金融、医疗等领域得到了广泛的应用[10]。

关联规则算法包括两个步骤:首先列出集合中所有高频项;然后根据高频项生成频繁关联规则。高频表示某一项的词频已达到或超过一定水平，该词频为支持度 (Support)。定义如下：

$$Support(A => B) = P(A \cap B) \quad (1)$$

当{A, B}的支持度大于或等于最小支持度时，则将{A, B}放入高频项组。

关联算法的第二步是关联规则的生成。根据第一步得到的高频项组，如果某条规则满足最小置信度，则该规则为关联规则。置信度(Confidence)定义如下:

$$Confidence(A => B) = P(B|A) \quad (2)$$

除支持度和置信度以外，提升度是常用的评价关联规则的指标，为了寻找疾病之间的潜在关联，将提升度也作为确定疾病关联的相关系数。

提升度(Lift) 被定义为:如果$P(A \cap B) = P(A) \cdot P(B)$，则项集 a 的出现与项集 B 无关;反之，项集 A 和项集 B 是相互依存、相互关联的。A、B 的升力可由下式计算:

$$Lift(A,B) = \frac{P(A \cap B)}{P(A)} \cdot P(B) \quad (3)$$

如果公式(3)的结果小于 1，那么 A 和 B 的出现是负相关的，即当一个出现时，另一个很可能不出现。如果公式的结果大于 1，那么 A 和 B 是正相关的，也就是说，如果有一个存在，那么另一个很可能存在。如果公式的结果等于 1，则 A 和 B 是独立的，两者之间不存在相关性。

（3）链路预测

链路预测拟解决的关键问题是根据观测到的链路预测给定网络中两个节点间存在链路的可能性[11]，其特点是在缺乏传感器观测数据和无法准确建模的情况下，利用网络拓扑结构代替实际数据进行预测。因此，链路预测方法可以应用于社交网络、生物网络、交通网络等领域。

（a）链路预测指标选取

为了预测和发现患者可能存在的并发症，使用链路预测技术进行更进一步的分析研究，本研究选取了分别基于节点和路径的 12 种指标，通过基于相似性的多项指标分析含权和不含权网络，并将数据分成训练集和测试集，使用训练集计算精度，选择最佳指标用于计算未链接节点之间出现链接的可能性[12]。具体指标及其计算公式如表 2 所示：

表 2 链路预测相似性指标

指标	基于节点局部信息的相似性指标 含权算法	不含权算法
AA	$S_y^x = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{2\lg(1+S_z)}$	$S_{xy} = \sum_{z \in (x) \cap \Gamma(y)} \frac{1}{\lg k(z)}$
PA	$S_{xy}^{PA} = \sum_{x' \in (x',x)} w(x',x) \sum_{y' \in (y',y)} w(y',y)$	$S_{xy} = k(x)k(y)$
RA	$S_{xy}^{RA} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{yz}}{S_z}$	$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k(z)}$
CN	$S_{xy}^{CN} = \sum_{z \in \Gamma(X) \cap \Gamma(y)} W_{xz} + W_{zy}$	$S_{xy} = \lvert T(x) \cap T(y) \rvert$
Salton	$S_{xy}^{salton} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{\sqrt{s(x)s(y)}}$	$S_{xy} = \frac{\lvert T(x) \cap T(y) \rvert}{\sqrt{k(x)k(y)}}$
Jaccard	$S_{xy}^{Ja} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{s(x) + s(y) - w(x,y)}$	$S_{xy} = \frac{\lvert T(x) \cap T(y) \rvert}{\lvert T(x) \cup T(y) \rvert}$
Sorenson	$S_{xy}^{so} = 2 \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{s(x) + s(y)} \cdot \frac{1}{2}$	$S_{xy} = \frac{2\lvert T(x) \cap T(y) \rvert}{k(x) + k(y)}$
HPI	$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{\min[s(x), s(y)]}$	$S_{xy} = \frac{\lvert T(x) \cap T(y) \rvert}{\min[k(x), k(y)]}$
HDI	$S_{xy} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{\max[s(x), s(y)]}$	$S_{xy} = \frac{\lvert T(x) \cap T(y) \rvert}{\max[k(x), k(y)]}$
LHN	$S_{xy}^{LHN} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{W_{xz} + W_{zy}}{s(x)s(y)}$	$S_{xy} = \frac{\lvert T(x) \cap T(y) \rvert}{k(x)k(y)}$
指标	基于路径的相似性指标	

LP	$S = A^2 + \partial A^3$	$S_{xy}^{LP} = \sum_{l_{xy}^n \in L(x,y)} l_{xy}^2 \partial l_{xy}^3$
Katz	$S = (I - \partial A^{-1}) - I$	$S = (I - \partial A)^{-1} - I$

（b）评价指标选取

应用上述指标对网络进行测度，可以计算出两两节点间的相似度，每对节点被赋予一个分数值，不同指标计算结果不同，为此还需评估指标的测度准确性[13]。本研究选取常用的评估指标 AUC 作为评价标准，AUC 测度值在 0 到 1 之间，值大于 0.5 表明所提算法优于二进制类的随机预测。在链路预测方面，AUC 是指随机选择的实际边缘得分的概率高于随机选择的不存在的边缘得分的概率[14]。AUC 的计算方法如公式(4)所示：

$$AUC = \frac{n' + 0.5n''}{n} \tag{4}$$

其中 n 为独立比较的次数，n'表示实际的边的分数比不存在的边的分数高多少倍，n''表示实际的和不存在的边的分数相等的次数。将 AUC 值最大指标的计算结果作为疾病间相似度值。

3. 结果与分析

3.1. 基于疾病网络的社会网络分析

本研究是针对疾病并发症网络进行社会网络分析，因此选择连线≥4 的节点，利用 VOSviewer[15]软件来绘制网络图，结果如图 3 所示。分别选用 Gephi、UCINET 等社会网络分析软件分别从网络密度、网络中心性、网络结构特征三方面展开分析，利用关联强度算法对疾病网络进行聚类，由 VOSviewer 软件实现可视化。由于疾病信息来源于文本的摘要和题目，所以出现在同一篇文献的不同疾病属于共现关系，并无明确指向性，因此导入数据后生成节点数为 206 和 21115 条边的无向图。

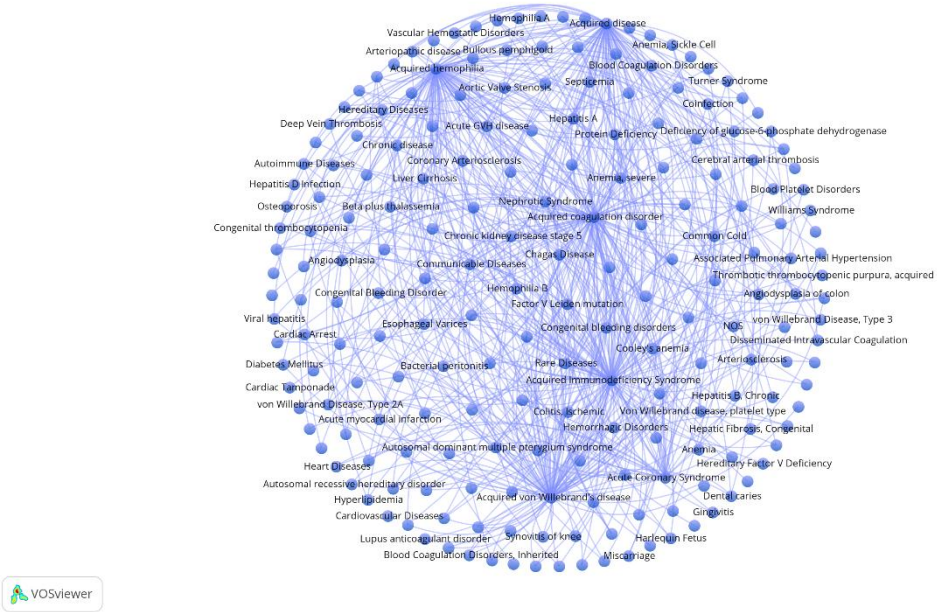


图 3 并发症疾病网络

（1）网络密度分析

运用 Gephi 软件运行分析出的网络密度结果如表 3 所示。根据运行结果可做如下分析，网络的平均度为 181.496，表明每个疾病大约与 181 个其他疾病产生联系，网络直径是指连接网络中任意

两个最远的节点所需要的边数。值越小，说明两者之间更容易建立起连接[15]，该网络直径为 3，这表明网络中各疾病之间建立联系相对容易。图密度表示各疾病之间的紧密程度，值越大则表示联系越紧密，本研究中的网络密度达到了 0.741，表明各疾病实体存在较为紧密的联系，生成的疾病关联规则可能较多，网络的整体稳定性较强。

表 3 网络密度分析结果

指标	数值
平均度	181.496
网络直径	3
图密度	0.741

（2）网络中心性分析

(a)度数中心性分析

通过 UCINET 软件得到疾病网络中度数中心性排名前十的数据如表 4 所示。从下表可以看出，Hemophilia A(血友病 A)是血友病众多类型中被研究最为广泛也是最常见的血友病类型，同时从表中也可以看出，目前血友病的相关研究中多伴随传染性和遗传性疾病一同出现。

表 4 度数中心性排名结果

排名	疾病	度数	排名	疾病	度数
1	Hemophilia A	46.38	6	Communicable Diseases	36.81
2	Hemophilia, NOS	44.09	7	Hepatitis C	36.61
3	Blood Coagulation Disorders	42.27	8	Hemophilia B	36.35
4	Malnutrition	37.27	9	Acquired hemophilia	36.03
5	HIV Infections	36.94	10	Hereditary Factor XI Deficiency	35.91

(b)中介中心性分析

上述疾病网络中介中心性排名前 10 得结果如表 5 所示，结果显示 Hemophilia A（血友病 A）、Blood Coagulation Disorders（血液凝固异常）、Malnutrition（营养不良）、HIV Infections（HIV 感染）等十个疾病中介中心性较高，其中一部分如 Hemophilia A（血友病 A）、Blood Coagulation Disorders（血液凝固异常）等属于血液系统疾病，另一部分疾病具有传染性，还有极少部分疾病直接与患者日常生活息息相关。

表 5 中介中心性排名结果

排名	疾病	中介中心性
1	Hemophilia A	189.000
2	Hemophilia, NOS	178.167
3	Blood Coagulation Disorders	167.500
4	Malnutrition	136.000
5	HIV Infections	135.500
6	Communicable Diseases	133.167
7	Hepatitis C	132.000
8	Hemophilia B	129.333
9	Acquired hemophilia	126.500
10	Thrombocytopenia	124.167

（3）网络结构特征分析

(a)小世界理论分析

小世界网络具有以下两个基本特征：①小的平均路径(Average Distance) ②大的聚类系数(Clustering Coefficient)[16]，运用 UCINET 软件分析得出，此疾病网络的平均路径为 2.027，聚类系数为 18.953，符合复杂网络理论的“小世界现象”特征。通过上述数据可以得出每个疾病在上述网络中只要平均通过 2.027 个疾病实体就能与其他疾病构建联系，说明此疾病网络的传播速度快且传播效率高，平均聚类系数为 18.953，说明网络中两个疾病之间可能存在联系的概率为 0.995，疾病之间的凝聚力较强。

(b)凝聚子群分析

应用 VOSviewer 软件对上述网络进行聚类分析，可视化结果如(图 4)所示，共形成了 4 个凝聚子群，不同子群有不同的颜色表示且度数越高，节点越大，节点之间联系越强，其连接线条越粗。

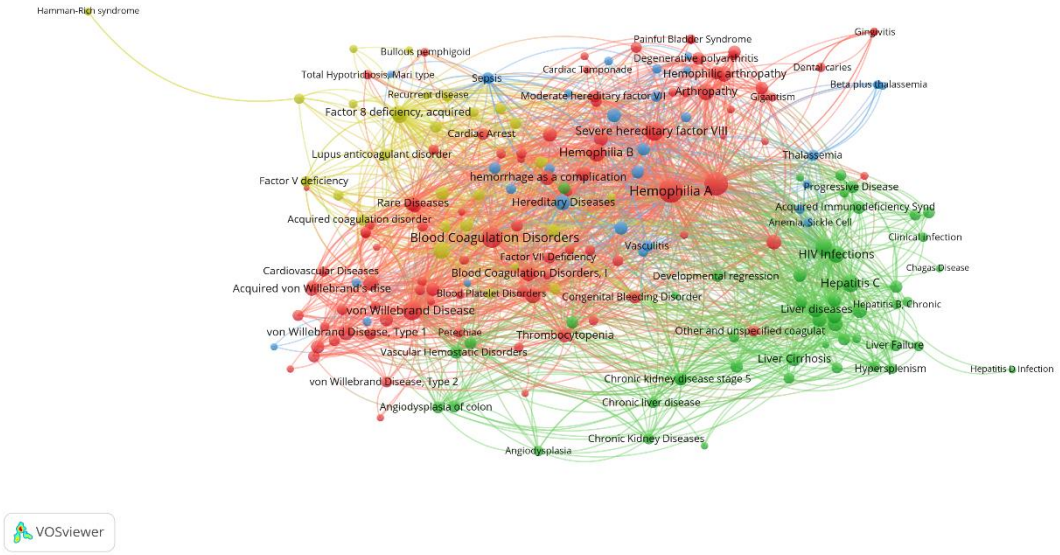


图 4 关联强度聚类结果

凝聚子群分析反映了子群的聚合层次，其中，子群 1 用红色表示，主要节点为 Acquired Hematologic Disorders（获得性血液系统疾病）、Acute posthemorrhagic anemia（急性出血性贫血）、Hemophilia A（血友病 A）、Hemophilia B（血友病 B）等 94 个疾病实体，Hemophilia A 作为整个网络的中心性排名最大的节点，同时也在子群 1 中连接着大多数其他疾病节点，这些疾病多数为出血性疾病，在发病原因上分为先天性出血，和后天性出血，在临床上表现为急性出血或慢性出血，同时也表明血友病 A 的研究是目前临床血友病研究的主要方向，说明在进行血友病 A 的研究时会与大多数其他出血性疾病会进行对比分析和诊断。

子群 2 用绿色表示，包括 Blood-borne infections（血液传播感染）、HIV Infections（Human Immunodeficiency Virus，人类免疫缺陷病毒感染）、Hepatitis B（乙型肝炎）、Hepatitis C（丙型肝炎）等 50 个疾病实体，这一子群的疾病大多具有传染性，与一般的出血性疾病不同，但是它们的传播渠道大多可以依靠血液进行传播。同时在现有医疗环境下和研究中与其他传染性疾病相比，血友病可能并未被列为医疗保健重点[17]，但血友病对于患者的伤害同样不能被忽视。

子群 3 用蓝色表示，主要包括 AA amyloidosis（二型淀粉样变）、Abdominal infection（腹部感染）、Atherosclerosis（动脉粥样硬化）、Ankle ankylosis（踝关节强直）、Chronic anemia（慢性贫血）等 32 个疾病实体，从上述疾病中不难看出，子群 1 中的疾病大多属于慢性疾病，这很可能与血友病患者长期的炎症反应、血液凝块异常、血管损伤和使用凝血因子制品等因素有关。

子群 4 用黄色表示，包括 Hypomenorrhea（月经过少）、Marfan Syndrome（马凡综合征）、Dyslipidemias（血脂异常）、Miscarriage（自然流产）、Malnutrition（营养不良）等 30 个节点，子群 3 中的节点大多代表机体自然生理过程和代谢异常现象等，这类疾病大多与家族遗传和

日常生活息息相关，这与血友病的遗传特征相似，且血友病也会伴随患者终身，影响患者生活质量，所以在血友病得相关研究中会大量出现。

3.2. 基于疾病网络的关联规则分析

通过 2.2.(2)节的关联规则指标分析，在本研究中，设置最小置信度≥0.8、最小支持度>0.01、最小提升度≥1，根据上述计算指标，总计生成 133 条关联规则，关联规则的定义存在三种关联类型，即基础关系、进阶关系和同级关系[18]，根据关联计算结果，在得出的所有关联规则中，有以下 3 种情况：

（1）若疾病对{A，B}只存在一条关联规则，即A→B，置信度为x，说明疾病A影响疾病B的出现，因此定义A为B的基础疾病，B为A的进阶疾病，即在发生疾病A的基础上，更倾向于产生疾病B。

（2）若疾病对{A，B}存在两条关联规则，即存在 A→B，置信度为 x，又存在 B→A，置信度为 y，且 x>y，则说明疾病 A 对疾病 B 出现的影响大于疾病 B 对疾病 A 出现的影响，因此舍弃 B→A 这条规则，定义 A 为 B 的基础疾病，B 为 A 的进阶疾病；反之若 x<y，则 B 为 A 的基础疾病。

（3）若疾病对{A，B}存在两条关联规则，即存在 A→B，置信度为 x，又存在 B→A，置信度为 y，且 x=y，则说明疾病 A 与疾病 B 具有同等影响，因此合并两条规则，定义 A 和 B 为同级疾病，即疾病 A 与疾病 B 可能同时出现。

根据上述规则，对上述的 133 条关联规则进行整理，其中同级关系有 12 组疾病对，如表 6 所示：

表 6 关联规则同级关系对

同级关系对	置信度
‘Hemophilia, NOS', 'Arthropathy'→'Hemophilia A'	0.96
‘Hemophilia，NOS', 'Hemophilia B', 'Blood Coagulation Disorders'→'Hemophilia A'	0.92
‘Liver diseases', 'Hepatobiliary Disorder'→'Hemophilia A'	0.91
‘Hepatobiliary Disorder', 'Hemophilia A'→'Hemophilia, NOS', 'Liver diseases'	0.90
‘Hemophilia，NOS', 'Hepatitis C，Chronic'→'Hemophilia A'	0.86
‘Hemophilia，NOS', 'Hepatitis A'→'Hepatitis C', 'Hemophilia A'	0.86
‘Liver Cirrhosis', 'Hemophilia A'→'Hemophilia，NOS', 'Hepatitis C'	0.86
‘ Liver diseases', 'Hepatobiliary Disorder' → 'Hemophilia, NOS', 'Hemophilia A'	0.82
‘HIV Infections', 'Hemophilia A'→'Hepatitis C'	0.80
‘Hemophilia，NOS', 'Liver diseases', 'Communicable Diseases'→ 'Hepatitis C'	0.80
‘ Hepatobiliary Disorder', 'Hepatitis C' → 'Hemophilia ， NOS', 'Hemophilia A'	0.80
‘ Liver diseases', 'Hepatobiliary Disorder', 'Hepatitis C' → 'Hemophilia，NOS',	0.80

同级关系中 Arthropathy（关节性病变）与 Hemophilia A（血友病 A）的研究属于临床常见研究通过国内范志娜[19]等和国外 Gualtierotti, R[20]等诸多学者的研究中不难验证，在其他同级关系中常见的血友病与其他血液疾病的共同研究也都有据可查，如 Blood Coagulation Disorders（血液凝结障碍）与 Hemophilia A（血友病 A）在 Johnsen, Jill M[21]等人的研究中都有具体表现，同时从上述关系对中不难看出与血友病临床研究交集最多的就是传染性疾病，比如在 Gharoonpour, A [22]等人的研究中就验证了 HIV Infections，Hemophilia A→Hepatitis C 这一研究所得同级关系对，Zabaleta, N[23]等人的研究也都出现了 Liver diseases（肝脏疾病）与 Hemophilia A（血友病 A）。

除了上述同级关系得到考证外，剩余的 53 条“基础→进阶“关系，也都是有据可查的，所以综上所述，关联规则分析得到的疾病对都具有一定的参考价值。

3.3. 基于疾病网络的链路预测

根据 2.2.(3)节中提到的链路预测各项指标，分别计算其 AUC 值，其对比结果如图 5 所示：

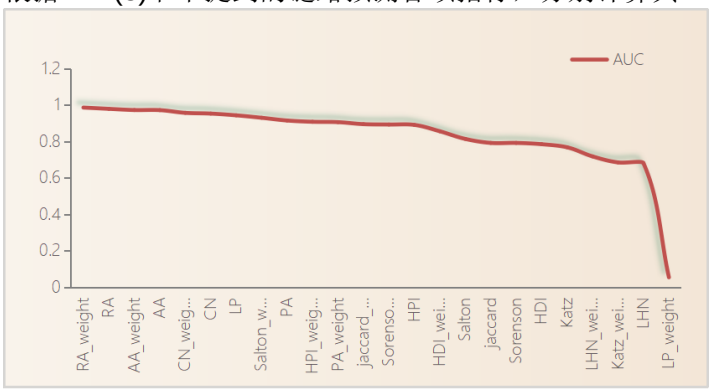


图 5 各指标 AUC 值对比(含加权)

（备注：Katz 指标，含权算法参数取值 0.05，不含权算法参数取值 0.01）

实验发现，AA、PA、RA、CN、Salton 五种链路预测指标在含权和不含权的两种网络中均表现出较好的预测效果，平均精准度高达 80%以上，然而 LP、Katz、LHN 指标在两种网络下，AUC 精度值都并不理想。

对于含权网络和不含权网络，综合各指标 AUC 均值的结果，发现含权网络在链路预测方面表现更好。在含权网络中，RA 的效果最佳。因此，本文采用含权的 RA 算法进行链路预测，以预测该网络节点未来产生关联的机会与可能性。

结合上文实验结果，拟选取 RA 预测指标值排名前十的疾病对进行分析，如表 7 所示。疾病对之间得分越高，两个疾病间产生关联的可能性越大。

表 7 链路预测得分结果(部分)

并发症预测	得分
Angiodysplasia of colon→Hemophilia, NOS	0.986940 82
Arthropathy→Hepatitis C	0.978281 19
Blood Coagulation Disorders→hereditary coagulopathy	0.957669 59
Ischemic stroke→Myocardial Infarction	0.952820 14
Hepatitis C→Progressive HIV Infections	0.932637 08
Hemophilia, NOS→Thrombophilia	0.927152 33
Ascites→Hepatobiliary Disorder	0.922510 46
Premature osteoarthritis→Prostatic Hypertrophy	0.912257 83
Hemophilia A→Vitamin K Deficiency	0.911437 59
Hemophilia A→Hyperammonemia	0.903046 55

根据表 7 可以看出，临床上的部分不同类型的血友病多伴随着其他的血液系统疾病或血管性疾病及部分凝血功能障碍等症状并发，Isabela Siloşi [24]等人的研究已经证实 Arthropathy（关节病变）和 Hepatitis C（丙型肝炎）的并发性关系，Ascites（腹水）和 Hepatobiliary Disorder（肝胆疾病）的并发也有多篇相关文献研究[25]，Ischemic stroke（缺血性卒中）和 Myocardial Infarction（心肌梗死）两种疾病的并发关系在临床治疗和研究中较为常见[26]，但是在血友病研究领域还未发现具体相关研究，Premature osteoarthritis（早发性骨关节炎）和 Prostatic Hypertrophy（前列腺增生）两种疾病在临床上虽然并无直接的相关性，但是在链路预测结果中得分却很高，经查阅相关资料发现，两种疾病虽病理上无相关性，但是从遗传因素角度分析，两种疾病多发于有家族疾病史的家庭中，若家族中有人患病，那么两种疾病的发病率都会显著提升。

经预测的大部分结果在现有文献数据库中能找到相应的研究成果，可证明利用链路预测的方式对临床患者的并发症预测在临床中切实可行。

4. 结论

并发症对血友病患者得治疗和康复有着非常明显的影响，一些潜在的并发症如果不进行及时发现，轻者会影响患者的日常活动，对患者的关节和神经系统造成损伤，严重时甚至会造成反复出血和神经系统受损，直接导致关节畸形、运动障碍、癫痫及瘫痪等症状。本研究首先通过社会网络分析得到在血友病研究领域主要的疾病及并发症，同时在整体结构上该网络满足小世界网络特征，并根据凝聚子群特征，该疾病网络分为4大聚类，并联合关联规则挖掘探究了疾病并发症之间的一般性规律，且通过链路预测的方式，发现了部分在现有血友病相关研究文献中未体现的疾病潜在的并发症，并进一步通过文献、发病机制以及外部因素等分析进行验证，证实了这一路线的可行性，为患者在住院过程中可能出现的并发症的预防和治疗提供了一定的参考价值。

本研究的不足之处在于，受到学科专业的限制，并未得到临床病例信息和医学实验对现有预测结果进行验证，但本研究已通过现有开放资料验证了部分预测结果的准确性，采用社会网络分析的方法仅从疾病网络的宏观静态结构进行分析，未能考虑网络的时态因素。

PU-107

基于知识发现工具 BITOLA 与 Arrowsmith 探索叶酸辅助治疗自闭症的潜在机制

钟小云 王康龙 吕艳华
山西医科大学管理学院

1 引言

随着社会的发展，科学技术的突飞猛进，专业分工越来越详细，跨学科的交流越来越困难，而一个专业领域的信息可能对另一个专业领域有价值，存在着一些隐秘和有价值的相关性，并且常规方式无法检索到这种相关性[1]。1986 年美国芝加哥大学 Swanson 教授提出了一种基于非相关文献的知识发现方法，并在 1991 年与他的合作者 Neil Smalheiser 开发了第一个非相关文献知识发现系统—Arrowsmith，依据闭合式知识发现过程思想开发，利用 Arrowsmith 可以对科学假设进行验证，为研究人员提供新颖并很可能真实的新知识或潜在关系[2]。BITOLA 系统是由斯洛文尼亚的生物统计与医学信息研究所的 Dimitar Hristovski 与 Borut Peterlin 研制开发，是一个基于文献的交互式生物医学发现支持系统，有闭合式发现系统和开放式发现系统两种模式。闭合式发现系统允许输入两个概念，并生成对两个实体之间关系的潜在解释，通过搜索已发表的文献来找到中间链接。开放式发现系统允许输入单个概念，然后是该概念的一阶亲属的类别，然后是那些一阶概念的亲属的类别。因此，它可以从疾病链接到相关药物，然后链接到与这些药物相关的基因，然后测试这些基因是否已被提及/测试与疾病相关，如果答案是否定的，那么该基因可能相关但尚未在文献中进行测试[3]。 自闭症，也称孤独症，是一种广泛性发展障碍，被归类为一种由于神经系统失调导致的发育障碍，以严重的、广泛的社会相互影响和沟通技能的损害以及刻板的行为、兴趣和活动为特征的精神疾病，是精神卫生领域的一大难题[4]。2006 年自闭症被列入我国的残疾人目录，在国内 ASD 患病率的最新报告指出中国目前的患病率与西方国家接近，学龄期约为 1%，目前在中国有超过一千万人罹患 ASD，ASD 个体大都精神残疾且生活不能自理，该疾病导致的家庭、社会及医疗费用非常高昂[5]。目前没有找到自闭症的确切病因，一般认为与大脑组织结构和神经发育异常、遗传因素及环境因素有关，而叶酸是一种水溶性 B 族维生素，具有促进神经系统、神经元生长发育和避免神经管缺陷的功效，故本文通过知识发现工具 BITOLA 和 Arrowsmith 探索研究叶酸辅助治疗自闭症的潜在机制。

2 资料与方法

2.1 数据来源

BITOLA 系统是基于 Medline 数据库进行开发的，而 Arrowsmith 系统是基于 PubMed 数据库开发，因此本文所挖掘到的数据资源来源于 Medline 和 Pubmed 数据库。

2.2 检索策略

登录 BITOLA 网站 https://ibmi.mf.uni-lj.si/bitola，首先点击 Open discovery system 进入开放式发现系统检索界面，在 Concept 内输入“Autistic disorder”为起始概念，点击“Find Starting Concept”，点击“C0004352”的编号，此概念属于的语义类型为“Mental or Behavioral Dysfunction”。然后语义组设为“Any”，语义类型设为“Gene or Gene Product”，其它保持默认，点击“Find Related Ys”，根据出现的频次降序排列检索出的相关概念 Y，筛选出的概念 Y 列表是符合“X→Y”关联规则的相关概念，得到 340 个 Y 概念集合。对概念 Y 列表进行语义筛选，勾选频次≥10 的概念 Y，语义组设为“Any”，语义类型设为“Pharmacologic Substance”，其它保持默认，再点击“Find Related Zs”，根据出现的频次降序排列检索出的相关概念 Z。本系统根据是频次(frequency)或可信度(confidence)对挖掘到的概念进行排名，频次是指 PubMed 中 X 和 Y(或 Y 和 Z)共现的文献数，可信度是指所有含有 X 的文献中含有 Y(或所有含有 Y 的文献中含有 Z)的百分比[6]。

3 结果

3.1 BITOLA 挖掘的候选药物

表 1 为与 X 相关的 Y 概念(共现频次≥10)。选定所有共现频次≥10 的 8 个 Y 概念之后，根据选定的 Y 寻找相关的 Z，得到 Z 概念列表，列表有 538 个 Z 概念。通过查阅资料分析 Z 概念列表的前 100 个概念，去除无关概念，通过对比筛选，发现 3 种维生素，有生物素（维生素 B7）、叶酸（维生素 B9）、维生素 B12 ，以及维生素 A 的代谢中间产物维甲酸，故猜测维生素可能有辅助治疗自闭症的潜在机制。所以更换 Z 概念限定语义类型为“Vitamin”，检索得到 Z 概念列表，有 13 个如表 2，从表 2 可以看出，根据频次降序排序后，Biotin、Folic Acid、Vitamin B12 位列前三位，生物素的频次对比其他维生素频次太高，研究意义不大，所以选择排在第二位的 Folic Acid 为研究目标。

表 1 BITOLA 中设定语义类型为 Gene or Gene Product 挖掘出频次≥10 的 Y 概念

序号	概念 Y	频次	置信度（%）
1	SH3PXD2A	26	0.322
2	DNTT	26	0.322
3	GABRB3	16	0.198
4	FMR1	16	0.198
5	MECP2	14	0.173
6	UBE3A	12	0.148
7	CARS	11	0.136
8	MLS	10	0.124

表 2 BITOLA 中设定语义类型为 Vitamin 挖掘出的 Z 概念

序号	概念 Z	频次	置信度（%）
1	Biotin	216	3.401
2	Folic Acid	48	7.582
3	Vitamin B12	40	11.762
4	Acetylcarnitine	24	2.784
5	Niacinamide	16	1.320
6	Leucovorin	16	3.562
7	Tetrahydrofolates	8	3.478
8	Retinaldehyde）	8	11.429
9	Vitamins	8	0.083
10	Calcitriol	8	1.236
11	Ascorbic Acid	8	0.083
12	6-Aminonicotinamide	8	1.236
13	Inositol	8	1.236

3.2 Arrowsmith 候选药物的闭合验证

使用 Arrowsmith 工具进行闭合验证，选择 Arrowsmith 工具进入其检索界面，以 Autistic disorder 作为集合 A 的检索词，以 Folic Acid 作为集合 C 的检索词进行检索。对 A、C 两个文献集合的中间关联词进行挖掘并进行语义筛选，保留属于“Chemicals＆Drugs”“Genes & Molecular Sequences”和“Gene & Protein Names”语义类别的关联词，得到筛选后的关联词列表 B。去掉同义的关联词，按照相关度的高低排列，并对相关度≥0.98 的词进行统计，共筛选出 30 个关联词，如表 3。

表 3 相关度≥0.98 的关联词

排名	关联词	相关度
1	wnt	0.99
2	akt	0.99
3	mthfr	0.99
4	mtor	0.99
5	stat3	0.99
6	toll receptor	0.99
7	gene c677t	0.99
8	mmr	0.99
9	pi3k	0.99
10	pten	0.99
11	dexmedetomidine	0.99
12	mapk	0.99
13	metformin	0.99
14	bdnf	0.99
15	single nucleotide	0.98
16	foxp3	0.98
17	transporter gene	0.98
18	fragile x	0.98
19	fetal alcohol spectrum	0.98
20	adiponectin	0.98
21	caspase-3	0.98
22	nrf2	0.98
23	valproic acid	0.98
24	inflammatory cytokine	0.98
25	vitamin d supplementation	0.98
26	sulforaphane	0.98
27	valproate	0.98
28	erk signaling	0.98
29	tsc2	0.98
30	wnt beta catenin	0.98

3.3 结果讨论

通过对以上关联词的分析检索，从相关度≥0.99 的关联词中选取“bdnf”关联词进行讨论。

脑源性神经营养因子（brain-derived neurotrophic factor，BDNF）是具有促进神经元分化、生长及突触形成的神经营养因子，主要在中枢神经系统内表达，其中海马和皮质的含量最高，被认为是自闭症患者神经重塑的关键因子[7]。有研究结果显示，BDNF 水平与自闭症患者病情程度呈显著相关性，高 BDNF 水平有助于缓解自闭症患者的病情，认为 BDNF 水平的升高很有可能是抵抗自闭症病理的一种代偿性机制[8]。低 BDNF 水平不利于自闭症的治疗，有报道称 BDNF 水平的降低与同型半胱氨酸（homecysteine,Hcy）诱导的氧化应激有关，而叶酸缺乏会导致血浆内 Hcy 酸浓度升高，使谷氨酸能受体活化诱发氧化应激，Mate[9]等也证实了高 Hcy 酸浓度会降低海马区 BDNF 含量，诱发氧化应激反应[10][11]。孙彩虹[12]等通过采用病例对照研究方法，发现叶酸干预前，自闭症儿童叶酸含量低于正常儿童，而血浆内 Hcy 酸含量高于正常儿童，叶酸干预后自闭症儿童血浆内 Hcy 酸含量明显降低。因此，自闭症患者可以通过补充适量叶酸，减少 Hcy 酸含量，从而弱化

氧化应激反应，升高 BDNF 水平，促进神经元的分化、生长以及重塑，缓解病情。

4 小结

随着信息技术的快速发展和信息资源的日渐丰富，非相关文献知识发现的应用将会越来越广泛，利用非相关文献知识发现的情报学方法，可以帮助研究人员发现一些新颖、有趣且可能真实的科学假设、关联等，尤其是在医学领域中，可以进行药物预测、发现以及重定位研究，为疾病治疗带来福音。而 BITOLA 和 Arrowsmith 工具为进行更有效的知识发现提供了平台。在前文中，我们利用 BITOLA 发现叶酸可能对自闭症的辅助治疗提供帮助，故又用 Arrowsmith 进行闭合验证，筛选出了相关度≥0.98 的 30 个关联词，以相关度≥0.99 的“bdnf”关联词为例进行讨论分析，发现自闭症患者给予叶酸干预，可以减小体内 Hcy 酸含量，从而弱化氧化应激反应，升高 BDNF 水平，高 BDNF 水平有助于缓解自闭症患者的病情，所以，在自闭症患者的临床治疗中，用其他对症药物联合叶酸治疗，可能会收获较好的治疗效果。当然，以上结果还需要科学实验进行验证，我们通过这种方法，可以发现更多有用、潜在的知识。

未来，通过机器学习、自然语言处理、人工智能等技术，知识发现工具会更加完善，可以快速、准确地进行非相关文献知识发现，从而为创新研究、跨学科的合作、不同领域之间的融合提供更多的支持，为科学问题提供更加完美的解决方案。

PU-108

医学语料库研究及其临床应用

赵浚植[1] 戴维[2] 赵文龙[2]
1.重庆医科大学第三临床学院
2.重庆医科大学医学信息学院

目的 了解国内外医学语料库的建设及运用现状。

方法 采用文献研究法

结果 语料库是一定量计算机可读文本的集合，取样的文本在最大程度上代表一种语言或变体 。语料库构建研究始于通用领域，语料库常常和本体、术语集等概念混用，其核心是术语。医学语料库最著名的当属美国国立医学图书馆研发的 UMLS （Unified Medical Language System）；医学系统命名法-临床术语（Systematized Nomenclature of Medicine -- Clinical Terms，SNOMED CT ）则是当前临床应用最为广泛的术语标准，涵盖如疾病、症状、操作、微生物、药物等方面的临床信息，支持实现临床数据的标引、存储、检索和聚合，便于计算机处理。面向生物医学领域国内学者也开展了卓有成效的研究，也建立了一定规模的医学语料库，如：CUMLS、中文 MedDRA、“七巧板”医学术语集、中医临床术语集、生物医学信息本体系统及各种单病种语料库等。医学语料库在临床科研中主要用于对电子病历的实体（概念、术语）的识别、分类、以及标准化。对电子病历的挖掘开发是目前临床研究的热点，如何高效地挖掘大量电子病历数据中的信息是一个重要的研究课题。语料库作为信息抽取的基础，其质量直接影响实体识别和关系抽取的结果，现有的语料库构建方法分为手动构建和自动构建两种。手动构建效率低下，自动构建大都依赖于 AI 去发现实体及其关系，准确率有待提高。

结论 已建成的医学语料库大多是单一疾病或特定用途的小规模数据集，缺乏鲁棒性。生物医学文本语料库在医学文本挖掘与知识发现领域的重要作用，整合更多生物医学研究领域词表或本体，借助众包、机构间协作等多种方式构建语义类型多样、语义关系丰富的生物医学文本语料库，可为命名实体识别、关系抽取等研究打下坚实基础。语料元数据的建设需要长期的积累和持续的更新，因此建立一种具有众包共建模式的半自动化更新的规模化的、完备的医学语料元数据库建设模式和生产系统很有必要。

PU-109

基于语义三元组和网络药理学的抑郁症药物知识发现

邰杨芳[2]，苏鑫[1]，华国旻[2]，吴娟[1]

1.山西医科大学 基础医学院

2.山西医科大学 管理学院

抑郁症是一种常见的精神疾病，其主要临床特征包括情绪低落、悲伤、活动减少、失眠等，给患者造成极大的痛苦甚至引起患者的自杀行为。同时，抑郁症患病率逐年上升，已经是造成全球健康负担的主要原因之一[1]。在治疗方面，当前临床上的一线抗抑郁药存在治疗效果不佳、副作用大、易产生耐药性等问题。因此，安全、有效的药物研发一直是抑郁症研究领域的重要课题。然而，传统的新药开发周期长，风险高，人力财力都耗费巨大[2]。

药物重定位是指对已上市药物寻找新适应症的过程，为解决新药研发的周期和成本问题提供了新途径，成为新兴的药物研发策略。基于实验方法的药物重定位需要大量的成本和较长的周期，若能采用计算的方法先对大量数据进行系统分析与预测，然后利用预测结果指导实验，则可提高实验成功率，药物重定位效率更高。网络药理学的产生和发展使药物重定位研究实现从基于实验的方法向基于计算或计算与实验协同方法的跃迁，它以数据分析、计算机虚拟计算、网络数据库检索为基础，从生物学和关联网络的角度对生物系统进行生物计算、系统整合，可以阐明药物作用于疾病的复杂分子机制，从而进行药物重定位、指导新药发现，成为新药发现和药物重定位研究的重要技术之一[3]。

药物重定位候选药物的线索信息最早源于临床观察和盲法筛选，偶然发现药物未知的适应证；基于计算的重定位药物线索主要来自对支持药物重定位的实验或临床等相关的各类生物医药数据库集合（如基因、疾病、药物等数据库）的计算、对生物医学文本的挖掘和推理。其中，基于生物医学文本挖掘的重定位候选药物的筛选不受已有生物医学数据库的限制，对于未有明确实验或机制说明的疾病或药物的知识发现可提供更多的关联信息，有利于发现隐含的药物知识假说[4]。随着自然语言处理技术的发展和生物医学自然语言处理工具和平台开发，基于自然语言处理的文本挖掘方法因能对文本中概念及其关系进行分析和提取从而实现基于语义的文本细粒度知识发现，被认为是优于基于文献主题或关键词、引文信息等传统知识挖掘方法的先进方法，在重定位药物的筛选中得到应用[5, 6]。

基于自然语言处理方法提取出的知识实体及其关系形成由主语-谓词-宾语（Subject-Predicate-Object，SPO）构成三元组结构，也称为 SPO 三元组。 SPO 中的主语和宾语来自文本中的同一句子，语义表达明确且易于理解，在隐含信息的挖掘方面比其他关键词共现方法准确度更高，更适合生物医学文献中的隐含知识发现[7]。

基于上述分析，本文采用基于计算的方法从现有药物中筛选出抑郁症的潜在药物，实现药物新知识发现。首先，采用自然语言处理方法从抑郁症相关的科研文献中提取 SPO 三元组并对其进行筛选，从中提取出可能抗抑郁症的重定位候选药物线索信息；其次，去除其中已获 FDA 批准的抑郁症治疗药物，得到抑郁症的潜在药物；然后，通过网络药理学方法分析潜在药物对抑郁症的作用机制（靶点和通路分析）；最后，通过计算药物-靶点的结合力来验证药物治疗抑郁症的潜在可能性，为后续的实验研究、临床试验和临床用药提供一些参考。

1 材料与方法

基于语义三元组和网络药理学分析的抑郁症药物发现框架包括两部分内容：①使用 SemRep 工具抽取 PubMed 相关文献，生成 SPO 三元组数据，筛选潜在的抗抑郁药物。②潜在药物的网络药理学分析。

1.1 潜在抗抑郁药信息的筛选

1.1.1 SPO 语义三元组的抽取

从全球使用最广泛的生物医学文献数据库 PubMed 中检索并获取抑郁症的相关研究文献；去除不相关文献及综述型文献后，采用美国国家医学图书馆开发的自然语言处理工具 SemRep 对文献的摘要文本进行分词、命名实体识别和实体间语义关系基出上的语义关系抽取，得到以句子为单元抽取出的 SPO 语义三元组。SemRep 是一个基于规则的系统，结合了句法、语义规则和统一医学语言系统（Unified Medical Language System，UMLS）的结构化生物医学知识，用于支持各种临床和转化应用，包括临床决策、医学诊断、药物重定位、基于文献的知识发现及假设生成等，在科学界影响广泛[8]。

提取出的 SPO 语义关系三元组中，Subject 和 Object 是统一医学语言系统 UMLS 叙词表中的规范化概念与术语，即语篇的知识实体。每一个实体在 UMLS 均有特定的语义类型； Predicate 则是以谓词形式展现的实体间的各类型语义关系。

1.1.2 SPO 语义三元组的筛选

UMLS 词库中的概念具有层级结构，映射 UMLS 抽取出的 SPO 三元组可能会出现非特指或无实质含意的概念；在纳入的文本数据量较大时，难免会出现与研究目的相关性较弱的三元组；因此要进行三元组的数据清洗，以便得到更准确的结果。

参考 Fiszman **错误!未找到引用源。**和蔡妙芝等提出的部分三元组清洗原则和研究需要[9, 10]，本文首先根据研究目的确定用于药物重定位分析的三元组的语义关系（即，谓词 P）和语义类型（即生物医学概念实体 S 和 O 的语义类型）；然后从中过滤掉过于泛指的无实质含义的实体概念。之后，统计所有符合条件的三元组频次（同一文献中的相同三元组只计一次）并按降序排列。

1.1.3 抑郁症潜在药物信息的提取

从筛选后的抑郁症药物相关语义三元组中提取其中“抑郁症”实体（S 或 O）的另一端(O 或 P)对应的药物实体集合，排除其中已被 FDA 批准的抗抑郁药，得到的即是有迹象表明的潜在抗抑郁药物，选择频次大于一定阈值的药作为有潜在价值的重定位药物。

1.2 抑郁症潜在药物的网络药理学分析

1.2.1 潜在药物的抗抑郁靶点获取

借助已经被科学实验和文献验证的 6 种生物医学及药物数据平台，获取每一种潜在药物的抗抑郁核心靶点。表 1 是对各数据平台的数据、功能等的简要说明。具体步骤如下：

（1）在 TTD、OMIM、Genecards（score≥2）、Clinvar 数据库中检索，获得抑郁症的相关基因；

（2）在 PubChem 数据库中检索并获取潜在药物的分子结构式 smile，将其输入到 Swiss Target Prediction 平台，设置 Probability＞0，获取潜在药物靶点，与在 Genecard（score≥2）上获取的潜在药物靶点合并去重后，得到潜在药物的作用靶点。

（3）利用 Jvenn web tool (http://jvenn.toulouse.inra.fr/app/example.html)计算得出潜在药物靶点与抑郁症靶点的交集，即为潜在药物的抗抑郁靶点。

表 2 生物医学及药物数据平台简介
Tab 1 Introduction to Biomedical and Drug Data Platforms

数据库	简介
TTD (http://db.idrblab.net/ttd/)	存储药物-靶点-疾病的相互作用信息的数据库
OMIM (https://www.omim.org/)	存储遗传疾病的基因、表型、遗传方式和参考文献的人类遗传疾病数据库
Genecards (https://www.genecards.org/)	存储人类基因的序列、表达、功能、关联疾病的基因信息数据库
Clinvar (https://www.ncbi.nlm.nih.gov/clinvar/)	存储基因变异、染色体变异等遗传变异信息的人类遗传变异数据库
SwissTargetPrediction (http://www.swisstargetprediction.ch/)	基于药物分子结构信息和靶点蛋白序列信息预测药物靶点
PubChem (https://pubchem.ncbi.nlm.nih.gov/)	存储化合物结构、性质、药效等信息的化合物信息数据库

1.2.2 潜在抗抑郁药关键靶点蛋白互作网络的构建

将交集靶点输入到 STRING 数据库，设定物种为“Homo sapiens”，置信度设置为 0.4 进行筛选，获取靶基因的蛋白质相互作用（protein-protein interaction，PPI）信息，将其以 TSV 格式导出文件。将其导入 Cytoscape 软件，并通过 cytohubba 插件进行网络分析，选取对应度值大于平均度值的视为核心靶点。

1.2.3 抗抑郁关键基因的功能与通路富集分析

利用 R3.6.0 软件编写代码对潜在抗抑郁药物作用的抑郁症核心基因进行 GO 功能富集和 KEGG 通路富集分析，将富集分数靠前的生物过程（biological process，BP）、细胞成分（cellular component，CC）、分子功能（molecular function，MF）和信号通路（pathway）分别绘制成直方图和气泡图，用于揭示不同基因功能注释结果的分布情况和不同代谢通路的富集程度，探测潜在药物治疗抑郁症可能的作用机制。

1.2.4 潜在药物成分-关键基因编码蛋白的分子对接

分子对接是一种计算机辅助技术，通过受体与药物分子之间的相互作用预测它们的结合方式和结合能力[11]，从分子水平上模拟验证有效成分作用于疾病的科学依据，为药物进一步的研发提供指导。

分别从每个潜在药物的 PPI 网络中选取一个核心蛋白基因，使用 PDB 寻找含有对应靶点的蛋白晶体结构，作为蛋白受体。蛋白质晶体结构通常会包含一些水分子但不包含氢原子，在进行分子对接时可能会影响结果的准确性。为了得到更稳定更利于分子对接的蛋白构象，要通过 AutoDockTools 对蛋白受体进行去水、加氢、能量优化等一系列处理，保存为 PDBQT 格式文件，再进行分子对接。

从 PubChem 中获取潜在药物（小分子配体）的 3D 结构，这个结构是某一构象下的分子结构，它的能量可能不是最低的，所以为了得到最稳定的构象，要对得到的 SDF 格式文件进行能量优化。将优化后的 SDF 文件导入 AutoDockTools（Version 1.5.7），检测其扭转中心和扭转键，并保存为 PDBQT 格式，作为配体。

使用 AutoDockTools 软件对受体蛋白和配体小分子进行分子对接。构建 20*20*20 对接盒子，设 Autodock Vina 对接模型输出为 20，其他参数保持默认值；将得到的分子对接结果通过 OpenBabel 转化为 PDB 格式，然后导入 PyMOL（3D， Version2.2.0）和 Discover studio visualizer 2022 对其可视化。

2 结果

2.1 SPO 语义三元组的抽取

为了检索出尽可能全的相关文献以便在后续的药物发现中找到更多的潜在药物，本研究设置了较为宽泛的检索词和较大跨度的时间范围。在 PubMed 中以"Depression"[Mesh]为检索词，去除综述型文献后，得到 2000 年 1 月—2021 年 12 月间发表的抑郁症研究文献 17391 篇，获取文献的 PMID 号、标题和摘要等数据。

使用 SemRep 对标题和摘要文本进行语义抽取，得到 654420 条三元组数据，使用 Python 3.7 将三元组非结构式文本转化为 Excel 结构化文本格式，如图 1 所示。

	A	B	C	D	E	F	G	H	I	J	K
406052	28190681.ab.10	relation	C0011860	Diabetes Mellitus, N	dsyn	dsyn	PROCESS_OF	C0025266	Male population group	popg,humn	humn
406053	28190681.ab.10	relation	C0025598	Metformin	orch,phsu	phsu	TREATS	C0025266	Male population group	popg,humn	popg
406054	28190681.ab.10	relation	C0025598	Metformin	orch,phsu	phsu	TREATS(INF	C0011860	Diabetes Mellitus, Non-Insul	dsyn	dsyn
406055	28190411.ti.1	relation	C0086132	Depressive Symptoms	sosy	sosy	PROCESS_OF	C0043210	Woman	popg,humn	humn
406056	28190411.ab.2	relation	C0086132	Depressive Symptoms	sosy	sosy	PROCESS_OF	C0599755	cohort	popg,humn	humn
406057	28190411.ab.2	relation	C0238711	AUSTRALIAN	fndg	fndg	PROCESS_OF	C0043210	Woman	popg,humn	humn
406058	28190411.ab.3	relation	C0086132	Depressive Symptoms	sosy	sosy	PROCESS_OF	C0043210	Woman	popg,humn	humn
406059	28190411.ab.6	relation	C0020699	Hysterectomy	topp	topp	compared_w	C0020699	Hysterectomy	topp	topp
406060	28190411.ab.6	relation	C0086132	Depressive Symptoms	sosy	sosy	PROCESS_OF	C0043210	Woman	popg,humn	humn
406061	28190411.ab.8	relation	C0020699	Hysterectomy	topp	topp	NEG_ADMIN	C0043210	Woman	popg,humn	humn
406062	28190411.ab.8	relation	C0086132	Depressive Symptoms	sosy	sosy	PROCESS_OF	C0043210	Woman	popg,humn	humn
406063	28190411.ab.8	relation	C0279025	Endocrine therapy	topp	topp	NEG_ADMIN	C0043210	Woman	popg,humn	humn
406064	28190411.ab.8	relation	C0279025	Endocrine therapy	topp	topp	NEG_ADMIN	C0043210	Woman	popg,humn	humn
406065	28190411.ab.8	relation	C0279025	Endocrine therapy	topp	topp	compared_w	C0020699	Hysterectomy	topp	topp

图 4 SemRep 输出结果转化后示例
Fig 1 Example of SemRep Output Result Conversion

其中，A 列是“PMID+表示抽取出的概念实体之间关系的来源位置”；B 列表示概念与概念之间有“relation（关系）”,G 列则表示此关系具体是什么关系；D 和 I 列是 SemRep 从文献中抽取出的两个共现概念，C 和 H 列分别为两个共现概念的概念代码 CUI 号，EF 和 JK 分别是在 UMLS 中两个实体的语义类型。

以第 406054 行数据为例，SemRep 从 PMID 为 28190681 的文献摘要第 10 条句子中抽取出三元组：药物 Metformin（二甲双弧）对疾病 Diabetes Mellitus, Non-Insulin-Dependent（糖尿病，非胰岛素依赖型）有 TREAT（治疗）作用。

2.2 潜在抗抑郁药的筛选

为提取有效的潜在抗抑郁药线索信息，需要来自文献标题及摘要文本的语义三元组在概念实体的语义类型、实体间的关系类型并出现频次上满足一定的要求。具体如下：

（1）谓词关系类型。为提高研究结果的相关性，本文参照 Rosemblet 和 Fiszman 等提出的药物[12, 13]、疾病相关的关系谓词，将谓语关系类型限定为：TREAT/TREAT（INFER）、INHIBITS、DISRUPTS、PREVENTS。

（2）实体语义类型。将主语语义类型设置为药物相关的：如 antb，bacs，elii 等，将宾语语义类型设置为疾病相关的：如 acab，anab，agab 等。具体如表 2 所示：

表 2 语义类型举例说明
Tab 2 Example of semantic types

类别	实体的语义类型
药物	Elii : Enzyme/ligand interaction-酶/配体相互作用 Phsu： Pharmaceutical Substance-在 UMLS 中表示所有种类的药物、制剂和化合物的概念 Clnd : Clinical Drug-临床药物 Antb : Antibiotic-抗生素
疾病	Chvf : Congenital Vascular Abnormality-血管疾病 Acab：Acquired Abnormality-获得性异常。该类型用于描述在生命周期中后期获得的疾病、症状、损伤或其他异常状态。 Anab：Anatomical Abnormality-解剖异常。该类型用于描述身体器官、组织或解剖结构的异常状态，包括先天畸形和后天病变。 Agab ：Age-related Abnormality-与年龄相关的异常。该类型用于描述随着年龄增长而出现的生理或病理异常。 Diso ：Disease or Syndrome-疾病或综合征。该类型用于描述医学上诊断出的疾病或疾病综合征，包括常见疾病、罕见病和遗传性疾病等。

（3）实体概念较为专指。实体概念越具体，表达的意义越明确。因此，需要去除主语和宾语中的宽泛概念，如：“Antidepressive Agents”、“pharmaceutical preparation”、“patient in hospital”、“patient Non-compliance”等等。

（4）在文献中出现的频次较高。在同一篇文章不同句子中出现的三元组，其频次只记一次。对符合限定条件的数据去重整理并统计其频次，得到 1116 条三元组数据。其中，频次≥30 的三元组数据有 49 条，涉及 15 种药物实体。15 种药物实体中，有 9 种是已被 FDA 批准的抗抑郁药。因此，其余 6 种药可作为药物重定位候选药物，它们是：Benzodiazepine、Curcumin、Hydrocortisone、Metformin、Nicotine、Risperidone，表 3 是潜在药物的概述。

表 3 抑郁症潜在药物概述
Tab 3 Overview of potential drugs

药物	上市时间	已知用途	作用机制
Benzodiazepine (苯二氮䓬类药物)	1960 年代	焦虑症、失眠症	通过增强神经递质 GABA 的抑制作用，减少神经兴奋性，从而产生镇静、抗焦虑和抗抽搐的效果
Curcumin (姜黄素)	天然产物	抗氧化、抗炎、抗癌	抑制炎症细胞的活化、减轻炎症细胞的浸润和减少炎症介质的释放；抑制肿瘤细胞的增殖、诱导肿瘤细胞凋亡和抑制肿瘤血管生成等
Hydrocortisone (氢化可的松)	1950 年代	各种炎症、过敏反应、自身免疫性疾病	抑制炎症介质的合成和释放
Metformin (二甲双胍)	1994 年	2 型糖尿病	通过降低肝脏糖原合成、增加胰岛素敏感性、抑制肠道糖吸收等途径
Nicotine (尼古丁)	1991 年	辅助戒烟	模拟神经递质乙酰胆碱，刺激中枢神经系统和交感神经系统
Risperidone (利培酮)	1993 年	精神分裂症和双相障碍	调节多巴胺和 5-羟色胺等神经递质的水平

2.3 潜在抗抑郁药关键靶点蛋白相互作用网络（PPI）

以潜在药物 Hydrocortisone（氢化可的松）为例，利用 Jvenn web tool 发现 Hydrocortisone 和抑郁症的交集靶点 153 个。将其导入 STRING 数据库，获取 PPI 网络如图 2 所示，图中共有 151 个靶蛋白节点，5500 条相互作用的边，平均度值为 72.8476821192053，平均节点紧密度（Closeness Centrality）为 0.551337096075086，平均节点介度（Betweenness Centrality）为 0.0057110686400877；选取对应度值大于平均度值的视为核心靶点，共有 33 个。其中排名靠前的 15 个靶点（表 4 所示）：INS、ALB、IL6、TNF、GAPDH、IL1B、LEP、IGF1、MAPK3、CXCL8、PTGS2、EGFR（表皮生长因子受体）、EGF、IL10、BDNF；这些靶点在 PPI 网络中起到了重要的作用，可能是 Hydrocortisone 发挥抗抑郁作用的关键靶点。

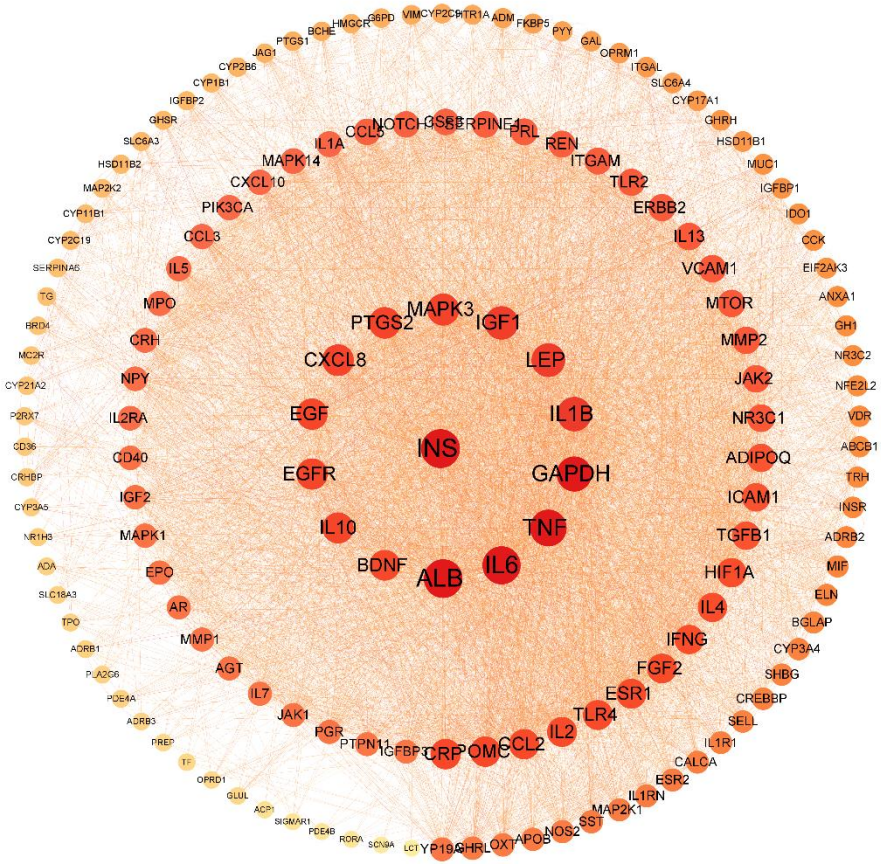


图 5 Hydrocortisone 的 PPI 网络图
Fig 2 Hydrocortisone's PPI network diagram

表 4 PPI 分析结果（degree 排名 top15）
Tab 4 PPI analysis results (degree ranking top 15)

靶点简称	中文全称	度值
INS	胰岛素	226
ALB	白蛋白	224
IL6	白细胞介素 6	218
TNF	肿瘤坏死因子	204
GAPDH	糖酵解催化酶	188
IL1B	白细胞介素 β	182
LEP	瘦素	174

IGF1	胰岛素样生长因子 1	172
MAPK3	丝裂原活化蛋白激酶 3	158
CXCL8	趋化因子	154
PTGS2	前列腺素合成酶 2	154
EGFR	表皮生长因子受体	150
EGF	表皮生长因子	150
IL10	白细胞介素-10	148
BDNF	脑源性神经营养因子	144

2.4 GO 富集分析与 KEGG 通路分析结果

2.4.1 GO 分析

从潜在药物与抑郁症的交叉靶蛋白 GO 富集分析结果选取生物过程、细胞组成及分子功能中 q-value 最小的前 10 条可视化，从这些可视化结果中，我们可以看到 Hydrocortisone(氢化可的松)，Benzodiazepine(苯二氮䓬类药物)，Risperidone(利培酮)，Curcumin(姜黄素)，Nicotine(尼古丁)和 Metformin(二甲双胍)这六种潜在药物在不同的生物过程、细胞组成和分子功能方面具有不同的表现。具体如图 3 所示。

从生物过程的角度来看，Hydrocortisone，Curcumin 和 Nicotine 都特别强调了炎症反应的调节。此外，Benzodiazepine，Risperidone 和 Nicotine 都与认知过程有关；Risperidone 和 Benzodiazepine 还涉及到神经递质水平的调节。而 Metformin 的富集结果显示其主要涉及葡萄糖代谢过程。

从细胞组成的角度看，Hydrocortisone， Metformin 和 Nicotine 都涉及到囊泡腔和血小板 α 颗粒。Benzodiazepine，Risperidone 在突触前膜和突触后膜的成分方面有较高的富集。

从分子功能的角度看，大部分药物都显示出受体配体活性和信号受体激活剂活性。此外，Curcumin 还显示出了磷酸酶结合和蛋白磷酸酶结合。

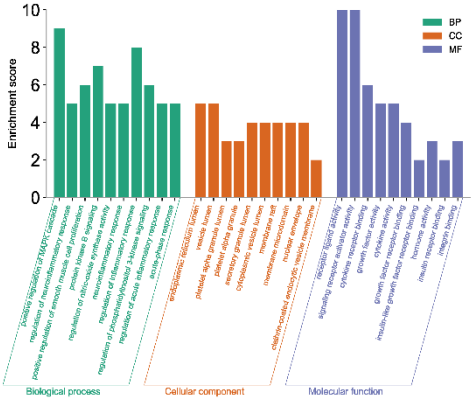
综上，这些药物在多种生物过程、细胞组成和分子功能中的富集结果揭示了它们与抑郁症病理生理机制的多个层面的潜在相互关系，这为我们理解这些药物如何对抑郁症产生作用提供了更深一步的理解。

2.4.2 KEGG 分析

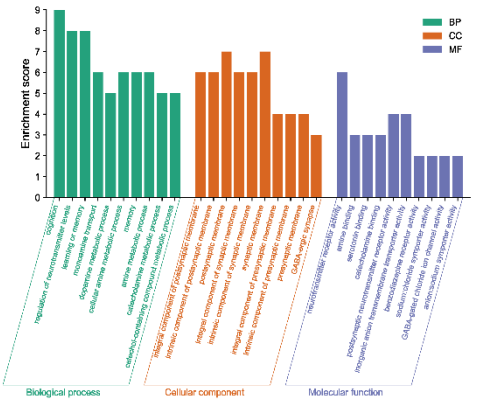
从潜在药物与抑郁症的交叉靶蛋白 KEGG 通路分析结果选取前 15 条可视化，如图 4 所示。通路富集结果发现靶点主要涉及丝裂原活化蛋白激酶（MAPK）信号通路、人巨细胞病毒感染、非酒精性脂肪肝、多巴胺能突触、肿瘤坏死因子（TNF）信号通路、白细胞介素-17（IL-17）信号通路。

从可视化结果来看，六种主要通路中，多巴胺能突触与 TNF 信号通路可能是最直接与抑郁症相关的，非酒精性脂肪肝和人巨细胞病毒感染与抑郁症的关联性可能较弱。由于抑郁症的复杂性，

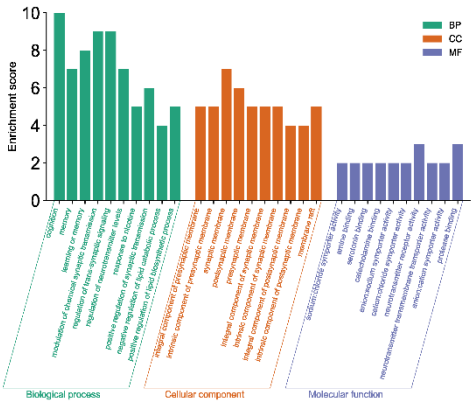
这些通路可能是各种因素在发病机制中相互作用的结果，而不仅仅是单一因素产生的影响，所以有效的治疗策略可能需要同时针对多个不同的通路。



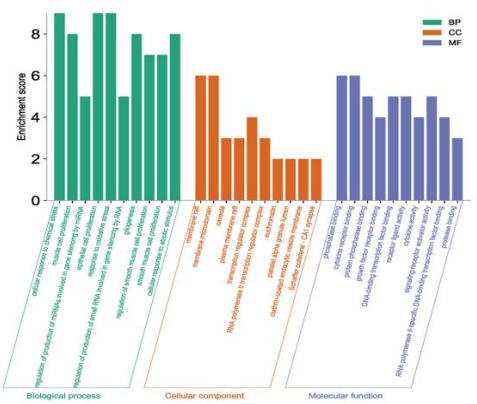
3a: Hydrocortisone



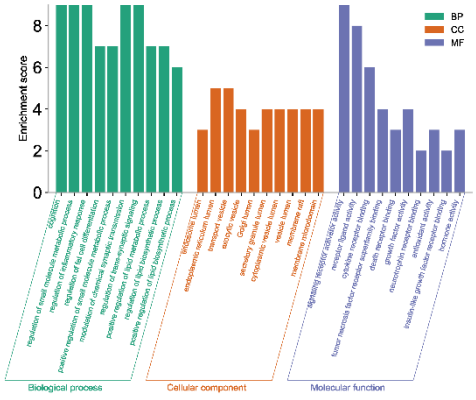
3b: Benzodiazepine



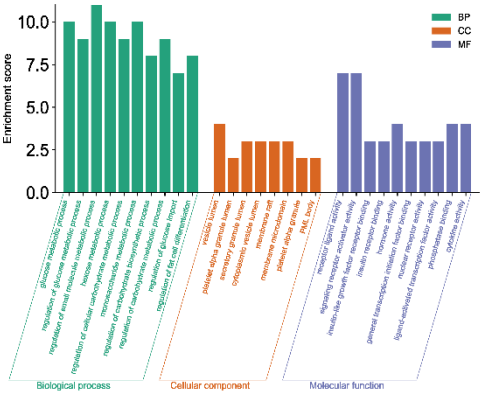
3c: Risperidone



3d: Curcumin

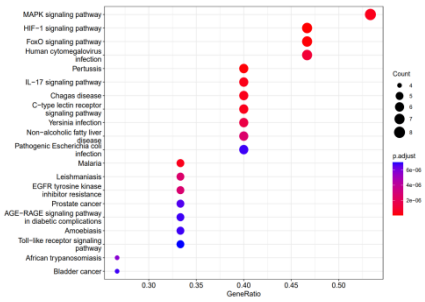


3e: Nicotine

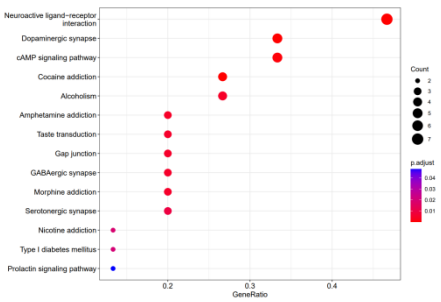


3f: Metformin

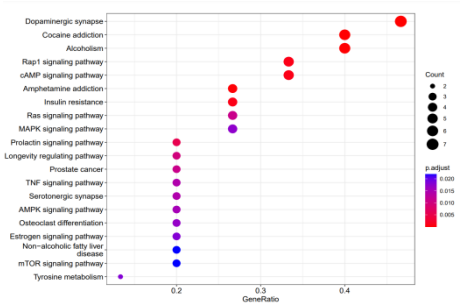
图 6 GO 富集分析结果可视化
Fig 3 Visualization of GO enrichment analysis results



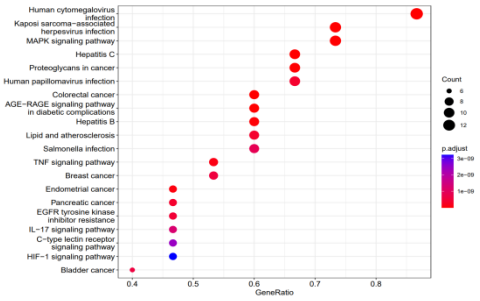
4a: Hydrocortisone



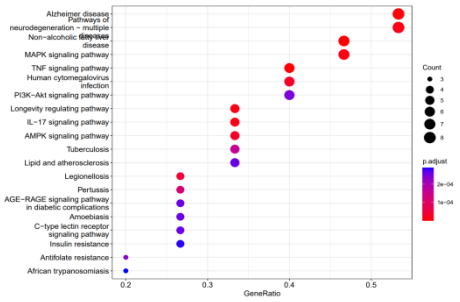
4b: Benzodiazepine



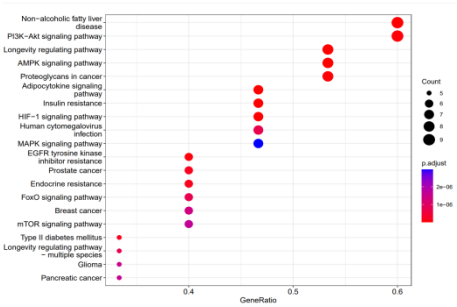
4c: Risperidone



4d: Curcumin



4e: Nicotine



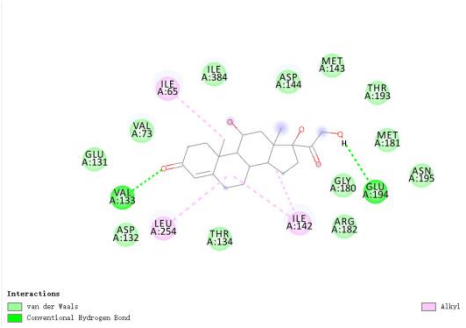
4f: Metformin

图 7 KEGG 结果可视化

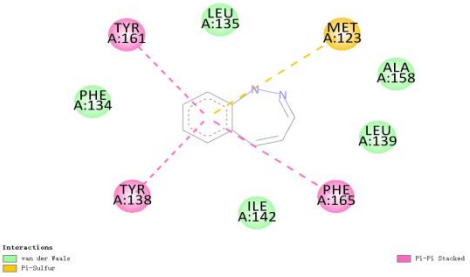
Fig 4 Visualization of KEGG results

2.5 分子对接结果

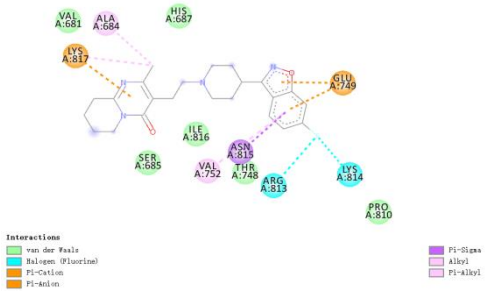
分子对接结果经 PyMOL(3D,Version2.2.0)和 Discover studio visualizer 2022 进行可视化后，如图 5（氢键由虚线表示）和图 6 所示：其中 Hydrocortisone（氢化可的松）与其蛋白受体的结合构象中存在两个氢键，Curcumin（姜黄素）、Nicotine（尼古丁）、Metformin(二甲双胍)各存在一个，而 Benzodiazepine（苯二氮䓬类药物）、Risperidone(利培酮)并未与其蛋白受体形成氢键。通常情况下氢键数量越多，表明结合地越稳定，对接结果越理想。经 AutoDock 计算得到各潜在药物与其对应蛋白受体的结合能，结合能数值越小表示结合力越强，结合力强弱排序如表 5 所示：其中 Hydrocortisone、Benzodiazepine 结合能数值＜-7 kcal/mol，但从结合能数值来看，表明它们具有较强的结合力；Risperidone、Curcumin 、Nicotine、Metformin 结合能数值＜-4 kcal/mol，这表明这些潜在药物与其对应的蛋白受体也有一定的结合力。



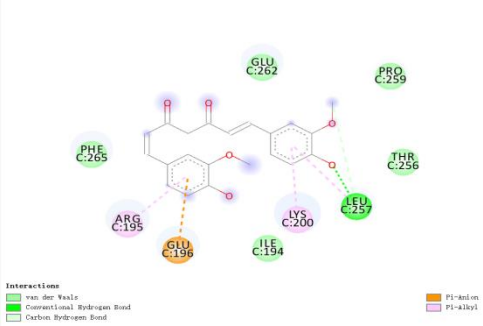
5a: Hydrocortisone



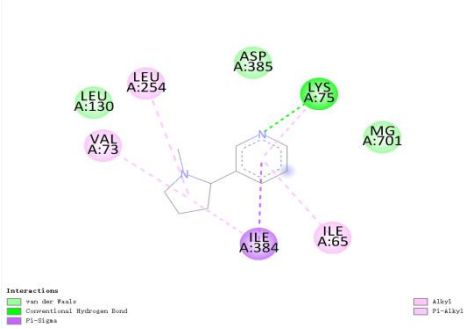
5b: Benzodiazepine



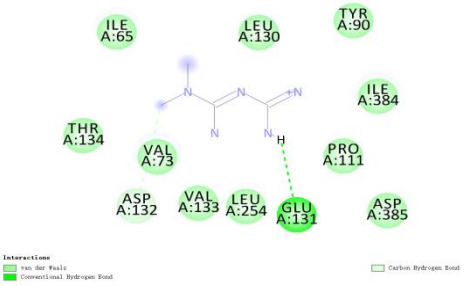
5c: Risperidone



5d: Curcumin



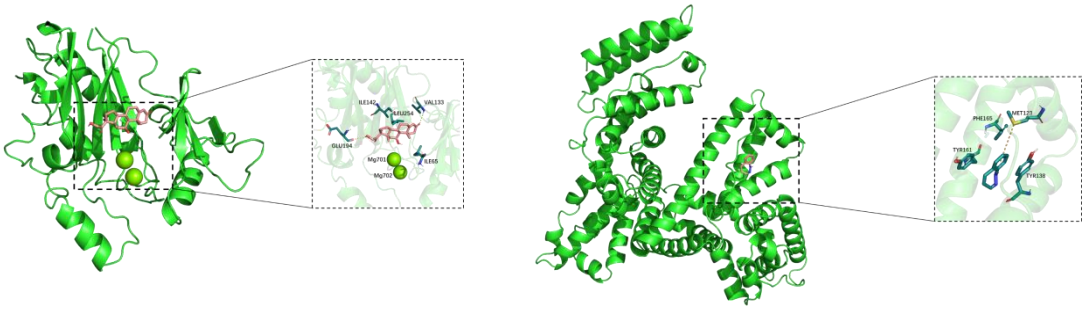
5e: Nicotine



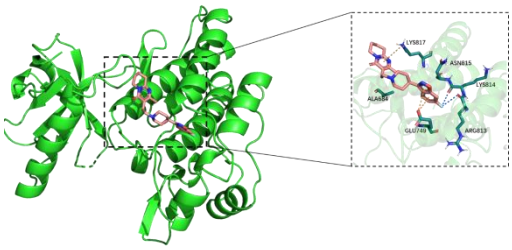
5f: Metformin

图 8 分子对接结果 2D 可视图

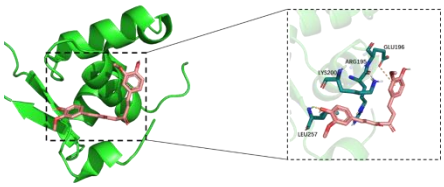
Fig 5 2D visualization of Macromolecular docking results



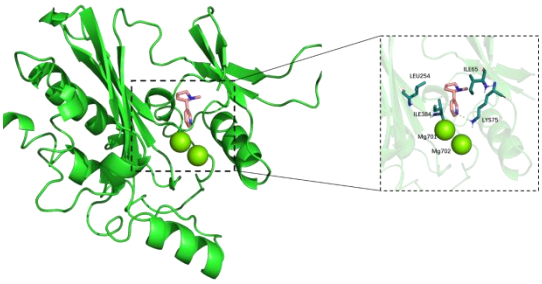
6a: Hydrocortisone



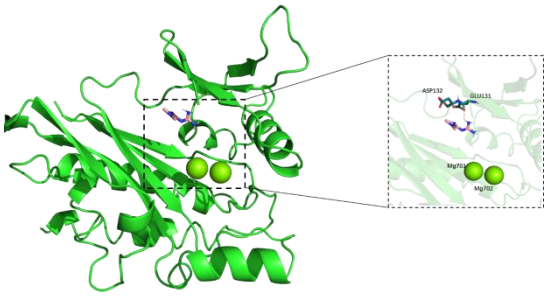
6b: Benzodiazepine



6c: Risperidone



6d: Curcumin



6e: Nicotine

6f: Metformin

图 9 分子对接结果 3D 可视图
Fig 6 3D visualization of Macromolecular docking results

表 5 分子对接结合能
Tab 5 Macromolecular docking Binding energy

receptor-ligand(受体-配体)	affinity(结合能)（kcal/mol）
INS-Hydrocortisone	-8.3
ALB-Benzodiazepine	-7.8
BDNF-Risperidone	-6.6
TP53-Curcumin	-6.4
INS-Nicotine	-4.7
INS-Metformin	-4

3 讨 论

本研究通过 SemRep 抽取 PubMed 文献中的语义三元组，得到 15 种药物实体，其中有 9 种是已被 FDA 批准的抗抑郁药，说明按此标准提取出的三元组所表示的药物-疾病关系存在合理性的概率较大。

其中，以 Hydrocortisone 为例做 PPI 网络分析发现 INS、ALB、IL6、TNF 可能是 Hydrocortisone 发挥抗抑郁作用的重要靶点，这在通路分析和分子对接中也得到了验证说明。Hydrocortisone 通路分析结果显示其靶点主要涉及信号通路包含 TNF，分子对接结果表明其与 INS 有较强的结合力。

对筛选后的 6 种潜在药物做 GO 分析发现，Hydrocortisone，Curcumin 和 Nicotine 都特别强调了炎症反应的调节，可能是炎性因子升高影响皮质醇的信号传导，进而导致皮质醇调控炎性因子

释放量增加，加重 HPA 轴功能紊乱，进而影响抑郁症的发生和发展；在抑郁症的研究领域，过多或过少的皮质醇水平都被认为可能与抑郁症的发病和病情严重程度相关[18]，这与抑郁症的炎症反应病理相关性一致。此外，Benzodiazepine，Risperidone 和 Nicotine 都与认知过程有关，正如多数的抑郁症患者都会出现记忆力和注意力等认知功能的下降现象；已有研究结果表明合适剂量的 Risperidone 可缓解大鼠的抑郁症状；Risperidone 可能是通过改善抑郁症引起的认知障碍、降低肿瘤坏死因子-α (TNF-α)水平、调节神经递质的平衡和促进神经保护因子的释放来起到治疗抑郁症的效果[21]。虽然 Benzodiazepine 可以减轻抑郁症的严重程度；Nicotine 可以改善抑郁症患者的睡眠质量和抑郁症状[22]。但长期使用 Benzodiazepine 和 Nicotine 可能会对患者造成严重的副作用，从而导致抑郁症患者的抗抑郁疗效降低[19,23]。Metformin 的富集结果显示其主要涉及葡萄糖代谢过程，这可能与抑郁症患者常见的体重增加和食欲改变有关。同时，Hydrocortisone， Metformin 和 Nicotine 都涉及到囊泡腔和血小板 α 颗粒，暗示着这些药物可能影响神经递质的释放。而 Benzodiazepine，Risperidone 在突触前膜和突触后膜的成分方面有较高的富集，可能反映了它们在调控突触传递和神经信号的作用。大部分药物都显示出受体配体活性和信号受体激活剂活性，这揭示了它们可能通过调节特定的受体来发挥抗抑郁作用。此外，Curcumin 还显示出了磷酸酶结合和蛋白磷酸酶结合，可能暗示了它在信号传递的调控中起作用。

通路分析结果显示靶点主要涉及 MAPK 信号通路、人巨细胞病毒感染、非酒精性脂肪肝、多巴胺能突触、肿瘤坏死因子（TNF）信号通路、白细胞介素-17（IL-17）信号通路。其中，MAPK 信号被认为与炎症反应、神经适应和神经保护有关[14]，而这些过程可能是导致精神障碍的关键。通路分析结果显示人巨细胞病毒感染通路也存在异常，这表明抗病毒治疗可能有助于缓解这些症状；有研究也表明人巨细胞病毒可以改变宿主的神经免疫反应，从而引发抑郁症[15]。多巴胺能突触在抑郁症的发病机制中起到关键角色，已有的一些抗抑郁药物就是通过影响多巴胺能突触改善患者的抑郁症状[16]，这和潜在药物的 KEGG 可视化结果也是一致的。 非酒精性脂肪肝与代谢疾病，如 2 型糖尿病和高血压密切相关，这些疾病被认为可能与抑郁症有关。因此，调控这个通路也可能有助于减轻与抑郁症相关的神经症状。TNF 信号通路和 IL-17 信号通路都是与免疫机制有关的通路，研究显示抑郁症和炎症反应之间存在很大的关联，TNF 和 IL-17 在炎症反应中起着关键作用[17]，因此它们的信号通路也可能是潜在的药物靶点。由于抑郁症的复杂性，这些通路可能是各种因素在发病机制中相互作用的结果，而不仅仅是单一因素产生的影响，所以有效的治疗策略可能需要同时针对多个不同的通路。

上述结果为进一步的实验验证和药物研发提供了理论基础和有价值的参考信息；本文的不足之处在于只在生物学基础上讨论了药物对疾病和基因的作用和影响，没有考虑药物副作用等问题。未来可以加入临床真实数据，构建多源知识网络，提高药物知识发现的准确度和完备度。

PU-110

新型冠状病毒肺炎医学影像学研究热点与研究前沿的知识图谱分析

[1]姜海 苑旸 [2]潘希娟
[1]大连医科大学图书馆
[2]首都医科大学宣武医院

1 引言

2020 年初突然爆发的新型冠状病毒肺炎肆虐全球，使得数以万计的民众失去了生命，同时造成了世界经济、文化上巨大的损失。在新型冠状病毒肺炎（2019-nCoV）的诊断和疫情防控过程中，医学影像学对新型冠状病毒肺炎的确诊、分型、分度的诊断发挥了至关重要的作用，相关研究文献和成果迅速发表和出版，形成研究热点，研究热点不断聚集，大量研究成果不断迭代，形成研究前沿问题，促进临床医生和医疗科研工作者进行深入性研究和探索。

本文以文献计量学和科学知识图谱作为理论基础，通过 CiteSpaceII 信息分析对新型冠状病毒肺炎的医学影像学研究数据进行可视化分析，绘制新型冠状病毒肺炎的医学影像学研究的知识图谱。通过对研究领域的研究热点和前沿进行客观数据的量化分析，辅以医学影像学专家判断的质性

研究，以可视化知识图谱的形式呈现出一个直观的研究热点和研究前沿分析，为医学影像学在新型冠状病毒肺炎诊疗过程的深入研究提供可以借鉴的思路。

2 研究方法

文献中的关键词作为作者对发表文献中核心观点、研究结果与研究方法的凝练与概括，能够客观、准确出分反映文献的核心研究内容[1]。通过共词分析方法对文献中出现的高频关键词进行频次、共现与聚类的分析，以知识图谱方式呈现新型冠状病毒肺炎的医学影像学的研究热点和研究前沿问题，揭示医学影像学在新型冠状病毒肺炎诊疗过程成果的总体结构特征、发展规律、研究热点和发展趋势等内容[2]。

陈超美教授团队开发的 CitespaceII 作为一款具有引文可视化分析功能的定量分析软件可以绘制研究领域的科学知识图谱，在无人工干预前提下自动聚类和标引，整个研究领域文献的关键词自动抽取，所反映的研究问题具有一定客观性。但因其聚类标识关键词过程中无法从专业角度描述研究问题，所以本文在自动聚类标识的基础上，对知识基础和研究前沿中的施引与被引文献进辅以专家解读和经验判断，发现新型冠状病毒肺炎的医学影像学的研究领域的最新进展和研究热点[3]。

3 数据来源及检索策略

以科睿唯安 WOS（Web of Science）平台核心合集(SCI-EXPANDED，SSCI，CPCI-S，CPCI-SSH，CCR-EXPANDED，IC)作为文献来源数据库并且进行分析，检索策略为：TS=(“2019-nCoV”OR“SARS-CoV-2”OR“COVID-19”) AND TS=“imag*”， 时间限定为 2019-2021 年，检索时间是 2021 年 1 月，共获得 6523 条记录。使用 CiteSpace 软件对下载数据进行分析，完成对论文数量、国家-机构-作者合作、研究热点与研究前沿的量化分析和知识图谱，同时下载“新型冠状病毒的影像学”的知识基础和研究热点文献，进行研读和梳理，辅以医学影像学专家和临床医生的实践经验判断，提取这一研究问题的关键信息。我们得出了新型冠状病毒肺炎在医学影像学领域研究热点、研究的知识基础和前沿问题，以及今后研究中新趋势的探索。

4 新型冠状病毒肺炎的医学影像学研究领域研究热点与新趋势

关键词是表达文献主题概念的自然语言词汇，一个学术研究领域较长时域内大量学术研究成果高频关键词的集合，可以揭示研究成果总体内容特征、研究内容之间的内在联系、学术的研究热点和发展趋势等。

利用词频分析法分析文献题录中的关键词/主题词重复出现的频次， 并借用信息可视化的方式表现关键词/主题词频次大小，从而确定新型冠状病毒肺炎医学影像学研究领域的研究热点。通过将检索数据导入 CiteSpace 软件进行分析（检索策略：取出现频次前 30 位的关键词），选择探索关键路径的路径算法（pathfinder），得到高频关键词知识图谱及高频词列表，如图 1、表 1 所示，并按照出现频次（Freq）、中心度（Centrality）二种方式进行排序，选取热点词中排名靠前的关键词，进行分析。

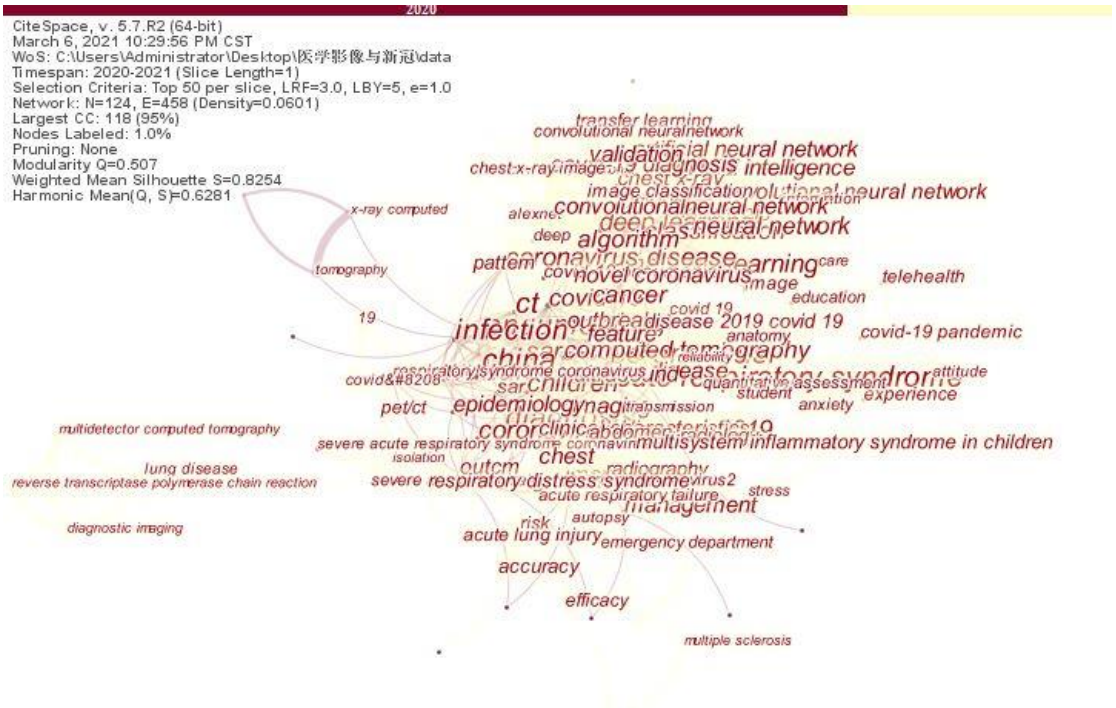


图 1 新型冠状病毒肺炎的医学影像学研究领域研究热点高频关键词知识图谱

以频次（Freq）排序前 10 位出现的关键词为：covid-19（COVID-19）、pneumonia（肺炎）、coronavirus（冠状病毒）、sars-cov-2（SARS-Cov-2）、wuhan（武汉）、computed tomography（计算机断层成像）、china（中国）、infection（感染）、ct（CT）、diagnosis（诊断），反映了新型冠状病毒肺炎的研究热点。

以中心度（Centrality）排序前 10 位出现的关键词为：coronavirus（冠状病毒）、diagnosis（诊断）、pneumonia（肺炎）、acute respiratory syndrome（急性呼吸综合征）、pandemic（大流行）、severe acute respiratory syndrome coronavirus 2（2 型冠状病毒引起的严重急性呼吸综合征）、ct（计算机体层成像）、china（中国）、infection（感染）、wuhan（武汉），反映出医学影像学研究领域研究热点中的一些内容参照指标。

表 1 新型冠状病毒肺炎的医学影像学研究领域研究热点高频关键词（前 20 个）

No.	Freq	Keywords	Centrality	Keywords
1	1361	covid-19	0.25	coronavirus
2	408	pneumonia	0.25	diagnosis
3	390	coronavirus	0.16	pneumonia
4	379	sars-cov-2	0.16	acute respiratory syndrome
5	167	wuhan	0.14	pandemic
6	163	computed tomography	0.12	severe acute respiratory syndrome coronavirus 2
7	140	china	0.11	ct
8	137	infection	0.1	china
9	130	ct	0.1	infection
10	127	diagnosis	0.09	wuhan
11	104	pandemic	0.09	deep learning
12	101	deep learning	0.09	mortality
13	89	acute respiratory syndrome	0.09	management
14	80	covid 19	0.09	lung disease
15	78	coronavirus	0.08	sars-cov-2

		disease 2019		
16	77	sar	0.08	coronavirus disease 2019
17	71	disease	0.07	children
18	71	lung	0.07	coronavirus disease
19	61	outbreak	0.06	computed tomography
20	59	imaging	0.06	cancer

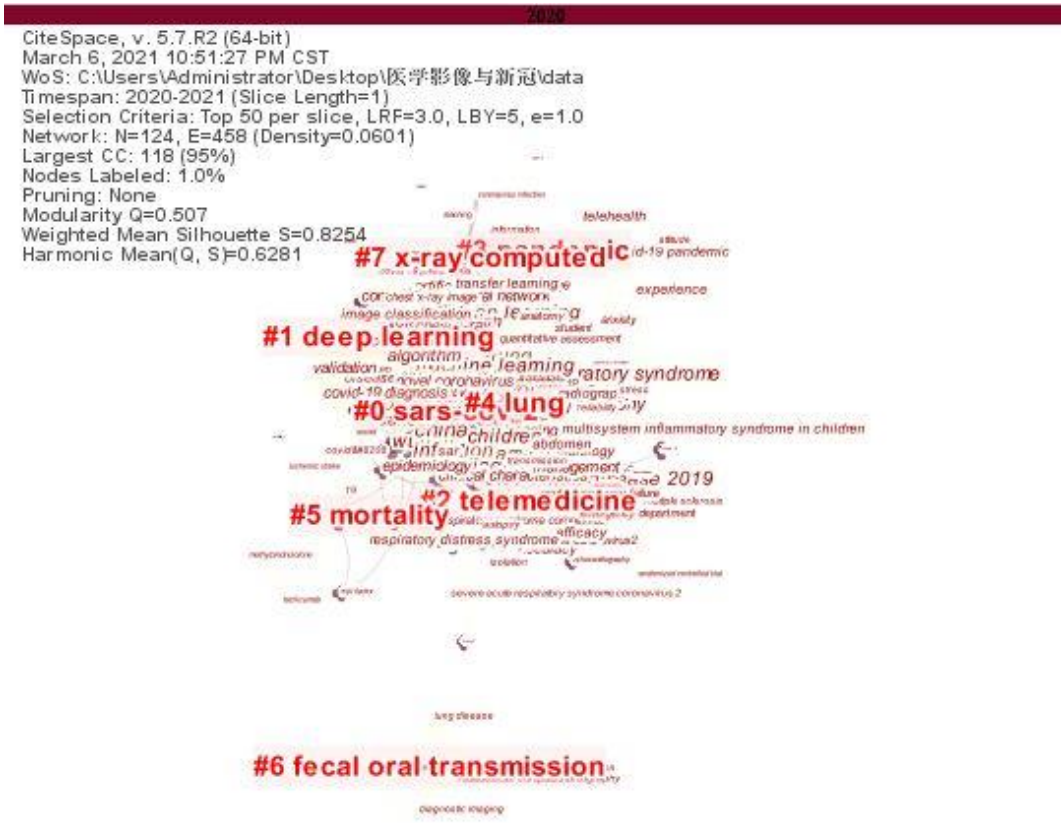


图 2 新型冠状病毒肺炎的医学影像学研究领域研究热点高频关键词聚类关系知识图谱

5 新型冠状病毒肺炎的医学影像学研究前沿

高影响力文献形成的网络，构成了学科的知识基础。高影响力论文施引文献的研究主题，构成了研究前沿。如果一篇文献同时引用两篇不同的文献，则认为这两篇文献存在共被引关系。文献共引分析是将一个研究领域的文献以研究方向来分类，这些被引文献则构成该领域的知识基础，高影响力文献形成的网络，构成了研究的知识基础。高影响力论文施引文献的研究主题，构成了研究前沿。进入 CiteSpace 软件（策略：阈值选择 top50），节点类型选择 Cited Reference，运行生成合并的知识基础和知识前沿的网络知识图谱，如图 3 所示。

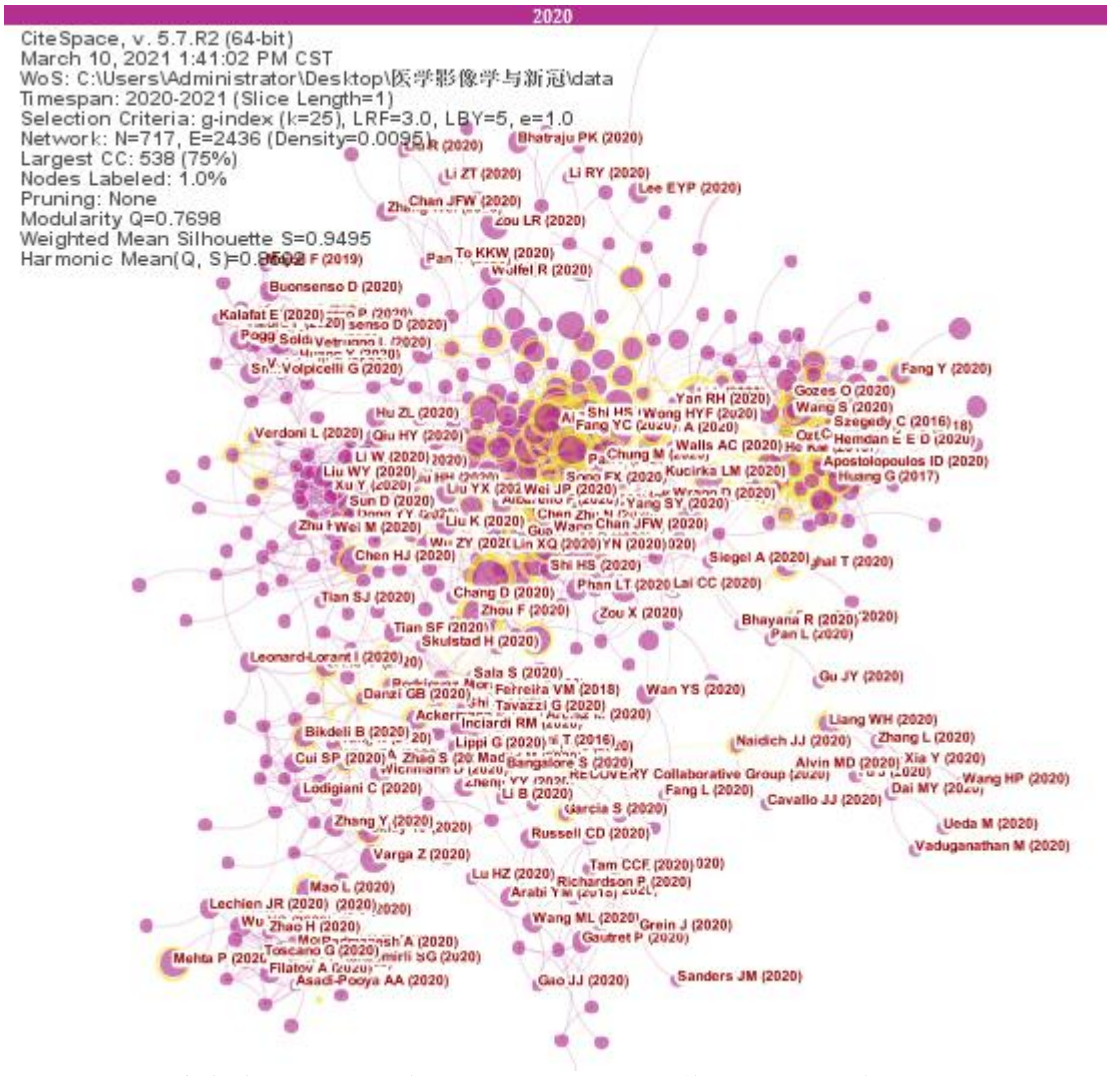


图 3 新型冠状病毒肺炎的医学影像学研究的知识基础及其研究前沿图谱

通过 keywords 聚类，得到新型冠状病毒肺炎背景下医学影像学的知识基础及其研究前沿聚类图谱和两种算法的研究前沿标识词，如图 4 和表 2 所示，聚类 13 个，其中较大的聚类 7 个，标识词采用 LSI 算法标签词。

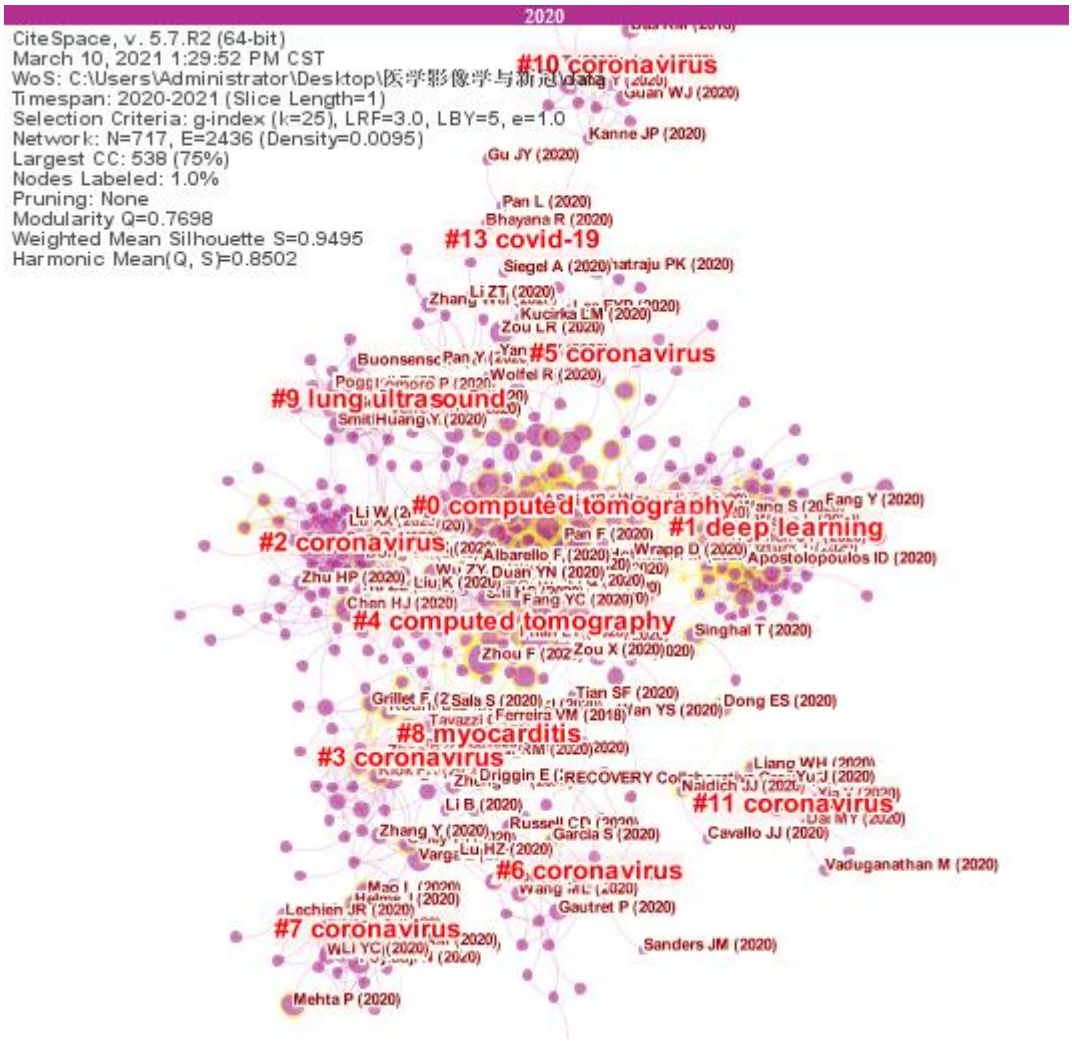


图 4 新型冠状病毒肺炎的医学影像学研究的知识基础及其研究前沿聚类图谱（采用 LSI 算法标签词）

表 2 两种算法的研究前沿标识词

类号	文献数量	聚类紧密程度	平均年份	LSI 算法标签词
0	8	0.932	2020	computed tomography; chest x-ray; radiographic technologies; quarantine station; predictive score; artificial intelligence; disease coronavirus disease; clinical data; computer-assisted quantification; patient outcome assessment | coronavirus; coronavirus infections; multidetector computed tomography; sub-pleural consolidation; ratio; ground glass opacification; high-resolution computed tomography; differential diagnosis; chest radiography; ultrasonography
1	7	0.971	2019	computed tomography; german; asymptomatic carrier; coronavirus infection; solitary pulmonary nodule; lung neoplasms; artificial intelligence; coronavirus pneumonia; covid-19 pneumonia; review | coronavirus disease; viral pneumonia; influenza; virus; interstitial pneumonia; covid-19 prognosis; acute respiratory distress syndrome; thoracic ultrasound; multimodality imaging; chest x-ray classification

2	7	0.866	2020	computed tomography; chest radiograph; german; image classification; artificial intelligence; thorax; steroid therapy; chest imaging; convolutional neural networks; diagnostic testing | coronavirus; radiologist; german; image classification; artificial intelligence; thorax; steroid therapy; chest imaging; convolutional neural networks; diagnostic testing
3	7	1	2020	coronavirus disease; myocardial injury; myocardial infarction; cardiac troponin; type; serological antibody detection; nucleic acid test; suspected case; cardiovascular dysfunction; immunosuppression | coronavirus; pseudovirus; covid-19 disease; asymptomatic carrier; prognostic factors; lung disease; predictive score; molecular imaging; epidemic characteristics; fetal death
4	5	0.914	2020	coronavirus; asymptomatic carrier; quarantine station; ace ii receptor; epidemiological history; lung disease; multimodal learning; epidemic characteristics; fetal death; artificial intelligence | coronavirus disease; lung ultrasound; ultrasonic imaging; critical care ultrasound; frequency control; volume quantification; computer tool; air bronchograms; quarantine station; acute respiratory disease
5	5	0.954	2020	computed tomography; coronavirus infections; sensitivity; specificity; virus; viral pneumonia; influenza; analysis; novel; coronavirus-infected pneumonia | coronavirus; coronavirus infection; syrian hamsters; respiratory distress syndrome; covid-19 pneumonia; prevention; risk factors; thin-section chest ct; nucleic acid; analysis
6	4	0.9	2020	coronavirus disease; computed tomography; artificial intelligence; biomedical imaging; prognostic implication; volumetric quantitative analysis; clinical data; berlin criteria; ischaemic stroke; steroid therapy | coronavirus; clinical features; pathogenesis; natural history; radiology images; emergency service; diagnosis; guidelines; vascular remodeling; pulmonary vascular diseases

6 研究结论与展望

通过基于 CiteSpaceII 信息分析软件的共现网络知识图谱分析和医学影像学临床医生对文献进行综合判读，我们将新型冠状病毒肺炎医学影像学研究热点与研究前沿的量化分析的进行知识图谱可视化方式展示，发现主要集中在四个方面：

6.1 人工智能技术在快速筛查方面的应用。基于 AI（人工智能技术）的高分辨率计算机断层扫描技术在医学影像学上的应用，快速完成病灶的自动检出、分割和精准量化，帮助医生实现对病例的医学影像学分型及分期的精准化判读，为后期的新型冠状病毒肺炎诊疗提供影像学上的指导和参考。同时，借助 AI 深度学习功能，将病例的影像样本数据与智能数据库的样本数据进比对，快速识别新冠肺炎影像的特征，提升病例的识别准确率[4]。

6.2 病灶变化趋势分析。利用构建医学影像学数据库，完成对全病程多期医学影像中的病灶变化（如全肺感染体积、实性密度）进行定量分析，对多期病灶变化状况进行对比分析，评估病情发展状况，预判患者的新型冠状病毒肺炎的发展趋势[5]。

6.3 病变筛查与诊断。通过对病变的总体情况和具体形态、密度及分布情况加以辅助性诊断和鉴别，及时发现病理改变，对比病灶变化的状况，高效精准的确定感染区域，形成可视化的诊疗评估参考，协助临床医生准确快速判断病情[6]。

6.4 整合建模的应用。完成对临床治疗、影像诊疗、实验室检查的三者整合建模，完成精准分层和对比，在整个治疗过程中为患者提供治疗效果评价以及病情发展过程的状况监测，实现临床医生在面对大量新型冠状病毒肺炎病况数据筛查过程的快速诊断，同时有效地避免漏诊情况[7]。

PU-111

基于知识图谱的高血压及肾病共病挖掘

洪怡敏[1] 刘金林[1] 张晗[1]
[1]中国医科大学健康管理学院

1 引言

近年来，随着生活节奏的加快，高血压患病人数逐年增加。高血压是一种常见的慢性病，在高血压持续 5-10 年后或出现肾脏的病理改变，而后出现蛋白尿、肾功损害等临床表现。肾损伤作为高血压常见并发症，可导致患者的肾功能及结构受损，且较多损伤不可逆转[1]。当患者出现高血压和肾病共病时，会增加其他疾病的发病风险，对患者的身心健康和生活质量造成严重的影响，甚至危害患者生命[2]。共病通常是指同一个体在某一特定时间段内，存在两种及以上疾病，且多以慢性病为主。王浩等[3]研究表明，2018 年我国中老年人的慢性病共病发生率为 55.77%。因此，对高血压和肾病的共病情况进行研究和管理具有重要意义。

伴随着计算机、互联网技术的发展，知识图谱在生物医学领域发挥着重要作用。知识图谱是以关系网络的方式来记录和表达知识体系中的概念、实体及其关系的知识集合[4]。在医学领域，知识图谱可以对于疾病诊断[5]与药物开发[6]提供帮助。羊艳玲[7]收集古今医案云平台软件中高血压疾病相关医案进行命名实体识别，构建高血压中医医案知识图谱并进行链路预测，为临床辅助诊疗提供决策依据。Kuan 等[8]利用英国住院患者护理数据集的医院发病统计数据构建不同种族、性别、年龄的患者发生原发性乳腺癌、哮喘、双相情感障碍相关共病的知识图谱，研究共病情况。

在构建知识图谱的过程中，知识抽取是非常关键的步骤之一[9]。知识抽取的方法包括基于传统规则的方法、基于机器学习的方法和基于深度学习的方法[10]。近年来，深度学习已经在自然语言处理，图像处理，语音识别，视频识别等方面取得了很好的表现[11]。相比于基于传统规则以及机器学习的方法，基于深度学习方法的知识抽取能获得更好的效果。

故本研究采用深度学习的知识抽取方法构建高血压及肾病共病的知识图谱，同时挖掘与高血压及肾病共同的并发症，了解高血压及肾病的关联，为高血压肾病的疾病推断提供依据，维护和推动全民健康。

2 知识抽取模型构建

本文的研究方法如图 1 所示，首先从 PubMed 数据库中获取有关高血压及肾病文献的标题及摘要，利用深度学习模型 BioBERT-Linear 抽取实体和关系，并利用规则生成三元组。对实体进行消歧后构建高血压及肾病共病的知识图谱。

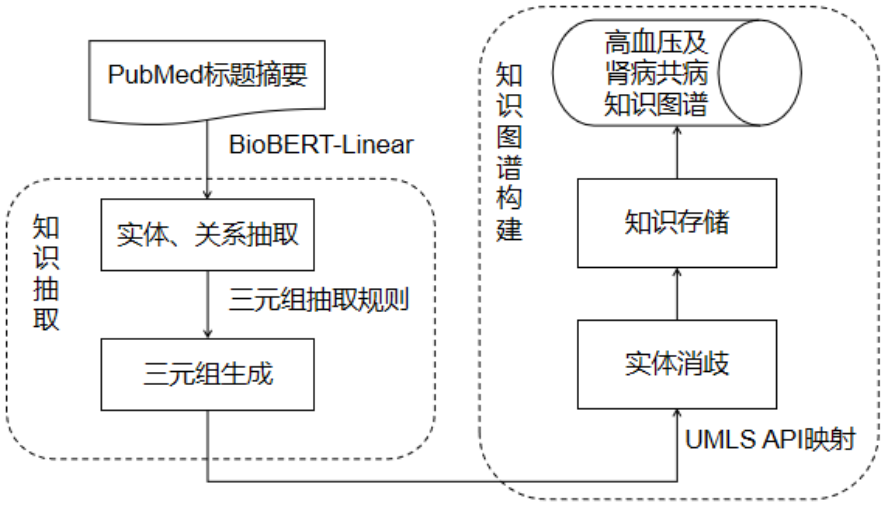


图 1 高血压肾病共病知识图谱构建流程

2.1 实体关系抽取

本研究利用深度学习模型同时抽取实体和关系，将抽取出的实体词语及关系词语按照预先定义的有逻辑的抽取规则进行三元组提取。共定义 3 种实体类型如表 1 所示，分别为疾病实体（包含疾病或症状）；化学物质实体（包含化学和药物）；生物靶点实体（包含基因或蛋白质）。共定义 4 种关系类型如表 2 所示，分别为“治疗”、“引起”、“并发”和“相关”。

表 1 实体示例

实体类型	示例
疾病	metabolic syndrome；hyperlipidemia；insulin resistance
化学物质	metformin；estradiol；acrylamide
生物靶点	interleukin-6；adiponectin；GLUT4

表 2 关系示例

关系类型	示例
治疗	treat；therapy in；therapeutic option for
引起	cause；lead to；risk factor for
并发	comorbidity of；common in；occur in
相关	associated with；influence on；relationship between

本研究利用 BioBERT-Linear 的模型架构如图 2 所示。BioBERT 是在 BERT 基础上采用 PMC 和 PubMed 语料库进行大量生物医学领域训练的基于 Transformer 模型搭建的知识抽取模型[12]，在 NER 任务中有较为优异的效果。为了表达词间的位置关系，BioBERT 加入了 Token Embeddings、Segment Embeddings 以及 Position Embeddings 三种嵌入，将以上三种 Embedding 的结果拼接作为下一步的输入。然后输入到 Transformer 层通过自制注意力机制获取上下文信息以及各个词语之间的关系。

Dropout 在训练中随机对神经网络结构中的节点进行隐藏[13]，使模型在训练过程中不会太依赖于某些特定特征[14]。避免出现过拟合现象，提高模型的泛化能力。本研究将 dropout 设为 0.5，即每次训练节点隐藏比例为 50%。

Linear 层为全连接层，实现对标签进行分类。Linear 层所有的神经元都与上一层的所有神经元连接[15]。在该层中，每个神经元都接收到了上一层所有神经元的输出，并使用这些输入计算出输出。

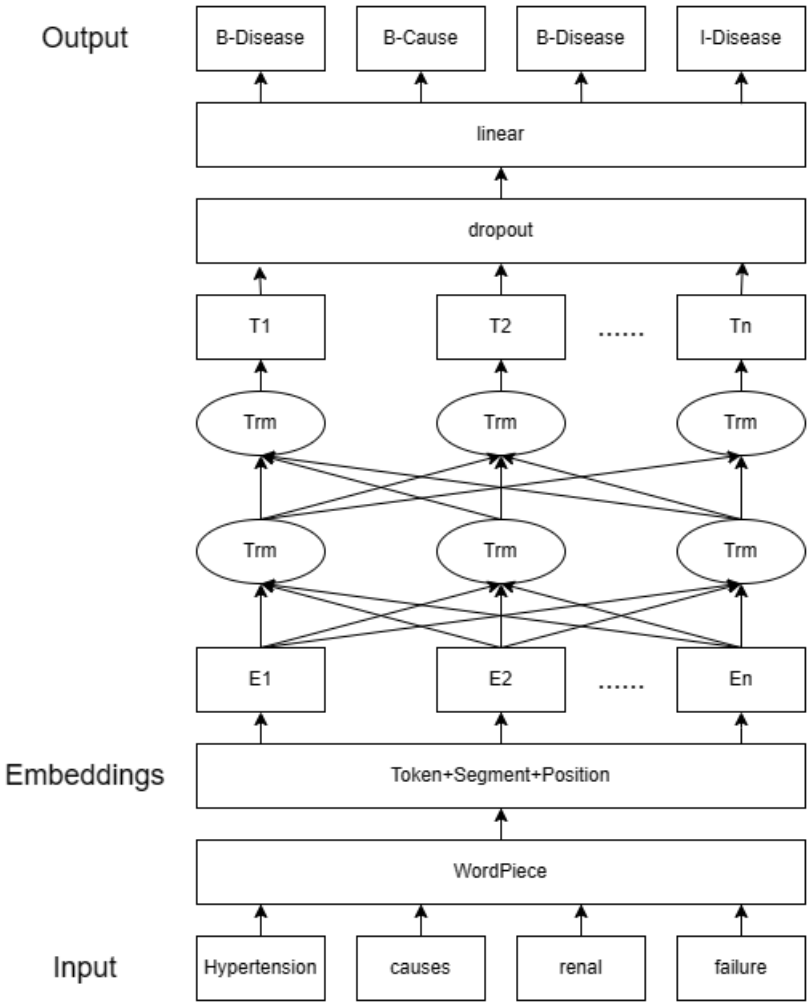


图 2 BioBERT-Linear 模型架构图

2.2 基于规则的三元组生成

规则也叫逻辑规则，是一种限制或约束，一般由专家进行手工定义[16]。通过规则可以对头实体和尾实体的位置以及它们之间的关系进行约束，从而生成“实体—关系—实体”三元组。本研究在规则的制约下提取句中具有逻辑与意义的三元组，共制定了 13 条规则，部分如表 3 所示。

表 3 抽取规则（部分）

头实体类型	关系类型	尾实体类型
化学物质	相关	生物靶点
化学物质	相关	化学物质
化学物质	引起	疾病
疾病	并发	疾病
化学物质	治疗	疾病
...	...	...

2.3 实体消歧

将获得的三元组进行实体标准化，可以将实体从自然语言文本映射到标准术语集上，给出其在受控词表中的唯一编码[17]。本研究旨在对共病网络进行研究，故对疾病类实体进行标准化，利用

UMLS 中的 API 将实体映射到 UMLS 术语表中的规范概念，保证了后续所构建的知识图谱中术语的一致性。

3 实验及结果

3.1 数据来源

检索 Pubmed 近 5 年以来有关高血压和肾病的文献，分别获得 28,025 和 61,894 篇文献，具体检索式如下：

Hypertension[Majr] AND Humans[Filters] AND 2019:2023[PDAT]

Kidney diseases[Majr] AND Humans[Filters] AND 2019:2023[PDAT]

3.2 知识抽取结果

本研究选取 PubMed 数据库中有关高血压和肾病的文献 1473 篇（23300 个句子）进行标注，内容涵盖高血压、肾病、糖尿病的并发症和症状、基因、药物等文献。对标题以及摘要中出现的概念实体和关系进行 BIO 标注，其中“B”表示实体或关系的第一个单词、用“I”表示该实体或关系剩余的词、用“O”表示与实体和关系无关的词。将标注后的数据集以 8:1:1 的比例划分训练集、测试集和验证集进行模型训练。模型实验涉及到的参数如表 4 所示。从知识抽取的结果中随机抽取 200 个句子进行模型效果评价，如表 5 所示。模型对疾病、“治疗”的抽取效果最好。相比于关系，模型对实体的抽取效果更好，可能是因为标注数据中关系类数量较少[18]，模型训练不足。或组成该关系的谓词语法较复杂，模型不能很好地学习[19]。

表 4 模型实验参数表

参数	值
max_len	75
batch-size	32
learning rate	3e-5
eps	1e-8
epochs	15
max_grad_norm	1.0

表 5 BioBERT-Linear 对不同实体及关系识别结果

类别	Precision	Recall	F1-score
疾病	93.4%	93.0%	93.2%
化学物质	80.0%	76.9%	78.4%
生物靶点	87.8%	82.7%	85.2%
治疗	100%	87.5%	93.3%
引起	75.0%	100%	85.7%
并发	33.3%	50.0%	40.0%
相关	88.6%	81.6%	85.0%
总体	89.0%	88.1%	88.5%

从 PubMed 数据库中获取高血压文献 28025 篇、肾病文献 61894 篇。利用模型进行知识抽取并依据规则抽取出高血压三元组 238568 条、肾病三元组 616936 条，其中涵盖的各类实体数量及三元组数量分别见表 6 和表 7。

表 6 实体数量分布

实体类型	高血压实体数量（%）	肾病实体数量（%）
疾病	3165（36.6%）	5981（29.6%）
化学物质	1828（21.1%）	4090（20.2%）
生物靶点	3665（42.3%）	10145（50.2%）
总计	8658	20216

表 7 三元组数量分布

关系类型	高血压数量（%）	肾病数量（%）
治疗	38807（16.3%）	23945（3.9%）
引起	5673（2.4%）	96171（15.6%）
并发	15670（6.5%）	34616（5.6%）
相关	178418（74.8%）	462204（74.9%）
总计	238568	616936

3.2 图谱结果展示

本研究共病指同一个体在某一特定时间段内，存在两种及以上疾病的情况，即一人多病的现象，所以本研究将关系类型限定为“并发”和“相关”。将高血压三元组中频次大于 80 以及肾病三元组中频次大于 95 的知识关联保留。最终构建了高血压及肾病共病的知识图谱。

3.2.1 高血压共病知识图谱

图 3 展示了与高血压共病且频次大于 80 的疾病，红色节点的实体频次大于 300，橙色节点的频次在 150-300 之间。蓝色节点的频次在 80-150 之间。可知高血压常见的并发症有非胰岛素依赖型糖尿病、肥胖、脂血异常、慢性肾脏病、脑血管意外等。

图 3 高血压共病知识图谱

3.2.2 肾病共病知识图谱

图 4 展示了频次大于 95 的与肾病共病的疾病。其中，红色节点频次大于 300；橙色节点实体频次在 150-300 之间；蓝色节点的实体频次在 95-150 之间。从图中可以看出感染、新型冠状病毒肺炎、脓毒症经常与急性肾损伤共同发生。糖尿病肾病、动脉粥样硬化、非酒精性脂肪肝、经常与慢性肾脏病共同发生。蛋白尿、炎症、高血压既与急性肾损伤又与慢性肾脏病共病。

图 4 肾病共病知识图谱

3.2.3 高血压及肾病共患病

图 5 展现了频次大于 40 的高血压及肾病的共同共患病，其中黄色节点表示高血压；蓝色节点表示肾病；粉色节点表示高血压及肾病的共同共患病。可以发现与高血压和肾病共同共患病包括 4 类：①代谢性疾病：非胰岛素依赖型糖尿病、肥胖、高尿酸血症以及胰岛素抵抗；②心血管疾病：心力衰竭、心房颤动以及冠状动脉疾病；③神经系统疾病：脑血管意外以及认知障碍；④其他：抑郁、新型冠状病毒肺炎、炎症以及蛋白尿。

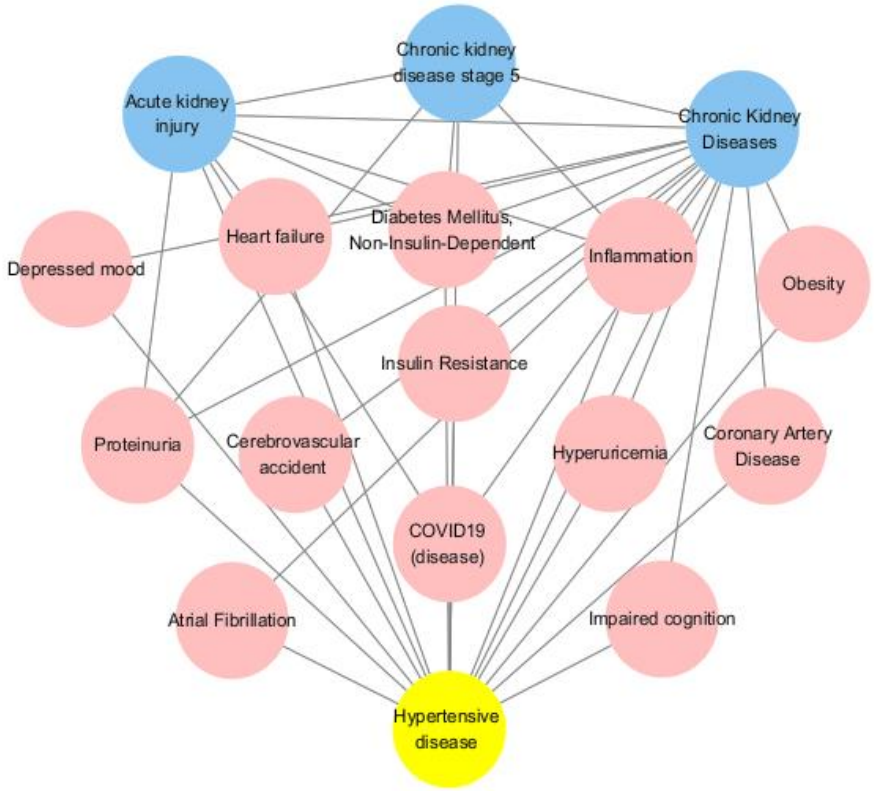


图 5 高血压及肾病共同共患病知识图谱

目前已有研究表明慢性肾脏病的发生和发展与非胰岛素依赖型糖尿病[20, 21]、肥胖、胰岛素抵抗和高尿酸血症[22]密切相关。D'Elia L 等[23]研究发现超重或明显的肥胖、胰岛素抵抗可以通过增加肾素-血管紧张素-醛固酮系统兴奋性、交感神经活性以及盐敏感性参与高血压的发生。

此外已有研究证实了心血管疾病与高血压和肾病的相关性，例如中国高血压防治指南报道了血压水平与心血管风险呈正相关关系[24]。刘毕君[25]的研究表明慢性肾脏病患者心血管疾病的发病率会相对较高，因为肾脏疾病时会出现各种代谢紊乱，如脂质代谢异常等，是心血管疾病的高危因素。

包括脑血管意外及认知障碍的神经系统疾病也已被证实与高血压及肾病相关。长期的高血压将会导致患者的脑动脉管腔变细，当患者的血压急剧上升时，其脑血管将会突然破裂导致脑血管意外[26]。王昀暄等[27]研究表明高血压导致认知障碍的发生发展与血压异常导致大脑皮层结构的损伤有关。

高血压及肾病还与抑郁、感染等多种疾病相关，陈玮等[28]研究显示慢性肾病患者病情越重，患者出现抑郁状态评分的比率越高。De Vos AC 等[29]研究表明抑郁可激活肾素-血管紧张素-醛固酮系统，使血管紧张素Ⅱ及醛固酮的分泌增加，引起血压升高，同时血管紧张素Ⅱ也会加重抑郁的症状。Hassanein 等[30]研究表示许多慢性肾脏病患者有糖尿病和高血压等多种合并症，这些合并症可使其易患新型冠状病毒肺炎。

总之，高血压合并肾脏疾病可能出现多种并发症，医务人员应当根据不同情况给出相应的治疗措施，降低患者不良结局的风险，保证患者的生命质量。

4 结语

本研究利用深度学习与基于规则相结合的方法进行知识抽取，构建了高血压及肾病共病的知识图谱，通过对知识图谱进行分析发现了高血压常见的并发症、肾病常见的并发症以及高血压与肾病的共患病。

构建知识图谱可以清楚地挖掘两种疾病的共病情况，使患者直观地了解相关共病，维护和推动全民健康。并为高血压以及肾病的疾病推断提供重要依据，为临床医生提供了一定的参考价值，具有一定的实用性。未来将对此模型进行优化并比较不同模型之间的结果，以提升命名实体识别和关

系抽取的准确率。此外，基于文献的知识图谱覆盖面较小，未来可以扩充数据来源并与多种类型数据源相结合，不断完善知识图谱。

PU-112

生物信息学分析非小细胞肺癌差异表达关键基因的研究

齐荣暄[1,2]，于 琦[1,2*]
1.山西医科大学管理学院
2. 临床决策研究大数据山西省重点实验室

据国家癌症中心报道，肺癌是我国发生率和死亡率均位列第一位的恶性肿瘤,无论在发达国家还是发展中国家，肺癌都是癌症死亡的主要原因。肺癌的病理学类型主要分为小细胞肺癌(SCLC)和非小细胞肺癌(NSCLC)两类，后者是主要的类型[1]。尽管肺癌在诊疗技术上有了很大的进展，但肺癌的预后仍不满意，5 年生存率小于 15%， 肺癌的发生发展涉及多因素、多步骤、多基因的复杂生物学过程，除了环境因素作用，肺癌的基因学说也越来越受重视。NSCLC 的发生与遗传因素和环境因素均相关，是一种复杂性疾病[2]。随着精准医学的发展，从分子水平阐明疾病发生发展相关的基因改变将有利于疾病的诊断、治疗和预后判断。在过去的几十年，对肺癌发生发展的分子生物学认识有了很大的提高，然而目前临床可用的基因数量有限，仅有 EGFR、ALK、 ROS1、BRAF、RET 和 C-MET 等几个基因，而且主要集中于肺腺癌中[3]。在肿瘤发生发展过程中，有大量伴随基因参与癌基因扩增过程，但是绝大多数并不是关键基因[4] 。而且 NSCLC 发病机制依然尚未明确，因此从基因组水平挖掘与 NSCLC 相关的基因改变将有可能对其诊断以及预后判断提供更多的分子标记[5]。本文运用生物信息

学的方法筛选出与 NSCLC 差异性表达有关的关键基因，可以为诊断和针对性治疗 NSCLC 提供一种新的思路。

方法

1.7 数据获取与预处理

从 GEO 数据库数据中下载了 GSE33532 芯片数据集和 GSE63113 芯片数据集，在进行研究之前，首先运用 R 语言中的 limma 包和 sva 包对 GSE33532 芯片数据集和 GSE63113 芯片数据集的基因取 unique，得到唯一基因，获取出两个芯片数据集的交集基因，之后在合并数据时，对重复的基因数据取平均值，只保留一行，最终对得到的基因进行批次校正，并将校正后的数据输出到一个文本文档中。

1.2 差异表达基因的筛选

DEGs 是通过利用统计学的方法来筛选出 NSCLC 对照组与 NSCLC 实验组的 DEGs。使用 R 语言的 limma 包，对所获取的样本合并基因中删除表达量为 0 的基因，并在此基础上得到对照组和实验组基因，在 NSCLC 对照组和 NSCLC 实验组之间构建矩阵和比对模型，平行比较基因在对照组和实验组中的表达量。随后使用 lmFit 函数进行线性模型拟合，接着进行贝叶斯检验，经由 fdr 校正，筛选出|log2FC|=2 且 adj.P.Val=0.05 的基因作为两组值的 DEGs,并输出 DEGs 的表达量。

1.3 功能富集分析

基因本体富集分析（Gene Ontology,GO）：GO 包括生物学过程（Biological Process,BP）、细胞定位（Cellular Component,CC）和分子功能（Molecular Function, MF）,各自描述了基因产物可能行使的分子功能、所处的细胞环境、以及参与的生物学过程。

使用 David 在线工具对 DEGs 进行 Goterm 富集度统计学分析，从而得到 DEGs 的 pVal（标准为：pVal<0.05），进而可以得到 DEGs 富集具有显著性的 Goterm。

京都基因与基因组百科全书通路富集分析（Kyoto Encyclopedia of Genes and Genomes，KEGG）：KEGG 信号通路图可以把输入的 DEGs 对照找到其所在的代谢通路并且标注出来。

1.4 蛋白互作网络的构建

蛋白质-蛋白质相互作用（protein-protein interaction，PPI）：是指在生物信号传递、基因表达调节、能量和物质代谢和细胞周期调节的各环节中，通过不同的分子间的交互作用形成的。从中可以发现高分模块，能够从分子层面把控疾病发生发展的机制，也能进而发现新药靶点。本文中利

用 STRING 在线工具对上文所得到的差异基因进行 PPI 网络的制作，将其中的与整个网络无关的作用节点剔除，最后生成蛋白互作网络。

1.5 蛋白互作网络图分析

由于从 STRING 在线数据库得到的 PPI network 也较为复杂。为了方便后续的研究，这里将对生物分子交互起决定性作用的核心基因进行筛选。

Cytoscape 是以一个网络为中心的，每一个结点都代表着不同的生物分子，而在结点间的线条则是它们的交互作用。接着使用其中的 cytoHubba 插件来查看其中得分位于前十的 Hub 基因，这些 Hub 基因就是蛋白互作网络的关键，可能与 NSCLC 发生转移密切相关。

1.6 关键基因的生存分析

运用 GEPIA 在线工具对得到的关键基因进行生存分析，因为疾病类型中没有 NSCLC 这个疾病，而肺腺癌（Lung adenocarcinoma，LUAD）属于 NSCLC ，所以在对关键基因进行生存分析时，选择疾病类型为 LUAD，点击选择总体生存期（Overall Survival ，OS）、组截止为 Median、风险比率为 yes、置信区间为 95%、坐标轴单位为 Months，即可以得到关键基因的生存分析，并将生存分析图下载保存。

研究结果的分析与解读

2.1 数据标准化处理与差异表达基因的筛选

通过使用 R 语言的 limma 包进行 DEGs 的筛选，条件为|log2FC|=2 且 adj.p.Value<0.05，最终共得到 321 个 DEGs，其中上调基因 96 个，下调基因 225 个，并绘制出了对照组和实验组比较的差异基因热图(见图 1)，并在此基础上绘制出了 DEGs 的火山图（见图 2）。在图 3- 1 中，红色代表高表达，蓝色代表低表达。(如 WIF1 在对照组中高表达，而在实验组中低表达)。

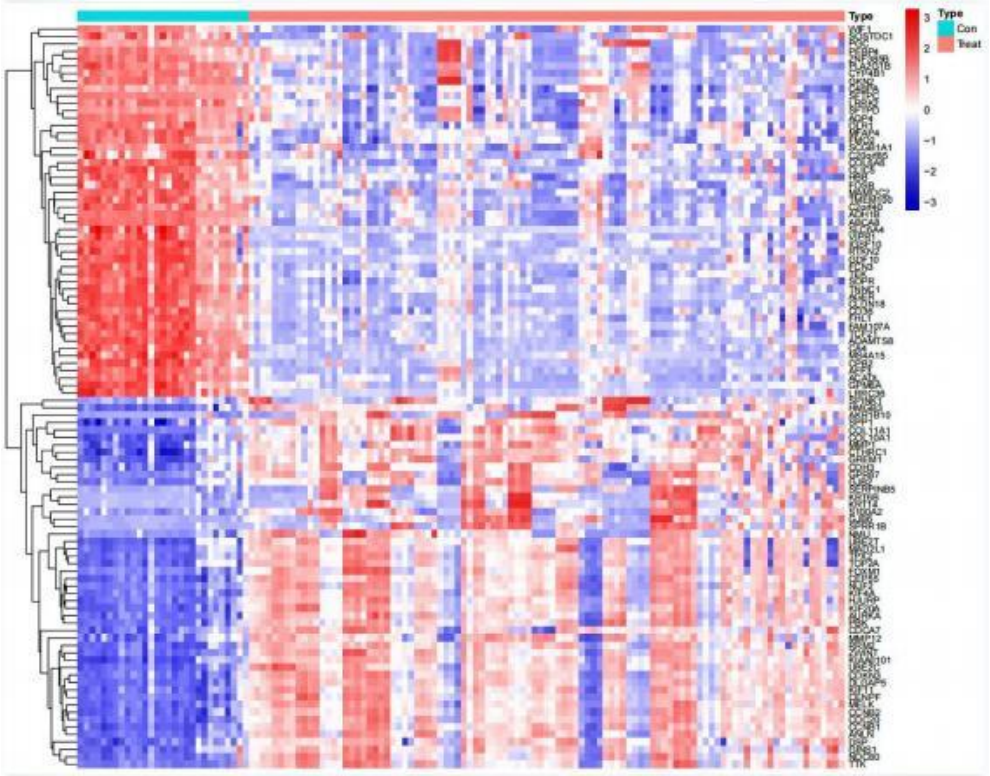


图 1 差异基因热图

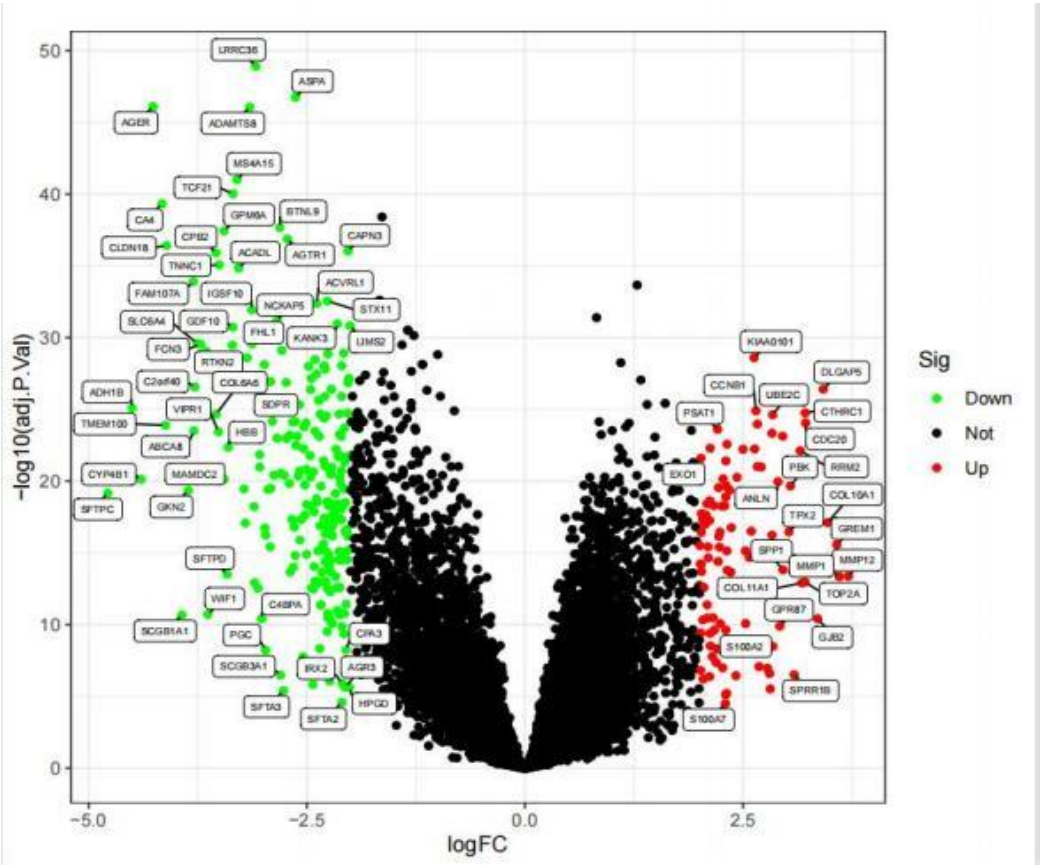


图 2 差异基因火山图

2.2 差异表达基因 GO 富集和 KEGG 富集分析

2.2.1 差异表达基因 GO 富集分析

使用 David 在线工具对筛选出的 DEGs 做 GO 功能富集分析，结果发现 DEGs 中有 287 个条目呈显著富集性(pValue＜0.05)。通过使用 R 语言做出部分 GO 功能富集气泡图（见图 3）,以及 GO 功能富集柱状图（见图 4）。对于 BP，主要参与细胞间粘附、细胞分裂和血管生成等生物学过程；对于 CC，其主要表达在胞外区、细胞外间隙、原生质膜等；对于 MF，其主要参与相同的蛋白质结合、蛋白结合和钙离子结合等生物学过程。

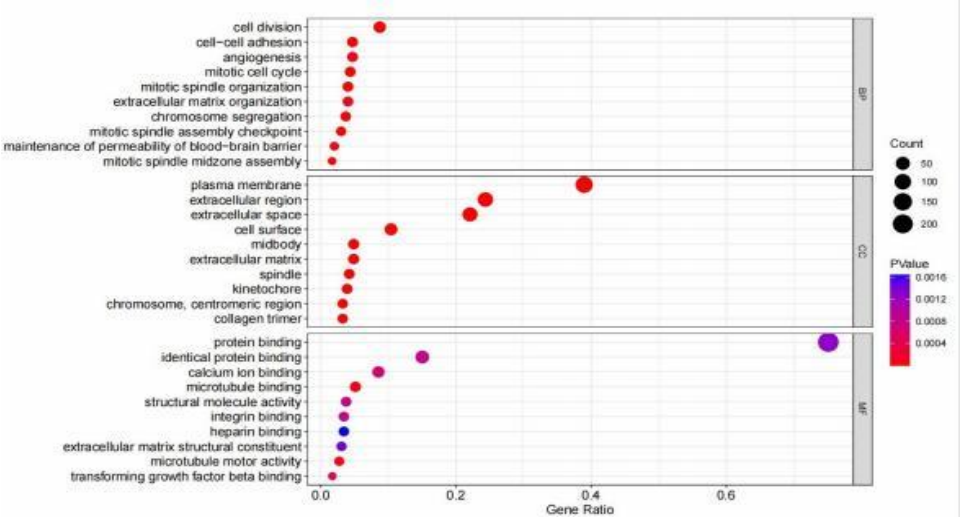


图 3 GO 功能富集气泡图

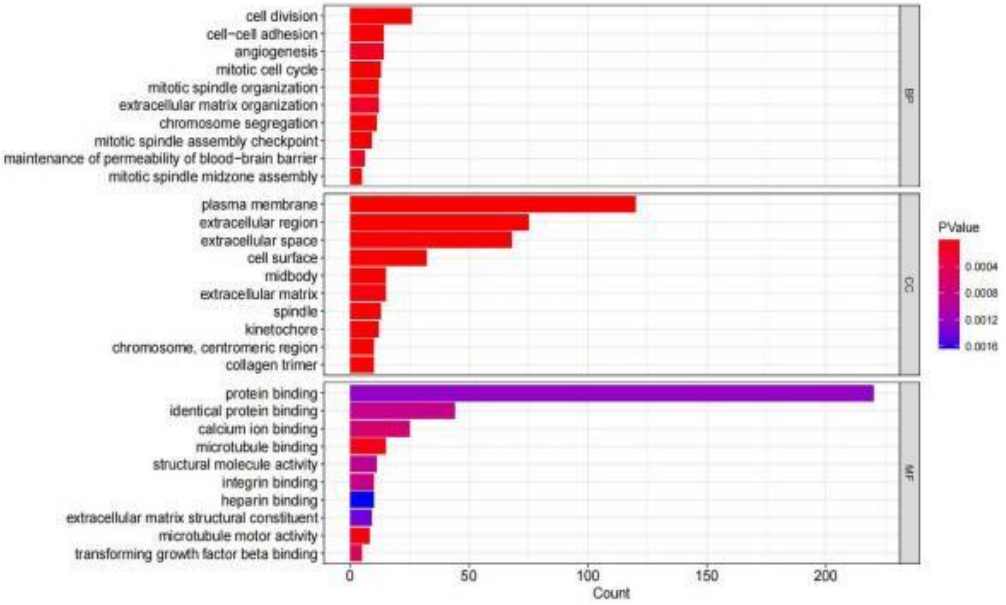


图 4 GO 功能富集柱状图

2.2.2 差异表达基因 KEGG 富集分析

使用 David 在线工具对 DEGs 进行 KEGG 通路富集分析，结果发现 DEGs 中有 11 个通路呈显著富集性（pValue＜0.05）。KEGG 通路富集气泡图（见图 5），及柱状图（见图 6）可视化 DEGs 在各个通道上的分布:富集显著性由高到低依次为细胞粘附分子、ECM-受体相互作用、细胞周期、蛋白质 消化和吸收、白细胞经内皮细胞迁移、PPAR 信号通路、补体和凝血级联、p53 信号通路、牛磺酸和次牛磺酸代谢信号通路。

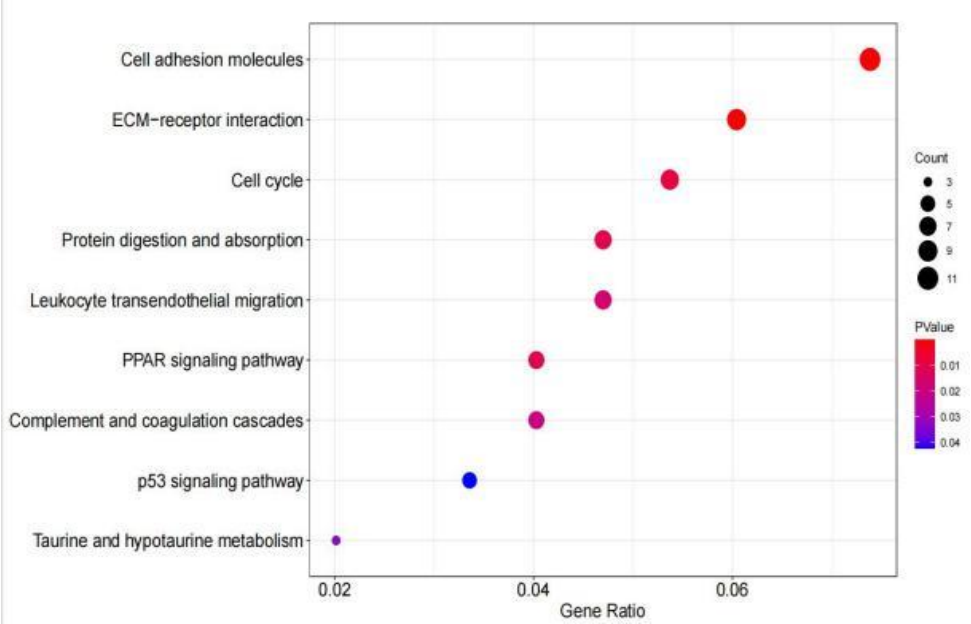


图 5 KEGG 功能富集气泡图

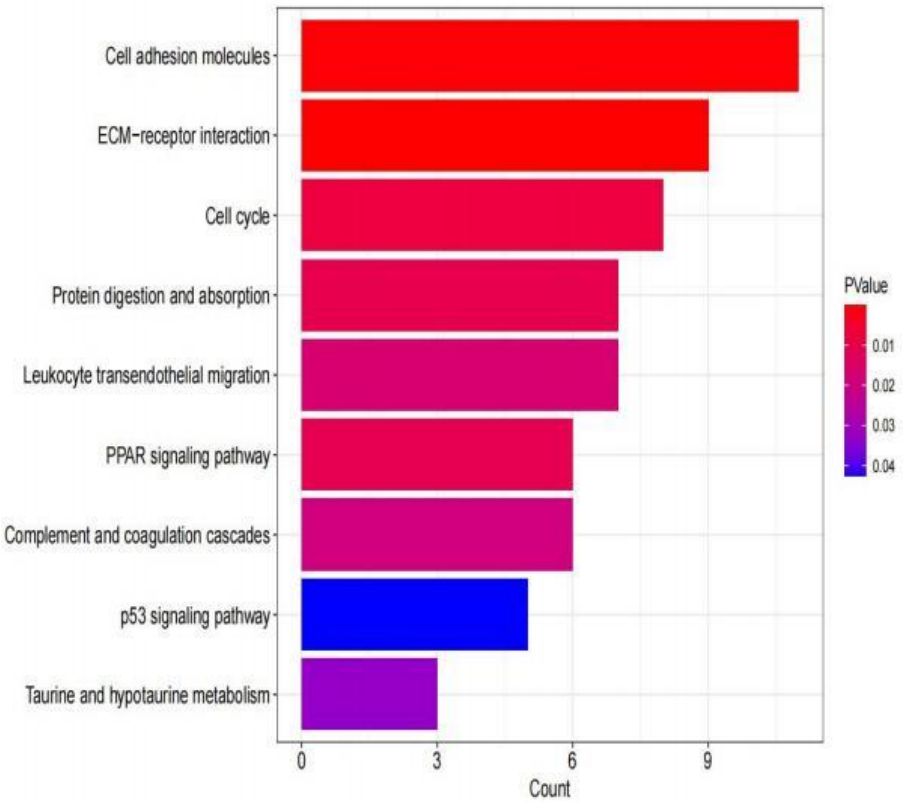


图 6 KEGG 功能富集柱状图

2.3 蛋白互作网络(PPI 网络)构建及关键基因筛选

将经过筛选得到的 321 个 DEGs 导入 STRING 在线工具制作出 DEGs 的 PPInetwork（见图 7）。通过 Cytoscape 对不同蛋白的交互作用进行了研究，从中找出了关键基因。利用 cytoHubba 插件，以 MCC 标准筛选出十个核心的 DEGs 以及邻接节点（见图 8），颜色越红则表示这些基因为 DEGs 蛋白互作网络的关键枢纽基因。

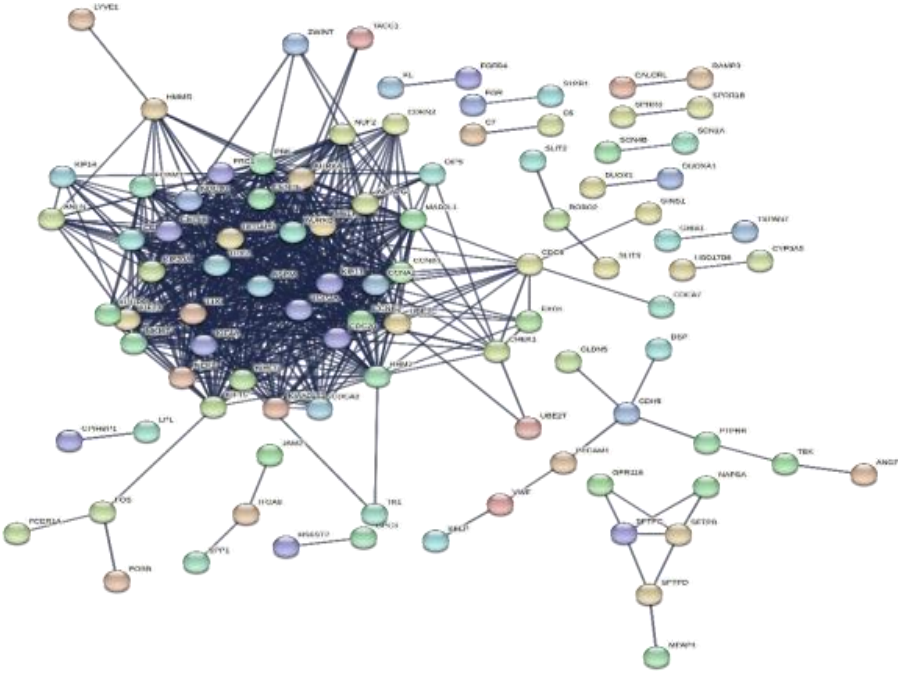


图 7 DEGs 的 PPI network

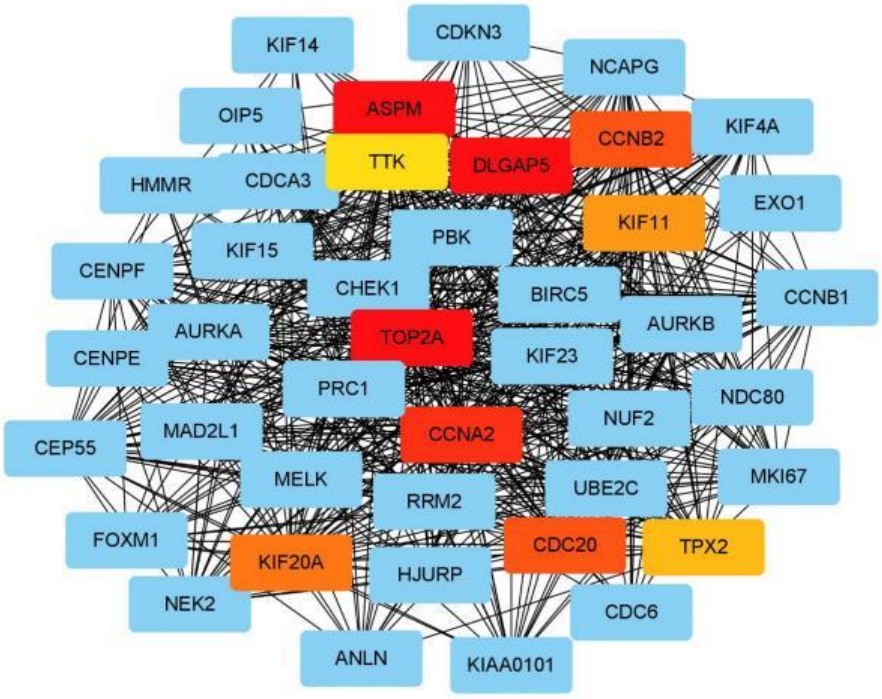
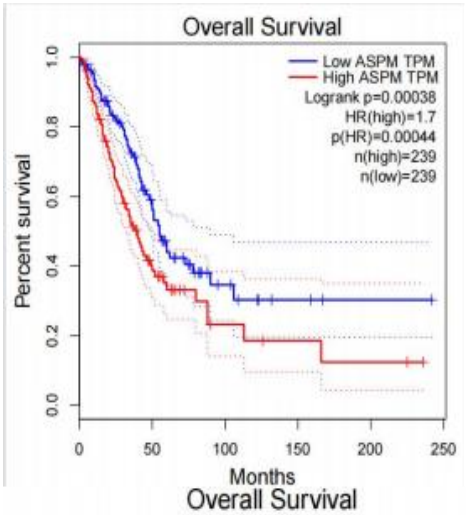
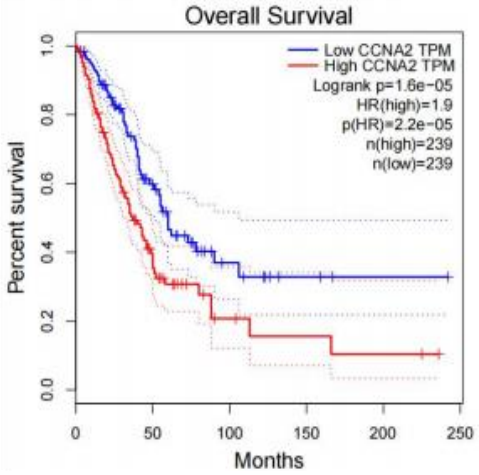


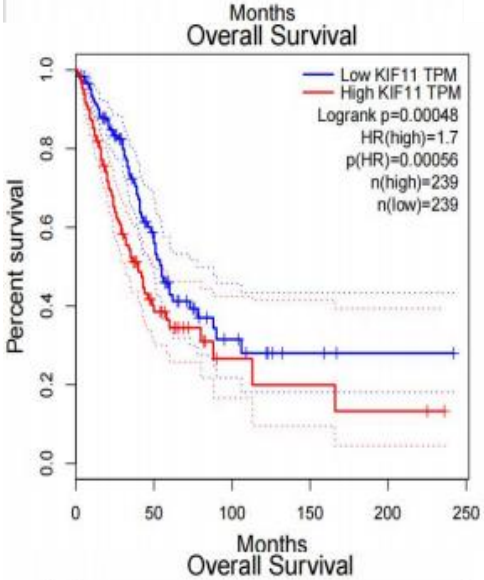
图 8 DEGs 的 PPI network 的前十个关键枢纽基因

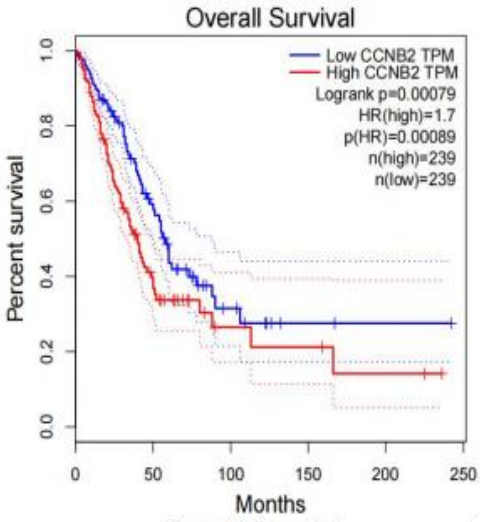
2.4 对差异表达基因进行生存分析

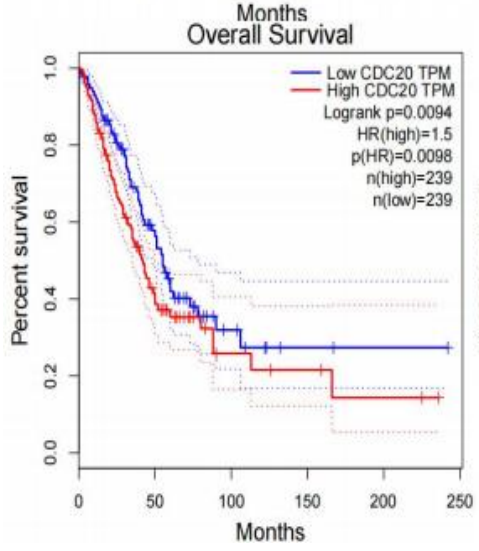
运用 GEPIA 在线工具对上述得到的 10 个差异基因分别进行生存分析，结果如图 9 所示。结果表明本文鉴别的十个与 NSCLC 相关的 DEGs 高表达与肺癌人群的总体存活时间减少存在显著的相关性(P＜0.05)。

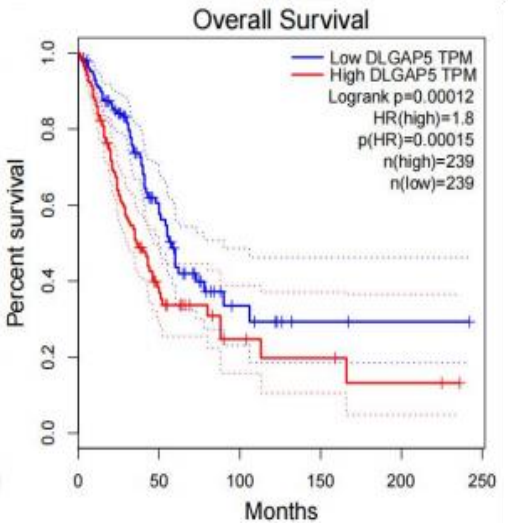

Overall Survival
Low ASPM TPM
High ASPM TPM
Logrank p=0.00038
HR(high)=1.7
p(HR)=0.00044
n(high)=239
n(low)=239
Percent survival
Months


Overall Survival
Low CCNA2 TPM
High CCNA2 TPM
Logrank p=1.6e−05
HR(high)=1.9
p(HR)=2.2e−05
n(high)=239
n(low)=239
Percent survival
Months


Overall Survival
Low KIF11 TPM
High KIF11 TPM
Logrank p=0.00048
HR(high)=1.7
p(HR)=0.00056
n(high)=239
n(low)=239
Percent survival
Months


Overall Survival
Low CCNB2 TPM
High CCNB2 TPM
Logrank p=0.00079
HR(high)=1.7
p(HR)=0.00089
n(high)=239
n(low)=239
Percent survival
Months


Overall Survival
Low CDC20 TPM
High CDC20 TPM
Logrank p=0.0094
HR(high)=1.5
p(HR)=0.0098
n(high)=239
n(low)=239
Percent survival
Months


Overall Survival
Low DLGAP5 TPM
High DLGAP5 TPM
Logrank p=0.00012
HR(high)=1.8
p(HR)=0.00015
n(high)=239
n(low)=239
Percent survival
Months

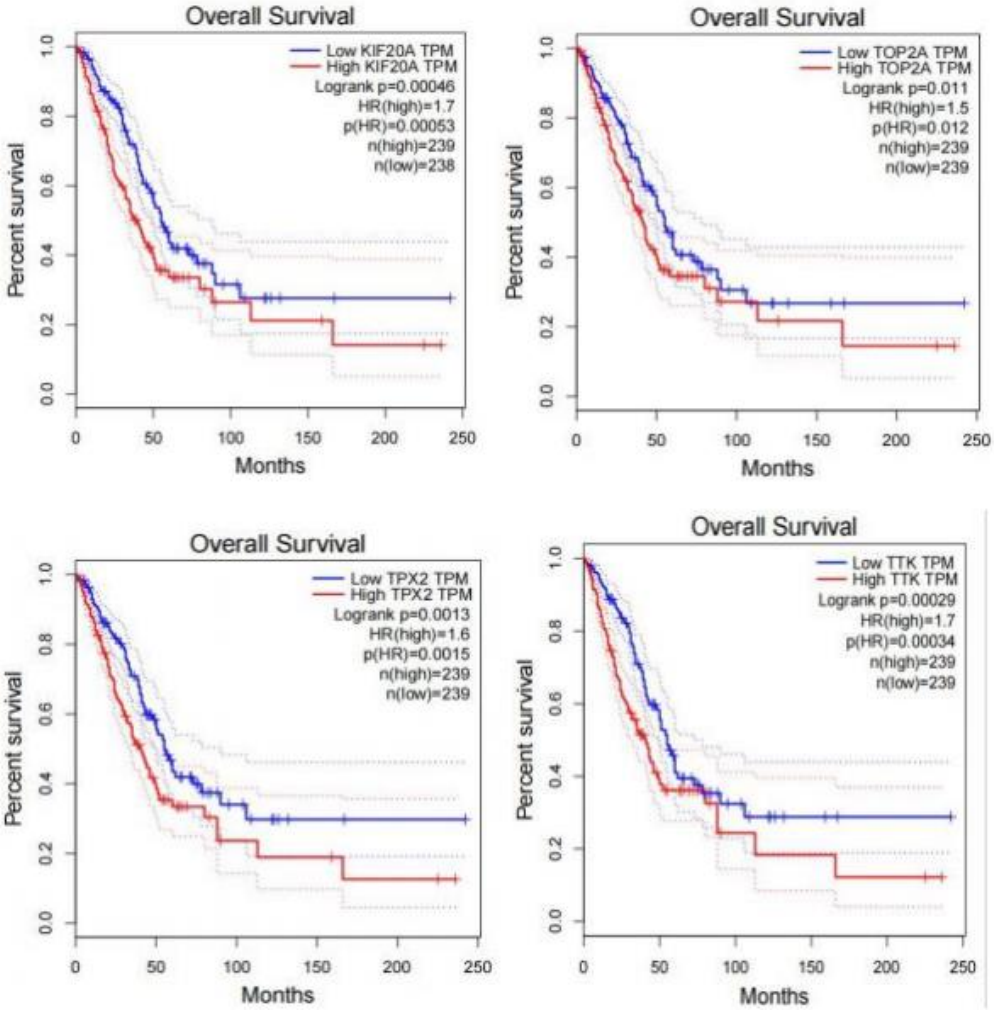


图 9 十个关键差异基因的生存分析结果

讨论

肺癌是一种发病率与致死率都很高的疾病,研究表明,每年约有 120 万人死于该病[6]。NSCLC 是肺癌的一种亚型，发病率占肺癌患者的 75%~80%[7]。几年来，虽然 NSCLC 的治疗方法不断改进，但其 5 年生存率依然很低，仅为 16%[8]。为探究 NSCLC 的发生机制，本文从 GEO 数据库中下载 NSCLC 对照组和实验组样本数据，随后运用生物信息学方法进行分析，包括 DEGs 筛选、GO 功能富集分析和 KEGG 通路分析、蛋白互作网络的构建与分析、关键基因的生存分析。最终筛选出的枢纽基因可能与 NSCLC 的差异化表达密切相关。

通过对芯片 GSE33532 和 GSE62113 的合并数据进行分析，结果共获得 321 个差异基因，发现 NSCLC 中上调基因数有 96 个，下调基因数有 225 个。GO 功能富集可以看出 DEGs 主要表达在胞外区、细胞外间隙、原生质膜等，同时参与细胞间粘附、细胞分裂、血管生成、相同的蛋白质结合、蛋白结合和钙离子结合等生物学过程。大量数据表明，NSCLC 的发生与有丝分裂纺锤体相关蛋白的调控有关[9]，并且 ASPM、DLGAP5、TOP2A、CDC20 、KIF20A、KIF11、TPX2、TTK 这八个差异基因均与细胞有丝分裂过程有关。而 KEGG 富集分析显示这些 DEGs 富集显著性由高到低依次为细胞粘附分子、ECM-受体相互作用、细胞周期、蛋白质消化和吸收、白细胞经内皮细胞迁移、PPAR 信号通路、牛磺酸和次牛磺酸代谢信号通路。在 NSCLC 发生过程中，细胞周期调节起着关键性作用。癌细胞分裂过程中常常伴有有丝分裂相关因子和蛋白的表达异常，细胞周期相关蛋白以及参与有丝分裂的成分的异常表达是肿瘤细胞增殖所必需的[10]。

综上所述，本文是基于生物信息学的方法从分子层面角度对 NSCLC 进行了分析，通过 DEGs 筛选、富集分析、关键基因的筛选和生存分析，最终得到十个关键基因，它们可能在 NSCLC 的发生发展中发挥重要作用，同时为进一步发现 NSCLC 潜在的生物标志物和生物学机制的研究提供新的线索，也为进一步确定 NSCLC 的诊断和治疗干预方法提供方向。

PU-113

融合多源数据与方法的疾病-基因-药物知识发现模型

郜杨芳，华国旻
山西医科大学 管理学院

1. 引言

一款药物从研发到被批准上市平均需要数十年时间和高额资金。但是，也只有不到 10%的药物能进入临床阶段，最终被批准上市[1]则是更小部分的药物。

随着大数据与人工智能（artificial intelligenc，AI）技术的发展，大批量的 AI 企业布局药物研发赛道，AI 在生物医药产业的各方面发挥着重要作用，并且在靶点发现等领域已经为企业带来收益。将信息学与医学、药学和生物信息学深度融合，采用大数据、人工智能为代表的先进方法与技术能够提升传统高通量筛选过程，缩短药物研发周期，降低药物研发成本，提升药物研发成功率[2]，因而成为目前生物医药领域进行药物研发的重要方向。

随着生物医学技术和科学研究的不断发展，研究成果不断涌现，相应地知识数据快速增长，包括科研文献、医药知识库、医学基础和临床数据等类型的生物医学大数据资源业已形成。运用计算机科学、生物学、信息学等各种技术,挖掘海量结构化或半结构化的多源数据中蕴含的知识和规律并揭示其意义，进而辅助生物医学领域的研究与应用，具有重要的作用与意义。

2. 基于信息学方法的医药知识发现研究现状

依据信息资源类型，医药领域基于信息学方法的知识发现的相关研究可分为两大类：基于文献的知识发现和数据的知识发现。

2.1. 基于文献的知识发现

基于文献的知识发现（Literature Based Knowledge Discovery，LBD），是一种发现跨领域知识之间隐性关联的特定类型文本挖掘方法。LBD 最初由 Swanson 提出,指从科技文献中以自动或半自动化的方式挖掘知识概念之间的隐性关联[3]，揭示尚未被发现的公共知识，用以制定或评估新的科学假设、辅助支持潜在的科学发现。目前医学领域的文本挖掘研究已经得到国内外学者的广泛关注。

已有的医学文本挖掘研究主要围绕医学实体识别和关系发现展开，从医学文本的有效利用与关系发现的角度出发，通过实体识别与信息抽取等技术，进行医学文本挖掘。在方法上多采用基于统计共现分析、基于语义关系分析、基于知识图谱分析和基于深度学习分析等方法[4]，对药物预测、共病关系、治疗前沿、靶点与疾病的关系等方面的研究起到较大的推动作用。总体上，现有基于文献的知识发现研究以知识概念的共现关联为分析对象，对概念及概念间关系进行语义关联分析的较少。同时，内容上多关注单一实体对间的关联，如疾病-基因[5]、疾病-药物，复杂、深层次的隐性知识关联难以发现；对文本挖掘结果的解读较浅，通常限于通过文献回溯法进行验证，进一步更有说服力的客观论证较少。

2.2. 基于数据的知识发现

基于数据的知识发现是指利用包括电子病历中结构化和非结构化的临床数据、生命体征与既往病史数据、生物样本组学及药品数据在内的覆盖基因、蛋白、通路、疾病、表型、药物等全过程事实型数据进行的医药知识的挖掘与发现。在方法上以生物信息学、网络药理学[6]及统计相关方法为主，基于现有的生物医学知识发现平台(如进行蛋白质互作网络分析的 String 数据平台) 和领域专用的软件或插件进行疾病与基因的生物信息学分析，在分子和细胞层面发现疾病、基因及生物功能和药物之间的关联，揭示致病基因、生物过程及机制层面的知识，并在此基础上结合机器学习、深度学习等方法预测可能的药物。该类研究在揭示当前生命科学问题的规律方面具有较大优势，加速了人类揭示生命健康疾病奥秘的进程。但由于所基于的数据为真实世界的客观数据，在新知识的发现方面与本身就包含科学假设和具有新颖性的科研文献相比不占优势。

将基于文献与基于数据的两类知识发现方法相结合，将会弥补各自不足，在生物医学规律的揭示和新知识的发现方法取得较好结果。目前将二者有效结合进行生物医学知识发现的相关研究相对缺乏。本课题综合利用文献数据及和类型生物医学数据，采用自然语言处理、生物信息学和数学计算的方法构建了一种集知识的组织、挖掘与利用于一体的集成化模型。目的在于发现疾病-基因-药物之间的语义关联并构建疾病知识图谱、确定疾病的核心基因以揭示疾病的生物学机制、为疾病预

测潜在药物。最后，通过实例验证的所提出模型的有效性，旨在为医学科研及临床诊治提供有价值信息，启发科研思路、促进生物医学及药物学的科研及实践发展的作用。

3. 疾病-基因-药物知识发现研究框架

融合多源数据与方法的疾病知识发现模型主要包括三个模块：知识图谱的构建、疾病-基因知识发现和疾病-药物知识发现。总体思路如图 1 所示。

3.1 基于文献的疾病知识图谱构建

（1）数据的获取

PubMed 是由美国国立医学图书馆所构建的全球最大的生物医学文献数据库，数据来源广泛、文献质量可靠并且标引规范，是医学知识发现常用的科技文献数据源。本文以 PubMed 数据库为文献数据源，获取某一疾病相关研究文献的题录数据,主要包括篇名、摘要、主题词，文献类型，文章 ID 号等字段。

（2）生物医学实体及关系抽取

本文以文献的标题及摘要为数据源，以句子为中心，基于命名实体识别、句法依存及语义类型分析工具 SemRep，识别同一句中实体及实体之间的关系，抽取出的结果为一个个主-谓-宾三元组结构。下文将其称为 SPO 三元组，其中 S 代表主语实体，P 代表关系谓词，O 代表宾语实体。

① 抽取 SPO 三元组

SemRep 是美国国家医学图书馆开发的一个 NLP 系统，依托于统一医学语言系统（UMLS）中的超级叙词表、语义网络、专家词典和词汇工具，能够对文献数据进行词法、句法、语义分析并通过与标准词表映射实现实体的归一化操作。经 SemRep 处理后的数据如图 2 所示。

图 2 SemRep 结果输出信息

Fig.2 SemRep Result Output Information

② SPO 三元组数据处理

首先采用自编程序对 SemRep 处理后的数据进行整理形成结构化数据。整理后，一条记录为一个来源于某一篇文献的 SPO 三元组数据，包括来源文献的 ID 号、命名实体的 CUI 号，实体名称、实体语义类型、实体间的语义关系等。

其次对 SPO 三元组数据进行清洗。进入知识图谱的应是具有一定可靠性并且指向明确、专指性的概念及关系。为此，需要去除三元组中含义宽泛的概念（如“Child”“People”“Human”等）、去除否定谓词（如“NEG_PREDISPOSES”“NEG_PRODUCES”“NEG_COEXISTS_WITH”）。同一个 PMID 下重复的 SPO 数据只计一次。

最后筛选与疾病相关的 SPO 三元组。疾病知识包括疾病的症状或异常等、相关联基因、治疗药物、诊治过程等五类相关内容，这些类别的实体对应的语义类型如表 1 所示。相应地，这些疾病相关实体间的语义关系则应由表示与疾病发生、发展、诊断、治疗、预防等相关的谓词(Predication)来充当，具体涉及到的关谓词关系类型[8]如表 2 所示。因此，进入疾病知识图谱的 SPO 三元组应满足的条件是三元组中的实体语义类型属于表 1 中的语义类型、实体间的关系谓词属于表 2 中所列的关系类型。

表 1. 疾病相关的实体语义类型

Table 1. Semantic Type Filtering Table

实体分类	语义类型
药物类	有机化学物质（orch），药理物质（phsu），生物活性物质（bacs），抗生素（antb），元素、离子或同位素（elii），指示剂、试剂（irda），免疫因子（imft），无机化学物质（inch）
靶点类	基因或基因组（gngm），氨基酸、肽或蛋白质（aapp）
疾病类	获得性异常（acab），解剖异常（anab），细胞或分子功能障碍（comd），疾病或综合征（dsyn），疾病实验模型（emod），检查/检验结果（fndg），受伤或中度（inpo），精神或行为功能障碍（mobd），瘤形成过程（neop），病理功能（patf），体征或症状（sosy）
诊疗活动类	保健活动（hlca），诊断过程（diap），治疗或预防过程（topp），分子生物学研究（mbrt），研究活动（resa），实验过程（lbpr）

表 2. 疾病相关的语义关系

Table 2. Semantic Relation Filtering Table

类型	语义关系
诊断治疗类	TREATS，DIAGNOSES，PREVENT，METHOD_OF，MANIFESTATION_OF，MEASURES
相关疾病类	PRECEDES，COEXISTS_WITH，COMPLICATES
疾病特征类	ISA，LOCATION_OF，PROCESS_OF
影响/关联类	AFFETS，ASSOCIATED_WITH，CAUSES，PREDISPOSES
药理作用类	AUGMENTS，DISRUPTS，INTERACTS_WITH，STIMULATES，INHIBITS

（2）疾病知识图谱的构建

知识图谱通过图模型来描述现实世界中各种实体之间的关系，图模型采用节点描述实体、节点之间的边来描述各实体之间的关系[9]。每个节点、边可包含也可不包含属性信息，可以有一个或多个属性信息。

Neo4j 是当前最主流的图数据构建和存储工具，其基于属性图模型，设计了节点、节点属性、边及边属性的存储方案[10]。本文采用 Neo4j 将筛选出的 SPO 三元组数据构建成知识图谱，形成可视化的疾病知识数据库。其中，图谱中的节点为 SPO 中的头实体和尾实体，每个实体具有语义类型、CUI 号等属性；图谱中的边对应 SPO 中的实体关系。

3.2 疾病-基因知识挖掘

文献中抽取出的疾病-基因（或基因-疾病）SPO 语义关系对只能说明基因对疾病的作用或影响关系，关于基因如何作用或影响该疾病的发生发展知识需要在生物医学知库的支持下进一步挖掘。

（1）疾病核心基因的获取

首先，从疾病知识图谱中检索包含目标疾病并且关系对中另一实体语义类型为基因的 SPO 三元组，获得与该疾病相关联的基因列表。其次，将获得的基因列表导入 String 数据库构建蛋白质互作用网络[11]（PPI）。String(https://string-db.org/)是一个在线搜索已知的蛋白质互作关系的数据库，提供了超过 2000 个不同物种的近 10 万多种蛋白质互作信息。通过该数据库对基因构建的 PPI 网络中,节点是基因所编码的蛋白质，节点间的连边表示蛋白质之间具有相互作用关系。然后，利用 CytoHubba 插件对 PPI 网络节点进行 Degree 分析获得该疾病的关键基因（Hub Gene）[12]。

（2）基因和通路富集分析

为揭示基因信息结构的复杂性与根本规律，采用R语言中的clusterprofiler包对关键基因进行基因本体（GO）富集分析和通路(pathway)富集分析,用于发现疾病关键基因的生物进程与蛋白通路，揭示其中隐含的生物学分子机制，从分子层面解读基因与疾病的关系，反映疾病发生发展和遗传规律。

3.3 疾病-药物知识的发现

（1）疾病的关联中药提取

将疾病的关键基因映射到Coremine Medical数据库，从其中的中药数据模块（traditional Chinese medicine）以$P < 0.05$为标准筛选出关键基因相关的治疗中药。

（2）疾病的核心中药挖掘

首先，以基因为中心，构建基于基因的中药共现网络并对其进行聚类。聚类结果中，同一类团内的药物作用于较多的共同基因或者具有相似的性能。不同类团间的药物区别较大。其次，采用代表度指标R提取每个类团中对类团贡献度大的指定数量的药物作为类团的核心代表。所有类团中的核心药物构成对目标疾病具有潜在意义的药物。

基于文本标引的TF-IDF技术的思想原理[13]，本文构造了药物对类团的代表度指标R。具体内容为：

药物节点对聚类类团的代表度反映该药物对类团的贡献大小或重要程度。如果一种中药在类团中出现的频次高而在整个中药网络中出现的频次相对低，则说明该药物对所在类团的代表度较高。

基于TF-IDF原理，本文通过类团中药物的中心度（与类团中其它节点共现的频次之和）和类团中药物连接的类团外药物的数量之比表示药物对所在聚类类团的代表度R。

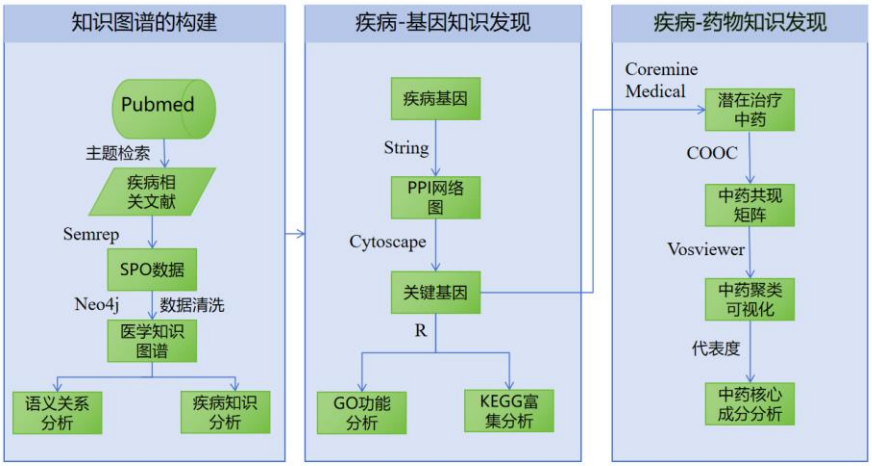


图 1 疾病知识发现模型构建框架

Fig.1 Framework of disease knowledge discovery model

4. 实证分析

尿毒症是一种慢性肾功能疾病。患有尿毒症之后，人体因不能通过肾脏排出代谢所产生的废物而造成机体内部紊乱。该病治疗周期漫长且不可逆转，患者生存质量差且医疗成本高。本文以尿毒症为例，基于多源数据挖掘疾病知识、预测潜在治疗药物，期望验证模型的可行性及有效性的同时，为尿毒症疾病诊治的相关研究和实践提供有价值的参考信息。

4.1. 数据获取与处理

在 PubMed 数据库中，采用检索式“Uremia[mesh]”检索尿毒症相关的文献，去除综述和人工判读删除不相关文献后得到文献 25,439 篇，检索时间为 2022 年 9 月 19 日。

使用 SemRep 对文献的标题和摘要进行句子级别的句法、词法和语义分析后得到包含“主-谓-宾”结构的 SPO 语义三元组原始信息，采用 Python 程序对分析后的数据进行结构化整理，最终从文献中抽取到 116444 个 SPO 三元组。

4.2. 尿毒症知识图谱的构建结果

按照表 1 和表 2 实体类型和关系类型约束进行筛选，提取出尿毒症的语义三元组 32180 个。利用 Neo4j 图数据库对 SPO 三元组构建成可视化网络知识图谱（图 3）。

4.2.1. 尿毒症知识图谱概况

所构建的尿毒症知识图谱由 7831 个节点、20140 条连边构成。三元组中的一组谓词和论元对应于知识图谱中的一条边及其连接的两个节点，构成尿毒症疾病的一个知识条目。按照连边代表的关系类型，统计尿毒症疾病知识中诊治类、相关疾病类、疾病特征类、影响/关联类和药理作用类知识条目数量分布，结果如图 4 所示。

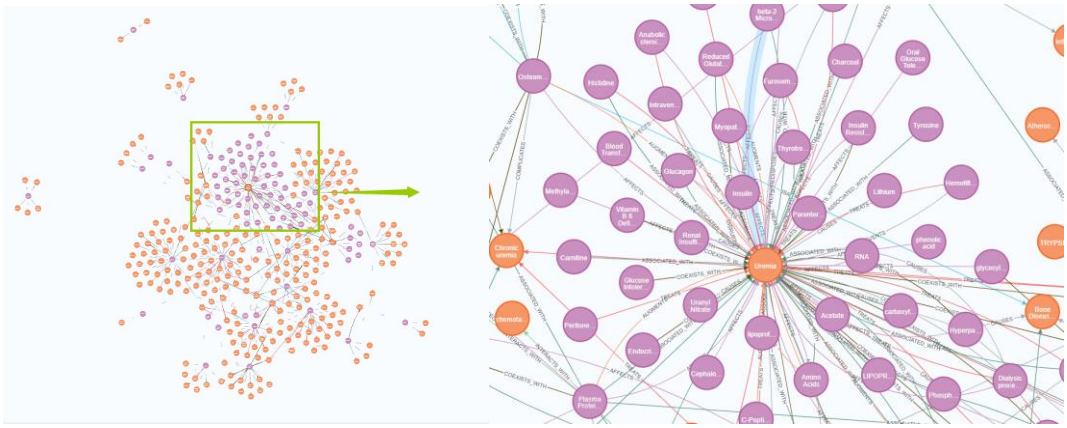


图 3. 尿毒症 SPO 数据可视化（部分）
Figure 3. Uremia SPO data visualization (part)

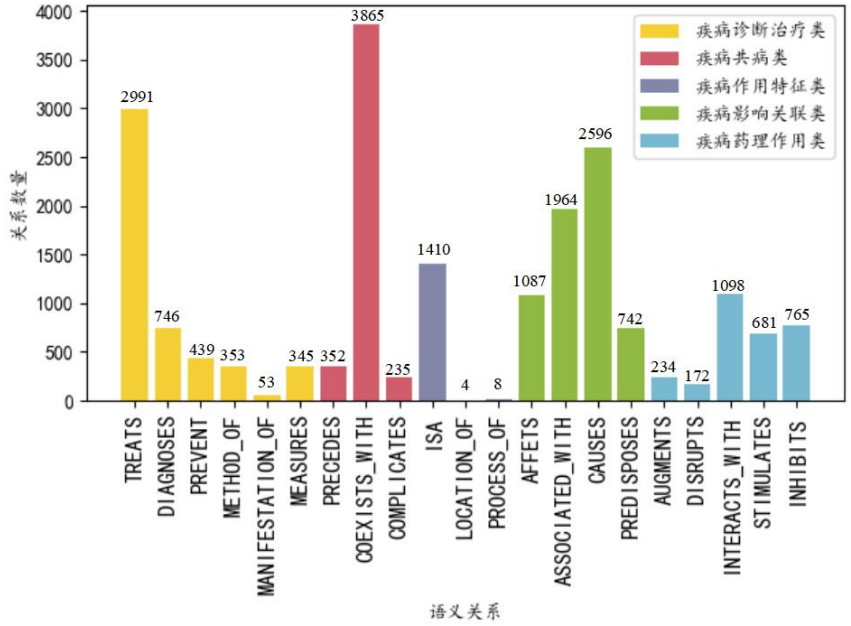


图 4. 尿毒症语义关系柱状图
Figure 4. Histogram of semantic relationships in uremia

诊断治疗类条目记录了尿毒症的诊断与治疗相关的知识，共 4927 条，占比为 24.5%。知识条目中的关系谓词包括表示尿毒症的治疗方法、药物和诊断疾病的检验、检查方法等的词语。谓词论元中主语语义类型主要为 Pharmacologic Substance(药效物质)、Therapeutic or Preventive Procedure（治疗或预防过程）、Health Care Activity（保健活动）等，宾语的语义类型主要为 Pathologic Function（病理功能）、Disease or Syndrome（疾病或综合征）、Finding（检查/检验结果）等。

相关疾病类条目共 4452 条，占比为 22.1%。该类知识主要揭示尿毒症的共病（多种疾病共同作用于个体）情况以及各疾病发生的先后关系，如尿毒症的并发症有血栓性血小板减少性紫癜、血小板减少症、微血管性溶血性贫血、肝病、糖尿病等，早于尿毒症的疾病有出血性腹泻等。相应的关系谓词有 PRECEDES（先于）、 COEXISTS_WITH（与...共存）和 COMPLICATES（使复杂），谓词论元均为疾病类语义类型的实体。

疾病特征类条目共 1422 条，占比为 7.1%。用于表示尿毒症相关实体的作用位置和疾病的发展过程。关系谓词包括 ISA（是）、LOCATION_OF（位置）和 PROCESS_OF（发生在...中）三种语义关系。ISA 主要是说明不同实体之间的上下位关系，如溶血性尿毒症综合征是一种肾脏肾小

球疾病。LOCATION_OF 则通常为疾病发生的身体部位，PROCESS_OF 则说明疾病发生的有机体（人体、鼠、生物模型等）。

影响/关联类条目共 6389 条，占比为 31.7%。用于表示疾病或者其他物质对尿毒症产生的影响。关系谓词包括：AFFETS（影响），表示疾病和其他疾病、功能的影响关系，如骨疾病会影响尿毒症；ASSOCIATED_WITH（与...有关联），表示疾病和其他物质、疾病之间的关联关系，如色氨酸和尿毒症的发生有关联；CAUSES（引起），用于说明疾病的致病因素和会造成的结果，如尿毒症会造成低血糖症；PREDISPOSES（诱发），表示疾病的诱发条件，如肾衰竭易导致血管钙化。

疾病药理作用类条目共 6389 条，占比为 14.6%。药理作用类知识表示物质与物质、疾病与物质之间的相互作用等，关系谓词包括 AUGMENTS（增加）、DISRUPTS（扰乱、使中断）、INTERACTS_WITH（相互作用）、STIMULATES（刺激）和 INHIBITS（抑制）。STIMULATES 和 AUGMENTS 反映正向关系、DISRUPTS 和 INHIBITS 反映负向关系，INTERACTS_WITH 反映两个物质之间的相互作用。如抗凝血酶与肝素相互作用。

4.2.2. 尿毒症疾病知识

尿毒症知识图谱可提供尿毒症知识的高效检索、导航和知识关联发现等功能，为科研人员、临床医师和患者提供有效帮助。下面从疾病的基因、共病关系、检测方式和治疗手段四个方面简述尿毒症的核心知识。

（1）尿毒症的关联基因

按抽取出的三元组频次，尿毒症相关的基因排在前三位的依次是有 CFH、CD46、GH1 等基因。研究证实，CFH 基因与宿主细胞表面的 C3b 或糖胺聚糖的结合受损会降低 CFH 对 C3 转化酶的抑制作用，从而促进血栓的形成，最终导致非典型溶血性尿毒症综合征（aHUS）进一步发展[14]；aHUS 发生发展和补体替代途径相关，aHUS 患者溶血期血红素的释放能诱导补体激活，使内皮细胞表达 CD46 减少[15]。

（2）尿毒症的共病关系

按照抽取出的在三元组频次，主要的尿毒症并病症状为血栓性血小板减少性紫癜、血小板减少症、溶血性贫血等。血栓性血小板减少性紫癜与溶血尿毒症综合征被统称为 TTP-HUS[16]，它们在症状和治疗方式都近乎相同，主要表现为微血管血栓导致的全身各系统损坏。由于尿毒症的主要方式是血液净化治疗，肝素抗凝是治疗的重要途径，治疗时会导致血小板减少因而产生血小板减少症，并形成其他病理过程，如脑血栓[17]。又由于频繁的血液透析会导致铁丢失，会使得尿毒症患者的机体免疫功能紊乱、红细胞破坏加速溶血超过骨髓代偿而引发溶血性贫血[18]。

（3）尿毒症的检测

尿毒症的检测方式主要有肾活检和超声。肾活检是使用超声显像仪对肾脏进行定位后，用自动活检枪机在肾脏使用穿刺针穿刺获得肾组织，穿刺的次数应该小于 5 次，将取得的组织通过显微镜检查得到结果的方法[19]。使用超声检测肱动脉内皮功能，得到的结果可以提示血管的变化，结果可以为患者进行早期干预治疗[20]。

（4）尿毒症的治疗

尿毒症的治疗包括药物治疗、替代治疗及干预治疗三种方式。

治疗尿毒症常用的药物有依库珠单抗、重组促红细胞生成素、环孢霉素等。依库珠单抗是一种抗体，用于治疗儿童原发性非典型溶血性尿毒症综合征。但此药有患脑膜炎球菌和其他细菌的可能性[21]。重组促红细胞生成素能够增加人体的造血功能，加快消耗体内的铁元素，治疗尿毒症患者的贫血问题[22]。环孢霉素能选择性抑制 T 细胞和辅助性 T 细胞，对肾移植抗排斥反应有作用[23]。

尿毒症的替代治疗主要有血液透析、腹膜透析、血浆置换、肾移植等，血液透析是利用透析器将患者血液中的毒素排除体外，并补充身体需要的物质以维持酸碱平衡。但透析容易引发其他并发症[24]。肾移植则是在配型成功后将功能健康的肾脏移植到尿毒症患者体内代替患者原本肾脏的方法，是当前最有效的治疗方法，术后会有免疫排斥反应[25]。

饮食干预、心理干预、护理干预可改善尿毒症患者的病情。通过科学决策指定合理的饮食安排，能够改善病人的病程，也能有效降低并发症的发生，使得病人的身心状况都得到很大的好转[26]。

4.3. 尿毒症基因知识发现

4.3.1. 基于蛋白质互作网络挖掘尿毒症核心基因

筛选以尿毒症及基因实体为论元的谓词三元组，将从中得到的 137 个尿毒症关联基因提交至 String(https: //string-db.org)平台，生物种类设置为“Homo sapiens”，置信度选择≥0.4，构建的 PPI 网络（图 5）包含 131 个结点和 1834 对互作关系。图 5 中，节点越接近中央，其中心性越高。节点的面积越大，说明与该接点连接的其它节点越多，节点越重要。将图 5 中的网络数据导入生物网络分析平台 CytoScape(v3.9.1)，利用 cytoHubba 插件提取 Degree 排名前 15 的节点构成尿毒症的核心基因网络（图 6），网络中的节点颜色越深，节点的度值越大。其中，度值最大的基因是 IL6（99 分），其次是 IL1B（91 分），其后依次是 CRP（75 分）、IL10（75 分）、TGFB1（71 分）等。

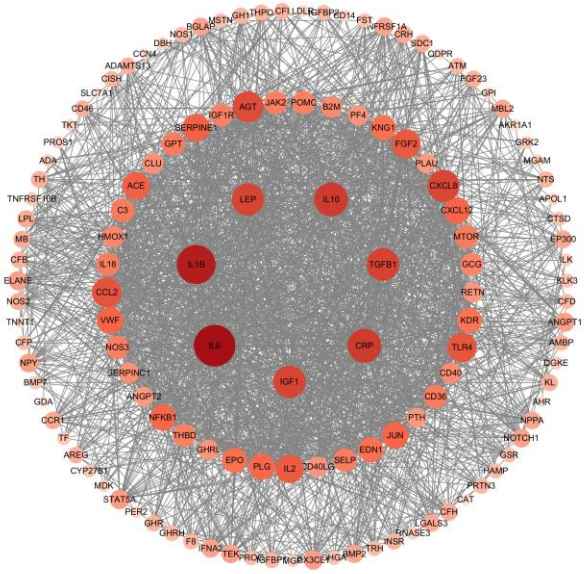


图 5. PPI 网络图

Figure 5. PPI Network Diagram

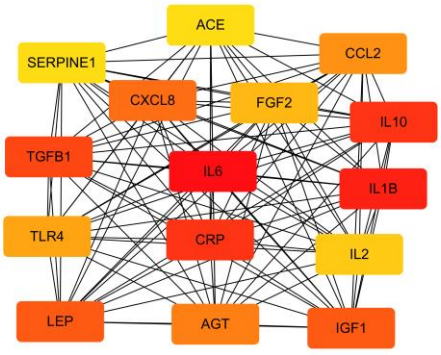


图 6. 核心基因互作网络图

Figure 6. Core gene interaction network

4. 3. 2. 核心基因的 GO 功能富集分析和 KEGG 通路分析

使用 R 语言 clusterProfiler 包对得到的关键基因进行 GO 分析。图 7 表明关键基因主要参与的生物过程主要有单核细胞增殖、淋巴细胞活化的正向调节、对脂多糖的反应等；图 8 揭示尿毒症核

心基因主要发挥的分子功能，主要包括受体配体活性、细胞因子活性、生长因子活性等；图 9 显示受尿毒症核心基因影响的细胞组分（CC）包括血小板 α 颗粒管腔、血小板 α 颗粒、分泌颗粒管腔、细胞质囊腔、囊腔等。

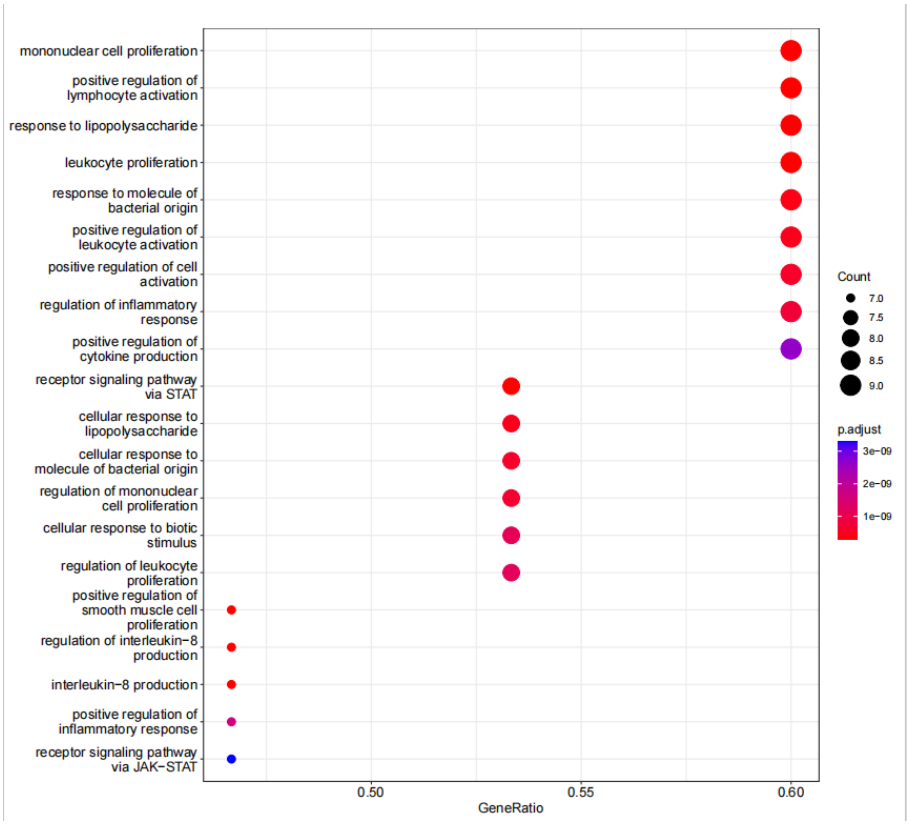


图 7. 生物过程分析图

Figure 7. Analysis Diagram of Biological Process

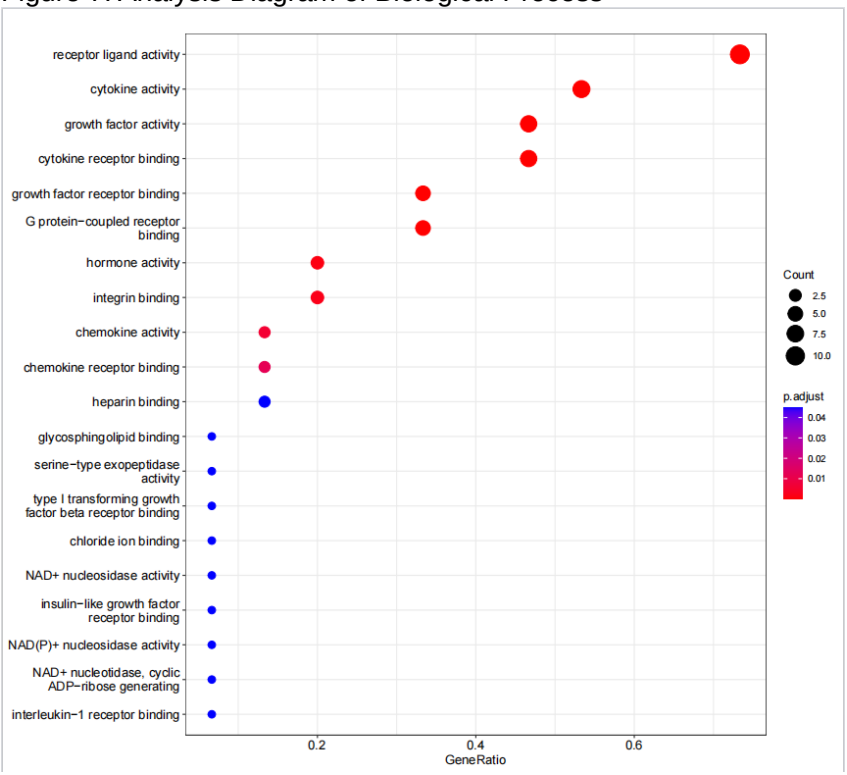


图 8. 分子功能分析图

Figure 8. Molecular Function Analysis Chart

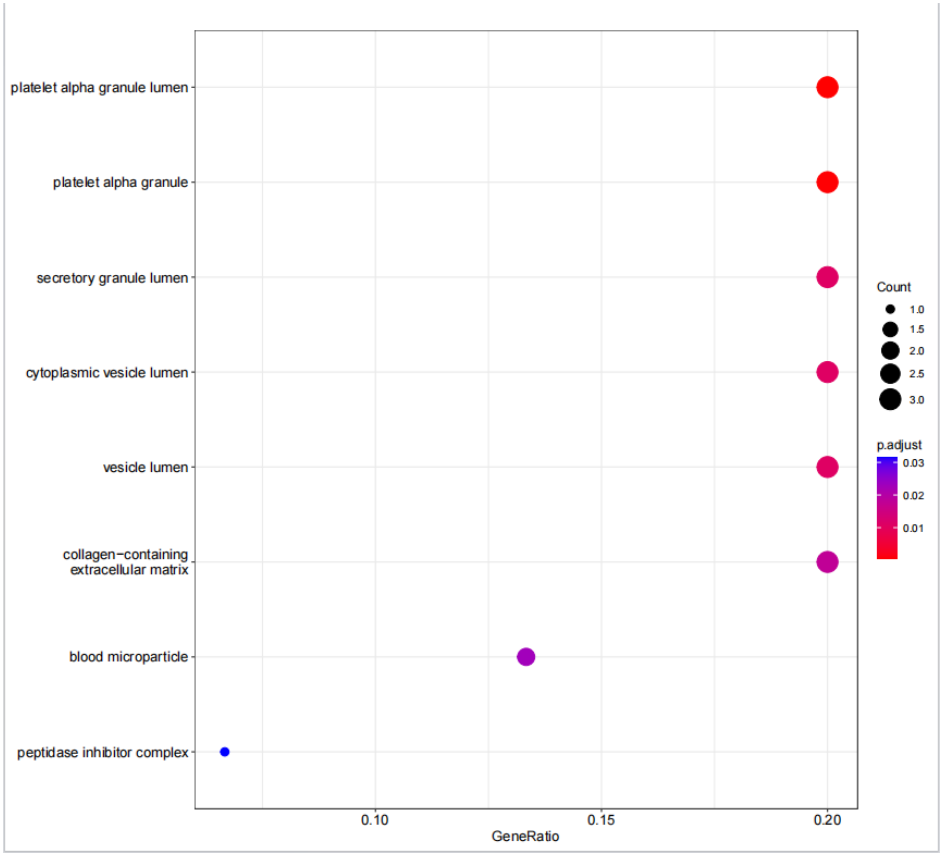


图 9. 细胞组分分析图
Figure 9. Analysis Chart of Cell Components

使用 R 语言分析关键基因富集的 KEGG 代谢通路，进一步了解基因生物学功能，明确尿毒症的潜在发病机制[27]。图 10 显示，恰加斯病、细胞因子-细胞因子受体相互作用、疟疾、糖尿病并发症中的 AGE-RAGE 信号通路、鼠疫感染等为主要通路。

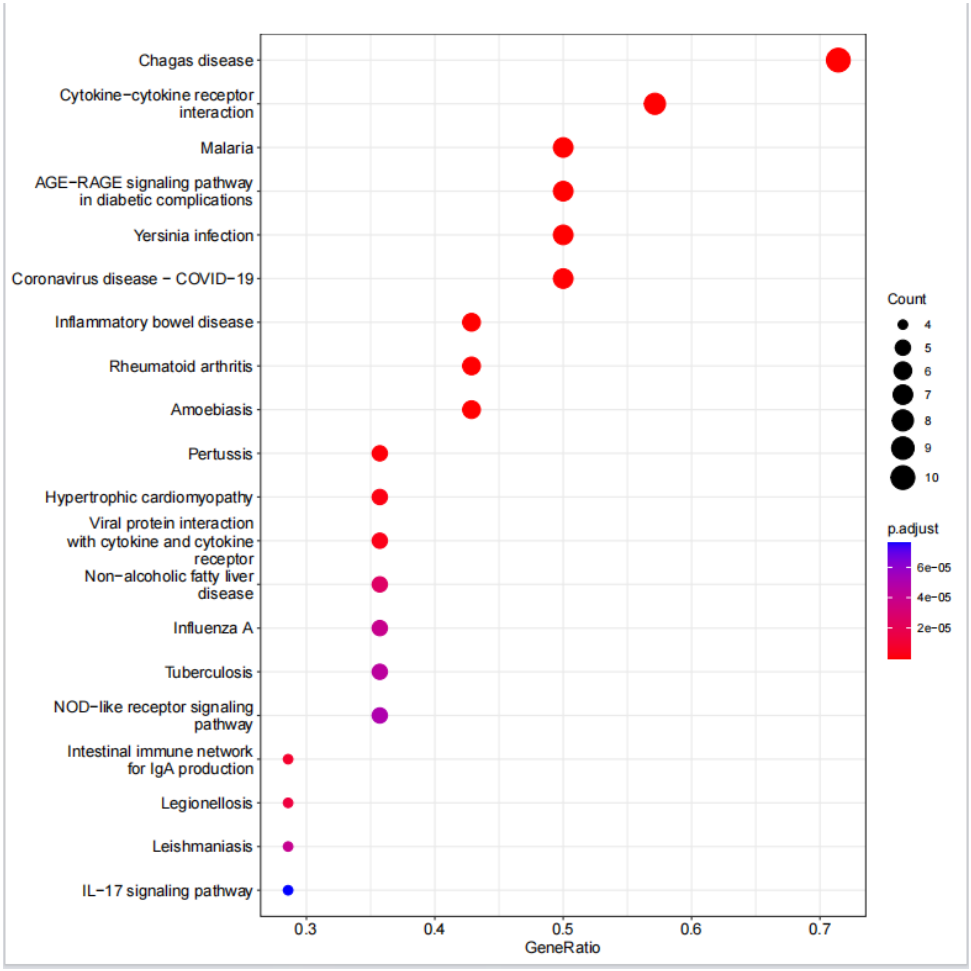


图 10. KEGG 通路分析图

Figure 10. KEGG Enrichment Analysis Diagram

将尿毒症、关键基因及基因的 MF、BP、CC 和代谢通路导入 Cytoscape，生成的尿毒症靶点-通路-生物过程-分子功能-细胞组分的全景式拓扑结构图（图 11），为多层次、全面了解尿毒症发生、发展机制提供便利。

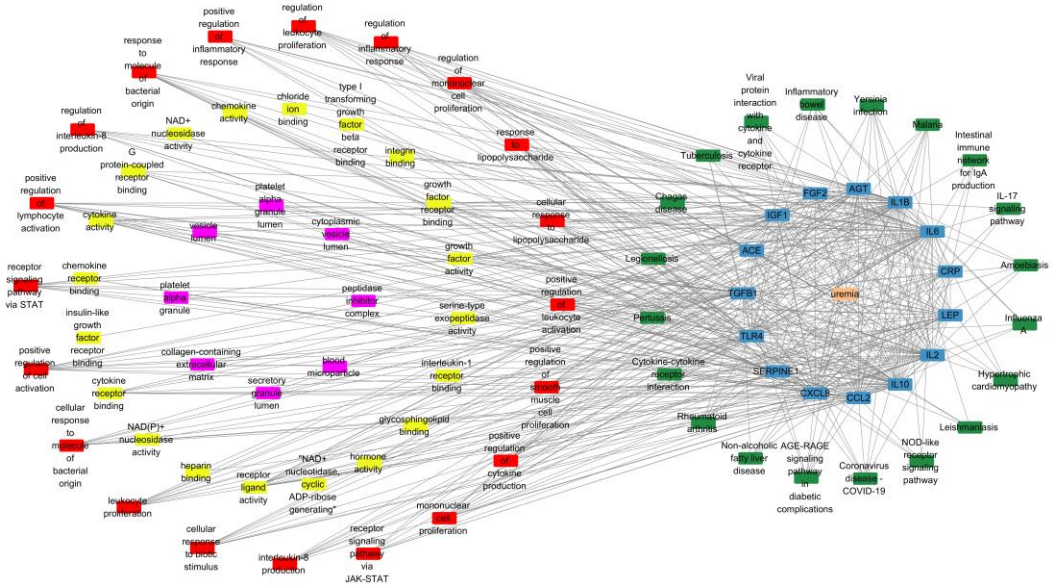


图 11. 尿毒症的靶点-通路-生物过程-分子功能-细胞组分拓扑图

Figure 11. Target pathway biological process topology of uremia

图中，左则蓝色圆圈对应靶点集合，绿色圆圈是对应的基因的代谢通路，红色、黄色和紫色圆圈分别对应于基因参与生物过程、分子功能和细胞组分。

4.4. 尿毒症药物知识发现

4.4.1. 尿毒症的中药信息预测

将 15 个尿毒症关键基因提交至 Coremine Medical 数据库，按显著性 P<0.05 标准筛选得到对关键基因具有调节作用的中药，可作为尿毒症的潜在治疗药物。表 3 是尿毒症关键基因、对应的潜在中药及其 P 值的信息示例。

表 3. 关键基因的潜在中药预测信息示例

Table 3. TCM prediction of key genes (partial)

序号	基因	预测的中药及 P 值
1	IL6	人参花（0.000521663）、人参（0.000525562）、黄芩（0.000640549）、连翘（0.001272144）.....
2	IL1B	膜荚黄芪（0.001324988）、白术（0.001536086）、三七（0.002489645）、白及（0.014858648）.....
..	...	...
15	SERPINE1	鹿筋（0.000904047）、关木通（0.001688026）、水芹（0.007352782）、三棱（0.016143062）.....

4.4.2. 尿毒症潜在中药的核心成分分析

（1）中药的共现网络构建及聚类分析

将预测出的 Hub Genes 调节中药按频次进行排序，获得 123 种频次大于 4 的中药，根据药物基于基因的共现关系构建成 123*123 的中药共现矩阵（表 4）。基于弦相似度对中物进行聚类分析并将聚类结果进行可视化。得到的中药聚类图谱（图 12）显示，尿毒症的潜在中药可聚类为 4 个类团，图中同一颜色的中药属于同一个类团。同一类团中，药物之间的关联紧密，相似度较大；不同类团间的药物之间区别较大。

表 4. 基于关键基因的中药共现矩阵(部分)

Table 4. TCM Co occurrence Matrix Based on Key Genes (Partial)

序号	特征关键词	膜荚黄芪	人参	人参花	人参芦	人参叶	丹参	…	菝葜
1	膜荚黄芪	9	8	8	8	8	7	…	4
2	人参	8	9	9	9	9	6	…	4
3	人参花	8	9	9	9	9	6	…	4
4	人参芦	8	9	9	9	9	6	…	4
5	人参叶	8	9	9	9	9	6	…	4
6	丹参	7	6	6	6	6	8	…	3
…	…	…	…	…	…	…	…	…	…
123	菝葜	4	4	4	4	4	3	…	4

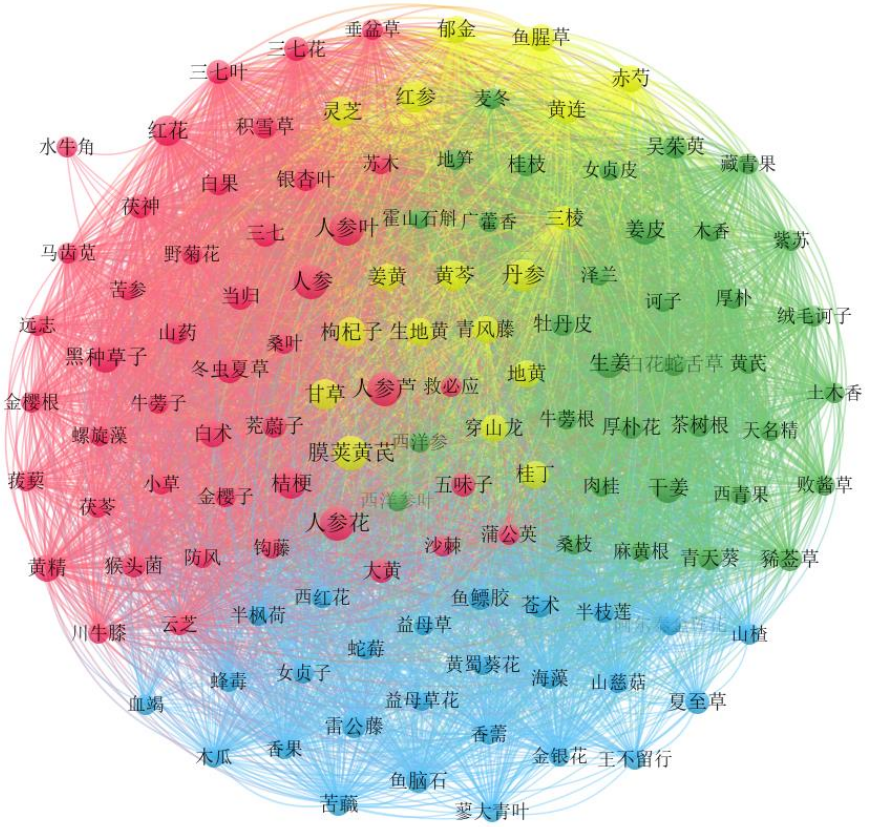


图 12. 治疗中药关键词聚类图(频次≥4)

Figure 12. Keyword cluster diagram of therapeutic Chinese medicine (frequency ≥ 4)

按顺时针方向，将图谱中的红色节点集合定义为第 1 个中药类团，第 2~第 4 个药物类团依次为黄色、绿色和蓝色节点构成的集合。

（2）中药类团中的核心药物提取

计算类团中每一种中药对所在类团的代表度，选取类团内代表度由高到低排名前 10 的中药作为该类团的核心药物。每个类团的核心药物及其对类团的代表度计算结果如表 5 所示。

表 5. 不同类团点度中心度排名前 10 的关键词

Table 5. Top 10 keywords for ranking centrality of different clusters

类团	核心中药及类团代表度
类团 1	人参（2.78）、人参花（2.78）、人参芦（2.78）、人参叶（2.78）、红花（2.47）、冬虫夏草（2.47）、白术（2.47）、黑种草子（2.47）、桔梗（2.47）、三七（2.36）
类团 2	膜荚黄芪（1.02）、黄芩（0.99）、丹参（0.99）、枸杞子（099）、灵芝（0.99）、甘草（0.99）、赤芍（0.96）、青风藤（0.96）、郁金（0.96）、姜黄（0.96）
类团 3	姜皮（1.60）、生姜（1.60）、干姜（1.60）、茶树根（1.47）、厚朴花（1.47）、青天葵（1.47）、白花蛇舌草（1.47）、吴茱萸（1.47）、豨莶草（1.40）、桂枝（1.38）
类团 4	鱼脑石（0.90）、鱼鳔胶（0.90）、苦藠（0.84）、雷公藤（0.80）、苍术（0.80）、阿尔泰金莲花（0.79）、山慈菇（0.79）、海藻（0.79）、金银花（0.79）、半枝莲（0.79）

5. 讨论与结论

5.1 抽取文本中的 SPO 语义三元组可实现疾病知识图谱的高效构建与利用

本文基于自然语言文本处理工具 SemRep 抽取文献摘要中的 SPO 三元组并构建了尿毒症知识图谱。将文献中蕴含的关联知识结构化表示并且明确了知识单元之间的关系的具体语义，实现了文献中细粒度知识提取并将其存入图数据库。构建的知识图谱是对原有文献中非结构化知识的形式化表示，可实现新疾病知识的可视化导航、查找、分析与利用等，可大幅度提升对疾病知识的管理与利用效率。

5.2 通过 SPO 三元组分析可有效发现疾病的影响基因

从文献信息中提取出三元组并进行筛选和生物关联分析得到的 IL6、IL10、IGF1 等 15 个尿毒症关键基因，经与已有基因数据库比对发现，均为数据库中确认的尿毒症靶基因，说明了本文基于 SPO 三元组信息抽取的疾病知识发现模型的有效性。这也说明，如果扩大核心基因提取的数据范围，就有可能发现现有知识库中尚未出现过的全新有效基因。

研究显示，尿毒症患者由于肾功能不全导致促炎细胞因子清除力减弱，血清中 IL-6 等炎症因子显著升高，导致肾脏组织血管堵塞，进一步加重血管内皮功能受限的程度，从而对肾脏功能产生不良影响[28]。IL10 是细胞因子合成抑制因子，可以抑制肾小球系膜细胞产生，进而产生炎症反应，保护肾功能[29]。VEGFA 表达减少导致肾小管毛细血管网萎缩，加重肾纤维化[30]。IGF1 能治疗肾间质纤维化，减轻肾小球进出动脉阻力，延长慢性肾衰竭的发展[31]。

5.3 对疾病关键基因的生物关联分析提供了对疾病发生、发展机制与规律的多角度、多层次的知识发现。

对关键基因进行 GO/KEGG 分析得到尿毒症发病的主要通路。其中，糖尿病并发症中的 AGE-RAGE 信号通路，通过氧化加重血管损伤，加重肾组织恶化等风险，导致死亡率增加[32]。IL-17 信号通路是炎症信号通路。IL-17 具有促炎功能，当机体受到外界刺激时，IL-17 刺激细胞产生 IL-6 和 TNF- α，激活 NF- κB 信号通路，NF- κB 信号通路是炎症相关的经典通路，而炎症反应能加速肾小动脉粥样硬化进而加快肾功能损伤[33]。

5.4 基于关键基因预测的中药性味归经与尿毒症病机相符

从中医理论层面来说，尿毒症发病原因多为肾元亏虚，湿浊内蕴。主要采用补气健脾益肾类中药进行治疗。在《中国药典》（2020 版）、云南中医药大学中医药文化传播平台数据库和中医资源网中检索和分析尿毒症潜在中药的性味归经等知识发现，这些尿毒症相关药物按功效可分为四大类：清热药（桔梗、黄芩、青天葵、金银花）、解表药（生姜、姜皮、赤芍）、温里药（干姜、吴茱萸）、理血药（丹参、三七、红花）和补益药（人参、膜荚黄芪、枸杞子）等可能治疗尿毒症的潜在药物的多种中药。其中黄芩和金银花均为性寒类中药，具有清热燥湿，泻火解毒的功效。生姜味辛性微温，具有解表散寒，温中止呕的功效。干姜味辛性热，具有温中散寒，回阳通脉等功效。丹参味苦性微寒，具有活血祛瘀，凉血消痈等功效。山慈菇味微辛，性寒，具有清热解毒、消肿散结等功效。这些药物归经多归为脾、肾、心经，功效也与尿毒症的病机相同。

5.5 预测得到的尿毒症核心中药经验证可作用于四类尿毒症相关疾病

经过文献检索和人工阅读相关文献发现，每个类团中均有被临床案例或研究文献报道的可用于治疗尿毒症相关疾病的核心药物。

第一类药物主要用来治疗糖尿病肾病。糖尿病肾病临床表现主要为蛋白尿，随着病情的发展会最终发展形成尿毒症。山药、当归、红花、大黄、人参等方剂构成的降糖益肾方能够延缓肾小球硬化，有效控制患者血糖水平，减少患者蛋白尿，保护肾功能[34]；使用糖肾康合剂（白术、三七、黄芪、山药、大黄等）能有效改善肾组织损伤，降低炎症反应进而保护肾脏[35]；由人参、白术、川穹、当归等方剂构成的康乐三号方结合西医常规疗法能有效降低患者蛋白尿，改善肾功能[36]。

第二类药物主要治疗慢性肾衰竭。慢性肾衰竭主要表现为肾脏功能衰退、蛋白质代谢紊乱、最终会发展为尿毒症。使用益肾活血汤（丹参、生地黄、赤芍、甘草、山药等）可以改善肾脏血管通透性，抑制肾间质纤维化，保护肾功能[37]；服用益肾汤（大黄、黄连、丹参、赤芍、甘草等）能够延缓肾衰竭速度，保护肾功能[38]；刘晓静运用保肾合剂（杜仲、女贞子、枸杞、丹参、赤芍等）治疗慢性肾衰竭，在临床中能改善患者临床症状，提高患者生活质量[39]。

第三类的中药基本上是用来治疗肾病综合征。肾病综合征是临床中常见的肾疾病，使用温阳利水化瘀方（制附子、干姜、黄芪、桂枝、泽泻等）可以增强患者免疫功能，减轻高凝态[40]；服用温阳补肾方（制附子、干姜、桂枝、黄芪、丹皮等）治疗肾病综合征，可以减轻肾损伤，减轻炎症反

应，改善肾功能[41]；由白花蛇舌草、黄芪、桂枝、牡丹皮、紫草、干姜等方剂组成的芪地固肾方治疗小儿肾病综合征能够清除炎症因子，有效降低蛋白尿[42]。

第四类主要用来治疗慢性肾疾病。慢性肾疾病如果不加以治疗的话会在后期发展为尿毒症。服用肾炎康胶囊（半枝莲、雷公藤、茯苓、川穹等）可以减轻肾炎的高凝血症，降低血尿和蛋白尿[43]；由苍术、益母草、山慈菇、牛膝等方剂构成的中药复方能够改善肾功能，能够发挥重要作用提高临床疗效[44]；使用自拟方（山药、苍术、金银花、车前子、丹参等）可以改善肾微循环，降低蛋白尿，有效改善肾功能[45]。

在临床上，上述四种疾病都是导致患者最终发展为尿毒症的主要疾病[46]。因此，挖掘出潜在中药和提取出的核心中药可形成尿素症及相关疾病的中药方剂候选组方药材推荐列表，并对方剂组方在配伍中的君臣佐使角色具有一定的指导意义。例如，在各类团中核心成分排名靠前的人参、人参花、膜荚黄芪、黄芩、姜皮、生姜、鱼脑石、鱼鳔胶等可能是药剂配方中的君药。

6. 结语

本文提出基于生物医学文献构建疾病知识图谱，进行“疾病-基因-药物”的药物预测模型并以尿毒症为例通过实证分析验证了模型的可行性和有效性。

基于自然语言处理等方法从生物医学文献中提取疾病相关的语义三元组构建的疾病知识图谱，实现了基于文献的疾病知识挖掘、存储、检索与利用。对筛选出的“疾病-基因”关系对进行的生物信息学分析，确定了疾病的关键基因，疾病靶基因参与的生物过程、分子功能、细胞组分和主要的代谢通路，揭示了疾病在细胞和分子层面的发生、发展规律。以尿毒症为例，预测出的尿毒症潜在中药在性味归经、当前主治的疾病类别两方面都与尿毒症的病机、尿毒症前期疾病高度相似和相关，说明它们对目标疾病具有较高的推荐价值。对各类团药物核心成份的提取方法为未来目标疾病的中药方剂组方的配伍提供了指导思想。

本研究的不足在于，本文的研究结果仅是在现有知识库和研究文献中得到初步验证，推荐的用药虽然在理论上具有一定的治疗意义，但在实际中是否确实适用，还需要进一步的实验和临床验证。

PU-114

基于生物医学文献和图挖掘算法发现糖尿病-糖尿病肾病复杂病因网络及路径分析

胡拯涌 段一凡 李佳明 周佳茵 胡万飞 陈振丽 吴思竹*
中国医学科学院医学信息研究所

引言

糖尿病是 21 世纪增长最快的全球突发卫生事件之一，全球成年糖尿病患者人数达到 5.4 亿，近年来增长幅度呈上升趋势。糖尿病进展期可能损害心脏、血管、眼睛、肾脏和神经等多系统，其中糖尿病肾病是糖尿病最严重的并发症之一，也是导致糖尿病患者死亡的重要原因，而糖尿病-糖尿病肾病存在交复杂的发病机制和表现网络，包含了大量中介、混杂因素，因此需要系统开发糖尿病-糖尿病肾病之间的因果关系路径，从而获取医学因果知识，提供更多临床诊疗线索。

生物医学文献包含海量因果机制相关信息，然而从文献的文本信息中构建因果路径是较为困难的，因此国内外也针对各类生物医学知识开发了可计算生物医学知识库，如中文医学知识图谱、UMLS[1]、SemMedDB 等，其中 SemMedDB 是基于 PubMed 文献摘要提取的（subject-predicate-object，SPO）三元组，提供了一个基于生物医学文献的可计算数据集[2]。目前已有研究基于 SemMedDB 数据开展了疾病预测[3]、药物重定位[4]等研究，但关于糖尿病的因果路径研究较少。

因此，本研究旨在利用 SemMedDB 数据构建糖尿病-糖尿病肾病知识图谱，并基于图挖掘算法构建糖尿病-糖尿病肾病的因果路径，为临床实践提供一定的研究启示和思路。

材料和方法：

数据收集和预处理：

本研究将糖尿病肾病（Diabetic Nephropathy）和其同义词与危险因素、评分（Risk factor、score）等在 PubMed 组合检索，共检索得 6771 篇文献，随后导出文献的 PMID，以 PMID 为输入

从 SemMedDB 中提取文献中的三元组，其中的实体和谓词均已映射到 UMLS 的语义网络，最终提取得 22069 个初始三元组，除名称外，还包括主语和宾语的 CUI、语义类型（SUBJECT/OBJECT_SEMTYPE），谓语的语义关系类型以及该三元组来源的句子，。

初始三元组的预处理内容如下，在统计主语和宾语的概念后，首先去除了主语和宾语相同（如糖尿病-影响-糖尿病）的三元组，共删去 50 个三元组，随后剔除了不在研究范围内的概念（如妊娠糖尿病的相关概念以及国家、种族等）和抽取错误、混乱的主语和宾语（如日期、数字和乱码等），共删去 116 个三元组，清洗后最终纳入了 21902 条三元组。

图谱构建和图挖掘

基于纳入研究的三元组，本研究以主语为源节点，谓语为边，宾语为目标节点，在 neo4j 中构建糖尿病肾病危险因素的有向知识图谱，每个节点的属性包括语义类型和 CUI，而边则包括 PMID、语义关系类型和来源句。随后我们通过图挖掘算法进行路径挖掘。

首先对于节点，即主语和宾语的概念应满足语义上的相关性，我们使用 BERT[5]模型计算两个概念的语义相关性，BERT（Bidirectional Encoder Representations from Transformers）是一种预训练的深度学习模型，用于自然语言处理任务，如文本分类、情感分析、语言生成等。其基于 Transformer 的模型，使用一个大型的无监督语言建模数据集来学习一个句子的表示，这个句子对单词的顺序是不变的。具体，即通过短文本向量的余弦相似度 $similarity$来表示，对于主语和宾语的两个概念 S，O，其文本向量夹角越小，余弦相似度 $similarity(S,O)$的值越大，越认为其语义相关性高，反之则两个概念越不相关。

其次对于边即谓语，本研究基于 UMLS 语义网络的语义类型进行分类，将 functionally_related_to 关系和所有子类以及他们的反义谓词（neg_functionally_related_to）作为因果类型的谓词，其他语义类型作为非因果谓词。设定主语和宾语为 S 和 O，其构成的三元组中谓词为因果谓词的三元组数量 $N(S,P_{causal},O)$，非因果谓词的三元组数量为 $N(S,P_{Non\_cau},O)$，因果谓词的比例越高，主宾对之间为因果关系的可信度越高，因果谓词比例 $causality_ratio(S,O)$为因果谓词三元组的数量减去非因果谓词数量，再除以因果谓词的数量[6]。

最后结合语义相关性和因果谓词比例进行计算，得到最终的因果相关性分数 $S_{Causal\_relatedness}$，即 S_{cr}，

$$S_{cr}=\frac{N(S,P_{causal},O)-N(S,P_{non\_cau},O)}{N(S,P_{\quad},O)}*similarity(S,O)$$

通过设置 S_{cr}、$similarity$的不同阈值对图谱修剪，随后使用 network 库中的简单路径算法实现图谱中路径的遍历，路径中不包含重复节点，将源节点设置为糖尿病，目标节点设置为糖尿病肾病，最终得到对应阈值下所有长度的糖尿病-糖尿病肾病的因果路径。

结果

通过预处理获得的 21902 个三元组，本研究构建了一个由 6889 个实体、31859 条边组成的糖尿病肾病危险因素的知识图谱。基于未修剪的图谱先使用简单路径算法遍历糖尿病-糖尿病肾病的简单路径，计算不同路径长度下路径数量的分布，结果发现，随着路径长度增加，对应的路径数量显著增多，路径长度为 10 时，生成的路径已超过一万条，但路径长度过长时概念间的因果关联减小[7]，且提供的信息量也逐渐减少[8]，所以本研究设置挖掘算法的路径长度为小于等于 8。

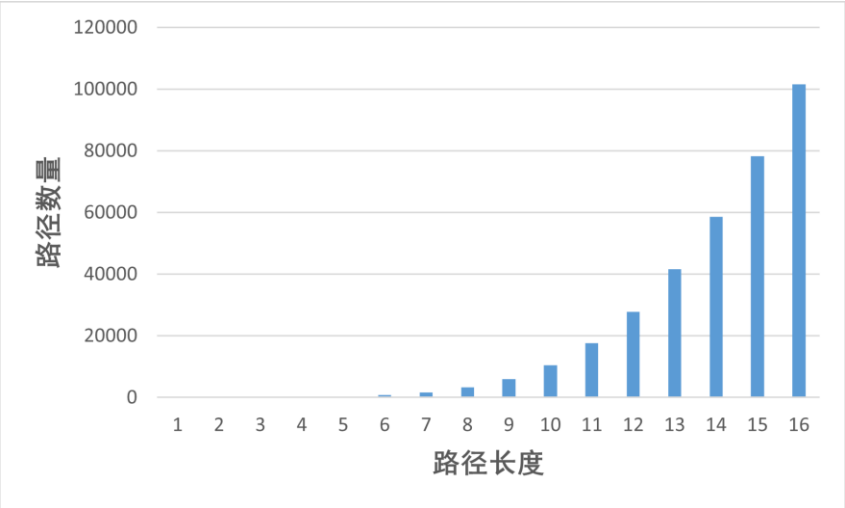


图 3.1-10 路径数量分布直方图

随后计算余弦相似度 $similarity$和因果相关性分数 S_{cr}实现对修剪。首先调用 BERT 预训练模型计算得到每个主语和宾语间的余弦相似度 $similarity$，其值越接近 1，语义相关性越强，为保证概念间基本的语义相关性，对 $similarity$设置不同阈值对比修剪效果。在设定余弦相似度阈值的基础上，计算出每个主宾对的因果相关性分数 S_{cr}，S_{cr}分数使用不同的百分位数设定阈值。根据以上两个参数的变化，列出不同阈值下的修剪结果。本研究希望在保证较高语义相关性的同时，避免过度修剪概念和三元组影响挖掘因果路径，故将 $similarity$阈值固定为大于 0.7。

表 3.1-1 部分阈值下概念和三元组修剪效果

阈值设置		图谱中的概念数和百分比	图谱中的三元组数和百分比
未修剪		5026	21902
基于 $similarity$ 修剪	>0.9	1215（24%）	2372（11%）
	>0.8	3546（71%）	10845（50%）
	>0.7	4752（95%）	17269（79%）
	>0.6	4996（99%）	19003（86%）
基于S_{cr}分数修剪	>85%	1536（31%）	3286（15%）
	>80%	1868（37%）	4381（20%）
	>75%	2219（44%）	5476（25%）

完成阈值设定后进行了糖尿病-糖尿病肾病的因果路径挖掘，在 $similarity$> 0.7，S_{cr}分数>95%的阈值设置下，没有挖掘到满足条件的糖尿病-糖尿病肾病的因果路径，而在 S_{cr}分数>85%、>80%、>75%3 个阈值下，分别挖掘出 2 条、19 条、53 条路径长度为 8 的因果路径，使用该路径 S_{cr}分数之和除以路径长度的平均值作为该路径的评分，S_{cr}分数>80%评分最高的五条因果路径和对应路径评分如下表。

表 3.1-2 S_{cr}分数>80%阈值下评分最高的五条路径

因果路径	路径评分
糖尿病→高血糖→肾功能损害→心血管疾病发病率→高血压疾病	0.906519

→慢性肾小球肾炎→终末期肾功能衰竭→肾小球病→糖尿病肾病		
糖尿病→糖尿病并发症→肾脏疾病→自主神经病变→视网膜疾病	227	0.903
→盲目的视力→终末期肾功能衰竭→肾小球病→糖尿病肾病		
糖尿病→糖尿病神经病变→肾脏疾病→自主神经病变→视网膜疾病→盲目的视力→终末期肾功能衰竭→肾小球病→糖尿病肾病	417	0.885
糖尿病→胰岛素→肾功能损害→心血管疾病发病率→高血压疾病	611	0.877
→慢性肾小球肾炎→终末期肾功能衰竭→肾小球疾→糖尿病肾病		
糖尿病→血管舒张障碍→高血压疾病→肾功能不全→心血管疾病	148	0.877
→跟进→终末期肾功能衰竭→肾小球病→糖尿病肾病		

表中的因果路径提供了从糖尿病到糖尿病肾病的可能发病机制，其中肾小球疾病到糖尿病肾病的路径在评分前五的路径均出现，这一路径的因果分数为 0.92，提示了其较高的因果关联，先前的研究中有认为肥胖相关肾小球疾病有可能加重糖尿病肾病[9]。除此之外还有慢性肾小球肾炎、自主神经病变等多个中间节点出现，心血管疾病、视网膜疾病等提示路径中可能出现的并发症。

讨论与总结

本研究提出一种基于生物医学文献知识发现和知识图谱用于挖掘糖尿病-糖尿病肾病致病因果路径的方法。具体来说，与聚焦于糖尿病-糖尿病肾病的特点临床数据不同，本文纳入了更广泛的生物医学文献中的知识，利用三元组过滤出实体间的因果关系，并基于因果分数和路径长度形成了最终的因果路径。尽管目前对糖尿病及其并发症的发病机制和治疗、预后等策略已经取得了一定的研究进展，但并发症因其复杂的发病模式，之间复杂的网络难以清理并形成整体的分析结果来指导下一步的研究，因此高质量的系统评价和专家的临床共识是目前可靠的汇总知识的有效方法之一，但由于生物医学文献的数量的快速增长，这种方式难以跟上文献知识的快速增长和更新。相比之下，本研究结果表明，我们提出的自动化因果路径分析框架，有效地聚合了以糖尿病-糖尿病肾病为中心的生物医学文献知识，并可以识别相似的生物医学概念，揭示实体之间的关系，可以从快速增长的生物医学文献中链接其中稀疏的知识，从而促进新知识的发现。

进一步，本研究存在进一步改进这一框架的可能性。首先，本文的知识图谱数据来自 SemMedDB, 这是一个包含从 PubMed 文章中提取的三元组形成的开源数据库，但该数据库是基于 SemRep 工具进行的关系抽取，整体的准确率较低，因此未来可基于 BERT 或其他自然语言处理技术进一步提高信息提取的准确性，提高数据集的质量。其次，除了从生物医学文献中提取知识以外，还可以从认可度较高的生物医学数据库中合并三元组，进一步丰富知识图谱。

总的来说，本研究提出的因果路径框架可辅助临床医生从糖尿病及其并发症的复杂疾病网络获得关键信息，并可以通过降低算法阈值，获取不同的信息。同时该框架具有较强的灵活性，可以应用于药物重定位、多组学应用以及其他疾病的临床决策支持。

PU-115

基于用户画像的情报跟踪服务医学科研探讨

李海生，张新颜，王梅，舒林，章敏，鲍兵，柴媛，孙学文
西部战区总医院医学信息数据室

1 引 言

大数据时代，医疗卫生领域的科技人员对知识的需求越来越精细、多元化[1]，传统的图书情报信息资源组织与服务方式，已无法满足他们需求，为此也促使医院图书馆尽快实现转型升级。目

前，医院图书馆已经逐渐从以纸质图书信息资源为主体的“有形实体”向以电子数据资源为主体的“无形虚体”转变，服务途径与方法也从原先的“面对面的线下交流”向“点对点的线上交互”转化，服务效率实现了质变。图书情报领域中用户画像式服务方式的引进，在一定程度上将情报精准服务推向了一个更高的层级。对医学科研实施情报跟踪服务是笔者多年来一直跟进的一个科研课题，本文将结合自身工作实际，从用户画像的概念、用户画像式精准服务及用户画像式情报跟踪科研定题服务系统、科研全程服务系统与科研成果服务系统等方面，就医院图书馆图情人如何采用用户画像式情报跟踪服务医学科研进行分析与探讨。

2 用户画像的概念

用户画像方式早先用于生产与服务领域。用户画像的概念最早是由被称为交互设计之父的 Alan Cooper 提出，他指出用户画像是真实用户的虚拟代表，是建立在用户真实数据之上的特殊模型[2]。Cooper 也是最早将用户画像方式应用于计算机软件领域，旨意改善产品与服务的用户体验。Teixeira C 等研究者认为，用户画像是一个独立的描述用户需求、偏好和兴趣的模型，是从海量数据中筛选、整合、提炼的个人信息的数据集[3] 国内学者曾建勋认为用户画像是指获取用户的专业背景、学历情况、获取知识习惯、兴趣偏好、特长技能等与用户趋向性需求相关的信息，以此为基础为用户制定模型化特定标签[4]。国内外学者对用户画像的概念都是基于用户真实数据而表述与概括的。

3 用户画像式精准跟踪服务

根据国内外学者对用户画像的概述，结合实际工作，笔者认为对医学科研实施用户画像式精准情报跟踪服务，是促进医院整体临床科研水平、提升临床一线技能的重要举措。

3.1 情报跟踪服务科研[5-6]

图书情报跟踪服务医学科研，即对用户从科研选题定题、科研过程、科研结题到成果转化与推广的全程全方位服务。在科研选题、定题、立项上，依据委托人提交的《医药卫生科技查新委托协议书》中项目主要内容及查新要求的查新点（创新点）、新颖性与方法、技术路线等，主动帮助科技人员检索国内外相关学术领域前沿文献信息，并与查新点进行比对，提示委托人（项目负责人）对项目创新点进行修订、补充、完善；科研中阶段性的跟踪科研进程，定期与课题（项目）负责人交流沟通课题进展情况，及时收集国内外相关学术领域前沿科技文献，并形成二次文献提供给项目科研团队，让科技人员同步了解掌握科研前沿信息，以对照调整科研方法与路径，同时协助科研人员进行项目研究阶段性小结与评估；科研成果转化与推广的情报跟踪，即将类似的国内外科研成果在临床中转化情况，经搜集、整理与分析形成二次信息，及时提供给成果转化人借鉴，让项目研究成果切实提升临床诊疗技能，让新技术、新方法、新专利和新药品及早惠及广大患者。

3.2 用户画像式情报跟踪服务

基于用户画像式图书情报跟踪服务，是建立在通过后台对用户（医药卫生科技人员）静态信息（姓名、性别、年龄、学历、毕业院校、职业、职称、研究方向、擅长特长等）、动态信息（注册内容、登录网页、搜索领域、浏览方向、访问数量、点击内容、跳转内容与次数、订单消息、下载频次、评价情况等）进行综合、提炼、挖掘，进行模型化而实施的特定标签式用户画像服务。

用户画像生成后，根据各数据间关联程度和用户需求进行爱好（偏好）与兴趣趋向分析，或通过网络爬虫技术、网页抓取工具、大数据等，跟踪搜寻用户在图书馆提供的中外文数据库等数字资源中的行为信息数据“痕迹”，如设置“胰腺炎”或“肝胆”等关键字，对用户已经浏览查阅过的专业领域的次数、频率及点击量等进行统计分析；通过这种数据挖掘的方式获取用户客观信息数据，从而实现用户画像式情报跟踪服务科研[7-9]，比如：选题定题服务中的二次文献推送、科研全程情报跟踪与成果转化推广情报服务。

4 用户画像式情报跟踪医学科研服务系统构建

4.1 数据采集与处理 数据采集是从数字图书馆原始服务系统和用户使用的网页中搜集用户静态与动态数据，纳入用户行为日志，形成用户数据集。利用数据分析模块对用户数据进行清洗、整理与分析，获得与用户特征、偏好与行为相关的权重标签，经整合处理后输送至用户画像模块，为绘制用户画像做足准备。在用户特征、偏好及行为分析过程中，采用统一标记的方式，依据标记规则设置用户对应的静态标签；动态标签可采用数据挖掘算法进行分析而设置，通过聚类算法区分不同性质用户，通过相应的数据抽取方式获得对应的用户动态信息标签[1，10-11]。

4.2 用户画像模型分析 医院图书馆建立用户画像后，对组成画像要素进行全方位分析，如认知偏差、兴趣偏好、行为习惯、擅长领域、研究方向等，经精准掌握用户需求，为用户提供最新、最

前沿的文献信息资源，提高图书情报精准服务效率。医院图书馆可以借助层次分析法，构建用户画像层次结构模型，逐层进行计算与检验，确定不同层次因素的对应权重。对比不同的权重因素，加权得到合成权重，掌握用户画像模型对于用户需求服务目标的影响，检验用户画像的服务效率，如用户浏览频率、查阅内容、信息咨询频次等，对用户画像的服务效能影响较大。医院图书馆可以将权值较大的标签作为提供个性化服务的重要依据[12-15]。

4.3 构建用户画像式情报服务系统

4.3.1 科研选题定题情报跟踪服务系统

用户画像式科研选题定题情报跟踪服务系统（见图 1）：图书情报人采集用户（医学科技人员）静态信息、动态信息数据，形成粗放型数据采集模块，通过聚类分析法，对静态数据与动态数据进行整合、挖掘与提炼，形成数据整合模块；再经过对整合模块的信息技术处理，即分类与组合，形成一类群体特征性“标识性数据”，对号入座式识别各类用户群体，从而绘制出各类群体画像。根据群体信息数据画像分别进行如下分析，首先针对用户兴趣爱好（偏好），提炼出相关关键词、主题词，在数据库或互联网中搜索出相关领域的原文献，经整理、筛选、过滤形成二次文献；其次是根据用户所擅长的学术领域、学术专长，检索出国内外相关领域的前沿文献信息，结合用户需求整理、挖掘形成二次文献；最后是从用户对学术的认知角度入手，分析其认知维度，即用户在其所擅长的学术领域提倡与质疑、支持与规避、倾向与减持等，搜索出用户所提倡、支持与倾向的学术领域前沿文献，利用文献分析方法，经筛选、过滤、挖掘形成二次文献。将经过信息挖掘形成的用户兴趣爱好、擅长领域与认知维度三个方面的二次文献，构建科研选题定题二次情报精准推送模块，形成科研选题方向、技术路线建议，便于用户参考借鉴与使用，同时进一步构建数据可视化模块推送用户[16-18]。

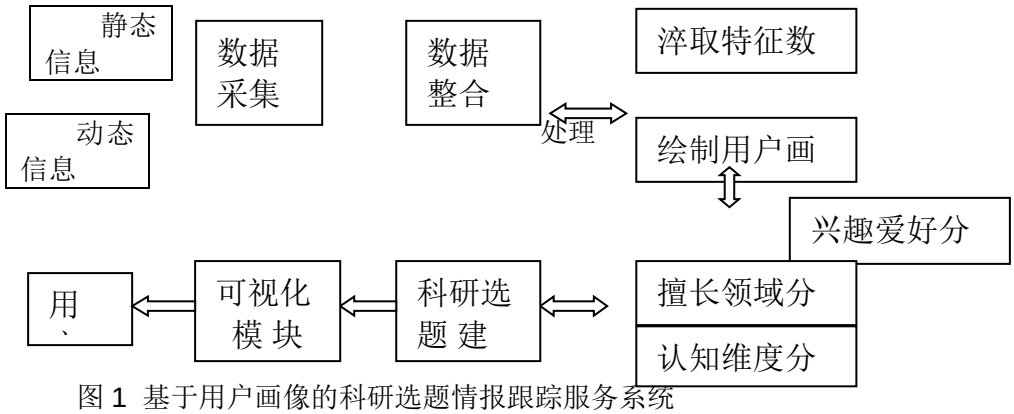
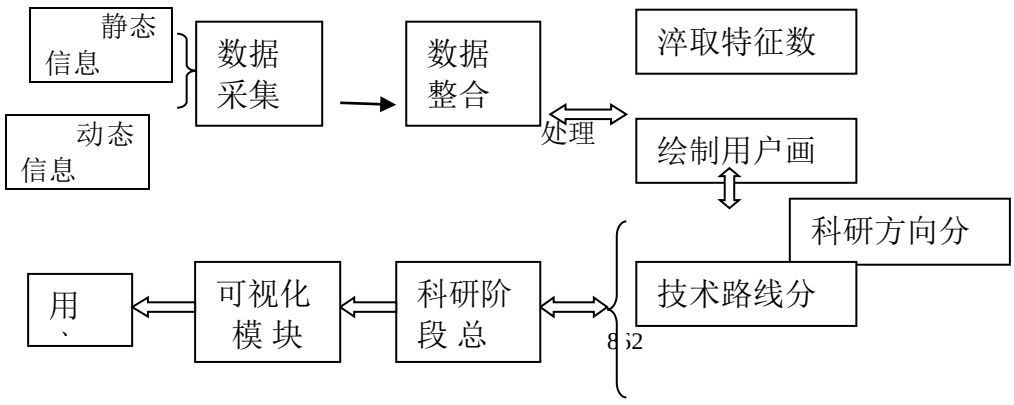


图 1 基于用户画像的科研选题情报跟踪服务系统

4.3.2 用户画像式科研全程情跟踪报服务系统

用户画像式科研全程情跟踪报服务系统（见图 2）：生成用户画像前的用户静、动态信息采集、数据整合、淬取用户特征数据等与用户画像式科研选题定题情报跟踪服务类同，对在研项目采用用户画像式对号入座后，针对科研全程实施情报跟踪服务，即主要在科研方向、技术路线及科研阶段性成果小结等方面进行情报服务，情报人利用自身检索知识、技能与技巧，针对科研进程搜集国内外相关领域科技前沿文献，形成二次文献及时提供给科技人员，让其在科研方向、技术路线等方面进行比对与借鉴，以便及时调整、完善，以免学浪费科研资源，造成科研项目低水平重复。同时，协助科技人员对科研项目进行阶段性或中期小结，总结前期科研成果，高效开启后期科研进程。



传输

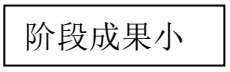


图 2 基于用户画像的科研全程情报跟踪服务系统

4.3.3 用户画像式科研成果转化及推广情报跟踪服务系统

用户画像式科研成果转化及推广情报跟踪服务系统（见图 3）：生成用户画像后，对科研成果转化及推广情报跟踪服务，主要体现在基于用户画像的成果转化与推广，亦即在用户科研成果模型化特征基础上，对成果转化、应用与推广进行分析、评估，包括成果转化的可能性、成果应用的可行性与成果推广路径、领域等的评价、预判。

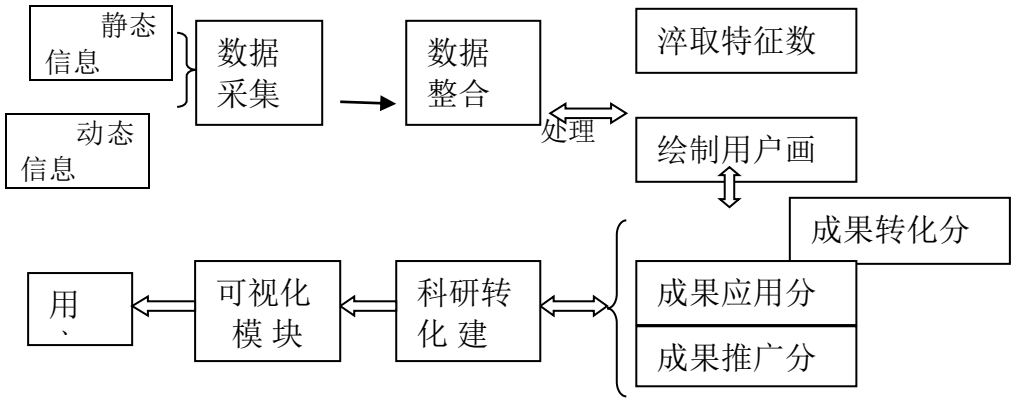


图 3 基于用户画像的科研成果转化与推广情报跟踪服务系统

5 讨 论

笔者研究的情报跟踪服务医学科研模式，缺乏对科技人员科研工作的“预见性”“前瞻性”引导，遵循的仅是“先有项目，后有服务”的较为“被动”研究思路，难免产生“应付式”的服务效率与服务质量。而此文研究的用户画像式图书情报跟踪服务医学科研模式，开启的是“主动服务，产生项目，主动跟踪”的“启发式”精准服务，亦即用户“有潜在需求即服务”[19-20]。

用户画像式情报跟踪服务医学科研，除了可以遵照选题定题服务、科研全程情报跟踪服务与科研成果情报跟踪服务系统开展工作外，还要求图情人具备一定的信息素养与科研素养，主要包括熟悉并掌握国内外主要医药卫生数据库的检索流程、方法、技巧，具备一定的相关专业知识背景、外语水平与科研工作经验，这也是我们进一步研究的方向。

PU-116

Faculty Opinions 最新评分的多维数据源下生物医学科技论文评价指标研究

许丹 韩爽 鞠昕蓉 王雯霞 徐爽

1 引言

Faculty Opinions 作为全球生物医学领域内最为广泛的科技论文同行评议数据库。世界范围内近 90%的顶尖生命科学研究机构都在使用 Faculty Opinions 进行文献推荐，这里汇集了超过近万名生命科学家的观点，突出他们认为重要的、开创性的、对科学家来说必不可少的文章[1]。推荐人员通过其丰富的专业知识进行严格筛选，可以保证推荐文献的高质量和高效能，因此不再需要浪费时间搜索最值得注意的影响你研究的发现。Faculty Opinions 不仅展示了最受欢迎的研究成果，也展示了那些不太为人所知的经常被遗忘的瑰宝[1]。

在 2021 年 4 月，Faculty Opinions 再次重新更新了评分计算标准，引入新的专家意见和建议评分并结合相对引用率这一指标重新进行的计算赋予新的得分[2]，并在结果页面显示出该篇文章在整个学科领域中的排名等级、领先占比情况。最新得分中包括了相对引用率（RCR）和星级权重（WSS）指标，加权星数代表着专家的的意见得分，相对引用率代表着文章的文献计量学性能如何，在统计运算中分别确定了两者的相对权重的比率，使得专家推荐优先于文献计量学数据[2]。如图 1 所展示：左侧显示了经过一系列计算后得到的最新评分 Score 为 325.6，这一评分主要包含两部分，其一是 RCR 为 1.57，其二是 WSS 为 70；下面显示了该篇文章所在领域中的排名情况，图

中显示该文排在基因组学和基因遗传学领域中 Top 0.001%的位置；右侧则显示了文章被推荐的总次数和不同推荐的评级情况，图中显示该篇文章一共被专家推荐了 20 次，推荐为 Exceptional 三星级的有 15 次，推荐为 Very Good 二星级的有 5 次。

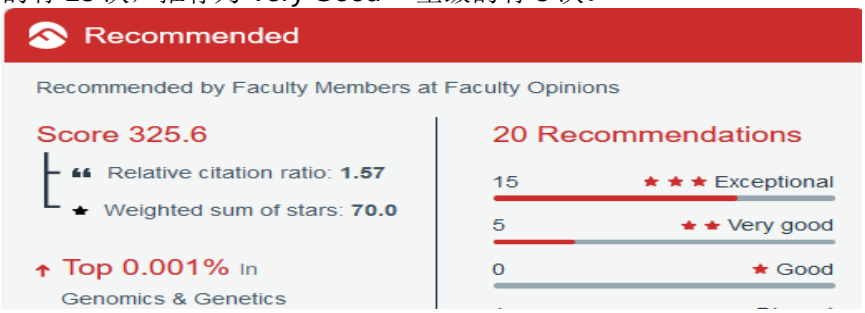


图 1 最新评分解读图片

RCR 最早是由学者 Hutchins BI [3]提出的一种运用被引率来衡量文章水平上的影响的一种新的文献计量学度量方法。相关研究报道较少，Bornmann L[4]将 RCR 指标与归一化指标和同行评议指标进行了调查研究，发现 RCR 与归一化指标高度相关，但与同行评议指标的相关性仅为低或中等。国内学者陈斯斯[5]也对 RCR 与 F1000 评分的相关性进行了验证研究，得出结论与 Bornmann L 一致。Faculty Opinions 推出的最新的评分将专家意见评价与文章的文献计量性能结合起来，增强了对于最终获得高水平的引用和有重大影响的文章的预测能力[2]。

当前有不少学者和业内同仁认为，在反“四唯”、不“唯 SCI”等意见指导下，我们应自觉淡化和抵制 SCI“标尺”。然而，不“唯 SCI”并不是完全摒弃不用 SCI，而是不把 SCI 当作唯一的、普适的评价标准。SCI 评价工具还是具有其文献计量学上的参考价值，尤其是在新的更为科学、全面、公平、公认的评价工具尚未完全确立之前，我们仍可通过它的某些科学应用来了解世界科技期刊的发展情况，帮助、鞭策我们获得更快的进步[6]。因而纳入更多不同类型的多维数据源，对于生物医学科技文献进行多元化的综合、全面评价则显得更为重要，同时对于不同评价类型指标之间的关联如何进行探析。

有研究指出，Scopus 作为与 WOS 同类的引用数据库，在单篇论文的相关计量数据中，不仅包含引用相关指标(CiteScore)等，还有学科标准化后的科研论文影响力特有指标 Field-Weighted Citation Impact-FWCI)，期刊的相关指标 SNIP（标准化影响系数）和 SJR（威望指数）等，同时具备 Plum Analytics 合作提供文献的 Altmetrics 社会影响力相关指标数据，这些指标都为科技论文评价的多元化提供助力[7]。SJR 全称 SCImago Journal Rank，是来自西班牙的 SCImago 团队基于 Scopus 数据库提出的一个指标，其定义是利用 Google 的 PageRank 算法来测量期刊的声望，并且考虑了期刊的选题和声望对其引文价值的影响，其作用是测量基于 Scopus 数据库的期刊声望[8-11]。SJR 算法较为复杂，是一个既考虑了期刊被引数量，又考虑了期刊被引质量的指标，反映了期刊的声望指数。这是除了期刊影响因子 Impact Factor 外对期刊的另一个评价指标，SJR 重引用数量，更重引用价值，对期刊的评估更为全面，尤其是非英语期刊，她指出一种期刊越多地被高声望期刊所引用，此期刊的声望才越高[8-12]。标准化影响系数（SNIP）是 Elsevier2012 年发布的影响因素，以 Scopus 数据库为基础评估期刊影响力。SNIP 根据某学科领域的总引用次数确定引用权重。因此，单一引用出现在不太可能引用的学科领域时会被赋予较高数值。SNIP 纠正了不同学科领域期刊的引用行为差异[8-12]。

我国有多位不同学者结合同行评议 F1000 评分指标和 Altmetric、WOS、学术迹等不同数据库中不同类型评价指标对科技文献进行评价指标之间相关性的分析报道研究[13-15]，均是对于改版之前的 Faculty Opinions 评分进行的研究，而基于 Faculty Opinions 最新评分的相关研究鲜有报道。本研究以 Faculty Opinions 推出的最新评分为基础，探讨该评分内部之间和与不同数据源下的不同类型评价指标之间的关系如何，从而发掘不同数据源下的多元化各类指标之间的隐含关联，为我国生物医学文献论文综合而全面的评价提出新的参考。

2 材料与方法

2.1 数据源选取

考虑到文献发表后各类引用指标数据出现的时间差和不同峰值情况 [16]，本研究选取了 Faculty Opinions 中推荐的发表于 2019 年的肿瘤 Oncology 和生物化学 Chemistry 两个学科文献进行研究。经查询获取到肿瘤学科推荐文献 380 篇，生物化学学科推荐文献 346 篇，用 PMID 作为唯一

标识符号。然后再分别获取每篇推荐文献的各类指标数据，包括如下：推荐得分(Recommendations Score，RS）、推荐次数(Recommendations Times，RT)、相对引用率（Relative citation ratio，RCR）、星级权重（Weighted sum of stars，WSS）、Top %（所占学科领域排名百分比）、Exceptional、Very Good、Good，同时对每篇文章给出的不同评价标签情况予以标注。相比于旧版的 Faculty Opinions 评分计算上，新版评分多提供了相对引用率、星级权重和文章所占学科领域的 Top 百分占比情况的这三个新的指标。

运用 PMID 与 WOS 入藏号进行标识链接，分别在 WOS 和 Inictes 数据库中，获取基于 WOS 的学科规范化的引文影响力(CNCI)、期刊规范化的引文影响力(JNCI)、期刊影响因子(IF)和被引频次（TC）指标。肿瘤学科 380 篇文献中有 11 篇文献不被 WOS 收录，WOS 覆盖率为 97.1%；生物化学学科 346 篇文献中有 9 篇文献不被 WOS 收录，WOS 覆盖率为 97.4%。在 Altmetric 平台中分别获取 Altmetric Score、Twitter、Mendeley 指标，在 Scopus 中分别获取 CiteScore 和 FWCI 指标，济南泉方云图书馆中分别获取 SJR 指数和 SNIP 指标，上述指标均以 PMID 做为唯一标引。

2.2 方法

使用 SPSS Statistics 23 软件对上述 4 个数据源下的不同类型指标数据进行斯皮尔曼相关性比较，并且分类进行对比分析得到研究结果。

3 研究结果

3.1 生物化学学科不同类型评价指标数据的斯皮尔曼相关性比较结果

笔者对 346 篇生物化学学科文献的不同类型评价指标进行相关性比较，结果见表 1。研究发现：同行评议指标内部，RS 与 RT、Exceptional 和 Very Good 成中度相关；与最新评分计算而引入的 RCR 也成中度相关，相关系数为 0.551,；RS 与 WSS 为高度相关，相关系数为 0.938，与 Good 为低度相关。RS 与 Top %显示为负的高度相关关系。RT 与 RCR 为低度相关，与 Very Good 为高度相关，而与其他指标均为中度相关关系。RCR 与 Top %相关系数为-0.469 ，和 Good 不相关，而与 WSS 和不同评价星级均为低度相关关系。WSS 与 Top %为负的高度相关关系，而与 RT、Exceptional、Very Good 成中度相关关系，与 RCR 和 Good 为低度相关。我们可以看到，Faculty Opinions 最新评分后，同行评议内部几个指标与星级权重（WWS）的相关性和所占学科领域排名百分比相关性的绝对值都要高于相对引用率(RCR)的相关性，这也说明了改版后的 Faculty Opinions 评分中星级权重对于同行评议得分更加重要。Top %与 RS 和 WSS 成负的高度相关，相关系数分别为-0.766 和-0.712，Top %与 RCR 和 RT、Exceptional、Very Good 均为负的中度相关，与 Good 不相关。推荐得分和星级权重与该文所在领域内的百分比成负的高度相关关系，这说明了推荐得分和星级权重越高，百分比分值越小，则说明该文章在领域内排名越靠前，越重要。

同行评议指标与外部其他类型指标相关性比较发现：RS 与社会计量学各指标均为中度相关，这与笔者之前研究结论一致[17]。社会计量学各指标与 RT、WSS、Exceptional、Very Good、Good 均为低度相关，与 Top %为负的低度相关关系；只有 RCR 与 Altmetric Score、Twitter 成中度相关关系，而与 Mendeley 相关系数为 0.821，为高度相关。CiteScore、FWCI、SJR、SNIP 均与 RS 为中度相关关系，与 RT、WSS、Exceptional、Very Good、Good 均为低度相关或不相关，与 Top %成负的低度相关。RCR 与 FWCI 为高度相关，相关系数达到 0.941，与 CiteScore、SJR、SNIP 均为中度相关。JNCI 与同行评议中各个指标均为低度相关或不相关，CNCI、IF、TC 与 RS 为中度相关，与 RT、WSS、为低度相关，与 Top %为负的低度相关，与各评价星级不相关。WOS 中的 CNCI 和 TC 指标与 RCR 均为高度相关，相关系数分别是 0.929 和 0.958，与 IF 为中度相关关系。

Altmetric Score、Twitter、Mendeley 与 Scopus 数据库中的 CiteScore、FWCI、SJR、SNIP 均为中度相关关系，除了 Mendeley 与 FWCI 为高度相关关系，相关系数为 0.802。Altmetric Score、Twitter 与 WOS 数据库中的 CNCI、IF、TC 均为中度相关，Mendeley 与 JNCI 为低度相关，与 IF 为中度相关，与 CNCI 和 TC 均为高度相关。CiteScore 与 FWCI 为中度相关，而与 SJR 和 SNIP 均为高度相关，相关系数分别为 0.930 和 0.954。SJR 和 SNIP 也为高度相关关系，相关系数为 0.878。CiteScore、SJR、SNIP 均与 JNCI 为负的低度相关，与 CNCI 和 TC 为中度相关，与 IF 为高度相关。FWCI 与 JNCI 和 IF 为中度相关，与 CNCI 和 TC 为高度相关。Scopus 数据库中的 CiteScore 是由根据一系列算法给出了的评分，与 WOS 数据库中一篇文章的 IF 值相类似，因此得到高相关性，也说明了两个数据库虽然数据源评价算法虽不同，但对于同一篇文章给出的评价

评分结论是一致的。SJR 是一个期刊的威望指数，与 WOS 中期刊影响因子也成高度相关性，两类数据库中对于期刊的评价也是相类似的。

3.2 生物化学学科不同评价标签与不同类型指标数据的斯皮尔曼相关性比较结果

本研究选取了生物化学学科文献中所标注的 6 种覆盖率较高的代表性评价标签进行研究，标签分别是 Confirmation、Good For Teaching、Interesting Hypothesis、New Finding、Technical advance 和 Novel Drug Target，研究其内部之间和与不同类型指标数据的相关性如何，具体结果见表 2。研究结果发现：Confirmation 与 Good For Teaching 相关系数为 0.494，是低度相关关系，与其他标签无相关性。Good For Teaching 与 Interesting Hypothesis、New Finding 也分别为中度相关。Interesting Hypothesis 与 New Finding、Novel Drug Target 为中度相关，New Finding 与 Technical advance、Novel Drug Target 也为中度相关。其他不同标签之间大多为中低度相关关系或者无相关关系。

RS、WSS 均与不同标签类型大多程中低度相关关系，相关系数介于 0.454 至 0.540 之间；TS 与不同标签类型多数呈中度相关，只有 New Finding、Technical advance 与 TS 为高度相关关系，相关系数分别是 0.854 和 0.736；RCR 与不同标签成低度相关或不相关；Top %与不同标签大多成负的低度相关性，只有 Good For Teaching、Technical advance 与 Top %为负的中度相关关系。Exceptional 与 Good For Teaching、Technical advance 为中度相关关系，与 New Finding 为低度相关，与其他标签不相关；Very Good 与 Confirmation、New Finding 、Technical advance 为中度相关，与 Good For Teaching、Interesting Hypothesis、Novel Drug Target 为低度相关；Good 与 Good For Teaching、Interesting Hypothesis、New Finding 分别为高度、中度、低度相关关系，与其他标签不相关。社会计量学各指标 Altmetric Score、Twitter、Mendeley 与不同标签大多均为低度相关关系，与少数标签之间为不相关。Scopus 数据库平台下的 FWCI、CiteScore、SJR、SNIP 大多与 Good For Teaching、New Finding、Technical advance 均为低度相关关系，与其他各标签不相关。各类标签与 WOS 数据库中的 CNCI、IF、TC 大多数为不相关，少数为低度相关；与 JNCI 均不相关。

3.3 肿瘤学科不同类型评价指标数据的斯皮尔曼相关性比较结果

笔者对肿瘤学科 380 条文献记录进行斯皮尔曼相关性比较分析，得到如下结果：见表 3。肿瘤学科的文献同行评议指标内部，RS 与 RT、Exceptional、Very Good、Good 均为低度相关，与 RCR 为中度相关，与 WSS 为高度相关，相关系数达到 0.930，与 Top %为负的中度相关关系，相关系数为-0.692。RT 与 RCR、WSS 为低度相关，与 Exceptional、Very Good、Good 均为中度相关。RCR 与 WSS 为低度相关，与 Exceptional、Very Good、Good 均为不相关，说明了肿瘤学科中相对引用率在同行评议的不同星级里没有相关性。WSS 与 Exceptional、Very Good、Good 均为中度相关，与 Top %为负的中度相关，相关系数为-0.646。Top %与 Exceptional、Very Good、Good 均为负的低度相关关系，可见肿瘤学科中文章在该学科领域中排名百分占比情况与评价为不同的星级没有太大相关关系。不同评价星级与 WSS 的相关性要高于与 RCR 和 Top %的相关关系。

同行评议指标与外部其他类型指标相关性比较发现，Altmetric Score、Twitter、Mendeley 与 RS 均为中度相关关系，与 RT 和 WSS、Exceptional 均为低度相关，与 Top %为负的低相关关系，与 Very Good、Good 不相关。RCR 与 Altmetric Score、Mendeley 均为高度相关关系，相关系数分别为 0.714 和 0.874，与 Twitter 为中度相关，相关系数为 0.657。Scopus 数据库中的 FWCI、CiteScore、SJR、SNIP 与 RS 均为中度相关，与 RT、WSS 为低度相关，与 Top %为低度负相关，与不同评价星级均为不相关。RCR 与 FWCI 为高度相关关系，相关系数为 0.926，与 CiteScore、SJR、SNIP 为中度相关。WOS 数据库中 JNCI 与 RCR 为中度相关，与 RS 为低度相关，与 Top %为负的低度相关，与其他不相关。CNCI、IF、TC 与 RS 均为中度相关，与 RT、WSS 均为低度相关，与 Top %为负的低度相关。RCR 与 CNCI 和 TC 为高度相关关系，相关系数分别为 0.941 和 0.969，与 IF 为中度相关，相关系数是 0.687。

社会计量学各个指标与 Scopus 和 WOS 数据库中不同类型指标相关性比较研究发现，Altmetric Score、Twitter、Mendeley 除了与 JNCI 均为低度相关外，分别与 CiteScore 、SJR、SNIP、IF、TC 均为高度相关。FWCI 和 CNCI 分别与 Altmetric Score、Mendeley 为高度相关，与 Twitter 为中度相关。Scopus 数据库中 CiteScore 与 FWCI 为中度相关，而与 SJR、SNIP 成高度

相关，相关系数分别为 0.963 和 0.955。FWCI 分别与 SJR、SNIP 成中度相关关系，SJR 与 SNIP 为高度相关。Scopus 数据库中各个指标与 WOS 数据库各指标相比较，除了 FWCI 与 JNCI、IF 为中度相关，JNCI 与 CiteScore、SJR、SNIP 不相关外，其他不同指标之间相关性均为高度，相关系数在 0.704-0.966 之间。WOS 数据库中，JNCI 与 CNCI、TC 为中度相关，与 IF 不相关。TC 分别与 CNCI、IF 为高度相关，CNCI 与 IF 为高度相关。

3.4 肿瘤学科不同评价标签与不同类型指标数据的斯皮尔曼相关性比较结果

本研究同样选取肿瘤学科文献中所标注的 7 个覆盖率较高的标签进行分析，分别是 Confirmation、Controversial、Good For Teaching、Interesting Hypothesis、New Finding、Technical advance 和 Novel Drug Target，研究其内部之间和与不同类型指标数据的相关性如何，具体结果详见表 4。研究结果表明：肿瘤学科的不同标签内部之间，Confirmation 与 Good For Teaching 和 New Finding 均为低度相关关系，Controversial 与 Good For Teaching 为中度相关，与 Technical advance 和 Novel Drug Target 相关系数为 1，这可能由于案例数极少产生的偏差。Good For Teaching 与 Interesting Hypothesis、New Finding 均为高度相关关系，相关系数在 0.8 以上，与 Novel Drug Target 为中度相关，相关系数为 0.528。Interesting Hypothesis 与 New Finding 和 Novel Drug Target 均为中度相关，New Finding 与 Novel Drug Target 相关系数是 0.748，其他标签之间无相关关系。同行评议指标中，RS 与 Confirmation、Good For Teaching、Interesting Hypothesis、New Finding 为低相关，与 Novel Drug Target 为中度相关；RT 与 Confirmation、Interesting Hypothesis 和 Novel Drug Target 为中度相关，与 Controversial 为低度相关，与 New Finding 和 Good For Teaching 为高度相关。RCR 只与 Good For Teaching、Interesting Hypothesis、New Finding 为低度相关，与其他标签不相关。WSS 与 Controversial、Good For Teaching、Interesting Hypothesis、New Finding 为低相关，与 Novel Drug Target 为中度相关。Top%与 Controversial、Good For Teaching、Interesting Hypothesis、New Finding 为负的低度相关，与 Novel Drug Target 为负的中度相关。Confirmation 和 Technical advance 与不同评价星级均无相关性，Controversial 只与 Good 评价星级相关系数为 1，与其他评价星级不相关。Good For Teaching 和 nteresting Hypothesis 均与 Exceptional、Very Good、Good 三个评价星级分别成高度、中度、高度相关关系。而 New Finding 与 Exceptional、Very Good、Good 三个评价星级分别成中度、高度、低度相关关系。Novel Drug Target 只与 Exceptional、Very Good 成中度相关。

社会计量学指标相关分析中发现，大多标签与三个代表性指标 Altmetric Score、Twitter、Mendeley 均无相关性，只有 Good For Teaching、Interesting Hypothesis、New Finding 与三个代表指标均成低度相关关系。Scopus 平台下，CiteScore 只与 Good For Teaching 和 New Finding 成低度相关关系，与其他标签无相关关系。FWCI 只与 Confirmation、Good For Teaching、New Finding 和 Interesting Hypothesis 为低度相关，与其他标签也无相关关系。SJR 只与 New Finding 为低相关，相关系数为 0.207；SNIP 与 Good For Teaching 和 New Finding 为低度相关，相关系数分别是 0.262 和 0.180。WOS 平台下各指标中，JNCI 与各个标签均无相关性，CNCI 和 TC 均只与 Good For Teaching、Interesting Hypothesis、New Finding 均成低度相关性，与其他标签无相关性，而 IF 只与 New Finding 为低相关，相关系数为 0.193。

4 讨论

4.1 Faculty Opinions 最新评分指标内部分析

Faculty Opinions 更改了最新评分标椎后，本研究对生物化学学科与肿瘤学学科的文献中 Faculty Opinions 内部各类指标进行对比分析发现：无论是基础科学类的生物化学文献，还是实践科科学类的肿瘤学文献，RS 与 WSS 的相关性是最高的，系数均在 0.9 以上，这也反映了星级权重在 Faculty Opinions 最新评分中的绝对地位。RS 与 Top%的相关性的绝对值在 0.7 上下，呈现出负的中度或高度相关关系。RS 与 RCR 相关度均为中等，相关系数在 0.5 左右，不同学科的 RS 与 Good 一星级均显示低度相关性，与 Very Good、Exceptional 的相关性要略高。不同学科的 RT 与 Very Good 的相关性最高，都要高于其与另外两个评价星级的相关性。不同学科中，RCR 与不同评价星级均显示低相关或不相关关系，这也表明了相对引用率对于评价星级来说，相关度不大；而 WSS 与 Very Good、Exceptional 均显示为中度相关性，这说明了星级权重的多少与该篇文章被评价为三星级和二星级具有中等程度的一致性。Top%这一指标在不同学科类别中，均显示了其与 RS

和 WSS 相关性的绝对值要高于其与 RT 和 RCR 相关性的绝对值，即 RS 和 WSS 越高，则 Top% 越小，标明该文章在学科领域中越处于领先排名地位。

不同评价标签类型对于 Faculty Opinions 内部各类指标进行对比分析发现：不同学科的文献，基础类学科中 RS 大多与各个标签成低度相关或中度相关关系，实践类学科中 RS 大多与各个标签成低度相关或不相关关系。基础类学科中 RT 与不同标签成高度或中度相关性，与 New Finding 和 Technical advance 为高度相关，其他为中度相关；实践类学科中，RT 与不同标签分别形成不同的相关性，与 Controversial 和 Technical advance 为低度相关，与 Confirmation、Interesting Hypothesis、Novel Drug Target 为中度相关，与 New Finding 和 Good For Teaching 为高度相关。不同学科中，RCR 分别与不同的标签大多数为低度相关或不相关关系。WSS 在基础类学科中与不同标签多数呈低度或中度相关性，在实践类学科中，除了与 Novel Drug Target 为中度相关，其他为低度相关或不相关。Top%在基础类学科中与 Technical advance 相关性绝对值最高；在实践类学科中与 Novel Drug Target 相关性绝对值最高。不同类学科中，标签 Good For Teaching 和 New Finding 与三个评价星级的相关系数所显示的趋势走向均一致，即 Good For Teaching 与 Very Good 相关性<Good For Teaching 与 Exceptional 相关性<Good For Teaching 与 Good 相关性;New Finding 与 Good 相关性<New Finding 与 Exceptional 相关性<New Finding 与 Very Good 相关性。

4.2 Altmetric 平台社会计量学各指标分析

不同学科类别中，社会计量学指标 Altmetric Score、Twitter、Mendeley 分别与 RS 呈中度相关，与 RT、WSS 为低度相关，与 RCR 相关性要高于与其他指标的相关性，Mendeley 与 RCR 相关系数是最高，与 Top %相关性是负低度相关。Altmetric Score、Twitter、Mendeley 与三星级 Exceptional 和二星级 Very Good 大多数呈低度相关关系，与一星级 Good 均不相关。

不同学科类别中，不同评价标签与 Altmetric Score、Twitter、Mendeley 各指标大多均显示为低度相关关系。生物化学学科中，Twitter 与 Good For Teaching 和 Novel Drug Target 相关性是最高的。肿瘤学科中，社会计量学各指标与标签 Confirmation、Controversial 均不相关，与标签 Technical advance 的相关性是最高的。

4.3 Scopus 数据库中各个指标分析

不同学科类型文献在 Scopus 平台下，CiteScore、FWCI、SJR、SNIP 均与 RS 成中度相关，与 RT 为低度相关，与 RCR 大多也为中度相关，其中 FWCI 与 RCR 是高度相关关系，相关系数在 0.9 以上。各指标与 WSS 均为低度相关，与 Top %为负的低度相关关系。Scopus 平台下的各个指标均与同行评议中的不同评价星级显示无相关关系，而与 Altmetric 平台下社会计量学各个指标显示出高度相关关系，这表明了 Scopus 平台与 Altmetric 平台对于论文评价给出了高度一致性的结论。

Scopus 平台下指标与 WOS 平台下各指标相关性比较发现：两个平台下的大多数评价指标之间的相关性较高。肿瘤学科类别中，除了 FWCI 与 JNCI 为中度相关关系，其他指标与 JNCI 之间无相关性；而 CNCI、IF、TC 与 CiteScore、FWCI、SJR、SNIP 均显示为高度相关关系。生物化学学科中，FWCI 与 JNCI 也为中度相关关系，JNCI 与 CiteScore、SJR、SNIP 大多数成负的低相关性，其他指标之间大多数显示中度或高度相关关系。特别是 FWCI 与 CNCI 和 TC 相关系数均在 0.9 以上，对于不同学科文献对比研究表明了在单篇论文的评价上，学科标准化后的科研论文影响力与学科规范化的引文影响力和被引频次均展现出高度一致的结论。IF 与 SJR、SNIP 相关系数也在 0.9 以上，表现出高度强相关关系。这也表明了 Scopus 平台中的期刊标准化影响系数和期刊威望指数和 WOS 中的影响因子在评价期刊上的一致性，即威望指数越高的期刊，其 IF 也越高，这也与之前学者研究的结论相一致[14-16]。

Scopus 平台下指标与不同评价标签比较发现：生物化学学科文献中，标签 Good For Teaching、New Finding、Technical advance 与 CiteScore、FWCI、SJR、SNIP 各指标大多数为低度相关关系，其他标签与各指标无相关性。肿瘤学科文献中，标签 Good For Teaching、New Finding 与 CiteScore、FWCI、SJR、SNIP 各指标大多数为低度相关关系，其他标签与各指标无相关性。这与不同标签在 WOS 平台下的相关性研究结果相类似，大多数标签与 WOS 各指标无相关系，少数显示为低相关性[17]。

4.4 WOS 数据库中各个指标分析

不同学科的文献中，JNCI 与 RS 为低度相关，CNCI、IF、TC 均与 RS 为中度相关关系，WOS 各指标与 RT 均为低度相关。RCR 与 CNCI 和 TC 显示为高度相关，相关系数在 0.9 以上，

RCR 与 JNCI 和 IF 显示为中度相关关系。WSS 与 CNCI、IF、TC 均显示低相关性，Top%与 WOS 各指标为负的低相关关系。WOS 各指标与不同评价星级均不相关。

WOS 平台中各个指标与不同评价标签大多数显示无相关性或低度相关。肿瘤学科中，CNCI 和 TC 与标签 Good For Teaching、Interesting Hypothesis、New Finding 显示低相关性。生物化学学科中，CNCI、IF、TC 与 New Finding、Technical advance 呈低相关性。WOS 平台中各个指标与 Altmetric 平台指标对比发现，生物化学学科中，CNCI、IF、TC 与 Altmetric Score 和 Twitter 为中度相关，与 Mendelay 为高度相关。肿瘤学科中，CNCI、IF、TC 与 Altmetric 各个指标大多数呈高度相关关系。JNCI 与不同学科的 Altmetric 各个指标大多数为低度相关或不相关。

5 结论

本研究基于 Faculty Opinions 最新评分对多维数据源下的生物医学科技论文各类评价指标开展了相关性的研究。相较已发表文献，过去的文献均是对 Faculty Opinions 数据库中原有评分和各类指标进行的研究，对最新更新后的评分如何未见相关的研究报道。同时本研究选取的数据源更多，涵盖了 Faculty Opinions、Altmetric、Scopus 和 WOS 四大数据源，涉及了同行评议、社会计量学、和分别以 Scopus 和 WOS 为基础的引文计量相关指标数据，研究范围更加全面和丰富，并且探讨了这些不同数据源下的各类指标相关性如何。通过本文研究得到以下结论：

（1）不同学科类别文章 Faculty Opinions 最新评分中，RS 与 WSS 的相关性最高，显示了星级权重在 Faculty Opinions 最新评分中的绝对地位。RCR 与 RS 相关度均为中等，RCR 与不同评价星级相关度较低低或不相关。RS 和 WSS 越高，则 Top%越小，表明该文章在学科领域中越处于领先排名地位。

（2）Scopus 平台下的各个指标均与 Faculty Opinions 中不同评价星级无相关关系，而与社会计量学各个指标显示出高度相关关系，这表明 Scopus 数据库评价指标与 Altmetric 评价指标对于论文评价给出了高度一致性的结论。

（3）不同学科类别的文章，Scopus 各指标与 WOS 各指标之间的大多相关性较高。在单篇论文的评价上，FWCI 与 CNCI 和 TC 均展现出高度一致的结论。在期刊的评价上，IF 与 SJR、SNIP 也展现了高度的一致性。

（4）不同学科类别的文献 WOS 平台中各个指标与同行评议指标和评价标签大多显示为低度相关或不相关关系，与 Altmetric 社会计量学指标大多为中度或高度相关关系，少数不相关或低相关。

本研究仍存在不足之处，选取了 Faculty Opinions 中 2019 年发表的生物化学学科和肿瘤学科的推荐文献进行研究，L. Bornmann 等[18-19]认为无论对于哪一个学科进行测度,必须要经历 3 年的引文积累周期。本文虽然排出了因发表年限造成的引用量偏倚，但总体样本数量还是略少，后期可选取更多不同学科类别的文献进行大量数据规模的样本研究，使研究结论更加可靠。同时对于 Top%的选定中，本研究只选取了%的占比数的数值，而没有对于每篇文章所占的百分比可能分属于不同具体学科类别进行分别统计。

但本文仍具有一定的研究意义，基于 Faculty Opinions 最新评分下开展不同数据源下各类评价指标研究，可以给与科研人员评价科技文献上一定程度的指导和借鉴，帮助其了解同一篇文章在不同数据源下的评价指标所展示出的评价结果如何，各类评价指标之间相关性如何，为建立多维度下多元化的科技论文评价指标体系奠定基础。

PU-117

从无效审查程序看中国生物医药领域专利质量的关键影响因素

单治易[1,2] 关陟昊[3] 王奎芳[1,2] 孙文君[1,2]
1.中国科学院文献情报中心
2.中国科学院大学经济与管理学院信息资源管理系
3.中国农业科学院农业信息研究所

1 引言

提高专利质量是加速中国实现从知识产权引进大国向知识产权创造大国的重要路径。专利权是国家以法律程序赋予发明人一定期限内的垄断权利，为了保证程序的公正和维护公众利益，任何单位或者个人都有权通过专利无效程序对专利权的有效性提出异议。专利无效宣告程序作为保护创新的重要方式之一，不仅是专利审查制度和专利权制度的延续，还是保护公众利益的必要手段。该制度是“为错误的专利授权而设置的一种纠正程序”，具有纠错作用，抗辩作用和防御作用。此外，专利无效程序还是优化专利质量的一项重要政策工具，其能够剔除不符合法定授权条件的专利，提升专利权利的稳定性和专利的价值。

从全球来看，各个国家都有相应的法律法规对“专利无效程序”或“异议程序”做出规定，如，美国的专利行政确权程序包括多种途径。国外通过专利无效程序解决低质量专利问题，在欧洲，专利无效审查制度使用率为 5%，无效成功率高达 68%。中国专利无效的最终结果包括专利的部分权利无效，专利权利全部无效，专利权全部维持有效。被宣告无效的专利权利要求所涉及的技术的使用不再有约束力，任何单位或个人可以无偿地使用它们，由此获得经济效益。对于专利权人而言专利权利的无效将带来巨大的损失，所以提高专利质量，保证专利权的稳定性有重要意义。

本研究选择中国生物医药领域的涉专利无效程序的数据，识别经过无效程序依然维持有效专利的质量特征，从多维度分析提出优化专利质量的策略，以期为专利申请人在提升专利文本质量方面提供参考质保。

2 数据、指标与方法

2.1 数据来源

专利数据来源于 incoPat 专利数据库。专利数据筛选步骤如下：

1）根据战略性新兴产业分类与国际专利分类参照关系表，用如下检索式检索：(IPC=(A61K31* OR A61K38* OR A61K39* OR A61K47* OR A61K48* OR A61K33* OR C07J* OR A61K9* OR C07K* OR A61P* OR (C07C* NOT (C07C1* OR C07C2/00 OR C07C2/30 OR C07C4/02 OR C07C4/12 OR C07C4/22 OR C07C5/333 OR C07C6/04 OR C07C7/13 OR C07C7/177 OR C07C9/10 OR C07C9/21 OR C07C9/22 OR C07C11* OR C07C13/12 OR C07C13/20 OR C07C13/50 OR C07C13/68 OR C07C15* OR C07C21/14 OR C07C27* OR C07C29* OR C07C31* OR C07C35/28 OR C07C35/36 OR C07C37/18 OR C07C37/84 OR C07C39/23 OR C07C41/28 OR C07C41/40 OR C07C41/44 OR C07C43* OR C07C45/49 OR C07C47/02 OR C07C49/00 OR C07C49/205 OR C07C49/258 OR C07C49/573 OR C07C49/713 OR C07C51* OR C07C55/12 OR C07C59/00 OR C07C59/11 OR C07C61/13 OR C07C63/24 OR C07C63/38 OR C07C67* OR C07C69* OR C07C71/00 OR C07C203/00 OR C07C205/05 OR C07C209/22 OR C07C209/44 OR C07C211* OR C07C215* OR C07C217/14 OR C07C217/30 OR C07C217/76 OR C07C219/08 OR C07C219/10 OR C07C229/68 OR C07C231* OR C07C233* OR C07C235* OR C07C237/32 OR C07C245/14 OR C07C251/20 OR C07C251/22 OR C07C253* OR C07C255/20 OR C07C255/55 OR C07C269/02 OR C07C271/02 OR C07C271/68 OR C07C275/06 OR C07C275/10 OR C07C309* OR C07C311/06 OR C07C311/49 OR C07C313/28 OR C07C319* OR C07C323/41 OR C07C333/20 OR C07C403/16 OR C07C409/08 OR C07C409/12)) OR (C07D* NOT (C07D201* OR C07D207/335 OR C07D209/76 OR C07D211* OR C07D213* OR C07D215* OR C07D223* OR C07D235* OR C07D239* OR C07D243/04 OR C07D249* OR C07D251/38 OR C07D255/04 OR C07D277/84 OR C07D279/32 OR C07D293/12 OR C07D295/037 OR C07D295/10 OR C07D301* OR C07D307* OR C07D311/26 OR C07D311/68 OR C07D313* OR C07D317* OR C07D319* OR C07D329* OR C07D333/10 OR C07D333/78 OR C07D341/00 OR C07D401/00 OR C07D405* OR C07D413/02 OR C07D421/14 OR C07D487* OR C07D495/08)) OR C12Q1/68 OR C12Q1/70)) AND (RI-TYPE=(无效口审 OR 无效审查决定))，专利类型选择发明授权，检索日期是 2023 年 3 月，一共检索到 829 件专利；

2）对检索到的 829 件专利进行数据预处理，处理方式包括：去除没有无效决定日、没有复审决定、没有复审无效法律依据等的数据，经过预处理后包括得到 642 件专利；

3）对预处理后的专利根据国家知识产权局的审查决定进行分组，具体地，将 642 件专利依据审查决定分为“维持有效”组和“宣告无效”组，其中“维持有效”组专利有 226 件，“宣告无效”组专利有 416 件。

2.2 指标选取

本文从专利的申请、确权和无效审查 3 个阶段中综合考量对专利质量产生影响的因素。在专利申请阶段，选择专利申请文件中的发明人数、申请文件页数 2 个指标；在专利确权阶段，选择授权

文本权利要求数、技术覆盖范围、引证专利数和引证论文数共 4 个指标；在专利无效审查阶段，选择审查决定的法律依据。后续将对以上所选择指标的原因进行说明。

2.2.1 发明人数

专利法所称发明人或者设计人,是指对发明创造的实质性特点作出创造性贡献的人。发明人数量的多少原则上反映参与发明创造投入的人力资源成本，那么耗费的人力资源成本越高技术复杂度和技术价值应该越大，邓洁等人对该结论做了实证研究。然而，在专利申请的实操时，由于不同的企业、高校、科研机构对发明人的管理方式有所不同，可能有些专利的发明人数量和专利的技术复杂度或技术价值之间没有直接相关性。但是，通常涉及无效程序的专利的技术价值更高，该专利技术受到了竞争者的关注或为竞争者难以规避，从这个视角出发，涉及无效程序的专利的发明人数量是否和专利质量有相关性值得进一步基于涉及无效程序的专利数据研究。

2.2.2 申请文件页数

研究表明，发明授权专利的技术复杂性越高，其整体篇幅会更长，附图也会较多，如，公开文本中说明书的总页数（包括附图）影响专利权的稳定性，进而影响专利质量。然而，在本文的“维持有效”组和“宣告无效”组专利数据的基础上，“维持有效”组专利稳定性经过了无效审查的考验，“宣告无效”组专利的稳定性没有通过无效审查，两组数据对比之下申请文件页数是否对专利的稳定性产生影响，需要进一步进行实证研究。

2.2.3 授权文本权利要求数

在发明专利实质审查的过程中，审查员通常至少发出一次审查意见通知书，在通知书中指出专利不符合授权条件的问题，如不符合新颖性，创造性等问题。为了克服审查员指出的问题，在审查意见答复时，通常对独立权利要求的范围进行缩限以争取授权，常用的修改方式包括将从属权利要求合并到独立权利要求形成范围更小的新的独立权利要求。由此，授权文本权利要求数量受到审查过程的影响，进而影响专利质量，在本文中，基于涉及无效程序的专利进一步对授权文本权利要求数量做实证研究。

2.2.4 技术覆盖范围

每个已公开专利都被赋予至少一个专利分类号，通常情况下，一件专利的分类号越多，专利技术覆盖范围越大，专利技术广度也更大。ERNST 等人将技术覆盖作为评价专利质量的指标。但 Harhoff 等人认为技术覆盖范围大，说明该专利技术分布不聚焦，没有较高的质量。面对当前研究的争议，聚焦在技术价值更高的专利基础之上，进一步实证研究技术覆盖范围对专利质量的影响。

2.2.5 引证专利数

Burke 等人认为引证专利数量越多，该专利技术对现有专利技术的依赖性越大，创新性越小，价值或质量也越小。在专利审查的过程中，审查员找到对比文件数量越多，则影响专利新颖性或创造性的现有技术越多，专利的授权可能性或权利稳定性更差，即使授权或权利稳定所获得的专利保护范围也比较小。所以，本文基于涉及无效程序的专利，进一步实证研究引证专利数对专利质量的影响。

2.2.6 引证论文数

专利引证网络是将文献计量引证网络的对象从科技论文到专利文献的直接延伸，专利与其引证论文或引证专利密切相关，反映了相同或相近的技术创新，这种相关性是由技术本身的传承和发展而形成的。据此，利用各种数学以及统计学的方法对施引的引证论文数分析可以反映专利的技术创新程度，进而实证引证论文对专利质量的影响。

2.2.7 援引法条

《专利法实施细则》第 65 条规定，无效宣告请求的理由，是指被授予专利的发明创造不符合专利法第 2 条、第 20 条第 1 款、第 22 条、第 23 条、第 26 条第 3 款、第 4 款、第 27 条第 2 款、第 33 条或者本细则第 20 条第 2 款、第 43 条第 1 款的规定，或者属于专利法第 5 条、第 25 条的规定，或者依照专利法第 9 条规定不能取得专利权。在无效审查决定书中，也存在一些专利不符合上述法条中的多个，援引的法条作为经过无效审查程序后，专利权维持有效和宣告无效的法律依据。分析无效审查援引法条就是分析当前涉无效审查的专利有争议的问题出现在哪些方面，归纳分析涉无效审查专利的争议点，启发之后的专利申请，在专利撰写阶段就有针对性的避免这些问题，对提升专利权的稳定性，提高专利质量具有重要意义。

表 1 指标选取及其统计方法

过程	指标	指标统计方法

专利申请	发明人数 申请文件页数	授权文件中发明人数 专利申请文件页数
专利确权	授权文本权利要求数 技术覆盖范围 引证专利数 引证论文数	主从权利要求数 专利中不同的 IPC 小组分类号数 专利引证数 文献引证数
无效审查	援引法条	无效审查决定中的法条

2.3 方法

专利计量学方法可用于对专利数据进行定性和定量分析。方法被广泛应用于评价技术质量、识别领域竞争对手、管理科技研发、识别潜在的技术机会和科研人员人力资源管理。戈一冰等人用专利计量学构建了规模、技术和法律 3 个维度 15 个指标的医学院专利评价指标体系。本研究采用专利计量学方法分析经无效审查程序后，“维持有效”组和“宣告无效”组专利之间的发明人数、申请文件页数、授权文本权利要求数、技术覆盖范围、引证专利数、引证论文数和援引法条指标差异。

独立样本 T 检验是用 T 分布来检验差异性发生的概率大小的统计学方法，常用来比较两组数据的差异性是否显著。王黎明等人使用独立样本 T 检验，检验了论文作者数与论文被引频次的关系。因为本研究使用的专利数据分为“维持有效”组和“宣告无效”组，所以以二分类 Logistic 回归模型进行研究，使用 SPSS26.0 软件进行描述性统计和独立样本 T 检验，对提出的除援引法条之外的其他指标进行分析，及两组所援引法条数量进行对比统计。在统计和分析后，识别“维持有效”组和“宣告无效”组专利的具有显著性差异的因素，挖掘这些因素对专利质量的影响。

3 无效审查专利概述

3.1 专利无效审查的趋势

从图 1 的整体趋势看，自从 2003 年起，中国生物医药领域被无效审查的专利数量波动上升，并且 2017 年相较于 2016 年增幅较大，增长 81%。

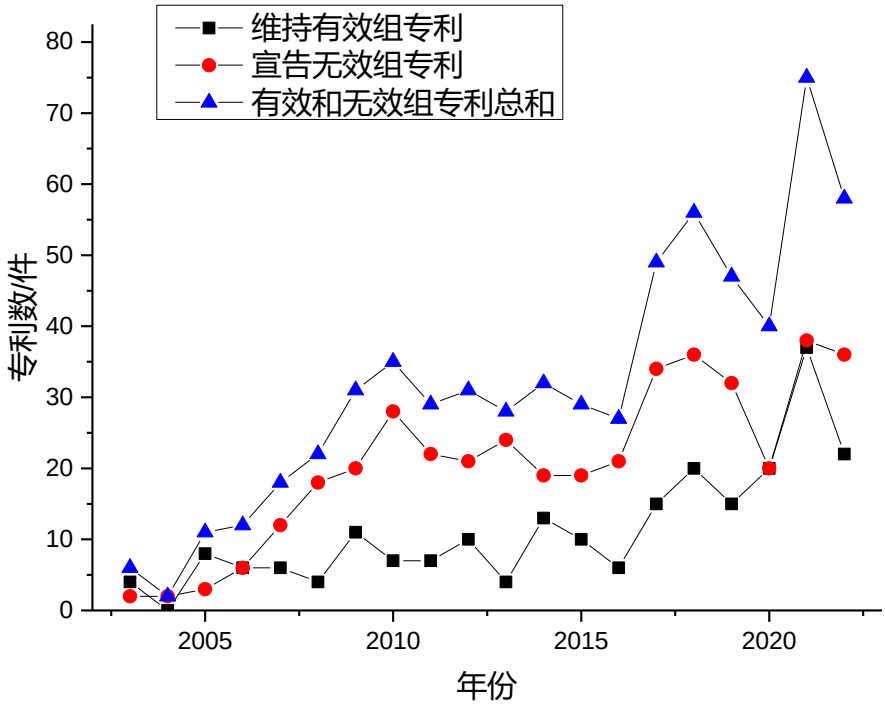


图 1 专利无效量趋势（2003 年至 2022 年）

3.2.1 发明人数

表 2 发明人数描述性统计

	无效决	个	平	标准偏	标准误差平均

定	案数	均值	差	值
维持有效	226	3.95	3.476	0.231
发明人数 宣告无效	416	3.81	2.678	0.131

根据表 2 发明人数描述性统计发现“维持有效”组发明人数量比“宣告无效”组发明人数量多。假设发明人数与专利质量呈显著正相关。

表 3 发明人数独立样本检验

		F	显著性	t	自由度	Sig.（双尾）
发明人数	假定等方差	15.175	0.000	0.573	640	0.567
	不假定等方差			0.531	372.497	0.596

发明人数独立样本检验 P 值小于 0.05，方差不齐性。t=0.531，P 值大于 0.05，不具有统计学意义。拒绝原假设，发明人数与专利质量不呈显著相关。

3.2.2 申请文件页数

表 4 申请文件页数描述性统计

无效决定	个案数	平均值	标准偏差	标准误差平均值
申请文件页数 维持有效	226	42.62	77.602	5.162
宣告无效	416	29.13	41.033	2.012

根据表 4 申请文件页数描述性统计发现“维持有效”组申请文件页数比“宣告无效”组申请文件页数多。假设申请文件页数与专利质量呈显著正相关。

表 5 申请文件页数独立样本检验

		F	显著性	t	自由度	Sig.（双尾）
申请文件页数	假定等方差	23.593	0.000	2.883	640	0.004
	不假定等方差			2.436	294.856	0.015

申请文件页数独立样本检验 P 值小于 0.05，方差不齐性。t=2.436，P 值小于 0.05，具有统计学意义。接受原假设，申请文件页数与专利质量呈显著正相关。

3.2.3 权利要求数

表 6 权利要求数描述性统计

无效决定	个案数	平均值	标准偏差	标准误差平均值
权利要求数 维持有效	226	11.31	10.437	0.694
宣告无效	416	13.50	13.508	0.662

根据表 6 权利要求数描述性统计发现“维持有效”组权利要求数比“宣告无效”权利要求数少。假设权利要求数与专利质量呈显著负相关。

表 7 权利要求数独立样本检验

	F	显著性	t	自由度	Sig.（双尾）

权利要求数	假定等方差	8.317	0.004	-2.116	640	0.035
	不假定等方差			-2.280	566.479	0.023

权利要求数独立样本检验的 P 值小于 0.05，方差不齐性。t=-2.280，P 值小于 0.05，具有统计学意义。接受原假设，权利要求数与专利质量呈显著负相关。

3.2.4 技术覆盖范围

表 8 技术覆盖范围描述性统计

	无效决定	个案数	平均值	标准偏差	标准误差平均值
技术覆盖范围	维持有效	226	8.03	8.343	0.555
	宣告无效	416	7.85	7.199	0.353

根据表 8 技术覆盖范围描述性统计发现“维持有效”组技术覆盖范围比“宣告无效”组技术覆盖范围广。假设技术覆盖范围与专利质量呈显著正相关。

表 9 技术覆盖范围独立样本检验

		F	显著性	t	自由度	Sig.（双尾）
技术覆盖范围	假定等方差	0.557	0.456	0.282	640	0.778
	不假定等方差			0.270	407.661	0.787

技术覆盖范围独立样本检验 P 值大于 0.05，方差齐性。t=0.282，P 值大于 0.05，不具有统计学意义。拒绝原假设，技术覆盖范围与专利质量不呈显著相关。

3.2.5 引证专利数

表 10 引证专利数描述性统计

	无效决定	个案数	平均值	标准偏差	标准误差平均值
引证专利数	维持有效	226	10.65	26.599	1.769
	宣告无效	416	10.86	27.064	1.327

根据表 10 引证专利数描述性统计发现“维持有效”组引证专利数比“宣告无效”组引证专利数少。假设引证专利数与专利质量呈显著负相关。

表 11 引证专利数独立样本检验

		F	显著性	t	自由度	Sig.（双尾）
引证专利数	假定等方差	0.222	0.638	-0.90	640	0.928
	不假定等方差			-0.91	468.845	0.928

引证专利数独立样本检验 P 值大于 0.05，方差齐性。t=-0.90，P 值大于 0.05，不具有统计学意义。拒绝原假设，引证专利数与专利质量不呈显著相关。

3.2.6 引证论文数

表 12 引证论文数描述性统计

无效决定	个案数	平均值	标准偏差	标准误差平均值

引证论文数	维持有效	2268	0.6	1.400	0.093
	宣告无效	416	0.4	0.948	0.046

根据表 12 引证论文数描述性统计发现“维持有效”组引证论文数比“宣告无效”组引证论文数量多。假设引证论文数与专利质量呈显著正相关。

表 13 引证论文数独立样本检验

		F	显著性	t	自由度	Sig.（双尾）
引证论文数	假定等方差	21.018	0.006	2.823	640	0.005
	不假定等方差			2.528	339.694	0.012

引证论文数独立样本检验 P 值小于 0.05，方差不齐性。t=2.528，P 值小于 0.05，具有统计学意义。接受原假设，与专利质量呈显著正相关。

3.2.7 无效审查援引法条

无效审查决定中援引法条频次如表 14 所示。

表 14 无效法律依据（前 5 条）

序号	无效法律依据	“维持有效”组每项法条专利数及所占比例	“宣告无效”组每项法条专利数及所占比例
1	第 22 条第 3 款	198(87.61%)	358(86.06%)
2	第 26 条第 4 款	128(56.64%)	118(28.37%)
3	第 26 条第 3 款	113(50.00%)	111(26.68%)
4	第 22 条第 2 款	69(30.53%)	118(28.37%)
5	细则第 20 条第 1 款	48(21.24%)	49(11.78%)

在维持有效的 226 件生物医药领域专利中，有超过 56.64%的无效审查请求援引了专利法第 26 条第 4 款。该法条体现在权利说明书撰写方面，权利要求书应当以说明书为依据,清楚、简要地限定要求专利保护的范围。有超过 30.53%的无效审查请求援引了专利法第 22 条第 2 款。该法条体现在新颖性方面，新颖性是指该发明或实用新型不属于现有技术，也没有任何单位或者个人就同样的发明或实用新型。有超过 21.24%的无效审查请求援引了专利法实施细则第 20 条第 1 款。该法条体现在保护范围方面，权利要求书应当说明发明或者实用新型的技术特征,清楚、简要地表述请求保护的范围。对于维持有效的专利争议主要围绕专利权利说明书撰写，新颖性，保护范围。

“宣告无效”组的专利与“维持有效”组的专利相比，援引法条的比例具有明显的差异性。“维持有效”组的专利在无效审查请求援引专利法第 26 条第 4 款占比 56.64%，而“宣告无效”组专利占比 28.37%。说明“宣告无效组”在专利权利说明书、新颖性和保护范围 3 个方面存在一定的问题。

4 结论及建议

4.1 结论

本研究通过对中国经历无效宣告的生物医药领域专利进行专利计量分析，识别影响专利权无效宣告结果的因素。研究结果表明，专利申请阶段的发明人数、申请文件页数，授权阶段的发明授权专利的权利要求数、技术覆盖范围、实用新型专利的引证专利数具有显著性差异。维持有效的专利在申请文件页数更多，权利要求数更少，引证论文数更多。申请阶段的发明人数，授权阶段的技术覆盖范围和引证专利数不具有显著性差异。

在无效审查过程中援引的法条显示，在专利权利说明书撰写，新颖性和保护范围 3 个方面，维持有效的专利比宣告无效的专利更好。宣告无效的专利在专利权利说明书、新颖性和保护范围 3 个方面存在一定的问题。

4.2 建议

生物医药领域专利的质量不仅和发明人的技术新颖性有关，一些共性的指标也起到了基础性作用。

第一，申请文件页数、权利要求数，引证论文数对提高专利质量具有重大意义。在同等条件下，适当增多申请文件页数，降低权利要求数，和增加引证论文数会提高专利的质量。

第二，在专利权利说明书撰写方面，更清晰地撰写权利要求，将专利的技术特征表达清楚，记载解决技术问题的必要技术特征。在保护范围方面，对专利保护的范围进行清楚的表达，不盲目扩大保护范围，必要时进行缩小，确定一个合适的保护范围，且该范围不和现有技术交叉，更有利于提高专利的质量。

第三，构建专利质量评价体系时，不应只关注专利技术本身，还应关注专利申请文件的撰写方面。本研究显示，即使专利已授权，并且可能具有较高的市场价值，但是稳定性较低，严重影响到它们的市场价值，导致其市场价值的降低。在专利评价体系纳入撰写指标，有利于科研人员加大对撰写权利说明书等的关注度，避免因表达不清楚、没说明必要的技术特征、盲目扩大保护范围等造成的专利质量降低。

第四，政府相关部门应重视无效审查程序的使用率。目前，中国的无效审查程序使用率不足1%，欧洲的无效审查使用率达 5%。促进无效审查程序的使用率，对培育高质量专利有一定的促进作用。同时，高市场价值的专利增加，也会促进无效审查程序使用的增加，达到相互促进的正反馈。政府相关部门可以通过降低相关费用或给予一定的补贴，具体来说如降低无效审查请求费或者由专利权人承担全部无效成功掉所产生的各种费用，来鼓励个人或企业提出对专利权存在问题专利的无效审查请求。

PU-118

河北省医学情报事业发展现状及对医学科技成果评价的构想

张秀菊，刘玉华，李向黎，翟俊霞
河北省医学情报研究所

引言

国内医药卫生科技查新咨询服务工作源于 20 世纪 80 年代，经过近 40 年的发展，在医学科技创新中起到了提供信息支撑的作用[1]。随着 2021 年国务院办公厅发布《关于完善科技成果评价机制的指导意见》[2]，科技成果评价的发展提上日程，各地积极发布相关的实施方案，探索科技成果评价，但对于医学类的科技成果还没有形成标准的科技成果评价，因此本文首先阐述河北省医学情报事业发展现状，在此基础之上对医学类科技成果评价方法提出构想，为医学科技成果评价的发展提供参考。

1 河北省医学情报事业发展现状

1.1 数据库资源

河北省医学情报研究所承担着全省的医学科技情报查新检索工作，目前本所订购的数据库资源包括中国生物医学文献服务系统（Sinomed）、中国知网中国医院数字图书馆数据库、万方医学数据库、维普网数据库、维普外文全文数据库等数据资源，数据库资源种类多、涵盖内容多、服务方式多样化，能够为全省医学科研工作者提供全面、系统及多层次的医学信息咨询服务。

1.2 查新检索数量

科技成果的水平和数量是衡量一个地区、机构和个人科研能力的重要标志[3]。以近三年本所登记的课题项目数据，2020 年、2021、2022 年每年查新量分别达到 2774 项、2790 项和 3162 项，每年查新总量稳步增加。2020 年、2021、2022 年中医、中药学查新课题数量三年分别为 386 项、411 项和 423 项，所占比例分别为 13.91%、14.73%和 13.38%，中医类课题项目较往年增多，说明中医药医务工作者科研意识大幅度增强。

1.3 查新申报单位分布

从申报单位来看，省级单位 1874 项，占总数的 59.27%，其中立项 1448 项，占 77.27%，鉴定验收 415 项，占 22.15%；市级单位 939 项，占总数的 29.70%，其中立项 625 项，占 66.56%，鉴定验收 311 项，占 33.12%；县级单位 94 项，占总数的 2.97%，其中立项 70 项，占 74.47%，鉴定验收 23 项，占 24.47%。

2 科技成果评价产生的政策背景

当前，我国坚定实施创新驱动发展战略，科技评估在推动创新环境建设、促进科技体制改革、支撑科技管理决策等方面发挥着越来越重要的作用。2016 年，科技部根据《国务院办公厅关于做好行政法规部门规章和文件清理工作有关事项的通知》对《科学技术成果鉴定办法》等规章予以废止。科技成果评价作为科技成果转移转化的重要环节，过去由政府科技主管部门对科技成果进行鉴定的做法，已经不能适应当前的形势发展需要，而通过第三方专业评价机构对科技成果的科学价值、技术价值、经济价值、社会价值进行客观、公正的评价，更有利于获得投资方和合作方的认可，更有利于技术交易的顺利进行，也更有利于获得政府支持。国务院印发的《“十三五”国家科技创新规划的通知》中也把第三方的评价结果作为财政科技经费支持的重要依据。为探索和建立以市场为导向的新型科技成果评价机制，科技部在 2009 年就启动了科技成果评价试点工作。

2021 年 7 月 16 日，国务院办公厅发布《关于完善科技成果评价机制的指导意见》[2]，指出：坚持科技创新质量、绩效、贡献为核心的评价导向。充分发挥科技成果评价的“指挥棒”作用，全面准确评价科技成果的科学、技术、经济、社会、文化价值。坚决破解科技成果评价中的“唯论文、唯职称、唯学历、唯奖项”问题。

3 医学科技成果评价的作用及必要性

医学科技成果评价更加注重科研项目本身所创造的科学价值，能够真正体现科研项目的创新程度，从而发挥第三方评价机构在科技成果评价中的作用。医学科技成果评价能够破解“唯论文、唯职称、唯学历、唯奖项”问题。医学科技成果评价将对科技成果的科学性、创新性、先进性等多维度进行评价，充分发挥科技成果评价“指挥棒”作用和价值发现作用，以科技创新质量、绩效、贡献为评价核心，全面准确反映科技成果创新水平、转化应用绩效和对经济社会发展的实际贡献，激发创新主体和科技人员积极性。破除“唯论文”和“SCI 至上”为突破口，不把论文数量、代表作数量、影响因子作为唯一的量化考核评价指标。

4 科技成果评价全国发展情况

随着国家对科技成果评价的重视及推进，全国各省积极响应。2021 年 11 月，湖北省出台《关于完善科技成果评价机制的实施意见》，介绍科技成果评价改革政策亮点和具体措施，配套出台《湖北省科技成果评价工作指引（试行）》，增强成果评价的操作性，提出同行评议、市场化评价、第三方评价等多种评价方法，满足不同成果评价需求。四川作为全国 4 个综合试点省之一，开展有针对性的改革试点，推动科技成果评价观念转变、方式方法创新和体制机制改革，探索形成可操作可复制的科技成果评价“四川模式”，建立完善四川科技成果评价体系。山东省人民政府办公厅出台《关于完善科技成果评价机制的实施意见》。2022 年 2 月 28 日，山西省出台《关于完善科技成果评价机制的实施意见》，通过评价激发科技人员积极性，推动产出高质量成果，有力促进创新链、产业链、价值链深度融合，为全方位推动高质量发展提供有力支撑。2021 年 12 月，河北省人民政府办公厅出台《关于完善科技成果评价机制的实施意见》，坚持科学多维的评价体系。健全多维度、分类的科技成果评价体系，科学确定基础研究和应用基础研究成果、应用技术成果、可转化和产业化成果等的评价标准，开展多层次差别化评价。2022 年 1 月，河北省科学技术厅印发《河北省科技成果评价工作指引（试行）》（冀科成市函〔2022〕3 号），同时启动开展评价改革试点工作。

5 河北省医学科技成果评价的构想

5.1 医学科技成果评价分类

医学科学技术的研究依据的是人体生理或疾病的某些本质和规律，科研课题的长期性和复杂性决定了评价的阶段性和综合性，科研的评价不是“一锤子买卖”，要遵循研究规律，把握渐进性和成果阶段性的特点。针对不同阶段的医学课题分为课题立项前评估、项目中期评估、项目完成评估和转化评估。将医学类科研成果分为四类，进行差异化评价：①基础研究：该医学领域的成果以学术价值评价为主，并进行潜在经济效益、社会效益及推广应用前景的评价；②临床研究：该类医学类研究成果以实用价值评价为主，并进行技术水平评价。③预防医学与卫生学研究:主要体现研究的社会价值及对当前发展形势下的社会影响力。④其它研究：除以上三类医学科研类型的成果研究划分到其它类。

5.2 医学科技成果评价方法

以科研项目完成评估为例，具体评价方法涵盖课题组成员评估、成果完成方的基础条件评估、项目研究的类型进行成熟度和先进度评估、项目新颖性评估、项目查引查证评估、项目效益评估、专家评价、综合加分项八个方面。具体评估方法见图 1。

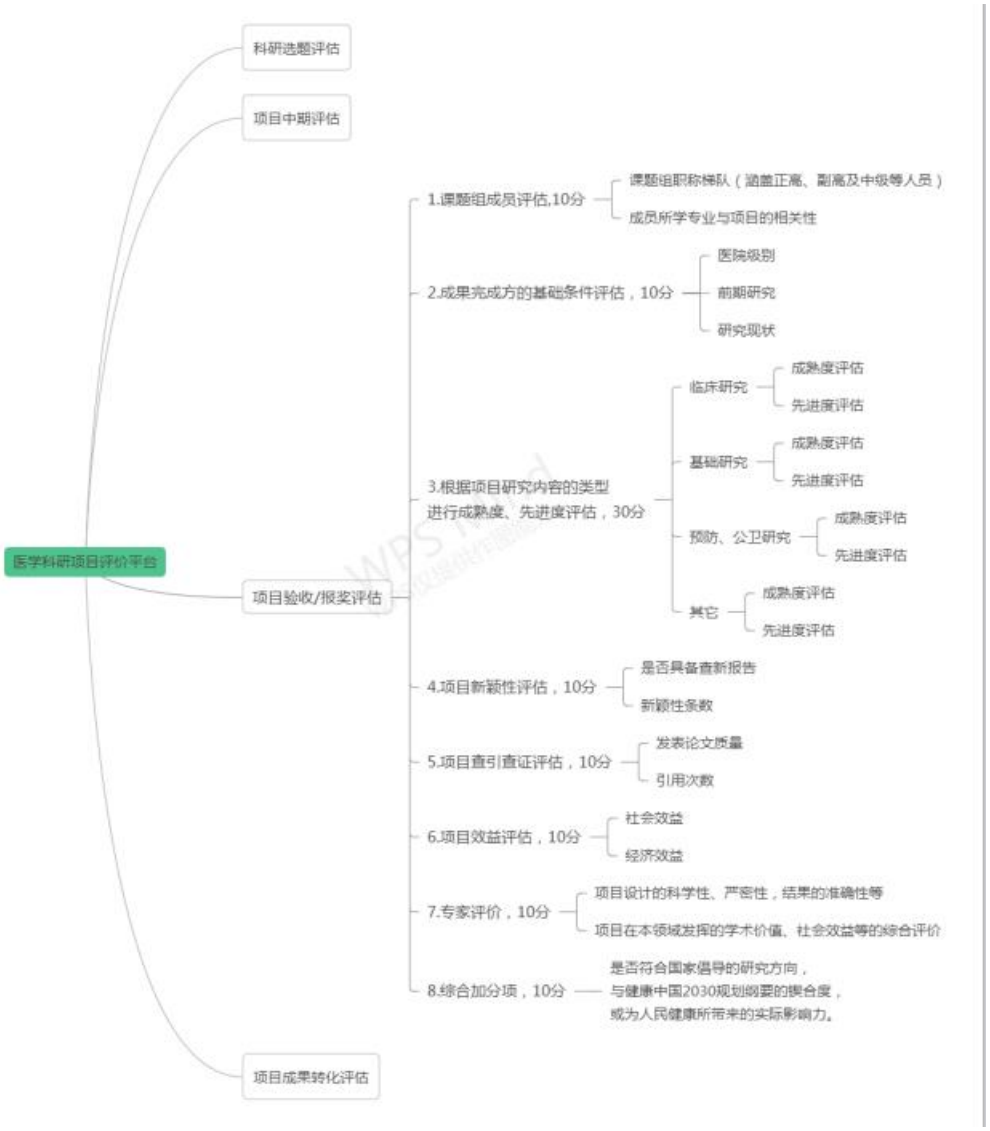


图 1 医学科技成果评价方法

5.3 医学科技成果评价流程

医学科技成果评价流程：1.客户委托。由评价委托方评价机构提出科技评价申请，提交评价申请表，确认评价目的、评价对象等基本信息。2.接受委托并签订合同。评价机构根据自身评价范围和评价能力确定是否接受委托。一旦接受委托，双发签订评价合同，约定有关评价要求、完成时间和费用等。3.评价原始材料收集。科技评估师指导委托方准备评价原始材料。4.聘请咨询专家。评价机构根据评价技术所在的专业领域聘请符合要求的咨询专家。5.科技查新。根据委托方提交的 创新点和专家对科技成果创新点的咨询意见，委托查新机构进行查新和情报文献分析。6.核心指标评价。科技评估师根据项目基本信息、查新报告、第三方检测报告、发票、应用证明、审计报告等材料，按照评价标准、规定和方法对科技成果进行标准化评价，在咨询专家的指导下，评定成熟度、创新度、先进度的等级。7.完成科技评价报告初稿。评估师根据评价结果撰写科技成果标准化评价报告，并由咨询专家进行审核确认。8.评价报告备案审核。评价报告确认无误后，评估师通过评价系统提交评价咨询专家信息、评价报告及关键证明材料。9.完成评价报告终稿。评价报告审核通过后，评价机构打印报告，并且由科技评价机构盖章，评估师和咨询专家签字，加盖河北省科技成果标准化评价专用章，评价报告生效，流程见图 2。

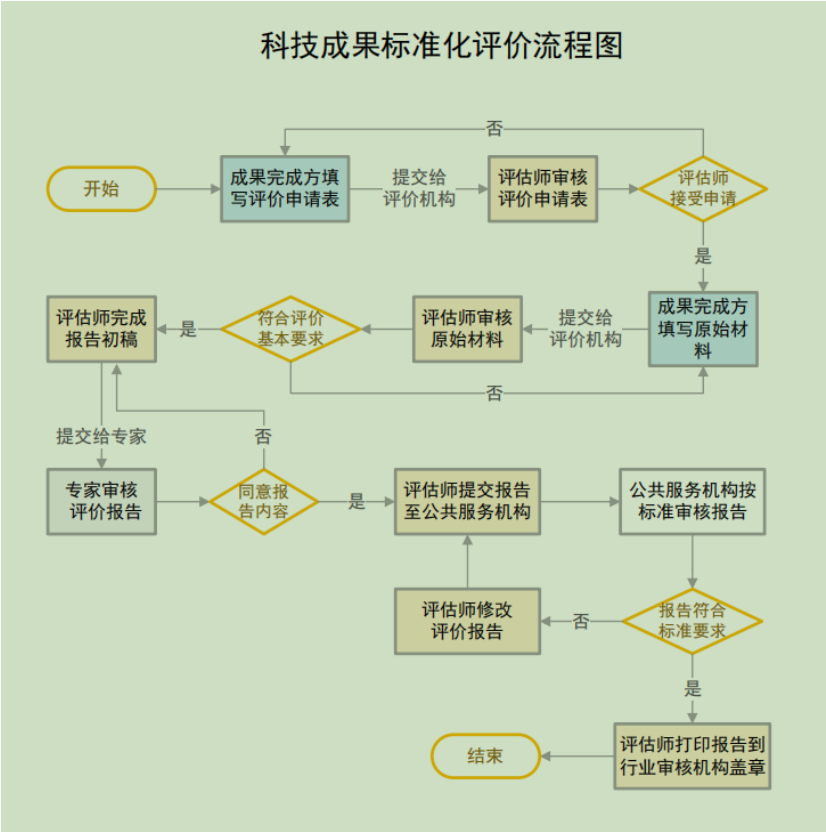


图 2 医学科技成果评价流程

6 小结

综上，医学科技情报在医学科研、医院管理中发挥着重要作用，医学情报事业随着社会的发展需要创新工作模式。河北省医学科技情报事业发展态势良好，随着科技成果评价政策的落地，如何设计医学科研课题的科技成果标准化评价是未来发展的趋势，值得进一步深入研究。

PU-119

中国护士社交媒体依赖影响因素及作用机制研究：基于湖南省的横断面调查

康春良 [1,2,3]，谢文照 [1,2,3]*

1.中南大学湘雅三医院

2.中南大学生命科学学院

3.医学信息研究湖南省普通高等学校重点实验室（中南大学）

引言

过去一年有 4.24 亿用户开始了他们的社交媒体之旅，相当于平均每天有超过 100 万社交媒体应用新用户[1]，在中国估计的 14 亿人口中，有超过 10 亿人在上网。平均而言，中国网民每天上网时间为 6 小时 25 分钟[2]。社交媒体成瘾被描述为无法控制使用互联网和社交媒体的冲动，导致一个人在不上网时过度关注社交媒体的状态。社交媒体成瘾的事实表明，由于他们将大量的精力和时间投入到社交媒体的渴望中，这种依赖性损害了他们生活的其他重要领域[3]。有学者研究表明中国医护人员有着相对较高的社交媒体成瘾倾向[4]，护士是一门高压力的职业，需要面对各种病人、医生和管理者的需求[5]，为了缓解工作压力，一些护士会利用空闲时间使用社交媒体，与朋友、家人或同行交流。然而，如果护士过度依赖社交媒体，会忽视工作和生活中的其他事务，在影响病人的就医体验，职业素养和健康状况都会受到损害。目前关于移动社交媒体成瘾的研究有了一些研究成

果，但对于护士社交成瘾的影响因素机制探究的研究文献尚少。本研究从 6 个人口统计学因素、职业倦怠、工作投入三个方面探讨护士社交媒体成瘾影响因素。

1.1SMA 的影响因素

1.1.1 人口学因素

随着智能设备的普及社交媒体使用的依赖，目前研究发现，社交媒体成瘾可能与性别，年龄，婚姻状况等其他人口因素相关。女性比男性 SMA 水平更高[6]，年纪轻的比年长的更容易形成社交媒体成瘾、单身未婚的人群比已婚人群更容易形成 SMA[7]。此次增加了技术职称、月收入、经历工作场所暴力事件的次数可能与护士群体密切相关的因素。因此我们提出假设 1。

H1：中国护士的 SMA 水平与人口因素（性别、年龄、婚姻状况、技术职称、月收入、经历工作场所暴力事件的次数）相关。

1.1.2 倦怠

职业倦怠指的是对工作的疲惫和愤世嫉俗的状态[8]， Maslach 和 Jackson 将职业倦怠定义为一种以情绪衰竭、去人格化和缺乏个人成就感为特征的综合症。情感衰竭是指因与他人接触而产生的情感枯竭的感觉是职业倦怠的核心维度。去人格化是指对接受服务或照顾的人产生的消极或过分冷漠的态度。个人成就感的降低是指一个人在工作中的自豪感和成功的成就感的下降[9]。职业倦怠综合症是对长期暴露在职业压力下的一种心理反应。情感枯竭、认知疲劳和身体疲劳是其主要症状。睡眠障碍、烦躁、不耐烦、脱离工作和对自己的工作持嘲讽态度，以及无效的感觉是其他常见但次要的特征[10]。在过去十年中，倦怠综合症逐渐被认为是影响医疗保健专业人员工作效率的重大问题[11]。任何职业都可能发生倦怠，但在直接与人打交道的专业人员中最常见，尤其是在医疗保健人群中[10]。Bereznowsk 等[12]研究发现社交媒体成瘾的程度增加导致疲惫的问题而导致产生倦怠，罗等[4]研究发现社交媒体成瘾与职业倦怠存在正向相关，考虑到职业倦怠对社交媒体成瘾的影响，我们提出假设 2。

H2：中国护理人员的职业倦怠水平与 SMA 水平呈现正向相关关系。

1.1.3 工作投入

工作投入被定义为一种积极的动力状态[8]，工作投入是与工作相关的积极情感和动机持续激活的心理状态，包括活力、奉献和专注三个方面。活力指工作时表现出高水平的能量和心理韧性，愿意为工作投入努力并克服困难。奉献指个体高度融入工作，在工作中能体验到意义、鼓舞和自豪感。专注指个体全神贯注于自己的工作，感觉时间过得很快且不容易从工作中分离。社交媒体的使用在工作场所极为普遍，曹雄飞等[13]学者指出过度使用社交媒体进行社交和娱乐会产生社交媒体使用和工作需求之间的冲突，会对工作绩效产生不利影响。 Hosgor[7] Haydar [14]等研究发现 SMA 水平和工作投入变量之间存在显著的负相关，因此，考虑到社交媒体成瘾对工作投入的影响我们提出假设 3。

H3：中国护理人员的工作投入水平与 SMA 水平呈现负向相关关系。

2 方法

2.1 研究参与者

本研究对参与者纳入标准包括：1.医院的在职注册护士（至少一年）；2.自愿参与调查；3.能够理解所有的问卷问题。排除标准包括：1.既往有精神病史和意识障碍史；2.护理学生。

2.2 措施

2.2.1 社交网络服务成瘾量表（SMA）

采用 Wang[15]开发的移动社交网络服务 SNS 成瘾量表，该量表共 18 个条目，分情绪调整、显著性、容忍、撤销、冲突、复发 6 个维度，移动 SNS 成瘾量表中各项目得分可用于计算用户的成瘾水平。总分的计算方法是将各个项目的分数相加，然后除以项目的数量。采用 Likert5 级评分，从“非常不同意”到“非常同意”分别赋值 1~5 分，总分越高表示成瘾程度越高。Cronbach’s α 系数值为 0.959。

2.2.2 Maslach 的职业倦怠调查（MBI-GS）

选择经李超平修订的 Maslach 职业倦怠调查普适量表(Maslach Burnout Inventory General Survey, MBI-GS)[16]，修订后的 MBI-GS 简表共 15 个条目,分为情绪衰竭、玩世不恭、成就感低落(该维度为反向计分)三个维度，采用 Likert7 级评分，从“从不”到“每天”分别赋值 0~6 分，计分方法：职业倦怠综合得分=[0.4×情绪衰竭条目均分+0.3×玩世不恭条目均分+0.3×成就感低落条目均分][16]。得分越高，职业倦怠水平越高。职业倦怠综合得分<1.5 分为无职业倦怠，1.5～3.5 分为疑似职业倦怠，≥3.5 分为有职业倦怠[17]。Cronbach’s α 系数值为 0.902。

2.2.3 工作投入量表（UWES-9）

UWES-9 由 Schaufeli 等 [18]开发，该量表共 9 个条目，分为活力、奉献和专注 3 个维度。采用 Likert 7 级计分法，从“从不”到“每天”分别赋值 0~6 分，得分越高，工作投入水平越高。作投入程度越高的员工，所得分值越高。各维度均分<2 分提示员工工作“低度投入”，介于 2—4 分之间提示员工工作“中度投入”，各维度均分>4 分则提示员工工作的“高度投入”[19]。Cronbach’s α 系数值为 0.976。

2.2.4 人口学变量

通过文献分析以及专家咨询，确定了 6 个可能对数据产生影响的人口统计学变量：性别、年龄、婚姻状况、技术职称、月收入、经历工作场所暴力事件的次数。

2.3 数据收集

问卷于 2022 年 12 月 4 日通过“问卷星”平台，对湖南省一二三级综合及专科医院护士发放问卷进行调查。此次调查共回收问卷 1194 分，经过筛选，最终回收到有效问卷 1141 份。

2.4 质量控制

调查完全匿名，对无效问卷筛查标准如下：1.答题时间小于 300 秒（低于正常答题时间）；2.所有题项回答一致或大量回答重复的问卷。所有程序均经中南大学湘雅三医院机构审查委员会批准。

2.5 统计学分析

对于数据的分析，采用 SPSS27.0 进行数据分析。计数资料采用例数描述，计量资料采用均数±标准差描述。使用独立样本 t 检验来确定性别是否与社交媒体成瘾的分数存在差异，使用方差分析来确定年龄、婚姻状况等是否与社交媒体成瘾的分数存在差异，以 P＜0.05 为有统计学意义。其次，采用 Pearson 相关性分析 SMA 与职业倦怠、工作投入之间的相关性。

3 结果

3.1 参与者的社会人口统计学特征和社交媒体使用习惯

表 1 列出了所有社会人口统计学特征以及 SMA 评分的相应均值和标准差。在此次调查的 1141 名护士中，女性 1112 人（97.5%），男性 29 人（2.5%）女性的数量明显多于男性符合我国护士职业特点，调查的年龄范围从 18-50 岁以上，其中 31-40 岁的护士人数最多，为 551（48.3%）。71.3%的参与者处于已婚状态，25.1%的参与者处于未婚状态。护士的技术职称主管护士人数最多为 593 人（52%），收入水平主要集中在 3000-6000 元人数为 469 人（41.1%），其他人口学统计资料如表 1 所示。

表 1 参与者的社会人口统计学特征和 SMA 得分

项目	变量	频率	构成比（%）	SMA 分数 $\bar{X}\pm S$
总计		1141	100	46.15±15.09
性别	男	29	2.5	48±13.11
	女	1112	97.5	46.1±15.14
年龄	18-30 岁	413	36.2	49.02±13.93
	31-40 岁	551	48.3	45.01±15.12
	41-50 岁	152	13.3	42.54±16.44
	50 岁及以上	25	2.2	45.68±17.11
婚姻状况	已婚	813	71.3	45.37±15.16
	未婚	286	25.1	48.98±14.23
	离异	39	3.4	40.13±15.98
	其他	3	0.3	65±11.14
技术职称	护士	188	16.5	48.68±14.41
	护师	299	26.2	47.91±14.26
	主管护师	593	52	44.62±15.18
	副主任护师	57	5	44.09±18.13
	主任护师	4	0.4	51±22.54
月收入	3000 元以下	42	3.7	49.64±13.27
	3000-6000 元	469	41.1	46.71±15.23
	6000-9000 元	423	37.1	45.28±15.17
	9000 元	207	18.1	45.93±14.97
经历工作场所暴力事件的次数	从来没有	632	55.4	44.42±15.47

	1-3 次	27	4	37.4	3	48.11±13.9
	3-5 次	7	3	3.2	1	49.32±14.6
	5 次以上	5	4	3.9	1	49.16±17.8
	很健康	87	2	25.2	2	41.55±16.0

3.2 参与者 SMA、MBI-GS 和 UWES-9 评分分布

如表 2 所示，所有研究对象 SMA 得分为（46.15±15.09），中位数为 48 分，SMA 量表 6 个维度中情绪调整维度平均得分最高（8.45±3.07），其次为显著性维度（8.00±2.88）分。目前 SMA 量表只能测试受试者 SMA 程度，没有明确的诊断标准，但本研究参与者平均的分均已超过量表总分的一半（45 分），其中 SMA 得分≥45 分人数为 658 人（57.67%），说明我国护士群体普遍存在社交媒体成瘾，得分占总分的 75%，67.5 分的有 76 人（6.66%）如表 3 所示，说明我国职业人群 SMA 总体得分比较高。参与者职业倦怠综合倦怠得分为（2.02±1.06），综合得分 1.5～3.5 分为疑似职业倦怠，本次调查中疑似职业倦怠 668 例，有职业倦怠 93 例，职业倦怠占比 66.7%，如表 4 所示。表明倦怠在中国护士群体中非常普遍。护士工作投入总分为（36.33±13.63），按照工作投入划分标准，1141 位护士低度工作投入 137 例,中度投入 428 例，高度投入 576 例，如图表 5 所示，总体来说我国护士工作投入程度整体得分较高。

表 2 SMA、MBI-GS、GSES 量表评分

变量名	平均值	标准差	中位数
社交媒体成瘾总分	46.15	15.09	48.00
倦怠综合得分	2.02	1.06	2.02
工作投入总分	36.33	13.63	37.00

表 3 护士社交媒体成瘾水平

分组	例数	百分比（%）
＜45 分	483	42.3
≥45 分	658	57.7
＞67.5 分	76	6.7

表 4 护士职业倦怠水平

分组	例数	占比（%）
无职业倦怠	380	33.3
疑似职业倦怠	668	58.5
有职业倦怠	93	8.2

表 5 护士工作投入水平

分组	例数	占比（%）
低度投入	137	12
中度投入	428	37.5
高度投入	576	50.5

3.3 人口学变量与 SMA 评分关系

如表 6 所示，对性别，进行 t 检验，发现 SMA 水平上与性别上无显著性差异，对年龄、婚姻状况、技术职称、月收入、平均每月值夜班次数进行方差分析，方差分析结果显示，SMA 总分与

年龄存在显著性差异（F=9.132，p=0.000***），SMA 总分与婚姻状况存在显著性差异（F=7.838，p=0.000***），SMA 总分与技术职称存在显著性差异（F=4.285，p=0.000***），SMA 总分与平均每月值夜班次数存在显著性差异（F=6.201，p=0.000***），SMA 总分与经历工作场所暴力事件存在显著性差异（F=6.398，p=0.000***）。结果显示性别、月收入不支持 H1，但社交媒体成瘾得分在年龄、婚姻状况、技术职称、平均每月值夜班次数方面的差异支持 H1。

表 6 人口学变量与 SMA 评分的关系

变量	SMA	情绪调整	显著性	容忍	撤销	冲突	复发
性别	-0.407 0.684	0.322 0.747	-0.732 0.464	-1.298 0.195	-0.478 0.633	-0.672 0.502	-0.08 0.937
年龄	9.132 0.000***	11.034 0.000***	14.164 0.000***	8.861 0.000***	4.708 0.003***	2.804 0.039**	1.121 0.339
婚姻	7.838 0.000***	7.101 0.000***	8.298 0.000***	7.209 0.000***	3.944 0.008***	3.745 0.011**	2.506 0.058*
技术职称	4.285 0.002***	3.834 0.004***	6.09 0.000***	3.992 0.003***	2.641 0.032**	2.343 0.053*	1.498 0.202
月收入	0.817 0.484	3.491 0.015**	1.646 0.177	0.457 0.712	0.925 0.428	0.495 0.686	1.154 0.326
经历工作暴力次数	6.398 0.000***	5.719 0.001***	4.372 0.005***	3.863 0.009***	5.543 0.001***	4.985 0.002***	5.67 0.001***

注：***、**、*分别代表 1%、5%、10%的显著性水平

3.4 SMA 评分与 MBI-GS 评分和 UWES-9 评分的关系

如图 7 所示，Pearson 相关分析显示，SMA 水平与职业倦怠水平呈正相关(r=0.423,p < 0.01)，与工作投入水平呈负相关(r=-0.395,p < 0.01)，职业倦怠得分及各维度与 SMA 各维度得分均呈正相关（p < 0.05），表明中国护理人员的职业倦怠水平与 SMA 水平呈现正向相关关系，支持 H2。工作投入得分及各维度与 SMA 各维度得分均呈负相关（p < 0.05），表明中国护理人员的工作投入水平与 SMA 水平呈现负向相关关系，支持 H3。

表 7 SMA 评分与其他条目的 Pearson 相关性分析

变量	社交媒体成瘾总分	情绪调整	显著性	容忍	撤销	冲突	复发
情绪衰竭	0.423(0.000***)	0.383(0.000***)	0.321(0.000***)	0.351(0.000***)	0.345(0.000***)	0.383(0.000***)	0.351(0.000***)
去人格化	0.43(0.000***)	0.347(0.000***)	0.305(0.000***)	0.35(0.000***)	0.338(0.000***)	0.42(0.000***)	0.384(0.000***)
缺乏个人成就	0.244(0.000***)	0.183(0.000***)	0.149(0.000***)	0.19(0.000***)	0.18(0.000***)	0.258(0.000***)	0.245(0.000***)

感							
倦怠综合得分	0.473(0.000***)	0.398(0.000***)	0.337(0.000***)	0.385(0.000***)	0.374(0.000***)	0.454(0.000***)	0.419(0.000***)
活力	-0.386(0.000***)	-0.342(0.000***)	-0.291(0.000***)	-0.322(0.000***)	-0.306(0.000***)	-0.355(0.000***)	-0.324(0.000***)
奉献	-0.391(0.000***)	-0.344(0.000***)	-0.307(0.000***)	-0.331(0.000***)	-0.297(0.000***)	-0.351(0.000***)	-0.335(0.000***)
专注	-0.37(0.000***)	-0.308(0.000***)	-0.253(0.000***)	-0.298(0.000***)	-0.287(0.000***)	-0.361(0.000***)	-0.341(0.000***)
工作投入总分	-0.395(0.000***)	-0.343(0.000***)	-0.293(0.000***)	-0.328(0.000***)	-0.307(0.000***)	-0.367(0.000***)	-0.344(0.000***)

注：***、**、*分别代表 1%、5%、10%的显著性水平

4 讨论

4.1 人口统计学特征与护理人员社交媒体成瘾之间的关系

本研究是目前中国样本量最大，的针对护士 SMA 的一项研究。研究发现，中国护士群体存在着较高社交媒体成瘾现象。与学者罗[4]等研究结果一致，SMA 量表 6 个维度中情绪调整维度平均得分最高（8.45±3.07），其次为显著性维度（8.00±2.88）分。这表明护士使用社交媒体来缓解工作上的焦虑并且对移动 SNS 的行为、认知和情感上的关注是非常明显的。如果他有机会，他就会随时随地访问移动社交媒体。其中 6 个人口统计学特征中“性别”的数据分析结果未能支持本研究的第一个假设。在性别上男护士和女护士的 SMA 分数没有明显的差异，尽管其他研究表明 SMA 可能有性别[20]有差异。由于本探究样本收集是针对部分区间及时段，存在一定的局限性，性别对社交媒体成瘾的影响还值得进一步探讨。

本文研究发现护士社交媒体成瘾与年龄、婚姻状况、技术职称、每月值夜班次数、经历工作暴力事件次数存在显著性差异。18-30 岁和 50 岁及以上的护士群体更容易产生社交媒体成瘾，因为年轻人的社交媒体成瘾得分年轻人普遍拥有 3 个及以上社交媒体账号，对网络信息敏感，乐于在网络上分享自己的生活，并期待能够网络上引起家人朋友的点赞和关注，不免会经常使用手机去浏览和增加社交媒体使用的时间，从而相比于年长群体的护士 SMA 水平更高，这一研究发现也进一步证实了前人的研究[21]，但本研究发现在 50 岁以下的护士群体社交媒体成瘾得分为 45.68，仅次于 18-30 岁的年轻护士，研究发现社交媒体平台已经成为老年用户强大的健康交流工具[22]，和社交媒体是老年人与现代社会连接的中介，使用社交媒体能够提高他们幸福感降低孤独感[23]从而使他们不愿意控制社交媒体的使用时间从而更容易产生社交媒体成瘾，人口老龄化在全球已经是一个大趋势，未来针对护士老年群体的社交媒体成瘾值得进一步研究。护士婚姻状况与社交媒体成瘾有显著性差异，未婚的要比已婚的社交媒体成瘾程度更严重，与前人[24,25]的研究结果一致。单身或离婚的人需要更多的社会化，会更多的倾向在社交媒体上表达需求，对于已婚人群下班后会花费时间精力陪伴照顾家人。

在技术职称上职位越低的护士社交媒体成瘾得分越高，我们认为一线护理中，护士需要花费大量的时间和精力来照顾患者。这样的工作环境让护士很容易出现疲惫、失眠、焦虑等问题，社交媒体成了他们缓解压力、放松身心的一种手段。在本次研究中也发现平均每月值夜班次数为 5—10 次和 10 次以上的 SMA 得分最高，研究发现上夜班次数越多的工作者身体健康会受到影响，会产生头痛[26]，失眠、心理健康、疲劳等一系列健康问题[27]，同时在一定程度上反映了 SMA 与身体健康状况相关，本研究的数据也证实了护士的 SMA 与身体健康状况存在显著性差异。本研究发现经历

工作场所暴力事件的次数的护士比没有经过的护士更容易产生社交媒体成瘾，护士由于其职业的特殊要求，包括与患者频繁接触和繁重的工作量通常面临更大的工作场所风险[28]，一般而言，与其他部门的工人相比，医疗保健专业人员（主要是护士）遭受暴力的风险增加 16 倍[29]，大量研究表明，经历过工作暴力事件的受害者的身心健康产生严重影响，干扰他们的工作能力从而可能带来长期的心理创伤和情绪压力[30]，这种情况下，社交媒体可能成为一个可以帮助缓解情绪的途径，也可能成为一个逃避现实、释放压力的渠道，那些经历过工作暴力的护士可能会觉得社交媒体提供的虚拟世界更安全、更舒适，因此会更倾向于在社交媒体上花更多时间。针对护士的暴力行为是全球普遍关注的问题，但经过文献调研发现尚未有把经历过暴力事件次数与护士社交媒体成瘾的研究，后续可以针对此展开更进一步的研究。

4.2 SMA 与职业倦怠、工作投入的关系

4.2.1 护士的 SMA 和职业倦怠

本研究发现中国护理人员的职业倦怠水平与 SMA 水平呈现正向相关关系，也就是说存在职业倦怠的护士更容易产生 SMA，与之前的一些[4,12,31]研究结果一致。在罗等[4]人的研究对象中包含医生护士，在本研究中对象只有护理人群，结果对护理人群的 SMA 解释更有针对性。本研究的数据显示，职业倦怠量表的情绪衰竭得分最高(11.7±6.95)，说明护理人员在工作中普遍感到精力不足、情绪疲惫、无法应付工作要求，心理资源消耗程度较高；其次是成就感低落得分（11.22±8.04），说明护理人员在工作中对自己的能力、贡献或价值感到不满意、不自信或不认可，自我评价程度较低[32]，反映了护理人员存在工作压力过大、工作满意度较低、心理健康受损的情况。目前只有少数研究关注到护士 SMA 可以与职业倦怠存在关联这个问题[25]，几乎所有职业都存在职业倦怠，而职业倦怠对员工自身以及他们所服务的对象，甚至对他们所在的组织机构都是有害的，未来对可对此进行更深入的研究。

4.2.2 护士的 SMA 和工作投入

经过文献分析发现目前对于 SMA 在工作场所中的研究多半是关于 SMA 对工作绩效、工作满意度的相关机制影响研究，而在对于工作投入 SMA 之间的关系研究还尚少。本研究与 Haydar[21]研究的土耳护士 SMA 与工作投入存在负向相关的研究结果有差异，在本研究中发现护士工作投入与 SMA 并无关系，一项基于美国职场人士的研究[33]也发现，工作投入与社交媒体使用频率没有显著相关。因此对于本研究来说工作投入和 SMA 可能没有直接的因果关系，而是受到其他变量的影响，后续可以采用多种方法、探索更多的中介或调节变量。

5 结论

在护士社交媒体成瘾研究方面已有部分文献把职业倦怠、工作投入进行研究，只是进行单独部分关系研究，因此本文将这些研究中分析的部分关系整合进行研究，同时我们在一个相对较大的样本中进行了调查，保证了数据代表性。在本研究中，首先我们发现中国护士普遍存在较高的社交媒体成瘾水平与职业倦怠水平，护士在医疗卫生系统中占重要地位，应得到相关关注。第二在护士群体中 18-30 岁和 50 岁以上的、单身、 职称低、值夜班次数多、经历过工作场所暴力事件的护士是 SMA 高风险人群。第三本研究进一步证实了 SMA 与职业倦怠之间的正相关关系，并且表面二者之间是相互影响的，并不是单向影响。第四本研究发现工作投入水平与护士社交媒体成瘾水平呈现负相关，护士对社交媒体成瘾直接影响到对工作的投入与工作的服务质量，在后续研究中值得进一步探究。

6 不足

尽管这项研究提供了重要的见解，但它有一些局限性。本次研究数据来源还不够全面，只选取一个地区，限制了结果对其他人群的可推广性。其次，数据是横断面的，因果推断是有限的。

PU-120

新形势下医学科技查新助推基层科研的对策

刘玉华 李向黎 张秀菊 翟俊霞
河北省医学情报研究所查新咨询室

党的十八大和十九大均提出了“健康中国”的战略目标，2016 年中共中央国务院发布《“健康中国 2030”规划纲要》,明确要全方位全周期维护和保障人民健康[1];以习近平同志为核心的党中央高度重视人民健康工作,把人民健康放在优先发展的战略位置，全民健康是健康中国的重要内容。人民健康水平的提升与医疗事业的发展密不可分，而医疗事业的发展离不开科技创新，科技创新离不开科研管理工作[2]。

基层医药卫生机构需承担对当地特色病、多发病及常见病等病种的机制、诊疗和预防保健等工作,其科研工作的重要性日益凸显，但科研力量相对较弱[3-5]，因此提高基层科研水平事关本省医学科技创新水平的大局。党和国家的卫生工作始终坚持一切为了人民健康的工作基点，不断促进医疗卫生工作重心下移和资源下沉，因此科研管理也要进行科技扶贫。

河北省医学情报研究所（以下简称本所）自 1987 年开始承担河北省医药卫生科技查新服务项目，新形势下努力作好医药卫生科技查新咨询服务,为“健康中国”的战略保驾护航[6-8]，本课题在充分调研本省基层医疗卫生机构科研的现状基础上，探讨医学科技查新助推基层科研的对策，以进一步提升本省科技创新水平，助力健康中国。

1 提升基层科研的迫切需求

1.1“健康中国”战略背景下基层医疗卫生机构的科研水平需要大幅度提升

党的十八大和十九大将“健康中国”作为战略目标纳入了国家基本方略，也促使了医学模式从“疾病模式”向“健康模式”的转变[9]。基层科研医疗卫生服务能力与科研水平息息相关。基层医疗卫生服务能力是维护人民群众健康的基础保障，基层能力明显提升，可体现人民群众看病的公平性、可及性和便利性,因此科技创新在基层医院中的作用日益凸显[10-11]。因此有必要进一步提升基层科研医疗卫生服务能力，为健康中国作出贡献。

1.2 国家对基层医药服务能力的日益重视对科研管理工作提出了挑战

卫生健康事业迈进新时代，国家对医学科学研究高度重视和大力支持。基层科研水平是评价其整体实力的重要指标之一。科研能力薄弱,一定程度上限制了基层卫生机构医疗技术的发展和为广大官兵服务的卫勤保障能力[12]。十四五”时期是实现中医药振兴发展的关键时期。2022 年 3 月，国家中医药局等 10 个部门联合印发《基层中医药服务能力提升工程“十四五”行动计划》，提出完善基层中医药服务网络、推进基层中医药人才建设等 7 个方面的重点任务，以及到 2025 年社区卫生服务中心和乡镇卫生院中医馆实现全覆盖等 5 个“全覆盖”的具体目标[13]。因此给科研管理工作提出了挑战。

1.3 对提高基层医疗卫生机构核心竞争力具有重要意义

科研已成为医学各学科积极进取，不断创新的动力，学科的建设和发展需科研的支持。科研能力是临床上解决实际工作问题、改革创新、探索真理的能力，是医学各学科发展的核心驱动力。基层医学各学科的发展需要科研工作者积极开展科研工作，科研工作的发展决定了基层医疗机构的核心竞争力。

2 河北省基层医疗卫生机构科研水平的现状

2.1 科研能力呈明显降低的趋势

科技成果的水平和数量是衡量一个地区、机构和个人科研能力的重要标志。本所 2017、2018、2019、2020、2021 年查新课题总量分别为 2472、2476、2496、2774 和 2790 项，其中基层医疗卫生机构查新课题数量分别为 420、351、289、249 和 187 项，所占比例分别为 16.99%、14.18%、11.58%、8.98%和 6.70%，呈逐年降低的趋势，提示本省基层医疗卫生机构科研能力显著降低。

2.2 申报项目呈明显升高的趋势

项目申报是科技项目实施的第一个阶段，也是关键的一步。2017、2018、2019、2020、2021 年基层医疗卫生机构立项查新课题数量分别为 232、199、165、148、124 项，所占比例分别为 55.24%、56.70%、57.09%、58.44%和 66.31%。

2.3 中医药学申报项目呈逐渐升高的趋势

2017、2018、2019、2020、2021 年基层医疗卫生机构中医药学查新课题数量分别为 92、102、105、98、43 项，所占比例分别为 21.90%、29.06%、36.33%、39.36%和 22.99%。

2.4 民营医疗卫生机构科研能力较弱

2017、2018、2019、2020、2021 年民营基层医疗卫生机构查新课题数量分别为 4、1、0、2、6 项，所占比例分别为 0.95%、0.28%、0、0.80%和 3.21%。

2.5 不同地区申请项目分布不均

石家庄地区、唐山地区、保定地区、邢台地区每年申请查新课题比例均超过 10%，秦皇岛、张家口、承德每年申请查新课题比例均未超过 5%，其它地区申请查新课题比例波动于 5%~10%，而雄安新区申请查新课题比例呈逐年升高的趋势，见表 1。

表 1 2017－2021 年不同地区基层医疗卫生机构申请查新课题数量的比较[项（%)]

年度	石家庄	唐山	秦皇岛	廊坊	保定	沧州	衡水	邯郸	张家口	承德	邢台	雄安新区
2017 年(n=420)	42(10.00)	81(19.28)	4(0.95)	23(5.49)	65(15.48)	48(11.43)	14(3.33)	50(11.90)	15(3.57)	13(3.10)	62(14.77)	3(0.71)
2018 年(n=351)	43(12.26)	57(16.24)	5(1.42)	31(8.63)	53(15.10)	30(8.55)	13(3.70)	52(14.81)	8(2.23)	7(1.99)	48(13.68)	4(1.14)
2019 年(n=289)	34(11.77)	37(12.80)	2(0.69)	28(9.69)	45(15.57)	29(10.03)	6(2.08)	43(14.88)	6(2.08)	5(1.73)	42(15.57)	9(3.11)
2020 年(n=249)	28(11.24)	46(18.48)	0	19(7.63)	33(13.25)	15(6.02)	9(18.07)	45 (18.07)	3(1. 20)	2(0.80)	41(16.47)	8(3.21)
2021 年(n=187)	19(10.16)	35(18.71)	0	10(5.35)	30(16.04)	16(8.56)	9(4.81)	11(5.88)	1(0.53)	6(3.21)	44(23.53)	6(3.21)

3 医学科技查新助推基层科研的对策

本课题立足“十四五”期间人社事业发展需要，重点围绕助力乡村医疗、破解基层卫生发展难题、推广应用新技术、培养培训人才、提高科研敏感性、培养论文撰写能力等主题，组建专家服务团开展服务活动、以达到提升民生服务目的。拟通过人力资源社会保障部部门的“牵线搭桥”，深入基层一线实地调研，摸清基层一线发展需求。

本文结果表明，本省基层医疗卫生机构科研能力呈明显降低的趋势，一方面原因可能与各级相关科研管理部分对科研重视程度不够，另一方面原因可能与基层医疗卫生机构科研人员科研意识有所降低有关。针对本省基层医疗卫生机构科研水平的现状，本课题提出以下医药卫生专家服务基层科研相关路径。

3.1 组建医学科技查新扶贫专家团队

本研究结合工作实际，通过线上发布、公开邀请、自愿报名和按需定向邀请等方式进行专家人才匹配。针对基层医疗卫生机构科研能力薄弱的现状，本省早在 2010 年就提出了科研管理应该积极为基层医务人员服务的思路[14]。组建河北省科技扶贫宣讲团队，团队成员包括科研究管理专家、统计学专家、期刊主编和查新专家，将医学科研管理知识多次送到“基层”，进行科技扶贫。

3.2 多个基层科研管理交流微信群建立

微信是目前国内移动互联网上发展较为迅猛的应用程序，是我国用户群体最大的移动即时通讯软件，微信作为互联网的新兴产业，不仅可以用作即时通讯软件，而且其快速的发布形式及便捷的沟通模式，可以为多个用户搭建一个沟通、交流的互动平台。本省已将微信用于提升医药卫生科技项目查新咨询服务的效果，得到了广大科研人员的一致好评[15]。

3.2.1 科研管理人员微信群建立

通过该群加强与各基层医学科研管理部门的联系和沟通，推送各科研管理机构发布的立项、成果验收、评审和其他有关科研申报的文件，及时了解科研管理部门对科研管理的需求，解答科研管理过程中的问题及疑问。

3.2.2 科研人员微信群建立

通过查新服务，建立科研人员微信群，基层医疗卫生机构的科研人员多数是科研的中坚力量，但是对于如何申报有创新点的项目不是非常清楚。查新机构可以将通过各种方式收集的各科研管理机构发布的有关立项或成果验收的文件发布在微信，提醒科研人员及早进行查新。通过微信群与科研人员保持互动，及时在群中发布科研管理申报课题的有关文件及查新需要准备的相关材料，让每一位科研人员提供高质量的查新咨询服务。

3.2.3 省级医药卫生专家微信群建立

省级医药卫生专家有丰富的科研经验，因此文献查阅能力、论文写作能力、科研实践能力、资料处理能力和科研设计能力均较强，通过医学专家咨询微信群，基层科研人员可以咨询科研过程中的专业问题，另外科研管理人员也可咨询作好基层科研的意见和建议，集思广益，进一步提高科研管理的服务质量。

3.3 加大对中医药学课题结题验收的指导力度

2019 年 10 月 25 日在北京召开的全国中医药大会，为新时代传承创新发展中医药事业指明了方向，也吹响了新时代中医药传承创新发展的号角，中医药振兴发展迎来大好时机[16]。

本研究结果表明，中医药学申报项目呈逐渐升高的趋势，这一成果的取得，与本省实施了中医药项目的大力扶持政策有关，河北省委、省政府印发《关于促进中医药传承创新发展的若干措施》，明确了 6 大方面 22 项具体措施，全面推进中医药传承创新发展[17-18]，这有力地促进本省整体中医药学科研水平；另外也与本省在作好中医药卫生科研管理在进行与时俱进的实践与探索有关[19-22]。

项目申报是科技项目实施的第一个阶段，结题验收完成才是一个科研项目。中医药学是中华民族的伟大创造，在维护和促进人民健康中发挥着独特作用，传承创新发展中医药是新时代中国特色社会主义事业的重要内容。因此应该加大对中医药学课题验收的指导力度，以提高验收率。实施的措施包括：定期推送结题验收的工作流程和相关材料，督导各基层医疗机构的科研处每年均需填写“科研项目进展情况表”，对项目执行情况、下一步工作安排和存在问题等进行总结, 由科教部进行审核, 对项目组存在的问题和遇到的困难进行及时地协调解决, 为顺利完成项目研究提供及时有效的帮助；项目到期前 6 个月，将项目结题验收书 ”发至项目负责人手中，及时提醒项目组办理结题验收;各项目撰写的结题验收材料需先交科教部审核把关, 审核通过后再提交上级部门，申请组织结题验收。

3.4 加大对民营医疗卫生机构科研的扶持力度

民营医疗集团作为我国医疗服务市场的重要组成部分,民营医疗集团作为企业和医院的综合体[23]其“百年”梦想一样离不开科学研究，科学研究水平不但是反映一个医院、一个学科以及一个团队实力的重点标志，也是决定其是否具有可持续发展潜力的决定性因素[23]，民营医疗集团的科研活动要与企业的生存发展挂钩，这就决定了民营医疗集团科研工作。

中共中央 国务院关于促进中医药传承创新发展的意见（以下简称《意见》）明确提出，大力扶持民营医疗机构，科技创新是民营医疗机构能立足的基石，因此民营医疗机构应该加强医药科研，以适应新时代我国医疗事业的发展[24]。本文结果表明，民营医院科研能力较弱，因此应该加大民营医疗机构科研的扶持力度，如在科研立项方面降低入围的标准，放宽成果结题验收的条件，鼓励民营医疗卫生机构对科研人员在基金和奖励方面给予照顾等；加大对民营医院科研相关知识培训[25]。

综上所述，科研管理工作者应该充分认识到基层科研在“健康中国”大局中的重要性，针对基层科研存在的问题，探讨行之有效的服务基层科研的对策，进一步提升基层科研水平，作好科技扶贫，在科研方面推动本省各地区乃于京津冀协同发展，助力健康中国。

PU-121

通过质量和疗效一致性评价药品的供应可视化分析—基于山东省医疗机构药品使用数据

孙雅群[1,2]，宋燕[1,2]
1.山东第一医科大学（山东省医学科学院）医疗保障学院
2.山东省医药卫生科技信息研究所

我国在 2012 年《国家药品安全“十二五”规划》中首次提出一致性评价的概念[1]，2016 年 3 月国务院办公厅印发《关于开展仿制药质量和疗效一致性评价的意见》[2]，标志着中国仿制药质量和疗效一致性评价工作的全面展开。仿制药质量和疗效一致性评价工作是提升我国仿制药整体质量、促进产业转型升级、在临床上实现原研替代、促进用药可及的重要举措。2018 年 3 月国务院办公厅印发《关于改革完善仿制药供应保障及使用政策的意见》中进一步强调了要加快推进仿制药质量和疗效一致性评价工作[3]。2019 年扩围至全国实行的国家集中带量采购，投标的条件之一就是通过一致性评价或视同通过一致性评价。随着一致性评价品种数的增加、带量采购政策的常态化，开展通过质量和疗效一致性评价药品的供应变化研究，以掌握过评药品在过评前后、集采前后的采购使用变化情况。研究结果可为促进优质仿制药优先采购使用等相关配套政策的制定提供参考依据。

1 数据来源

本研究数据主要来源于中国医药工业信息中心，数据内容为 2017 年过评的 12 个品种在 2017-2022 年山东省代表性医疗机构的药品采购信息，包括药品名称、规格、剂型、采购数量等。

2 分析方法

分析药品的采购量和配送到货率情况。采购量以用药频度 DDDs 表示，计算方法为 DDDs=该药品采购数量×规格/该药品的限定日剂量（DDD）。DDD 值从世界卫生组织药物统计方法协作中心查询或以药品说明书规定的成人常用日剂量为准。配送到货率=最终到货入库量/订单下单量×100%。

4 研究结果

4.1 过评药品的供应情况分析

4.1.1 各品种采购量的变化情况

过评品种五年采购量情况如图 1 所示。过评后部分品种（富马酸替诺福韦二吡呋酯片、吉非替尼片、盐酸帕罗西汀片）的采购量呈现明显上升的趋势，但整体变动趋势不明显。2019 年第 4 季度末山东省开始执行带量采购，集采后只有两种药品（厄贝沙坦片、氯沙坦钾片 50mg/100mg）的采购量有明显上升趋势。部分品种采购量出现下降或持续低迷的原因一方面可能是受疫情的影响，另一方面是过评企业过少，影响了供应量，比如福辛普利钠片、吉非替尼片和赖诺普利片的五年累计过评企业仅有 1-2 家。2022 年之后部分药品供应量发生较大变化（比如厄贝沙坦片、氯沙坦钾片 50mg），可能与国采协议期满后接续工作调整相关。

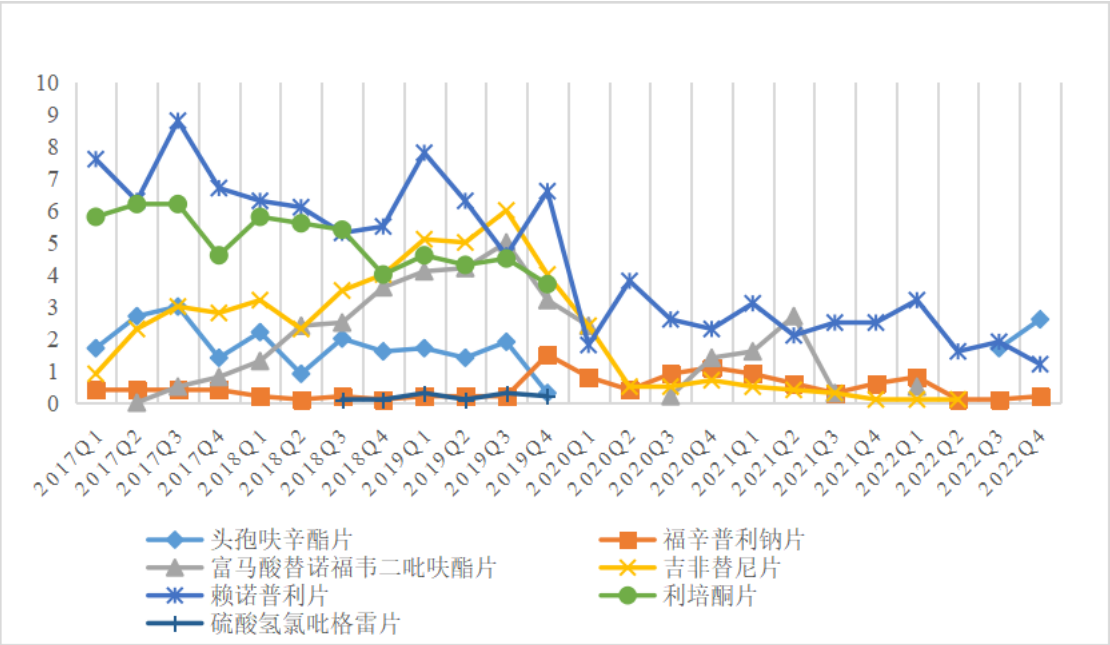
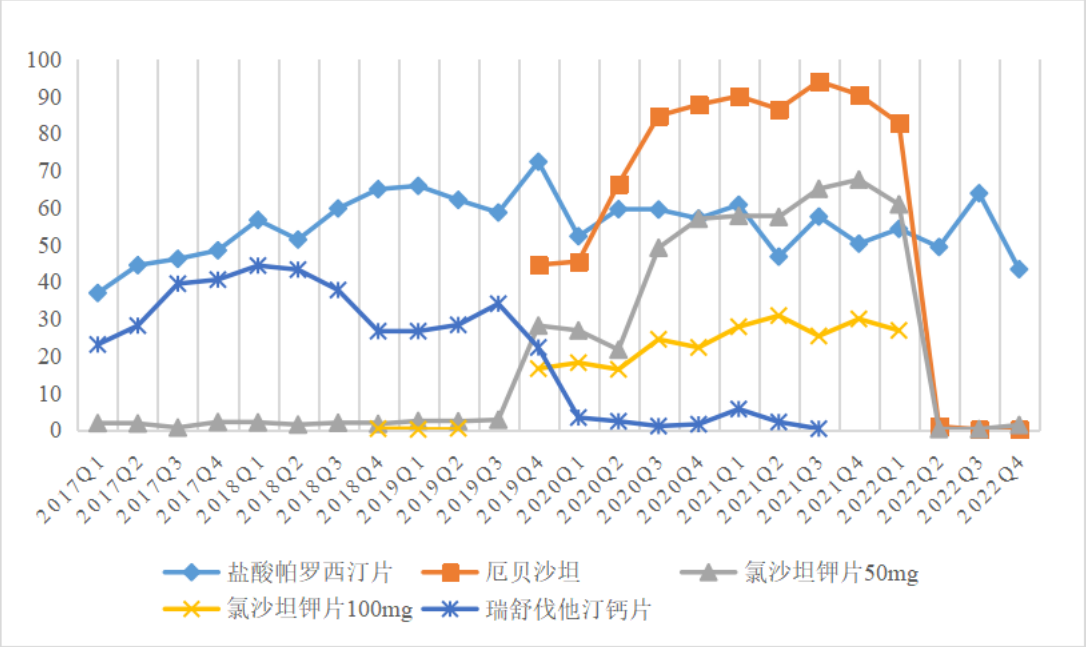


图 1(a) 2017-2022 年过评药品采购量变化



*厄贝沙坦氢氯噻嗪在山东没有 2017 年过评仿制药品种的使用数据，因此未作分析。

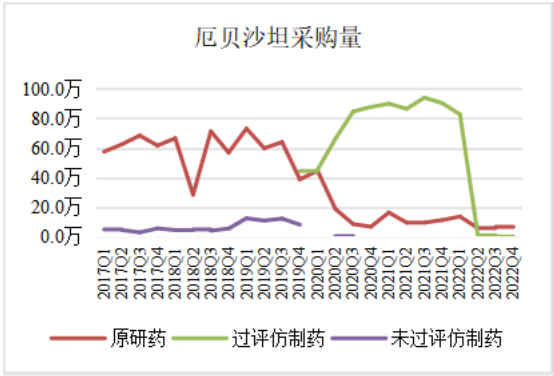
图 1(b) 2017-2022 年过评药品采购量变化

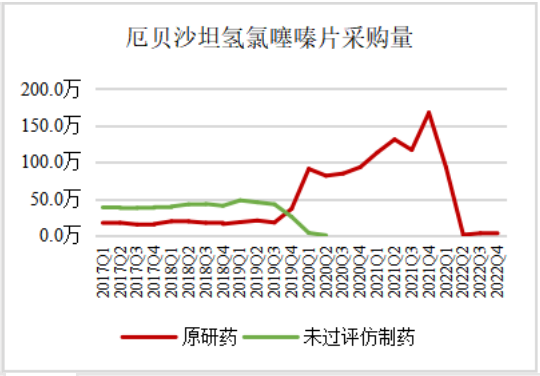
4.1.2 过评药品与对应原研药及未过评药品的采购量比较

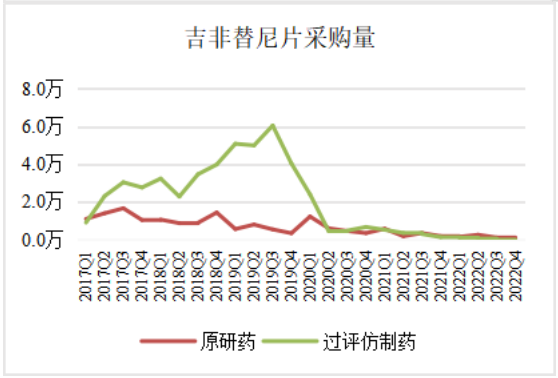
从具体品种的情况看，除厄贝沙坦氢氯噻嗪片和硫酸氢氯吡格雷片之外，其余品种原研药的采购量在 2017-2022 年之间均呈现波动下降的趋势，并且有 5 个品种原研药品的采购量在 2020 年之后低于过评品种的采购量，同时未过评仿制药的采购量数据大多集中在 2020 年之前（图 2）。这些情况都表明集采后，过评仿制药的市场份额正逐渐增大。

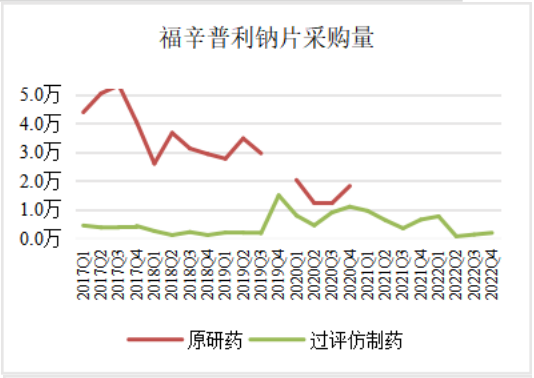
但是，对比过评药品价格后发现，部分过评仿制药品的采购量并没有随价格的大幅下降而上升（比如盐酸帕罗西汀片），甚至有的过评药品的采购量是波动下降的（比如吉非替尼片、福辛普利钠片、富马酸替诺福韦二吡呋酯片、赖诺普利片和瑞舒伐他汀钙片），其原因可能是原研药在集采前就已经占据相当大的市场份额，所以集采后的一段时间内过评仿制药的采购量仍受其影响（比如

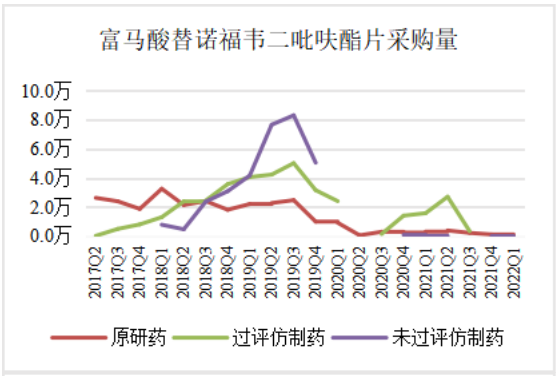
福辛普利钠片、瑞舒伐他汀钙片），也有可能是医生和患者对过评仿制药认可度不高，尽管原研药价格高，依然倾向使用原研药。

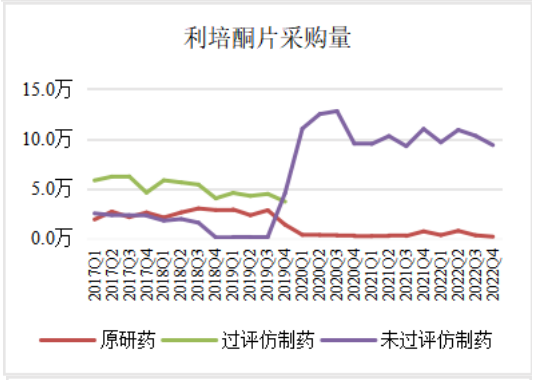


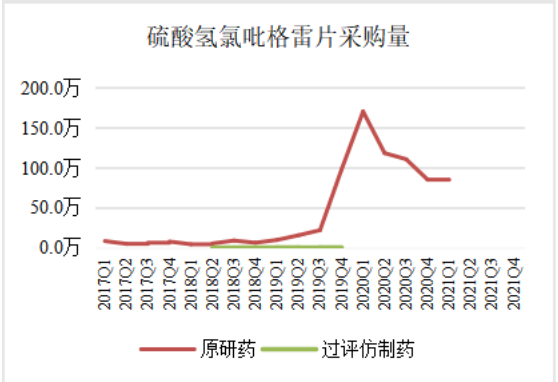


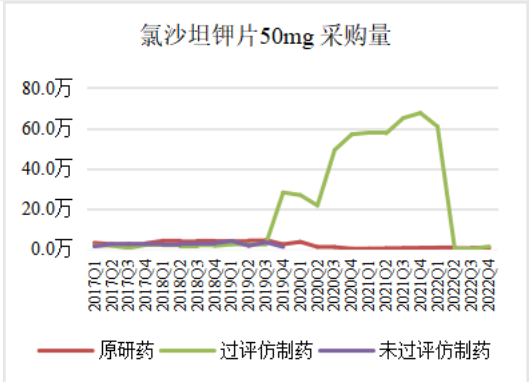


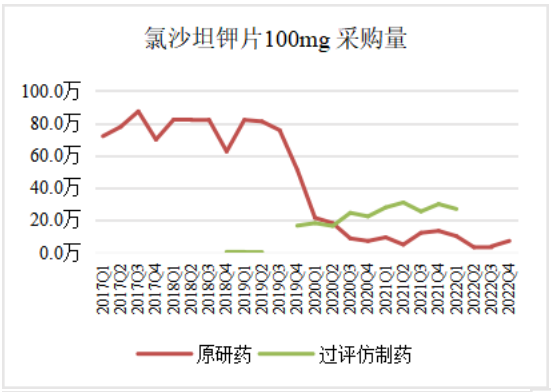


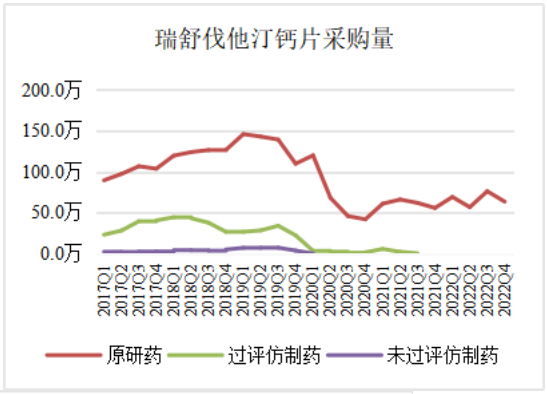


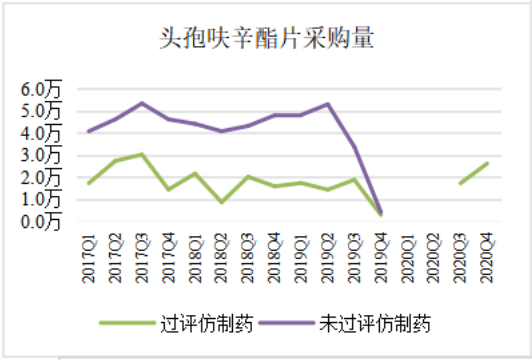


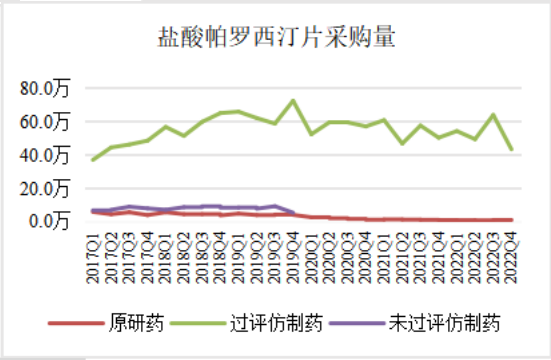


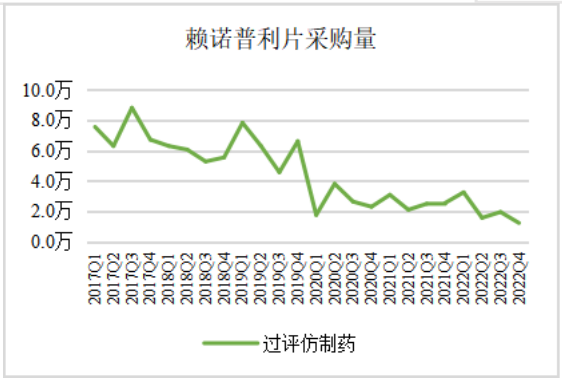












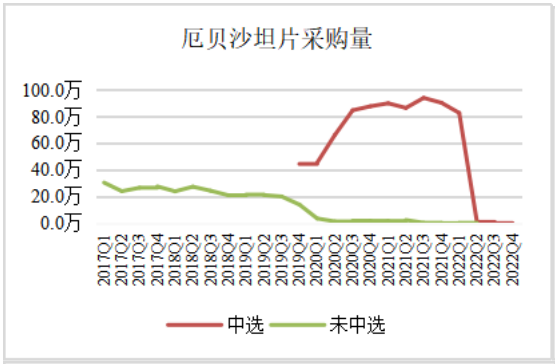
*赖诺普利片在山东仅有过评仿制药的数据，因此无法做原研药、过评仿制药和未过评仿制药的对比分析

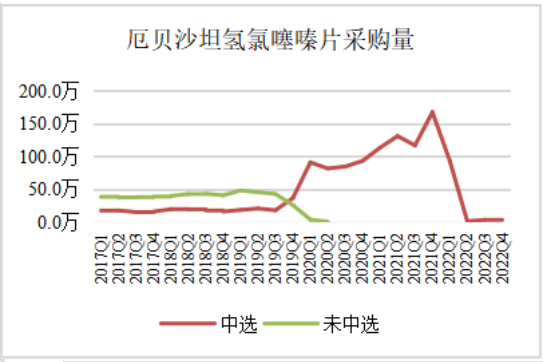
图 2 过评仿制药、原研药、未过评仿制药采购量对比

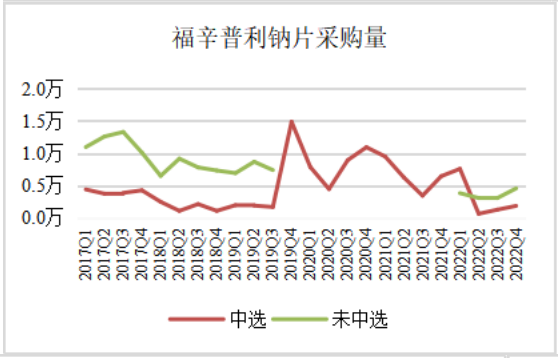
4.1.3 集采中选品种和未中选品种的采购量变动情况

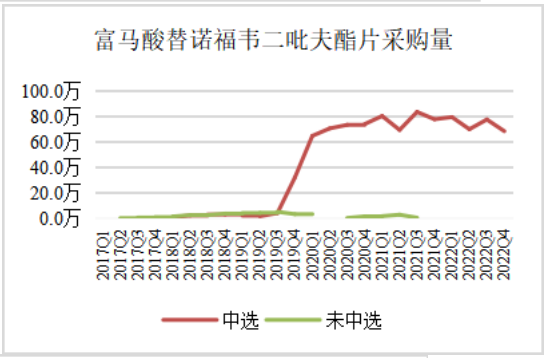
2017-2022 集采品种采购量变动趋势如图 3 所示。13 个药品中，有 10 个药品的中选品种采购量在 2019 年第四季度和 2020 年第一季度后有一个上升的趋势，其中瑞舒伐他汀钙片增幅最大（472.93%），硫酸氢氯吡格雷片增幅最小（74.08%）。药品集中采购显著提高了中选品种的临床使用量。

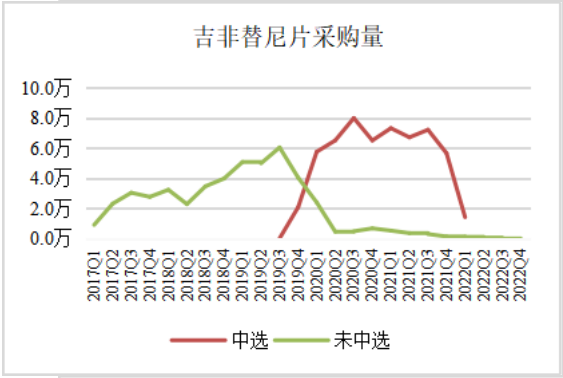
集采中选后采购量在 2019 年第 4 季度到 2022 年第 4 季度期间呈波动下降趋势的药品有福辛普利钠片、赖诺普利片。这两种药品的价格在这段时间内一直保持稳定，药品采购量的下降可能与需求减少或者供应不及时有关。相较集采中选品种，未中选品种的采购量全部呈现下降趋势，且大多在 2020 年之后出现断点或降为 0。

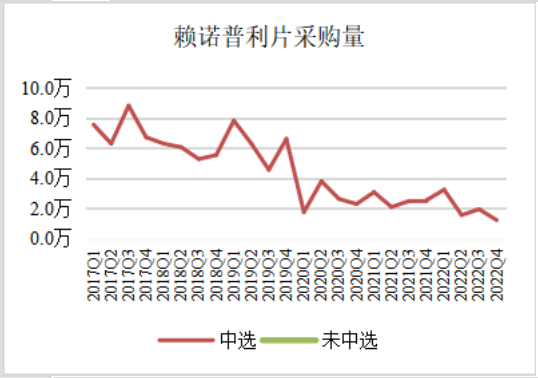

厄贝沙坦片采购量
100.0万
80.0万
60.0万
40.0万
20.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

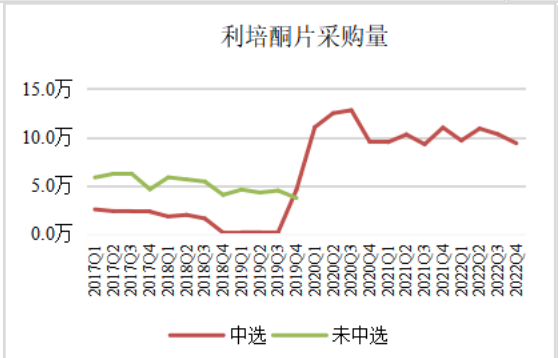

厄贝沙坦氢氯噻嗪片采购量
200.0万
150.0万
100.0万
50.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

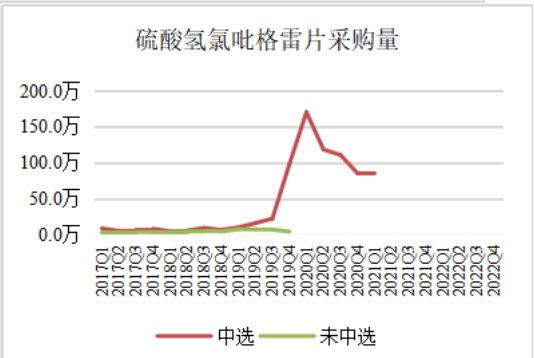

福辛普利钠片采购量
2.0万
1.5万
1.0万
0.5万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

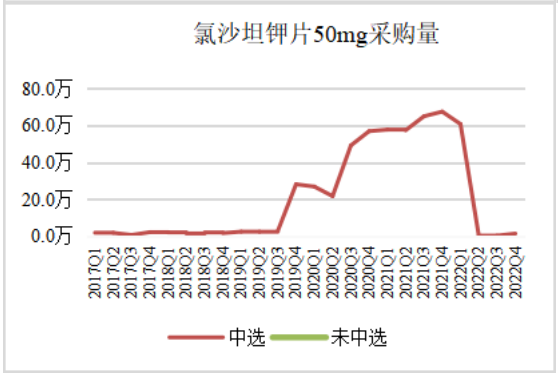

富马酸替诺福韦二吡夫酯片采购量
100.0万
80.0万
60.0万
40.0万
20.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

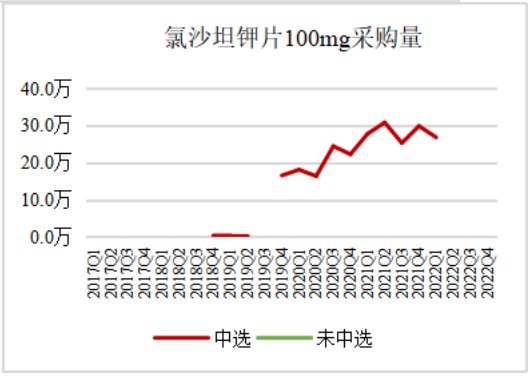

吉非替尼片采购量
10.0万
8.0万
6.0万
4.0万
2.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

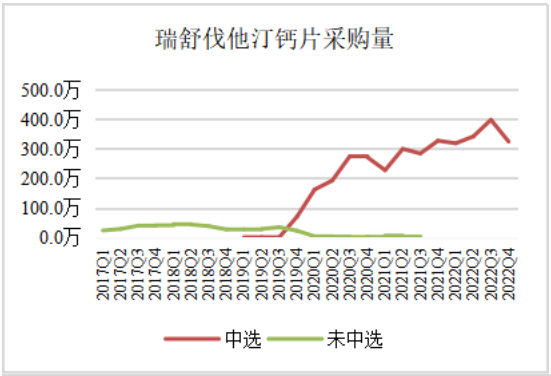

赖诺普利片采购量
10.0万
8.0万
6.0万
4.0万
2.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选

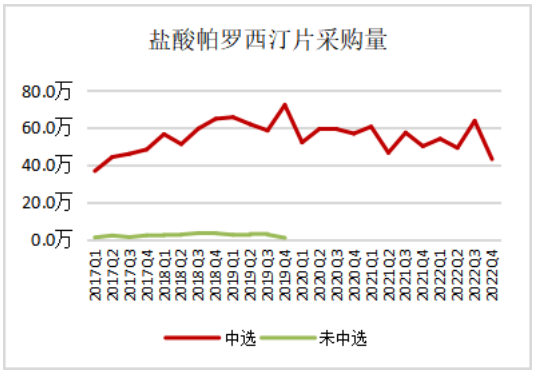

利培酮片采购量
15.0万
10.0万
5.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选


硫酸氢氯吡格雷片采购量
200.0万
150.0万
100.0万
50.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选


氯沙坦钾片50mg采购量
80.0万
60.0万
40.0万
20.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选


氯沙坦钾片100mg采购量
40.0万
30.0万
20.0万
10.0万
0.0万
2017Q1 2017Q2 2017Q3 2017Q4 2018Q1 2018Q2 2018Q3 2018Q4 2019Q1 2019Q2 2019Q3 2019Q4 2020Q1 2020Q2 2020Q3 2020Q4 2021Q1 2021Q2 2021Q3 2021Q4 2022Q1 2022Q2 2022Q3 2022Q4
中选
未中选





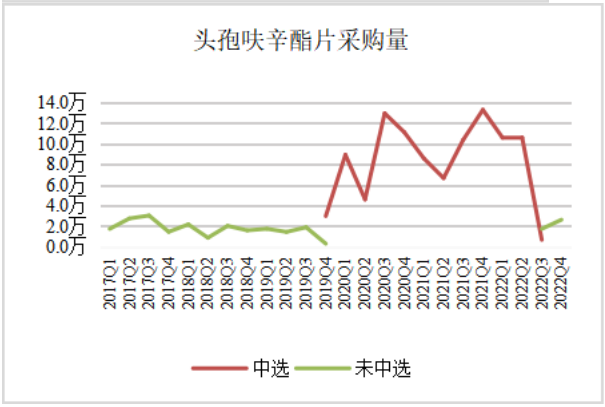


图 3 2017-2022 年集采品种采购量变动趋势

4.2 各品种配送到货率情况

2021 年和 2022 年，12 个品种在山东有采购记录的所有过评企业总配送到货率均大于 80%，且基本稳定在 90%以上（表 1）。山东省医疗保障局规定连续两年配送到货率为零的药品，采购平台予以撤网，两年内不得重新挂网[4]，同时 2023 年省医保局发布，目前中选产品供应配送率达到 80%以上[5]，与本研究数据结果一致，说明山东省配送到货率总体情况较好。

总配送到货率最低的为吉非替尼片 84.46%，赖诺普利片、硫酸氢氯吡格雷片的配送到货率近两年均为 100%。但是分析发现，有一半的药品出现了 2022 年配送到货率低于 2021 年的情况，其中下降幅度最大的为厄贝沙坦氢氯噻嗪片。两年配送到货率上升幅度最大的药品为福辛普利钠片 18.18%。按医疗机构等级看，配送到货率最高的为基层医疗机构（表 2）。

表 1 过评药品配送到货率情况（%）

药品名称	2021 年到货率	2022 年到货率	总到货率
厄贝沙坦片	97.47	94.70	96.09
厄贝沙坦氢氯噻嗪片	100.00	91.06	95.53
福辛普利钠片	81.82	100.00	90.91
富马酸替诺福韦二吡呋酯片	90.47	92.25	91.36
吉非替尼片	86.36	82.56	84.46
赖诺普利片	100.00	100.00	100.00
利培酮片	93.83	96.41	95.12
硫酸氢氯吡格雷片	100.00	100.00	100.00
氯沙坦钾片 0.1g	90.00	100.00	95.00
氯沙坦钾片 500mg	92.50	89.12	90.81
瑞舒伐他汀钙片	94.53	93.72	94.13
头孢呋辛酯片	98.66	97.51	98.09
盐酸帕罗西汀片	97.85	97.18	97.52

表 2 2021/2022 年不同等级医疗机构过评药品平均配送到货率

医院等级	配送到货率（%）
基层医疗机构	96.32
二级医疗机构	94.53
三级医疗机构	91.76

5 讨论

药品通过仿制药一致性评价后，采购量明显上升的只有三个品种（富马酸替诺福韦二吡夫酯、吉非替尼和盐酸帕罗西汀），且只有富马酸磷酸二吡呋酯片的采购量呈持续上升趋势，其余品种整体变动趋势不明显或呈下降趋势，可能是由于仿制药过评后价格升高，受价格的影响，药品采购量出现下降或无法持续的上升趋势。2019 年 12 月 1 日起，山东开始执行国家组织药品集中采购和试点扩围，部分品种（福辛普利钠片、富马酸替诺福韦二吡夫酯片、吉非替尼片、赖诺普利片和盐酸帕罗西汀片）在此期间出现了一个使用量的峰值，但在随后几个季度中其采购量又出现不同程度的下降或波动下降的趋势，仅厄贝沙坦片和氯沙坦钾片 50mg/100mg 有明显持续上升的趋势。使用量持续上升可能有两个原因，一是这两种过评药品均无完全可替代品种，也就意味着其在临床使用上竞争较小；二是这两种药品的过评企业数均较多，药品供应量可以得到保障，厄贝沙坦片和氯沙坦钾片五年累计的过评企业数分别为 10 家和 7 家。

值得注意的是，过评药品在集采期间价格均下降，然而其使用量并没有随价格的下降而上升，反而出现波动下降的趋势，分析其中的原因可能与三年特殊时期有关，人口流动量下降，就医需求相对较低会直接导致用药量下降。另外，福辛普利钠片、吉非替尼片和赖诺普利片这三种产品在 2017 年只有 1-2 家过评企业，在这之后连续五年均无一家企业过评，药品过评企业数过少也可能会引出供应上的问题。过评药品使用量的下降也提示有可能存在临床同类替代现象，目前虽然药品实行零差率政策，公立医疗机构的收入方式发生转变，但还是能从药品供应链业务中获取一定的利益，因此医疗机构缺乏主动使用中选品种的动力[6]，比如富马酸替诺福韦二吡呋酯片的完全可替代品种丙酚替诺福韦，该药在 2020 年第 2 季度到 2022 年第 4 季度，其使用量一直高于富马酸替诺福韦二吡夫酯片，说明该过评品种在临床上大部分已被其完全可替代品种替代。对仿制药的认知和临床用药习惯，也会影响过评仿制药的使用量，在一项中国居民对仿制药的认知调查中显示，我国居民对仿制药的了解程度较低，对仿制药的认知也存在一定的偏差，认为即使通过一致性评价的仿制药，其临床疗效和安全性仍然不及原研药[7]。

另外，在本研究中还发现 2021-2022 年山东省不同等级医疗机构的平均配送到货率虽然都在 90%以上，到货率从高到低依次是基层医疗机构、二级医疗机构、三级医疗机构。分析可能的原因是不同等级医疗机构间的采购品种和数量不完全相同，从基层医疗机构、二级医疗机构到三级医疗机构，总体上药品采购品种数和采购量都是依次递增的，品种数越多出现不能足量配送的情况越多，而当配送企业备货量不足时，对基层医疗机构的需求或许可以满足，而对于二、三级医疗机构的大量采购可能就无法全部满足，从而到货率较低[8]。

针对以上可能存在的问题，可以从以下方面入手改善过评药品的供应保障，首先要制定仿制药研发激励机制，提供专项资金给予支持，鼓励制药厂商参与仿制药一致性评价，特别是优先开展过评企业数较少的品种的一致性评价和审批工作；其次要完善过评仿制药的宣传教育体系，医生、药师和患者三方缺一不可，可以落实经济激励措施，合理引导医生和药师主动使用集采中选品种，同时加强对居民的宣传教育，提高患者对仿制药的认知和信任度；最后在药品生产、流通到销售的各个环节，医疗机构都应加强与药品生产企业的沟通，以提高药品的供应保障。

6 结论：

本研究从过评药品的采供使用情况和配送到货率情况这两个方面进行分析，发现部分药品的采购量并没有随价格的下降而上升，配送到货率方面总体良好。下一步政策应针对性的向过评仿制药的采购监测方面倾斜，以切实提升患者的用药保障水平。

PU-122

基于 BERTopic 与 Citespace 的中医药领域研究热点与主题演化分析

张文婷[1] 王安莉[2]
1. 中南大学
2. 中南大学湘雅三医院

中医药是中华民族的智慧结晶，在漫长的历史长河中发展出独特的诊疗体系，具有重要的挖掘潜力和开发、推广应用价值。[错误!未找到引用源。] 探索中医药领域的研究热点与发展趋势，有助于更深入地把握该领域的发展动态。毛学韫等[错误!未找到引用源。] 利用 CiteSpace 梳理了中医药现代化发展进程的各个阶段，探索中医药现代化可能的未来发展方向。陈娜等[错误!未找到引用源。] 采用 CiteSpace 可视化分析、关键词共现和聚类分析的方法分析了中医药健康服务领域的研究现状、热点与趋势。

目前中医药领域热点研究的文献，落脚点主要是具体的某一细分领域，研究方法主要是借助 Citespace 等文献计量学工具，从关键词词频分布特征等信息来挖掘研究热点，而通过主题挖掘来研究中医药领域热点的文献较少。陈琦等[错误!未找到引用源。] 采用 LDA 主题模型对中医药人工智能领域的文献进行挖掘，发现了主题演化路径并进行了可视化展示。主题模型能够发现海量文本中隐含的抽象主题，并呈现主题间的关联关系，对单一的关键词分析方法是一个有力的补充。

基于此，本研究采用主题挖掘与知识图谱可视化分析结合的方法，通过 BERTopic 模型进行主题建模，从全局把控中医药领域总体研究内容，再聚焦于细分的子主题，采用 Citespace 对中医药发展主题的文献进行计量学分析,以探索中医药领域发展过程中的研究热点与演化脉络。

1 数据来源与研究方法

1.1 数据来源与预处理

本文的研究数据来源于万方数据库中关于中医药领域近 10 年的文献，以“(关键词:(中医药)) and Date:2013-2022”为检索式，文献来源选择 CSTPCD、北大核心、CSSCI、CSCD。导出文献标题、摘要、出版时间等信息，删除摘要重复及摘要或时间缺失的文献，最终获得 12683 条有效数据。

对于中文内容，在进行主题建模之前需要根据语义进行分词。本研究结合中医药领域相关专业词汇，构建自定义的停用词、同义词和词典，采用 python 的 jieba 模块进行分词。

1.2 研究方法

本文采用主题模型 BERTopic 对论文摘要进行主题抽取分析,BERTopic 是一种基于 BERT 词向量的主题建模技术。BERT 是谷歌新推出的 NLP 模型,它以 Transformer 作为特征抽取器,并采用双向语言模型作为预训练任务, 提供预编译的模型,为现今最为常用的、获得研究者一致认同的预训练词向量模型。[错误!未找到引用源。]

本研究采用“paraphrase-multilingual-MiniLM-L12-v2”模型作为嵌入，将每个文档转换为高维向量表示，这些向量保留了文档的上下文信息和语义表示；接着通过 UMAP 降维、kmeans 聚类算法以及 c-TF-IDF 来创建密集的集群。[错误!未找到引用源。] 利用主题模型生成“文档-主题-关键词”矩阵，统计各主题的出现频次以得到主题强度。然后引入时间戳，计算热点主题在 2013-2022 年间的主题强度，绘制主题强度演化趋势图。

在此结果的基础上，深入研究中医药领域的发展这一子主题，通过 Citespace 绘制关键词共现图谱与关键词突现图谱，并使用 Timeline View 功能绘制时间线图，分析该领域发展过程中的研究热点与演化路径，以期为中医药领域的研究提供一些思路和参考。

基于主图模型与知识图谱的中医药领域热点分析的流程图如图 1 所示。

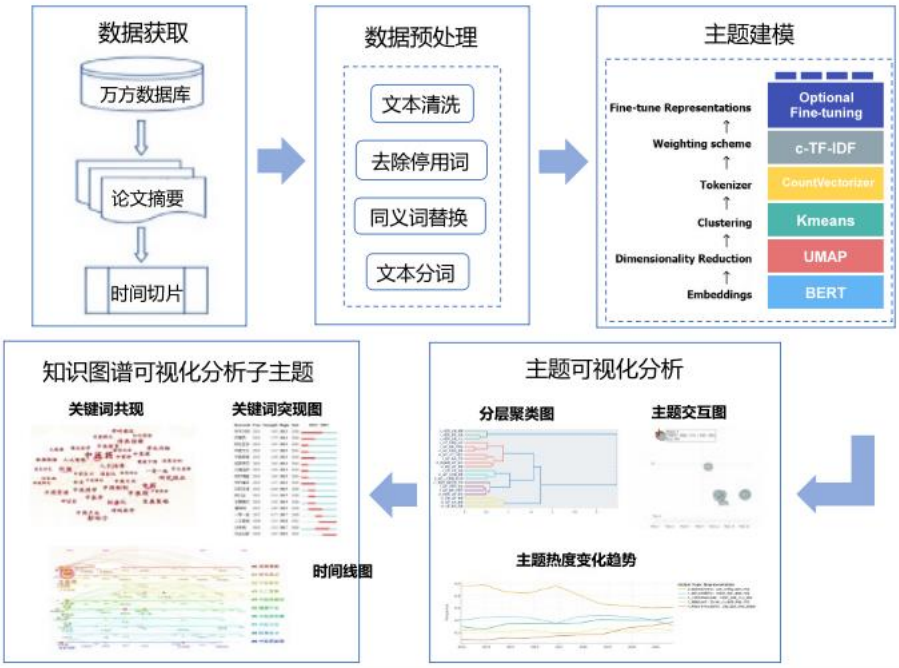


图 1 基于主图模型与知识图谱的中医药领域热点分析流程图

2 研究过程与结果

2.1 主题建模

采用默认参数进行主题建模，模型将自动选择合适的主题数量。通过 Umap 将主题的 c-TF-IDF 表示嵌入到二维中，然后将其可视化，生成的交互式图如图 2 所示。错误!未找到引用源。滑动页面下方的滑块，相应的主题会被标红，以便查看该主题的信息，包括主题的大小及其对应的关键词。图中圆圈的大小与主题对应的文档数呈正相关，相似度越高的主题之间距离越近，重合率极高的可以视为一类主题，据此可以从图中看出主题数大致为 20 个。

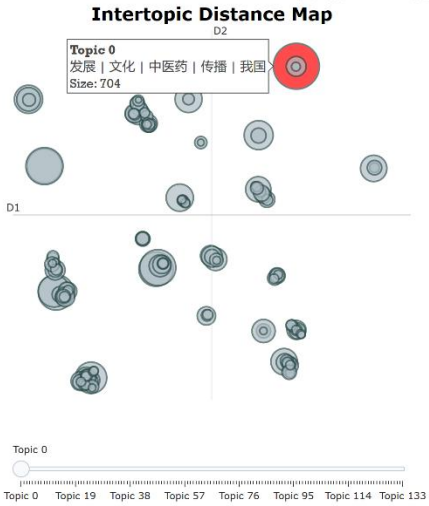


图 2 2013-2022 年中医药领域研究主题交互图

为了全局性地概览这 20 个主题，采用 kmeans 聚类方法，手动地将主题缩减至 20 个，得到的结果如图 3 所示，从每个主题的主题词集合可以识别出主题内容，例如，通过主题 1 中的主题词“分析、中药、文献、中医药、治疗、药物、临床、数据库、患者、纳入”,判断出该主题内容为“中医药在临床治疗中的药物分析与文献研究”。

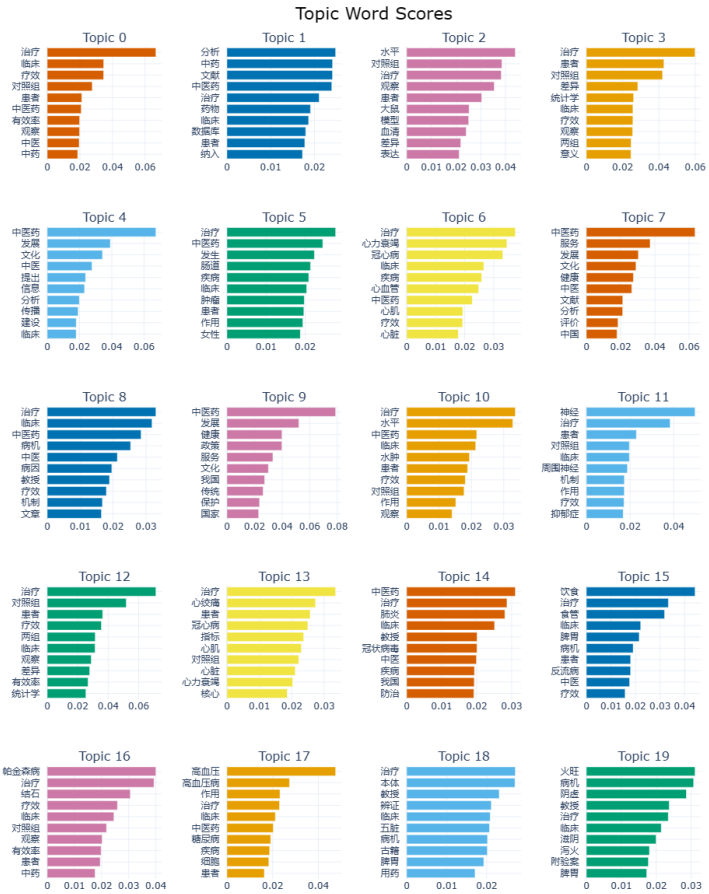


图 3 2013-2022 年中医药领域研究主题

模型训练好后，为了进一步了解主题的潜在层次结构，使用 scipy.cluster.hierarchy 创建聚类，以直观地显示不同主题之间的联系，如图 4 所示。**错误!未找到引用源。** 应用余弦相似度创建相似度矩阵，其可视化热图如图 5 所示。

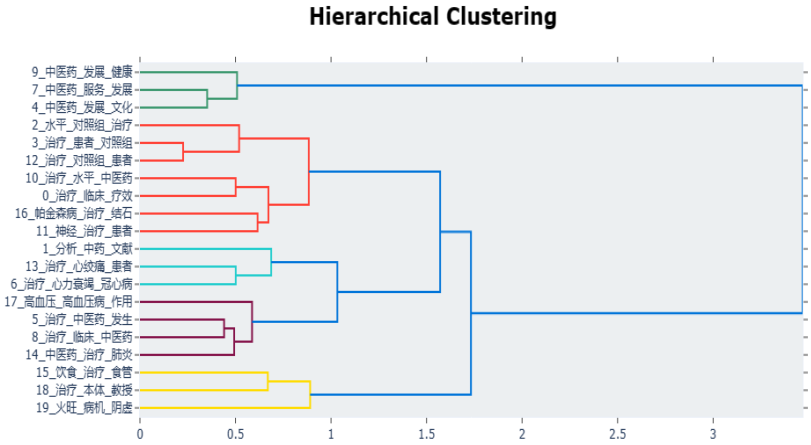


图 4 2013-2022 年中医药领域研究主题分层聚类图

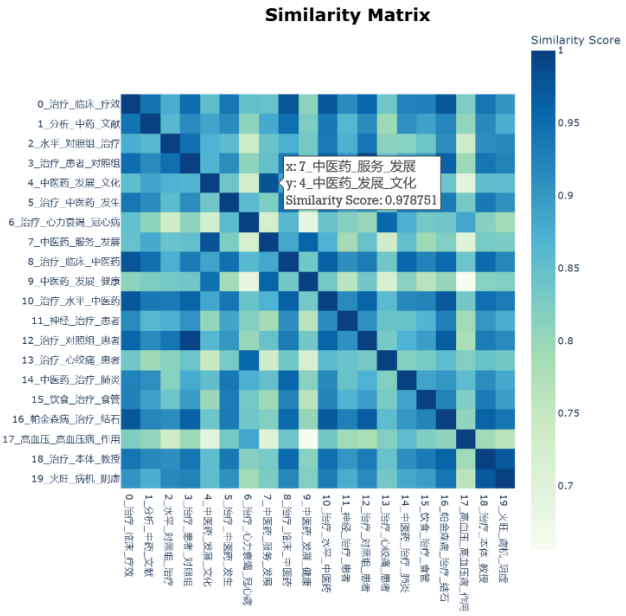


图 5 2013-2022 年中医药领域研究主题相似度热图

结合图 3 中的关键词内容，以及图 4 与图 5 所显示的主题之间的相似度与关联情况，可以从全局来把握中医药研究文献的总体内容。主题 9、7、4 的话题是中医药领域的发展与建设，聚焦于中医药文化的传承与推广、国家政策、健康服务等内容。主题 2、3、12、10、0 的内容是中医药临床实验与疗效，学者们通过建立大鼠模型、进行随机对照实验等研究中医药治疗的临床疗效，探讨中医药应用于各类疾病的干预效果。**错误!未找到引用源。** 主题 16、11、13、6、17 涉及的高频关键词有神经，抑郁症，心绞痛、冠心病、心力衰竭、心血管、高血压、糖尿病，帕金森等，主题是中医药在心血管疾病、神经系统疾病、内分泌系统疾病等的治疗。相较于西医，中医对这类慢病有着独特的诊疗方式,并广受认可，如食疗、运动疗法、针灸、拔火罐、推拿按摩、中药沐足、药酒疗法、药茶疗法等等。**错误!未找到引用源。** 主题 5、8、14、15、18、19 的相似关键词有“病机、病因、教授、疗效、辩证，饮食”等等，中医讲究辨证论治，有着悠久的历史和丰富的经验积累的同时，也具备个体化诊疗的特点与经验性传承的特色。名中医和专家会进行总结和归纳，以分享他们的经验和见解，探索病因与病机理论、中医药的作用机制和调控途径等。**错误!未找到引用源。**

根据相似度合并近似的主题后，得到的五个主题如表 1 所示。

表 1 2013-2022 年中医药领域研究主题

主题	关键词	主题解释	文档数
0	治疗_对照组_临床_疗效	临床实验与疗效	5054
1	中医药_病机_教授_作用	病机与药理研究	2701
2	中医药_发展_文化_服务	中医药领域的发展	2280
3	冠心病_心力衰竭_疾病_疗效	慢病的治疗	1470
4	文献_临床_药物_数据库	药物分析与文献研究	1178

2.2 主题演化分析

为了进行主题演化分析，在文档-主题矩阵中引入时间维度，然后基于文档-主题矩阵和时间信息，计算每个主题在不同时间点的强度，即在当前时间（年份）所有文档数中该主题所占的权重。最后利用得到的主题强度数据，绘制主题演化趋势图。图 6 为 5 个主题随时间演化的趋势图，图中横轴代表时间(年份),纵轴代表主题强度值。

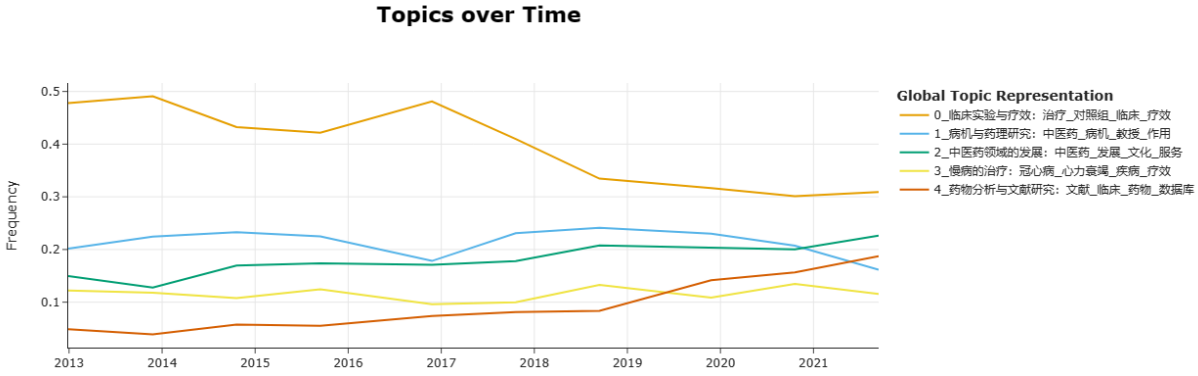


图 6 2013-2022 年中医药领域研究主题热度变化趋势

研究结果显示,主题 0（临床实验与疗效）研究热度呈下降趋势，但主题强度一直居高，其涉及中医治病疗效的核心内容，具有重要的理论和实践价值，是中医药领域中重要的研究方向。错误!未找到引用源。可能因为学者们已经在此领域进行了广泛的研究，加上其他研究方向（如中医药领域的发展）在中医药领域受到更多关注，所以此主题的文档所占比例随着时间有所下降。主题 1（病机与药理研究）和主题 3（慢病的治疗）的研究热度整体呈平稳波动趋势。而呈上升趋势的主题有主题 2（中医药领域的发展）和主题 4（药物分析与文献研究），这两个主题在中医药研究领域越来越受到关注。事关中医药领域的建设、文化传承与推广等内容，有国家政策的关注和扶持。错误!未找到引用源。随着时间的推移，政府、研究机构等会加大对中医药领域的投入，进一步推动其建设与发展。

2.3 研究热点分析

根据主题模型聚类得到的结果，'中医药领域发展'这一主题近年来备受关注，主题热度逐渐增长的同时，发文数量也居于前列。为了探索中医药领域发展的脉络，有必要对这一主题的文献进一步分析。在主题建模得到的'中医药领域发展'主题文档基础上，整合从万方数据库检索得到的'中医药发展'主题的补充数据。采用 CiteSpace 软件对文献的关键词进行可视化分析，并通过 Timeline View 功能绘制时间线图。

（1）关键词共现分析

关键词能够对论文主题和论文内容进行归纳总结，准确的关键词能够对论文进行高度概括。中介中心性主要衡量节点在网络中的重要性，是评价文献重要性的主要指标错误!未找到引用源。。中医药领域的发展主题关键词分布如图 7 所示，高频关键词中心性如表 2 所示，可见中医教育、人才培养、传承创新、文化传播、国际化、标准化、人工智能等都是中医药领域发展的研究热点。

图 7 中医药发展主题关键词共现图（频率≥8）

表 2 中医药发展主题关键词及其中心性

序号	关键词	频次	中心性	序号	关键词	频次	中心性
1	中医药	251	0.39	16	研讨会	15	0.06
2	中药	75	0.23	17	中药复方	14	0.06
3	标准化	33	0.2	18	治未病	13	0.06
4	传统医学	23	0.19	19	中药制剂	33	0.05
5	中医学	24	0.18	20	中药资源	26	0.05
6	国际化	16	0.17	21	人工智能	22	0.05
7	中医院	54	0.15	22	针灸	19	0.05
8	人才培养	30	0.14	23	数据挖掘	16	0.05
9	研究现状	36	0.12	24	健康中国	14	0.05
10	学科建设	17	0.09	25	教学改革	11	0.05
11	发展策略	37	0.08	26	课程思政	11	0.04
12	传承创新	37	0.07	27	中药炮制	10	0.04
13	中药产业	21	0.07	28	中医文化	20	0.03
14	发展模式	14	0.07	29	中医教育	19	0.03
15	中医药学	28	0.06	30	循证医学	14	0.03

（2）研究热点及演化分析

关键词突显图谱是高度活跃研究热点的指示器，高突发强度的关键词在某一时期的出现，表明相应主题在特定时期引起了业界的特别关注。错误!未找到引用源。关键词突现图谱如图 8 显示，自 2015 年屠呦呦女士获得诺贝尔奖，“中药资源”、“研讨会”、“发展模式”这些词一时大热。随着互联网和信息技术的发展，各行各业都被前沿技术所惠及，其中也包括有海量医疗数据积累的中医药领域。自 2018 年起，“人工智能”这一具有高突发强度的主题词出现，至今仍是中医药领域新兴的研究热点之一。2017 年出现的“一带一路”与 2019 年出现的“守正创新”均与时政密切联系，国家政策的制定积极促进着中医药传承创新发展。错误!未找到引用源。

图 8 中医药发展主题关键词突现图谱

绘制关键词时间线图，如图 9 所示，关键词的字体大小与其出现频率成正比。根据关键词突现图与时间线图，将中医药领域发展研究分为三个阶段，第一阶段（2013-2014 年）突出表现为关键词密集，表明中医药领域正处于积极扎根的发展时期，研究焦点集中在中医教育、人才培养以及临床研究等方面，为中医药的进一步发展奠定了坚实的基础。第二阶段（2015-2017 年）迎来中医药领域的高速发展，热点主题为“屠呦呦”、“中药资源”、“发展模式”、“一带一路”等。自 2015 年屠呦呦女士携青蒿素享誉海内外，中医药也迎来各方更广泛的关注与扶持，迈入蓬勃发展阶段。关键词“大数据”首次出现在 2016 年，此后与新兴技术的应用有关的关键词不断涌现。第三阶段（2018 到 2022 年）热点主题为“人工智能”、“守正创新”等，在政策和技术的双重加持下，中医药领域步入了一个新的高速发展阶段。

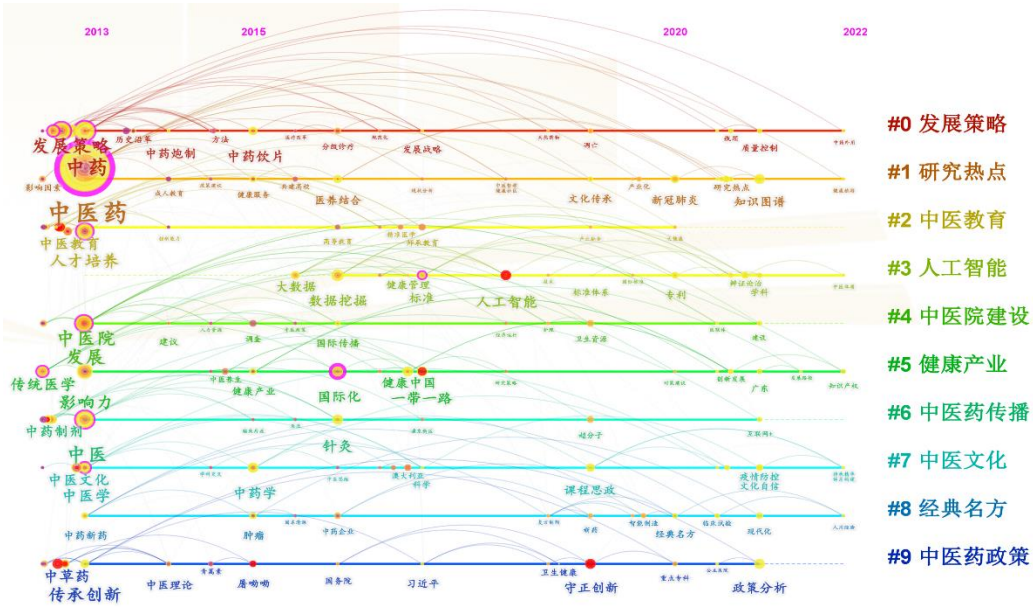


图 9 中医药发展主题关键词时间线图

3．结论与讨论

本文利用 BERTopic 主题模型抽取了万方数据库中有关中医药领域文献的研究主题，经过分析归纳后共得到临床实验与疗效、病机与药理研究、中医药领域的发展、慢病的治疗和药物分析与文献研究这五个主题，结合相关文献与专业知识，根据主题强度值对各主题随时间演化的趋势进行了分析讨论，其中“中医药领域发展研究”这一主题热度呈增长趋势，为了进一步了解中医药领域发展的研究热点与演化路径，本文基于 Citespace 对此主题的文献进行计量学分析，结果显示，人才培养、文化的继承与传播、发展政策、新兴技术等是中医药领域发展绕不开的重点话题。

中医药文化的传承与推广离不开政策的推动，自实施《中国人民共和国中医药法》，中医药的传承和发展已成为国家战略。**错误!未找到引用源。**加快推进中医药现代化、产业化，推动中医和西医相互补充、协调发展，推动中医药走向世界，充分发挥中医药防病治病的独特优势和作用成为新时代的新命题。**错误!未找到引用源。**

中医药教育与人才的培养是另一个重要的议题，政府、社会和高校都需要共同参与构建中医药拔尖人才培养体系，通过改革教育模式和管理机制，提高培养能力，以推动中医药传承创新发展。**错误!未找到引用源。错误!未找到引用源。**

在大数据时代这一背景下，人工智能、互联网+等风潮与中医药领域的深度融合趋势已然显现，这些新兴的技术将在中医临床诊疗辅助、中医健康管理、名中医经验传承、治未病等的发展中起到卓越的推动作用。深入挖掘中医诊疗技术的科学内涵，借助跨学科的信息科学技术有效利用数据信息，将提高中医诊疗模式的科学性和效能，为中医药现代化、国际化、标准化发展带来机遇。

PU-123

基于网络药理学探讨定喘宜肾宜肾汤对肾透明细胞癌的作用机制

李渊 于琦* 齐荣暄 王灿 薛丹阳 纪之琳
山西医科大学

肾透明细胞癌(ccRCC)是肾细胞癌中最常见的一种组织学分型，占肾细胞癌的 75%以上[1]。全球每年死亡人数在癌症死亡人数中位居前列[2]。有氧糖酵解是肿瘤中重要的代谢特征之一，糖酵解过程可以为肿瘤细胞提供大量能量与其他代谢通路所必需的前体物质[3]。ccRCC在常氧条件下，会优先进行有氧糖酵解，并抑制线粒体的功能，促进细胞增殖且抑制其凋亡[4]。ccRCC作为一种恶性肿瘤，常常伴有代谢通路的改变，通常体现在糖酵解通路的上调和三羧酸循环的下调[5]。

手术切除是治疗肾癌的主要手段，但 20%~40%的术后患者会发生复发或伴有远距离侵袭[6]。现有针对肾癌化疗药物主要有：VHL肿瘤抑制剂、酪氨酸激酶和VEGF抑制剂等[7–9]，但药物的耐药性使得治疗效果逐渐变差，患者术后 5 年生存率仍然很低。大量研究表明，中医药代谢调节对癌症治疗有很好的治疗潜力，其多药理作用在抑制癌细胞代谢过程方面具有优势，同时中医药具有多成分、多靶点、协同增效的作用，能够有效抑制肿瘤的扩散和转移[10]。段万里等[11]发现姜黄素可能参与氧化应激途径抑制肾细胞癌的增殖。高俊杰等[12]发现天然虾青素可以抑制肾透明细胞癌的增殖并抑制其细胞中PGF的表达。刘青松等[13]发现小檗碱可以通过抑制细胞增殖、迁移和血管生成来实现对肿瘤的治疗。

定喘宜肾汤载于《临证医案医方》，有益肾定喘宜肾，化痰，痰鸣，腰腿酸软无力，脉沉细无力的功用。清肺益肾能使肺肾互补，达到治疗的目的。遗憾的是，尚未有定喘宜肾汤治疗肿瘤的潜在作用机制的相关报道。因此，本研究旨在通过网络药理学方法探讨定喘宜肾汤治疗ccRCC的潜在靶点并进一步探讨其在糖酵解途径治疗ccRCC的潜在作用机制。

1 材料与方法

1.1 ccRCC转录组数据及临床数据收集

从癌症基因图谱（TCGA）[14]数据库中筛选获取ccRCC患者二代测序全转录组数据集和患者临床特征数据集。

1.2 定喘宜肾汤活性成分以及药物靶点筛选

利用中药系统药理学分析平台（TCMSP）[15]数据库检索巴戟天、茯苓、枸杞子、核桃、款冬花、苦杏仁、人参、五味子、淫羊藿、紫苏子所含的中药成分，根据口服生物利用度（oral bioavailability，OB）≥30%和类药性（drug likeness，DL）≥0.18 筛选出定喘宜肾汤相关的主要活性成分，并筛选药物靶点。

1.3 ccRCC风险基因及潜在作用靶点筛选

基于TCGA数据集，利用perl语言程序包，针对二代测序数据和临床特征数据进行整理，获取矩阵文件，筛选风险基因。利用Venny2.1 平台（https://bioinfogp.cnb.csic.es/tools/venny/）整合药物靶点与风险基因，获得潜在作用靶点。

1.4 差异分析

基于潜在作用靶点，随后利用R（4.2.1）软件Bioconductor中的Limma程序包筛选差异表达基因，筛选标准为P<0.05，|logFC|≥1.2。

1.5 富集分析

基于潜在作用靶点的差异表达基因利用R（4.2.1）软件Bioconductor中clusterProfiler包、org.Hs.eg.db包、enrichplot包、ggplot2 包进行基因本体论（Gene Ontology，GO）富集分析以及

京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes，KEGG）通路富集分析，随后采用R语言进行可视化展示。

1.6 蛋白互作网络构建以及核心分析

利用STRING（https://string-db.org/）平台，构建蛋白相互作用（Protein-protein interaction，PPI）网络，本研究使用的可靠性阈值为＞0.8。利用Cytoscape（3.8.0）中的CytoHubba插件计算PPI网络中每个节点的Degree值，将Degree值Top10 的mRNA作为Hub基因。

1.7 肾透明细胞癌糖酵解通路关联基因筛选

基于MSigDB数据库（http://www.gsea-msigdb.org）中HALLMARK基因集，使用GSEA（4.2.3）软件对TCGA数据库ccRCC数据进行GSEA富集分析，筛选ccRCC糖酵解相关基因。并利用Venny2.1 整合差异表达基因与糖酵解相关基因，获得糖酵解通路关联基因。

1.8 富集分析

利用R（4.2.1）软件，对酵解通路关联基因进行GO富集分析以及KEGG通路富集分析，随后采用R语言进行可视化展示。

1.9 受试者工作曲线验证

以TCGA数据库ccRCC患者肿瘤样本与癌旁样本为验证组与对照组，使用pROC包绘制受试者工作曲线，统计灵敏度、特异度和曲线下面积（Area Under Curve，AUC），来评估关键基因的预测效能。

1.10 分子对接

根据上述筛选出的活性成分和糖酵解通路关联基因，进行分子对接验证。从 RCSB PDB数据库（https://www.rcsb.org/）下载核心靶蛋白受体的晶体结构，运用PyMoL软件删除靶蛋白受体无关配体和非蛋白分子。利用AutoDockTools（1.5.6）软件对靶蛋白受体和配体小分子进行处理，利用PYMOL（2.5.4）软件进行结果可视化。

2 结果

2.1 定喘宜肾汤活性成分及药物靶点的筛选

从TCMSP数据库中筛选获取定喘宜肾汤主要成分共 184 种化合物，其中巴戟天 20 个、茯苓 15 个、枸杞子 45 个、核桃 4 个、款冬花 22 个、苦杏仁 19 个、人参 22 个、五味子 21 个、淫羊藿 23 个、紫苏子 16 个，得到 246 个药物靶点。并构建成分-靶点网络图，见图 1。

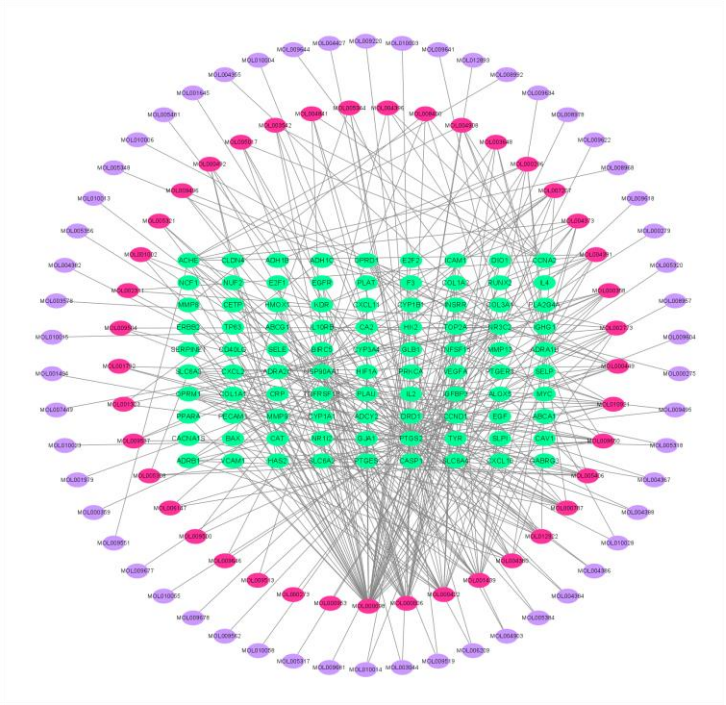


图1 成分-靶点网络图

2.2 潜在作用靶点以及差异表达基因筛选

将药物靶点与风险基因导入Venn2.1 在线分析平台得到潜在作用靶点 232 个，见图 2。以 *P*<0.05，|logFC|≥1.2 为筛选条件，得到差异表达基因共 75 个，其中上调基因 49 个，下调基因 26 个。

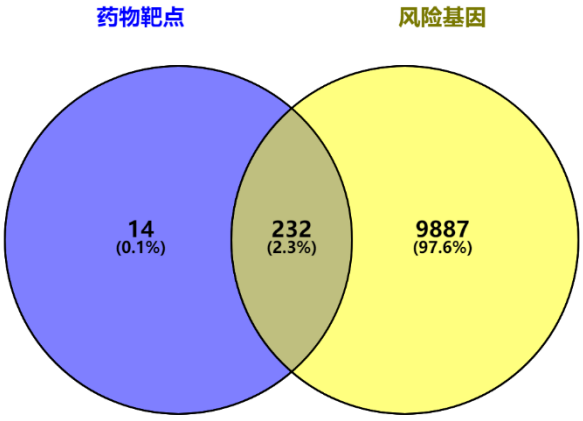


图2 潜在作用靶点韦恩图

2.3 差异表达基因的功能及通路富集分析

对于GO分析，生物学过程（BP）的变化包括对外源性刺激、营养水平、酮、氧化应激、辐射以及活性氧的反应；细胞组分（CC）表达过程中，主要涉及膜阀、膜微区、突出后模以及囊泡腔；分子功能（MF）相关的过程中，主要有核受体活动，配体激活的转录因子活性以及DNA结合

转录因子结合，见图 3。KEGG富集分析发现差异表达基因主要富集于TNF信号通路、AGE-RAGE信号通路、HIF-1 信号通路、PI3K-Akt信号通路等，见图 3、4。

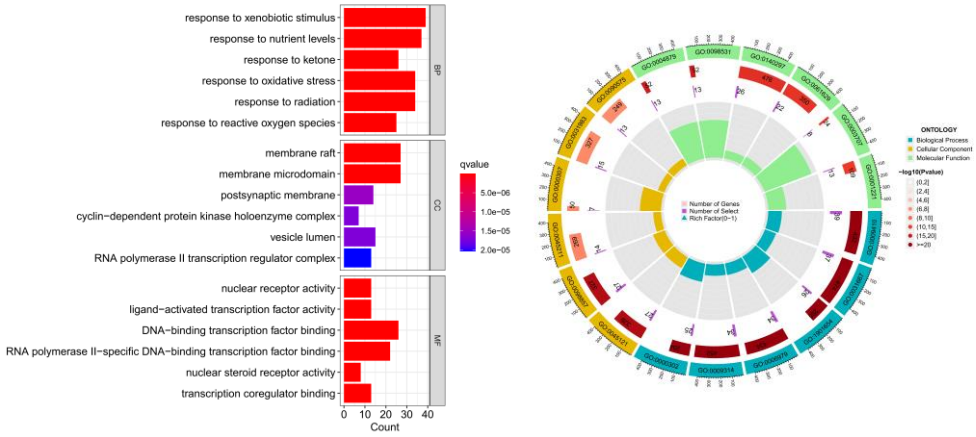


图3 GO 富集分析

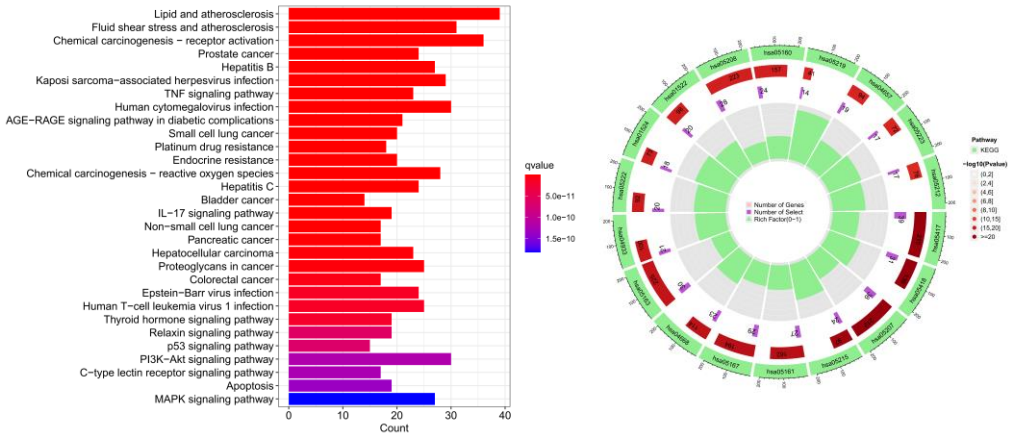


图4 KEGG 富集分析

2.4 蛋白互作（PPI）网络与核心基因

利用STRING平台，构建PPI网络，利用Cytoscape软件，筛选出核心基因（hub gene）为：VEGFA，ESR1，CCND1，CTNNB1，JUN，MYC，HIF1A，EGFR，AKT1，CASP3，见图 5。

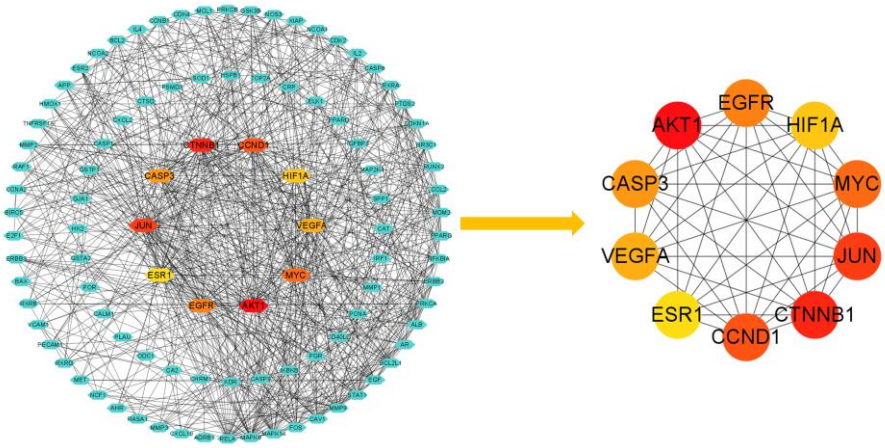


图5 PPI 网络及 HUB 基因图

2.5 糖酵解相关的疾病靶点筛选

对TCGA数据库ccRCC患者的转录组数进行GSEA富集分析，并筛选得到糖酵解相关基因 165 个。

2.6 糖酵解通路关联基因筛选及其功能和通路富集分析

通过整合差异表达基因与糖酵解相关基因，筛选得到糖酵解通路关联基因共 4 个：EGFR，HK2，IGFBP3，VEGFA，见图 6。为了探索潜在作用靶点在糖酵解通路相关的生物学功能和通路，进行了GO和KEGG富集分析。对于GO富集分析，BP的变化主要包括一氧化氮介导的信号转导的调节、肽酰酪氨酸自磷酸化、细胞生长的调节以及腺体发育；MF相关过程中，主要包括纤连蛋白结合、生长因子结合、葡萄糖结合、血管内皮生长因子受体结合等；CC表达过程中，主要体现在囊泡腔、多泡体、血小板α颗粒腔等；KEGG富集分析则主要富集在HIF-1、MAPK、PI3K-Akt等信号通路上，见图 7、8。

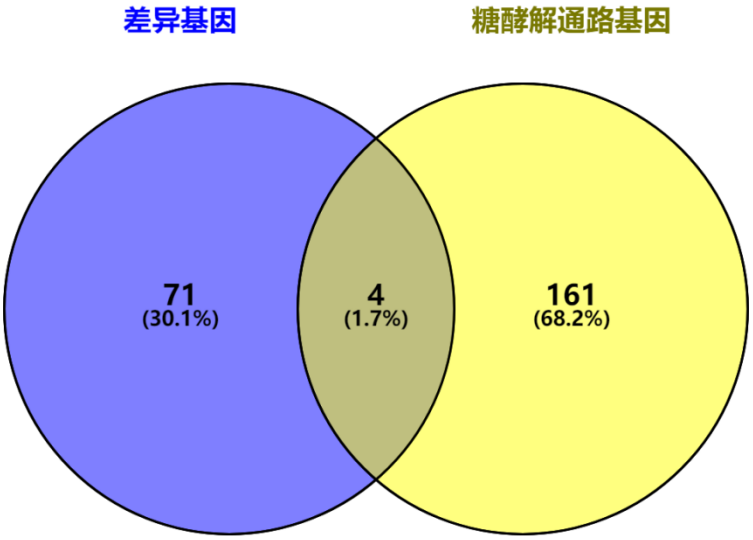


图6 糖酵解通路关联基因韦恩图

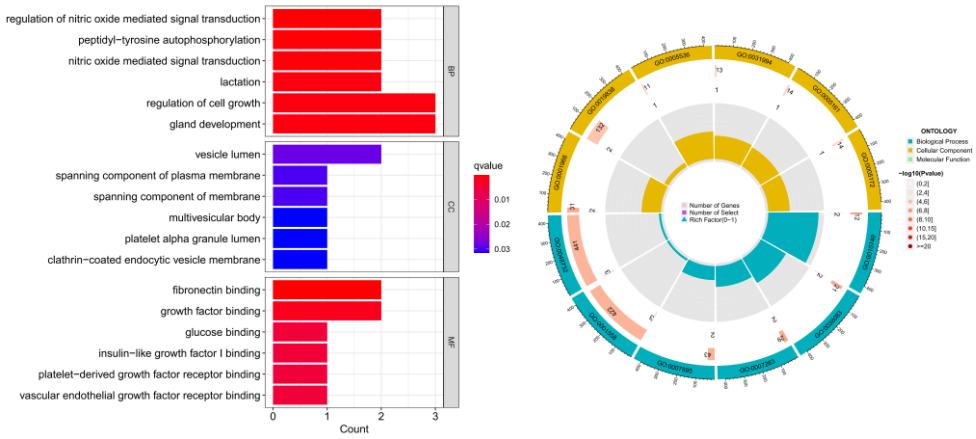


图7 GO 富集分析

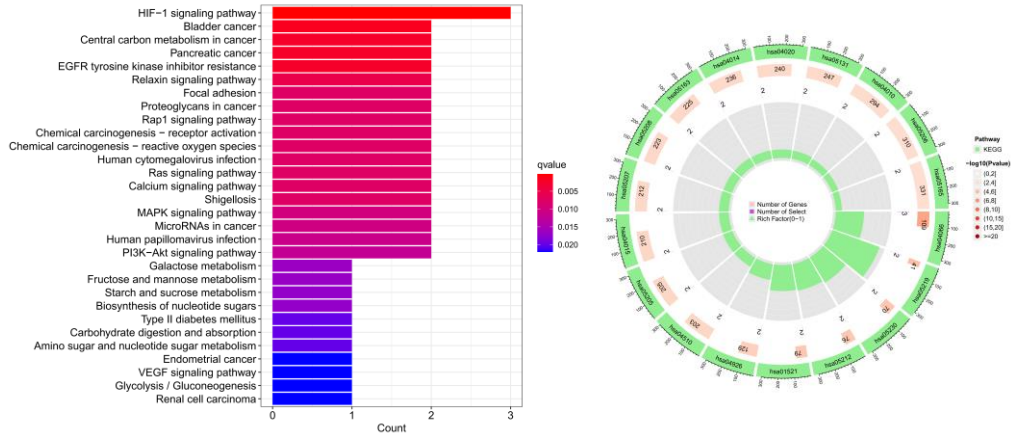


图8 KEGG 富集分析

2.7 受试者工作曲线验证

对糖酵解信号通路关联基因进行受试者曲线分析，见图 9。IGFBP3、VEGFA、EGFR和HK2的AUC值分别为 0.952、0.964、0.868 和 0.958。本研究认为这 4 个基因与ccRCC发病密切相关。

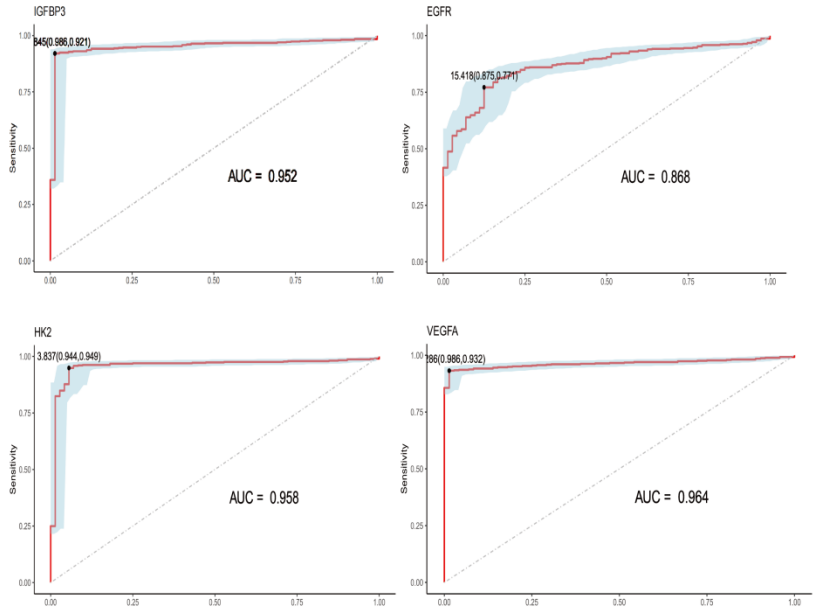


图9 受试者工作曲线图

2.8 vina分子对接

选取糖酵解信号通路关联基因与活性成分进行分子对接验证，糖酵解信号通路关联基因为：EGFR（PDB ID：7tvd），HK2（PDB ID：5hfu），IGFBP3（PDB ID：7wrq），VEGFA（PDB ID：6d3o）。选取自由能最小的三位配体分子为：鞣花酸（ellagic acid）、木犀草素（luteolin）、槲皮素（quercetin），根据成分-靶点网络筛选最佳的七种对接情况，见表1。

表1 活性成分与靶点结合能

配体	受体	结合能/(kcal/mol)
quercetin	7tvd	-7.9
	6d3o	-7.2
	5hfu	-8.8
	7wrq	-7.3
luteolin	7tvd	-7.6
	6d3o	-7.3
ellagic acid	6d3o	-7.2

3 讨论

手术切除是治疗ccRCC的主要方法[16]，但肾癌患者术后复发率高[17]，5年生存质量差，5年生存率不到20%[18–20]。肾癌作为以代谢重编程为特征的恶性肿瘤[14]，突变的基因涉及许多不同的通路，影响细胞代谢的各个方面[21]。在ccRCC中，缺氧诱导因子（HIF）在常氧条件下也会被激活，能够通过糖酵解通路代谢使肿瘤细胞快速成长[22]。此外，肿瘤细胞适应缺氧环境的一种途径是提高糖酵解速率，另一种就是形成多血管体系[23]。

本研究可知定喘宜肾汤有多种活性成分，调控232个治疗ccRCC的潜在作用靶点。研究结果表明，定喘宜肾汤中的木犀草素、槲皮素、鞣花酸都在糖酵解中有着重要的作用。木犀草素作为一种黄酮类化合物，存在于水果、蔬菜和药材中[24]，有效抑制HIF-1α/VEGF相关信号通路的表达，从而降低上皮-间质转化（EMT）和血管生成[25]。其糖苷形式LUT-7G则显著减少了糖酵解早期步骤中的中间代谢物，如6-磷酸葡萄糖（G6P）和6-磷酸果糖（F6P），以及最后阶段的3-磷酸甘油酸（3PG）和磷酸烯醇丙酮酸（PEP），通过减少ATP的合成达到抑制ccRCC发生发展过程的作用[26]。槲皮素作为一种黄酮类化合物，广泛存在于各种水果和蔬菜中，可以抑制多种癌细胞类型的增殖[27]，其通过调节HSP70的表达来诱导肾肿瘤细胞凋亡[28]，通过调节STAT信号通路抑制肿瘤细胞

活性以及其侵袭能力[29]，可以有效降低HK2 的蛋白水平并抑制AKT/mTOR、HIF-1A途径抑制有氧糖酵解，从而拮抗葡萄糖波动引起的肾损伤[30,31]。鞣花酸是一种天然多酚化合物，已被证明对细胞功能有多种影响[32]，通过抑制PI3K和Akt信号通路，不仅能够诱导肿瘤细胞凋亡，还能够抑制其侵袭，其在多肿瘤中都有一定的抑制效果[33,34]。

本研究中VEGFA、EFGR、HK2、IGFBP3、HIF-1A等靶点均证实是ccRCC的关键靶点。血管内皮生长因子A（vascular endothelial growth factor A，VEGFA）是最主要的血管内皮生长因子，在肾癌中，VHL基因的缺失会刺激HIF1，从而诱导VEGF的异常表达[22]，导致ccRCC肿瘤微环境中高血管生成和炎症反应[35]。VEGFA受体抑制剂能够抑制肿瘤血管的生成，在肿瘤的生长和转移中发挥作用[36]。表皮生长因子受体（Epidermal growth factor receptor，EGFR）是表皮生长因子受体（HER）家族成员之一，对细胞的生长、增殖和分化等生理过程发挥重要的作用。EGFR的突变会导致MAPK途径的活化，从提升ccRCC细胞的增殖和分化能力[37]。己糖激酶 2（Hundred Kilobytes Kernel，HK2）是一种转移酶激酶，可将线粒体外膜处的葡萄糖磷酸化为葡萄糖-6-磷酸，是糖酵解途径的起始酶[38]，是催化葡萄糖代谢的关键激酶，对维持肿瘤细胞的生长有重要作用[39]，抑制肿瘤细胞中高表达的HK2 能够对肿瘤进行有效的治疗[40]。胰岛素样生长因子结合蛋白 3（Insulin-like growth factor binding protein 3，IGFBP3）是高亲和力结合蛋白家族的成员，在调节胰岛素样生长因子（IGFs）的促有丝分裂和抗凋亡作用中发挥重要作用[41]。IGFBP3 在肿瘤组织中的过表达与肿瘤的增长、侵袭以及预后恶化有显著相关性[42]。能够通过增加鞘氨醇激酶-1（SphK1）的表达和鞘氨醇-1 磷酸（S1P）的形成来增强细胞生长，从而反式激活表皮生长因子受体（EGFR）或IGF1R信号通路，起到肿瘤启动子的作用[43]。通过阻断IGFBP3-AKT /MAPK-Snail通路能够有效抑制ccRCC细胞的增殖与侵袭[44,45]。低氧诱导因子-1A（hypoxia inducible factor-1A，HIF-1A）作为一种重要的氧感知调控因子，导致VEGF上调，可激活包括血管生成、葡萄糖代谢和侵袭等多种生物学过程关键基因与通路[46,47]。其调控葡萄糖转运蛋白（GLUT1）、己糖激酶II（HKII）、乳酸脱氢酶A（LDHA）和丙酮酸脱氢酶激酶 1（PDK1）来增强肿瘤细胞中的糖酵解[48]。

本研究中发现MAPK、AGE-RAGE、HIF-1、PI3K-Akt等信号通路与ccRCC糖酵解过程有密切关联。MAPK信号通路作为VEGF的相关下游通道，MAPK信号通路的抑制能够有效减少磷酸化的改变，同时能够抑制VEGF介导的肿瘤增殖[49]。瘤内缺氧和基因改变可导致HIF-1A过度表达，HIF-1A诱导的PDK1 可通过阻止丙酮酸进入三羧酸循环来调节葡萄糖氧化，并促进糖酵解[50]。当RAGE在与AGE结合后被激活时，AGE-RAGE轴在肿瘤生长和转移中起着关键作用，会诱导VEGF的激活，促进血管生长[51]。其靶向METTL3 通过VEGF介导的PI3K-Akt信号通路促进血管形成，能够影响下游的多种效应分子而发挥其抗凋亡作用的[52]。

综上所述，本研究基于网络药理学初步揭示了定喘宜肾汤在糖酵解途径治疗ccRCC的活性成分、作用靶点以及信号通路，初步发现定喘宜肾汤最重要的 3 个活性成分（木犀草素、槲皮素、鞣花酸）和 5 个重要靶点（VEGFA、EFGR、HK2、IGFBP3、HIF-1A），主要调控MAPK、AGE-RAGE、HIF-1、PI3K-Akt等信号通路。本研究日后需和更多的临床实验相结合，为进一步挖掘新功能及拓宽其应用提供研究思路。

PU-124

热点主题探测方法应用现状及在医学情报领域的实践

徐 源
上海市卫生和健康发展研究中心（上海市医学科学技术情报研究所）

随着医学信息资源在数量和种类上的急剧增长，学科间的交叉融合不断涌现，科研人员及情报分析人员在把握学科领域研究热点的发展变化时面临越来越多的挑战。主题探测（Topic Detection），作为一项旨在帮助人们应对信息过载问题的研究，对学科主题进行探测和识别，实现按主题查找、组织和利用来自多种媒体的多语言信息具有十分重要的现实意义，有助于提高及时把握科学研究动态的能力，快速追踪学科前沿，同时有助于应对信息过载问题，提高工作效率。本文基于热点主题探测相关研究的回顾，对采用的数据源、研究方法和工具进行系统梳理，探索主题探测在全领域和医学情报领域的应用现状，为热点主题探测在医学情报领域的进一步实践提供经验总结，助力新研究点的探寻。

一、研究方法

（一）数据来源

采用检索式 TI=('研究热点'+'研究前沿'+'新兴趋势')OR TI= ('研究主题'+'热点主题'+('主题'*'热点')) OR TI= ('主题探测'+('主题'*'探测'))在 CNKI 全部分类和“医药卫生科技”分类分别检索与热点主题探测相关的文献。对所检索到的文献进行筛选，排除编者按、投稿须知、征文通知等非学术论文以及“主题班会”、“主题党日”、“主题活动”等与研究主题无关的论文，最终在全库和医药卫生科技库分别筛选得到 5715 篇、771 篇文献。由于部分论文在关键词、摘要中并未明显地表现出数据类型、研究方法与工具，我们在数据清洗过程中人为标注数据类型、研究方法与工具，大部分沿用文献中的表述，只对不同表达方式进行归并，以便全面、客观地呈现所应用的研究方法与工具。

（二）分析方法

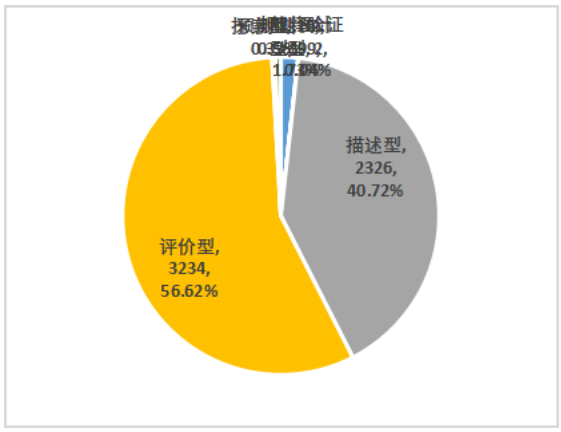
通过对文献内容的评阅、归纳整理和分析鉴别，梳理和分析热点主题探测相关情报研究的类型特征，及采用的数据类型、研究方法与工具，采用 Excel 软件进行描述统计分析。

二、研究结果

（一）热点主题探测情报研究类型分析

根据情报学研究的类型特征，将文献分为描述型、探索型、解释验证型、预测型、评价型、规划设计型 6 种类型（由于分类不是严格意义上排他的，在统计的过程中只选择最为主要的类型进行频次的计算）。当前全领域和医学领域热点主题探测相关文献的类型以评价型（56.62%，64.20%）和描述型（40.72%，34.89%）为主，存在一定比例的规划设计型（1.73%，0.78%），探索型、预测型和解释验证型研究较少，见图 1、2。

分别看期刊论文、学位论文和会议论文的研究类型分布，学位论文的研究类型最为丰富，且规划设计型、预测型和解释验证型研究的主要出现在学位论文，评价型研究的比例远高于其他两类论文，描述型研究则远低于其他两类论文，见图 3。



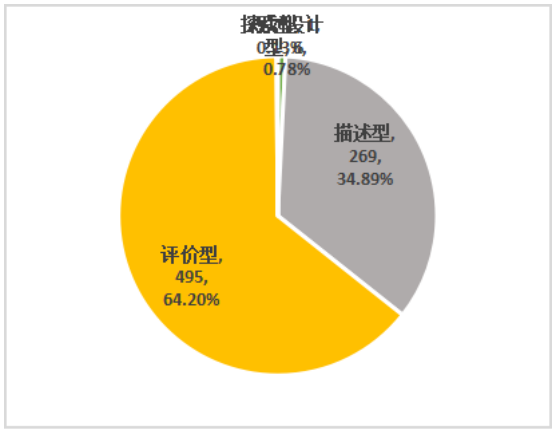


图 1 主题探测相关文献研究类型分布　图 2 医学领域主题探测相关文献类型分布

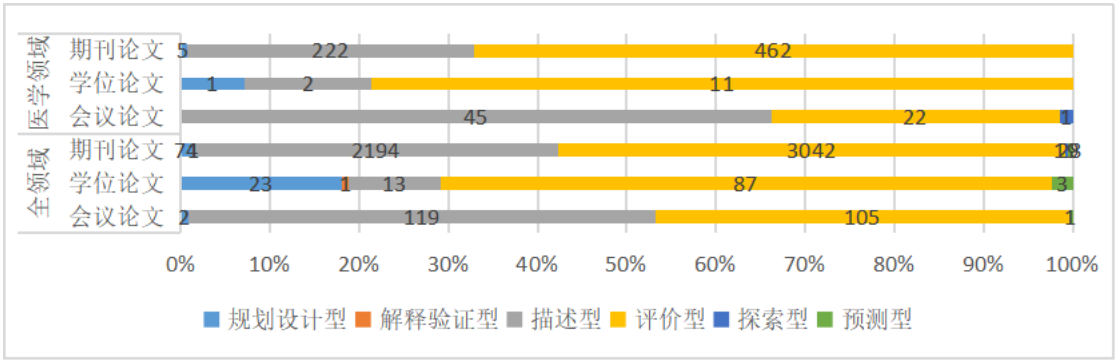


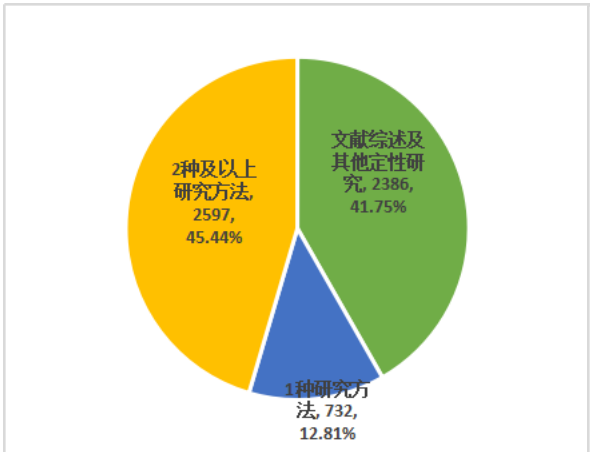
图 3 主题探测相关不同文献研究类型分布

（二）热点主题探测情报研究的常用研究方法和工具

全领域和医学领域主题探测相关研究中，分别有 35.28%和 45%的研究为文献综述及其他定性研究，可见医学领域相关研究中定量分析的比例高于全领域，其中运用了 2 种及以上方法的研究占比分别为 42%和 38%，比例远高于仅采用 1 种方法的研究（图 4、5）。

图 4 全领域主题探测研究方法分布　图 5 医学领域主题探测研究方法分布

具体对除文献综述及其他定性研究以外的其他研究方法进行分析发现，主题探测相关文献采用的研究方法具有多样性，包括统计分析方法、数学模型分析方法以及计算机辅助分析方法，研究方法使用过程中呈现较明显的集中与分散的趋势，少数的大众研究方法受到大多研究者的青睐，而大部分方法的使用频率很低（表 1），2010 年后研究方法丰富程度和使用频率快速增长（图 6）。



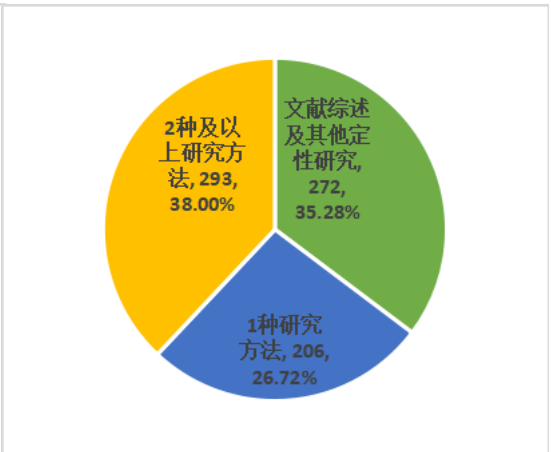


表 1 主题探测相关文献采用的研究方法统计

研究领域	研究方法（使用频数）
全领域	词频分析（1571）知识图谱（1482）描述性分析（1259）聚类分析（1053）共词分析（811）社会网络分析（573）多维尺度分析（408）引文分析（301）因子分析（151）可视化分析（150）共词聚类分析（133）战略坐标分析（114）突变检测（72）内容分析法（55）主题分析（35）双向聚类分析（34）文本挖掘（29）耦合分析（22）机器学习（21）自然语言法（10）回归分析（8）数据挖掘（7）社团检测（6）数理统计法（6）专利分析（6）相关性分析（5）关联分析（4）相似度分析（4）合著关系分析（3）效度分析（3）映射分析（3）比较分析（2）多元统计分析法（2）假设检验（2）曲线拟合（2）信息采集（2）综合指数（2）Lucene 算法（1）meta 分析（1）编辑距离算法（1）层次分析法（1）词袋模型分析（1）词重叠法（1）单因素分析（1）仿真分析（1）活跃度指数（1）计盒法（1）技术挖掘（1）空间测度法（1）灵敏度分析（1）谱分解法（1）区块分析（1）融合分析（1）时间序列预测法（1）时间因子（1）时序变化曲线（1）数据包络分析（1）数据流分析（1）双图叠加分析（1）思维导图（1）谓词树法（1）线性回归（1）遗传算法（1）重叠社群（1）最大概率法（1）最小生成树算法（1）
医学领域	聚类分析（258）共词分析（140）知识图谱（140）词频分析（70）社会网络分析（63）描述性分析（62）双向聚类分析（32）引文分析（29）战略坐标分析（23）可视化分析（14）多维尺度分析（10）主题分析（9）内容分析法（7）文本挖掘（4）回归分析（3）因子分析（3）突变检测（3）综合指数（2）PFNET 分析（2）词重叠法（1）单因素分析（1）合著网络分析（1）基于共词的逆文档频率量化法（1）假设检验（1）耦合分析（1）思维导图（1）文献耦合（1）相关性分析（1）纵向映射分析（1）热点区域分析（1）

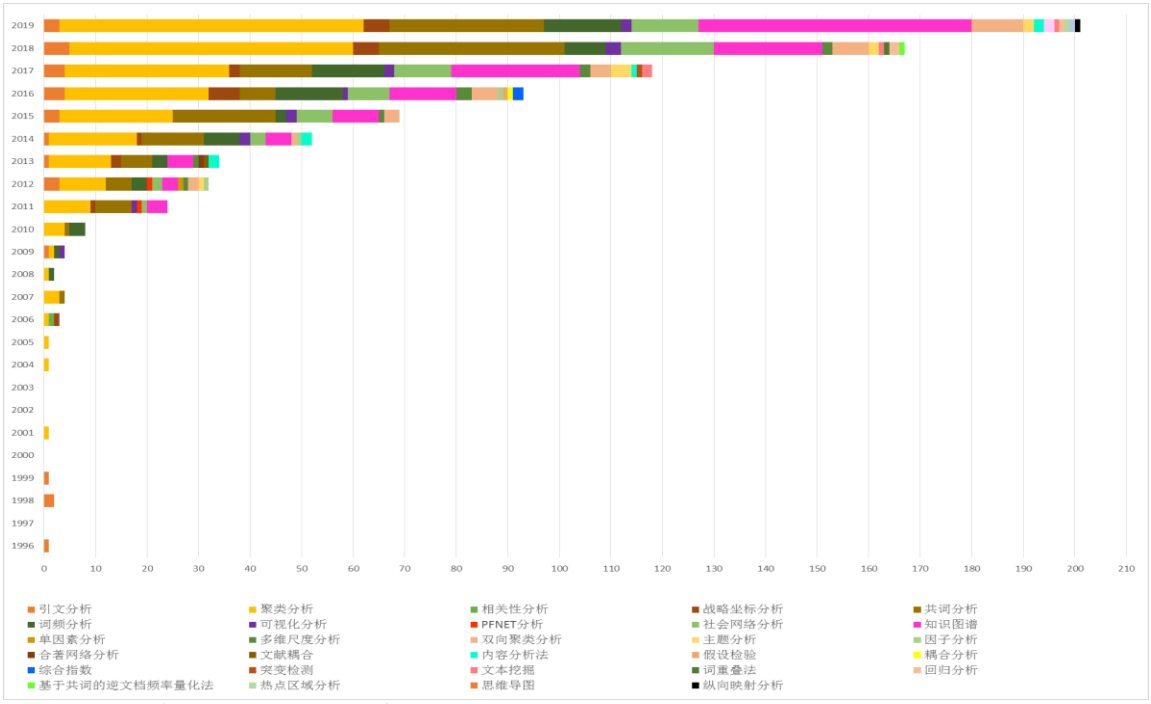


图 6 医学领域主题探测研究方法时间分布

（三）热点主题探测情报研究的常用数据来源

在全领域和医学领域有明确数据源的 3505 篇、503 篇文献中，分别仅有 9 篇、3 篇同时使用了多种数据来源。学术文献仍是最主要的数据来源，除此之外还包括国家自然科学基金、国家社科基金、国家哲学社会科学基金等一些大型基金的科研项目资助情况，微博、微信、Twitter、论坛等社交媒体，一些开放的网页信息，以及专利数据和杂志等作为数据源。

表 2 热点主题探测相关文献采用数据来源统计表

数据源	频次	
	全领域	医学领域
学术文献	3371	499
科研项目	89	3
网页信息	27	2
社交媒体	27	1
专利数据	7	1
杂志	1	0
药物研发实验	0	1
总计	**3522**	**507**

从时间趋势上来看，全领域采用学术文献作为数据源的研究最早从 1991 年开始，并逐年增加，从 2006 年起增速加快，2016 年起增速又达新的高度；2002 年、2008 年采用科研项目和采用互联网和网络社交媒体的信息作为数据源的研究开始出现，并逐年波动上升；2009 年采用专利数据作为数据源的研究开始出现，而后出现空白，2016 年开始重新开始发展（图 7）。医学领域 2016 年开始有研究以科研项目作和专利数据为数据源，2018 年网页信息开始被用来作为数据源。2019 年，李玲琍 等基于论文、专利、药物研发及临床实验等多源数据分析了造血干细胞领域发展趋势、研究主题、重点药物及重大病症。

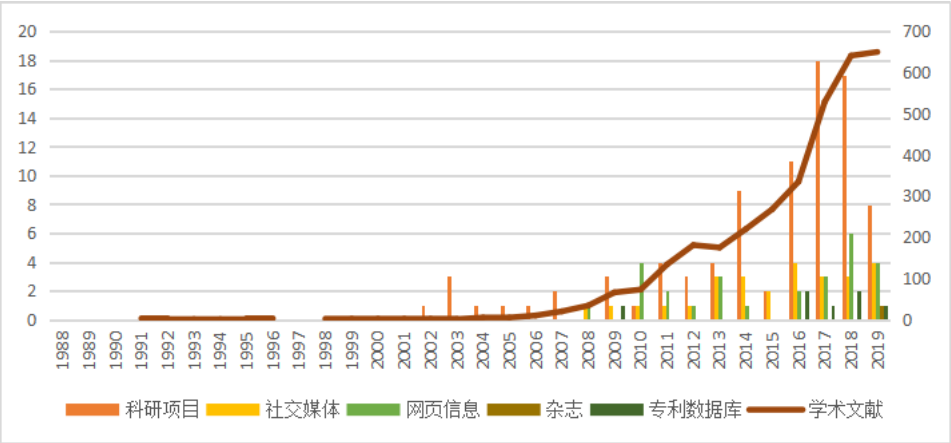


图 7 主题探测相关文献数据来源时间趋势

主题探测相关研究采用的研究工具可以归类为信息计量学应用软件，文献管理软件及文献/专利数据库分析功能，专门用于网络分析的软件，统计学及数学通用软件，服务器、计算机程序设计、数据库管理相关的计算机辅助软件，数据挖掘软件、文本挖掘软件等大数据分析软件，以及一些可视化绘图软件（表 3）。各种不同类型的研究工具都均与研究方法呈现类似的趋势，即少部分工具使用的频次较高，而大部分工具使用频次相对较低。

表 3 主题探测相关文献研究工具统计

软件类型	软件名称（频数）	
	全领域	医学领域
信息计量学应用软件	Citespace（1360）Ucinet（562）Bicomb（552）Netdraw（510）SATI（163）gCluto（118）VOSviewer（102）Bibexcel（92）BICOMS（46）HistCite（24）Gephi（22）ICTCLAS（17）SciMAT（10）Sci2（7）NodeXL（6）自编文献计量学统计分析软件（7）SCIVAL（4）NLPIR（3）BibStats（2）itNetExplorer（2）CitNetw（1）CO_OC（1）ctclas（1）Yacsi（1）生物医学文献共现分析系统（1）	BICOMB（179）Citespace（133）gCluto（101）Ucinet（66）Netdraw（45）VOSviewer（24）BICOMS（23）Thomson Data Analyzer（13）Bibexcel（11）SATI（10）Gephi（3）HistCite（3）自编文献计量学统计分析软件（2）Sci2（1）生物医学文献共现分析系统（1）引文分析软件（1）
文献数据库分析功能	CNKI（8）PubMed（4）INCITES（2）WoS 平台（1）超星知识发现系统（1）Elsevier Fingerprint Technology（1）PROQUEST（1）	PubMed（3）Sci MAT（2）
文献管理软件	NoteExpress（22）Endnote（16）Notefirst（1）Refviz（1）	NoteExpress（9）Endnote（3）
专利数据库分析功能	专利分析平台（5）IncoPat（1）INNOGRAPHY（1）	
网络分析软件	Pajek（43）Cytoscapc（1）Map Equation（1）PFNET（1）	Pajek（3）Cytoscapc（1）Node XL（1）The Knowledge Network Organizing Tool（1）
统计学及数学通用	SPSS（979）Excel（474）R 语言（24）Matlab（15）PASW（7）MySQL（4）Nvivo	SPSS（173）Excel（70）PASW（6）R 语言

软件	（4）Mathematica（3）LibSVM（1）SAS（1）	（5）MySQL（2）Matlab（1）SAS（1）
服务器软件	Apache（2）Ajax 异步传输技术（1）	
计算机程序设计软件	Python（58）Java（35）C 语言（6）VB（4）Regex Buddy（3）API（2）Eclipse（2）Fox base（2）JDK（2）GoWeb（1）MyEclipse（1）nodejs（1）PHP（1）Visual Studio（1）	java（9）Python（4）Regex Buddy（3）Fox base（2）API（1）
数据库管理软件	SQLServer（3）access（1）FoxPro2.5 for Windows（1）Hive（1）Oracle 数据库（1）SQLite（1）	
数据挖掘软件	TDA（39）Knime（10）BlueMC（2）SOMToolbox（2）Carrot（1）Cfinder（1）GOOSEEKER（1）Larbin（1）rapidminder（1）Weka（1）WordNetLemmatizer（1）网络神采（1）网络智能采集系统（1）新闻网页采集系统（1）智能网络主题搜索引擎（1）信息共现挖掘系统（1）战略情报分析平台（1）	
文本挖掘软件	ROST（27）JGibbLDA（5）MetaMap（5）GibbsLDA（4）Word Smith（4）Ant Conc（3）jieba（2）LDA 算法包（2）Alchemy API（1）ATLAS.ti（1）Google Cloud Natural Language API（1）Han LP（1）IKAnalyzer（1）Leximancer（1）LTP（1）Mallet（1）NLTK（1）PDFMiner（1）PLDA 软件（1）SemreX（1）TDT 系统（1）Topic Extract Tool（1）UMLS（1）VantagePoint（1）Word Stat（1）词频分析器（1）结巴分词系统（1）热点主题探测系统（1）图悦词频分析工具（1）主题模型工具包（1）stammer 分词系统（1）	MetaMap（5）ROST Content Mining（3）UMLS（2）信息共现挖掘系统（1）NLPIR 汉语分词系统（1）
可视化绘图软件	Origin（5）Mindmapper（4）Echarts（2）NEViewer（2）Simulink（2）Alluvial flow（1）ArcGIS（1）Graphpad Prism（1）agxedo creator（1）ThemeRiver（1）Word Art.com（1）	ArcGIS（1）Graphpad Prism（1）

三、讨论与结论

（一）热点主题探测的研究方法具有多样性，但仍处于探索阶段

热点主题探测的研究以描述型和评价型为主，预测型和探索型比较少，相比之下，解释验证型和规划设计型相对较多。学位论文研究类型丰富，且探索型、预测型和解释验证型研究的比例远高于期刊和会议论文。可见主题探测领域研究仍处于方法模型的摸索完善和应用探索阶段。热点主题探测相关研究除了采用情报研究常用方法外，还采用了统计分析方法、数学模型分析方法以及计算机辅助分析方法，构建模型、编写代码、开发程序，进行数据和文本挖掘。但受到大多研究者青睐的仍以传统的情报研究方法为主，其他方法使用频率很低。医学领域研究在方法的多样性和工具的丰富性上与全领域整体水平存在差距，特别是数据库管理软件、服务器软件、计算机程序设计软件等计算机辅助软件，以及数据挖掘和文本挖掘等大数据分析软件的应用。

（二）热点主题探测研究的数据来源有所丰富，但以文献为主有待多元化和综合化

虽然一些大型基金的科研项目资助情况、网络社交媒体信息作等等新兴数据源开始被采用，并不断波动发展，但学术文献是主题探测相关研究最早采用也始终是最主要的数据来源，并在在持续迅速增长，这一定程度上是由于情报学本身脱胎于文献学有关。不同类型的数据来源有着不同的特点，科技文献，专业性高但从研究到结果产出在到成文发表的过程使得科技文献的即时性不足；基金项目代表前瞻性、先导性和探索性的学科研发方向，与期刊论文、专利信息等数据源相比，能更早反映出学科领域未来发展方向与结构部署；互联网的产生与发展形成更多综合、即时和高效的知识信息，但同时，网络信息的鱼龙混杂，给主题探测提出了更高、更精和更专的要求。相比于研究方法的多样性，现有研究的数据来源仍显单一，且综合采用多种数据来源的研究较少。特别是医学领域，虽然定量研究的比例高于全领域的平均水平，但综合多种数据源的研究还是较少。

（三）大数据环境带来的机遇，不断丰富热点主题探测研究方法学和在医学情报领域的实践

大数据环境下人们得以以极低成本和极快速度，从超大体量、繁多结构的数据中采集、捕获、发现、分析和提取多元、丰富且有价值的数据，给主题探测情报工作带来极好的机遇。同时，基于大数据环境的工作流程与内容与传统情报工作极其相似，这也奠定了大数据利用大数据进行主题探测的可行性。不同类型的数据来源有着不同的优势和缺陷，也从不同角度来反映问题。情报研究的数据源本身就具有多样性，除了文献型信息源外还包括非记录形式的面谈、问卷、实地考察、小组座谈，组织方面的管理过程中形成的，以及公共渠道等获得的非文献信息源。除此之外，医药卫生领域有着丰富的特色数据来源，如实验数据、医学指南、卫生政策文件、各种量表和指标体系等等。随着数据来源的多元化以及数字化科研环境的形成，热点主题探测不应局限于文献和传统情报研究方法，而要在方法学层面进行探索和丰富。未来热点主题探测在医学情报领域的研究实践应首先明确医学领域可获得和可利用的数据源及其特征，选择合适的研究方法，综合使用多元数据，形成互补，以保证主题探测的全面性和科学性。满足“多样化、高效化、个性化、专深化”卫生信息需求升级，为管理决策、技术预研、产品研制、条件建设等提供技术引领和决策支撑，进而凸显情报研究的作用和价值的需要。

PU-125

文献计量视角下国家谈判药品落地政策分析

强滟

山东第一医科大学医疗保障学院

目的 基于政策工具视角对国家谈判药品落地的相关政策文本进行分析，为后续政策制定和优化提供参考。

方法 从政策工具和相关利益者视角构建二维分析框架，对国家谈判药品政策文本进行内容分析。

结果 纳入分析的 18 份政策文件共形成 120 条政策编码；政策工具维度中，命令与规制、权威、能力建设、激励、劝告 5 类政策工具使用频次占比分别为 60.83%、12.50%、10.83%、8.33%和 7.50%；利益相关者维度中，从政策工具的绝对数量来看，政府相关部门最高，医疗机构在共现网络中权重最大，几乎所有的政策工具都有涉及。

结论 持续优化政策工具内部结构，提高激励性政策工具的使用比例，促进国家谈判药品政策的落地。

PU-126

基于 CiteSpace 的指纹识别技术发展应用趋势分析

何阿妹[1] 王思成[2] 杨山石[1] 徐源[1] 宋捷[1] 沙小苹[1]
1 上海市卫生和健康发展研究中心（上海市医学科学技术情报研究所）
2 华中科技大学医药卫生管理学院

一、引言

社会和经济的发展使人们对身份鉴别的准确性、安全性以及实用性提出了更高的要求。基于信物或口令的传统方式存在安全隐患，如容易丢失、遗忘、被复制及盗用，但通过辨识人的生理和行为特征进行身份认证的生物识别技术可改善这种情况[1]。指纹识别技术是一项集计算机、网络、光电技术、图像处理、智能卡、数据库技术于一体的综合技术[2]，利用人体所固有的生理特征或行为特征来进行个人身份鉴定。指纹中有许多特征点，特征点提供了唯一的辨认信息，这是指纹识别的基础，分为总体和局部特征，总体特征包括核心点、三角点、纹数，局部特征是指纹的细节特征，在特征点处的方向、曲率、节点的位置，这都是区分不同指纹的重要指标[3]。指纹识别技术的应用在生活中已随处可见，如银行确认系统、手机解锁、门禁、考勤机等等[4]。我国已进入信息化高速发展阶段，生物特征识别（包括个人面部识别特征、虹膜、指纹、声纹等）所具有的便捷与安全等优点以及其在身份认证识别和网络安全领域所发挥的作用，国家在政策和标准制定上一直持续支持生物特征识别技术发挥其作用，并建立健全生物识别技术相关法律法规、制度体系和标准规则，这也是推动生物识别技术应用健康发展的重要保障。2020 年 3 月，国家标准化管理委员会出台《信息安全技术基于生物特征识别的移动智能终端身份鉴别技术框架》，新国标对指纹、面部识别等生物特征识别技术制定了统一标准规范，适用于生物特征识别身份鉴别服务与产品的开发、测试和评估等。为保护生物特征识别信息安全，规范生物特征识别信息处理活动，2021 年 10 月，国家市场监督管理总局、国家标准化管理委员会颁发《信息安全技术 生物特征识别信息保护基本要求》，并 2022 年 5 月 1 日正式实施。

2016 年 10 月，中共中央、国务院印发“健康中国 2030”规划纲要，提出要建设健康信息化服务体系，“智慧医疗”逐渐成为医院信息化建设的趋势，即将智能技术广泛应用于医院各个科室和部门。目前指纹识别技术在医疗卫生领域也有相关应用，包括考勤管理、电子处方开具系统、贵重药品存放门禁系统、患者识别、手术室等。指纹识别技术在医院应用，可以优化各科室工作流程，提高医院精细化管理水平、减少人力成本的支出，提高工作效率，通过无纸化节约经费，有利于智慧医院的建设与患者管理，增加经济效益。本研究旨在探究指纹识别技术应用现状及未来发展趋势，并分析该技术在生物医药领域应用情况，为指纹识别技术相关产业更好的发展提供参考。

二、基于知识图谱的国内外指纹识别技术应用可视化趋势分析

（一）数据来源与检索策略

选择中国知网（CNKI）的“中国期刊全文数据库”和 Web of Science 核心合集作为文献来源，时间均限定为“2012-2022”（数据检索时间为 2022 年 8 月 20 日）。在 CNKI 检索中选择文献类型为“期刊”，检索途径为“专业检索”，检索条件：SU='指纹识别'*('技术' + '应用' + '系统')，检索出相关原始文献 1682 篇。在 Web of Science 检索中选择文献类型为“论文”、“会议录论文”、“综述论文”，检索条件：TS=(Application *(' fingerprint recognition ' or ' fingerprint identification '))，检索出相关原始文献 2103 篇。去除重复文献后，使用 CiteSpace 科学文献分析可视化软件，通过选择不同的网络节点类型，分析文献共现热点呈现、主题剖析以及时间脉络演进趋势，对国内外指纹识别技术现状进行分析及预判。

（二）研究结果

1.学科分布情况

国内指纹识别技术相关研究的文献学科分布，如图 1 所示。国内指纹识别技术研究主要集中在计算机软硬件、自动化技术、公安、电信技术等学科，占比超过 70%，研究大多为技术研究与技术开发。同时也可以看出，指纹识别技术应用研究中与医疗卫生和医学相关的研究较少，约为 0.73%。

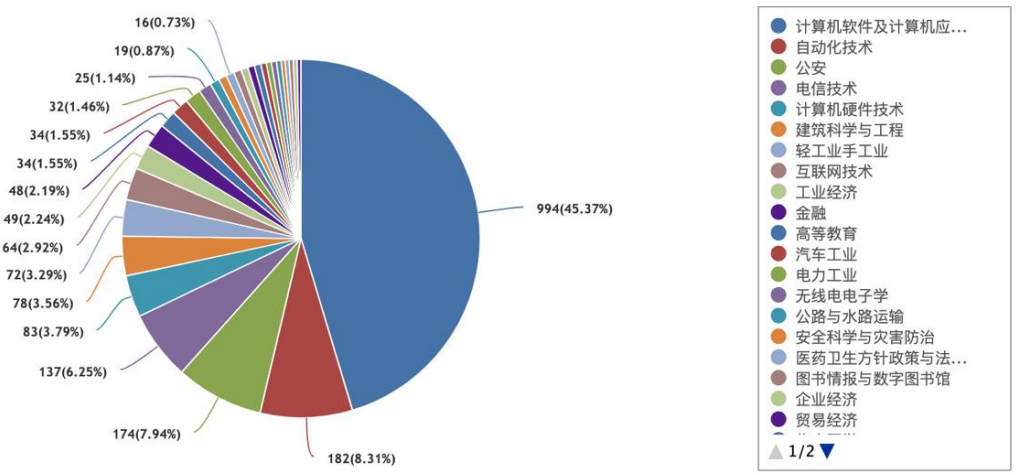


图 1 国内研究学科分布图

将学科限定为医药卫生后，主题主要与患者身份识别、医院信息管理相关，如图 2 所示。结果表明指纹识别技术在医疗健康情境下的研究主体主要集中在医院管理的各个方面，与身份识别和安全相关的问题是医疗健康情境下需侧重关注的内容。

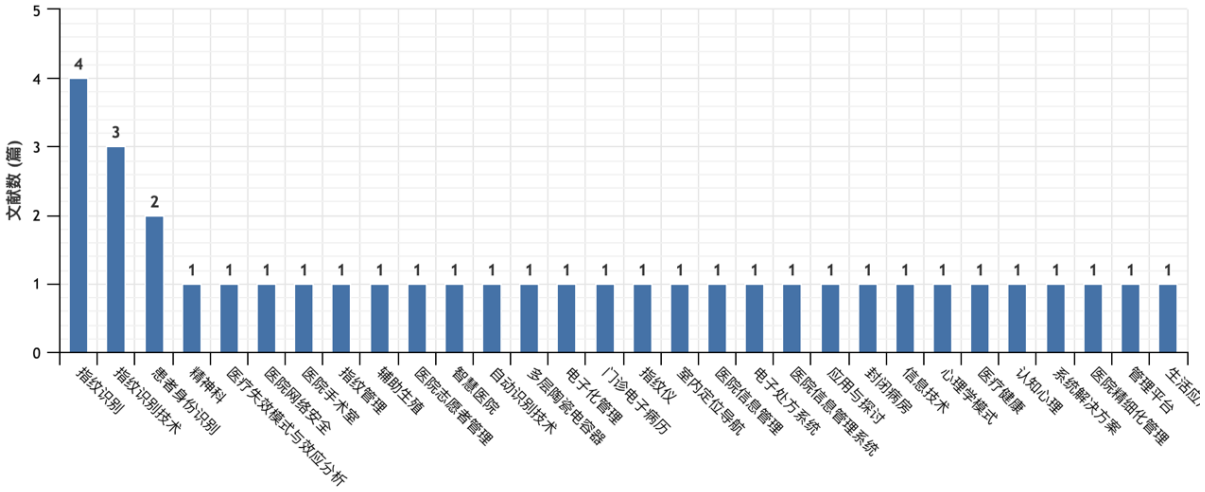


图 2 医药卫生—主要主题分布图

国外指纹识别技术相关研究的文献学科分布，如图 3 所示。国外指纹识别技术研究主要集中在计算机科学技术、电子工程、生命科学与生物医学等学科，占比超过 80%。其中，与医疗卫生关联较大的生物医学研究占比约为 23%，但其研究多为指纹图谱技术在生物医学领域和药理学领域的使用，涉及指纹识别技术的文章较少。

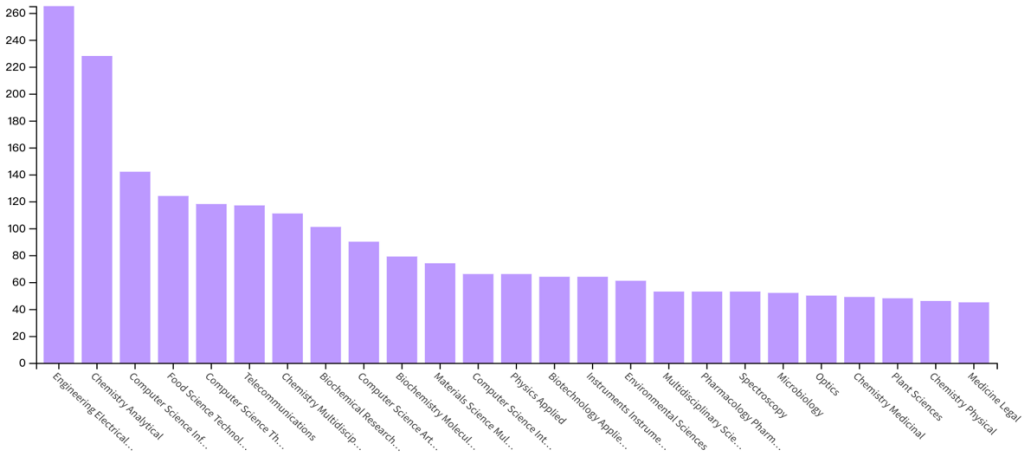
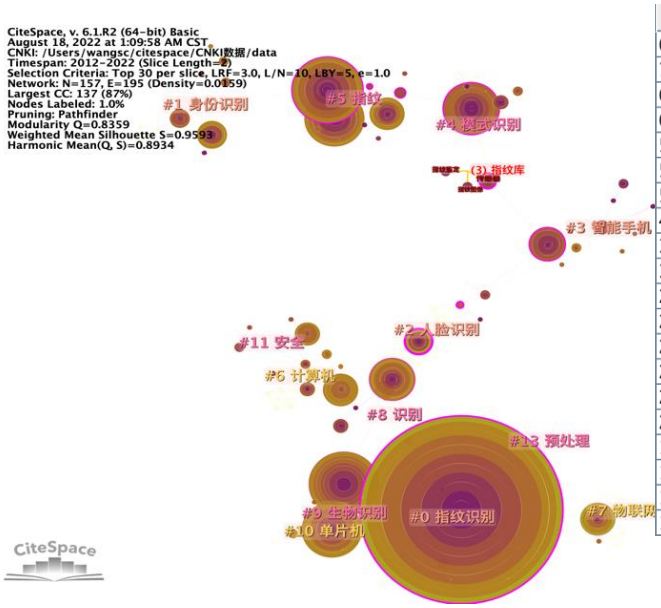


图 3 国外研究学科分布图

2.关键词共现热点呈现分析

根据 CNKI 得到的关键词共现热点分布图谱见图 4，可以看出与指纹识别技术与应用相关的文章最常使用的关键词为“指纹识别”、“身份识别”、“特征提取”等，大部分高频关键词的中介中心性都大于 0.1，说明其重要性较为突出。出现最多的关键词“指纹识别”与“特征提取”、“传感器”、“射频指纹”等关键词有紧密联系，是我国指纹识别技术的主要研究内容。通过对关键词的合并处理，可以看出不同时期我国指纹识别技术与应用研究的相关热点，聚类分析发现研究内容主要集中在主体识别和安全技术两个大类，2012 年学者们主要围绕指纹匹配、指纹采集等大方向展开研究，2013 年关注图像处理、生物特征的研究，2016 年后更加关注指纹识别技术的实际应用，转向对安全监控、支付系统、指纹分类比对效率、考勤管理进行研究，2019 年后对技术和应用的研究更加细化，对控制系统、支付技术、实时认证、智能家居的研究较多，2022 年后学者们更加注重指纹识

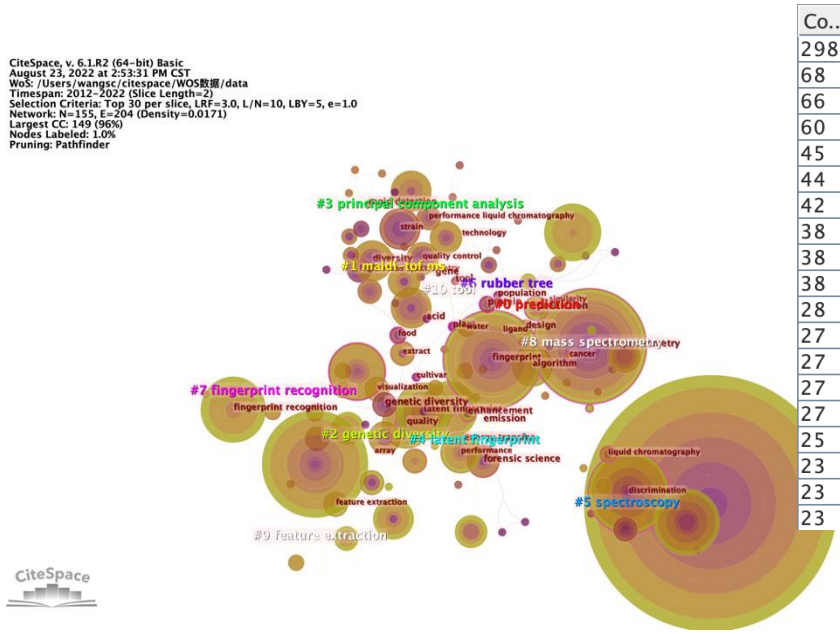


Co...	Centr...	Year	Keywords
676	0.38	2012	指纹识别
75	0.64	2012	指纹
65	0.21	2012	生物识别
63	0.11	2012	单片机
56	0.21	2012	应用
53	0.35	2012	人脸识别
52	0.38	2012	身份认证
40	0.11	2012	门禁系统
38	0.06	2013	智能手机
34	0.13	2017	人工智能
29	0.08	2012	身份识别
23	0.46	2012	物联网
23	0.14	2012	特征提取
23	0.01	2012	图像处理
22	0.14	2016	计算机
22	0.05	2013	信息安全
19	0.30	2012	数据库
18	0.15	2012	识别
16	0.23	2015	考勤
16	0.11	2013	传感器

别技术的智能化应用，包括多模态、系统设计矩阵和关键词智能安防、机智云、智慧管理等。

图 4 国内研究关键词共现网络

根据 Web of Science 核心合集得到的国外研究关键词聚类与共现热点分布图谱见图 5，文章最常使用的关键词为“identification（识别）”、“classification（分类）”等，国外指纹识别主要围绕识别技术展开，与国内侧重点不同在于，指纹只是多模态识别技术中的一环，没有特别强中心性的方向。经过关键词聚类分析可以看出，在 2012—2022 年期间，与“识别技术”相关的关键词贯穿于多个主题，研究方向发散性很强，更注重底层技术的拓展和延伸。



Co...	Cen...	Year	Keywords
298	0.05	2012	identification
68	0.09	2012	classification
66	0.24	2012	mass spectrometry
60	0.27	2012	fingerprint
45	0.02	2012	nanoparticle
44	0.22	2012	discrimination
42	0.10	2014	spectroscopy
38	0.10	2018	fingerprint recogn...
38	0.29	2012	recognition
38	0.00	2018	machine learning
28	0.00	2012	system
27	0.03	2012	principal compon...
27	0.11	2012	algorithm
27	0.15	2012	latent fingerprint
27	0.08	2012	diversity
25	0.12	2013	quality
23	0.11	2012	strain
23	0.15	2016	performance
23	0.08	2015	quantification

图 5 国外研究关键词共现网络

3.关键词突现分析

国内指纹识别技术与应用研究的关键词突现图见图 6，Year 为关键词首次出现的年份，Strength 为突现强度，右边的红条表示关键词突现的持续时间。从图中可以看出指纹鉴定和指纹图像突现时间早、突现程度高，是国内早期指纹识别的核心研究主题；随着移动互联时代的到来，智能手机、考勤、移动支付突现时间长、突现程度高，成为指纹识别技术应用的主要领域；当前技术应用开始向智能化发展，转向矩阵键盘、智能门锁等方面，指纹识别技术在不断迭代更新，由指纹图像的表层特征研究转向模式识别和人工智能的系统设计。

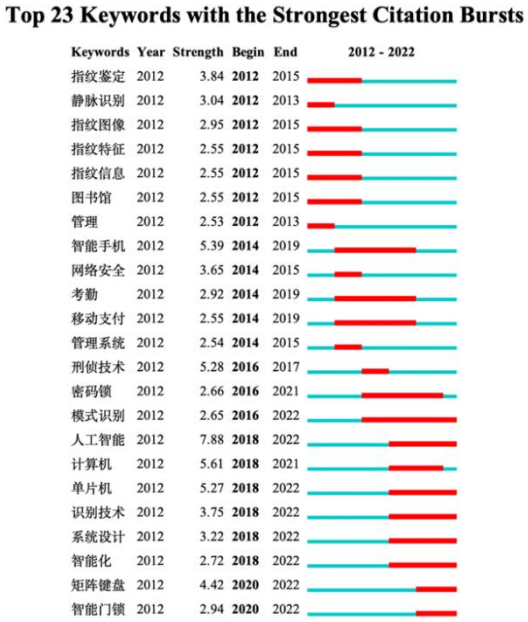


图 6 国内研究关键词引用突现

国外指纹识别技术与应用研究的关键词突现图见图 7，从图中可以看出指纹识别技术随着时间推进研究关注的焦点从图像逐渐向微观发展，随着智能化技术的推广，人工智能加入到指纹识别技中，对技术进行了扩充，“deep learning（深度学习）”于 2020 年开始突现并保持至今，且具有较高的突现程度，表明其为近年来学者较为关注的内容。

Top 13 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2012 - 2022
flight mass spectrometry	2012	4.2	**2012**	2015	
sequence	2012	3.68	**2012**	2015	
dna	2012	4.65	**2014**	2017	
performance liquid chromatography	2012	4.55	**2014**	2019	
bacteria	2012	3.7	**2014**	2017	
machine learning	2012	7.65	**2018**	2022	
nanoparticle	2012	5.43	**2018**	2022	
security	2012	3.75	**2018**	2019	
prediction	2012	3.71	**2018**	2019	
fingerprint recognition	2012	7.44	**2020**	2022	
feature extraction	2012	5.47	**2020**	2022	
deep learning	2012	4.86	**2020**	2022	
photoluminescence	2012	4.35	**2020**	2022	

图 7 国外研究关键词引用突现

三、讨论及趋势预判

指纹识别技术仍然是现阶段研究的热点，不同时期研究的方向随着信息化更迭而转变。国内外指纹识别技术发展轨迹大致相同，都是朝着智能化方向发展，国外研究相比而言，更加深入和底层，涉及的技术和应用范围相比更宽泛。根据以上研究结果，可以看出在指纹识别技术与应用的研究中，与医疗健康领域相关的研究占比较少，针对性的研究较为缺乏。而现有的在医疗健康领域应用指纹识别技术的应用能够改善医疗质量，也能提升医疗流程和管理效率。因此，加强医疗健康相关的指纹识别技术应用研究很有必要，使更多的医院能够根据自身的流程和管理要求把指纹识别技术融入各个环节中去，提升现代化医院管理水平。

随着信息技术的不断发展，智新型指纹图像采集技术和传感器不断涌现，可以预见指纹识别的应用场合将不断增加，指纹识别在以下四个领域将具有广阔的市场前景。

一是活体指纹识别，指纹采集仪很容易被精心伪造或复制的指纹样本欺骗，故指纹识别设备对于活体识别的需求显著提升。活体指纹识别是以电容传感为原理基础，电子信号能够感应识别对象的生物特征，利用人体真皮组织的电特性，获取真人手指持续有效的指纹特征值数据的一门新兴技术，只对真人活体指纹产生识别反应，对其他一切物质不作识别[5]。

二是多重特征融合，既多模态生物特征识别，除易被盗用问题外，特殊状态指纹（如手指潮湿、受伤等）的识别问题也亟需解决。单一生物识别无法支撑越来越复杂化、多样化的身份验证场景，结合其他生物特征（如虹膜、视网膜血管网络等）的多种生物识别技术，能根据不同的应用需求和场景变化，选择合适的结合方式和权重决策，这将成为身份识别与认证领域的发展趋势。

三是可穿戴式产品，近两年随着物联网技术的突飞猛进，可穿戴技术迅速崛起，由于可穿戴设备具有“贴身”特性，而新型传感器可以捕捉并生成更为复杂、可靠且难以篡改的生物认证身份，生物认证将被更广泛地应用于可穿戴设备。

四是人工智能模式识别，自动化指纹识别是人工智能技术在图像识别领域的一个具体应用，具有一定的易用性，精确度较为良好，容易被用户接受，而且稳定性也很高，和其他的人体生物特征的结合分析的效果也很高，具备直接性特点、友好性特点、便利性特点[6]。

PU-127

医学信息多维服务系统集群建设实践及研究

段永璇* 甄天民 宋燕
山东第一医科大学，山东省医药卫生科技信息研究所

目的 医学信息是多学科、多领域深度交叉和融合的研究，医学信息服务的多样性和专业性导致服务形式的复杂性，涉及多类系统及数据服务的搭建和深度整合，单一系统或平台较难实现医学信息系统建设的体 系化以及服务的全面化。本文通过医学信息多维服务系统集群建设实践及研究，从强化群体性宏观决策服 务的前瞻性和预测性，以及个体性微观决策服务的系统性和精准性等维度，系统性介绍了系统集群的建设 模式和各平台特点，为医学信息服务提供平台建设及发展思路。

方法 “医学信息多维服务系统集群”项目是在国家健康医疗大数据建设战略指导下，分析医学信息大数据建设及服务的应用要求，结合医学科研活动的实际需要，以及医学信息学的发展趋势，以服务医学科研群体为宗旨，以建设医学信息综合服务平台为目标，以提供优质医学信息服务为核心，开展与医学相关的科技信息数据及综合服务体系及平台建设研究。利用计算机软硬件、大数据、人工智能等技术，联结医学领域研究与科研人员的实际需求，将科研信 息、文献检索、咨询服务、信息推送、数据处理等信息服务模式有效的结合起来，建设集数 据汇聚、管理共享、信息推送、综合展示、知识服务为一体的智能化医学信息综合服务平台， 从而实现医学信息情报的融合汇聚、高效利用、规范管理，逐步形成符合实际的医学信息数据知识网络以及大数据服务体系。

结果 本项目以山东省医学科学院创新工程建设为契机，以提升医学信息服务及创新团队建设 为目标，自 2018 年起，经过 4 年的持续研发及建设，开发医学信息服务子系统 10 余个，包括医学信息科研云平台、医学学术信息推送系统、医药卫生信息动态实时平台、信息舆情分 析系统、医学信息知识共享空间、机构库系统、医药卫生决策平台、医药卫生学科评价平台、 医药卫生成果展示系统、医学信息数据可视化系统等，建成具备医学信息实时搜集、数据处 理、知识应用、预测分析、精准服务、决策支持、共享展示等多项服务能力的医学信息及数 据服务的系统集群。

结论 通过医学信息多维服务系统集群项目的建设与研发，搭建了专业度高、信息量大、针对 性强、学术氛围浓厚的医学信息服务平台，对科研、教育、基层等医疗机构科研活动的开展 具有指导意义，在加强医学信息和大数据综合利用，提高医学情报管理水平和工作效率，培养医学高端人才和打造科研创新团队，以及支撑健康山东建设等方面具有积极作用。借助平台，研究团队已开展多项发展战略研究和调研活动，涵盖山东省医养健康发展、医疗保障制度改革、管理标准及重大专项等方面，现已依托集群产出多项研究成果，包括咨询报告、实 施建议及调研报告等，取得较好的经济及社会效益。

PU-128

基本心理需要理论下微博用户信息行为的人格特质预测研究

刘帅兵, 张云秋
吉林大学公共卫生学院

前言

人格是人类不同个体特征的高度概括，一个人的人格是其展现的特质和风格，心理学上常使用人格来定义人们的个性，解释人们的行为和偏好[1]。由于人格不同，即便在同样的环境中，不同的人也会采取不同的行为。伴随着网络实名制的推进，用户的网络行为和现实行为具备较强的一致性[2]。利用社交网络用户在网络中留下的数据来预测用户的人格特征，在个性化服务、舆情管控、商品推荐等方面都有着重要意义。微博作为近些年快速发展的社交网络平台，微博社交网络中包括大量的用户行为信息，这些信息不仅直观呈现了微博用户的行为特征，还反映出用户当时的心理活动和性格特征。

当前，随着新型信息技术的快速发展和社交网络的大范围应用，利用社交网络进行用户人格预测已经成为研究热点。但现有研究多是针对预测方法及预测效果进行优化，如 Globeck 等[3]为精准衡量用户的人格特质，通过采集 Twitter 上用户公开的个人信息进行分析和建模，采用机器学习方法预测用户的人格，预测结果为实际值的 11%-18%；Oberlander 等[4]利用个人网络日志中不同的 n-grams 集合作为特征，以朴素贝叶斯为学习算法，发现二值分类和自动特征选择返回最佳分类结果；Bai Shuotian 等[5]从新浪微博中提取了静态特征和基于时间窗格统计的动态特征，并使用高斯核函数的 SVM 模型和 Pace 回归模型预测人格评分，认为大五人格与状态数、域名长度、粉丝数、互粉粉丝比相关；杨洁等[6]基于用户情感和网络关系构建人格预测模型，考虑到大五人格特质之间的相关性影响，采用了多任务学习方法进行建模预测，并与单任务学习算法进行比较，结果显示多任务学习方法优于单任务学习效果；Suman Maloji 等[7]进行了支持向量机、随机森林、朴素贝叶斯和逻辑回归四种机器学习模型对人格预测的比较，支持向量机凭借其分类机制获得了 88%的准确率；Hans Christian 等[8]引入了一种基于文本的人格预测方法，该方法使用多模型深度学习体系结构，结合 BERT、RoBERTa 和 XLNet 等多个预训练语言模型，采用基于模型平均的决策方法进行预测。

已有文献为本研究提供了重要的理论基础与经验证据，但仍有一些问题值得关注与思考。总体上看，现有人格预测研究大多集中于对人格预测方法及预测效果的优化上，缺少对于内在机制的讨论，即行为是如何对人格产生预测作用的。因此，本研究针对微博用户，基于基本心理需要理论，引入情绪心理指标作为媒介，探讨用户行为特征对人格特质产生预测作用的潜在机理。

1 理论模型与假设

1.1 理论基础

1.1.1 基本心理需要理论

基本心理需要理论来源于 Deci 和 Ryan 所提出的自我决定理论(Self-Determination Theory, SDT)，是 SDT 的核心概念与重要的理论分支。个体的基本心理需要由胜任需要(competence，也译作能力需要)、关系需要(relatedness，也译作关联需要)和自主需要(autonomy)构成[9]。人类的行为是在人们内在心理需要得到满足的基础上而形成的，心理需要的满足能助力个体心理的成长、发育、成熟和完善，是个体行为的基本动力[10]。只有人类的三大基本心理需要得到满足，才能实现自我激励和体验幸福[11]。

1.1.2 大五人格理论

大五人格特质是个体在不同时间或不同情境下表现出的某种特质，在以往对大五人格的研究中

我们不难发现，大五人格同样与人类心理及情绪之间存在关联。

大五人格将人格特质划分为宜人性（Agreeableness）、开放性（Openness）、外倾性（Extraversion）、尽责性（Conscientiousness）和神经质（Neuroticism）5 个维度[12]。外倾性是评定内外向性格的指标；宜人性是评判能否与他人和睦相处、团结协作的指标；尽责性是用于衡量责任感，对工作的敬业以及认真程度的指标；神经质是评测情绪化程度的指标；开放性则用来反映个体对经验的开放性、智慧和创造性。大五人格模型在德国、荷兰、西班牙、匈牙利等国都得到了验证，获得了大量跨文化研究的支持，因此本研究采用其探讨微博用户行为对人格特质的预测机制。

1.1.3 情绪心理指标

情绪心理指标是指能够侧面反映人心理情绪的指标，本研究中纳入主观幸福感、生活满意度以及生命意义感作为研究的情绪心理指标。主观幸福感指的是个体基于自身要求对生活质量的整体性评估，是衡量个人生活质量的综合性心理指标，反映主体的社会功能与适应状态[13]。生活满意度是指个体根据自己的选择标准对自己大部分时间或持续一定时期生活状况的总体性认知评估[14]，是主观幸福感重要的认知成分[15]。生命意义感是指个体感受、领会或理解自己的生命意义、生命目的或人生使命的程度[16]，生命意义感不仅可以帮助个体形成良好的人际关系以维护身心健康和提升幸福感，还可以引发个体的积极应对态度、修复心理创伤并促进个体心理成长[17]。缺乏生命意义感的个体容易产生消极情绪，陷入焦虑、抑郁、自杀意念等不利处境，甚至会产生自杀、反社会等问题行为[18]；拥有较高生命意义感的个体往往具备较好的心理应对能力，在应对不利处境时自我效能感较好、态度积极，较少体验到焦虑、抑郁等消极情绪[19]。

1.2 理论关系

1.2.1 基本心理需要理论与情绪心理指标的关系

基本心理需要的满足，会促使个体做出更多具有目标指向性的行为，从而使自己的生命充满意义感[20-21]。有大量研究基本心理需要理论在健康领域探索的应用，结果表明个体通过对胜任、自主、关系三种需要的满足可以获得幸福感与满足感，进而表现出积极的行为[22-23]。同时，基本心理需要对于精神上的健康及良好关系的建立起着积极的预测作用，当个体的基本心理需要得到满足时，个体的心理健康水平会更高，也更容易与他人建立良好的关系[24]。相应地，在基本心理需要无法被满足时，其与自杀、抑郁等问题行为存在着负向相关，会直接或间接地导致躯体化疾病、焦虑、压抑等等[25-27]。综上可以看出基本心理需要的满足与否与人的情绪心理息息相关，研究认为当用户满足基本心理需要时会导致情绪心理朝向积极的方向发展。

1.2.2 大五人格与情绪心理指标的关系

现有研究表明：外倾性、宜人性及尽责性与主观幸福感、生命意义感、生活满意度与积极情绪显著正向相关[28-35]；神经质与积极情绪、生活满意度、生命意义感及主观幸福感显著负相关，与消极情绪呈现正相关[28-35]；开放性作为评判创造力及想象力的指标，其与心理和情绪的关系在不同的情境下结果不一，没有证据证明其与情绪心理指标之间的关系[32,35]，因此，在实际研究中单独对其进行假设及验证。基于此，本研究将微博用户人格分为三类，即正向人格、负向人格及中性人格。其中，正向人格包括外倾性、宜人性及尽责性，负向人格为神经质，中性人格为开放性。研究所定义的正负向并不存在褒贬之分，仅仅基于已有研究与情绪心理指标的相关性分析结果，按其情感倾向划分进相应类别之中。

情绪心理指标作为与基本心理需要和大五人格同时存在相关性的指标，为微博用户行为特征对人格特质的预测构建了桥梁，将行为与人格联系起来。基本心理需要的三种需要互相补充而非互相排斥，一种需要的满足一定程度上也会带来其他需要的满足[36]。本研究认为当微博用户行为满足其中一种心理需要或多种需要时，这一行为指标对用户人格存在预测性，具体表现为对正向人格存在正向预测作用，对负向人格存在负向预测作用，对中性人格存在预测作用，但不同指标的预测倾向存在差异。本文据此提出研究假设及各假设研究模型。

1.3 研究假设

基于前文对理论基础及理论关系的梳理，研究认为当微博用户的行为特征满足其基本心理需要时，会对他们不同倾向的人格特质产生相应的预测作用，故本文提出的假设如下。

1.3.1 满足胜任需要的行为特征对人格特质的预测

胜任需要指个体预期有能力实现目标，成功完成挑战任务后所体验到的一种价值感和认同感，反映对所处环境的掌控能力。这一指标与微博用户关注认识的人/关注数、发布微博文本中情感倾向为正的微博占比、含图片帖子比相对应。

微博用户关注认识的人/关注数是指在用户关注的人中，现实中认识的人所占比例，这一指标值越高，证明用户倾向关注他们在现实中认识的人，由此看出这类用户对微博掌控能力越强。因此本研究提出假设

H1a：关注认识的人/关注数对用户正向人格存在正向预测作用；

H1b：关注认识的人/关注数对用户负向人格存在负向预测作用；

H1c：关注认识的人/关注数对用户中性人格存在预测作用。

发布微博文本中情感倾向为正的微博占比反映微博用户在微博平台表达的情感偏向于正向情感，由于这一指标越高反映用户对微博这一环境的掌控感和舒适度越高，满足了用户的胜任需要，因此提出假设

H1d：发布微博文本中情感倾向为正的微博占比对用户正向人格存在正向预测作用；

H1e：发布微博文本中情感倾向为正的微博占比对用户负向人格存在负向预测作用；

H1f：发布微博文本中情感倾向为正的微博占比对用户中性人格存在预测作用。

含图片帖子比是微博用户发布的帖子中含图片的帖子占比情况，能在微博中发布含图片的帖子体现微博用户对微博的发文功能充分了解，并对微博环境有一定的适应性，研究认为这一指标值越高越满足微博用户的胜任需要，基于此提出假设

H1g：含图片帖子比对用户正向人格存在正向预测作用；

H1h：含图片帖子比对用户负向人格存在负向预测作用；

H1i：含图片帖子比对用户中性人格存在预测作用。

图 1 所示为满足胜任需要的行为指标对人格预测作用的研究模型图。

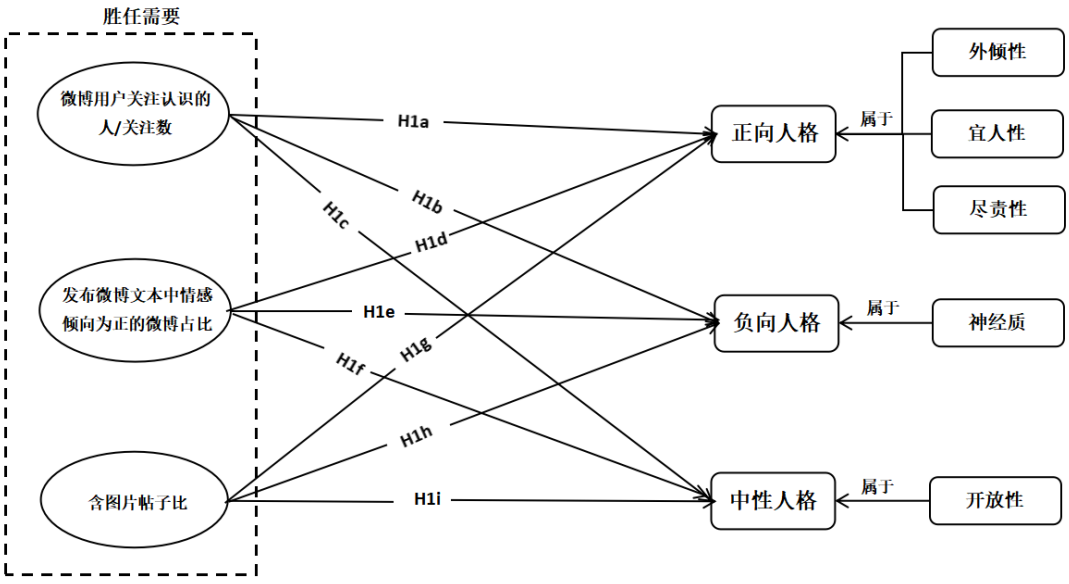


图 1 满足胜任需要的行为指标研究模型

1.3.2 满足自主需要的行为特征对人格特质的预测

自主需要指个体行动出于自由意志，而非外在逼迫，对生活有足够的选择权并能从心理上体验

到自由。与微博用户的原创微博数、原创微博占比相对应。

微博用户的原创微博数是用户发布的原创微博数量。原创微博作为用户个人的首创微博文本，能记录用户的个人经历和情绪，发布数量多是用户自主需要得到满足的体现，基于此提出

假设 H2a：原创微博数对用户正向人格存在正向预测作用；

H2b：原创微博数对用户负向人格存在负向预测作用；

H2c：原创微博数对用户中性人格存在预测作用。

微博用户的原创微博占比是用户发布的原创微博数量占全部微博数量的比例，与原创微博数相同都反映了自主需要的满足，基于此提出

H2d：原创微博占比对用户正向人格存在正向预测作用；

H2e：原创微博占比对用户正负向人格存在负向预测作用；

H2f：原创微博占比对用户中性人格存在预测作用。

图 2 所示为满足自主需要的行为指标对人格预测作用的研究模型图。

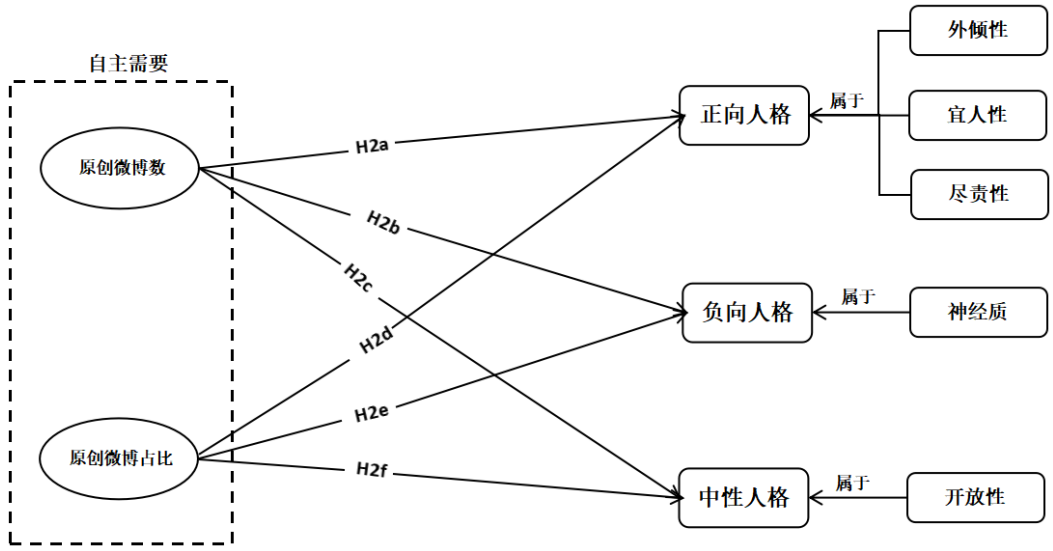


图 2 满足自主需要的行为指标研究模型

1.3.3 满足关系需要的特征对人格特质的预测

关系需要指个体渴望和他人建立情感联结，感受到归属感、关怀、理解、支持和爱。这一需要与微博用户简介字数、转发帖子数量、互粉粉丝数、关注数相对应。

微博用户通过在微博平台设置个人简介表达自我，字数越多，越能反映用户寻求他人认同感，其关系需要并未得到满足，因此本研究提出假设

H3a：简介字数对用户正向人格存在负向预测作用；

H3b：简介字数对用户负向人格存在正向预测作用；

H3c：简介字数对用户中性人格存在预测作用。

微博用户转发帖子数量是用户与他人交往频次的体现，用户通过转发他人微博的方式满足其关系需要。基于此本研究提出假设

H3d：转发帖子数量对用户正向人格存在负向预测作用；

H3e：转发帖子数量对用户负向人格存在正向预测作用；

H3f：转发帖子数量对用户中性人格存在预测作用。

微博用户的关注数是微博用户关注全部博主的数量，这一指标与转发帖子数量相似，都反映了用户试图借与他人建立联系满足自己的关系需要。因此本研究提出假设

H3g：关注数对用户正向人格存在负向预测作用；

H3h：关注数对用户负向人格存在正向预测作用；

H3i：关注数对用户中性人格存在预测作用。

微博用户的互粉粉丝数是微博用户与微博好友间建立的实际联系，体现用户的微博交友现状，这一指标越高表示用户越满足关系需要。因此本研究提出假设

H3j：互粉粉丝数对用户正向人格存在正向预测作用；

H3k：互粉粉丝数对用户负向人格存在负向预测作用；

H3l：互粉粉丝数对用户中性人格存在预测作用。

图 3 所示为满足关系需要的行为指标对人格预测作用的研究模型图。

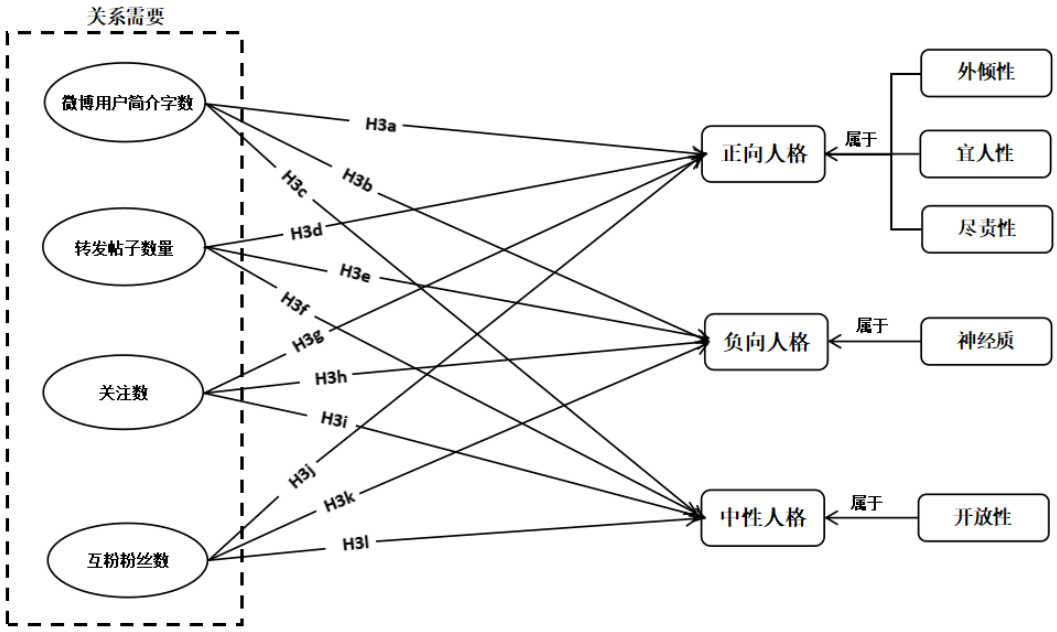


图 3 满足关系需要的行为指标研究模型

1.3.4 其他不满足基本心理需要的行为指标对人格特质的预测

研究在进行指标选取时注意到，存在已有研究中证明对微博用户人格存在预测作用的指标，但并不满足用户的基本心理需要，具体体现为指标无法体现用户主观感受和指标本身受平台功能限制。考虑到研究的全面性，本研究对这一类指标同样纳入研究，意在探索其实际的预测作用。

本研究各维度变量的对应情况如图 4 所示。

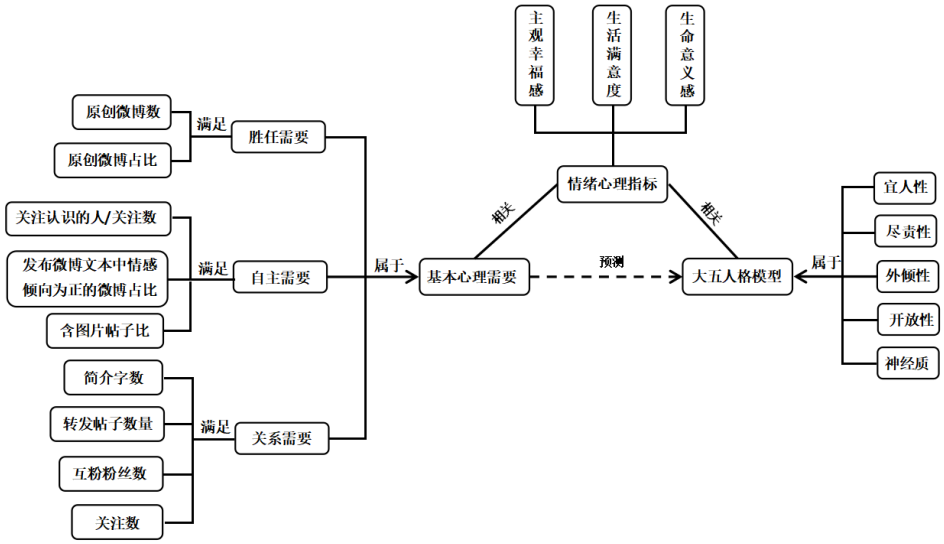


图 4 维度变量对应图

2 资料来源与方法

2.1 资料来源

采用滚雪球抽样法，以用户发布微博数大于 30 认定为活跃状态为标准发放问卷。研究回收问卷 314 份，综合问卷及微博用户行为特征数据结果，剔除无效、不完整、无法配对问卷后，最终得到有效问卷 225 份。

2.2 研究方法

本文通过问卷调查法获取用户基本信息及大五人格特质，数据采集软件 GooSeeker 获取用户行为特征数据。其中大五人格问卷采取的是目前国际上心理学界认可的 NEO-PI-R 简版大五人格问卷[37]。

2.3 统计学方法

采用 spss 25.0 对数据进行统计分析。选取最大值、最小值、均值及标准差进行描述性统计分析；采用逐步回归法进行多元回归分析，检验水准 α = 0.05，$P<0.05$ 为差异有统计学意义。

3 结果

3.1 大五人格特质描述性统计分析

在本研究中，NEO-PI-R 问卷的五个维度得分见表 1。

表 1 微博用户大五人格特质描述性统计分析结果

	Min	Max	M±SD
神经质	12	60	38±8.65
外倾性	15	53	34±9.06
开放性	25	60	44±7.18
宜人性	18	50	38±6.13
尽责性	22	58	40±5.96

由表 1 可知，被试微博用户的神经质 M=38，SD=8.65；外倾性 M=34，SD=9.06；开放性 M=44，SD=7.18；宜人性 M=38，SD=6.13；尽责性 M=40，SD=5.96。

3.2 逐步回归结果

研究使用 spss 25.0，选择逐步回归算法进行微博用户人格预测，算法以人格分值作为因变量，微博用户行为特征作为自变量，回归结果如表 2 所示（仅对存在预测作用的结果进行展示）。

表 2 微博用户行为特征对其人格逐步回归预测结果

因变量	自变量进入顺序	R	R^2	F	B	sig
外倾性	关注认识的人/关注数	0.336	0.117	27.821	5.950	0.000
	转发帖子数量	0.418	0.173	16.483	−0.001	0.000
	用户简介字数	0.452	0.204	7.333	−0.158	0.007
宜人性	原创微博数	0.208	0.043	9.870	−0.002	0.000
	关注认识的人/关注数	0.279	0.078	8.166	2.317	0.009
	互粉粉丝数	0.322	0.102	5.976	0.000	0.011
	发布微博文本中情感倾向为正的微博占比	0.349	0.122	4.816	5.402	0.029

尽责性	关注数	0.19	0.04	9.072	—	0.00
	互粉粉丝数	0.26	0.06	6.577	0.00	0.01
神经质	用户简介字数	0.20	0.04	9.541	0.16	0.00
	关注认识的人/关注数	0.27	0.07	7.906	—	0.00
	含图片帖子比	0.32	0.10	6.558	—	0.01
开放性	关注认识的人/关注数	0.32	0.10	25.67	—	0.00
	发布微博文本中情感倾向为正的微博占比	0.36	0.13	6.269	— 7.027	0.01
	原创微博占比	0.39	0.15	5.645	—	0.01

3.2.1 胜任需要预测人格特质结果

（1）关注认识的人/关注数对微博用户的外倾性（β=5.951，P<0.01）、宜人性（β=2.317，P<0.01）存在正向预测作用，对神经质（β=－3.416，P<0.01）及开放性（β= －4.779，P<0.001）存在负向预测作用，验证了假设 H1a，H1b，H1c。

（2）发布微博文本中情感倾向为正的微博占比对宜人性存在正向预测作用（β=5.402，P<0.05），对开放性存在负向预测作用（β=－7.027，P<0.05），验证了假设 H1d 和 H1f，而假设 H1e 不成立。

（3）微博用户含图片帖子比对神经质存在负向预测作用（β=－2.882，P<0.05），验证了假设 H1h，而假设 H1g 和 H1i 不成立。

3.2.2 自主需要预测人格特质结果

（1）微博用户的原创微博数对宜人性存在负向预测作用（β=－0.003，P<0.01），这一结果与假设 H2a 相冲突，因此假设 H2a，H2b，H2c 都不成立。

（2）微博用户的原创微博占比对开放性存在负向预测作用（β=－2.165，P<0.05），因此假设 H2f 成立，而 H2d 和 H2e 不成立。

3.2.3 关系需要预测人格特质结果

（1）微博用户简介字数对外倾性存在负向预测作用（β=－0.158，P<0.01），对神经质存在正向预测作用（β=0.165，P<0.01），因此假设 H3a，H3b 成立，而假设 H3c 不成立。

（2）微博用户转发帖子数量对外倾性存在负向预测作用（β=－0.001，P<0.01），因此假设 H3d 成立，而假设 H3e 和 H3f 不成立。

（3）微博用户的关注数对微博用户的尽责性存在负向预测作用（β=－0.004，P<0.01），因此假设 H3g 成立，假设 H3h 和 H3i 不成立。

（4）微博用户的互粉粉丝数对宜人性（β=0.009，P<0.05）和尽责性存在正向预测作用（β=0.009，P<0.05），因此假设 H3j 成立，而假设 H3k 和 H3l 不成立。

3.2.4 其他指标对人格特质的预测结果

研究中纳入的其他指标包括粉丝数、注册时长、微博总数、更换头像个数、点赞数、昵称字数及发布微博文本中情感倾向为负的微博占比，逐步回归结果发现以上指标对人格特质都不存在预测作用（P>0.05）。

4 讨论

4.1 胜任需要预测人格特质结果讨论

关注认识的人/关注数对微博用户的外倾性、宜人性存在正向预测作用。用户关注认识的人占关注数比例越大，反映用户在使用微博时更多与认识的人交流沟通。证明微博这一环境对用户来说是熟悉且具有掌控感的，这类用户更愿意将真实自我暴露给熟悉的人，因此其更倾向于是高外倾性、高宜人性、低神经质人群。但同时伴随的是其不愿意对不认识的微博用户分享自我，因此这类用户的开放性较低。

发布微博文本中情感倾向为正的微博占比反映微博用户在微博平台表达的情感偏向于正面的情感，这一类用户满足胜任需要的同时表现出易相处易沟通的特点，因此倾向于高宜人性。

微博用户含图片帖子比对神经质存在负向预测作用，含图片帖子比越高反映微博用户在表达想法及分享生活时能合理利用微博提供的功能，这一指标越高越满足微博用户的胜任需要，因此假设 H1h 成立。而假设 H1g 和 H1i 不成立的原因可能是研究只考虑了含图片帖子的比例，并未将图片所含有的情感倾向考虑进去，因此会对最后的预测产生影响。

4.2 自主需要预测人格特质结果讨论

微博用户的原创微博数对宜人性存在负向预测作用，与假设 H2a 恰恰相反且不满足假设 H2b、H2c，这一结果出现的原因可能是原创微博文本中包含用户情感，这一指标不仅仅要考虑数量的影响，同时可能存在其文本自身情感倾向存在的影响，即可能存在原创微博数高但其中情感倾向为正的微博占比低的情况，这一指标同时对自主需要和胜任需要存在影响，导致对人格的预测作用与假设不同；与原创微博数相似，H2d 和 H2e 不成立的原因可能因为原创微博占比这一指标的预测作用同样受其文本情感倾向的影响。原创微博占比越高，反映微博用户更倾向于在微博中分享个人经历、发表个人见解，而并非通过转发他人的微博文本与人进行沟通交流，因此这一指标越高越可能是低开放性人群，对开放性存在负向预测作用。

4.3 关系需要预测人格特质结果讨论

微博用户简介字数多，反映用户希望通过用户简介而非直接地沟通交流向他人展现自我，更多地是通过剖析内心表明观点和态度，因此这类用户偏向于含蓄内敛，对外倾性存在负向预测作用，对神经质存在正向预测作用，假设 H3a，H3b 成立；微博用户转发帖子数量是用户与他人交往频次的体现，转发帖子数量越多反映用户越希望引起他人的注意，这一类用户试图用转发的方式满足其关系需要，因此微博用户转发帖子数量对外倾性存在负向预测作用；微博用户的关注数与转发帖子数量指标相似，都反映了用户试图借与他人建立联系满足自己的关系需求，关注数越多表示用户越希望关系需求得到满足，反映实际的关系需求满足率越低，因此这一指标对尽责性存在负向预测作用假设成立；微博用户的互粉粉丝数是微博用户与微博好友间建立的实际联系，反映了用户的微博交友现状，这一指标越高表示用户越满足关系需要，因此微博用户的互粉粉丝数对宜人性和尽责性存在正向预测作用假设成立。H3 对转发帖子数量、互粉粉丝数和关注数提出的假设中，对于负向和中性人格的预测假设不成立的原因可能是以上行为指标更多是不含情感的行为，而神经质更多是用户情绪的体现，因此影响其预测效果。而开放性更多表现出用户的开放性思维和观点，与转发及关注行为的关联性较低，因此假设不成立。

5 结语

本文以基本心理需要理论视角研究用户信息行为对人格特质的预测作用，引入情绪心理指标作为媒介，在已知基本心理需要及大五人格特质都与情绪心理指标存在相关性的基础上，提出当用户信息行为满足其基本心理需要时其大五人格特质倾向于正向人格，解释了用户信息行为预测用户人格的潜在原因，弥补了现有信息行为预测人格特质研究在内在机理方面的缺陷。

本文仍存在一些局限，研究对象类型单一且样本数量有限，在后续的研究中将增加实验数据，拓展研究对象类型，以便进一步验证研究结果。

PU-129

融合多源数据特征及语义关系的创新前沿识别

陈逸菲，李永洁，孙轶楠，杨雪梅，于诗睿，唐小利
北京协和医学院/中国医学科学院医学信息研究所

1 引言

习近平总书记曾强调：“建设世界科技强国，不是一片坦途，唯有创新才能抢占先机。”自2022 年我国进入创新型国家行列以来，政府对于基础研究、技术创新、成果转化等各环节的投入力度不断加大，整体创新能力不断增强，国家科技实力迈上新台阶[错误!未找到引用源。]。创新推动发展，前沿则代表了创新过程中最具前瞻性的研究方向，因此基于已有的科研学术成果精准识别创新前沿、预测未来发展方向，对于完善国家整体战略布局、优化资源配置，抢占技术先机至关重要。鉴于此，本文将在梳理创新前沿基本概念及识别方法的基础上，总结当前存在的问题和局限，并在此基础上提出一套融合多源数据特征和语义关系的创新前沿识别体系，以期为精准识别提供一些参考。

2 相关研究

2.1 概念界定

“创新前沿”这一概念包含了两大主体，即“创新”和“前沿”。托马斯·库恩认为科学革命是一种范式向另一种范式的转换[错误!未找到引用源。]。范式是一种公认的理论体系，也是科学研究的基础，当科学研究中出现反常时，科学家们便开始批判和超越旧范式，解决反常背后的问题，最终探索和开创出新的范式，带来新的观察和新的研究成果，因此创新从根本上讲是范式的转换[错误!未找到引用源。]，即为满足某种需求改进或创造出新的理论体系，并以新产品、新技术、新方法等作为常见的表现形式。

学术领域的“前沿”一般指研究前沿，与之相近的概念还有研究热点、新兴研究等，当前研究普遍认为研究前沿由新兴研究发展而来，是最活跃、最具前瞻性和发展潜力的研究领域[错误!未找到引用源。错误!未找到引用源。错误!未找到引用源。]。2011 年，创新前沿首次作为一个完整概念被日本科学技术振兴机构 JST 提出，并通过基于专利引用的高被引论文聚类识别创新前沿[错误!未找到引用源。]，相较于研究前沿具有更高程度的创新性和前瞻性。因此本文要精准识别的创新前沿即为各领域新兴研究成果中最具前瞻性和发展潜力的那一部分，并且受范式不可预见性、不确定性和复杂性的影响，创新前沿同样具备这些特性，对其实现精准识别是当前一个值得探索的方向。

2.2 创新前沿的识别方法

创新前沿虽然在概念上与研究前沿、研究热点、新兴研究等有细微差别，但本质都反映了领域内前沿的思想理论或尖端技术方法，因此本文不对识别方法作精细的区分，主要可总结为以下几类：专家判读法、文献计量法、基于多源数据和多维指标的识别方法。

专家判读法指利用专家的领域知识和经验对创新前沿进行预测，通常借助德尔菲法、专家咨询法、头脑风暴法等获得专家的意见。这种定性方法虽权威、便捷，但主观性较强，可能存在主观与事实的偏差，于是学术界更多通过定量方法得出结论，再咨询专家意见验证识别结果的准确性。

文献计量法是定量识别前沿的重要方法，大致可分成引文分析和计量指标分析法。基于引文的分析主要立足于文献与文献之间，多从直接引用、共被引分析、文献耦合三种视角展开，但以单篇论文为节点的网络颗粒度较大，较难精准识别前沿主题，于是学者们将目光转向粒度更细的知识单元，采用词频分析（突发词监测）和共词分析法找寻科学概念间的关系[错误!未找到引用源。错误!未找到引用源。]，但由于句子是最小的语义单元，词和共词都不能揭示科学主题的全貌，且会忽略文章的语义信息[错误!未找到引用源。]。

早期的创新前沿识别通常以单一的知识源为研究对象，如仅围绕单一的论文数据或专利数据挖掘适合的测度指标，构建评价指标体系对各主题的创新程度进行测度。随着第四范式的兴起、科学

数据的数量和种类愈加丰富，学者们开始探究基于多源数据和多维指标的识别方法，如许晓阳等[错误!未找到引用源。]在关键词共现的基础上整合论文与专利的数据，发现了基础研究与应用研究相结合的研究前沿；杜建等[错误!未找到引用源。]认为重大创新存在于科学与技术的交叉处，并构建了专利与论文的交叉象限模型，将二者重叠的科学主题视为创新前沿，但基于引用关系的数据源整合和集成还不够深入。

综合已有研究可以发现，融合专利、论文等多来源数据和多维度指标进行创新前沿识别已成为趋势，然而多数研究选取的研究对象粒度较粗，且不同数据间缺乏语义关联，识别准确性受到局限，因此，本文从科学与技术的交叉处出发，以论文被专利引用作为科学与技术交叉的基础测度指标，基于 ESI 高被引论文和高价值专利数据构建引用关系网络，使用社区发现算法和相似度算法叠加进行主题聚类，并借助 SemRep 工具提取专利和论文的 SPO 三元组（Subject-Predicates-Object triple），从细粒度语义层面挖掘创新前沿，最后绘制创新前沿图谱。

3 研究设计

综合当前问题和局限，本文提出一套新的创新前沿识别体系，并将结果绘制为图谱，根据图谱可将识别出的前沿主题分为三大类，分别是文献代表的科学创新前沿、专利代表的技术创新前沿，以及二者的交叉——科技创新前沿。技术路线如图 1 所示。

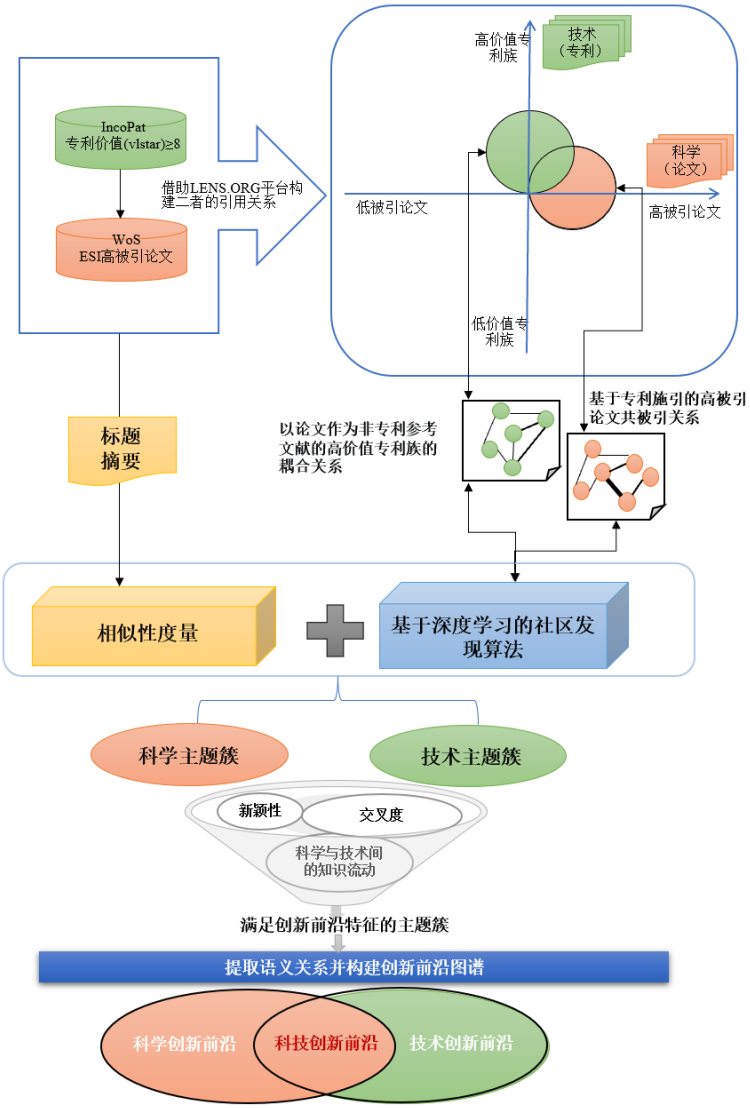


图 1 研究技术路线

具体步骤如下：

（1）采集高被引论文和高价值专利，并构建二者的引用关系网络。本文基于 Web of Science

数据库获取 ESI 高被引论文，基于 Incopat 专利数据库获取高价值专利（vlstar>8），并借助全球专利引用论文的数据平台 Lens.org 构建专利-论文引用的映射关系。

（2）构建论文共被引网络和专利耦合网络。本文将高价值专利引用高被引论文作为论文和专利的交叉测度指标，分别构建基于专利施引的高被引论文共被引网络和以论文作为 专利参考文献的高价值专利族的耦合网络，并分别作为科学创新前沿和技术创新前沿的研究基础数据。

（3）叠加社区发现算法与文本相似度算法，对步骤（2）的网络进行主题聚类。常见社区发现算法有 Leiden[错误!未找到引用源。]、CPM[错误!未找到引用源。]、AGDL[错误!未找到引用源。]及 Node Embedding+XMeans[错误!未找到引用源。]算法等，经过小样本人工判读后发现 Node Embedding+XMeans 算法的社区划分准确率最高，更适用于当前场景，故选择该算法。考虑到仅通过引用关系划分的簇准确性低，本文将论文及专利标题和摘要的文本相似度纳入考量范畴，与社区发现算法叠加进行主题聚类，经计算，叠加后的划分准确率从 77.85%提升到 79.88%。

（4）引入创新前沿特征指标打分，遴选前沿主题簇。郑彦宁[错误!未找到引用源。]、范少萍[错误!未找到引用源。]等学者对研究前沿的属性进行了探究，并提出一系列测度指标，本文认为其中新颖性反映了科学主题的时效性，学科交叉度（Rao-Stirling Index）意味着创新与进步，是最能反映创新程度的指标，此外，考虑到科学与技术的交叉融合和成果转化在前沿探测中具有重要地位，还应增加学术界向产业界的知识流动、技术相关的科学知识密集程度指标。因此科学前沿和技术前沿的度量指标如表 1 所示。

表 1 科学前沿和创新前沿的度量指标

类别	指标	计算方式	符号表示
科学前沿	论文所属学科的交叉度	Leydesdorff L 等[错误!未找到引用源。]	X_1
	论文新颖性	参考文献平均发表年与当前时间的距离	X_2
	学术界向产业界的知识流动	论文平均被专利引用的次数	X_3
技术前沿	专利领域交叉度	吴菲菲等[错误!未找到引用源。]	Y_1
	专利新颖性	参考文献平均发表年与当前时间的距离	Y_2
	技术相关的科学知识密集程度	专利平均引用论文的篇数	Y_3

采用熵权法为三个指标赋予一定权重，将得分高的主题视为创新前沿。

（5）基于 SemRep 提取语义关系，绘制创新前沿图谱。为从更细粒度的语义层面识别创新前沿，本文采用生物医学领域发展较为成熟的工具 SemRep 进行领域知识抽取，将论文和专利的摘要导入工具，经过其自动化处理形成 SPO 三元组，其中主语和宾语都是 UMLS 中的概念，谓语为 UMLS 语义网络中的语义关系。语义关系提取完成后，进一步遴选出创新前沿指标得分高的专利簇和论文簇，利用 Neo4j 工具构建可视化语义网络。

4 实证分析

本文将肺癌领域作为应用分析领域，以“Pulmonary Neoplasm、Pulmonary Neoplasms、

Lung Neoplasm、Lung Neoplasms、Lung Cancer、Lung Cancers、Pulmonary Cancer、Pulmonary Cancers、Lung carcinoma、Lung carcinomas、Pulmonary carcinoma、Pulmonary carcinomas”为关键词分别从 Web of Science 数据库和 Incopat 数据库检索相关论文和专利，并通过 Lens.org 平台确定论文的施引专利不为空值，得到符合条件的论文共 649 篇，专利 2289 篇。

4.1 创新前沿识别结果

对采集到的论文和专利数据建立二者的引用关系，分别构建高被引论文共被引网络和高价值专利耦合网络，并用社区发现算法叠加相似性算法进行主题聚类，聚类后得到专利有 15 个簇，论文有 13 个簇。再分别计算每个簇中每篇论文或专利的创新前沿指数，该簇所有论文或专利创新前沿得分的平均值即为该簇的创新前沿得分。

分别计算每篇论文的创新前沿指标，再通过熵权法赋权重，经计算，论文的创新前沿得分计算公式为：

$$S\ paper\ =0.052\ X\ 1\ +0.012\ X\ 2\ +0.936\ X\ 3$$

专利的创新前沿得分计算公式为：

$$S\ patent\ =0.440\ Y\ 1\ +0.065\ Y\ 2\ +0.495\ Y\ 3$$

取专利和论文创新前沿得分前 5 名的簇，利用 SemRep 工具提取摘要中的语义信息，并借助 Neo4j 工具构建 10 张可视化的语义网络。为进一步叠加专利语义网络和论文语义网络，本文分别计算了网络间的相似度，相似性高的前 5 个组合如表 2 所示。

表 2 相似度前五的专利与论文组

组合	I	II	III	IV	V
专利	#0	#0	#0	#0	#2
论文	#0	#1	#3	#5	#0

本文以组合 I 为例挖掘三种类型的创新前沿，专利簇 0 和论文簇 0 叠加之后的创新前沿图谱如图 2 所示（图中只显示了科学创新前沿与技术创新前沿重叠的部分，即科技创新前沿）。

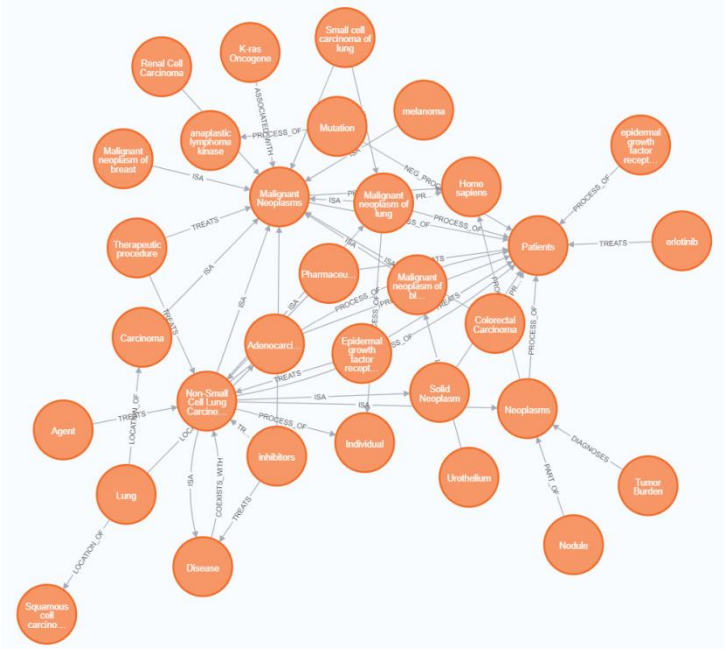


图 2 专利#0 与论文#0 叠加后的创新前沿图谱

4.2 结果分析

结合图谱，统计三类前沿的高频实体名称及其之间连线代表的关系可以发现：

科技创新前沿的高频词为：非小细胞肺癌、恶性肿瘤、抑制剂、埃罗替尼，推测当前的科技创新前沿为埃罗替尼等抑制剂在治疗非小细胞肺癌等恶性肿瘤上的治疗效果；科学创新前沿的高频词为：非小细胞肺癌、免疫疗法、PDCD1 gene，推测当前的科学创新前沿为针对 PDCD1 gene 的非小细胞肺癌免疫治疗；技术创新前沿的高频词为：非小细胞肺癌、药物制剂、抗体、肽，推测当前的技术创新前沿为非小细胞肺癌相关的多肽类抗体药物的研发。

5 结论

本文从科学与技术的交叉处出发，构建了一套融合多源数据特征和语义关系的创新前沿识别体系，并以肺癌领域的专利数据和论文数据开展应用研究，不仅描绘了对相关技术有重大影响的基础科学创新前沿，而且描绘了以相关科学知识为基础的原始发明创新前沿。此外，为克服实体间细粒度语义关联无法仅用参考关系网络表征的局限性，本文将文本相似关系与引文关系重叠，融合文献语义特征构建的加权网络可以更准确地识别创新前沿。然而当前仍存在一些局限：基于引用关系的交叉测度指标在度量专利与论文是否有内容上的交叉时存在一些偏差，后续将进一步探究能够直接反映专利和论文内容一致性的指标。

PU-130

2021 年度上海市区级医院科研竞争力分析

李娜 朱婷婷 牛玉宏

上海市卫生和健康发展研究中心（上海市医学科学技术情报研究所）

1 背景

公立医院高质量发展对区级医院科研管理提出了更高的要求，而其中的重要环节——科研评价需要更加专业、客观和准确。自 2013 年起，上海市卫生和健康发展研究中心（上海市医学科学技术情报研究所）每年运用科研竞争力评价体系对上海市医院的整体科研实力及不同学科的科研竞争力进行评价，为各医院及政府科研管理部门提供了重要参考。随着各类政策的不断落地及科研评价工作的不断变化，科研竞争力评价指标应随之进行动态调整。因此，根据以往 8 年的工作总结，并在广泛调研以及结合各方专家意见的基础上，对现有科研竞争力评价指标体系作进一步完善，运用完善后的科研竞争力评价指标体系对上海市区级医院的科研状况进行更加精准、客观、动态的评估，便于管理部门更好地掌握上海市区级医院的整体科研状况。

2 资料来源与方法

2.1 指标体系

基于原有上海市医学科研竞争力评价指标体系，结合文献综述、利益相关者调查、专家咨询、对比分析等适当增减相关指标，对基于层次分析法获得的权重结构作相应调整，得到完善后的上海市医学科研竞争力评价体系，并检验其信度和效度。

具体指标体系分为投入和产出两个一级指标。投入包括人力资源、科研项目、人才培养计划、科研基地与学科建设；产出分为产权性产出和奖励。

2.1.1 投入

2.1.1.1.人力资源

人力资源指标细分为获得博士学位的职工数、具有高级职称职工数、院士数、留学归国人员数、中华医学会学术任职数、中华口腔医学会学术任职数、中华中医药学会学术任职数、中国中西医结合学会学术任职数、上海市医学会学术任职数、上海市口腔医学会学术任职数、上海市中医药

学会学术任职数、上海市中西医结合学会学术任职数等三级指标。

2.1.1.2.科研项目

科研项目细分为国家级科研项目、省部级科研项目、厅局级科研项目等三级指标。

2.1.1.3.人才培养计划

人才培养计划细分为国家级人才项目、省部级人才项目、厅局级人才项目等三级指标。

2.1.1.4.科研基地与学科建设

科研基地细分为国家级学科建设项目/基地与工程中心、省部级学科建设项目/基地与工程中心、厅局级学科建设项目/基地与工程中心等三级指标。

2.1.2 产出

科研产出方面从产权性产出和奖励两方面进行衡量和评价。

2.1.2.1.产权性产出

产权性产出包括论文和专利。本文将论文数量和质量都引入评价指标。专利通过专利授权数和专利转化 2 个指标衡量。

2.1.2.2.奖励

奖励包括获奖科研项目和人才奖励。获奖科研项目包括国际奖项、国家级奖励、上海市级奖励、中华医学会奖励、上海医学会奖励等指标。人才奖励包括国家卫生健康委员会“有突出贡献中青年专家”、陈嘉庚奖、何梁何利奖、上海青年科技英才、上海科技精英、上海市科技功臣以及自然科学牡丹奖等指标。

2.2 数据来源

纳入本次评价的医院共 30 家，宝山区吴淞中心医院（简称“吴淞医院”），宝山区仁和医院（简称“仁和医院”），新华医院崇明分院（简称“新华崇明”），奉贤区中心医院（简称“奉中心”），上海市嘉定区中心医院（简称“嘉中心”），金山医院，第六人民医院金山分院（简称“金中心”），闵行区中心医院（简称“闵中心”），第五人民医院（简称“五院”），中山医院青浦分院（以简称“青中心”），松江区中心医院（简称“松中心”），浦东医院（简称“浦东医院”），浦东新区人民医院（简称“浦人民”），浦东新区公利医院（简称“公利医院”），浦东新区周浦医院（简称“周浦医院”），浦东新区浦南医院（简称“浦南医院”），同仁医院（简称“同仁医院”），第四人民医院（简称“四院”），瑞金医院卢湾分院（简称“瑞金卢湾”），第九人民医院黄浦分院（简称“九院黄浦”）

，静安区中心医院（简称“静中心”），静安区市北医院（简称“市北医院”），静安区闸北中心医院（简称“闸中心”），普陀区中心医院（简称“普中心”），普陀区人民医院（简称“普人民”），普陀区利群医院（简称“利群医院”），徐汇区中心医院（简称“徐中心”）

，第八人民医院（简称“八院”），杨浦区中心医院（简称“杨中心”），杨浦区市东医院（简称“市东医院”）。

3 结果

3.12021 年度上海市区级医院综合实力得分情况

上海市 30 家区级医院中，投入产出综合实力得分第一名是同仁医院，名列其后的 9 家医院及得分情况见图 1。

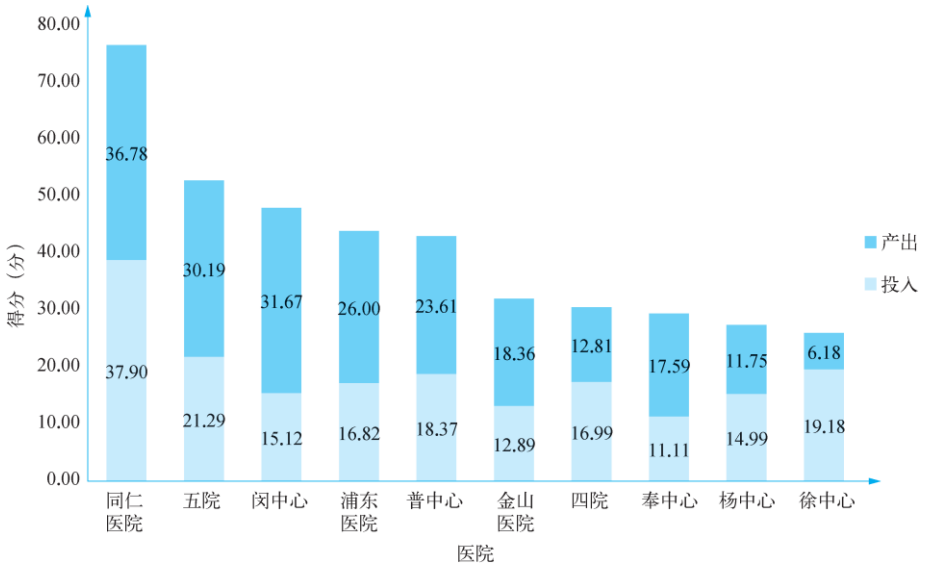


图1 2021年度上海市区级医院科研竞争力总得分排名（前十名）

3.2 各学科综合实力得分情况

将各医院科室按照教育部的学科分类合并处理，统一标准化后计算科研竞争力分值，部分学科排名前五名的医院及得分见表2。数据显示，在区级医院中，胃肠病学、神经病学、普通外科学、心血管病学、儿科学等学科的科研能力相对较强；重症医学、病理学、老年医学、核医学、风湿病学与自体免疫病学等学科的科研实力相对较弱，是医院未来需要重点督促和扶持的学科。此外，口腔医学、营养学、核医学、风湿病学与自体免疫病学等弱势学科除了人力资源外，其他指标得分几乎为0，故涉及上述学科的医院均不纳入前五名排名，并在表2内标注“除人力资源外几乎不得分”。

表2 2021年度上海市部分学科排名前五名的区级医院（单位：分）

学科	第一名		第二名		第三名		第四名		第五名	
	医院	得分	医院	得分	医院	得分	医院	得分	医院	得分
胃肠病学	同仁医院	52.57	闵中心	22.17	松中心	12.46	静心	7.39	浦东医院	6.88
神经病学	四院	37.98	闵中心	12.56	公利医院	11.62	杨心	9.80	同仁医院	9.66
普通外科学	闵中心	33.96	浦东医院	31.09	同仁医院	26.40	奉心	20.52	普中心	18.68
心血管病学	徐中心	29.17	普中心	25.89	周浦医院	13.19	同仁医院	12.66	五院	10.65
儿科学	同仁医院	23.87	金山医院	5.44	徐中心	1.77	浦人民	1.62	市东医院	1.54
肿瘤学	浦东医院	22.65	金山医院	17.28	周浦医院	7.56	同仁医院	7.45	普中心	4.04
药剂学	浦东医院	21.53	同仁医院	10.16	奉中心	9.83	五院	5.48	闵中心	4.80
血液病学	同仁医院	20.36	闸中心	9.93	闵中心	2.75	五院	2.74	青中心	2.11
妇产科	奉中	2	同	1	吴淞	6	五院	5	闵	5

学	心	0.09	仁医院	2.99	医院	.65		.97	中心	.96	
泌尿外科学	五院	19.71	公利医院	14.89	静中心	6.91	浦东医院	6.86	同仁医院	5.52	
呼吸病学	五院	18.92	青中心	9.26	普中心	8.79	徐中心	8.70	静中心	7.73	
骨外科学	浦东医院	18.87	杨中心	16.70	周浦医院	16.11	五院	15.71	同仁医院	11.88	
麻醉学	四院	18.27	同仁医院	15.78	公利医院	11.38	吴淞医院	7.80	松中心	5.64	
实验诊断学	奉中心	17.72	普人民	15.77	同仁医院	8.99	浦东医院	8.37	松中心	7.03	
内分泌病学与代谢病学	五院	17.46	浦东医院	14.33	周浦医院	12.83	奉中心	11.52	普中心	10.88	
急诊医学	闵中心	15.27	浦人民	14.67	松中心	7.14	普中心	6.75	金山医院	6.24	
全科医学	同仁医院	15.23	杨中心	10.93	五院	2.09	公利医院	1.17	松中心	0.93	
中医学	金山医院	14.21	普人民	6.61	普中心	6.25	公利医院	4.43	九院黄浦	4.41	
医学影像学	金山医院	13.68	同仁医院	10.27	松中心	6.63	五院	5.89	周浦医院	3.96	
康复医学	徐中心	13.42	周浦医院	6.38	普人民	5.21	新华崇明	2.50	四院	2.04	
肾脏病学	同仁医院	11.30	青中心	10.28	普中心	8.25	浦东医院	6.71	闸中心	6.50	
胸外科学	同仁医院	8.13	青中心	3.11	周浦医院	2.40	闵中心	2.05	新华崇明	1.54	
眼科学	杨中心	8.11	市北医院	5.52	周浦医院	4.15	同仁医院	2.32	普中心	2.18	
耳鼻咽喉科学	公利医院	8.11	同仁医院	2.05	普中心	1.82	金山医院	1.63	五院	1.30	
临床放射学	瑞金卢湾	7.38	同仁医院	4.92	公利医院	1.23	奉中心	1.22	九院黄浦	0.61	
神经外科学	普人民	6.91	浦南医院	5.61	闵中心	5.24	同仁医院	3.60	浦人民	2.31	
口腔医学	闵中心	6.81	五院	2.74	徐中心	2.61	浦人民	1.58	除人力资源外几乎不得分		
护理学	闵中心	5.26	同仁医院	3.76	杨中心	2.79	浦东医院	2.73	奉中心	2.64	
营养学	普人民	5.20	奉中心	1.83	徐中心	1.54	除人力资源外几乎不得分				
皮肤病学	闵中心	4.69	同仁医院	3.84	周浦医院	3.34	浦人民	1.97	四院	1.40	
感染性疾病学	普中心	4.25	五院	3.57	松中心	2.27	同仁医院	2.09	金山医院	1.29	
重症医学	闵中心	3.39	静中心	3.22	五院	2.15	市东医院	2.02	公利	1.34	

病理学	五院	3.29	浦东医院	1.63	浦人民	1.45	金山中心	1	嘉中心	1.38		
老年医学	闵中心	3.21	徐中心	3.16	杨中心	2.98	普中心	2.40	市北医院	2.12		
核医学	浦东医院	2.10	四院	0.71	除人力资源外几乎不得分							
风湿病学与自体免疫病学	同仁医院	1.70	普人民	0.97	除人力资源外几乎不得分							

3.3 部分重要指标情况

3.3..1 国家级科研项目

获得国家重点研发计划方面，同仁医院的儿科学、徐中心的心血管病学分别有 1 项项目获得国家重点研发计划资助。在国家自然科学基金项目方面，四院获得 1 项重点项目，其余均为面上和青年项目。2021 年上海市区级医院获得国家自然科学基金项目共 74 项，比 2020 年增加了 17 项，见表 3。

表 3 2021 年上海市区级医院获得国家自然科学基金项目的整体情况（单位：项）

单位	重点项目	面上项目	青年项目	专项基金项目	合计
同仁医院	0	9	8	1	18
四院	1	1	5	0	7
闵中心	0	4	3	0	7
浦东医院	0	4	2	1	7
五院	0	5	0	0	5
周浦医院	0	1	3	0	4
公利医院	0	3	0	0	3
浦人民	0	2	1	0	3
普中心	0	2	1	0	3
金山医院	0	1	2	0	3
普人民	0	1	2	0	3
杨中心	0	1	2	0	3
奉中心	0	1	1	0	2
新华崇明	0	1	1	0	2
静中心	0	0	2	0	2
嘉中心	0	0	1	0	1
瑞金卢湾	0	0	1	0	1
总计	1	36	35	2	74

3.3.2 论文发表情况

1.《科学引文索引》（*Science Citation Index*，SCI）论文发表情况

2021 年度上海市区级医院发表 SCI 论文共 1121 篇，比 2020 年增加了 113 篇。其中，发表 SCI 论文最多的是同仁医院，共计 134 篇。在 SCI 论文的单篇影响因子方面，单篇影响因子在 40~70 分的论文共计 3 篇、30~39 分的共计 5 篇、在 20~29 分的共计 6 篇，见表 4、表 5。

表 4 2021 年上海市区级医院 SCI 论文发表情况（前十名）

排名	单位	总影响因子	SCI 收录数量/篇
1	同仁医院	892.460	134
2	浦东医院	668.508	81
3	闵中心	579.351	111
4	五院	563.997	97
5	奉中心	393.397	57
6	四院	342.027	46
7	金山医院	300.054	68
8	周浦医院	278.386	66
9	杨中心	234.893	49
10	普中心	225.886	47

表 5 2021 年上海市区级医院 SCI 高影响因子论文单位排名

单位	科室	篇数	总篇数	单篇影响因子
浦东医院	肿瘤科	1	1	69.504
五院	呼吸与危重症医学科	1	1	46.297
浦东医院	内分泌科	1	1	40.706
	药剂科	1		32.086
奉中心	检验科	1	3	32.086
	内分泌代谢科	1		30.083
四院	内分泌代谢科	1	1	38.079
五院	中心实验室	1	1	38.079
同仁医院	普外科	1	2	23.655
	消化内科	1		23.168
浦东医院	普外科	1	1	25.832
静中心	呼吸内科	1	1	25.476

	四院	脑功能与人工智能转化研究所	1	1	24.798
院	金山医	肿瘤科	1	1	20.121

3.3.3 中文核心论文发表情况

2021 年度上海市区级医院发表核心期刊论文共 1322 篇。其中，发表中文核心论文最多的是浦东医院，共计 87 篇；其次为普陀区中心医院、新华医院崇明分院，各发表 84、80 篇，见表 6。

表 6 2021 年度上海市区级医院中文核心期刊论文发表情况（前十名）

排名	单位	中文核心期刊论文数量/篇
1	浦东医院	87
2	普中心	84
3	新华崇明	80
4	杨中心	79
5	五院	73
6	周浦医院	60
	青中心	60
7	同仁医院	59
8	公利医院	58
9	嘉中心	55
10	闵中心	53

3.3.4 专利情况分析

在专利授权方面，2021 年度上海市区级医院授权发明专利为 44 项，比 2020 年增加了 23 项。其中，发明专利最多的是金山医院，为 7 项。在专利转化方面，2021 年上海市区级医院中有 5 家医院成功实施转化，专利转化数量为 11 项（发明专利转化 3 项、实用新型专利转化 8 项），比 2020 年增加了 2 项，呈现上升趋势，见表 7、表 8。

表 7 2021 年度上海市区级医院国内发明专利授权数

单位	数量/项
金山医院	11
普人民	7
松中心	5
同仁医院	4
五院	3
市东医院	3
奉中心	2
普中心	2
仁和医院	2
杨中心	2
浦东医院	1
吴淞医院	1
公利医院	1

表 8 2021 年度上海市区级医院国内专利转化情况

转化类型	单位	科室	转化数量/个	转化金额/万元
发明专利	五院	骨科	1	5.000
	奉中心	普通外科	1	2.000
	周浦医院	皮肤科	1	1.357
		口腔科	2	4.000
		内镜中心	1	3.000
	闵中心	妇产科	1	2.000
实用新型专利		骨科	1	1.000
		护理部	1	1.000
	五院	骨科	1	3.000
	静中心	手及上肢外科	1	销售额提成、暂未实现销售

3.3.5 获奖情况分析

2021 年上海市区级医院各类科研项目获奖共 11 项，比 2020 年减少 3 项（表 9）。30 家医院均没有获得人才奖项，说明区级医院相对缺乏高端人才。

表 9 2021 年度上海市区级医院获得成果奖励情况

奖项类别	奖项级别	数量/个	获奖单位（数量/个）
上海市医学科技奖	三等奖	6	五院（2） 闵中心（1） 新华崇明（1） 同仁医院（1） 闸中心（1）
上海中医药科技奖	二等奖	1	普中心（1）
	三等奖	1	公利医院（1）
	科技成果推广奖	1	普中心（1）
上海中西医结合科技奖	三等奖	2	普中心（2）
总计		11	

4 讨论

从 2021 年度的科研相关数据来看，上海市 30 家区级医院科研竞争力较以往有一定进步。对比分析科研竞争力得分排名前五位的区级医院的得分发现，第一名（同仁医院）的科研竞争力总得分与第二名（五院）相差 23.1 分，第二名（五院）的得分与第五名（普中心）相差 9.51 分，说明区内或各区之间医院科研竞争力水平参差不齐。

从投入指标来看，大部分区级医院承担的科研项目数量较少，层次不高。虽然区级医院每年在 SCI 论文发表数量上都有所提升，但篇均影响因子仍处于较低水平，说明 SCI 论文发表质量仍有欠缺。从产出指标来看，区级医院在成果转化、获得科研成果奖励等方面能力较弱。在成果转化方面，虽然区级医院的专利转化数量较以往有所增加，但总体转化数量偏低，在 30 家区级医院中只有 5 家区级医院成果实施成功转化，总数仅有 11 项。在科研成果奖励方面，区级医院获得各类科研项目仅 11 项，比 2020 年少了 3 项，由此反映了区级医院对现有的科研成果没有进行系统性的

归纳总结；同时，在医院内部也缺少高层次人才培养和引进，没能够效利用地科研创新研究平台，进行“产学研医”融合，鉴于上述问题，提出如下建议。

一是加强科研成果系统梳理，完善学科布局。医院今后想要获得更多高层次项目，就需要对现有论文产出做一个系统性归纳总结，设计规划重点研究方向，按照该研究方向，有针对性地发表高质量 SCI 论文，并集中优势力量，寻找新的突破口，根据学科特色申报更高层次科研项目，而后遵照学科特色、补短板、发挥优势、差异发展的原则分类分层指导发展。

二是加强临床研究人员能力培养。无论是高质量 SCI 论文发表，还是高层次科研项目申报，成功的关键因素依然是人才队伍建设。只有加强人才队伍培养，才能更有力促进学科发展。建议医院可以利用现有科研平台和新型智库，打造一支在临床工作的专职从事科学研究的人才队伍，专职科研队伍与一线医生密切配合开展研究工作，可以缓解临床学科科研人手不足的问题，还能加速解决来自医疗卫生实际工作的科学问题，并将基础研究的成果更快运用于临床，促进转化医学工作，加大前瞻性、创新性的产出。

三是优化临床研究人才评价和激励机制。对承担临床研究工作的医疗卫生机构，适当提升其高级职称结构比例，优化临床研究岗位结构比例。进一步完善临床研究人才评价体系，优化临床研究职称评定制度，保障临床研究类人员职称晋升通道。把临床研究工作和取得的成果纳入各类人才评价和绩效考核体系。

四是有效利用现有科研平台，推动产学研医融合。医院需要进一步利用现有科研平台，加强与研究机构、大学研究院等协调配合，聚焦地区居民的健康需求，钻研关键核心技术，集智联合攻关，提高区域临床诊疗能力和科研水平。

PU-131

中医药治疗原发性肝癌的用药规律及研究热点分析

秦三利，石昕，林静，张仲男，丁颖，张渊厚，李娟，韩斌孝
甘肃省肿瘤医院（甘肃省医学科学研究院）

原发性肝癌是临床常见的恶性肿瘤之一，具有起病隐匿、进展迅速、易转移复发、生存期短等特点。中医药在肝癌防治、控制复发转移、改善中晚期患者临床症状、缓解西医治疗不良反应、提高生存质量、延长生存期等方面具有明显优势，是肝癌综合治疗中不可缺少的手段之一[1]。

近年来中医药治疗肝癌方面的相关文献较多，包括肝癌病因病机、辨证分型、专方专药、中药制剂，中医药在肝癌手术、化学治疗、放射治疗、介入治疗等综合治疗中的作用，苦参碱、人参皂苷、当归红芪超滤物、参芪抑瘤方、参雄抗癌方、当归贝母苦参丸等中药提取物与中药组方对肝癌细胞凋亡、增殖、侵袭、细胞周期、信号通路、自噬、辐射增敏、DNA 损伤、端粒酶活性影响以及对肝癌荷瘤小鼠抑瘤率、免疫功能、侵袭转移、血管生成影响的研究[2-5]。系统总结分析中医药治疗肝癌的研究成果，将为肝癌的中医药治疗提供参考依据。本文主要对中医药治疗肝癌方面用药规律及研究热点进行分析，旨在为中医药治疗肝癌方面提供理论依据。

1 资料与方法

在国内主要生物医学文献数据库中，检索中医药治疗肝癌方面的临床研究文献，同时辅以手工检索本单位订购的历年期刊进行补充。检索方式：[主题=（肝癌 or 肝肿瘤）and（中医药 or 中西医 or 推拿、按摩、针灸等中医适宜技术）] and [发表时间=2006 年 1 月 1 日-2021 年 12 月 31 日]。将文献题录导入 Endnote 软件，排除会议论文、重复发表、无作者及无全文文献，进一步根据全文内容确定纳入文献，利用 Excel 提取文献详细信息，运用 SPSS 22.0 对文献类型、关键词、中医证型、治则治法以及中药应用情况等进行频次统计，并在此基础上对关键词进行 SPSS 聚类分析，显示树状图聚类结果。

2 结果

2.1 文献概况

共纳入 2279 篇文献用于统计分析，其中期刊论文 1982 篇，学位论文 270 篇，科技成果 27 项。研究显示，中医药治疗肝癌方面文献基本呈现逐年上升趋势。

2.2 中医证型及治则治法分布情况

常见的肝癌中医证型包括肝郁脾虚型、气滞血瘀型、肝肾阴虚型、肝胆湿热型、脾虚型及湿热蕴结型，常用的治则治法包括健脾理气、活血化瘀、清热解毒、扶正祛邪扶正祛邪、软坚散结及疏肝健脾。

2.3 用药规律

2.3.1 中药方统计 共涉及 615 首中药方，其中煎服中药方 371 首，应用较多的是茵陈蒿汤、逍遥散及小柴胡汤；口服中药 117 种，应用最多的是金龙胶囊；静脉滴注中药 56 种，应用较多的是艾迪注射液、复方苦参注射液及华蟾素注射液；介入中药 19 种，应用较多的是鸦胆子油乳及华蟾素注射液；外敷中药 52 种。

2.3.2 不同证型用药 肝郁脾虚型、肝肾阴虚型及脾虚型肝癌多应用逍遥散加减（基本组方：柴胡、川楝子、白芍、当归、茯苓、白术、黄芪、白花蛇舌草、半枝莲等）和一贯煎加减（基本组方：沙参、麦冬、当归、生地黄、枸杞子、川楝子等）；气滞血瘀型肝癌多应用膈下逐瘀汤加减（组方：桃仁、丹皮、香附、红花、枳壳、丹参、白花蛇舌草等）和柴胡疏肝散加减（组方：白术、茯苓、柴胡、八月札、莪术、甘草等）；肝胆湿热型肝癌多应用茵陈五苓散加减（基本组方：茵陈、生大黄、栀子、茯苓、泽泻、猪苓、白术、黄柏、白花蛇舌草、半枝莲等）；湿热蕴结型肝癌多应用龙胆泻肝汤（基本组方：龙胆草、黄芩、栀子、泽泻、木通、车前子、当归、柴胡、甘草、白芍、生地黄、山药等）和茵陈蒿汤加减（组方：茵陈、栀子、大黄、田基黄、虎杖、柴胡、玄参、浙贝、牡蛎、甘草、莪术等）。

2.3.3 内服方中单味中药分布及功效分类 内服方共涉及 593 味中药，按累计频率达 75%及以上且相对频率由高到低依次为：茯苓、白术、半枝莲、柴胡、白花蛇舌草、黄芪、党参。对内服方中应用单味中药以补虚药应用最多，占药物总应用频次的 21.55%，主要包括白术、黄芪、党参、甘草、白芍、当归、鳖甲等；其次为清热药、活血化瘀药、利水渗湿药，分别占 10.30%、6.35%、5.58%。

2.3.4 外敷方中单味中药分布及功效分类 共涉及 193 味外敷中药，按累计频率达 75%及以上且相对频率由高到低依次为：冰片、乳香、没药、大黄、莪术、延胡索、白芷、三棱。外敷方中应用的单味中药以活血化瘀药最多，占药物总应用频次的 19.75%，主要包括乳香、没药、莪术、延胡索等；其次为开窍药，占 8.21%，主要为冰片和麝香；再次为泻下药，占 4.58%，主要为大黄和芒硝。

2.4 关键词聚类分析

依据词频界定公式 $T=(-1+\sqrt{8\times II})/2$（$II$ 是词频为 1 的关键词个数，T 为最低频次值）结合论文实际情况，共确定了 39 个高频关键词，对其进行聚类分析，共得到 4 组聚类组合，即 C1：肝肿瘤；C2：原发性；C3：TACE、癌痛、化学治疗、中晚期、临床疗效、介入疗法、晚期、免疫功能、生存质量、中西医结合疗法、肝功能、生存期、艾迪注射液、复方苦参注射液；C4：中医药、华蟾素注射液、腹水、辨证论治、栓塞后综合征、毒副反应、放射治疗、康艾注射液、中药外敷、射频消融、 癌痛、转移性、血管内皮生长因子、AFP、金龙胶囊、呃逆、术后、名医经验、护理、耳穴疗法、穴位贴敷、乙肝病毒、槐耳颗粒等。

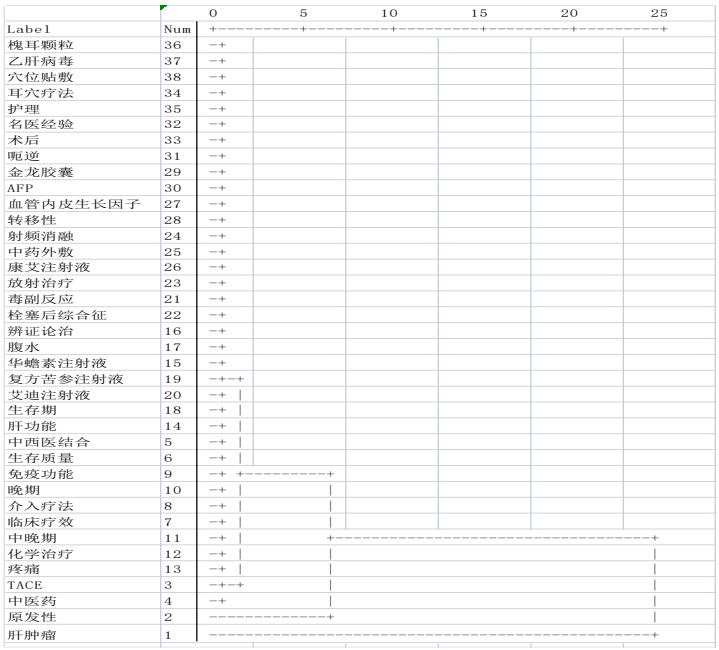


图 2 2016－2021 年中医药治疗肝癌文献高频关键词聚类树状图

2.5 研究热点总结

依据高频关键词的聚类结果结合论文实际情况，概括了十几年间中医药治疗肝癌方面的几方面研究热点：

（1）中医药联合 TACE、放射治疗、射频消融对中晚期原发性肝癌患者临床疗效、疼痛、免疫功能、生存质量、肝功能、毒副反应、VEGF 的影响；（2）中医药辅助化学治疗对原发性肝癌并发腹水患者生存质量、免疫功能、毒副反应等的影响；（3）中药外敷对原发性肝癌患者癌痛的疗效；（4）中药对肝癌介入术后栓塞综合征患者毒副反应、肝功能、生存质量的影响；（5）穴位贴敷、穴位按摩治疗对肝癌患者介入术后呃逆的疗效；（6）复方苦参注射液、艾迪注射液、康艾注射液等中成药持续静脉滴注辅助治疗对原发性肝癌介入治疗患者生存质量、毒副反应、肝功能、生存期的影响；（7）纯中医治疗：中成药持续静滴、中药联合中成药介入（包含鸦胆子油乳、华蟾素注射液介入等）、中药联合针刺或穴位按摩等对原发性肝癌的疗效以及中医药减少中成药介入治疗后毒副反应的研究；（8）名老中医的经验总结。

3 结论

3.1 中医药治疗肝癌的累计文献量呈现线性增长趋势，表明中医药治疗肝癌的研究处于稳定发展期，发展趋势良好。

3.2 原发性肝癌的常见中医证型为肝郁脾虚型、气滞血瘀型、肝肾阴虚型、肝胆湿热型、脾虚型及湿热蕴结型，且不同证型应用中药方剂不全相同；常用治则治为健脾理气、活血化瘀、清热解毒及扶正祛邪；药物应用方面以补虚药、清热药、活血化瘀药、利水渗湿药为主，提示治疗原发性肝癌时，适当配合健脾益气、活血化瘀、清热利湿类中药加减治疗，可取得较好疗效。本文通过文献整理、分析，初步探讨中医药治疗原发性肝癌的用药规律，可为中医药辅助治疗原发性肝癌临床和科研组方用药提供参考。

3.3 研究热点主要集中在中医药联合 TACE、放射治疗、射频消融等治疗对中晚期原发性肝癌及并发腹水患者疗效，中医药辅助治疗对肝癌介入术后毒副反应影响，纯中医药治疗对原发性肝癌患者疗效及减少中成药介入术后毒副反应以及名老中医经验总结的等方面。

PU-132

基于第三方评价数据的国家临床医学研究中心遴选方法构建与实证研究

牟燕[1,2] 何有琴[1,2] 温楠[2,3] 宋奎勐[1,2] 吴敏[1,2]
1.山东第一医科大学（山东省医学科学院）医疗保障学院
2.山东省医药卫生科技信息研究所
3.山东第一医科大学（山东省医学科学院）通识教育部

前言

国家临床医学研究中心[1]是由科技部会同国家卫生健康委（原国家卫生计生委）、军委后勤保障部和食品药品监管总局共同开展建设，面向我国疾病防治需求，以临床应用为导向，以医疗机构为主体，以协同网络为支撑，开展临床研究、协同创新、学术交流、人才培养、成果转化、推广应用的技术创新与成果转化类国家科技创新基地。国家临床医学研究中心是我国医学科技创新体系建设的重要内容，也是目前唯一面向医学领域设立的国家级科技创新基地。自 2013 年我国启动首批国家临床医学研究中心评审认定工作，截至 2022 年底分四批在全国 12 个省/市、20 个疾病领域和临床专科布局建设了 50 家国家临床医学研究中心。按照《国家临床医学研究中心五年（2017-2021 年）发展规划》的总体目标“在主要疾病领域和临床专科统筹建成 100 家左右的中心”，国家临床医学研究中心评审认定工作还将持续深入推进。

根据《国家临床医学研究中心管理办法（2017 年修订）》，国家临床医学研究中心的评审认定有严格的申报和评审程序。管理部门负责对材料进行形式审查，并对形式审查合格的申报单位开展进一步的能力评估和综合评审。同时，国家临床医学研究中心的申报工作面向所有的三级甲等医院法人单位，要求申报单位在该疾病领域和临床专科达到国内领先的临床技术水平，并有相关科研项目、领军人才、学术团队等方面的支持与保障。根据《中国卫生健康统计年鉴 2022》显示 2021 年全国有 1651 家三级甲等医院，与国家临床医学研究中心规划布局的数量相比，每年申报单位数量相对较多，管理部门如何综合衡量申报单位的科研水平、研究能力和条件资源等基础条件，做到优中选优，快速精准定位是国家临床医学研究中心评审认定工作的重中之重。

本文瞄准国家临床医学研究中心尚未布局和布局不足的疾病领域和临床专科，采用定性定量相结合方法，利用第三方评价数据构建针对疾病领域和临床专科的医院遴选评分方法并进行实证研究，快速有效遴选出符合国家临床医学研究中心要求的申报单位，为未来科学精准布局提供参考依据。

资料与方法

一、国家临床医学研究中心主要疾病领域和临床专科的未来布局空间

根据《国家临床医学研究中心五年（2017-2021 年）发展规划》，“十三五”期间国家临床医学研究中心建设主要疾病领域和临床专科有 28 个，其中重大疾病领域有老年疾病、代谢性疾病、恶性肿瘤、呼吸系统疾病、神经系统疾病、心血管疾病、精神心理疾病、感染性疾病。国家临床医学研究中心五年发展规划布局建设 100 家，对不同疾病领域和临床专科布局数量“根据疾病防控的实际需求，原则上各疾病领域和临床专科建设 1-3 家中心，重大疾病领域建设 3-5 家”[1]。

假设将重大疾病领域的国家临床医学研究中心最大布局数量空间设定为 5，非重大疾病领域的设定为 3，那么去除已布局数量，则为未来布局数量。经统计，除老年疾病（重大，6 家）和口腔疾病（非重大，4 家）已超出最大布局数量，慢性肾病、妇产疾病、消化系统疾病、血液系统疾病、眼耳鼻喉疾病已完成布局外，其他 21 个疾病领域和临床专科还有 1-4 家的未来布局空间（见表 1）。

表 1 国家临床医学研究中心主要疾病领域和临床专科的未来布局空间

编号	疾病领域和临床专科	批次	评审年度	是否重大疾病领域（是√，否×）	最大布局数量	已布局数量	未来布局数量
A1	代谢性疾病	第一批	2013	√	5	2	3
A2	恶性肿瘤	第一批	2013	√	5	2	3
A3	呼吸系统疾病	第一批	2013	√	5	3	2
A4	神经系统疾病	第一批	2013	√	5	1	4
A5	心血管疾病	第一批	2013	√	5	2	3
A6	精神心理疾病	第二批	2014	√	5	3	2
A7	感染性疾病	第四批	2018	√	5	3	2
A8	**老年疾病**	**第三批**	**2016**	**√**	**5**	**6**	**-1***
B1	**慢性肾病**	**第一批**	**2013**	×	**3**	**3**	**0**
B2	**妇产疾病**	**第二批**	**2014**	×	**3**	**3**	**0**
B3	**消化系统疾病**	**第二批**	**2014**	×	**3**	**3**	**0**
B4	**口腔疾病**	**第三批**	**2016**	×	**3**	**4**	**-1***
B5	**血液系统疾病**	**第四批**	**2018**	×	**3**	**3**	**0**
B6	**眼耳鼻喉疾病**	**第四批**	**2018**	×	**3**	**3**	**0**
B7	皮肤与免疫疾病	第四批	2018	×	3	2	1
B8	放射与治疗	第四批	2018	×	3	1	2
B9	骨科与运动康复	第四批	2018	×	3	1	2
B10	儿童健康与疾病	第四批	2018	×	3	2	1
B11	医学检验	第四批	2018	×	3	1	2
B12	中医	第四批	2018	×	3	2	1
B13	出生缺陷与罕见病	/	/	×	3	0	3
B14	危急重症	/	/	×	3	0	3

15	B	麻醉医学	/	/	×	3	0	3
16	B	影像医学	/	/	×	3	0	3
17	B	病理诊断	/	/	×	3	0	3
18	B	职业病	/	/	×	3	0	3
19	B	地方病	/	/	×	3	0	3
20	B	医学营养	/	/	×	3	0	3
合计						100	50	50

备注：A1-A8 为重大疾病领域和临床专科，B1-B20 为非重大疾病领域和临床专科。*表示已布局数量超出最大布局数量空间，故未来布局数量出现了负数（不排除存在撤并的可能）。

二、基于第三方评价数据的国家临床医学研究中心遴选方法

（一）研究技术路线和主要步骤

从《国家临床医学研究中心管理办法（2017 年修订）》关于国家临床医学研究中心申报条件入手，确定合适的第三方评价数据，在此基础上采用定性定量相结合的方法，开展国家临床医学研究中心遴选方法设计与实证分析，主要研究技术路线见图 1：

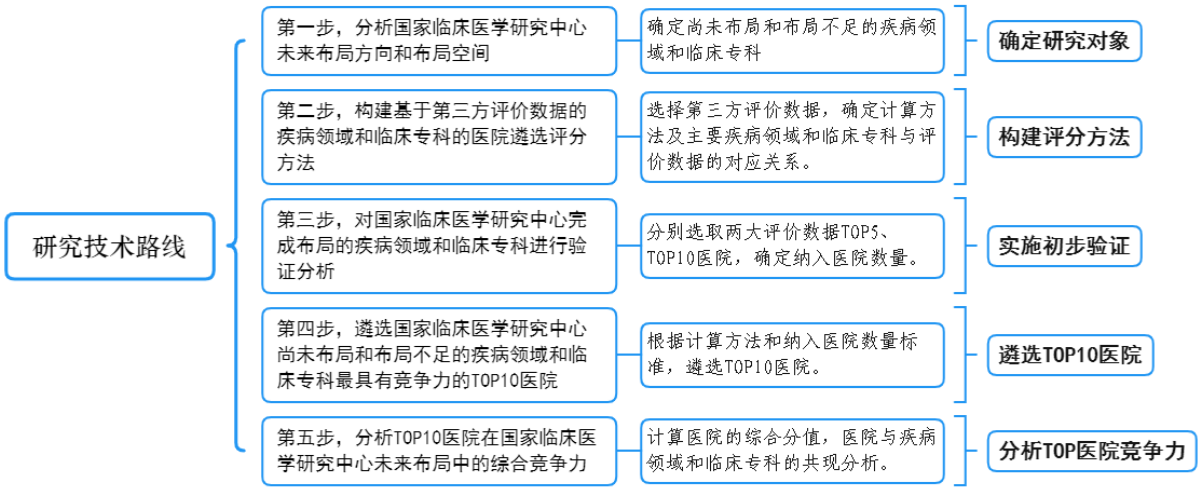


图 1 研究技术路线思维导图

主要步骤是：（1）确定研究对象，明确国家临床医学研究中心主要疾病领域和临床专科的未来布局方向和布局空间。（2）构建评分方法，对国内有关疾病领域和临床专科的第三方评价数据进行全面系统的定性分析，对其评价指标、评价维度和评价手段进行比较，对比国家临床医学研究中心申报条件的相关要求，选取适合的第三方评价数据。（3）实施初步验证，构建基于第三方评价数据的有关疾病领域和临床专科的医院遴选评分方法。（4）遴选 TOP10 医院，利用本方法遴选相关疾病领域和临床专科最具竞争力的 TOP10 医院。（5）分析 TOP10 医院竞争力，分析 TOP10 医院在国家临床医学研究中心未来布局中的综合竞争力。

1.第三方评价数据的遴选和确认

国家临床医学研究中心布局[1]遵从领域布局和区域布局相结合的原则。领域布局根据我国卫生与健康事业发展的需求，在充分掌握我国疾病谱特点的基础上，重点围绕慢性非传染性疾病和常见多发疾病进一步完善。《国家临床医学研究中心五年（2017-2021）发展规划》提出“十三五”国家临床医学研究中心建设主要疾病领域和临床专科 28 个，并细化了所包含的部分重点病种和技术领域。因此，单个疾病领域和临床专科可能涵盖多个学科或临床专科，如眼耳鼻喉疾病从学科角度

包括眼科学、耳鼻喉科学两个学科，从专科角度包括眼科、耳鼻喉科两个临床专科。

国家临床医学研究中心申报要求侧重于医疗机构在主要疾病领域和临床专科方面科技创新发展能力和水平，因此本研究应选取能够反映医院科研学术的相关评价数据，如中国医学科学院医学信息研究所研制的“中国医院科技量值（STEM）”（以下简称“STEM 排行榜”）是从科技产出、学术影响和科技条件三个维度进行评价，其中科技产出维度包括期刊论文、专利和标准以及论文应用等方面，学术影响维度包括科技奖励以及学术任职等方面，科技条件维度包括科研项目和科研平台等方面。复旦大学医院管理研究所研制的 “中国医院及专科声誉排行榜”（以下简称复旦排行榜）的专科综合排行是按照声誉标化值和科研标化值之和的大小进行排序，其中专科科研标化值主要是体现医院临床专科的 SCI 和科研奖项情况[3]。因此，STEM 排行榜的学科科技量值和复旦排行榜的专科科研标化值能够真实客观地反映医院相关疾病领域的研究基础和条件以及可持续研究能力，与申报国家临床医学研究中心所要求的研究基础、平台建设和临床资源情况等内容要求完全一致。

有研究者[4]以中国医学科学院医学信息研究所“中国医院科技量值（STEM）”、复旦大学医院管理研究所“中国最佳医院”以及香港艾力彼医院管理研究中心发布的“中国医院竞争力 • 顶级医院 100 强”3 个排行榜 2020 和 2021 年度的数据为基础，分析三大排行榜不同年度医院排名结果的稳定性并计算同一医院在 3 个排行榜中排名结果的差异性，结果发现“中国医院科技量值(STEM)”与“中国最佳医院”差异性不显著，但两者均与“中国医院竞争力 • 顶级医院 100 强”榜单差异性显著。由此可知虽然 STEM 排行榜和复旦排行榜的评价指标、评价维度、评价手段存在差异，但是在医院排名结果上差异性不显著。因此，对两个权威榜单的医院排名并进行标准化处理，能够为国家临床医学研究中心的申报单位遴选提供参考范围。

除此，两个排行榜均为我国目前公认的最权威、最受关注的榜单，并且已在医院科研评价和学科发展研究领域得到一定应用[5-6]。本研究从学科和专科两个层面综合评价医院在有关疾病领域和临床专科的科技创新能力和水平，所使用的数据为 STEM 排行榜 2013-2020 年度学科科技量值和复旦排行榜 2017-2020 年度专科科研标化值。

2.医院遴选的评分方法

首先确定国家临床医学研究中心的疾病领域和临床专科与 STEM 排行榜的学科和复旦排行榜的专科的对应关系。然后按照学科科技量值和专科科研标化值大小选取 TOP10 医院，对其进行量化赋值，并最终计算医院的综合分值。

设定“YY”释义为“医院”、“XK”为“学科”、“ZK”为“专科”，则“$XK_{排名}$”为某家医院的学科排名，“$ZK_{排名}$”为某家医院的专科排名，“$XK_{赋值}$”为某家医院的学科赋值，“$ZK_{赋值}$”为某家医院的专科赋值，“$YY_{综合}$”为某家医院的综合分值。

赋值原则是根据医院学科和专科排名进行逆向赋值，每家医院综合分值计算方法有以下四种特殊情况：

第一种：当某个疾病领域对应的学科数量为“1”且专科数量也为“1”时，

那么，$XK_{排名} = i$（$i = 1,\ldots,10$），$XK_{赋值} = 11 - i$

$ZK_{排名} = j$（$j = 1,\ldots,10$），$ZK_{赋值} = 11 - j$

当一家医院同时出现在学科排名和专科排名中时，其综合分值为：

$YY_{综合} =$（$XK_{赋值} + ZK_{赋值}$）/2

=（11 - i +11 - j ）/2

= 11 - (i + j)/2

否则，$YY_{综合} = XK_{赋值}$ 或 $ZK_{赋值}$。

第二种：当某个疾病领域对应的学科数量为“1”且专科数量为“2”时，

那么，$XK_{排名} = i$（$i = 1,\ldots,10$），$XK_{赋值} = 11 – i$

$ZK_{排名1} = j_1$（$j = 1,\ldots,10$），$ZK_{赋值1} = 11 - j_1$

$ZK_{排名2} = j_2$（j = 1,…,10），$ZK_{赋值2} = 11 - j_2$

当一家医院同时出现在两个专科排名中时，其专科的综合分值为：

$ZK_{综合} =$ （$ZK_{赋值1} + ZK_{赋值2}$）/2

= （$11 - j_1 + 11 - j_2$ ）/2

= 11 -（$j_1 + j_2$） /2

当一家医院同时出现在学科排名和合并后的专科排名中时，其综合分值为：

$YY_{综合}$=（$XK_{赋值}$+ $ZK_{综合}$）/2

= [（$11 – i + 11 -$（$j_1 + j_2$） /2] /2

= 11-（$2i + j_1 + j_2$） /4

否则，$YY_{综合}$= $XK_{赋值}$ 或 $ZK_{综合}$

第三种：当某个疾病领域对应的学科数量为“2”且专科数量为“1”时，

那么，$XK_{排名1}$= i_1（i = 1,…,10），$XK_{赋值1} = 11 – i_1$

$XK_{排名2} = i_2$（i = 1,…,10），$XK_{赋值2} = 11 – i_2$

$ZK_{排名}$= j（j = 1,…,10）， $ZK_{赋值} = 11 - j$

当一家医院同时出现在两个学科排名中时，其学科的综合分值为：

$XK_{综合} =$（$XK_{赋值1}$+$XK_{赋值2}$）/2

= （$11 - i_1 + 11 - i_2$ ）/2

= 11 -（$i_1 + i_2$） /2

当一家医院同时出现在专科排名和合并后的学科排名中时，其综合分值为：

$YY_{综合}$=（$XK_{综合}$+ $ZK_{赋值}$）/2

= [11 -（$i_1 + i_2$） /2 + 11 - j] /2

= 11 - （$2 j + i_1 + i_2$）/4

否则，$YY_{综合}$= $ZK_{赋值}$ 或 $XK_{综合}$

第四种：当某个疾病领域对应的学科数量为“2”且专科数量为“2”时，

那么，$XK_{排名1}$= i_1（i = 1,…,10），$XK_{赋值1} = 11 - i_1$

$XK_{排名2} = i_2$（i = 1,…,10），$XK_{赋值2} = 11 - i_2$

$ZK_{排名1}$= j_1（j = 1,…,10），$ZK_{赋值1} = 11 - j_1$

$ZK_{排名2} = j_2$（j = 1,…,10），$ZK_{赋值2} = 11 - j_2$

当一家医院同时出现在两个学科排名中时，其学科的综合分值为：

$XK_{综合} =$（$XK_{赋值1}$+$XK_{赋值2}$）/2

= （$11 - i_1 + 11 - i_2$ ）/2

= 11 -（$i_1 + i_2$） /2

当一家医院同时出现在两个专科排名中时，其专科的综合分值为：

$ZK_{综合} =$ （$ZK_{赋值1} + ZK_{赋值2}$）/2

= （$11 - j_1 + 11 - j_2$ ）/2

= 11 -（$j_1 + j_2$） /2

当一家医院同时出现在合并后的学科排名和合并后的专科排名中时，其综合分值为：

$YY_{综合}$=（$XK_{综合}$+ $ZK_{综合}$）/2

= [11 -（$i_1 + i_2$） /2 + 11 -（$j_1 + j_2$） /2] /2

= 11 -（$i_1 + i_2 + j_1 + j_2$ ）/4

否则，$YY_{综合}$= $XK_{综合}$ 或 $ZK_{综合}$

依据国家临床医学研究中心的疾病领域和临床专科，查找与 2020 年度 STEM 排行榜 31 个学科和复旦排行榜 42 个专科的对应关系，并根据对应学科和专科数量确定适用的计算方法。经统计，20 个疾病领域和临床专科均能找到对应或包含的学科和专科，其中 6 个为已布局完成。

表 2 国家临床医学研究中心主要疾病领域和临床专科与两大排行榜学科和专科的对应关系

编号	布局状态	疾病领域	STEM 排行榜学科	复旦排行榜专科	计算方法
B1	布局完成	慢性肾病	肾脏病学	肾脏病	第一种
B2	布局完成	妇产疾病	妇产科学	妇产科	第一种
B3	布局完成	消化系统疾病	消化病学	消化病	第一种
B4	布局完成	口腔疾病	口腔医学	口腔科	第一种
B5	布局完成	血液系统疾病	血液病学	血液学	第一种
B6	布局完成	眼耳鼻喉疾病	眼科学、耳鼻咽喉科学	眼科、耳鼻喉科	第四种
A1	布局不足	代谢性疾病	内分泌病学与代谢病学	内分泌	第一种
A2	布局不足	恶性肿瘤	肿瘤学	肿瘤学	第一种
A3	布局不足	呼吸系统疾病	呼吸病学	呼吸科	第一种
A6	布局不足	精神心理疾病	精神病学	精神医学	第一种
A7	布局不足	感染性疾病	传染病学	传染感染	第一种
B8	布局不足	放射与治疗	肿瘤学	肿瘤学	第一种
B9	布局不足	骨科与运动康复	骨外科学	骨科	第一种
B15	未布局	麻醉医学	麻醉学	麻醉科	第一种
B10	布局不足	儿童健康与疾病	儿科学	小儿内科、小儿外科	第二种
B13	未布局	出生缺陷与罕见病	儿科学	小儿内科、小儿外科	第二种
A4	布局不足	神经系统疾病	神经病学、神经外科	神经内科、神经外科	第四种
A5	布局不足	心血管疾病	心血管病学、心血管外科学	心血管病、心外科	第四种
B7	布局不足	皮肤与免疫疾病	皮肤病学、风湿病学与自体免疫病学	皮肤科、风湿病	第四种
B14	布局不足	危急重症	急诊医学、重症医学	急诊医学、重症医学	第四种
A8	布局完成	老年疾病	/	老年医学	不适用
B11	布局不足	医学检验	/	检验医学	不适用
B16	未布局	影像医学	/	超声医学	不适用
B	未布	病理诊断	/	病理科	不适

17	局				用	
12	B 布局不足	中医	/	/	不适用	
18	B 未布局	职业病	/	/	不适用	
19	B 未布局	地方病	/	/	不适用	
20	B 未布局	医学营养	/	/	不适用	

备注：A2 与 B8、B10 与 B13 对应相同学科和专科

结 果

一、基于第三方评价数据的国家临床医学研究中心遴选方法的实证分析

（一）对完成布局的 6 个疾病领域和临床专科进行验证分析

为了验证遴选方法的有效性，本研究首先将完成布局的 6 个疾病领域和临床专科作为研究对象，按照基于第三方评价数据的国家临床医学研究中心遴选方法，分别选取 STEM 排行榜学科科技量值和复旦排行榜专科科研标化值排名为 TOP5 医院、TOP10 医院，通过赋值计算并排序，筛选综合分值 TOP5 医院、TOP10 医院。

通过数量差判断赋值前后的 TOP5 医院、TOP10 医院是否涵盖已认定为国家临床医学研究中心的医院。当数量差为 0，说明赋值前后的 TOP5 或 TOP10 医院名单涵盖已认定为国家临床医学研究中心的医院。当数量差为 1，说明有 1 家已认定为国家临床医学研究中心的医院未出现在赋值前后的医院名单中，以此类推（见表 3）。

结果发现：当且仅当分别选取 STEM 排行榜学科科技量值和复旦排行榜专科科研标化值 TOP10 医院并形成赋值后 TOP10 医院，6 个疾病领域和临床专科的所有数量差均为 0。

表 3 6 个布局完成的疾病领域和临床专科与赋值后医院名单之间的数量差

编号	疾病领域	评审年度	局数量	科研标化值	科技量值	赋值后		科研标化值	科技量值	赋值后	
				TOP5	TOP5	TOP5	TOP10	TOP10	TOP10	TOP5	TOP10
1	B 慢性肾病	2013		1*	0**	0	0	0	0	0	0
2	B 妇产疾病	2014		2	1	0	0	0	0	1	0
3	B 消化系统疾病	2014		0	2	2	0	0	1	1	0
4	B 口腔疾病	2016		1	0	0	0	0	0	0	0
5	B 血液系统疾病	2018		0	0	0	0	0	0	0	0
6	B 眼耳鼻喉疾病	2018		2	1	2	1	0	0	1	0

*因“中国医院及专科声誉排行榜”于 2017 年才发布专科科研标化值，故 2017 年之前评审认定的国家临床医学研究中心采用的科研标化值均为 2017 年。

**采用的“中国医院科技量值（STEM）”科技量值年度与国家临床医学研究中心评审认定年度一致。

（二）14 个尚未布局和布局不足的疾病领域和临床专科的遴选分析

1. 对 14 个尚未布局和布局不足的疾病领域和临床专科进行遴选

由表 2 国家临床医学研究中心主要疾病领域和临床专科与学科/专科的对应关系来看，适用于本研究计算方法的疾病领域和临床专科共有 14 个，其中未布局的有 3 个。以 2020 年度“中国医院及专科声誉排行榜”专科科研标化值和 2020 年度“中国医院科技量值（STEM）”学科科技量值为准，按照本研究的量化赋值方法计算分值，筛选 Top10 医院。因存在分值相同、并列第 10 的情况，故最终确定 145 个名额，分别是代谢性疾病（11）、恶性肿瘤（11）、呼吸系统疾病（10）、神经系统疾病（10）、心血管疾病（10）、精神心理疾病（10）、感染性疾病（10）、皮肤与免疫疾病（10）、放射与治疗（11）、骨科与运动康复（10）、儿童健康与疾病（10）、出生缺陷与罕见病（10）、危急重症（11）、麻醉医学（11），初步形成了由 69 家医院构成的 14 个未布局和布局不足的疾病领域和临床专科 Top10 名单，详见表 4。值得注意的是，除中南大学湘雅二医院在代谢性疾病和精神心理疾病认定 2 家国家临床医学研究中心外，其他 15 家也在有关疾病领域和临床专科被认定为国家临床医学研究中心。

表 4 14 个未布局和布局不足的疾病领域和临床专科的 Top10 医院

编号	疾病领域和临床专科	排名	医院名称	分值	备注
1	代谢性疾病（11）	1	**上海交通大学医学院附属瑞金医院**	9.50	**已认定**
		2	中国医学科学院北京协和医院	9.00	
		3	**中南大学湘雅二医院**	7.50	**已认定**
		4	上海市第六人民医院	7.00	
		5	四川大学华西医院	6.50	
		6	复旦大学附属中山医院	5.00	
		7	中国医科大学附属第一医院	4.50	
		8	山东省立医院	2.50	
		9	中国人民解放军总医院	1.50	
		9	中山大学附属第一医院	1.50	
		9	华中科技大学同济医学院附属同济医院	1.50	
2	恶性肿瘤（11）	1	中山大学肿瘤防治中心	9.50	
		2	**中国医学科学院肿瘤医院**	9.00	**已认定**
		3	复旦大学附属肿瘤医院	8.00	
		4	四川大学华西医院	6.50	
		5	北京大学肿瘤医院	5.50	
		6	复旦大学附属中山医院	4.00	
		7	中国医学科学院北京协和医院	3.00	
		8	天津市肿瘤医院	2.50	

			9	**天津医科大学肿瘤医院**	**2.00**	**已认定**
		10	1	山东省肿瘤医院暨山东省肿瘤防治研究院	1.50	
			1	中南大学湘雅医院	1.50	
			1	**广州医科大学附属第一医院**	**10.00**	**已认定**
			2	中日友好医院	8.50	
			3	四川大学华西医院	7.50	
			4	复旦大学附属中山医院	5.50	
			5	首都医科大学附属北京朝阳医院	4.50	
3	A	呼吸系统疾病（10）	5	中国医学科学院北京协和医院	4.50	
			5	华中科技大学同济医学院附属同济医院	4.50	
			8	上海市肺科医院	3.50	
			9	上海交通大学医学院附属瑞金医院	2.50	
			1	华中科技大学同济医学院附属协和医院	2.00	
		10	1	**首都医科大学附属北京天坛医院**	**9.50**	**已认定**
			2	首都医科大学宣武医院	8.25	
			3	四川大学华西医院	6.75	
			3	复旦大学附属华山医院	6.75	
			5	中南大学湘雅医院	5.00	
4	A	神经系统疾病（10）	6	浙江大学医学院附属第二医院	4.00	
			7	天津医科大学总医院	3.75	
			8	中国医学科学院北京协和医院	3.50	
			9	上海交通大学医学院附属瑞金医院	1.50	
			9	华中科技大学同济医学院附属同济医院	1.50	
			1	**中国医学科学院阜外医院**	**10.00**	**已认定**
			2	**首都医科大学附属北京安贞医院**	**8.25**	**已认定**
5	A	心血管疾病（10）	3	复旦大学附属中山医院	6.75	
			4	华中科技大学同济医学院附属协和医院	4.25	

			5	广东省人民医院	3.75	
			6	四川大学华西医院	3.50	
			7	浙江大学医学院附属第二医院	2.25	
			8	上海交通大学医学院附属瑞金医院	2.00	
			8	北京大学第一医院	2.00	
			8	武汉大学人民医院	2.00	
			1	上海市精神卫生中心	8.50	
			2	**中南大学湘雅二医院**	8.50	已认定
			2	**北京大学第六医院**	8.50	已认定
			4	四川大学华西医院	7.50	
6	A 精神心理疾病（10）		5	**首都医科大学附属北京安定医院**	6.00	已认定
			6	广州医科大学附属脑科医院	4.00	
			7	西安交通大学第一附属医院	3.50	
			8	重庆医科大学附属第一医院	2.50	
			9	南京脑科医院	1.50	
			9	成都市第四人民医院	1.50	
			1	**浙江大学医学院附属第一医院**	10.00	已认定
			2	首都医科大学附属北京地坛医院	8.00	
			3	首都医科大学附属北京佑安医院	7.00	
			4	中国人民解放军总医院	4.00	
7	A 感染性疾病（10）		4	华中科技大学同济医学院附属同济医院	4.00	
			6	北京大学第一医院	3.50	
			6	华中科技大学同济医学院附属协和医院	3.50	
			8	南方医科大学南方医院	3.00	
			8	武汉大学中南医院	3.00	
		0	1	复旦大学附属华山医院	2.50	
7	B 皮肤与免疫疾病（10）		1	**中国医学科学院北京协和医院**	**7.50**	已认定
			2	中南大学湘雅医院	6	

				.00	
		2	中南大学湘雅二医院	6.00	
		4	四川大学华西医院	4.25	
		5	上海交通大学医学院附属仁济医院	4.00	
		6	复旦大学附属华山医院	3.50	
		6	北京大学人民医院	3.50	
		8	空军军医大学西京医院	3.25	
		9	南京大学医学院附属鼓楼医院	2.50	
		9	上海市皮肤病医院	2.50	
		1	中山大学肿瘤防治中心	9.50	
		2	中国医学科学院肿瘤医院	9.00	
		3	复旦大学附属肿瘤医院	8.00	
		4	四川大学华西医院	6.50	
		5	北京大学肿瘤医院	5.50	
8	E 放射与治疗（11）	6	**复旦大学附属中山医院**	4.00	**已认定**
		7	中国医学科学院北京协和医院	3.00	
		8	天津市肿瘤医院	2.50	
		9	天津医科大学肿瘤医院	2.00	
		10	山东省肿瘤医院暨山东省肿瘤防治研究院	1.50	
		10	中南大学湘雅医院	1.50	
		1	上海市第六人民医院	10.00	
		2	北京大学第三医院	8.50	
		3	四川大学华西医院	6.00	
		4	中南大学湘雅医院	5.50	
9	E 骨科与运动康复（10）	4	北京积水潭医院	5.50	
		6	中国医学科学院北京协和医院	4.00	
		7	**中国人民解放军总医院**	3.50	**已认定**
		7	河北医科大学第三医院	3.50	
		9	海军军医大学第二附属医院	3	

						.00	
		0	1	浙江大学医学院附属第二医院	.50	2	
			1	首都医科大学附属北京儿童医院	.75	8	
			2	复旦大学附属儿科医院	.00	8	
			3	**重庆医科大学附属儿童医院**	.75	7	**已认定**
			4	**浙江大学医学院附属儿童医院**	.25	7	**已认定**
10	E 儿童健康与疾病（10）		5	广州市妇女儿童医疗中心	.00	6	
			6	上海交通大学医学院附属新华医院	.75	5	
			7	上海交通大学医学院附属上海儿童医学中心	.50	3	
			8	首都儿科研究所附属儿童医院	.50	2	
			9	四川大学华西第二医院	.25	2	
		0	1	中国医科大学附属盛京医院	.50	1	
			1	首都医科大学附属北京儿童医院	.75	8	
			2	复旦大学附属儿科医院	.00	8	
			3	重庆医科大学附属儿童医院	.75	7	
			4	浙江大学医学院附属儿童医院	.25	7	
13	E 出生缺陷与罕见病（10）		5	广州市妇女儿童医疗中心	.00	6	
			6	上海交通大学医学院附属新华医院	.75	5	
			7	上海交通大学医学院附属上海儿童医学中心	.50	3	
			8	首都儿科研究所附属儿童医院	.50	2	
			9	四川大学华西第二医院	.25	2	
		0	1	中国医科大学附属盛京医院	.50	1	
			1	四川大学华西医院	.00	7	
			1	中国医学科学院北京协和医院	.00	7	
			3	浙江大学医学院附属第二医院	.25	4	
14	E 危急重症（11）		3	山东大学齐鲁医院	.25	4	
			5	东南大学附属中大医院	.75	3	
			5	武汉大学中南医院	.75	3	

			7	首都医科大学附属北京朝阳医院	3.50
			8	郑州大学第一附属医院	3.00
			8	中日友好医院	3.00
		0	1	上海交通大学医学院附属瑞金医院	2.50
		0	1	山东省立医院	2.50
			1	四川大学华西医院	9.50
			2	华中科技大学同济医学院附属协和医院	9.00
			3	华中科技大学同济医学院附属同济医院	6.50
			4	中国医科大学附属盛京医院	4.50
			5	中国医学科学院北京协和医院	4.00
15	E 麻醉医学（11）		6	郑州大学第一附属医院	3.50
			6	浙江大学医学院附属第一医院	3.50
			8	上海交通大学医学院附属仁济医院	3.00
			9	复旦大学附属中山医院	2.50
		0	1	空军军医大学西京医院	2.00
		0	1	中南大学湘雅医院	2.00

2.14 个未布局和布局不足的疾病领域和临床专科的 Top10 医院综合竞争力分析

进一步分析医院在国家临床医学研究中心未来布局中的综合竞争力，即在 145 个名额的基础上，剔除已经认定的 17 家国家临床医学研究中心，分析医院在 128 个名额中的竞争力情况。经统计删除重复项，发现 128 个名额涵盖 63 家医院。

第一步统计 63 家医院在 14 个疾病领域和临床专科上的综合分值。由表 4 可知，每家医院在疾病领域和临床专科都有一个分值，同时还存在一家在多个疾病领域和临床专科有分值，因此如果将医院对应的所有分值加起来，即综合分值，见图 2。根据综合分值情况，大致分为五个分数段，分别是 70 分以上（四川大学华西医院）、30-40 分（中国医学科学院北京协和医院）、20-30 分（复旦大学附属中山医院、中南大学湘雅医院）、10-20 分（中山大学肿瘤防治中心等 13 家医院）及 10 分以下（中国医学科学院肿瘤医院等 46 家医院）。

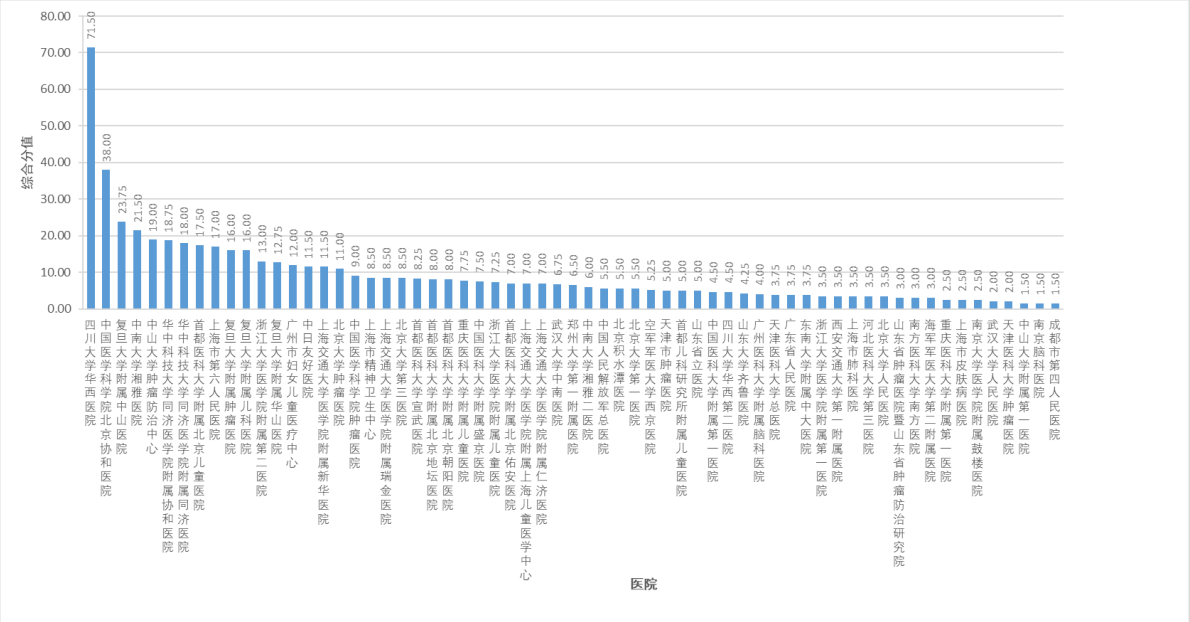


图 2 63 家医院在 14 个疾病领域和临床专科上的综合分值

第二步分析 63 家医院与 14 个疾病领域和临床专科之间的数量关系，绘制可视化图谱。对入选学科和专科排名 Top10 的医院与疾病领域和临床专科的对应关系构成 63*14 共现矩阵，利用 NetDraw 可视化工具绘制医院——疾病领域和临床专科共现矩阵的可视化图谱。结合图谱发现，四川大学华西医院（11）、中国医学科学院北京协和医院（8）、中南大学湘雅医院（6）、华中科技大学同济医学院附属协和医院（5）、复旦大学附属中山医院（5）、华中科技大学同济医学院附属同济医院（4）、上海交通大学医学院附属瑞金医院（4）、浙江大学医学院附属第二医院（4）、复旦大学附属华山医院（3）、中国医科大学附属盛京医院（3）10 家医院在 3 个及以上疾病领域和临床专科具有较强竞争力（如图 3）。

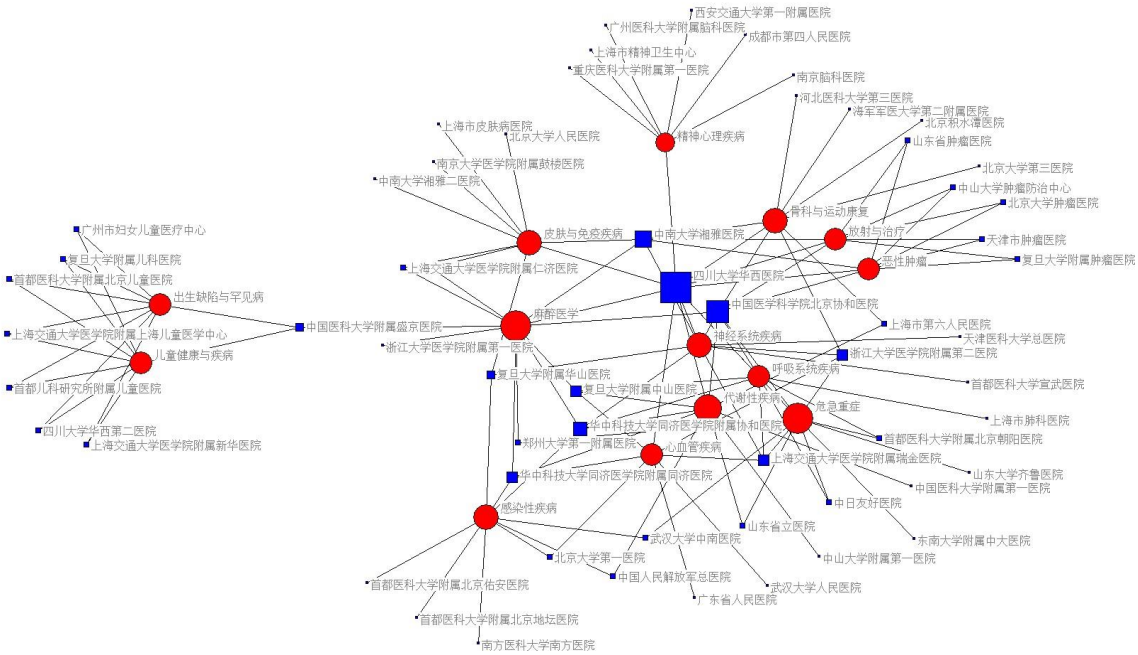


图 3 TOP10 医院——疾病领域和临床专科共现的可视化图谱

（蓝色“□”代表医院，红色“○”代表学科，大小代表疾病领域和临床专科/医院数量）

讨 论

一、基本形成国家临床医学研究中心 14 个疾病领域和临床专科的 TOP10 医院名单

借助第三方评价数据筛选国家临床医学研究中心 14 个疾病领域和临床专科 TOP10 医院，基本形成了由 69 家医院构成的 145 个名额，希望可以对国家临床医学研究中心未来布局提供参考。此外，通过计算 63 家医院综合分值以及分析医院与疾病领域和临床专科之间的关系，四川大学华西医院、中国医学科学院北京协和医院、复旦大学附属中山医院、中南大学湘雅医院、中山大学肿瘤防治中心、华中科技大学同济医学院附属协和医院、华中科技大学同济医学院附属同济医院、首都医科大学附属北京儿童医院、上海市第六人民医院为综合分值 TOP10，并且四川大学华西医院、中国医学科学院北京协和医院、中南大学湘雅医院、华中科技大学同济医学院附属协和医院、复旦大学附属中山医院、华中科技大学同济医学院附属同济医院、上海交通大学医学院附属瑞金医院、浙江大学医学院附属第二医院、复旦大学附属华山医院、中国医科大学附属盛京医院 10 家还在 3 个及以上疾病领域和临床专科具有较强竞争力。综上，四川大学华西医院在 11 个疾病领域和临床专科以综合分值 71.5 分位居第一，成为国家临床医学研究中心内未来布局综合竞争力最强的医院。

二、构建基于第三方评价数据的国家临床医学研究中心遴选方法的初衷

（一）研究结果能够为国家临床医学研究中心管理部门提供评审遴选的参考依据。

国家临床医学研究中心的组织申报与评审认定工作备受社会关注，通过材料审核和形式审查精准遴选出在临床诊疗技术、临床医学研究、领军人才和创新团队等方面具有明显优势的入围单位，对管理部门而言有一定的难度。本研究构建的遴选方法是借助第三方评价数据，可以帮助管理部门在缺乏评价指标体系及准确原始数据的前提下，快速精准缩小遴选范围。本方法基于定性定量相结合的原则，综合考量国家临床医学研究中心的申报要求和第三方评价数据的针对性、权威性，设计了基于医院学科和专科排名的相关计算方法。经验证，本方法遴选出来的医院名单已经完全覆盖了完成布局的疾病领域和临床专科、已认为国家临床医学研究中心的医院，充分说明本研究方法的准确性，同时科学利用第三方评价数据能够辅助国家临床医学研究中心评审认定工作并达到事半功倍的成效。

（二）研究结果能够为积极申报国家临床医学研究中心的医疗机构提供创建思路。

一方面，医院应精准对照国家临床医学研究中心的申报要求，寻找在临床诊疗技术、人才团队建设、资源平台建设、重大科技项目、国家级科技奖励等方面存在的短板和不足。国家临床医学研究中心是国家级科技创新平台，冠以“国家”之名的大型研究平台必定是在全国同类医学学科和临床专科研究方面达到国内领先、甚至国际先进或领先水平的，具备众多优秀的学科带头人和技术骨干和一支科研业绩卓越的研究团队，承担过国家重大科技计划项目，并获得过国家级科技奖励；建有临床数据和生物样本资源库，开展过多中心临床研究，并已经建立临床研究协同创新网络和信息平台。另一方面，医院应积极参考 STEM 排行榜和复旦排行榜两大排行榜的评价数据，根据各年度医院学科和专科排名及波动情况，合理布局、精准施策。STEM 排行榜和复旦排行榜以其长期稳定的数据来源和系统科学的评价指标而广受医学界认可，并较好呈现了国家临床医学研究中心遴选建设的重要内容。因此，对于医院及时了解全国医院学科和临床专科发展的总体概况，以及掌握医院自身学科和临床专科发展的优势逆势有重要的参考价值。医院也能通过比较每年学科和临床专科排名及其波动情况，从中寻找可以进行重点培育的优势学科和学科集群，全面实施学科战略性布局优化策略，采取针对优势学科、学科集群和扶持学科阶梯式发展的建设措施[7]。医院学科和临床专科建设无法达到立竿见影的成效，对于医院而言必定是一项任重道远的系统工程。

目前我国还有 20 个省/市/自治区尚未建立国家临床医学研究中心，经网络查询和关键词搜索发现，山东、安徽、广西和黑龙江等地已将国家临床医学研究中心的创建纳入其国民经济和社会发展

第十四个五年规划和 2035 年远景目标纲要中，作为学科建设和科技创新平台建设提质增效的重要目标。

三、本文存在的不足

一方面，本文是根据 STEM 排行榜和复旦排行榜两大排行榜的评价数据，分别对其学科科技量值和专科科研标化值的排名进行重新赋值，其目的是为了形成可比性、便于进一步同质计算并筛选出 Top10 医院，但在一定程度上也弱化了医院之间的真实差距。另一方面，因 STEM 排行榜和复旦排行榜与国家临床医学研究中心疾病领域和临床专科之间的对应关系，本文仅筛选出其中 14 个疾病领域和临床专科的 TOP10 医院，对其它 7 个疾病领域和临床专科不具有借鉴意义。

PU-133

基于 t-SNE 降维算法和 K-means 聚类的 HIV 感染者/AIDS 患者群体在线健康社区信息需求分析

钟华[1] 夏松[2] 陈红[1]

1 四川大学华西医院/四川大学华西护理学院

2 杭州海康威视数字技术股份有限公司

信息技术的快速发展和互联网的普及使在线搜索健康信息成为大部分用户寻求健康信息的主要途径之一[1–3]。因互联网虚拟空间具有便利性高、覆盖范围广和匿名性的特点[4–7]，其逐步成为 HIV 感染者/AIDS 患者的重要信息交流场所[3]。中国疾病预防控制中心于 2020 年发布的《互联网+艾滋病干预工作指南》中指出，通过互联网平台开展需求评估可快速掌握艾滋病防治工作重点方向及难点问题[3]。因此本文通过大数据挖掘及分析 HIV 感染者/AIDS 患者活跃的在线健康社区信息，了解需求，以期为开展和改进针对 HIV 感染者和艾滋患者需求的工作提供依据。

1 资料与方法

1.1 数据来源 在线健康社区主要指在开放式的互联网环境下，网络用户对健康与医疗等相关信息进行社区成员信息交互、诊疗经验分享、医疗专家咨询问诊等行为的在线健康平台[8]。目前我国在线健康社区的应用形式丰富多样，本文选择了“患—患型”在线健康社区：HIV 吧，艾滋吧及微博超话；“医—患型”在线健康社区：春雨医生、求医网、丁香医生、百度健康。

1.2 方法

1.2.1 数据收集 使用 python 的 scrapy 框架，爬取相关内容。（1）在线健康社区：HIV 吧、艾滋吧、微博相关超话采用全局爬取；（2）在线健康社区：春雨医生、求医网、丁香医生、百度健康均以“HIV”、“艾滋”、“AIDS”为关键词进行搜索，爬取包含问题的标题、描述、评论。

1.2.3 数据分析 包含数据预处理和主题聚类。数据预处理流程见图 1，因本研究中需要进行分词的文本会涉及医疗健康领域的专业术语，在采用基于 Python 语言环境下 Jieba 分词工具去除语气词、标点符号及其余无意义词汇的同时配合 CNKI 提供的医学主题词表以提高分词准确性。

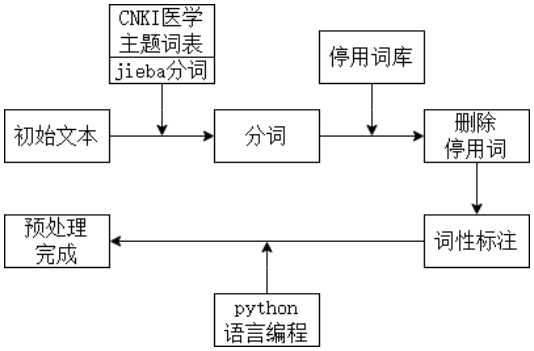


图 1 数据预处理流程

预处理后使用 glove 训练获得词向量，因训练后词向量维度过高，因此使用 t-SNE 算法将数据降维至 2 维并聚类。为获取更详细信息，使用小批量 K-means 主题聚类。词云图使用 PyEchart 的 WordCloud 子包绘制，高频词位于图像中心，字号指示词语频度。情感分析在前期完成了向量化表示、特征选择之后，使用 litNlp 下 TextCNN 情感分析模型训练及构建。

2.结果

2.1 一般特征　本研究共纳入 6 个在线健康社区，共采集 258379 条数据，分词得 40860 个词语，整体词云图如图 2 所示。高频词前十分别为加油、吃药、希望、知道、感染、朋友、医院、对象、感觉、喜欢。

图 2 词云图

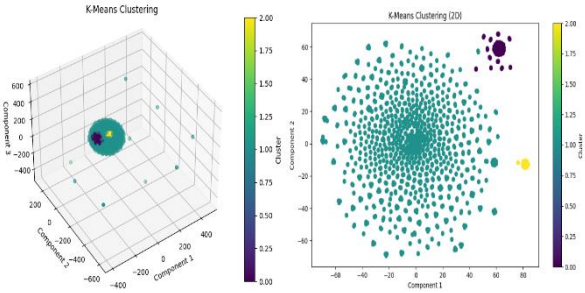


图 3 t-SNE 聚类图(左：三维；右：二维)

2.2 聚类结果

t-SNE 降维后聚类后共探测 3 个类别（见图 3），小批量 K 均值聚类后，根据获取信息，将主题归纳为社会陪伴需求、情感支持需求、健康信息需求，详细信息见表 1。①社会陪伴需求。该主题是三个主题中涵盖词语量最大的主题，可细分为线上交友、线下交友。其中线上交友相关数据量最大，也可见微信、qq 等常见社交平台软件词汇与其关联出现。线下交友多见，常与地域词同

时出现，其中排名前五地域词分别为广州、北京、山东、上海、四川。②情感支持需求。情感支持需求包括正向情感支持和负向情感宣泄，由图 4 所示情感占比图可得，正面情感支持较负向情感宣泄占比更高。③健康信息需求。健康信息需求涵盖 HIV 感染/艾滋病的疾病全过程，其中治疗阶段相关信息占比最多。

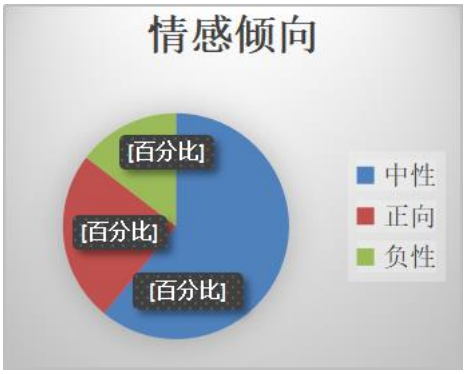


图 4 情感倾向分析图

主题	定义	Top 5 关键词	子类目	Top 5 关键词	原文范例
社会陪伴需求	寻求他人介入生活也以减少个人孤独感、无助感和脆弱感，并为日常生活中的事件提供意义[9]。	朋友、对象、喜欢、交友、交流	1.1 线上交友	朋友、对象、交友、网友、聊天	愿意的话可以加入我们微信群，相互支持，抱团取暖，交流，聊天。
			1.2 线下交友	朋友、对象、交友、北京、广州	坐标深圳，要找对象、找组织、找线下活动的联系我。
情感支持需求	与能够共情地倾听自我说话的其他人分享个人感受[9]。	加油、希望、感染、我们、朋友	2.1 正向情感支持	加油、希望、交友、家人、接受	加油！好好工作，好好生活，无论发生什么家人都会支持我们的！
			2.2 负向情感宣泄	晚上、害怕、压力、担心、歧视	有谁可以陪我聊聊天吗？很害怕，真的很害怕。
健康信息需求	患者为恢复身心健康,需要做出某些疾病相关决策时,对于自身缺乏的、不完善的和有缺陷的信	吃药、感染、生活、医院、医生	3.1 预防	预防、艾滋病、性行为、安全套、阻断	想了解一下如何预防艾滋病?
			3.2 诊断	检查、检测、高危、试纸、感染者	有没有高危行为？最好还是买试纸检测一下。
			3.3 治疗	吃药、感染、症状、副作用、体检	吃依非有什么副作用吗？

息 的 需 求 [10]。

表 1 HIV 感染者/AIDS 患者群体信息需求类目及关键词

3 讨论

通过分析 HIV 感染者富集的在线健康社区相关信息，本研究发现 HIV 感染者/艾滋病患者会借助在线健康社区以寻求相关支持。因此在大数据时代，以病人需求为导向的数据挖掘技术可在驱动临床疾病管理等方面具有一定意义[11]。

聚类后发现与社会陪伴相关的信息占比较高，这提示 HIV 感染者/艾滋病患者社交需求突出，与国内外的研究结果一致[12,13,5]。这可能与艾滋病污名化现象仍存在[14,15]且公众对艾滋病的恐惧及排斥仍占主导心理[16]，同时 HIV 感染者/艾滋病患者感知歧视及病耻感较强，会出现社交恐惧乃至自发社会疏离有关[17,18]。而本文选取的在线健康社区中的用户绝大部分为 HIV 感染者，具有相同或相似的治疗经历，彼此之间具有一定的归属感和认同感，更容易产生共情[19,20,9,21]且互联网兼具便利性、隐私性[7,22,10]。因此部分 HIV 感染者/艾滋病患者会依托在线健康社区建立社会联系，获得社会支持[20]。这提示我们 HIV 感染者/艾滋病患者社交需求应引起医护人员的重视。针对 HIV 感染者/艾滋病的线下社交需求，需引导公众正视 HIV 感染者[14]，同时可借助社会组织力量，开展同伴教育和交流分享，帮助其恢复社会交往[23]。针对其线上社交需求，在充分发挥中文互联网健康社区的社交属性的同时，平台应当加大监管力度 [12,24,25]，实时更新敏感词汇，同时利用大数据分析、机器深度学习等人工智能手段开发风险预测模型，约束用户不良行为，降低 HIV 传播风险。

此外 HIV 相关在线健康社区用户所发表文字中，情感宣泄占比也较多，其中很多都抒发了对身体健康、身份暴露、人际关系和社会适应等问题的担心，甚至伴随着对自我评价的降低等现象，这与曾雷霄等人研究结果一致[22,13,26]。其可能的原因在于 HIV 感染者/艾滋病患者面对压力来源多且宣泄途径少。除疾病本身带来的生理压力外， HIV 感染者还面对因污名化与道德层面的歧视带来的心理压力和社会性压力[22,27]。而感染者们且担忧身份一旦曝光，将会遭受异样的眼光，若告诉已披露的父母或者亲友，他们未必能够立刻给予一定的支持，甚至会遭到他们的“审视”，抑或是他们出于关心和不了解产生心理压力[22]。基于上诉因素，互联网成为了 HIV 感染者/艾滋病患者心理痛苦的重要宣泄口。与蒋金丽不同的是[28]，本研究在对爬取内容进行情感分析后发现，正向情感支持占比多于负向情感宣泄，其主要原因可能在于本研究爬取内容中患-患间在线健康社区平台内容占比多，互相加油打气是其内容重要组成部分，部分用户抒发自己心理痛苦之后，大量的病友也会回馈正向情感支持。这提示 HIV 感染者/艾滋病患者普遍存在心理痛苦乃至心理问题[29]，同时“互联网+心理干预”具有庞大的用户基础[30]。在线健康社区可发挥社会与心理支持功能，打通患者间、患医间、患者与社会组织间沟通壁垒，帮助用户提升情绪认知管理能力及社会支持寻求能力，拓展医疗资源可及性[31,28,32]。同时未来研究也可通过自然语言识别和深度学习建立心理健康高危人群预警机制，以识别潜在人群并提高心理干预实施的精准性。

由用户提出的健康相关信息分析可得具有健康信息需求的用户群体广泛，除 HIV 感染者外，还包含疑诊人群、高风险人群，且处于疾病不同阶段的不同人群其健康信息需求不同，和 ZIZZA 等人的研究一致[28,33]。高风险人群和疑诊患者更多关注艾滋病传染与临床诊断,包括艾滋病病毒检测等,对治疗等信息关注较少，而已确诊人群更多关注疾病治疗相关信息与日常生活信息，包括用药及其不良反应，复诊、职业、生育等方面[26]。这提示艾滋病相关健康信息传播不仅要针对 HIV 感染者/艾滋病患者人群,也要面向广大群众，要根据不同人群特点开发针对性和可接受性强的宣传材料。其中“患-患”型在线健康社区其内容复杂、多样，从生活闲聊到医疗知识均有涉及，但医疗相关内容占比不高且质量较差，可通过化模块设置，对患者发布的有效信息实行进一步分类，提供更为便捷的信息搜寻渠道并推广信息分享激励制度，鼓励患者分享治疗经验，同时对优质好帖加精置顶等促进患者的浏览与阅读，增强信息的易得性和识别度；此外还可吸引权威医疗机构入驻，提高健康信息的质量和可信度[16,15]。“医-患”型在线健康社区以健康咨询为主，但内容较少，权威性存疑，可通过完善平台专家认证机制、扩大宣传提高平台知名度，拓展 HIV/艾滋相关知识板块以提高其健康信息权威性、丰富度及流通性。此外市场尚缺乏针对 HIV 感染人群的信息整合的传播平台[34]，未来可进一步开发兼具健康信息传播功能的疾病管理平台以满足不同阶段 HIV 感染者/艾滋病患者的多样信息需求。

综上所述，随着互联网技术的不断发展，HIV 感染者/艾滋病患者富集于在线健康社区以寻求社会陪伴需求、情感支持需求、健康信息需求。同时在线健康社区具有成为开展艾滋病毒防控工作重点场所的巨大潜力，提供从艾滋病宣传教育、干预等防治工作的实施场地。目前在线健康社区发展自由，流通的信息多样、复杂、质量良莠不齐而社区健康平台缺乏系统性、科学性规范化的设计，因此平台尚需进一步优化，以提高在线建设社区的实用性、便利性、互动性、科学性、专业性。

PU-134

大数据时代情报跟踪服务对医学科技的影响

王刚
哈尔滨医科大学附属第四医院

近年来，我国科学技术迈入飞速发展时期，涌现出大量科学成功，医学科技体制也不断进行深入改革，传统生物医学模式已经转变为生物-心理-社会医学模式，对医学领域产生非常深刻的影响，医学情报人员也面临着巨大的工作挑战，如何更快、更准确的进行医学情报收集，满足医学科研人员情报需求成为医学情报人员急需解决的问题。在科技管理中医学情报查新工作变革势在必行，对医学情报人员工作提出更高要求。大数据时代的来临给医学科技提供了一个良好的机遇和平台，同时，医学情报人员也面临着如何在海量信息中快速、有效的将有用的数据找出来。本研究通过对大数据时代医学科技中情报跟踪服务的重要性进行分析，并就大数据时代中如何开展情报跟踪服务进行摄入探讨。

1.大数据时代情报跟踪服务的重要性

医学科技研究主要是对未知专业领域进行不断探索，基于现在已知医学科技基础不断进行研究，使其得到完善、创新。而当今时代信息化技术发展速度越来越快，大数据时代的来临方便医学科技人员及时掌握和获取数据，在医学课题研究中可以保持创新性和新颖性，还能避免低水平数据出现重复获取的情况[1]。但是因为检索手段及方法技巧的限制，医学科研人员在获取各个领域中有价值的信息数据时难度还是比较高，而图书情报人员通过情感跟踪可以为医学科研课题提供更多有价值的数据情报。

2.医学科技对情报跟踪服务的需求分析

在医学科技研究中不论是科学研究选题还是成果转化推广都需要“科研全方位情报跟踪”，也就是科研选题、立项、科研过程、成果转化、推广阶段均需要提供情报信息咨询服务。

2.1 选题、立项阶段情报跟踪服务

在医学科技研究中科研选题、立项是第一步，科研选题、立项直接关系着医学科研项目的成败。受到各种客观因素影响，以往医学科技研究工作人员在科研选题学术领域情报信息获取时主要是通过网络、纸质图书，医学科技工作者在选题时难度比较高，如果没有专业图书情报机构为其提供前沿同领域国外相关研究现状、进展资料，医学科技工作者在选题思考时往往比较局限，容易导致差物力、人力、财力的浪费，甚至错失医学科研机会。

在医学科研工作者选题、立项前，就应该向具备科技查新资质医学图书情报机构发出查新委托，评估拟立项课题科学性及新颖性，这也是科研全方位情报跟踪服务开展的第一步，是一个非常关键性的步骤[2]。通过对拟立项课题进行查新，可以为科研管理部门、评审专家提供国内外相关领域研究现状、进展的客观文献信息，在开提前方便医学科研工作人员对课题前沿情报有一个更加全面的掌握，进一步对科研项目设计进行优化，这样医学科技人员也能少走弯路，使科研立项、成果时间缩短。在确定科研选题后到立项阶段中信息跟踪更是显得尤为重要，医学科研人员需要根据课题对相关信息、新进展进行进一步信息跟踪，避免研究过程中再次发生重复情况。

2.2 医学科研阶段情报跟踪服务

虽然课题在立项时具备创新性和新颖性特征，但并不能保证课题在研究阶段仍旧保持新颖，如果在医学科研中没有进行全程情报跟踪很难确保科研技术的领先性。而医学科研阶段情报跟踪服务应从科研背景、内容、技术路线、科研结果各个方面着手进行科研全程情报跟踪。

2.2.1 科研背景的情报跟踪

科研背景情报跟踪又可以细化为选题立项、开题科研背景，立项背景主要涉及科技查新，通过科技查新为项目新颖性提供更多可靠依据。许多医学科技研究工作人员在开题后往往没有重视国内

外相关研究背景调查，但其实通过开题后科研背景调查非常有必要，可以避免医学科技研究工作人员出现项目低级重复的情况，可以进一步确保科研课题“含金量”[3]。因此，医学科技研究工作人员在确定课题时可以向图书情报人员了解国内外同领域研究背景文献资料，不断对研究内容进行完善更新，提高医学科学研究路线的可靠性，同时，也能为科研步骤拟定提供更多可参考已经。

2.2.2 科研内容情报跟踪

在科研课题立项完成后，医学科技研究工作人员应及时向图书情报人员了解课题科研进度，不断对其中存在的问题进行调整，随时注意国内外相关领域新的研究动态情报，图书情报人员也可以将以上数据整理为二次文献，方便医学科技研究工作人员参考[4]。从而确保医学科技研究工作人员思路开阔性，为课题研究内容新颖性提供更可靠保障，如果有需要，在医学科研实践过程中图书情报人员也可以参与其中，负责对各种相关信息资料进行收集，帮助医学科技研究工作人员进行整理、分析，并完成深层次加工，为医学科技研究工作人员提供更全面、准确、及时、具有针对性的情报资料，进一步提升医学科技研究工作人员科研效率和治疗[5]。通过情报跟踪服务可以确保信息同步性、及时性，医学科技研究工作人员能及时了解到更新的科研信息，不断对研究内容、方法进行完善、修整、充实，确保课题独创性、新颖性，提升课题优势及竞争力。

2.2.3 技术路线情报跟踪

要想顺利完成医学科技研究内容必须对研究方法、技术路线进行合理的设计，研究方案、技术措施设计包括理论分析、工作步骤、实验方法计划与安排。在制定研究方案的阶段，图书情报人员可以对信息资源进行收集、分析及研究，帮助医学科技研究工作人员对以往同领域他人研究经验、成果进行总结，为研究方案的设计提供更加正确具有创新意义的信息。而进入到实验研究、实验观察阶段，图书情报人员在情报跟踪中需要做到快、准、广、新，为医学科技研究工作人员收集同类课题研究手段和方法，使本组医学科技研究工作人员获得得多的启发，不断对实验手段、实验技术、观察方法进行完善和改进，这样研发过程也能得到进一步完善和深入，可以确保研究成果新颖、细致，科研活动可以走在该领域最前沿，取得突破性研究成果[6]。另外，图书情报人员还可以对国内外科研课题相关研究方法、技术路线等信息情报进行收集，分析整理成三次文献，这样医学科技研究工作人员能从资料中获得科研思路启发，不断对原定技术路线进行完善、修订，从而确保科研过程顺利开展。

2.2.4 科研成果情报跟踪

科研成果阶段情报跟踪主要是由科技查新机构客观的对科研成果创新性进行评价。图书情报人员充分发挥自身专业检索能力，检索国内外医学专业数据库系统文献，对和课题相关文献资料进行查找，综合分析该课题，客观评价课题成果创新性，判断课题成果是否存在重复问题，并完成书面报告。科研成果阶段情报跟踪主要是科研成果向社会效益、经济效益转化情报咨询服务，科学评估课题申报成果奖。

3.大数据时代如何开展新型情报跟踪服务

3.1 大数据时代跟踪服务特点分析

大数据时代情报跟踪服务也就是信息跟踪服务，是医学信息咨询服务中不可或缺的一部分，在大数据时代信息跟踪服务发展迅速，已经成为新型智能产业，可以掌握更多信息，并为医学科研工作人员提供信息咨询服务。如今，医学科研越来越重视信息情报时效性、专指性及决策性，对决策情报、竞争情报需求量也越来越大，在大数据时代情报跟踪服务中更是要注意以下几点[7]：（1）重视情报跟踪服务的针对性。图书情报人员在提供情报跟踪服务必须针对一个课题、一个项目进行情报跟踪，这样才能确保专指性。（2）重视情报跟踪服务的预测性。图书情报人员在提供情报跟踪服务必须根据课题动态进行预测，提高结论的可靠性。（3）重视情报跟踪服务的决策性。图书情报人员必须确保所提供的情报资料可以为医学科技研究工作人员提供决策性依据。

3.2 将信息跟踪服务和现代信息技术联合起来

3.2.1 借助网络、数据库追踪科研课题

通过数据库可以快速实现文献检索，图书情报人员可以帮助医学科技研究工作人员完成检索，并未医学科技研究工作人员提供文献检索指导。特别是近年来，信息化网络技术发展越来越快，图书情报人员已经可以通过 Archie、FTP、Gopher 等多种途径完成医学科技研究课题网络信息检索，还能通过 WorldWideWed 对全球信息进行获取[8]。通过 Internet 中各种种类繁杂医学文献数据库及医学网址收集所需要的信息，在同一数据库上借助多网址访问技术，可以完成网上文献数据库二次文献检索，获取多所需要的资料，并根据主体进行加工整理，从而为科研课题提供重要的信息

资源。

3.2.2 帮助医学科技研究工作人员获取原文

信息跟踪服务包括为医学科技研究工作人员提供课题研究所需要的原始文献，图书情报人员可以通过对图书馆馆藏目录、联合目录进行刊名、书名、著者研究，可以帮助医学科技研究工作人员查找到原文。通过联合目录还能获取到多个图书馆书刊收藏情况，方便医学科技研究工作人员其他馆收藏的文献资料，借助馆际互借途径获取到原文。另外，借助网络技术也可以帮助医学科技研究工作人员获取到原文，图书情报人员可以借助网络和文献作者取得联系，或者通过网上电子期刊对原文进行查找[9]。在 OCLC、Uncover、Medline 等大型文献数据库中也能获取到原文。

4.大数据时代医学情报跟踪服务需要注意的问题

4.1 重视医学科技研究工作人员与图书情报人员协同作用

医学科技研究工作人员应紧跟时代步伐，利用现代信息技术，掌握信息获取手段、技术，重视医学科技研究工作人员与图书情报人员配合的重要性，积极寻求图书情报人员帮助，降低文献检索误差。

4.2 打造网络互动平台

借助互联网技术，通过钉钉、QQ、微信等方式实现医学科技研究工作人员与图书情报人员双向互动，及时对课题研究进展、理论、实践进行沟通。

4.3 定期进行情报跟踪

医学科技研究工作人员与图书情报人员应随时联系，密切关注科研近况，图书情报人员可以充分发挥自身检索优势，收集课题研究相关领域最新信息，及时为医学科技研究工作人员提供情报，确保医学科技研究工作人员能随时掌握专业领域国际国内科研动态。

4.3 实时信息推送

图书情报人员应根据医学科技研究工作人员需求对各类中外文数据库、纸质、互联网自愿进行查阅，对专业领域报道前沿理论、方法、技能进行搜集，并进行整理、分析，深入挖掘，及时向医学科技研究工作人员提供信息推送微服务[10]。

小结

大数据时代情报跟踪服务是医学科研必要需求，在信息化进入飞速发展的时期，图书情报人员只有积极肩负自身职责，发挥能力，借助定期跟踪、信息推送服务，根据医学科技研究工作人员个性化定题服务，密切关注课题研究进展，总结阶段性数据，每个阶段为医学科研工作人员提供全方位情报跟踪服务，才能促进我国医学科研进一步发展。

PU-135

不同角色预设下的人工智能模型文献筛选

刘宇扬[1] 田倩男[1] 朱伟林[1] 周仪[1] 伍丽群[1*]
深圳市卫生健康发展研究和数据管理中心

1. 研究背景

近年来，全世界出版的研究成果迅速增加，这使得传统的研究综合方法面临挑战。一方面，大量的文献发布速度要求研究综述更快地更新；另一方面，大量的科研输出已超出了人工筛选的能力。文献检索是研究者开展一项新领域知识或是总结当前研究和知识进展的一项重要手段，[1] 并且良好的文献检索过程及系统总结助于深入回答特定的研究问题。但文献检索大多都需要研究者系统地评估和总结当前的知识，并且会极大的受到研究者的主观影响从而对研究过程造成影响。同时进行系统综述的第二步需要阅读大量文章标题及摘要，受发表文章质量的影响，通常这一步需要花费大量的时间与精力。[1-3]

现有的研究指出，只要正确使用并确保方法论质量，AI 工具如现有的自然语言模型(natural language processing，NLP)对当前的提供了加速系统评审方法，并在多个方面等到了证实 [4-6]。Perlman-Arrow 等利用 NLP 工具筛选 SARS-CoV-2 血清流行病学相关的文章，研究结果发现每篇摘要的筛选时间减少了 45.9%，总体筛选过程减少了 70%，并且筛选的准确度与人工筛选相当 [4]。Oude 等人 [5] 使用人工智能工具筛选文章的结果显示该技术在系统性评审中大幅减少筛选时间的潜

力。但多数的研究聚焦在文献筛选的快捷性，研究思路基于已经发表的文章结果或者是人工判断的结果进一步验证人工智能筛选的文献结果，人工智能筛选的逻辑仍然是“黑匣子”。[1,3]。并且 NLP 工具筛选文献多基于特定研究者提前设置的参数（如角色预设或排除纳入参数）等筛选文章，进一步未考虑不同角色预设下对于同一批文章的纳入标准筛选的结果。

本文目的拟基于同一数据集，基于现有的自然语言 NLP 模型，验证不同预设角色的前提下对同一批文献筛选结果，进而说明不同角色预设对于文献筛选的区别，为后续进行大量系统综述或文献筛选研究提供新的角色预设思路。

2. 研究方法

2.1. 数据准备

基于 Pubmed 数据库，通过关键词如“生态学”， “队列”等筛选（详见表 1）进行文献筛选，时间截至在 2023 年 8 月，最后纳入文献共 4494 篇。

表 1 文献数据集筛选过程

序号	搜索检索词			检索量
1	"ecologic*"[Title/Abstract]			160,636
2	"cohort"[Title/Abstract] "longitudi*"[Title/Abstract]	OR	OR "follow-up"[Title/Abstract]	2,090,119
5	1 AND 2			4,494

2.2. 模型参数与角色设定

基于现有开源的 NLP 模型及 API 工具，在 python 工具下进行识别，给 NLP 应用程序接口的提示如下[7]：

> ***System:*** *You are an ecologist/ecologist/policy maker would like to research articles related to ecology and epidemiology. Now you are searching for articles based on your aim.*
>
> ***Instructions =*** *"I will now provide you with information about an article, including its title, abstract, and keywords. Your task is to decide whether it aligns with your research objective. You will evaluate the titles and abstracts of scientific papers for inclusion or exclusion. If you believe the article is relevant to your research, please type 'Decision: included' and provide a detailed reason (Reason_decision:). If you do not consider the article pertinent to your research, type 'Decision: excluded' and provide a detailed reason (Reason_decision:). Additionally, please provide five keywords that describe the article (Keywords:). Finally, rate the relevance of this paper to your research on a scale of 0-5 (0 meaning not related at all, and 5 meaning highly related; for example: Rating: 0), and provide a detailed explanation for your rating (Reason_rate:). Use 'NA' if there's no information for a section. Please stick to this format: Decision:/reason_decision:/keywords:/rating:/reason_rate: and avoid adding extra details.",*
>
> ***Title:*** *{title}*
>
> ***Abstract:*** *{abstract}*
>
> ***Decision:***

其中设定三个角色类型：流行病学学家（Epidemiologist, EPI），生态学家（Ecologist，ECO）与政策制定者（Policy maker，PM），“Decision:”是指是包括还是排除这篇文章“Results”是指包含文章的打分以及排除纳入原因等。

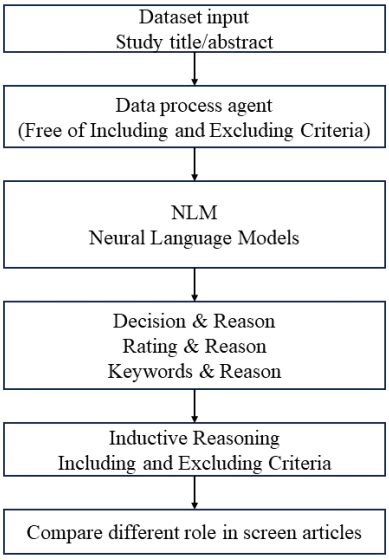


图 1 自然语言模型筛选流程

2.3.数据分析

针对不同预设橘色下的文章观察排除和纳入的文献结果，并比较三个视角下选择的文章重合程度以及排除纳入结果。所有结果分析在 Python 3.2 以及 R 4.3.2 中完成。

3. 结果

3.1 文章排除纳入结果

不同角色的排除纳入结果见表 1，ECO，EPI 和 PM 分别纳入 4058 (90.3%)，4001 (89.0%)，4062 (90.4%)，卡方检验 P = 0.057，三者间的文章纳入结果无显著统计学差异。不同年份下的文章如图 2 所示，ECO，EPI 和 PM 的纳入率也相近，均无统计学差异（卡方检验 P > 0.05）。

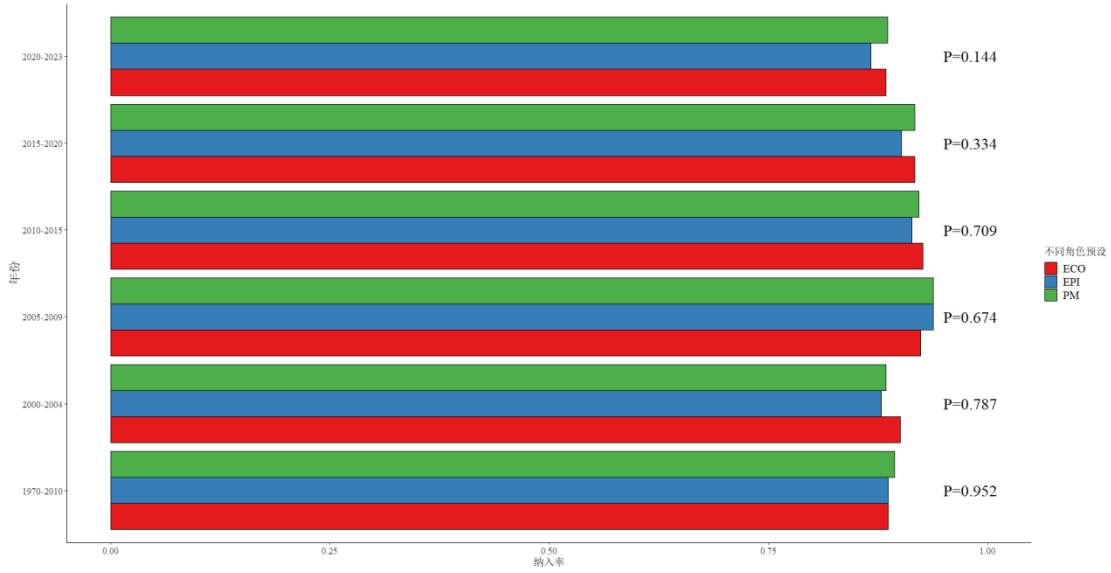


图 2 不同角色预设下的文章纳入率

3.2 文章排除纳入一致性

如表 2 所示，三种角色预设下的文献一同纳入 3873 篇文章，共排除 307 篇文章，不一致的文献共有 314 篇。两两比较结果显示 EPI 与 PM 排除和纳入卡方检验 P 值< 0.034,说明两种角色预设下的文章排除纳入结果有所差异。

表 2 三者文献纳入数比较

文献数	ECO	EPI	PM
N = 3873	Included	Included	Included
N = 307	Excluded	Excluded	Excluded
N = 314			-
N = 57	Included	Included	Excluded
N = 82	Included	Excluded	Included
N = 49	Excluded	Included	Included
N = 58	Excluded	Excluded	Included
N = 22	Excluded	Included	Excluded
N = 46	Included	Excluded	Excluded

三者的 kappa 系数 0.743（P < 0.05）,三者对于文献的筛选一致性高度一致。三种角色预设下，Kappa 如表 X 所示，Kappa 系数均高于 0.73（P 值均小于 0.05）。

表 3 三者文献纳入数比较

Kappa	ECO	EPI	PM
ECO	-		
EPI	0.761	-	
PM	0.732	0.736	-

4. 讨论

本文主要探讨了不同预设角色下对于统一数据集文献筛选的结果。研究发现不同角色情景预设下对于同一文献数据集筛选结果相近（kappa 系数为 0.743）。

现有研究已经发现了提示方法（Prompt Learning）在自然语言模型中的对于模型训练以及识别具有改善效果，而角色预设也是提示方法应用的一种。[8] 现有的 NLP 模型多是基于输入关键词或者设定相应的标准，NLP 模型基于相关性、字符出现次数进而判断其相关特性。[9,10] 本研究通过预设三个研究者的提示方法进行模型训练，在一定程度上改善了结果。并且三者对于同一文献数据集的结果不完全一致，展示了提示方法的有效性。

现有文章多基于节省时间的角度去使用 NPT 工具筛选文章，但未考虑不同条件下的 NPT 筛选的文章真实性。现有的文章已经证明了 NPT 工具可以有效地减少所花费的时间，并且在一定程度上开展大型的文献筛选工作。对于文章的识别更多的是对于关键词的识别来判断其是否符合排除纳入标准 [11]。本在筛选过程中并未设定排除纳入标准，文章的排除全由 NLP 模型与文章本身的标题与摘要所呈现的内容决定，这在一定程度上解释了三种角色预设下识别出来的文章 Kappa 系数达到 0.73 的原因。并没有考虑未角色设定的筛选结果与角色预设之间的差别。系统综述一般由相关背景的研究者进行，所以在比较未设定角色和角色预设之间的差别无实际意义。

本研究的优势在于首次考虑了在文献筛选的过程中考虑到了角色预设对于文献筛选的影响，并且验证了基于同一文献数据集下，筛选的结果的一致性和差异性。但本文也存在一些限制。首先本研究考虑到了三个角色预设，但并未考虑三者之间的交叉性。更多的限定词有助于更加精准的进行文献筛选。其次本研究仅针对生态学与队列研究相关的文章，文章内容相对局限单一，未来进一步可以针对大量文章，基于不同的角色预设进行文献识别与分类。最后本研究仅纳入了英文文章，这也是因为生态学与队列研究相关的文章尚未有中文文章，未来将进一步扩大研究面，测试不同语言

预设背景下对于文献筛选的结果影响。

5. 总结

本研究探讨了不同预设角色对于统一数据集文献筛选结果的影响，发现不同角色情境下的预设条件下，对于同一文献数据集的筛选结果具有相似性（Kappa 系数为 0.743）。未来研究可进一步探讨不同条件下的文献筛选，以提高准确性。此研究首次考虑了角色预设对于文献筛选的影响，为进一步研究提供了有益的参考。

PU-136

基于医疗社交媒体用户评论主题及情感的药物推荐策略研究——以高血压为例

邰杨芳* 魏宇虹 吴娟 付颖 杜沛泽 姬航琦 安鹏伟 陈天麒
山西医科大学

目的 挖掘国内外高血压常用药物的患者观点及情感倾向，提出高血压药物推荐策略，为高血压药物选择提供指导。
方法 本文提出一种基于 LDA-TF-IDF 和 BERT 的药物推荐方法。将 TF-IDF 算法与 LDA 主题模型结合，以挖掘国内外网站的高血压药物用户评论潜在主题，采用 BERT 模型训练评论数据，预测特定药物情感分类标签，分析特定药物整体及分别在各主题下的情感倾向，提供药物选择策略及建议。
结果 共获得国内外平台 12566 条评论文本，通过主题分析得出六大主题（高血压药物效果、药物对心血管方面的影响、药物副作用、药物对日常生活的影响、疼痛与身体不适、剂量调整与身体感受）。总体情感分布相对均匀，各主题中，药物效果主题情感倾向相对积极，其他主题情感倾向偏消极。
结论 高血压药物选择应关注主题及情感视角，建议药物选择从治疗效果和心血管健康，副作用、不适及适应性，疼痛缓解和不适症状，个人感受和剂量调整四个方面情感倾向进行综合考量。

PU-137

浅谈医院图书馆在助力临床及科研发展服务中的作用

顾显伟 何丽
上海市儿童医院 上海交通大学医学院附属儿童医院

引言：近年来，在多方面利好因素的影响下我国社会经济的发展取得了巨大的成绩，与此同时科学技术水平随之不断提升，大量的新兴学科知识文献应时而生，为各个学科知识综合水平的提升起到了积极的作用。在医学领域中医院图书馆在医院科研活动、临床研究工作以及医院医护人员专业水平提升和综合能力发展方面都具有重要的作用。鉴于此，这篇文章主要针对医院图书馆在临床以及科研信息服务领域中的实践运用进行深入分析研究，希望能够对我国医疗事业的未来良好发展有所帮助。

1 医院图书馆特色服务的思想积淀

1.1 纸质文献资源特色

在历经了多年的累积，普通医院图书馆内汇集了大量的印刷类的医学文献资源，这些文献资源可以说是医院图书馆专业化、特色化发展的重要条件，这类文献资源在我国医疗事业的发展方面也具有重要的保障作用。以往老旧的纸质版的文献资源与当代电子文献资源二者存在明显的差别，但是二者也是相互交融相互补充的，纸质版文献的信息价值是电子文献所不具备的。从某种层面上来说，纸质版文献资源是电子资源形成的前提条件，并且也是电子信息资源的有效补充，纸质版文献

资源是进行信息资源开发的关键基础资源。

1.2 多媒体数据库资源特色

现如今，我国医院图书馆内的医学情报服务工作通常都历经了长时间的发展和积累，不但拥有丰富的印刷型文献资源，并且也为医院科研工作的实施提供了需要的电子文献。我国医院知识资源总库 CH-KD 已经在国内大部分医院加以实践运用，以 CHFD 为代表的医院数字图书馆有效的发挥出了医院图书馆医学文献传播的作用，为用户可以提供实时文献查阅和检索的服务。在社会不断发展的过程中，各个医院图书馆在综合了各方面实际情况和需要的基础上，创设了满足医院自身发展需要的医学文献数据库，在数据库中将本院内病人的基础信息以及病情治疗过程中所涉及到的信息进行统一的收集和存储，这类数据库能够为医生临床诊断工作的实施给予必要的辅助，也可以使得医生对医疗全过程进行全面的监督和管理，这类数据库数据结构模式相对较为复杂，可以说是传统数据库资源的一种补充。当下我国存在大量的可以共享的医学资源，医院图书馆应当切实的结合医疗科研发展的实际情况和需要，切实的对医学文献信息数据库进行优化完善，为医疗事业的发展研发出更多的医疗专业技术给予需要的高品质的文献信息。现下，国内所正在创建的医学数字图书馆多媒体信息管理系统具有较强的综合性和实用性，医务工作人员可以结合自己的实际需要来选择使用适合的信息资源，主要涉及到各类文献以及各类视频资料，能够从多个角度对多媒体视频资料加以良好的运用。数字图书馆的创设对于促进我国医学信息服务行业的稳步健康发展，提升医学文献资源的利用效率，深入挖掘服务潜能都可以起到良好的助动作用。

1.3 突出以人为本的特色

人才资源可以说是各个不同事业稳步发展壮大的重要基础，医院图书馆要想将自身所具有的优越性充分的发挥出来，那么还需要对高素质人才资源进行合理的调配。首先，医院图书馆需要对复合型人才的培养工作加以侧重关注，医院图书馆工作人员不但需要具备良好的专业能力，并且还应当掌握充足的医学知识。其次，医院图书馆还应当积极的将医院专家资源所具有的优越性加以切实的利用，组建医院专家团队，为医院图书馆信息服务的发展提供需要的支持。

2 当前医院图书馆为临床及科研服务现状分析

2.1 意识不足，限制了医院临床和科研发展

医院在医疗临床服务以及医学科研活动方面都拥有十分重要的作用，在医院快速发展的过程中，医院临床研究与科研活动对于图书馆的服务工作水平提出了更高的要求。之前医院图书馆对于临床服务给予了更多的关注，将大量的注意力都放在了实现教学目标，提供文献资源方面，所以服务意识十分的被动。大部分时候，图书馆都是被动的运营，甚至逐渐的丧失了原有的图书馆服务和科研活动以及临床研究的作用。其次，大部分医院的图书馆往往都具有无法深入进行科研工作的问题，甚至很多医院图书馆工作人员认为医院深层科研工作都是由专业的科研工作人员参与的，而并非自己的工作职责，这样都会对医院临床和科研工作的实施以及未来发展形成一定的制约。

2.2 信息服务水平不足，管理人员素质有待提高

当下，医院图书馆管理工作人员团队结构存在明显的不合理的情况，管理工作人员专业水平较差，综合能力高低不齐，不具备完善的专业知识。再有，很多医院图书馆管理工作人员年龄相对较大，尽管拥有较强的工作责任心，但是自身的专业技能水平较低，并且也没有积极的对科研课题进行综合分析，信息化服务长期以来都是我国图书馆建设中最为重要的一个基本要求，但是因为受到基础条件的限制，所以造成医院图书馆信息服务整体水平较为地下，最终导致图书馆无法为临床研究以及科研工作的实施提供必要的辅助。预习同时，医院规模逐渐的扩展，科研活动数量不断增加，医院图书馆如果没有创设完善的信息化服务机制，那么必然会对科研工作以及临床研究工作的实施形成诸多的限制。

2.3 数字化、网络化的发展水平较低

医院图书馆管理工作牵涉到的层面较多，具有较强的复杂性和综合性。首先，图书馆管理工作中所需要运用到的数据量十分的巨大，在医院规模逐渐扩展，员工数量不断增加的形势下，图书馆服务对象范围随之不断蔓延，再加上文献数量的不断引入，图书馆管理系统中所融入的数据数量不断增加。服务工作可以说是医院图书馆管理工作中的重点，针对服务工作加以优化创新是保证医院图书馆管理工作整体效率和效果的重要基础，并且也是信息资源服务的重要条件。但是就现如今情况来说，当下医院图书馆在新型科技的引入方面相对较差，医院图书馆系统综合性能较低，所以为了切实的满足医疗事业的发展需要，高效的将网络信息资源加以运用，那么还需要我们切实的提升医院特色图书馆信息管理工作的整体水平，不断的推出新型的服务模式。

3 医院图书馆在临床及科研信息服务中的作用

3.1 医院图书馆是临床及科研信息保障中心

在科学技术飞速发展的带动下，计算机技术以及信息化技术水平得到了全面的提升，将诸多先进的科学技术引用到医院建设之中在促进实践工作效率和效果方面起到了良好的作用。当下，医院图书馆建设工作将更多的关注力都集中在医学文献的收集、整理以及存储管理方面，从而为医院临床以及科研工作的有序高效的开展带来了良好的助力。医院在组织实施临床科研工作的时候，医院图书馆的作用是非常重要的，切实的将网络技术加以运用，可以对医院图书馆医学文献的获取方式进行充实，从而可以在最短的时间内赶越需要的国内外数据文献。再有，医院图书馆还需要切实的完成对各类文献进行分类处理，对于不同类型的文献采用有效的方法来加以管理。对于所获得的文献资料，医院图书馆应当积极的对文献进行综合加工，从而为临床科研工作的实施给予必要的保障。医院图书馆对于所有相关文献资料实施统一的整理和加工，这样才可以为后续的资料使用创造良好的基础，为临床和科研工作的实施给予高品质的服务。

3.2 医院图书馆为临床及科研发展提供动力

就医院图书馆建设发展实际情况来说，最为重要的就是需要秉承立足实践、面向科研的原则，从多个角度来不断提高自身的服务理念，加大力度全面实施建设工作，确保医学文献数据查找、整理和存储工作具有良好的实用性。医院图书馆也应当积极的与医院内部临床科研工作人员进行交流，对于前沿研究发展趋势进行全面的掌握，随后实施相关数据收集工作，尽可能的保障医学文献资料能够为科研工作给予必要的帮助。其次，医院图书馆还需要切实的对现有图书文献资料进行合理的种类划分，结合各个不同医务岗位以及临床科研的实际需要，切实的对医学文献进行整合利用，为医院各项实践工作的实施以及发展提供高品质的服务，推动各项工作都能够按照既定的流程有序的开展。医院在实施临床科研工作的时候，不断地提升研究层次，将临床与实践工作进行充分的融合，保证其与科研未来发展方向保持良好的统一，医院图书馆也应当结合研究方向来给予切实的服务，这样对于促进医院临床和科研工作的稳步持续发展都是非常有帮助的。

4 提升医院图书馆为科研服务的措施和方法

4.1 做好信息收集整理和传递工作

医护人员在实践工作开展中不但需要重视工作的效率和质量，并且还会承担巨大的工作任务，所以高水平的医学文献数据整理工作的实施是医院图书馆工作中的重点内容。医院图书馆应当不断地充实医学文献数据的数量，并且还需要针对期刊类文献资料加以重点整理，这样才可以为读者的实际需要提供帮助。其次，对于各个科室实践工作的需要，不同专业临床科研人员对文献的需要加以全面的了解，对医院图书馆文献资料进行补充，从而切实的推进医院临床以及科研工作的有序开展。再有，医院图书馆应当切实的设立数据库，保证医院文献资料具备良好的针对性，确保各类文献收集的完整性，从而为医院临床以及科研工作的实施起到积极的促进作用。

4.2 加快医院图书馆硬件建设

全面的落实医院图书馆基础环境改造工作，转变以往图书馆老旧落后的思想理念，将以人为本的思想加以实践运用，定期对陈旧的书籍进行替换，充分结合实际情况和需要对图书馆流程进行优化完善，将自助借阅系统加以实践运用，为读者创造良好的阅读环境。

4.4 加强医院图书馆信息系统建设

在当前大数据时代中，医院图书馆也应当紧跟社会发展趋势，将前沿的数据技术加以积极的运用，这样才可以保证服务工作整体水平的不断提升。医院也应当将图书馆信息系统建设划分到医院信息化建设综合框架中加以整体规划，适当的将国内外先进的图书馆信息管理系统以及相关数字化技术加以合理的运用，充分的结合医院数据中心建设的实际情况，创设出完整的一体化信息服务机制，从而为医院职工利用内网来进行需要的信息数据的查阅工作。为了适应当下碎片化阅读的实际需要，医院可以利用手机当做移动终端，为图书馆的建设工作的实施起到良好的辅助作用，尽可能的为职工提供高品质的服务。图书馆在 APP 客户端与数据库进行连接，保证医护人员能够随时随地的进行资料的阅读。图书馆也可以运用客户端结合实际需要来为客户推送适合的读书目，并且可以进行对书目文献信息需求调查问卷活动，切实的对图书馆服务工作进行优化和创新。

4.5 提高图书馆员综合素质

医院图书馆在医院人员和医学文献之间起到了桥梁的作用，所以医院图书馆工作人员综合能力以及服务质量的提升是具有较强的现实意义的。图书馆工作人员应当积极的利用各种途径和方式方法增强自身的专业水平和实践能力，并且在工作中也需要具备良好的工作责任心。在当前信息化的

时代中，图书馆工作人员应当对先进的科学技术操作技能进行学习，从而为实践工作效率和质量的提升起到积极的作用。

4.6 全面树立为临床科研服务的主体导向

医院图书馆需要担负起临床科研服务的职责，对于自身职责加以全面的掌控，切实的扭转以往老旧落后的服务理念，全身心的参与到临床科研工作之中。对于医院学科重点以及未来实际发展趋势进行全面的掌握，结合实际情况和需要对文献收集的方向和重点进行调整，并且及时发布临床医学动态信息。

4.7 发挥医学信息在管理决策中的作用

医学科技成果的测量标准往往是以 SCI 收录情况为评价依据。我国医学科技成果的数量位居世界第一，但我国卫生体制改革中的软科学实力较差，这也是非常可怕的。我国在卫生信息化建设及卫生政策研究方面，与世界发达国家相比，差距甚远，关于卫生政策的研究仍是弱项。目前。在我国医学领域中医患纠纷事件频频加剧，视营利为第一目的，把救死扶伤只是放到了嘴边，在很大程度上丧失了实际行动。这些行为完全背离了“以人为本”的医疗规律。现如今，医疗市场的发展是迅猛而复杂的。医院图书馆要利用专业知识，分清当前医疗卫生事业中改革的真正目标与方向。要明确把握医疗卫生事业发展的核心任务，把握当今医疗发展的具体事实，科学而有目的地搜集该领域有关文献信息，对所需要的信息进行科学加工整理，以用文献专题调研、文献报告、文献摘要、文献综述等形式，提交有关部门或领导，为科学决策提供依据。

4.8 加强馆际合作以提升信息综合保障水平

医院图书馆的信息服务能力和信息支撑能力一定程度上影响着医院的医疗科研创新能力，是医院核心竞争力的有力体现。要提高医院图书馆的信息服务水平，一方面需要强化自身建设，另一方面也要加强与国内医学文献信息服务专业机构，以及高校、知名医院的图书馆等之间的馆际合作。在条件许可的情况下，可建立图书馆资源共享联盟，借助互联网实现馆藏资料的优势互补和资源共享，有力提升图书馆的综合信息服务水平。

4.9 充分利用医院数字图书馆为科研提供平台

随着社会的不断进步，医院图书馆急需完善自身建设，不断丰富馆藏文献数据，以此促进临床以及科研工作的发展。通过构建专题数据库，以此为医护人员提供更加快捷和及时的信息服务，比如说可以建立电子阅览室，免费供医护人员使用。不仅如此，不同的科室通过网络技术，实现数据信息的实时传送，帮助医护人员第一时间获取数据资料。

结语

总之，如果医院想要获取持续稳定的发展，必须提高自身临床科研能力，以此推动自身向现代化综合性医院方向发展。医院图书馆作为科研工作资料信息的主要获取场所，需要不断提升自身服务质量，为医院科研工作提供有力的保证。此外，医院全体员工也需要共同奋进，强化自身素质，推动医院图书馆事业的稳步发展。

PU-138

医学图书馆医学情报服务发展现况及趋势

杨山石，王常颖，何阿妹，宋捷，朱琳，马骋，沙小苹（通讯作者）
上海市卫生和健康发展研究中心（上海市医学科学技术情报研究所）

1 引言

中共中央和国务院印发《知识产权强国建设纲要（2021－2035 年）》提出了全面提升知识产权创造、运用、保护、管理和服务水平，充分发挥知识产权制度在社会主义现代化建设中的重要作用[1]。随着医学科技的发展和信息技术的快速发展，医学情报已经成为医疗、科研、教育等多个领域的关键资源，我国医学情报机构受到了巨大的冲击[2]。医学图书馆作为医学情报的重要存储和传播机构，其角色也日益凸显。如何适应信息技术的发展变革，提供更加高效、个性化的服务，是当前医学图书馆和信息科学领域亟需解决的问题。本研究旨在全面分析医学图书馆的发展现状和未来趋势，重点从医学科技查新、情报研究、知识产权服务三方面的通过文献研究和实地调研的方法，总结出机构发展存在的一些问题并提出相应的发展趋势及建议。

2 国内重点高校医学图书馆能力评估

2.1 纳入研究的高校医学图书馆

本研究聚焦于中国医学科学院“医学院校科技量值（STEM）” 前 10 的高校医学院[4]的图书馆，包括：北京协和医学院、北京大学医学部、上海交通大学医学院、复旦大学上海医学院等。10 所高校的医学院图书馆的数据库资源和电子刊物资源情况如表 1 所示。

表 1 各医学院图书馆电子资源及人员编制概况

图书馆机构名称	数据库资源	电子刊物资源
北京协和医学院医学信息研究所	国内外医药领域专业数据库 100 余种	2200 余种独家刊
北京大学医学图书馆	电子数据库 97 个	中外文电子期刊 61409 种
复旦大学图书馆医科馆	500 多种各专业数据库	有记录 7872 条期刊
上海交通大学医学院图书馆	医学类电子文献数据库 35 个	医学及相关学科电子全文期刊 4000 余种（中文期刊约 1500 种）
首都医科大学图书馆	86 种文献数据库（官网显示）	电子期刊 5 万余种，电子图书 129 万余册
四川大学华西医院图书馆	/	电子图书 18 万余册，生物医学类外文电子期刊超过 5000 种
中山大学医学院图书馆	十余种（中山大学总馆）	医学类外文电子期刊 7000 多种
浙江大学医学图书馆（农医馆）	各类文献数据库（含子库）779 个（浙江大学总馆）	中外文电子图书 244.6 万册，中外文电子期刊 8 万余种
华中科技大学同济医学院附属协和医院图书馆	190 多种	1 万多种中外文生物医学电子期刊
中南大学湘雅医学院图书馆	中文数据库 80 种 外文数据库 121 种 （中南大学总馆）	2.5 万余种

注：1）“/”官方表示未披露明细信息。2）北京协和医学院图书馆，亦挂牌“中国医学科学院医学信息研究所”、“北京协和医学院医学信息研究所”。3）部分数据采用总馆指标。

医学数据库与电子资源为各大医学图书馆建设的重点，多家医学院图书馆提供的数据库超 100 种，为科技查新、情报研究、知识产权服务提供了坚实的基础。各家图书馆电子刊物资源亦拥有广泛覆盖，但具体期刊种类和数量之间存在一定差距，多的医学图书馆可达数万种期刊。

2.2 查新业务

各医学院如书馆查新业务与查新资质情况如表 2 所示。

表 2 各医学院图书馆查新业务概况

图书馆机构名称	查新业务	查新资质
北京协和医学院医学信息研究所	申请科技立项；科技成果鉴定、成果奖励；发表论著；申请专利；其它（如博士论文开题、评职称等）[3]	科技部和(原)卫生部指定的国家一级科技项目查新机构

北京大学医学图书馆	科研立项及成果鉴定查新、定题服务	科研成果查新定点单位 北京大学医学信息咨询中心 教育部科技查新站
复旦大学图书馆医科馆	科研立项、申报奖励、成果鉴定、技术咨询、项目结题、论文发表等。医科馆专门接受医学类查新委托	查新工作站拥有 20 余位具有图书馆学、情报学、计算机、化学、物理、环境、数学、医学等专业背景的查新员和审核员。
上海交通大学医学院图书馆	课题检索和查新服务，开设文献检索课。对全社会开展课题检索和查新服务	教育部科技查新站（医学部） 科技部和卫生部以及上海市科委认可的查询点
首都医科大学图书馆	科技项目查新、论文收录及引用检索和查重服务	2006 年 1 月 24 日卫生部科教司正式下文确认为卫生部医药卫生科技项目查新咨询单位，具备为国家级或部级以上科研项目提供查新咨询服务的资格[5]
四川大学华西医院图书馆	高质量的信息支持与信息保障，包括教学、科研、临床及管理活动	2003 年 11 月，四川大学成为全国 11 个综合类“教育部部级科技查新工作站”之一（序号 Z05）。 2013 年，四川大学科技查新工作站被评为首届教育部科技查新工作站先进集体。[6]
中山大学医学院图书馆	科技查新服务	1993 年，全国首批“卫生部医药卫生科技项目查新咨询单位” 2003 年，教育部科技查新站
浙江大学医学图书馆（农医馆）	面向浙江省以及邻近省份的高校、政府机构、工矿企业、科研院所开展科技查新与咨询服务	教育部科技查新工作站(Z09)
中南大学湘雅医学院图书馆	申请科技立项；科技成果鉴定；成果申报奖励；申请专利；技术引进；新产品开发；其它	1993 年，全国首批“卫生部医药卫生科技项目查新咨询单位” 2003 年，教育部科技查新站

各大医学院图书馆查新业务范围趋同，主要面向高校师生、附属医院的学术研究、临床服务等需求。业务广度拓展以上海交大医学图书馆、浙大农医馆为代表。图书馆服务范围逐步拓展到全社会服务，如上海交大医学图书馆面向全社会提供“课题检索和查新服务”、浙大农医馆面向浙江省以及邻近省份的高校、政府机构等开展科技查新与咨询服务。业务深度拓展以复旦大学医学图书馆为代表，查新服务的深度延伸至科研立项、申报奖励、成果鉴定、技术咨询、项目结题等项目。

2.3 情报研究业务

医学院图书馆的情报研究业务一般归于“信息咨询部”等职能部门，较少单列为“情报研究中心”层级的独立架构，主要业务面向院校学科服务。各医学院图书馆负责的情报研究服务如表 3 所示。

表 3 各医学院图书馆情报研究概况

图书馆机构名称	情报研究
北京协和医学院医学信息研究所	医学信息研究 医学科技情报研究 科技评价研究 卫生政策研究
北京大学医学图书馆	医学文献研究中心、 医学信息咨询中心（包含 CALIS 全国医学文献信息中心、北京大学医学信息咨询中心） 数据与资源研究中心
首都医科大学图书馆	信息咨询部
中山大学医学院图书馆	学科服务（包括学科服务工具的使用、学科信息推送、送培训到科室、高端学科咨询服务等）。文献检索、调研、文献计量分析、战略情报分析、前沿综述、动态跟踪等
浙江大学医学图书馆（农医馆）	为科学研究、科研管理和战略决策提供情报支持。服务内容包括但不限于：各类评估数据支持、对标分析、人才库构建、规划调研支撑、排行榜支持、院系贡献度分析、研究热点及前沿探测、研究主题及学科布局变迁等。
华中科技大学同济医学院附属协和医院图书馆	/
中南大学湘雅医学院图书馆	/

其中北京协和医学院医学信息研究所较为突出，依托“中国医学科学院医学信息研究所”研究背景，在情报研究方面覆盖了医学信息分析评价研究、医学科技战略研究和大疾病防治与公共卫生信息研究三大类业务。

2.4 知识产权服务

一般挂牌大学层级“知识产权信息服务中心”，具体业务在图书馆总馆承担，不特定隶属于医学院层级。各医学院图书馆承接的知识产权服务及知识产权资质现况如表 4 所示。

表 4 各医学院图书馆知识产权服务概况

图书馆机构名称	知识产权服务	知识产权资质
北京协和医学院医学信息研究所	学科化知识服务、信息咨询	医学信息资源与知识服务中心
北京大学医学图书馆	知识产权咨询、分析、信息订阅、培训与讲座、申请、转移转化	北京大学知识产权信息服务平台
复旦大学图书馆医科馆	专利申请、成果转化、专利战略，地方经济、知识产权管理部门、企业、教育科研机构的专业知识产权信息服务	复旦大学知识产权信息服务中心
上海交通大学医学院图书馆	知识产权信息素养培训、专利检索、专利分析评议	上海交通大学知识产权信息服务中心
首都医科大学图	/	/

书馆		
四川大学华西医院图书馆	/	四川大学知识产权信息服务中心
中山大学医学院图书馆	专利咨询、专利信息素养教育、专利申请前检索、专利分析、专利查新[7]	中山大学知识产权信息服务中心
浙江大学医学图书馆（农医馆）	知识产权管理体系建设与产学研协同创新，科研项目选题、立项、实施、结题、成果转化等各个环节的知识产权信息服务。 包括专利导航、专利分析等工作，知识产权信息利用培训和推广工作，为政府部门、地方经济、企业发展提供知识产权信息服务。	浙江大学知识产权信息服务中心
华中科技大学同济医学院附属协和医院图书馆	/	华中科技大学专利中心
中南大学湘雅医学院图书馆	知识产权基础知识咨询 专利资源检索咨询 专利法律状态咨询	中南大学知识产权研究院

其中复旦大学知识产权信息服务中心设有医学部门，于 2019 年批准成立，由图书馆、科研院、上海市药物研发协同创新中心等校内机构共同组建，挂靠图书馆，以理、工、医科学科馆员为主[8]。

3 医学图书馆发展问题及趋势

为深入了解相关单位业务发展情况，挖掘行业共性问题，识别行业创新机遇点发展，本研究对某市重点医学情报研究机构与医学图书馆进行了深入调研。综合调研结果，总结出了相应的发展问题与趋势。

3.1 科技查新领域发展问题及趋势

经过调研发现，科技查新领域主要存在着系统无法准确处理的环节带来手动的工作量，严重降低了业务效率的问题：如文献的出版日期格式不统一导致出版日期字段需要手动处理，访问外国数据库；数据传输不稳定，在查新和查引操作中，可能出现数据传输不稳定，无法找到文章、获取引用数据的问题等。建议进一步推进查询服务自动化，实现查询服务的自动化，以提高工作效率和服务质量，同时进一步减少手动环节：图书馆的工作人员在未来的发展中，会进一步提高查新查引的工作效率，减少手动操作的环节，提高系统的自动化处理能力；最后推动业务转型，从查新业务转为知识产权信息服务。

3.2 情报研究领域发展问题及趋势

当前机构中面临的情报研究领域的发展问题主要为新型项目主要源于高校体系的衍生需求，如报奖类查新项目主要来源于复旦大学附属系列医院申报需求；学科评估类情报研究项目，来源于高校学科发展规划。针对该问题，各机构应拓展学术影响力评估业务[9]：可将情报研究领域作为部门的一个重点探索课题，为学校的学科建设和人才培养提供数据支持，同时针对课题、讲座的需求进行相关的学科服务，满足学校不同学科的特定需求。当前已有相关机构开拓了新的情报研究服务，如同济大学图书馆的绩效分析服务、学科发展报告服务。

3.3 知识产权领域发展问题及趋势

知识产权领域发展问题主要为业务范围狭窄，面向社会、企业的服务业务较少。被调研机构中，多家图书馆均提到企业端的项目，但总体处于尝试摸索阶段。目前以高校内部科研成果、专利转化等项目为主，承接企业需求较少。建议扩大知识产权服务合作范围[10]：寻求与更多高校以及社会各界的合作，举办相关会议，深化对知识产权的研究和应用。于此同时，深入研究和服务社会经济：研究战略新兴技术，满足市场、企业和政府部门的需求，支持科技发展和高质量社会发展。

PU-139

中华优秀传统文化体验提升医学生人文素养实践研究

刘 强
大连医科大学图书馆

党的十九大报告把文化自信上升到一个从未有过的高度：“没有高度的文化自信，没有文化的繁荣兴盛，就没有中华民族伟大复兴[1]。”图书馆具有是保存和传承人类文化遗产、开展社会教育的职能。新形势下，面临 95 后，00 后为主体的大学生群体，图书馆如何弘扬中华优秀传统文化，发展和创新人文素养教育内涵与形式，吸引大学生积极参与文化活动，提升大学生人文素养，正成为大学图书馆建设所面临的一项新课题。

长期以来，医学院校着重追求医学能力和技术的提高，注重“学以致用”，在人文素质、沟通服务和情绪调节等方面培养不够，从而，在处理复杂医患关系过程中处理问题能力不足[2]。大连医科大学图书馆开展调查研究，设计以“中华优秀传统文化”为核心内容，图书馆作为体验中心，以大学生为体验主体，以“文化体验”为教育形式的文化体验活动，开展人文素养教育，为医学生人文素养教育实践提供了一个新的探索。

1 中华优秀传统文化是人文素养教育的重要支撑

人文素养是指人在生活、学习、工作中所吸收的人文知识、人文思想以及在实践中由此内化而成的人文精神，具体表现为人丰富的精神世界，涵盖人生态度、道德修养、心理素质、文化品位、审美情趣和创新能力等方面[3]，其主要内容包括：人文思想内化、人文知识储备、人文方法运用和人文精神践行。

大学生人文素养的养成源于多方面共同影响，包括原生家庭、校园、职业、社会等，其中最为核心的是五千年中华传统文化的熏陶。中华优秀传统文化是中华民族的文化根脉，其蕴含的思想观念、人文精神、道德规范，不仅是我们中国人思想和精神的内核，对解决人类问题也有重要价值[4]。总的说来，中华优秀传统文化作为中华文明历史发展过程中所凝结的思想、观念和习惯等特质文化的总和，其具体表现为哲学思想、艺术审美、语言文字、手工技艺和民风民俗等方面。将中华优秀传统文化融入到医学生人文素养教育过程中，就是将中华民族世世代代在生产生活中形成和传承的世界观、人生观、价值观和审美观，通过知识传授、技能训练、习惯养成和环境熏陶，使之内化为大学生的人格塑造、气质修养和文化基因[5]。

2 医学生人文素养教育现状调研分析

通过采用资料调查、问卷调查和访谈法对大连医科大学人文素养课程设置、医学生阅读状况和人文活动参与情况三个方面进行调研，调研数据和结果如下：

2.1 人文知识课程设置比较单一

受传统应试教育模式的影响，医学院校对人文素养教育的认识存在着不足。选取最具代表性的临床医学专业进行调研，在大一至大四的四个学年中，开设的与人文相关的课程有 10 门，（占总课程数的 15%），其中公共课 4 门，涉及医学人文方面有 3 门，涉及传统文化历史方面的 3 门，学分及课时（占总课时数的 10%）均不多，均为考查课，见表 1。可以看出，人文素养类课程设置无论从内容上还是学时上存在不足现象，学生在课堂上的人文素养教育系统性不足。

表 1 学校开设的与人文素养相关课程

课程类别	课程名称	学分	学时	考核类型	开课年级
公共课	形势与政策	0	8	考查	大一
	大学生职业规划	.5	23	考查	大一
	军事理论	1	36	考查	大一
	马克思主义概论	.5	80	考查	大二
		2			
		5			
医学人文	医学文献检索	1	20	考查	大二
	医学伦理学	2	26	考试	大四
	卫生法规	1	13	考试	大四
传统文化	思想道德与法律	3	48	考查	大一
		3	48		

历史	基础	1	13	考查	大一
	中国近代史纲要			考查	大四
	人文基础				

2.2 阅读数量和质量不足

2019 年 11 月，图书馆对 340 名医学生进行阅读现状问卷调查。统计发现，在阅读数量和时间方面，70%学生每学期课外阅读量不到 5 册，20.5%的学生不到 1 册；70%学生每天课外阅读时间不足 1 小时，14%学生不进行课外阅读。在阅读目的方面，62.5%学生以休闲娱乐为目的，34.8%学生目的是拓宽知识，这其中一半是专业知识拓展。在阅读内容方面，77.7%学生阅读内容为文学类（小说、诗歌、散文等），其次是生活类（旅游、励志、健康等）。在图书推荐方面，40%学生荐购小说，特别是玄幻、推理、科幻、言情和畅销书。由此看出，医学生在人文素养的图书阅读无论数量上还是质量上都有待提高，且深度较浅。

2.3 人文活动参与程度浅

通过访谈发现，医学生对学校人文讲座和文化活动是持有积极参与态度的，但是，参与活动往往是一次性的，而且对于活动本身表达的内涵，缺乏深入的了解和进一步探索的愿望，这就导致活动的效果有限。另一方面，人文活动本身形式和内容设计缺乏吸引性，医学生课业任务重，只注重活动的表面的实用性，活动参与程度浅，收获少。

3 中华优秀传统文化体验提升医学生人文素养的四个维度设计

3.1 把握时代特征

党的十九大报告指出，要深入挖掘中华优秀传统文化蕴含的思想观念、人文精神、道德规范，结合时代要求继承创新，让中华文化展现出永久魅力和时代风采[1]。医学生人文素养教育过程中必须强调大学生具有崇高的使命感和担当意识，我国古代先哲“穷则独善其身，达则兼善天下”、“先天下之忧而忧，后天下之乐而乐”，始终怀有以天下为己任的家国情怀和担当意识。未来医务工作者需要树立这种以天下为己任的崇高使命感，从中华优秀传统文化中蕴含的古老智慧中，铸造睿智的灵魂，提升人文素养。

3.2 阅读经典文献

中国古代经典《大学》提出“八目”，即“格物、致知、诚意、正心、修身、齐家、治国、平天下”，认为人的教育从自我修养到社会担当的全面培育，强调人的品格塑造。《论语》中 “君子和而不同，小人同而不和”的包容思想在国家、民族、个人之间所广为接受，这些古典经典文献成为人文素养教育的重要源泉。图书馆通过设计经典阅读书目，将中华优秀传统文化历经五千年积淀的哲学思想、价值理念、辩证思维和科学智慧的经典书籍向医学生推荐和导读，为医学生体验中华优秀传统文化体验打下良好的理论基础。

3.3 设计体验流程

中华优秀传统文化体验提升医学生人文素养是一个系统性的过程，需要从总体上进行规划、设计和实施，从而保障医学生文化体验的效果。首先要有理念和总体构想，体验什么、为什么体验、体验效果预期。其次，体验要素的调研与分析，通过调查研究获得图书馆现有资源状况，进行“SWOT”分析，为科学制定文化体验方案做好准备。最后，明确文化体验内容和形式，设计过程方案、确定组织人员，并对活动整体进度控制，活动结束后对体验效果进行评估[6]。

3.4 跨界整合协作

“推动中华优秀传统文化创造性转化、创新性发展”[5]，让我们意识到中华优秀传统文化体验单单依靠图书馆这一单一主体是不够的，大连医科大学图书馆采用“跨界-整合-协作”的视野、思路和方法，立足人文素养教育背景，图书馆、博物馆、美术馆等多主体共同参与，在人员、资源、技术和环境等方面进行整合，以跨界体验内容提升兴趣和参与度，共同协作提升医学生人文素养。

4 中华优秀传统文化体验提升医学生人文素养的具体实践

图书馆是高校文化传承中心，具有第二课堂职能，作为课堂教学辅助，有责任开展人文素养教育，大连医科大学图书馆以中华优秀传统文化体验活动为依托，开展了一系列实践工作。

4.1 中华优秀传统文化基地建设

图书馆与护理学院共建美育素质教育基地，将资源、空间、人力进行有效整合，对医学生进行阅读、书法、朗诵、音乐、绘画、影视等方面美育教育，培养未来护理人员的审美情趣和审美素养，提升对自然美、艺术美的欣赏研究，懂得社会美、人格美的精髓，从而使得医学生在今后的工作中为患者、医院和社会提供高水准的护理服务，实现学校人才培养的目标。在基地建设过程中，

扩大覆盖面，与学校工会深入合作，活动群体延伸至教职工，推广中华优秀传统文化，2018 年被授予“大连市优秀中华传统文化传播示范基地”。

4.2 书法学习推广优秀传统文化

中华优秀传统文化以儒家文化为核心，吸纳诸子百家为补充，以古文、诗词、民乐、国画、书法和对联等形式予以表达[7]。书法艺术作为中华优秀传统文化的重要组成部分之一，其本身具有陶冶情操、修身养性和传承文化等功能。2016 年，图书馆建立书法爱好者社团，具有书法特长的馆员讲授指导，成员包括学生和教师，零基础学习，大批书法爱好者参与其中，并以此成立了校书法协会。社团长期坚持传承中华经典和弘扬新时代精神，师生参与讲座、展览和比赛，近距离感受中华优秀传统文化魅力[8]。

4.3 中国传统习俗塑造节日仪式感

在传统节日，中国人把共同的生活理想、人间愿望与审美追求融入节日的内涵与种种仪式中，因此，传统节日是中华民族对理想生活的极致表现，中国的民族精神和文化传统就是通过这样代代相传的节日习俗传承下来[9]。图书馆坚持在中国传统节日举办体验活动，开展图片、实物展览，荐读节日相关书目，亲手制作节日物品，系统地对中国传统节日习俗学习和了解。比如，新年举办写春联和猜灯谜，端午节斗蛋和手工包粽子，中秋节诗词会，2008 年开展至今参与活动 7000 余人，每一项体验活动激发医学生对中国优秀传统文化的兴趣和热爱，增强生活中的仪式感。

4.4 “图-博”跨界开启新模式

新时代的大学生对中国传统的手工制作，器物器具文化等方面陌生，图书馆与博物馆在保存人类文化遗产、传播文化、开展社会教育等职能上具有共同性，在深化文献资源建设和展示资源方面具有互补性。从 2017 年开始，大连医科大学图书馆文化体验活动提出先从“认识脚下的这片土地”开始，联合国家一级博物馆——旅顺博物馆开展了“古字画的修复与装裱”、“古老的技艺——拓片”和“中国古代四大发明之造纸术”等非物质文化遗产体验活动，将许多中国传统文化元素带入校园，向全校师生进行展示和体验，观摩和实践古老技艺，3 年来，医学生参与活动 2000 余人。作为一种新的人文素养教育模式，“图-博跨界”联合体验中华优秀传统文化更能激发医学生好奇心和专注力，对历史文化、文物知识能近距离地了解，有效地拓展了医学生人文知识面。

5 结束语

文化体验是对中华优秀传统文化继承模式的创新，是一条提升医学生人文素养的途径。现代化的医疗模式势必对医学生全方位的素质提出更高更新要求，提升医学生人文素养必须加快脚步。中华优秀传统文化蕴含的人文思想和人文知识为提升医学生人文素养提供了丰富的养料和内涵。医学生人文素养提升是一个涉及面广、复杂的系统工程，同样也是医学生终身学习的过程，在高等教育阶段，图书馆有责任传承中华优秀传统文化，从多层面、多维度为医学生打造中华优秀传统文化教育模式，以丰富多彩的外在形式呈现出来。通过理念传播、亲自体验、习惯养成、以点带面将中华优秀传统文化进行普及、内化，力求双向实现中华优秀传统文化的传承和医学生人文素养提升。

PU-140

数字环境下医学图书馆推进特色读者服务的思考

周楠* 王瑶 王维岩 牟燕
山东第一医科大学（山东省医学科学院）

1. 引言

《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》中多次提到“数字”：加快数字发展，建设数字中国，加快数字社会建设步伐，营造良好数字生态。[1]

《纲要》第 16 章第一节，提供智慧便捷的公共服务中指出：“积极发展在线课堂、互联网医院、智慧图书馆等”，国家政策指明了图书馆的进一步发展方向，大数据和人工智能等信息技术的快速发展正在影响着图书馆用户的行为偏好。 智慧图书馆是通过先进技术为用户提供全面感知、高效互联、可持续发展的智慧化管理和服务的图书馆模式。[2]

2. 数字环境下医学图书馆读者服务开展的现状和问题

2.1 文献研究，注意紧扣数字环境。从文献里总结，数字环境下图书馆信息服务或者读者服务都有哪些研究方向。

2.1 文献研究，以中国知网学术期刊数据库作为国内文献数据分析来源，时间跨度为 2012—2022，检索日期为 2023 年 4 月 3 日。CNKI 的检索条件为主题：数字环境+图书馆+服务，文献类型为中文文献，得到文献 117 条，通过对所得文献进行阅读筛选，总结出数字环境下图书馆读者服务的研究方向有以下几个方面：

2.1.1 借助互联网和通信技术的参考咨询服务

数字环境下的参考咨询服务是图书馆服务能力与服务水平的重要体现，数字环境为图书馆的参考咨询服务提供了更为便捷地不受时间空间限制的有利条件。图书馆在开展个性化、特色化咨询服务方面应坚持系统性、完整性，努力保持特色服务的连续性，根据图书馆实际和市场需求，为用户提供特色信息，满足读者个性化、特色化的信息服务需求。[3]

2.1.2 图书馆数字人文科研支持服务研究。数字人文就是大数据技术与人文科学进行良好互动之后所催生出的一个新的研究领域，面向数字人文科研提供支持的服务是图书馆新的服务生长点。[4] 医学图书馆的信息资源优势和空间优势是提供医学人文精神培养的重要支撑，为读者提供服务是图书馆的中心任务，医学图书馆与读者之间在互动中重视人文关怀和人文精神的传播对读者来说是更高质量的服务。

2.1.3 数字环境下的图书馆嵌入式阅读推广服务。在数字环境下，图书馆可以给读者提供比传统文字更多的阅读形态，比如音频、视频，可以更精准地提供多方位的阅读推广服务。中国科学院国家科学图书 馆、厦门大学图书馆等都是嵌入式阅读推广服务的示范机构，图书馆嵌入式阅读推广服务工作是数字时代很多图书馆的发展趋势和工作重点。[5]

2.1.4 数字环境下医学图书馆的健康服务提升研究。

2021 年 3 月 6 日，习近平在看望参加政协会议的医药卫生界教育界委员时指出，人民健康是社会文明进步的基础，是民族昌盛和国家富强的重要标志。医学图书馆可以提供专业权威的健康信息服务，通过 PCPA 服务策略满足公众健康信息服务需求，提高医学图书馆健康信息服务效能，提升公众健康风险认知水平与健康行为自律程度，从而促进健康知识传播与健康文化形成。[6]

2.2 数字环境下图书馆读者服务的实践路径

根据对北美地区大学图书馆 CDS 的调研可知，数字学术主要利用数字工具来进行基于学科的学术研究，并以数字形式进行展示，利用大数据、数字询证方法、数据挖掘分析、数字出版和保存来实现研究目标。[7]大数据环境下，英国格拉斯哥大学图书馆不仅面向本校师生服务，还面向全英高校联盟体服务，将社交媒介引入到阅读推广服务中，成为与读者沟通交流的主要方式。[8]中国医科大学图书馆在微信公众号上就可以为读者提供 39 个中文数据库、48 个外文数据库以及 27 个仅限校内网使用的试用数据库；首都医科大学图书馆创新信息服务，在微信公众号上开设周三讲座链接，通过线上培训形式开展图书馆信息资源利用专题讲座；南方医科大学图书馆拓展新的服务方式，增加知识服务的内涵，与多家单位建立文献服务联系，促进资源共进共享。

2.3 数字环境下医学图书馆读者服务遇到的问题

2.3.1 数字环境下，海量信息中的高效率获取问题。随着读者获取信息的渠道的扩展，与传统模式下的图书馆相比，读者接触到的信息数量更多，信息量更大，来源更广泛，在海量信息中也裹夹着虚假信息，快速提取精准有效信息的难度增加了，广泛的以资源为中心的服务难以满足读者需求。此外，在加工信息的过程中，很多时候，馆员或专家在整合与处理信息时会按照自己的知识经验来进行，对于用户认为是至关重要的信息，他们可能会忽视掉，进而导致遗失掉一些宝贵的研究信息。[9]

2.3.2 医学图书馆阅读推广工作的专业化还需要进一步提高和加强。近些年，图书馆在读者服务中所开展的阅读推广工作蓬勃发展，已经成为图书馆主要服务内容之一，阅读推广的效果也成为医学图书馆的重要服务指标。医学图书馆是专业性清晰的图书馆，服务对象的专业性需求明显，但是在阅读推广工作的开展中，除荐书导读这部分服务依赖于馆藏并且可以提高馆藏的利用率，是非常专业化的图书馆服务内容外，其他形式的阅读推广工作比如棋牌、手工、书画展等很难展现专业服务的优势。图书馆开展的这些阅读推广活动,因为脱离自身专业知识和业务流程的支 撑,往往很难展现自身专业服务的优势。因而,阅读推广面临的最大的专业化挑战就是如何将阅读推广活动纳入现有图书馆业务流程中。[10]

2.3.3 自动化、智能化管理中遇到的新问题

伴随着 AI 智能、5G、大数据在图书馆发展中的应用，图书馆的管理也更多地使用智能设备，数字资源和多媒体媒介，以期实现更便捷高效的服务。然而在智能设备的长期使用中很多设备容易

损坏、出现一些故障，或者存在着一些数据错误的问题，需要随时或定时进行数据修改或修理，一定程度上影响读者的正常常规使用。

2.3.4 主动开展的健康服务中的问题

目前，我国高校图书馆健康信息服务研究涉及的内容较少，且主要集中在医学院校图书馆领域。[11] 现有的研究主要从服务主体、服务内容、服务方式、政策环境、服务对象等基本要素出发，忽视了要素间相互关联、相互作用所产生的动态要素。[12]

3. 数字环境下医学图书馆开展读者特色服务的路径

3.1 首先说一下数字环境下特色服务的方向。

数字环境下的医学图书馆服务应是以用户为中心，我们从不同的服务对象需求来分析读者服务的方向，医学图书馆的服务对象主要是医学生、医生、科研工作者、还有教辅部门的教职工，乃至社会大众，他们的需求是不完全相同的。随着数字环境和信息技术的发展，医学图书馆也在不断地尝试改进服务方式，调整服务方向，以期提高读者群体的满意度和需求。

3.2 数字环境下面向医学生的人文精神培养

医学图书馆是医学生教育的重要支撑机构，医学生也是医学院校图书馆服务的重点人群，是人数占比最高的读者群体。医学人文素养是医务工作者在临床医疗工作中品质和人文情怀的体现，培养医学生的人文素质，医学院校图书馆应发挥重要作用。医学人文素质教育有利于提高医学生的职业修养，有利于和谐医患关系，有利于医疗卫生事业的长久发展。 利用自身优势，提高医学生人文素养水平，建设与各自校史文化等相联系的医学特藏资源，有望成为医学图书馆转型发展的新方向。[13]

3.3 数字环境下面向医生和科研人员的信息服务

信息技术和数字环境带来了医学临床和科研教学的数字化电子化，医学图书馆也在积极通过调整图书馆电子资源库 IP 服务、账户服务等支持临床和互联网教学，在数字平台的不断完善中，打破时空的壁垒，服务于医学院校的课程教学和医务科研人员的参考咨询服务也在不断提升过程中。我国医学文献资源共建共享主要依靠以全国高等医药院校图书馆为主体的国家、地区、省（直辖市、自治区）3 级医学文献资源共享网络、中国高等教育文献保障系统（Ｃｈｉｎａ Ａｃａｄｅｍｉｃ Ｌｉｂｒａｒｙ ＆ Ｉｎ－ｆｏｒｍａｔｉｏｎ Ｓｙｓｔｅｍ， ＣＡＬＩＳ）和国家科技图书文献中心（Ｎａｔｉｏｎａｌ Ｓｃｉｅｎｃｅ ａｎｄ ＴｅｃｈｎｏｌｏｇｙＬｉｂｒａｒｙ， ＮＳＴＬ）以及部分区域性图书馆联盟。[14]

3.4 数字环境下医学图书馆的校园健康教育和社会公众健康服务

3.4.1 医学图书馆开展校园健康教育的方式很多，主要有依托于文献资源开展的讲座、活动展览，微信公众号上的信息推送等，依托于阅读推广活动中的健康教育宣传以及以关注医学生心理健康为方向的阅读疗法等。医学图书馆开展校园健康教育的认同度比较高，但是仍然存在参与率比较低的问题，怎样提高校园健康教育的参与率，让更多的读者主动走进健康教育的大课堂是今后要解决的问题。

3.4.2 我国医学图书馆在社会层面的健康信息服务尚未形成体系、规模，一方面是医学图书馆的从业人数有限、馆员结构方面缺乏为开展健康服务的人才支撑等自身限制问题，另一方面，缺乏政策性的行业指导，基本处于探索研究阶段。面向公众开展健康信息服务无法仅依靠医学图书馆自身力量完成，应联合医疗机构、公共图书馆、社会团体、社区居委会、健康信息商业网站等多方力量，协同合作、共同发展。[15]

4 数字环境下医学图书馆开展特色服务应具备的能力

大数据环境下的医学图书馆的特色服务、创新服务最终需要落实到图书馆的自身能力建设上，用发展来解决读者服务中出现的各种问题和需求。

4.1 通过培训学习提高现有馆员的服务能力，从读者需求出发，以问题为导向，着力解决读者服务中出现的不足，提高读者满意度。

4.2 引进人才，优化馆员结构。当前很多医学图书馆都存在人才储备不足的短板，有些图书馆退休人员多于新进人员，甚至退而不进。现有馆员队伍中图书情报专业比例少，能够支持开展数据环境下的个性化参考咨询、健康服务等方面的人才缺乏。

4.3 随着人工智能和信息技术的发展，医学图书馆也应尽快地进行更高效的尝试，解决读者服务和需求之间的矛盾。图书馆亟需进行馆藏资源与网络资源、实体资源与数字资源的整合、筛选与关联，并将信息服务范围从物理空间向虚拟空间、从 PC 端向移动端延伸扩展，实现图书馆资源、

服务的一体化变革。[16]

5 结语

总体而言，医学图书馆做为医学教育和医学研究的重要保障机构，在新的数字环境形势下，需要通过自身发展、合作协作来更好地实现人文精神培养、健康教育服务、参考咨询服务等特色读者服务，更好地实现社会价值。

简单说一下我们的研究有什么意义。

PU-141

河南省洛阳正骨医院（河南省骨科医院）科技论文产出计量分析

刘瑞霞 古想花 杜丽萍 丰锐
河南省医学情报研究所

科技论文是具象化的科研产出，某种程度上可以反映机构的研究水平和科研能力。虽然不能只用论文评价机构或个人的研究能力和学术水平，但论文作为科研产出的重要组成部分，用于宏观层次的整体科研评价是有意义的【1】。河南省洛阳正骨医院（河南省骨科医院）始建于 1956 年，是一所集医疗、教学、科研、产业于一体、省卫生健康委直属的三级甲等骨科专科医院【2】。本研究从论文产出视角对河南省洛阳正骨医院（河南省骨科医院）近 10 年的科研发展水平进行评价，以了解该院的科技实力和综合影响力，为制定和完善医院的科研管理、学科建设和人才培养提供参考依据和政策建议[3]。

1. 资料与方法

以中国资源总库 (CNKI)数据库为统计源，在“作者单位”项中以“洛阳正骨 or 河南省骨科医院 or 河南省正骨研究院”为检索词；或 第一作者单位或者通讯作者为“洛阳正骨 or 河南省骨科医院 or 河南省正骨研究院”，下文（统一）简称“洛阳正骨”。 时间限定 2013-2022 年，检索洛阳正骨近 10 年发表的科技论文，检索时间 2023 年 9 月 12 日。根据检索结果，从论文发文量、基金论文量、核心期刊发文量、主要主题分布、期刊分布、被引频次、下载频次、作者分布为评价指标，利用 Excel 统计软件进行整理、归纳、分析，对河南省洛阳正骨医院（河南省骨科医院）近 10 年发表的科技论文、科研发展变化进行评价。

2. 结果与分析

2.1. 总发文量、基金论文及核心期刊发文情况

洛阳正骨 2013～2022 年发文量、基金论文量及核心期刊发文量统计见表 1， 2013-2022 年共发表科技论文 3873 篇，其中第一著者单位为“洛阳正骨” 发表论文 2779 篇(71.75%)；基金论文 1485 篇，基金论文占比 38.34%；核心期刊发文量 468 篇，核心期刊发文占比 12.08 %。通过发文量趋势图（图 1）发现，从 2013 年起，论文发表量呈逐年增加趋势，到 2018 年到达顶点，然后逐年下降。 与我们课题组同期统计的其它医院相比有明显差异。近 10 年基金论文发文量、核心期刊发文量呈稳步上升趋势。

表 1 洛阳正骨医院 2013-2022 年发文量、基金论文量及核心期刊发文量统计表

年份	发文总量（篇）	基金论文量（篇）	基金论文占比（%）	核心期刊发文量（篇）	核心期刊占比（%）
2013	293	30	10.24	27	9.22
2014	291	55	18.9	35	12.03
2015	399	76	19.05	42	10.53
2016	332	111	33.43	45	13.55

2017	377	126	33.42	54	14.32
2018	492	182	36.99	43	8.74
2019	480	202	42.08	51	10.63
2020	449	235	52.34	59	13.14
2021	422	233	55.21	62	14.69
2022	338	235	69.53	50	14.79
合计	3873	1485	38.34	468	12.08

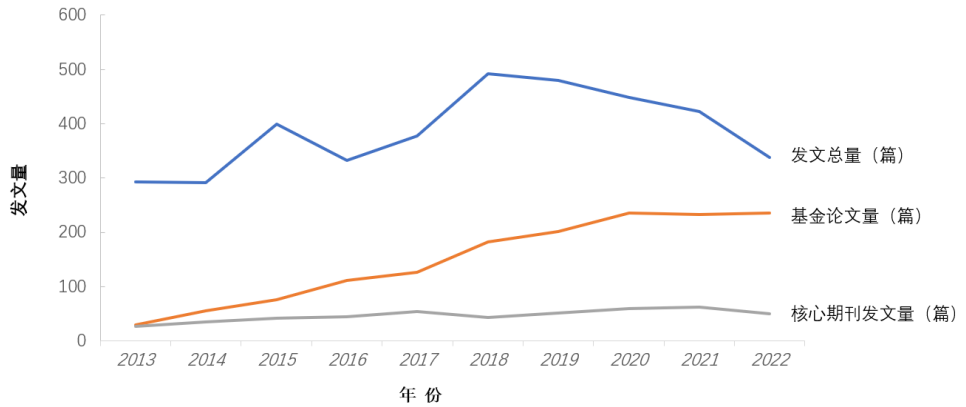
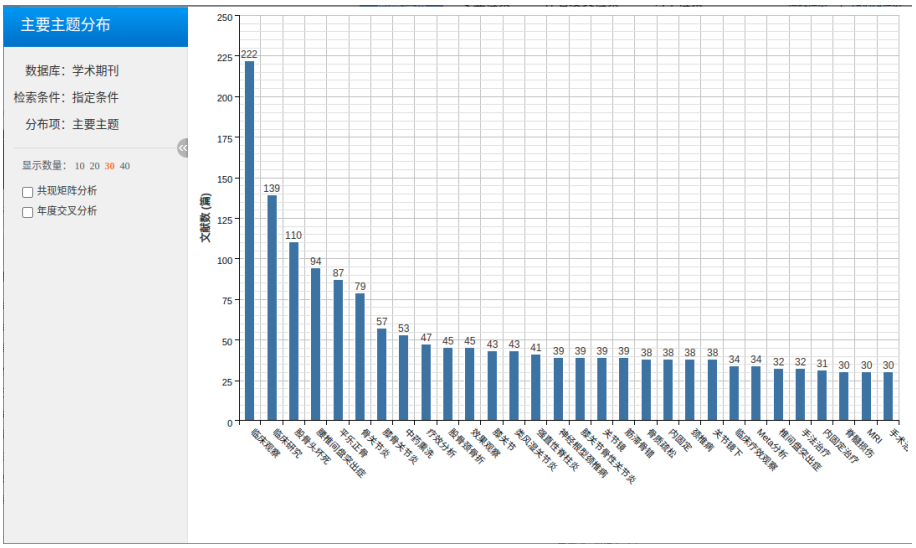


图1 洛阳正骨2013-2022年发文量、基金论文量及核心期刊发文量趋势图

2.2 主要主题分布

- 洛阳正骨医院 2013-2022 年发表论文主要主题分布情况见图 2，发文量最多主要主题的是临床观察(222 篇），其次是临床研究（139 篇），股骨头坏死（110 篇），然后依次是腰椎间盘突出症（94 篇），平乐正骨（87 篇），骨关节炎（79 篇），膝骨关节炎（57 篇），中药熏洗（53 篇）。

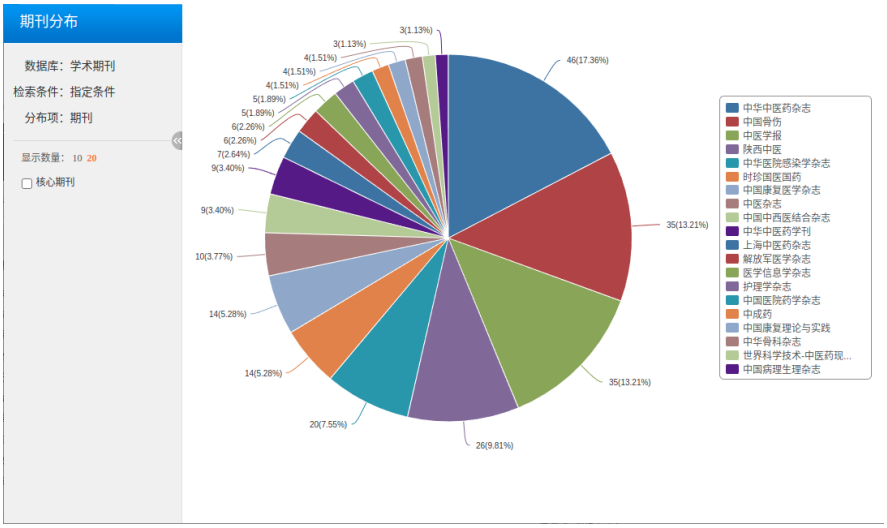
图 2 洛阳正骨医院 2013-2022 年发表论文主要主题分布图



2.3 期刊分布

洛阳正骨医院 2013-2022 年论文发表期刊分布情况见图 3，发表论文最多的期刊是中华中医药杂志 46 篇（17.36%）,其次是中国骨伤 35 篇（13.21%）、中医学报 35 篇（13.21%）、陕西中医 26 篇（9.81%）、中华医院感染学杂志 20 篇（7.55%）、时诊国医国药 14 篇（5.28%）、中国康复医学杂志 14 篇（5.28%）、中医杂志 10 篇（3.77%）。其中中华中医药杂志、时诊国医国药、中医杂志、中国中西医结合杂志、中华中医药学刊、中成药、护理学杂志、针刺研究为北大核心期刊。

图 3 洛阳正骨医院 2013-2022 年论文发表期刊分布图



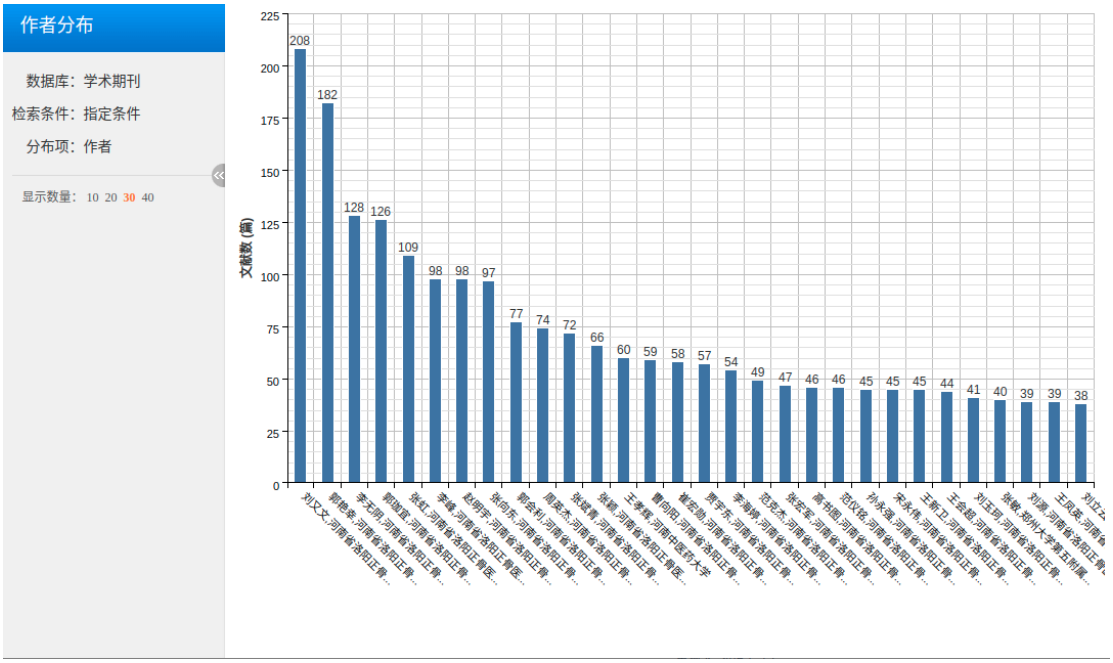
2.4 论文被引及下载情况

洛阳正骨医院 2013-2022 年论文单篇被引频次前 10 名的文献见表 2。被引频次最高的论文是张萌萌的“骨代谢生化指标临床应用专家共识（2020）”，被引频次为 233 次，下载频次为 5484 次；其次是黄宏兴的“肌肉、骨骼与骨质疏松专家共识”，被引频次为 186 次，下载频次为 4221 次。；除专家共识外，被引频次最高的论文是赵明宇（通讯作者）的“膝骨关节炎发病机制及治疗进展”被引频次为 170 次，下载频次为 4667 次。；其次是秦洛宜的“姜黄、莪术、郁金的化学成分与药理作用研究分析”,被引频次为 97 次，下载频次为 3132 次。

2.5 作者分布

洛阳正骨医院 2013-2022 年发表论文作者分布情况见图 4，发文量在 100 篇以上的作者有 5 位，最多的作者是刘又文（208 篇），其中核心期刊 44 篇，其次是郭艳杏 182 篇，其中核心期刊 26 篇；第三李无阴 128 篇，其中核心期刊 26 篇；第四郭珈宜 126 篇，其中核心期刊 22 篇；第五张虹 109 篇，其中核心期刊 20 篇。发文 38 篇以上的前 30 位作者，发文量为 2187 篇，占同期全院发文量的 56.47%，为本院较为稳定的核心作者。

图 4 洛阳正骨医院 2013-2022 年发表论文作者分布图



3.讨论

科技论文的统计分析，可从一个侧面反映出科研机构和学科团队的科研产出、学术影响力和活力等情况[4]。基金资助论文能较多地反映本专业领域的新动向、新成果，有助于发挥科研导向作用，论文通过核心期刊发表，无疑能增强其传播力度与效果[5-6]。

本文通过对洛阳正骨医院 2013-2022 年论文产出的分析，粗略了解了该院近 10 年的科研发展状况。总发文量前 6 年逐年波浪上升，后 4 年呈逐年下降趋势；这可能是 3 年新冠疫情影响了科研及论文量。近 10 年基金论文量、核心期刊发文量均呈稳步上升趋势，显示该院的发文质量逐年提高。发表论文最多主要主题的是临床观察(222 篇）；发表论文最多的期刊是中华中医药杂志、且该期刊为北大核心期刊，说明发表论文的质量和学者水平相对较高；除专家共识外，论文被引频次最多的是张荣的“膝骨关节炎发病机制及治疗进展”。高被引论文可以代表某一学科中质量较高、影响力较大的高水平论文[7]，论文被引用从某种意义上可以体现研究成果在领域内的引领作用和被关注、认可的程度[8]。从被引频次和下载频次可以反映出当前研究的热点和发展方向。从近 10 年发文的作者分布可看出该院核心作者群基本形成。上述统计结果也可见，核心期刊论文数量不高(12.08%)。医院虽然在各期刊发表论文增加，但在核心期刊发表论文数量相对不多。该医院论文质量还有很大的提升空间，让本院科研人员将发表高质量论文作为努力的方向。

为了有效提升医院科研实力，促进科研发展，提出以下提升策略。（1）科研管理部门应该加大科研人员的扶持以及对科研成果的奖励力度，建立良好的学术环境，及时向医务人员推送国内外医学领域的最新发展学术动态，增强学术氛围，激发医务人员的科研热情及产出科研论文的积极性；（2）定期举办各类培训，聘请专家对科研选题、科研设计、期刊的种类、论文写作格式、投稿技巧等进行详细讲解，帮助医务人员提高科研设计能力和产出论文的创新性[9]，促使其提高自身医学信息素养，助力其科研产出。（3）鼓励、引导医务人员多向高影响力或核心期刊投稿，因为这类期刊期刊论文受到学术团体的普遍认可与关注。

总之，通过对论文产出的统计分析，从侧面了解了近 10 年河南省洛阳正骨医院（河南省骨科医院）的科研水平现状，希望能对该医院科研管理起到一定的帮助，助力其提高医院诊疗水平及学术影响力。

PU-142

叙事医学视角下医院图书馆阅读推广服务模式研究

秦方 侯胜超 胡鸿*
华中科技大学同济医学院附属同济医院

全民阅读是增强文化自信和建设文化强国的重要举措，图书馆作为社会公益服务机构，肩负着培养用户阅读兴趣和提高阅读素养的重任，在促进全民阅读中发挥着重要作用。本文围绕医院图书馆阅读推广服务遇到的问题，从叙事医学的视角，以高质量服务为抓手，构建医院图书馆阅读推广服务模式。

1.图书馆服务与叙事医学概述

1.1 图书馆+叙事：“叙事元素”为图书馆阅读活动注入新活力

在医院图书馆阅读推广活动中引入“叙事元素”，融入外化、改写、见证等叙事技巧，结合国内外优秀叙事案例或有叙事元素的活动案例，运用“讲故事”的方法，对图书馆阅读活动提出叙事化策略：从叙事学、叙事心理学等学科领域的视角，确立故事主题、故事主体和故事主线[1]，让深藏在图书馆中的具有叙事价值的作品真正“活起来”，通过叙事吸引更多读者走近图书馆，爱上阅读，为医院图书馆阅读推广活动提供新的思路。

1.2 图书馆+医学人文：医学人文教育助力全民阅读高质量发展

早在 1927 年，我国著名图书馆学家李小缘先生在他的《图书馆学》一书中就提出: 图书馆即是教育，阅读是社会教育的基石[2]。医学人文教育是医学教育的重要组成部分，医院图书馆开展人文教育更具优势。医学人文是医学的意识形态、上层建筑，医学人文教育教学要有知识、有经典、有故事、有共识、有参与、有讨论，同时还要结合医学实践。图书馆工作与医学人文教育教学活动的融合，有利于公众获取更多阅读机会，帮助弥补我国教育体制中长期存在的“重理轻文”现象，使整个社会人员的医学综合素质得以提升。

1.3 图书馆+叙事医学：叙事医学让图书馆阅读推广更有温度

叙事医学（Narrative Medicine）是 2001 年由美国哥伦比亚大学内科教授丽塔·卡伦（Rita Charon）提出，2011 年被北京大学医学部郭莉萍教授引入国内后，其理论与实践方法受到学术界的广泛认可，被认为是医学人文在临床的落地工具。近年来，叙事医学作为医务工作者提高沟通能力的新方法和新手段，已受到国内外众多医学院校和研究机构的重视，在国内医院图书馆也有相关实践和应用。在中国知网中以“叙事医学”和“图书馆”为检索词，检索发现已有 6 篇文献报道。“有时去治愈，常常去帮助，总是去安慰”，图书馆开展叙事医学、叙事护理等活动，对改善医患沟通以及促进整个社会的和谐发展具有重要的意义。

2.医院图书馆阅读推广存在的问题

医院图书馆和公共图书馆、高校图书馆因归属领导不同，在馆藏资源、功能定位、服务对象也有很大差别，医院图书馆的阅读推广活动应该区分于公共图书馆的“公益性”定位，也应区分于高校图书馆的“学术性”定位，体现医院图书馆的专业性和医学优势与特色，在阅读推广活动中主要存在以下问题：

2.1 阅读推广机制不健全

医院图书馆阅读推广工作开展的基础是用户需求，有些馆在举办阅读推广活动时，仅从图书馆角度出发，根据以往的经验组织开展活动，没有对用户的人群特征、关注点及个性化需求展开实地调查，没有结合各自的馆藏优势进行推陈出新，呈现用户持续参与意愿降低，阅读推广活动服务质量不高等问题。当前医院图书馆阅读推广服务缺乏统一的科学规划，没有形成健全完善的阅读推广制度，缺乏推广体系、保障机制、激励机制、评价机制等。

2.2 阅读推广模式单一，同质化明显

随着用户阅读环境的不断变化， 5G、人工智能、物联网等新兴信息技术的层出不穷给阅读推广带来了挑战[3]。医院图书馆在阅读推广实践中推广模式单一，推广主体单一，大部分由馆员自己担任，没有发挥医疗领域的人才优势招募合适的阅读推广人。各医院图书馆之间存在相互借鉴，导致阅读推广活动同质化明显、区分度低等问题，难以满足读者实际的阅读需求。

2.3 阅读推广内容创新性不足

医院图书馆要科学整合馆藏资源和服务，加大经典阅读的内容引领，加强优质阅读内容供给，打造高品质的内容阅读品牌，使阅读推广的内容社会化、生活化、实践化，真正被广大读者感知、

认同和接受，从而推进医院高质量发展的进程。当前，医院图书馆阅读推广活动仍然依赖多年以来的线下传统媒体开展，如读书会/读书沙龙、阅读打卡、有奖征文、写书评、图书漂流等，线上开展阅读活动依托腾讯会议互动效果欠佳，这些活动多在 4.23 世界读书日、读书月以及特殊纪念日开展，活动形式与公共图书馆相似，缺乏创新。

2.4 阅读推广联合不够

医院图书馆在开展阅读活动时没有注重外部和内部的联合，一方面外部联合不够，大部分医院图书馆开展阅读推广活动主要面向的是本院读者，极少采用“院院联合”、“馆馆联合”、“院企联合”等方式向院外读者开放，仅少部分面向患者开放，这样的推广方式影响力弱，无法形成叠加效应。另一方面内部融合不够，医院图书馆在开展阅读推广活动时与医院其他职能、教学部门的联系不够，很多活动仅由图书馆自己组织开展，没有利用与医院其他部门的关系扩大阅读的覆盖面，不能有效激发用户主动参与到医院图书馆的阅读活动中去。

3.基于叙事医学构建医院图书馆阅读推广服务模式

近些年，公共图书馆和高校图书馆都涌现出了丰富的阅读推广活动服务案例，且各具特色。但医院图书馆的相关案例并不多见。而医院图书馆的服务对象区别于公共图书馆和高校图书馆，具有显著的医疗特征。因此，有必要构建一个适用于医院图书馆的阅读推广活动框架，以指导医院图书馆更好地开展相关实践活动。

3.1 构建“叙事医学”素材库

国内外叙事医学教育普遍缺乏系统的叙事素材，而图书馆阅读推广活动的本义即以培养一般阅读习惯或特定阅读兴趣为目标而开展的图书宣传推介或读者活动。图书馆可以通过整合相关资源，建立适用于用户阅读的医疗叙事素材库。素材库以卡里茨库斯(Kalitzkus)和马提森(Matthiessen)将疾病叙事分类作为分类框架：患者叙事、医生叙事、医患接触叙事和宏大叙事/元叙事，为阅读推广活动的开展和叙事医学教育提供高质量的内容基础。

3.2 搭建叙事医学阅读推广服务框架

丽塔·卡伦教授提出了将细读、创意写作、关切讲述和倾听来提升和构建用户叙事能力的特色教学方法，此维度下的“细读与写作”等同于图书馆阅读推广工作中提倡的“深阅读与书写阅读感悟”。本文从图书馆学理论来解读叙事医学的教学方法，初步以反思性接触、反思性写作（平行病历）、心得分享与讨论三个方面搭建医院图书馆叙事医学阅读推广服务框架。

3.3 拓展阅读推广形式

大数据、云计算、人工智能和区块链等新技术的广泛应用，使得阅读形式变得更加多元，新的阅读形式不断刷新着人们的阅读体验。在新媒体环境下，图书馆利用视觉艺术、表演艺术和音乐，借助抖音、短视频等自媒体、社交媒体实现图书馆与外部的跨界融合，用叙事故事让传播增值。用创造性的表现方式实现分享人身份多元，被叙维度也更加多元化，打通线上线下沟通的壁垒，使阅读更具沉浸感，提高叙事医学阅读推广的有效性。

3.4 开展推广效果评价

阅读推广首先是一种服务，活动是它的主要展现形式。图书馆主动研究用户参与叙事医学阅读推广活动的影响因素，注重每一次阅读推广活动的反馈，提升图书馆开展阅读推广活动的力度和广度。阅读推广活动要评估活动内容的吸引力、互动性、方案可行性、形式多样性等，同时也要考量活动时间、活动宣传、叙事经济成本等方面的因素，图书馆要为这些隐形指标分配权重进行量化评定，从而实现阅读推广活动的可持续性。

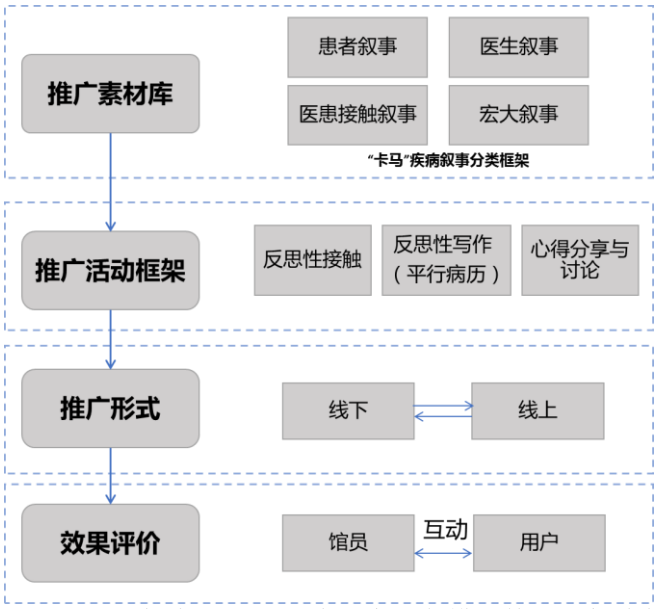


图 1：基于叙事医学的阅读推广服务模式构建的初步框架

4.基于叙事医学的医院图书馆阅读推广服务实现路径

图书馆是阅读推广工作的前沿阵地，叙事医学与日常医疗工作息息相关，医院图书馆利用医院的优势资源，开展有深度有厚度的叙事，通过叙事架起沟通的桥梁。本文从图书馆和用户两个方面来探讨叙事医学阅读推广服务模式的实现路径。

4.1 图书馆层面实现路径

4.1.1 组建阅读推广服务团队

医院图书馆要积极组建成立专业的阅读推广服务团队，考察服务团队的人员构成、学科背景和组织结构招募图书馆阅读推广人。除馆员外，邀请有医学背景的医生护士及其他医务人员作为图书馆的阅读推广大使，努力打造叙事医学专家团队，包括叙事医学教学团队，叙事医学护理团队.....充分利用现代信息技术，借助新媒体平台的优势创新开展图书馆智慧服务，积极开展叙事医学实践探索。

4.1.2 强化馆员叙事能力培养

国外叙事医学教学体系将细读、创意写作、关切讲述和倾听称为叙事医学的特色教学方法，其中标志性方法是细读和创意写作，此维度下的“细读与写作”等同于图书馆阅读推广工作中提倡的“深阅读与书写阅读感悟”[7]。韩启德院士提出现代医学不仅要学会“找证据”还要学会“讲故事”。阅读推广馆员的能力尤其是叙事、倾听、共情能力是高质量阅读推广服务的基础。

4.2 用户层面实现路径

用户是阅读推广活动的主要群体，用户的积极参与是阅读推广服务持续开展的生命力，让有叙事能力的读者参与进来，提升用户参与度是开展阅读推广活动的目标。叙事医学是医生、护士、社会工作者应该共同践行的医学。医院图书馆面向的用户群体主要是医务人员和患者，在特定时代背景下，用户个体叙事作品的传播增强了医学相关主体之间的关系，开放、包容，充满反思精神的叙事行为与路径，影响用户在文学和现实生活找到着力点，强化换位思考和共情的能力。

5.结语

医院图书馆阅读推广服务不应局限在对图书、信息、文献的推广上，还包括对图书馆员、图书馆空间、图书馆资源甚至是对图书馆本身的推广上。把叙事引入到图书馆阅读推广工作中，将叙事医学与图书馆服务工作结合起来，对于唤醒医学人文、深化阅读推广、促进医患沟通和营造亲密关系都具有积极推动作用[8]。面向医院高质量发展要求，叙事医学不仅是医学人文落地的有效工具，也是精准医学时代医学实践不可或缺的组成部分，未来，通过不断探索更多“图书馆+”阅读推广新形态，创新阅读推广活动方式，提升阅读推广专业化水平，努力实现图书馆阅读服务与各类读者阅读需求的有效对接，扎实推进医院图书馆全民阅读推广工作再上新台阶。

PU-143

智慧图书馆视角下新一代图书馆服务平台的研究现状和思考

冯研
大连医科大学图书馆

伴随着物联网、云计算、人工智能等技术的发展，智慧图书馆成为近年来图书馆界备受关注的话题，它是信息技术与图书馆传统服务模式相互融合的产物。图书馆服务平台作为智慧图书馆建设的重要组成部分，为智慧图书馆提供重要的数据基础、平台服务和资源支撑，国内新一代图书馆服务平台建设刚刚起步，很多图书馆也在积极考察或早期实施阶段，本文将调查分析新一代图书馆服务平台在国内的应用现状，总结目前国内应用的实践案例，讨论在智慧图书馆建设的视角下开展新一代图书馆服务建设以及实施的一些思考与展望。

1.引言

1.1 智慧图书馆的发展

2003 年，芬兰奥卢大学图书馆的 Aittola 等首次提出 Smart Library 的概念[1]，从概念提出之日起，围绕智慧图书馆的定义和如何建设，业界一直有着多种声音，其内涵也是在不断演变，总结智慧图书馆主要有以下观点。

一方面，从智慧应用和智慧服务的角度来看，智慧图书馆主要是通过物联网等感知技术为用户提供智慧化的服务和管理。具体而言，它们使用 RFID 技术实现自助借还书、图书盘点等功能；通过“一站式”搜索服务，使用户能够快速找到所需资源；利用语义技术和社交网络技术提供更具有感知性的服务，如虚拟咨询、个性化推荐等；同时，移动数字图书馆还可以通过手机、平板电脑等设备，不受时空限制为用户提供移动阅读服务；数据挖掘则能够从大量数据中提取有用的信息，帮助图书馆更好地了解用户需求，优化服务。

另一方面，从系统平台建设的角度来看，智慧图书馆需要构建智慧图书馆系统平台，形成智慧图书馆服务体系或服务架构。这需要硬件和软件两个方面的支持。硬件方面包括各种设备、传感器等，用于采集数据、实现自动化管理等功能；软件方面则需要开发各种应用软件、算法等，以实现对数据的处理、分析、挖掘等功能，从而支持智慧化的服务和管理。智慧图书馆系统平台还需要具备良好的开放性和可扩展性，能够适应不断变化的需求和技术发展。

1.2 新一代图书馆服务平台

图书馆管理系统一直是图书馆的核心应用，是信息化建设的主要技术体现，一直随着图书馆的发展而不断升级迭代。第一代图书馆管理系统始于 20 世纪 50 年代，主要是通过计算机与通信技术的应用使图书馆的 流通业务实现自动化，也被称作图书馆自动化系统，解 决的是从手工操作到计算机管理的自动化转变。至 20 世 纪 70 年代末，图书馆自动化系统集成了包括采访、编目、典藏、流通及连续出版物等传统图书馆的主要业务，开始向图书馆集成系统（Integrated Library Systems， ILS）转型。20 世纪末至 21 世纪初，伴随着信息技术的 发展，数字技术和存储技术得到大量应用，图书馆进入“数字图书馆”时代，一批新的服务与系统被投入使用，与传统图书馆集成系统共同组成第二代图书馆管理系统，完成了为图书馆提供更多服务形式和内容的数字化转变。2008 年随着“智慧”这一概念的普及，第三代图书馆——“智慧图书馆”成为图书馆建设的新理念， 其核心工作已经发生变化，传统的自动化系统已不能满足图书馆新的业务与服务发展的需要，亟需更加符合图书馆发展需求的“新一代”图书馆管理系统，下一代图书馆服务平台（Library Services Platforms，LSP）走进视野，目的是实现资源、技术、服务深度融合的智慧化。图书馆服务平台（Library Service Platform，简称 LSP）这一概念最初由 Marshall Breeding 于 2011 年 8 月提出[2]，是相对于上一代“图书馆集成管理系统 ILS”而言，应用于全媒体全流程资源管理并提供资源发现功能的图书馆 IT 解决方案，对应于过去 C/S 或 B/S 架构，一般都是以 XaaS 方式提供云服务，因此被称为“平台”。

2.新一代图书馆服务平台的研究现状

2.1 国内新一代图书馆服务平台

经过近十年的发展，国外的图书馆系统供应商推出了各自的下一代图书馆服务平台产品，目前国外市场上图书馆服务平台产品主要有 Alma、Sierra、BLUEcloud、FOLIO 等。

国内图书馆系统供应商在 LSP 的研发上虽然起步较晚，但一直在积极地研发自己的下一代 LSP 产品，基于对国外同类产品的学习以及对国内图书馆实际需求的了解，已推出部分成型产品，如超星集团与江苏图星合作的 LSP、江苏汇文的 META 和维普资讯的 DALIB 等。

根据目前国内目前 LSP 市场发展现状来看，已经推向市场并且比较适合国内图书馆的 LSP 产品主要有 4 款：国外的 Alma，国内的 META、LSP 和 DALIB。其所使用的主要技术对比见表 1.

表 1.国内新一代图书馆服务平台主要技术对比

	体系架构	部署方式	主要技术	合作数据商
Alma	SOA	云服务	框架：Spring Loud,数据库 Oracle	ProQuest
META	微服务	云、本地	框架：Vue.js VertX 数据库 Mysql	EBSCO、维普
LSP	微服务	云、本地	框架：Reac,Java 数据库 Mysql	超星
DALIB	SOA	本地混合	Hadoop、ElasticSearch	维普

2.2 新一代图书馆平台实施情况

国内一些高校和公共图书馆也实施了新一代图书馆服务平台，主要实施情况见表 2。

表 2. 国内新一代图书馆服务平台实施情况

	名称	公司	主要学校
Alma	**Almas**	Exlibs	北京师范大学、清华大学、北京大学、东北师范大学
	Meta	汇文	哈尔滨工业大学、中国海洋大学、大连理工大学、哈尔滨工程大学
	LSP	超星	中国矿业大学、南京大学、南京邮电大学、JALIS
DALIB	**DALIB**	维普	重庆大学、西南交通大学
Sierra	**Sierra**	Innovative interface	华中科技大学
WMS	**WMS**	OCLC	苏州图书馆

2.3 国内新一代图书馆服务平台的具体实践情况

国内高校虽然有些升级为新一代服务平台，但各馆使用的深度不一样，基本存在以下情况。

1）完成服务平台升级，还未全面管理电子资源

有些图书馆虽然升级为新一代服务平台，由于各种原因，未全面管理电子资源，电子资源的全生命周期管理，包括：定位电子资源包记录，订购完成订单信息，激活和测试电子资源包，完成发票信息以及每年的续订来实现对电子资源的选择、采访（订单、发票）、激活、编目和全文链接及访问服务，将原本分散且工作量较大的工作进行了集成和整合。新一代服务平台为电子资源全生命周期的系统化管理提供了机遇，但随之而来的也是图书馆电子资源相关业务工作内容的变革及流程的变化，需要采访、订购、参考资源多个部门馆员在后台协同工作，这项工作可想而知需要重新制定工作流程和工作规范，完成对电子资源的管理，

2）管理电子资源，未和发现系统融合

已经开展电子资源的管理，在馆藏书目查询页面上能查找电子书、电子期刊等，能显示全部订购电子资源中的刊名列表等，但检索级别不能达到篇章级别，不能提供类似发现系统的检索。

3）和发现系统融合，提供更准确全面的发现系统检索功能

在馆藏书目查询页面，发现系统替代了原来的 OPAC 系统，提供统一检索入口，能查询到系统内所有纸本，电子资源，而且检索级别提供到篇章级的检索。

4）管理多种自建资源

新一代服务平台除了管理电子资源，同时也提供对多种自建特色数字资源的管理，如学位论文，非书资料，研究数据等多种文献类型，前提需要馆员对多种类型格式数据制定格式标准，标引和整理上传数据。这样实现多种类型数据的统一管理，在页面提供统一查询。

3.国内主要应用案例

通过文献研究、网页访问、线上沟通等多种形式，对国内新一代服务平台的应用情况，做调查研究，可看到如下应用实践且非常有参考意义的案例。

3.1 和发现系统融合

1）上海图书馆使用云瀚系统，采用网络级发现服务实现统一检索[3]。上图基于 FOLIO 平台的微服务架构，采用开源发现工具 VuFind 作为资源发现服务门户，中文学术资源采用万方元数据仓储的中文发现服务，外文学术资源采用 EDS 外文发现服务。从而整合自购、自建、共享和开放获取的所有资源，通过统一、快速的搜索服务，支持对图书馆海量纸质馆藏资源与电子馆藏资源进行一体化检索,实现一站式的发现与获取服务。

2）采用 Alma 系统的清华大学图书馆、北京师范大学图书馆、东北师范大学图书馆、南方科技大学图书馆等，都实现了馆藏书目和发现系统的融合，完成一站式检索。

3）中国海洋大学图书馆利用汇文公司的 META 系统实现与发现系统的融合。

3.2 开展在新一代服务平台中电子资源库编目的实践

北京师范大学图书馆白雪等老师探讨了在 Alma 系统中尝试采用 MARC21 标准对数据库信息进行编目的实践过程。设定适用于电子资源库的主要 MARC21 字段著录说明探讨并列举电子资源库本身特有的著录字段和著录项目，包括头标区、007 和 008 控制字段、其他特殊字段等，对字段标准进行规范，以及实践过程的步骤，通过实例展示了经过编目后的电子资源库列表在发现系统读者界面的优化效果[4]。

3.3 制定新一代平台中电子资源管理的工作流程和实践

新一代图书馆服务平台的使用为电子资源全生命周期的系统化管理提供了机遇，但随之而来的也是图书馆电子资源相关业务工作内容的变革及流程的变化，只有因势利导，在对工作内容的变化和平台功能有清晰了解的基础上，图书馆才能制定出最适合本馆的电子资源管理流程，实现对电子资源全生命周期的管理，真正体现出系统管理的优势。北京师范大学图书馆在 Alma 实施的过程中从流程、方式和内容三个方面探讨和总结了电子资源管理工作的变革，并提出了电子资源管理工作的建议和策略[5]。

4 如何开展新一代服务平台建设的一些思考

国内高校图书馆在选择和实施新一代服务平台应该主要考量那些因素，特别在当前智慧图书馆建设的大环境下，如何开展和实施新一代服务平台的建设呢？应该主要从以下因素来考量。

4.1 新一代服务平台中的核心功能应完备

1）关注统一资源管理中的电子资源管理流程和内容，原有图书馆管理系统较好实现了对纸本资源的管理，目前要增加对电子资源，本馆自建特色数字资源的统一管理，特别是电子资源的全生命周期管理。电子资源的全生命周期管理是新一代服务平台中的重点，当前各服务平台中虽然都支持电子资源的管理，但有的比较简单，还应该与图书馆一起加强细节的设置，不断优化。各图书馆也应该积极开展电子资源管理，并在实践中制定新的工作流程和规范。

2)中央知识库元数据管理应完整且准确，利用中央知识库，对所有元数据管理。中央知识库中的元数据是纸电统一管理的数据基础，考量新平台，对中央知识库的元数据的完整准确性是重要的，考察各服务平台的数据合作商，同时也要关注元数据的版权。

3）与发现系统的融合，考虑与发现系统的融合，很少有学术图书馆会像 10 年前那样，单独采购发现服务。现在更接受与其服务平台捆绑在一起的发现产品，如 Alma 与 Primo 的组合，或 WorldShare Management Services 与 WorldCat 发现服务的组合，或利用与其 ILS 供应商有合作关系的发现服务，通常为 EBSCO 的 EDS 服务。

4）开放平台，利用 SOA 或微服务架构，开放 API 数据接口，实现与多个系统平台的交互。

4.2 能扩展外延和其他系统融合

在智慧图书馆建设的背景下，新一代服务品台应支持大数据平台建设，能整合各平台数据，实时监测统计，并能支持 可视化分析和智能决策。同时也能够与座位系统、空间预约、读者活动管理等多个系统相融合。

4.3 适合自身应用需求

“没有完美的软件，只有合适的系统”， 适合自己的才是最好的，新一代图书馆服务平台的应用直接关系到图书馆的资源整合与揭示程度。因此，图书馆在选择平台时，不能照搬其他高校的经验，不能一味地推崇最先进、最强大的系统。在选择国外系统时要充分考虑性价比、售后服务及本地化策略，在选择国内系统时要重点考虑技术支撑与发展前景，再根据校情、馆情，梳理好用户、资源、服务与系统之间的关系，选择最适合自己的系统。

PU-144

基于 DEA-Malmquist 的高校图书馆电子资源采购效率评价

郭雪梅
天津医科大学

0 引言

近年来，随着信息爆炸时代的来临，高校图书馆电子文献资源建设逐步发展与成熟，越来越多的读者对于电子资源的需求越来越高。与此同时，各个高校图书馆在电子资源的经费投入逐渐增大。然而，面对经济的变化以及电子资源每年持续的涨幅，图书馆用于电子资源采购的经费显得捉襟见肘。那么如何在经费削减而需求却逐年攀升的情况下，合理配置电子资源，成为高校图书馆面临的最棘手的问题。因此，对于所采购的电子资源的使用效率进行评估，调整采购结构，优化电子资源配置显得越来越重要。

1 相关研究

任俊霞[1]等通过分析哈尔滨工业大学面向的 14 个学科建设的数字资源绩效，以学科数据以及经费作为投入量，学科全文下载量以及论文发表量作为输出量，构建动态 DEA-Malmquist 模型绩效评价方法；沈思[2]等运用 DEA-CCR 及 DEA-BCC 模型分析西安科技大学图书馆 2013 年所购买的 8 种数字资源；贺意林[3]应用 DEA 模型中的 CCR 模型、BCC 模型以及超效率模型对四川轻化工大学图书馆 2018 年购买的 7 种中外文电子资源进行绩效评价；郭雪梅[4]对图书馆所购买的 6 个外文电子资源在 6 个不同学科中的实用性效率进行评价；马云辉[5]等通过使用 ERS 电子资源利用绩效分析平台，通过分析苏州大学 2018 年在 Web of Science 数据库中发文情况对图书馆所购电子资源进行绩效分析；蒲筱哥[6]等运用网络分析法实证分析了江苏师范大学图书馆采购的 4 中外文期刊数据库，其中设置了 5 个准则指标层，及 22 个主要影响绩效评价的因素指标层；齐春泽[7]通过建立以资源内容、检索、使用及成本 4 方面构成的评价指标体系，利用熵权 TOPSIS 法对某高校的电子资源绩效进行评价；李育平[8]通过构建“投入—状态—产出”模型，对图书馆电子资源服务绩效进行评价；Muhammad[9]以读者满意度为基础，对巴基斯坦大学图书馆的电子资源与服务的有效性进行评估；Noh[10]根据 DEA-CCR 模型和 DEA-BCC 模型，建立一个效率为 100%的参考群体，然后分析了导致决策单元效率低下的因素，分析了电子资源的引入与拒绝对效率的影响。并根据不同的图书馆管理机构比较了效率的差异；Bernardo[11]采用网络两阶段数据包络分析来研究图书馆的基本投入、中间产出和最终产出之间的关系，来对意大利国家公共图书馆进行整体绩效评估。Ramón[12]，通过利用三阶段 DEA 法分析对教学中的技术效率进行评价；Hsieh 等[13]结合台湾大学图书馆所建立的两种电子图书馆的投入与产出，构建一个高校电子图书馆的成本效率与服务效益模型，以合理分配电子资源成本，提高服务质量，满足读者对现有电子资源的需求。

2 研究方法及数据指标选取

对比国内外学者在对电子资源采购效率评价时都会选取篇均成本作为评价指标之一，或是单一通过静态或动态指标对某一年的采购效率进行评价。本文将使用 DEA-Malmquist 模型结合静态分析及动态分析，对 2015-2019 年天津医科大学图书馆购买的 8 个主要外文出版社的电子资源采

购效率进行定量实证分析，以期建立更具说服力的评价体系。

2.1 研究方法

（1）数据包括分析法

数据包括分析法（Data Envelopment Analysis,简称 DEA）是一个对多投入多产出指标的决策单元进行效率评价的方法。本文拟采用 DEA-BCC 模型对电子资源采购的效率进行静态分析。

假设要评价 *j* 个决策单元（DUM），x_{ij}为第 *j* 个决策单元的第 *i* 个投入量，且$x_{ij} \geq 0$；y_{rj}为第 *j* 个决策单元的第 *r* 个产出量，且$y_{rj} \geq 0$；对于第$j(j = 1,2\ldots., j)$个单元，θ表示目标规划值，ε表示非阿基米德无穷小，δ_j表示决策变量，s_i^-、s_i^+表示投入与产出的松弛变量。若$\theta = 1$，$s^- = 0, s^+ = 0$，则决策单位 DEA 有效；若$\theta < 1$，则决策单元 DEA 无效。若$\theta = 1$，$s^- \neq 0, s^+ \neq 0$，则决策单元 DEA 弱有效。表达式为

$$min\left[\theta - \varepsilon\left(\sum_{i=1}^{i} s^- + \sum_{r=1}^{r} s^+\right)\right]$$

$$s.t. \quad \sum_{j=1}^{J} x_{ij}\,\delta_j + s^- = \theta x_i^j \qquad i=1,2\ldots,I$$

$$\sum_{j=1}^{J} y_{lj}\,\delta_j + s^+ = y_r^j \qquad r=1,2\ldots,R \qquad (1)$$

$$\delta_j \geq 0 \quad j = 1,2\ldots,J$$

（2）Malmquist 指数法

Malmquist 指数法是用来衡量生产力的变化，并将这种生产力变化分解为技术变化和技术效率的变化。其表达式为

$$m_o(y_{t+1}, x_{t+1}, y_t, x_t) = \left[\frac{d_o^t(x_{t+1},y_{t+1})}{d_o^t(x_t,y_t)} \times \frac{d_o^{t+1}(x_{t+1},y_{t+1})}{d_o^{t+1}(x_t,y_t)}\right]^{1/2} \qquad (2)$$

$$Effch = \frac{d_o^t(x_{t+1},y_{t+1})}{d_o^t(x_t,y_t)} \qquad (3)$$

$$Tech = \left[\frac{d_o^t(x_{t+1},y_{t+1})}{d_o^t(x_t,y_t)} \times \frac{d_o^{t+1}(x_{t+1},y_{t+1})}{d_o^{t+1}(x_t,y_t)}\right]^{1/2} \qquad (4)$$

$$Tfpch = Effch \times Tech = (Pech \times Sech) \times Tech \qquad (5)$$

这表示生产点(x_{t+1}, y_{t+1})相对于生产点(x_t, y_t)的生产率，其价值大于一个将表明 TFP 从周期 *t*

到周期 *t+1* 的正增长。该指数是基于 Malmquist TFP 的两个输出指数的几何意义。一个指数使用周期 *t* 技术，另一个指数使用周期 *t+1* 技术。若$m>1$，表示效率提高，若$m<1$，表示效率降低。

2.2 数据来源及指标选取

考虑到指标数据的准确性以及可行性，本文对 2015-2019 年所购买的 8 个外文出版社电子资源（后称电子资源）采购效率进行测算分析。最终选取包括 2 项投入指标和 3 项产出指标（见表 1）作为效率分析指标体系。投入指标包括电子资源每年采购的金额，及采购电子资源包含期刊种数（因电子资源采购模式的不同，有的电子资源是通过购买学科包的形式进行采购，所以不同学科包中可能会有重复期刊，重复期刊将进行去重统计），该研究数据均根据中国高校图书馆数字资源获取联盟（digital resource acquisition alliance,DRAA）[14] 网站 2015-2019 年数据整理计算得出；产出指标包括各电子资源年下载量、及以第一单位或通讯单位发表 WOS 文章归类于不同出版社的篇数，这里选取 2015-2019 年天津医科大学为第一单位或通讯作者单位（含共同通讯作者单位）发表的英文论文被 WOS 的收录情况。经过检索，共有 9562 篇文献在这 5 年被收录，其中作为第一单位发表的文献 5804 篇，作为通讯作者单位发表的文献 4882 篇。去重后共有 6119 篇文献是作为第一单位或共同通讯作者单位被收录于 WOS。通过对期刊的分析发现，6119 篇文献所在期刊，其中 611 种期刊属于 OA 期刊，896 种期刊属于非 OA 期刊。由于 OA 期刊属于开放获取类期刊，读者在阅读时不受时间以及空间的限值，顾此次研究暂不对 OA 期刊进行讨论分析。再对非 OA 期刊的 2905 篇文献及其所在期刊的出版社进行分析，得到所覆盖出版社文章篇数及期刊数。

表 1 电子资源采购效率评价体系

Table 1 Evaluation system on the efficiency of purchase digital resources

类别层	指标层
投入指标	电子资源采购金额 采购电子资源包含期刊种数
产出指标	文献下载量 以第一单位或通讯单位发表 WOS 文章所覆盖电子资源篇数 所覆盖电子资源期刊数

3 外文电子资源采购效率分析

3.1 DEA 模型静态分析

利用 DEAP2.1 对 2015 年至 2019 年天津医科大学图书馆购买的 8 个电子资源进行采购效率测算分析，分别得到综合效率值、技术效率值、规模效率值，以及规模收益，结果见表 2。

（1）从综合效率分析 从 2015-2019 年整体来看，综合效率值呈总体上升，中间年份下降的趋势，但均未达到 DEA 有效，并且可以看出，自 2016 年起综合效率在大幅上升，并 2017 年达到增长峰值，2018 年又急速下降，而在 2019 年又稍有回升。从单一电子资源来看，Springer 和 Science 在 5 年中综合效率值始终为 1，说明 DEA 有效，采购效率合理；Wiley 的综合效率值是逐年增长的，在 2018 年达到 DEA 有效，并保持有效，采购效率逐步趋于合理。以上三个电子资源说明配置达到或逐步达到最优，采购资金投入及采购资源配置的结构合理，投入与产出达到了最优效果,最终体现为采购效率合理。Cell 在前四年保持了 DEA 有效，但从 2019 年开始效率大幅降低，同出现大幅降低的还有 SD 及 Ovid；Nature 及 ACS 则是经历了效率触底再反弹的过程。这说明这些电子资源在购买投入方面可能存在采购资金或采购资源配置的结构不合理的情况，应重新调整采购结构，以重新达到最佳效率。通过整体均值得出，采购效率最优为 2017 年。

（2）从技术效率分析 从 2015 年到 2019 年纯技术效率呈总体上升，并趋于平稳的态势。Springer、Science 的纯技术效率始终保持在 1，属于 DEA 有效，说明其采购效率合理。Wiley、Nature 的纯技术效率在 5 年中逐步上升，最终达到 DEA 有效。Ovid 仅是在 2016 年达到 DEA 有效，而在其他年份均无效。SD 在 2015 年、2018 年、2019 年及 Cell 在 2019 年技术效率均为 1，

而综合效率都小于 1，这种情况属于 DEA 弱有效，就是说产出相对于投入达到了最大化，在采购资金投入不变的情况下，应考虑调整规模，即采购资源结构、优化资源配置，提升规模效益，从而达到综合效率的提升。

（3）从规模效率分析 与综合效率相似，在 2016 年和 2017 年达到效率峰值，自 2018 年开始大幅下降，并且可以看出 2017 年 SD、Springer、Cell、Science、ACS 均达到效率最优，Wiley、Ovid、Nature 则都是规模效率递增的态势；而在其他 4 年里，几乎都是处于规模效率递减，并且存在技术效率与规模效率均小于 1，属于 DEA 无效，这说明电子资源存在采购资金或采购资源配置的结构不合理的情况，未得到有效利用，应重新进行所采购资源的结构优化，提升采购的有效性。

表 2 电子资源 2015-2019 年采购效率值

Table 2 Efficiency index of purchase digital resources in 2015-2019

年份	电子资源	综合效率	技术效率	规模效率	规模收益
2015	SD	0.761	1.000	0.761	drs
	Springer	1.000	1.000	1.000	-
	Wiley	0.519	0.617	0.841	drs
	Ovid	0.239	0.474	0.505	drs
	Nature	1.000	1.000	1.000	-
	Cell	1.000	1.000	1.000	-
	Science	1.000	1.000	1.000	-
	ACS	/	/	/	/
	均值	0.789	0.870	0.872	
2016	SD	1.000	1.000	1.000	-
	Springer	1.000	1.000	1.000	-
	Wiley	0.666	0.668	0.996	irs
	Ovid	1.000	1.000	1.000	-
	Nature	0.722	0.842	0.858	irs
	Cell	1.000	1.000	1.000	-
	Science	1.000	1.000	1.000	-
	ACS	0.826	0.827	0.999	drs
	均值	0.902	0.917	0.982	
2017	SD	1.000	1.000	1.000	-
	Springer	1.000	1.000	1.000	-
	Wiley	0.932	0.935	0.996	irs
	Ovid	0.765	0.769	0.995	irs
	Nature	0.837	0.971	0.862	irs
	Cell	1.000	1.000	1.000	-
	Science	1.000	1.000	1.000	-
	ACS	1.000	1.000	1.000	-
	均值	0.942	0.959	0.982	
2018	SD	0.373	1.000	0.373	drs
	Springer	1.000	1.000	1.000	-
	Wiley	1.000	1.000	1.000	-

	Ovid	0.385	1.000	0.385	drs
	Nature	0.419	0.686	0.611	irs
	Cell	1.000	1.000	1.000	-
	Science	1.000	1.000	1.000	-
	ACS	0.167	0.858	0.195	drs
	均值	0.668	0.943	0.695	
2019	SD	0.389	1.000	0.389	drs
	Springer	1.000	1.000	1.000	-
	Wiley	1.000	1.000	1.000	-
	Ovid	0.396	0.797	0.497	drs
	Nature	1.000	1.000	1.000	-
	Cell	0.670	1.000	0.670	drs
	Science	1.000	1.000	1.000	-
	ACS	0.331	1.000	0.331	drs
	均值	0.723	0.975	0.736	

3.2 Malmquist 指数动态分析

在 DEA 静态分析的基础上，为了更好分析电子资源采购效率的动态变化，通过利用 DEAP2.1 的 Malmquist 指数法测算，分析天津医科大学 2015-2019 年购买的 8 个电子资源采购效率的动态变化。

（1）从时间变化分析 从整体上来看（表 3），2015-2019 年 8 个电子资源的采购 Malmquist 指数平均下降了 8.8%，说明 8 个电子资源采购效率未达到应有的水平，未得到有效的提高。技术效率指数上升了 5.1%，技术进步指数下降了 13.3%，技术进步严重不足是导致 Malmquist 指数下降的主要原因，表明在采购资源配置存在不合理，要重新对所采购的电子资源配置进行合理更新。

（2）从各个年度来看,2017-2018 年度的技术效率指数为 0.600，技术进步指数为 5.210，说明该年度技术进步指数是当年 Malmquist 指数提升的主要因素，当年在资源配置方面做到了最优。其他年度，技术效率指数均大于 1，而技术进步指数小于 1，说明资源资金投入与采购资源配置的结构没有达到最优效率。各年份中纯技术效率指数均值为 1.032，规模效率指数均值为 1.019，均大于 1，都呈上升趋势，而在规模效率中，只有 2017-2018 年度，规模效率指数小于 1，由此可见技术效率受到了规模效率的制约。

表 3 2015-2019 年电子资源采购 Malmquist 指数
Table 3 Malmquist index of purchase digital resources in 2015-2019

年份	技术效率	技术进步	纯技术效率	规模效率	Malmquist 指数
2015-2016	1.660	0.155	1.060	1.567	0.257
2016-2017	1.052	0.820	1.052	1.000	0.863
2017-2018	0.600	5.210	0.979	0.613	3.127
2018-2019	1.165	0.856	1.039	1.122	0.998
均值	1.051	0.867	1.032	1.019	0.912

（3）从各个电子资源效率变化分析 从各个电子资源采购 Malmquist 指数（表 4）看出，SD、Springer、Wiley、Ovid、Science 这 5 个电子资源 Malmquist 指数均大于 1，说明这 5 个电子资源

的采购效率在不断提升。Springer 和 Science 技术效率指数和技术进步指数均大于 1，说明该两项指标共同推动了采购效率的提升；Wiley、Ovid、Nature 的技术进步指数均小于 1，技术效率指数均大于 1，说明采购效率的提升均来自于技术效率的提升；SD 则来自于技术进步指数的提升。而 Nature、Cell、ACS 的 Malmquist 指数小于 1，说明这 3 个电子资源的采购效率没有达到预期效果。

表 4 不同电子资源采购 Malmquist 指数

Table 4 Malmquist index in different purchase digital resources

电子资源		技术效率	技术进步	纯技术效率	规模效率	Malmquist 指数
	SD	0.973	1.109	1.000	0.973	1.079
Springer		1.239	1.029	1.000	1.239	1.275
	Wiley	1.434	0.975	1.128	1.271	1.398
	Ovid	1.258	0.965	1.139	1.104	1.214
Nature		1.000	0.992	1.000	1.000	0.992
	Cell	0.905	1.035	1.000	0.905	0.936
Science		1.000	1.213	1.000	1.000	1.213
	ACS	0.759	0.24	1.000	0.759	0.182
	平均	1.051	0.867	1.032	1.019	0.912

4 结论、建议与不足

4.1 结论

根据 DEA 静态分析，自 2015-2019 年所采购的 8 个电子资源中，Springer 和 Science 始终是处于采购效率合理，Wiley、Nature 则经过纯技术效率的提升，使得采购效率趋于合理。SD、Ovid、ACS 则需要通过重新调整采购资金或采购资源配置的结构，来达到采购效率的最合理状态。Cell 在前四年中一直保持采购效率最优，但自 2019 年由于变化了采购资源的配置，使得采购效率大幅下降，所以应重新调整资源配置，以期重新达到资源采购效率的合理化。

根据 Malmquist 指数动态分析，在 2017-2018 年度资源采购效率因为在资源配置按购买做到了最优从而得到很好地改善。而在其他年度中资金投入与资源配置的结构不合理，则需要通过改善结构，使其效率合理化。从单个电子资源的 Malmquist 指数来看，SD、Springer、Wiley、Ovid、Science 这 5 个电子资源采购效率在不断提升，而 Nature、Cell、ACS 则需要通过改变资金投入及资源配置的结构进行调整，以提高效率。

4.2 建议

采购电子资源的时候应在充分了解电子资源所包含期刊学科及期刊质量的同时，系统分析该电子资源的采购模式。打包采购的期刊应先掌握各学科包之间的重复期刊；可单独选刊采购的电子资源，应对每本刊的访问量、拒访量以及是否对应学校相关重点学科进行比较、分析，科学合理的针对学科进行电子资源的采购，充分发挥电子资源对学科的作用。对于包库采购的电子资源，应首先考虑受众群体的需求，并在采购后针对受众进行宣传与指导，使资源得到充分利用。

4.3 不足

由于相关数据获取的难度限制，在选取指标时，一是由于技术所限，产出指标中除下载量外，还应考虑到文献的使用情况；二是指标数据选取的较为客观，并没有反映出读者对电子资源满意度这项主观因素。故今后的研究重点将该两部分内容加入研究指标中，使得针对采购效率的评价是否更为精确合理。

PU-145

改进物元的开放存取资源质量综合评价模型

杨筠 宋蕊 黄晨晨 孙艳 吴涛
第四军医大学图书馆

1 引言

开放存取（open access，简称 OA）是一种通过互联网将科研成果进行自由传播的技术手段，它又可以被命名为开放存取、公开获取等，目前在科技研究领域、学术出版领域、信息传播领域普及效果较好。OA 的起源是二十世纪 90 年代美国 Los Alamos 国家重点实验室自主创建的一个用以公布原创高能物理学论文研究成果的网站。之后，随着网络的高速发展，OA 这一新的学术理念受到国内外研究者更多关注并得以发展等[1-2]。

OA 资源的“阅读免费”的形式，得到高校图书馆的广泛关注。另一方面 OA 资源以其成本低廉、出版及时、免费获取、加速知识传播、加速研究周期等优势吸引着科研人员。但也有不少业界人士对ＯＡ资源持保留、怀疑甚至反对的态度，其原因很大程度上是ＯＡ资源的质量问题。因此合理有效的开放存取资源质量评价对 OA 资源的发展具有极其重要的意义。

相对于传统的学术期刊，OA 资源具有自身的特点和性质，因此传统学术期刊的评价体系并不适用于 OA 资源，且其发展历史较短，因此尚未有标准的评价体系衡量 OA 资源的质量情况。定性评价与定量评价是传统期刊常用的两种方法，两者相辅相成，形成合理的评价体系。定量评价通常是指利用数学统计的相关知识，通过对采集的数据进行分析，对评价对象给予科学合理的衡量与归类，常用的定量评价方法包括综合评价法、层次分析法等。定性评价基于宏观角度，通过建立科学的指标体系，利用认为手段进行评价，评价结果具有科学性和权威性。常用的定性评价方法有同行评议(peer review)等。定性评价与定量评价两者通常交替使用，互为统一，各具特质。相比定性评价，定量评价的操作过程更简易、方便，评价结果更具体、准确，定性评价方法是分析事物的另一方面，考虑范围更加广泛、全面，定性评价往往会借助定量评价的手段，定量评价也是为了更精确的定性。基于以上考虑，OA 资源的评价过程的建立应充分结合两者的优势，从而实现客观、合理的评价[3]。

国外方面，英国生物医学期刊出版中心(Biology Medicine Central)综合考虑了每一篇论文的引用和被引频数变动以及审稿专家的评价结果，以此来判断其实际影响力水平。Steve Lawrence（2001）、Hajjem（2006）、 Mukherjee（2009）， Lansingh（2009）等多位学者采用引文分析方法，分析 OA 论文的被引用次数进而对相关 OA 期刊加以评价[4-7]。Sombatsompop N（2005）等首次定义了一个将论文自身的影响因子平均值、论文所在学科领域的影响因子、发表期刊的影响力三个变量综合的计算公式[8]。 Hyun Hee Kim 等（2007）建立的综合评价模型从资源系统与网络、操作习惯、使用者及提交者以及管理与政策四个角度对系统加以评价[9]。进而，又采用用户实地调研和实验研究等手段对机构库系统软件加以评价，构建了由满意度、支持度、有用性和实效性组成的 4 维评价框架[10]。国内方面，马翠嫦等（2009）依据 OA 论文引用优势的研究发展历程，整理并学习了目前有关开放获取出版方式(OA)与学术论文被引次数之间关系的权威研究成果[11]。曹兴等（2010）分析互联网上公开发表的科技论文的相关特点，采用层次分析法（AHP），首先从作者、专家和读者三个角度入手建立准则层，进而通过统计检验的方法确定 14 个参与评价的指标，并计算其相应的权重系数，最终构建出论文影响力评价指标体系[12]。沈小玲等（2013）的研究思路是将网络科技论文影响力评价与论文自身排名进行关联，选取 WOS 中数学类论文作为数据样本，采用文献信息方法提取论文特征属性，最后基于 Lasso 等算法建立 14 个论文网络传播的科技论文学术影响力评价指标[13]。

目前，学者们在 OA 学术论文质量控制研究领域已取得了初步的成果，但还未能建立成熟完备的评价机制，主要体现在主观性强，评价方法单一，而且容易忽略单个指标对评价结果的影响。为克服上述不足，本文采用物元模型进行了综合评价尝试，采用熵权法和层次分析法优化组合计算各指标权重，结合实际开放存取资源背景构建改进物元评价模型从而为后续进行 OA 质量提高及广泛应用提供理论支撑与决策依据。

2 开放存取的类型

研究人员对目前开放存取资源进行分类，将 OA 资源主要分为两类，分别是 OA 期刊和 OA 知识库。OA 期刊（open access journals）,是一种论文经过同行评审的、网络化的免费期刊，在文

章评审、文章出版和编辑部资源维护中会产生相应的费用，这些费用的承担者并非读者本身，而是在出版过程中涉及的作者或相关机构，OA 期刊也称开放存取期刊、金色 OA。

OA 期刊具有与传统期刊相同水平的论文质量，为了保证期刊的质量，所有向期刊投稿的学术论文都必须同样进行严格的领域专家评审，但是相较于传统印本期刊，OA 期刊最大的不同是 OA 期刊多以网络电子形式为主，因此可有效降低出版与传播成本。OA 期刊的运行模式与传统期刊不同，其付费方式为“ 作者(或机构)付费出版, 读者免费使用”。除此之外, OA 期刊也逐渐的获得了传统文摘索引服务商的关注和认可，发展为它们收录的对象。比如在 PubMed 检索系统中可以对 PMC、HighWire Press 等 OA 资源进行检索，浏览题录和文摘，并提供全文链接[14]。

根据 OA 期刊资源的免费时间不同，可以分为延时 OA 期刊（Delayed Open Access Journals）,它的无偿使用时间是指必须在期刊出版之后的一段时间;部分 OA 期刊（Partial Open Access Journals）,它的特点是其中收录的论文资源只有一部分是可以无偿使用或是无偿被一些国家地区所使用的；完全 OA 期刊（Open Access Journals），它的不同之处就是其中收录的全部的论文资源在任何时间地点都是可以无偿使用的。作者付费，读者免费是开放存取期刊的最基本特点。开放存取期刊依托网络技术、有严格的同行评审制度、有些有相关机构的赞助、处于不断发展和壮大阶段、数量不断增加，目前著名的开放存取期刊来源主要集中在生物和医学领域。

开放存取知识库（open access archives）是开放存取的新兴实践方式，也称 OA 知识库、e 印本文库（或称为 OA 仓储、OA 存档库等），绿色 OA，是为用户提供电子版学术文献存储和检索的数据库。在资源质量方面，OA 期刊有质量保证，OA 知识库有经济优势。OA 期刊比 OA 知识库的运行费用要高得多，这是因为 OA 期刊通过同行评审(peer review)模式来对其资源质量进行控制，因此用于支付同行评审的费用也成为了 OA 期刊的主要成本。与此不同的是，OA 知识库主要是采用公众审查(public peer review)的方式，即学者们可以自由的进入 OA 知识库进而表达自己的看法，与原创作者间互相交流想法和意见。这种方式相比于 OA 期刊的同行评审模式来说，评审经费得到了一定的节省，但对其质量的控制还处于探索阶段。因此针对学者对开放获取资源质量的要求，开放存取知识库尚不能满足学者对其质量的追求，因此对于开放获取资源质量评价，本文主要从 OA 期刊入手，对开放存取资源质量综合评价模型进行探索[15]。

3 评价指标

评价指标是评价模型的重要部分，若要实现合理有效的开放存取资源质量评价结果，需要设定合理的评价指标对开放存取资源质量进行评价。已有学者对此进行研究，但可能存在指标信息冗余、覆盖面不全等问题，因此本文基于目前已有的评价指标体系，构建了充分完备的开放存取资源质量综合评价指标体系。指标体系涉及三部分，分别是 OA 期刊自身影响力因素、OA 期刊读者影响力因素、OA 期刊网络影响力因素[16-17]。

3.1 OA 期刊自身影响力因素

（1）期刊规划

任何期刊出版商都必须制定短期的计划和长期的发展战略规划，既要能够充分反映期刊特点，又要迎合市场规律，做到持续长久的发展。

（2）出版机构的权威性

任何期刊出版商都必须具有良好的社会信誉，聘用声望较高的主编和国际编委。

（3）严谨的审稿规则

OA 期刊必须对论文的质量进行严格的把关控制，和传统学术期刊一样，通过权威领域专家或同行业内人士进行评审，确保论文质量。

（4）标准规范化

严格遵守国际国内通用的标准和规范。

（5）版权保护问题

和常规商业化出版形式有所不同的是 OA 可以允许作者“保留部分版权”。

（6）载文量

OA 期刊通过刊载的相关学科领域论文数量的高低来表征期刊的质量差异，载文量越大就代表期刊的稿件来源越丰富、影响力越大。

3.2 O A 期刊读者影响力因素

（1）页面浏览量：它所指的就是公开发表的论文被查阅的次数，其数量的高低可以充分表征 OA 学术论文的影响力，是衡量论文质量高低的重要指标之一。

（2）读者 IP 访问量：论文在所发表网页上的绝对访问人数，每一个 IP 访问者只对应着一个独立用户。

（3）摘要点击量。它所指的就是读者在阅读 OA 期刊网站上发表的论文时，对摘要的浏览次数。

（4）全文下载量：它所指的就是在 OA 论文所发表的网络中，读者对论文全文的下载次数，相当于传统学术期刊的发行量，也是衡量论文影响力高低的重要依据。

（5）读者参与互动评论程度：它所指的就是读者对 OA 论文的互动水平，这一指标值的可以表征 OA 论文的影响力高低、读者的回应力高低以及学术信息生产和消费的水平。

3.3 OA 期刊网络影响力因素

（1）他引频次：它所指的就是其他论文对该篇 OA 论文的引用次数。

（2）链接量：它所指的就是 OA 期刊网站站内和站外的链接数量总和，也可以间接表征论文的重要程度。

4 评价模型

物元模型是由中国学者蔡文在二十世纪八十年代初期首次提出，采用形式化模型研究事物拓展的可能性和开拓规律，通常适用于处理不相容的复杂系统问题，适合于多指标综合评价[18]。针对状态评价等级这一概念本身具有模糊性，且状态评价过程中所需的指标较多时，每一个独立指标的评价结果通常存在不相容现象，因此可采用物元分析法构建 OA 期刊综合评价模型，探索建立合理有效的评价方法。

4.1 开放存取资源物元模型

物元分析中所描述的对象 N 及其特征向量 c 和特征量值 v 共同构成开放存取资源物元，如下所示。

$$R=(N,c,v) \qquad (1)$$

如果对象 N 有 n 个特征向量 $c_1,c_2,\cdots c_n$ 及其对应的量值 $v_1,v_2,\cdots v_n$，则称 R 为 n 维开放存取资源物元模型，可以表示为：

$$R=\begin{vmatrix} N & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{vmatrix}=\begin{vmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{vmatrix} \qquad (2)$$

4.2 经典域与节域物元矩阵

开放存取资源的经典域物元矩阵定义为为：

$$R_{oj}=(N_{oj},c_i,v_o)=\begin{vmatrix} N_{oj} & c_1 & (a_{oj1},b_{oj1}) \\ & c_2 & (a_{oj2},b_{oj2}) \\ & \vdots & \vdots \\ & c_n & (a_{ojn},b_{ojn}) \end{vmatrix} \qquad (3)$$

式中，R_{oj}——经典域物元；N_{oj}——所划分开放存取资源质量等级的第 j 个评价等级；c_i——第 i 个评价指标；(a_{oji},b_{oji})——第 i 个评价指标对应等级 j 的划分区间，即经典域。

开放存取资源的节域物元矩阵定义为：

$$R_p=(N_p,c_i,v_p)=\begin{vmatrix} N_p & c_1 & (a_{p1},b_{p1}) \\ & c_2 & (a_{p2},b_{p2}) \\ & \vdots & \vdots \\ & c_n & (a_{pn},b_{pn}) \end{vmatrix} \qquad (4)$$

式中，R_p——节域物元；(a_{pi},b_{pi})——节域物元关于特征 c_i 的量值范围。

4.3 待评价物元

把待评价对象 N_x 的物元表示为 R_x：

$$R_x=\begin{vmatrix} N_x & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{vmatrix} \qquad (5)$$

4.4 改进关联函数计算关联度

依据经典的物元模型理论，皮肤指标 c_i 对某一等级 j 的关联函数 $K_j(c_i)$ 的定义为：

$$K_j(c_i)=\begin{cases}\dfrac{-\rho(c_i,X_{oji})}{|X_{oji}|} & c_i\in X_{oji}\\ \dfrac{\rho(c_i,X_{oji})}{\rho(c_i,X_{pi})-\rho(c_i,X_{oji})} & c_i\notin X_{oji}\end{cases} \quad (6)$$

式中 ρ 为距，可用下式表示：

$$\rho(c_i,X_{oji})=\left|c_i-0.5(a_{oji}+b_{oji})\right|-0.5(b_{oji}-a_{oji})$$
$$\rho(c_i,X_{pi})=\left|c_i-0.5(a_{pi}+b_{pi})\right|-0.5(b_{pi}-a_{pi}) \quad (7)$$

其中，$\rho(c_i,X_{oji})$为点 c_i 与经典域区间 $X_{oji}=[a_{oji},b_{oji}]$ 的距离；$\rho(c_i,X_{pi})$为点 c_i 与节域区间 $X_{pi}=[a_{pi},b_{pi}]$ 的距离；$|X_{oji}|=|b_{oji}-a_{oji}|$，$X_{oji}$、$X_{pi}$ 分别为待评对象开放存取资源物元模型的第 i 个指标的经典域和节域。

显然，当 $c_i=0.5(a_{oji}+b_{oji})$ 时，关联度值会达到最大，这表明在经典的综合评判物元模型中默认为最佳点位置就是各等级划分区间的中点。然而实际上，对开放存取资源研究中的大多数指标，其最佳的位置可能越小越好，也可能越大越好，并非一定是中值最好。因此，本文结合应用特点，提出构造非中值最佳点位置，如图 1 所示。

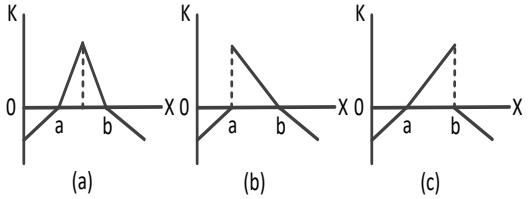


图 1 开放存取资源最佳位置点及关联度函数图

图中（a）为经典物元模型的最佳点位置及关联度函数图；（b）为指标值越小越好情况下的最佳点位置及关联度函数图；（c）为指标值越大越好情况下的最佳点位置及关联度函数图。依据上述分析，给出适合开放存取资源质量评价的改进物元模型关联度函数[19-20]。

①对值越小越好的指标

$$\rho(c_i,X_{oji})=\left|c_i-a_{oji}\right|-(b_{oji}-a_{oji}) \quad (8)$$

②对值越大越好的指标

$$\rho(c_i,X_{oji})=\left|c_i-b_{oji}\right|-(b_{oji}-a_{oji}) \quad (9)$$

显然，应用中可依据评价指标最佳点为居中型或极值型，动态调用 ρ 距计算公式。同理可以得到改进模型中 $\rho(c_i,X_{pi})$ 的表达式。

4.5 优化指标权重

指标权重的计算在综合评价算法中尤为关键，现有的计算权重的方法很多，主要包括层次分析法（AHP）、主成分分析法、德尔菲法（Delphi）、熵权法等。其中，熵权法的计算原理是通过信息熵求解出各指标间的差异系数，进而分析差异系数对整体的影响来判断指标的权重值，结果相对客观；AHP 是一种基于专家的主观经验和知识，对定性问题进行定量研究的多属性决策方法，主观随意性相对较大[21-22]。因此，为了综合考虑主客观信息对开放存取资源的影响，本文对熵权法和 AHP 进行优化组合赋权以确定各指标权重。

假设熵权法计算出的各指标权重为 $(v_1,v_2,\cdots v_n)$，AHP 计算出的权重为 $(u_1,u_2,\cdots u_n)$，根据组合权重值与各单一赋权法所求得的权重值之间的偏差应尽量一致的思想，建立基于熵权法和 AHP 的权值组合优化模型：

$$\min f=\sum_{i=1}^{n}[(w_i-u_i)^2+(w_i-v_i)^2] \quad (10)$$

约束条件为：

$$w_i\geq 0$$
$$\sum_{i=1}^{n}w_i=1 \quad (11)$$

此模型可以转化为一个条件极值问题的求解过程，首先建立拉格朗日 Lagrange 函数：

$$L(s_i,\lambda)=\sum_{i=1}^{n}[(w_i-u_i)^2+(w_i-v_i)^2]+\lambda(\sum_{i=1}^{n}w_i-1) \quad (12)$$

然后根据极值存在的必要条件，分别对 w_i， λ 求一阶偏导数，并分别令所对应的偏导数为零，得到：

$$\begin{aligned}&\frac{\partial L}{\partial w_i}=4w_i-2u_i-2v_i+\lambda=0\\&\frac{\partial L}{\partial \lambda}=\sum_{i=1}^{n}w_i-1=0\end{aligned} \quad (13)$$

其中， $i=1,2,\cdots,n$ 。求解公式（12）可得：

$$\begin{aligned}&\lambda=\frac{1}{n}\left[2\sum_{i=1}^{n}(u_i+v_i)-4\right]\\&w_i=\frac{1}{4}(2u_i+2v_i-\lambda)\end{aligned} \quad (14)$$

最后根据公式(13)，求得组合权值向量为 $W=[w_1,w_2,\cdots w_n]^T$ 。

4.6 计算综合关联度确定评价等级

待评价对象 $N_x(x=1,2,3,\cdots m)$ 关于等级 j 的综合关联度 $K_j(N_x)$ 为

$$K_j(N_x)=\sum_{i=1}^{n}w_iK_j(c_i) \quad (15)$$

其中， $K_j(c_i)$ 为 N_x 关于等级 j 的单指标关联度。若 $K_{jx}=\max[K_j(N_x)]$，则 N_x 处于开放存取资源质量的第 j 等级；若 $K_{ji}=\max[K_j(c_i)]$，则 N_x 的第 i 指标处于开放存取资源质量的第 j 等级。上述两个 K 值越大则说明该评价指标（或对象）在相应的等级中越稳定；反之若 K 值较小，则该评价指标（或对象）就具有向仅次于该等级 K 值的等级转化的可能，且二者的值越接近，转化的趋势越明显。若 K 值在所有等级中均为负值，说明评价对象开放存取资源质量在已设定好的标准等级之外，并具有向最大 K 值对应等级转化的趋势。

5. 模型应用分析

我们选取几个常用的代表性 OA 期刊，将 DOAJ 收录的 5 种生物医学期刊为研究对象，按照本文提出的改进物元模型进行分析，给出以上期刊大致的分类和排序结果，从而实现开放存取资源质量综合评价。

首先，针对测试样本已构建的指标体系，分别采用 AHP 和熵权法对各指标赋权，并参照 4.5 节所述的优化组合赋权法原理计算融合主客观信息后的各指标权重，结果见表 1。

表 1 开放存取资源质量评价指标体系及权重

目标	因素	指标	权重
开放存取资源质量	OA 期刊自身影响力因素	c_1 期刊规划	0.0366
		c_2 出版机构权威性	0.0658
		c_3 严格的审稿制度	0.1289
		c_4 标准规范化	0.0839
		c_5 版权保护问题	0.0528
		c_6 载文量	0.0320
	OA 期刊	c_7 页面浏览量	0.0742

读者影响因素		c_8 读者 IP 访问量	0.0958
		c_9 摘要点击量	0.0795
		c_{10} 全文下载量	0.1230
		c_{11} 读者互动评论度	0.0775
OA 期刊网络影响因素		c_{12} 他引频次	0.0705
		c_{13} 链接量	0.0796

本研究将开放获取数据库 DOAJ 收录的 5 种生物医学期刊交由权威的 OA 领域专家进行打分，并统计相关的指标数据，判断出相应的质量等级。根据指标量值建立待评复合物元矩阵 。

$$R_{mn}=\begin{bmatrix} & 1 & 2 & 3 & 4 & 5 \\ c_1 & 80.28 & 35.44 & 52.63 & 65.42 & 96.89 \\ c_2 & 11.52 & 20.46 & 17.32 & 15.64 & 5.36 \\ c_3 & 34.15 & 23.53 & 25.48 & 29.78 & 29.97 \\ c_4 & 109.12 & 235.49 & 185.53 & 189.83 & 100.55 \\ c_5 & 136.69 & 333.67 & 288.58 & 259.36 & 396.87 \\ c_6 & 69.14 & 69.44 & 63.84 & 66.79 & 60.5 \\ c_7 & 25.96 & 13.97 & 16.59 & 20.52 & 9.78 \\ c_8 & 25.49 & 19.46 & 13.97 & 10.36 & 13.69 \\ c_9 & 6.44 & 5.02 & 19.46 & 3.36 & 2.03 \\ c_{10} & 0.89 & 0.53 & 0.79 & 0.78 & 0.65 \\ c_{11} & 0.016 & 0.02 & 0.017 & 0.6 & 0.1 \\ c_{12} & 0.066 & 0.094 & 0.53 & 0.26 & 0.26 \\ c_{13} & 0.126 & 0.452 & 0.202 & 0.309 & 0.509 \end{bmatrix}$$

将待评物元输入到物元模型，便可以求得相应的评价结果。

表 2 开放获取数据库评价等级比较

综合关联度	I级	I I级	II级	IV级	I 级	等
1	0.157	0.108	-0.111	-0.06	C 级	I
2	-0.285	0.051	C.302	C.249	II级	I
3	-0.143	0.366	C.173	C.289	I 级	I
4	-0.265	0.105	C.299	C.201	II级	I
5	-0.302	0.059	C.108	C.315	V 级	I

为了证明模型评价结果的准确性，我们邀请了多位化学研究领域的学者对入选核心期刊范畴的 5 类期刊展开评价。总体来说，本文所建立的改进物元模型的定量评价得到的核心期刊列表获得了专家的一致认可，认为选出的核心期刊在研究工作的系统性及论文的实用性、创新性等方面均具有很大优势，充分反映了信息用户的实际需求。

6 结语

开放存取资源质量综合评价涉及指标众多，其结果可作为佐证期刊发展程度及未来改进方向的重要依据。本文综合考虑开放存取资源质量的影响因素，构建的 OA 期刊多元评价指标体系具有全面性的优点，采用熵权法和层次分析法优化组合的方法确定各指标权重，并结合实际应用特点，修正距计算公式建立了改进物元模型以对开放存取资源质量进行综合评价。这一算法克服了评价过程中主观人为因素影响较大的弊端，对参评因子进行量化处理，提高了开放存取资源质量评价的精度，表明本文方法的实际应用与理论参考价值，简洁、高效，易于推广，为后续期刊管理水平的改善提供方向性的指导。

PU-146

NSTL 开放获取文献资源保障研究

王勛 王巍洁 文怡力 任慧玲*

北京协和医学院/中国医学科学院医学信息研究所/图书馆

1.引言

二十一世纪初，基于我国科技文献资源匮乏问题国务院批准成立国家科技文献信息资源服务与保障网络虚拟机构——国家科技图书情报中心（Nation Science and Technology Library,NSTL）。九家文献信息研究单位展开密切合作，系统化建设国家科技文献资源战略保障体系，面向全国提供公益化的科技文献信息服务[1] ，其资源建设状况可以从一定程度上反映出我国科技文献资源建设的战略布局。NSTL 适应全球科技信息环境的变化，随着数据和资源的数字化、网络化发展，以开放获取（Open Access,OA）为代表的开放科学（Open Science）运动席卷全球，联合国教科文组织发布《开放科学建议书》，深刻地影响着全球范围内科学研究活动，推动了全球范围内的开放科学发展进程[2]。近年来世界各国积极探索开放科学新途径，我国作为科研大国也多次明确提出要形成具有全球竞争力的开放创新生态。NSTL 于 2016 年开展开放资源的建设工作，并于 2017 年 10 月 26 日，正式签署开放获取 2020（Open Access 2020 Initiative,OA2020）倡议的《关于大规模实现

学术期刊开放获取意向书》（Expression of Interest in the Large-scale Implementation of Open Access to Scholarly Journals,OA2020 Eol）[3]，这一举动表明了我国正积极探索开放获取的资源建设并推动文献资源的国际化合作。

近年来受开放获取的影响，业界出版模式发生了根本性变化。开放获取推动了传统出版模式的变革，科技文献资源在世界范围内实现了大规模的共享，但高程度的开放模式也随之带来一些问题。为应对开放获取模式下发生的一系列变化，保障我国文献资源的安全性和连续性[4]，本文将基于 NSTL 开放资源建设情况，探索 NSTL 在开放获取文献资源的建设路径，并提出相关策略。

2.相关概念概述

开放科学运动兴起至今发展时间较短，其定义在学界也尚未达成定论。Daryl E.Chubin 于 1985 年首次提出开放科学的概念[5]。作为一种科学研究新范式[6]，开放科学涵盖了多方面的内容，学者们基于不同的视角在研究中阐述对其的理解。从流程上来看，黄磊等人认为开放科学是开放数据、开放获取和开放可重复研究等一些系列活动的总称[7]，是以合作为基础，利用先进技术和协作工具，基于信息网络推动信息流动的过程；从方法论来看，开放科学被认为是一种基于合作研究的科学实践新方法[8]，是一种寻求外部投入和公众参与的各种策略总称[9, 10]，通过利用数字技术和新的协作工具促进共同努力，尽可能广泛地分享成果和知识，强调成果共享、过程透明；从文化传播来看，开放科学是一种协作文化，这种文化通过技术实现，使科学界和更广泛的公众能够开放地共享数据、信息和知识，从而加速科学研究和理解[11, 12]。

开放获取是开放科学运动的重要产物。布达佩斯开放获取计划（Budapest Open Access Initiative,BOAI）首次提出了开放获取的概念，指出开放获取是学术内容可以在网络上被公开免费获取，允许任何用户阅读、下载、拷贝、转发、打印、搜索、索引以及链接到这些文章全文的模式[13]。基于全球出版和传播环境的变化，我国致力于以科学和可持续的方式推动传统文献资源出版模式向开放获取转变。2004 年，国家自然科学基金委和中国科学院签署开放获取相关文件《柏林宣言》[14]。2017 年，以 NSTL 和中国科学院文献情报中心为首的国内研究机构签署“开放获取 2020”（OA2020 initiative, OA2020）倡议意向书[15]。2018 年，我国自然科学基金委与 NSTL 公开表示支持由 cOAlition S 联盟颁布的 Plan S 计划。目前开放获取已成为国内外广泛认同的出版模式之一[16]。

3.NSTL 开放获取资源保障概况

3.1 数据采集

数据采集截止至 2023 年 9 月 5 日的 NSTL 书目系统中的全部开放资源，共计 59630 条记录，其中科技图书 21426 册，科技期刊 15453 个，科技报告 11811 篇，会议文献 10651 篇，文集汇编 289 册。

3.2 数据分析

3.2.1 资源类型

NSTL 全方位对理、工、农、医四大领域的科技文献进行资源建设工作，全部文献资源类型包括学术期刊、会议录、学位论文、科技报告、图书、专利、标准和计量规程等[17]，同时大力开展国外开放资源的统一揭示和集成管理，形成了采购文献与开放获取文献协同服务的资源保障新格局。

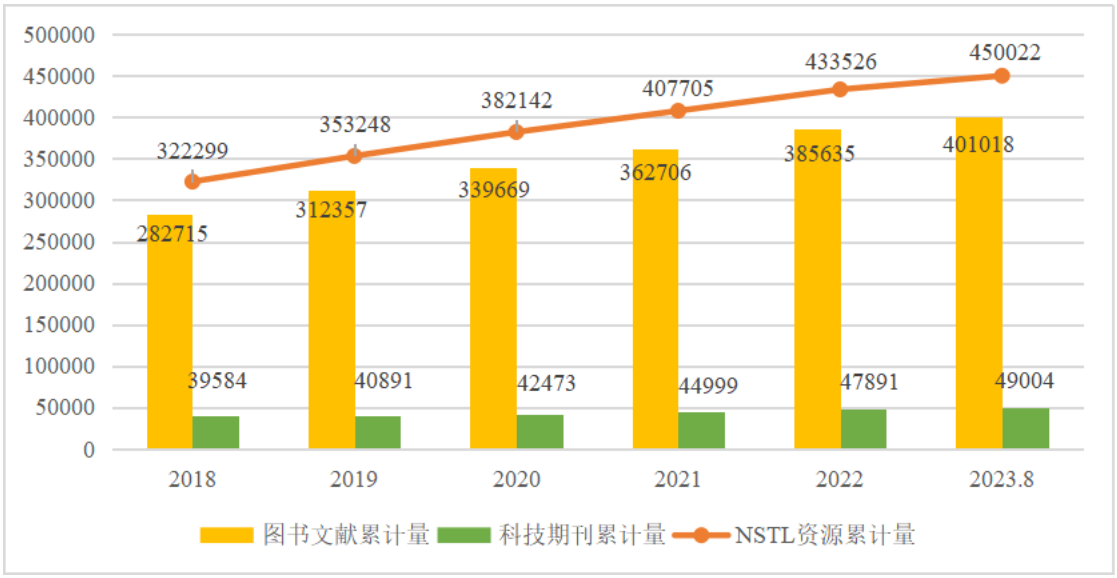


图 1NSTL 收录资源历年总量

截止至 2023 年 9 月，NSTL 收录各种类型资源共 433526 种，其中图书文献资源有 385635 种，科技期刊资源有 47891 种。总体来看，NSTL 文献资源保障数量呈逐年增长趋势，其中图书文献资源总量远高于科技期刊。资源的电子化和开放共享是随着计算机和大数据技术的发展产生的因此 NSTL 收录了大量的印本资源，保障数量远高于电子资源与开放资源。开放资源收录科技图书 21426 册，科技期刊 15453 个，科技报告 11811 篇，会议文献 10651 篇，文集汇编 289 册。

3.2.2 资源来源

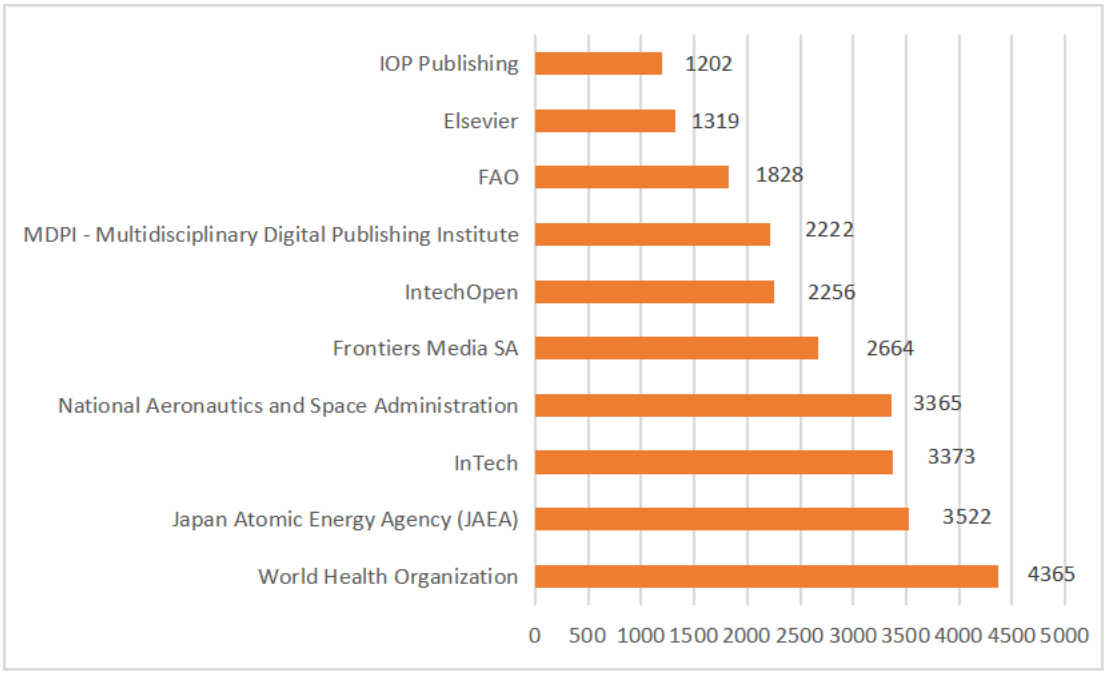


图 2 开放资源出版商分布

知名出版机构所出版的刊物在业内往往较为知名和权威，因此对出版商进行分析更能反映出 NSTL 保障开放资源的质量。开放资源来自国外多家著名出版商 Elsevier、IntechOpen、Springer 等，也包括研究协会出版机构 MDPI - Multidisciplinary Digital Publishing Institute、World Health Organization 等。如图 2 所示，其中来自 World Health Organization 期刊最多为 4365 种。

3.2.3 语言分布

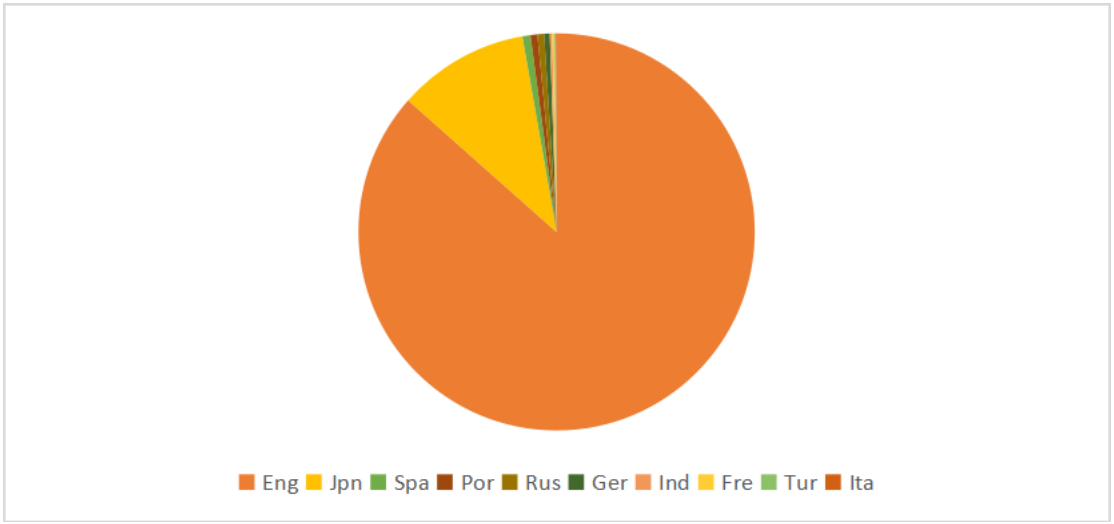


图 3 开放获取期刊语种分布

从期刊语种分布上看，NSTL 的开放资源主要有十个语种，分别为英语、日语、葡萄牙语、西班牙语、俄语、德语、法语、土耳其语和意大利语，如图 3 所示，NSTL 开放资源采集语种以英文为主，英文资源数量为 50157 种，占开放资源总量的 86.16%，其余资源以亚洲语种与欧洲地区小语种为主，仅占开放资源总量的 13.84%。

3.2.4 学科领域分布

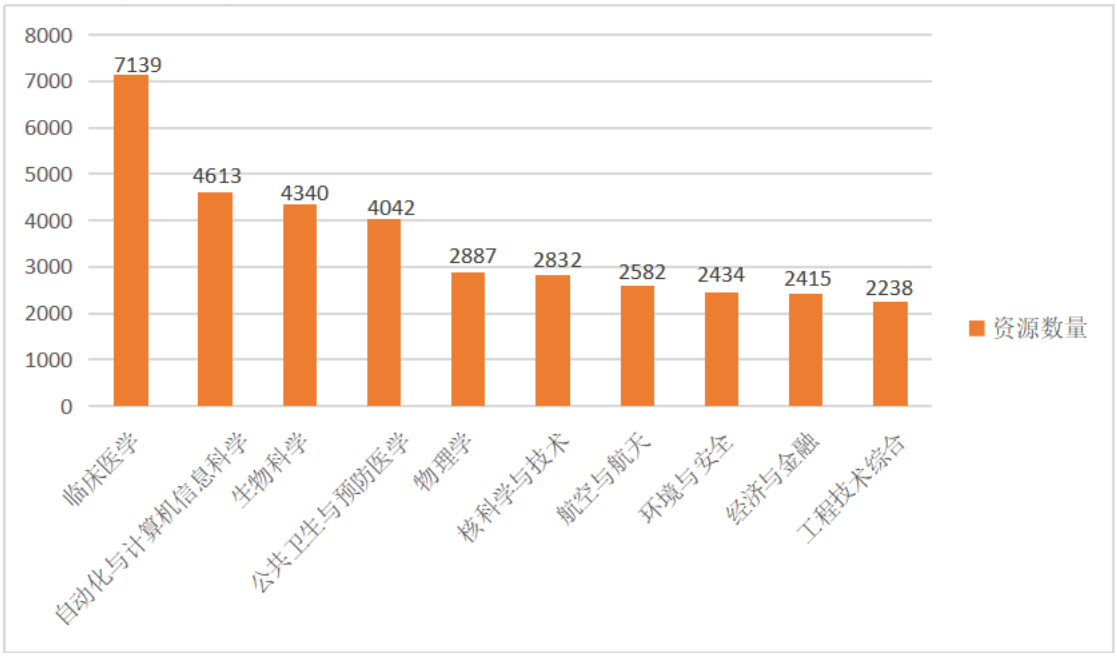


图 4 开放获取期刊主要学科领域分布

NSTL 将学科分为基础科学、工程科学、农业科学、医学科学和社会科学五大类，每个一级学科分别下设二级学科，涵盖了基础学科和交叉学科，基本可以满足科研人员的需要。如图 4 所示，资源保障数量最多的前十学科多为理、工、农、医学领域，其中临床医学学科占比最高。

3.2.5 主题分析

（1）模型构建

LDA 主题模型是 David Blei 提出一种无监督的机器学习方法，它是一个贝叶斯概率模型，主要用于大规模文本的主题挖掘和分类[18, 19]，本文利用 python 编写程序分别对开放资源书目进行主题分析，模型构建思路如图 5 所示。

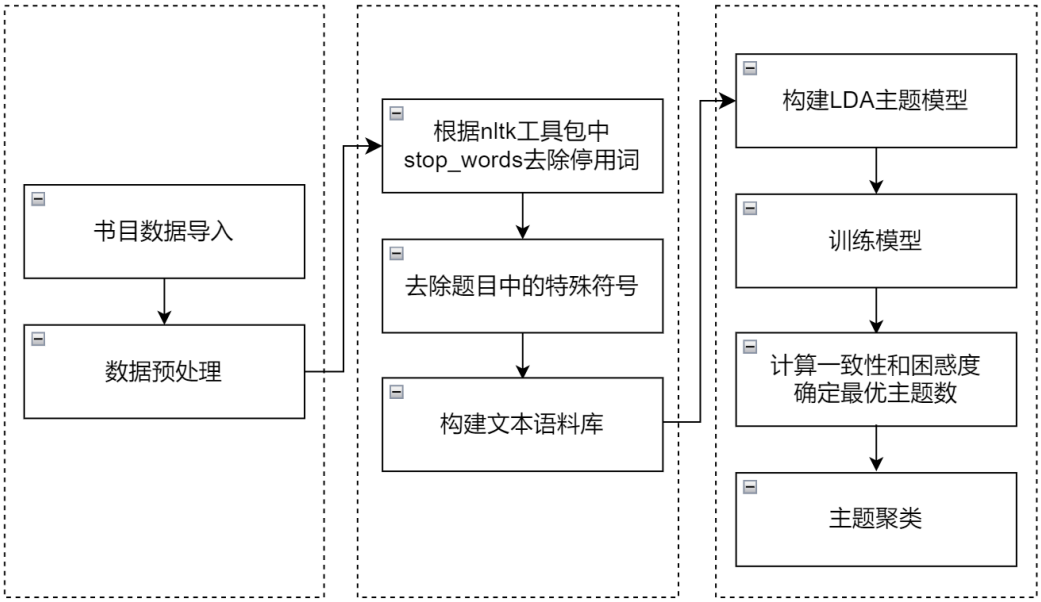


图 5LDA 主题模型构建思路

（2）主题聚类

通过计算选择一致性较高同时困惑度较低的数值作为主题数，利用 LDA 主题模型分别对四种不同的开放资源类型进行主题分析。

①科技图书

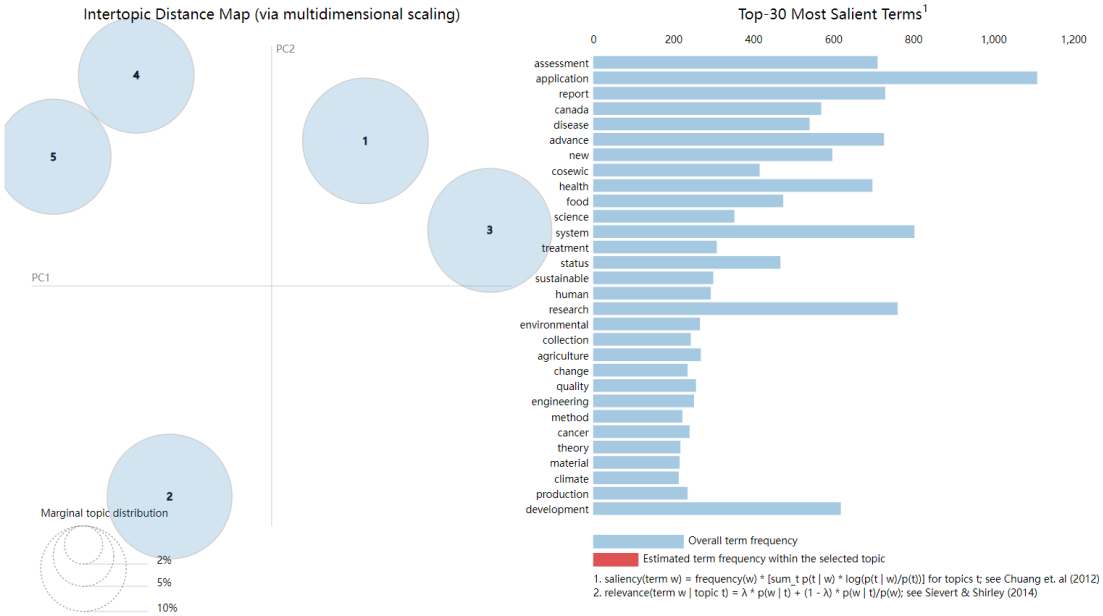


图 6 科技图书主题-主题词分布情况

根据一致性和困惑度结果确定最优主题数为 5，科技图书主题-主题词分布情况如图 6 所示，主题词出现频率较高的分别为 assessment、application、report、Canada、disease。

表 1 科技图书主题数为 5 时分布情况

序号	关键词	主题	归属学科
Topic 1	system、food、development、management、sustainable、agriculture	农业/食品管理系统可持续发展	农学、理学

Topic 2	disease、new、research、treatment、human	人类疾病新的治疗方法	医学
Topic 3	report、assessment、canada、health、cosewic	加拿大卫生评估报告	医学
Topic 4	science、application、research、collection、change、theory、climate	收集天气变化科学理论与应用研究	理学
Topic 5	application、advance、technology、material、advanced	在研与先进的科技材料应用	工学

②科技期刊

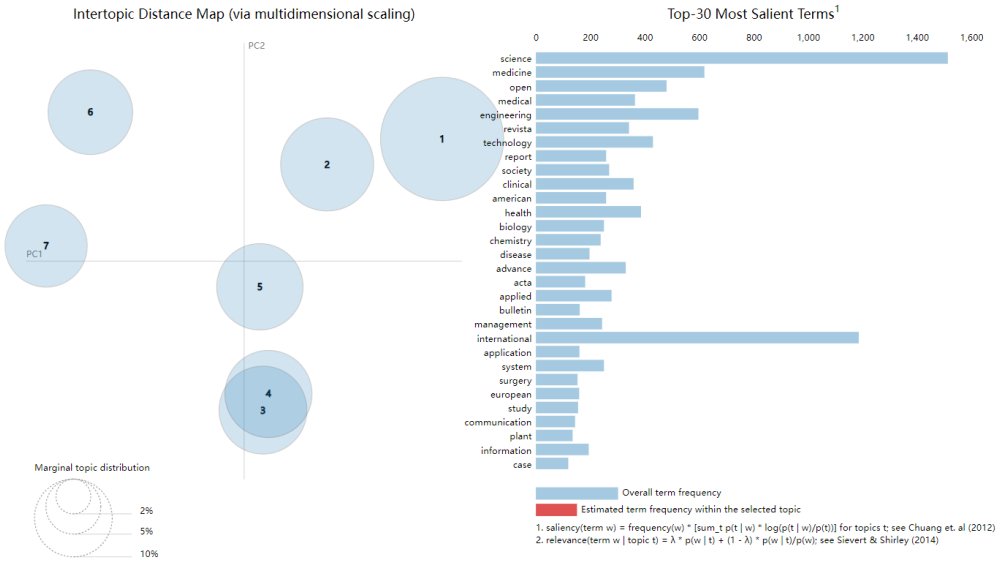


图 7 科技期刊主题-主题词分布情况

根据一致性和困惑度结果确定最优主题数为 5，科技期刊主题-主题词分布情况如图 7 所示，主题词出现频率较高的分别为 science、medicine、open、medical、engineering。

表 2 科技期刊主题数为 7 时分布情况

序号	关键词（前五位）	主题	归属学科
Topic 1	science、international、engineering、research、technology	国际范围内工程技术的科学研究	工学
Topic 2	medicine、advance、management、European、study	欧洲先进的医学管理学习经验	医学、管理学
Topic 3	open、chemistry、research、communication、plant	植物化学研究的开放交流	理学
Topic 4	clinical、society、research、india、molecular	印度社会的临床研究	医学
Topic 5	biology、disease、application、surgery、international	国际上生物学疾病手术	医学
Topic 6	revista、America、acta、bulletin、review、iotechnology	美国生物医学报/公报评论	医学、理学
Topic 7	medical、report、case、Internet、letter	医疗在线报告/案例	医学

③科技报告

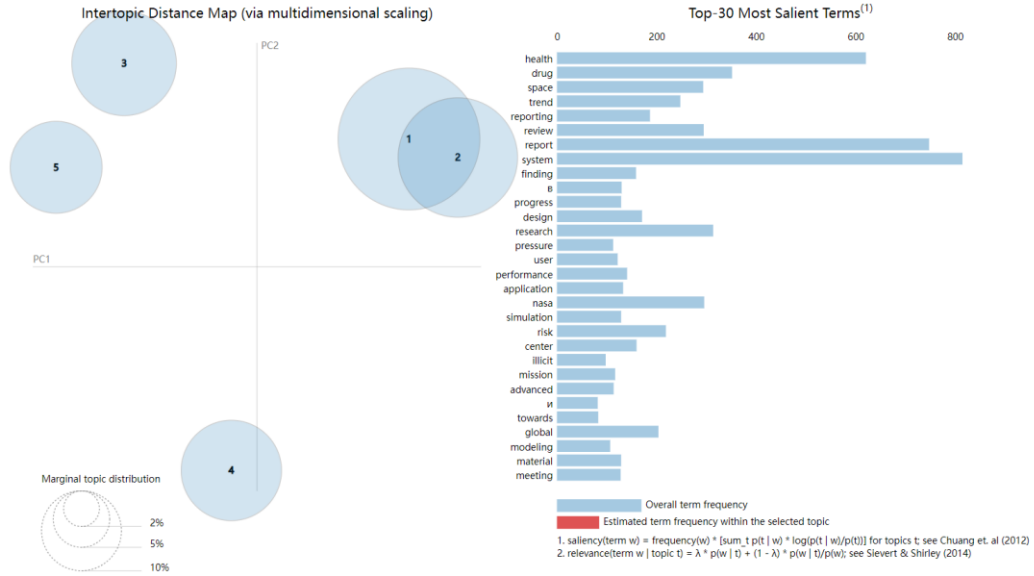


图 8 科技报告主题-主题词分布情况

根据一致性和困惑度结果确定最优主题数为 5，科技报告主题-主题词分布情况如图 8 所示，主题词出现频率较高的分别为 health、drug、space、trend、reporting。

表 3 科技报告主题数为 5 时分布情况

序号	关键词（前五位）	主题	归属学科
Topic1	health、report、system、review、research	健康系统研究与审查报告	医学、工学
Topic2	system、analysis、test、design、report	系统设计测试与分析报告	工学
Topic3	user、meeting、influenza、engineering、experiment	用户流感工程实验会议	工学
Topic4	drug、space、system、finding、research	空间环境药物发现系统研究	医学
Topic5	trend、reporting、progress、pressure、global	压力全球趋势进展报告	理学

④会议文献

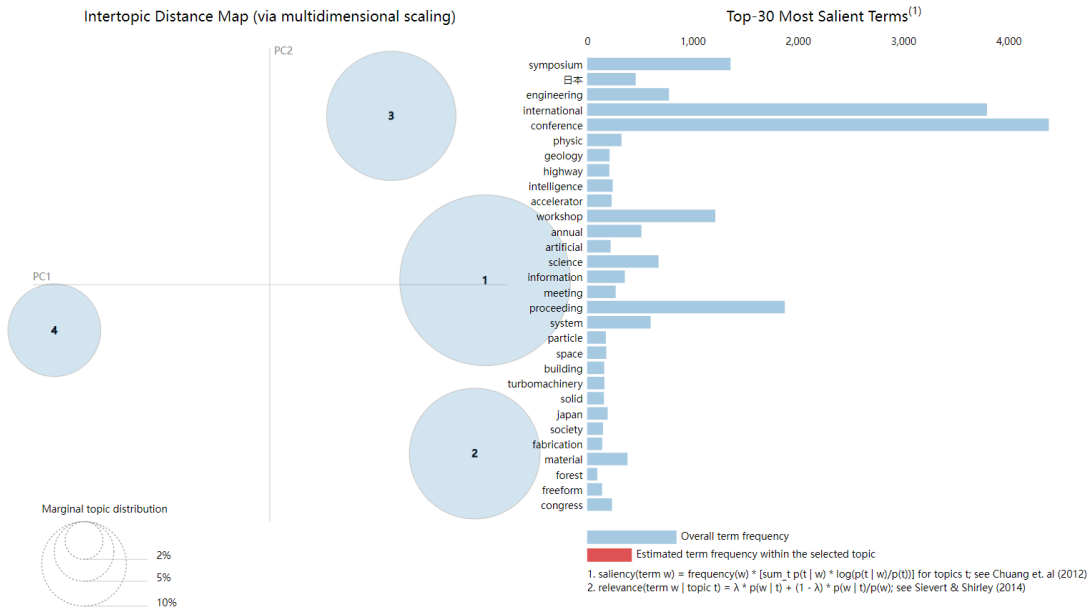


图 9 会议文献主题-主题词分布情况

根据一致性和困惑度结果确定最优主题数为 4，科技报告主题-主题词分布情况如图 9 所示，主题词出现频率较高的分别为 symposium、日本、engineering、international。

表 4 会议文献主题数为 4 时分布情况

序号	关键词（前五位）	主题	归属学科
Topic1	conference、international、proceeding、engineering、science	工程科学进展国际会议	工学
Topic2	symposium、international、conference、proceeding、workshop、physic、intelligence	物理学/人工智能国际会议的专题研讨会与论文集	理学、工学
Topic3	conference、international、workshop、proceeding、accelerator、particle、material	粒子在材料制备中的进展的国际研讨会	理学、工学
Topic4	日本、geology、highway、annual、forest	日本地质学公路森林的年生长	理学、农学

4.NSTL 开放资源保障困境及问题

上世纪九十年代，我国经济、科技处于高速发展阶段，但极低的科技文献资源保障率无法为科研人员提供可靠的文献资源支撑科研活动，NSTL 应运而生。NSTL 成立之初适逢开放获取运动兴起，文献资源获取的新模式为其保障工作增加了一重挑战。近年来，开放获取运动在业界引发了学术出版和商业模式的大变革，对文献资源的生产、传播和消费等环节均产生了重大影响[20]。

4.1 开放资源自身存在安全风险

科技文献资源是国家发展的重要战略资源，为经济发展、科技进步和文化传播提供可靠支撑。目前，我国作为世界上最大的发展中国家对国际开放科技文献资源仍有较大的需求，而在使用开放资源的过程中也承担着较大的安全风险，著作权问题、信息安全等已成为业内讨论和亟待解决的问题[21]。

4.1.1 著作权问题

开放资源是一种版权属性极强的资源。开放科学运动兴起前，几家国际出版商巨头垄断学术资源，导致学术资源价格持续走高，极大增加了科学研究人员的研究成本。为打破这一局面，以开放获取为代表的开放科学运动兴起，广泛推动学术资源的流动和传播。在版权归属方面，传统出版中

作者会将自己的部分权利转让给出版商，而开放获取论文的版权则全部由作者所有，实际上当论文以开放获取形式出版时，读者已获得了阅读、下载、传播、保存等相关权利；开放资源由于自身具有高度的开放共享性，许多机构和个人在浏览和使用过程中忽视了版权问题，导致资源被滥用。NSTL 作为开放资源的整合者和服务提供方，用户如无法正确使用资源，难免会遭遇相关纠纷。

4.1.2 使用安全

在开放获取的模式下，科技文献以数字化信息的形式存储和传播。文献资源订阅机构（如高校、图书馆、科研机构等）通过与出版商进行谈判获取服务，将原本的订阅费用转化为开放获取出版的费用，即开放获取转换协议（transformative agreements）[22]，实现了文献订阅和开放出版整体成本控制，同时促进从订阅模式向开放获取过渡，最终实现完全的开放获取，但这种制度下采购的仅仅是使用权而非所有权。且随着我国科技水平和地位的不断提升，欧美国家对我国的监视与限制也与日俱增，目前世界范围内的出版商巨头主要集中在欧美国家，一旦这些国家对我国采取制裁政策，我国开放资源使用安全将无法得到保障，不仅资源的使用路径会被跟踪，未本地保存的资源也将无法获取[23]。

4.2 开放资源订购成本增长

传统论文出版中作者需要向出版商缴纳版面费，读者获取论文资源也需向出版商缴纳一定费用，开放获取下读者可以免费获取所需要的文献资源。在费用支出方面，传统的出版模式下作者需向出版商缴纳版面费，读者获取资源也需缴纳购买纸质文献资源的费用，而在开放获取的环节中，作者向出版商缴纳版面费外，科研机构代替了读者原本的位置向出版商缴纳资源服务费用，通过其他形式将文献资源开放给读者。

在这个过程中，出版商虽仍可从作者和科研机构获取利润收入，但实际上远不如传统出版模式下订阅费的收入[24]，一旦开放获取全面普及，出版商的利润将大幅度被压缩。为应对这种困境，出版商将原本文献订阅费用的利润空间转嫁给投稿作者，其中有部分作者可得到工作单位、项目基金的支持，而对个人研究作者来说无疑加大了其研究的经济压力。此外，近年来由于开放程度的扩大，科研机构购买资源服务的费用也逐年增加。因此开放获取出版模式的推广为各方都带来了较大经济压力。

4.3 开放资源信息部分缺失

NSTL 在建设开放资源时会为根据一定规范对该条信息进行描述，如学科分类、分类号、语种、出版年份等。通过阅读，可以发现大多数开放资源的信息完整，用户可能够通过过滤器对选择自己所需的数据资源。但仍有一部分出版时间久远、小语种的开放资源缺失描述详细信息，数据缺失也会导致使用过滤条件无法对这部分信息进行筛选，影响用户使用。

4.4 小语种与交叉学科资源建设不足

为使资源被更多的人阅读和使用，各学科领域的优秀期刊大部分以英语为主。根据 3.2 统计结果，NSTL 收录的开放资源约 86%都为英文资源，仅有约 14%的资源为其他语种，NSTL 对小语种开放资源建设重视程度远不如英文资源，事实上这些小语种开放资源也极具研究和学习的价值。

NSTL 拥有九家成员单位，分别覆盖理、工、农、医等学科领域，各单位发挥专业优势，集中建设具有学科特色的开放资源。NSTL 建设开放资源以理、工、农、医四大学科为主，少量收录了经济学、管理学以及交叉学科。根据《研究生教育学科专业目录（2022 年）》[25]，我国共设置了 14 个一级学科门类， NSTL 目前收录的开放资源学科覆盖范围相对集中，资源建设的全面性仍有待提高。

5.NSTL 开放资源保障未来建设路径

5.1 加强用户使用开放资源培训

目前对开放资源版权理论的研究较少，相关版权纠纷案例可参考性不强，因此在日常工作中要严格遵守开放获取协议与相关法律规定，合理合法建设开放资源。同时可在开放资源检索页面增加用户培训模块，发布正确浏览、阅读、使用规范，提醒用户正确应用资源，避免侵权行为。

5.2 加强信息资源信息揭示

完成订阅的科技图书、科技期刊、科技报告、会议文献等目次、摘要等增值信息加工。如若是常年订购的资源可按时间进行资源连接，方便用户的拓展查询，补充完善开放资源上下关联信息，全面揭示开放科技文献馆藏书目信息，保障开放资源的连续性和完整性。

5.3 探索开放资源共建共享模式

NSTL 自成立之初便与国内多所文献情报研究机构进行合作，多年来 NSTL 加强合作同时在全国各地、各行业设立分中心，以支撑 NSTL 文献资源保障体系。在此基础上，NSTL 应继续积极探索开放资源建设共享模式。研究科技文献数字出版开放出版趋势，加强与出版商谈判，加强 NSTL 与各成员单位内部、地区和行业等资源的共享模式。在推进开放资源采集的同时也需强化开放资源的连续性和完整性，拓展文献资源的使用权利。

5.4 加强开放资源保障和质量评估

随着纸质印本不断向数字化形式转变和开放获取运动的开展，文献资源的订购逐渐从传统购买转变为服务许可，这对文献资源的使用和保存造成了较大的阻碍[26]。同时开放获取的资源种类和数量逐年上升，但内容质量良莠不齐。近年来各机构文献资源订阅的经费在不断缩减，为提高经费的利用效率，NSTL 在近期的工作报告指出，对引入资源应加强质量评估，聚焦国家科技创新与行业发展趋势，跟踪国际 STM 出版态势，对高质量、新方向等开放资源、智库资源予以续订和保存，对停更或质量下降的文献资源可予以剔除或替换，定期调整资源续订、停订清单。目前 NSTL 开放资源以英语资源为主，根据用户需求还可适量增加对小语种资源的订购。同时完善资源评估方法，从内容、用户、出版商三个主体出发，多维度对在订资源进行分析评估。

PU-147

蒙医药文献数字化平台建设的现状研究

赵宏亮
内蒙古医科大学图书馆

1 引言

党中央国务院历来重视和支持民族医药事业的发展，不断加大对民族医药的扶持和政策倾斜。近年来，内蒙古自治区党委、政府也高度重视并制定出台了《蒙医药中医药发展战略规划纲要（2016—2030 年）》、《促进中医药（蒙医药）传承创新发展的实施方案》等一系列推进蒙医药事业发展的地方性政策法规，极大地促进了全区蒙医药事业的快速发展。随着信息技术不断更新，传统蒙医药文化传承与发展必将要跟上数字化时代潮流。蒙医药文献的数字化平台建设就是顺应新时代发展，运用现代化信息手段继承和发展蒙医药的必然趋势。目前，蒙医药文献数字化建设工作经过多年的发展，达到了一定水平，取得了一定成就。本文主要通过分析蒙医药文献数字化建设过程中存在的问题，展望蒙医药文献数字化建设的发展趋势，指出蒙文蒙医药文献数字化处理的关键问题。

2 蒙医药文献数字化建设现状

数字资源是国家重要的战略资源，是国家信息技术设施的重要组成部分，对促进我国经济和社会发展发挥了重要的推动作用。目前，数字资源已经成为社会信息资源的主流，联合国教科文组织在《数字遗产保存宪章》中指出，数字资源应理解为一种有利于子孙后代活动的资本或遗产。但数字资源本身较为脆弱，受制于载体老化、技术过时等因素影响，数字资源生命周期比较短暂，如果缺乏有效的长期保存策略，势必造成数字资源的永久消逝，给民族记忆和人类文化传承带来重大损失。近年来，国家科技部联合中科院等多家机构启动建设国家数字科技文献长期保存体系，2015 年 9 月，国家科技图书文献中心（NSTL）邀请国内多个图书馆共同发起签署《数字文献资源长期保存共同声明》，强化对本土化重要文化和技术资源的长期保存。为了系统保存日益增长的数字内容知识资源，迫切需要加强对新型数字内容资源保存研究。近年来，蒙医药文献数字化建设工作有了初步的进展，但同时，基于蒙医学传承发展的目的，对蒙医数字资源如何长期保存提出了更高的技术和管理层面的要求，此项研究可以切实推动在蒙医药领域建立可靠的数字内容资源长期保存机制，确保重要技术和资源可靠保存和永续利用，并在此基础上推动其传承发展。

3 蒙医药文献数字化发展存在的问题

3.1 蒙医药资源共享问题

蒙医药文献数字化建设资源、技术、信息难以共享，部分技术、平台和研发成果存在重复投入

的问题。蒙医药文献的数字化建设不仅仅是单纯文字的数字化，还是我国蒙医药文化跨时代发展，在新的技术和平台下找寻新发展空间的必由之路。但是，我国就蒙医药文献数字化建设没有明确组织机构，没有统一的领导，导致蒙医药文献数字化建设缺乏管理机构，不能形成建设合力。除此之外，蒙医药文献数字化建设没有统一的标准，各类数据间难以共享和移植。不利于蒙医药文献的保存和利用。

3.2 蒙医药文献数字化建设人才缺乏

由于蒙医药应用的地域性，我国使用地域主要在内蒙古、新疆以及周边地区。研究蒙医药的专家和学者比较少，蒙医药文献数字化建设需要兼具蒙医药专业、计算机专业、蒙汉互译专业的专门人员。蒙医药文献的数字化建设不仅仅需要信息化建设人才，更需要蒙医药研究的人才，蒙古语翻译人才，复合人才的缺乏已成为蒙医药文献数字化建设的短板。

3 蒙医药文献数字化平台建设发展方向

蒙医药文献数字化平台建设是蒙医药教学、研究和开发的基础性工作，对促进蒙医药学事业发展有着重要的现实意义，是加快蒙医药学文献建设，提升蒙医药学文献利用率的根本途径。为了更好的传承发展蒙古族医学成果，系统保存蒙医药学数字资源，当前需要对其资源的发展形态与规模、内容结构与数据格式、描述元数据规范、权益管理机制、专门保存技术、保存系统体系、保存活动组织与运行机制等内容进行研究，探索建立可靠的蒙医药学数字内容资源长期保存机制，确保其在我国本土可靠保存、永续利用并加以传承。蒙医药文献数字化平台研究具体包括以下几方面内容：

4.1 数字化平台的技术目标与构架；

建设一个功能强大、数据完善、资源统一的蒙医药文献数字化平台，就必须将相对独立、分散的技术参数进行整理，将存在于数字化平台中各数据库没有统一且科学的建设规范的数据资源进行搜集、整理、归类、加工和著录，并搭建一套能实现这些宝贵资源的发布、展示、检索、管理和提供个性化服务的蒙医药资源服务数字化平台，以便于专家、学者的研究利用。为了实现蒙医药的发展研究、文化遗产保护和传承，要明确数字化平台的总体目标要求，实现总体目标要求对应分解相关技术目标。

利用数据库技术进行分类，形成分类导航。导航栏目对应相关数据库数据，包括：蒙医药学图书全文数据库、蒙医药学期刊全文数据库、蒙古文工具书全文数据库、蒙医药学博硕士论文数据库、蒙医药书目数据库、藏医药全文数据库、民族医药书目数据库、蒙医药视频数据库、蒙医药图片数据库及蒙医药研究等十个数据库。并实现蒙汉双语平台跳转检索。

数字化平台支持蒙医药整合资源跨库检索技术、简单检索技术、字段限定检索技术、高级组合检索技术、二次检索技术。在各类检索过程中计算机语言支持汉、蒙、藏、英等多语种信息处理，支持资源管理和发布等技术。多语种信息处理技术包括多语种网络资源采集技术、多语种搜索引擎应用技术、多语种资源在线阅读展示技术等。

在检索结果中通过计算机语言技术制定显示模型实现：表格模式、列表模式、详细模式。

通过模块化设计以及接口程序的设计，用户也可根据需求将内蒙古医科大学现有蒙医药特色自建数据库完整迁移至本平台中，也可新增自建数据库平台相关功能，因此该平台可创建多个数据库，且每个数据库支持多资源类型。

4.2 数字化平台的软硬件技术支撑；

蒙医药文献数字化平台系统包括计算机硬件和计算机软件两个部分。计算机硬件可以看成是数字化平台的可见部分，是数字化平台的运行基础。计算机软件帮助用户使用硬件，以完成数据资源的输入、处理、输出以及存储等活动。因此，软硬件技术一直贯穿于蒙医药文献数字化平台的系统规划、分析、设计、开发、实施以及运行始终，在此过程中相关硬件是实现数字化平台之本，相关软件是实现数字化平台之源。

平台可使用虚拟机搭建。虚拟化可以在同一台物理机上独立、并行，运行安装不同的操作系统和应用程序的多台虚拟机。使用虚拟化可以根据需要动态移动资源和处理能力。这是由于 VMware 将操作系统和应用程序与它们赖以运行的硬件分离开而得以实现。

数字化平台的 Oracle 数据库支撑 ，数据库的设计是数字化平台搭建、运行的重要组成部分，它是在选定了硬件、操作系统环境下，准确地表达用户的需求，并将其转换为有效存储数据的数据

模型过程。可以说是一个从现实世界向计算机世界转换的过程（见图 1）。

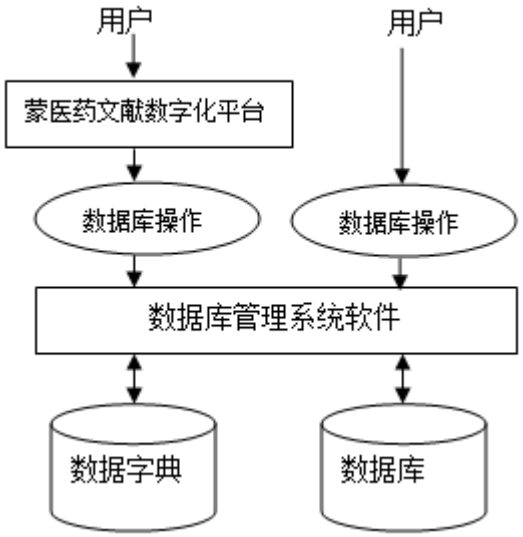


图 1 数据库系统示意图

4.3 数字化平台的后期技术运维与升级

数字化平台部署完成后就要投入日常运行、管理、维护和升级。数字化平台的运行必须遵守管理信息系统、网络和信息安全相关的国家法律法规以及学校相关规定，除此之外，平台的运行管理需要建立相应的技术管理、维护和升级作为保障。

蒙医药文献数字化平台系统运行主要采用传统医学文献信息部和技术服务部双重管理，工作人员主要针对用户管理、基础数据管理以及文档管理等。

系统平台运行中各类人员是各项制度和具体操作的实际执行者，蒙医药文献数字化平台中有大量蒙文资源，因此在具体操作和日常管理中必须有蒙古语背景的工作人员进行操作，系统管理与维护人员必须是具备相关专业背景工作人员进行管理。用户管理是用于管理员管理系统的用户信息，包括管理员设置和会员管理。

基础数据的质量是数字化平台发挥功用的基本保障，在系统平台运行管理中需要高度重视。基础数据的管理包括数据收集、数据录入、数据更新、数据共享和数据备份几个方面。

数据收集　明确规定了不同工作人员收集哪方面的数据、如何收集、判定标准和采集口径等。明确了数据分类整理的原则（见图 2）。

K	L	M	N	O	P	Q	R	S	T
其他责任	其他责任	责任方式	关键字	关键字（蒙	民族	出版社	出版社（蒙	出版地	出版地（蒙
内蒙古文艺		译	蒙古文药		蒙古文古籍	内蒙古文艺		赤峰	
内蒙古文艺		译	蓝琉璃		蒙古文古籍	内蒙古文艺		赤峰	
内蒙古文艺		译	蓝琉璃		蒙古文古籍	内蒙古文艺		赤峰	
内蒙古文艺		译	蓝琉璃		蒙古文古籍	内蒙古文艺		赤峰	
内蒙古文艺		译	罗布桑却		蒙古文古籍	内蒙古文艺		赤峰	
（蒙古文）		译	龙日格丹		蒙古文古籍	内蒙古文艺		赤峰	
			蒙古文药		蒙古文古籍	内蒙古文艺		赤峰	
旺钦扎布		校勘	药剂诸品		蒙古文古籍	内蒙古文艺		赤峰	
呼和禄		编译	饮膳		蒙古文古籍	内蒙古文艺		赤峰	
喜杰		校注	蒙古文医		蒙古文古籍	内蒙古文艺		赤峰	
布和巴特尔		校注	白晶鉴		蒙古文古籍	内蒙古文艺		赤峰	
宝龙		校注	蒙古文医；		蒙古文古籍	内蒙古文艺		赤峰	
白翠兰		校注	蒙古文医；		蒙古文古籍	内蒙古文艺		赤峰	
哲里木盟		译	蒙古文医；		蒙古文古籍	内蒙古文艺		赤峰	
钢卓力克；		校注	蒙古文医；		蒙古文古籍	内蒙古文艺		赤峰	
钢卓力克；		校注	蒙古文医；		蒙古文古籍	内蒙古文艺		赤峰	

图 2 数据收集（部分）

在蒙医药文献数字化平台运行过程中需要不断的进行技术维护，技术维护具体工作包括对数字化平台进行各项参数修改、状况调试、设备维修、服务启动等活动的统称。蒙医药文献数字化平台技术支持最终目的就是要实现根据读者需求进行资源的正常使用、数据的及时更新，对组织和系统中已经过时的、失效的或者冗余无用的数据和信息制定更新和处理规定，同时进行技术处理，保证信息的有效性和一致性。最终达到读者流畅、友好的使用该数字化平台。

PU-148

网络时代如何有效提升图书馆信息化建设水平

罗燕玲
空军军医大学

0 引言

高校图书馆是学生获取信息的重要场所，随着信息化技术以及数字化程度的提升，高校图书馆由传统的纸质图书馆向数字化、电子化方向发展，并且由于高校图书馆电子信息技术水平的提高，信息交流的方便程度也大幅上升，学生能够更好的利用电子信息资源，构建自我认知体系，但在高校图书馆信息资源利用方面，发展中仍然存在一定的问题，高校图书馆电子信息资源的安全以及合理利用也仍存在一定的局限性，如何突破高校图书馆电子信息利用掣肘，摒弃高校图书馆电子信息资源利用存在的缺陷，是当前高校图书馆信息技术发展中需要研究的问题。

1 我国图书馆发展现状

1.1 2017-2021 年我国图书馆机构情况

从数据调查显示中可以发现，从 2017 年开始，我国图书馆的机构数量逐年增加，具体可见图 1 所示。

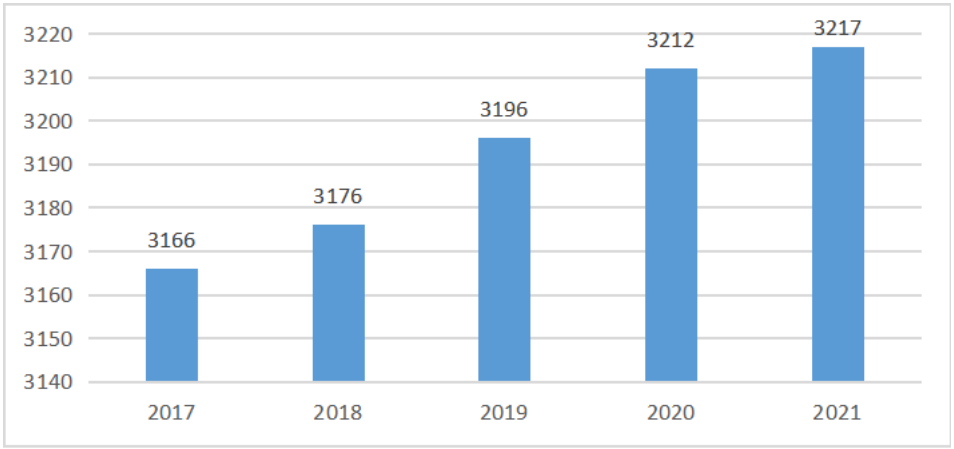


图 1：2017-2021 年我国图书馆机构数量情况

1.2 2016-2020 年我国图书馆资源使用情况

在受到大环境影响的情况下，我国图书馆资源的利用情况仍然在不断上升，从 2016 年到 2020 年期间，图书馆的资源利用率也在不断增加，平均利用次数超过千万人次，具体见表 1 所示。

表 1：2016-2020 年我国图书馆资源使用情况

年份	发放有效借书证（万个）	图书馆流通人次（万人次）	书刊文献外借人次（万人次）	书刊文献外借册次（万册次）
2016	5593	66037	24892	54725
2017	6736	74450	25503	55091
2018	7263	82032	25814	58010
2019	8627	90135	26609	61373
2020	10251	54146	17467	42087

2 图书馆信息化管理要求

2.1 精度要求

准确性要求意味着所有文件必须在短时间内获取，工作文件必须能够专业地记录某些参数和信息。例如，对于某些书籍文件，它们是公开的并且必须准确。记录图书当前存放位置，包括库房信息、书架信息、图书种类信息等要求，所有信息必须经过综合评估后录入专有资源系统。准确性要求是图书馆资源管理的关键指标，即使硬件和软件系统不能满足要求，也可以牺牲存储效率和管理规范来提高准确性。

2.2 快速需求

另一方面，及时性要求是指所有文件都可以在短时间内被记录、分析和存储，特别是需要履行这些责任和义务[1]。例如，一本书进入图书馆后，图书馆管理员必须立即探索和合并诸如图书目录号、适当的存储位置、图书内容和图书本身类型等参数。创建的文件系统。形成的工作系统具有更高的可靠性。此外，在快速需求阶段，需要明确现有工作模式存在的不足，对硬件和软件系统进行升级，使现有工作系统更加稳定，以保证实际执行质量。信息资源管理任务满足当前要求的要求[2]。

2.3 协同要求

图书馆资源的协同要求意味着当新的文件被添加到图书馆文件管理系统时，这些信息必须存储在一个准确的数据库中，并且可以在实际情况下向公众提供。图书馆资源管理系统根据实际使用和运行形式，可分为图书自身信息资源、人力资源管理信息资源、图书馆运营财务系统信息资源等多种形式，从而形成管理系统可以更好地运行。另外，一些部署了云计算的管理系统，需要在存储数据后，将本地化的文件与云存档关联起来，以便整个文件管理系统的所有存档都能在第一时间更新[3]。

2.4 改进要求

改进要求是指无论是文件管理系统还是任务管理系统，都必须有相应的改进指标，包括对现有管理机制的全面创新、管理方式的进一步落实等，可用于其他系统用户，根据自己的需要，在全管理系统中输入关键字，就可以得到你要查找的文件类型。此外，已完成的工作管理系统应具有更高的运行效能，在运行开发阶段，根据不同的输入指令和工作系统，从对应工作的数据库中调取信息，防止数据库管理系统无法工作[4]。

3 图书馆信息化建设水平提升路径

3.1 加强对高校图书馆管理人员的信息技术培训与管理

高校图书馆进行电子化服务升级和转型，在图书馆进行数字化转型与智能化转型的过程中，传统的图书馆管理人员已经不能适应新型图书馆的服务管理模式，需要对人员进行服务内容的培训和提升。对于图书馆管理工作人员的管理，可以从以下几个方面进行，一是，招聘更多能够适应于智能化图书馆服务的工作人员，在上岗前并对员工进行必要的操作培训，利用讲座或者进修的方式，引导图书馆工作人员实现自身技能水平的提升和转化，配合高校图书馆电子化升级转型的需求，并能够对智能化图书馆的相关设备进行自如的运用，当图书馆的设备出现故障时，能够独立的进行解决；二是加强自身技能的提升，除了对新招聘人员进行培训之外，还需要加强对已有的图书管理工作人员进行培训，所有的工作人员都需要提升自身的工作技能，为了满足图书馆的多元化服务对象以及服务类型，图书馆的工作人员还应该掌握相应的外语水平，并且掌握一定的基础知识，能够在用户遇到问题求助时，及时迅速的为用户提供解决方案[5-7]；三是图书管理人员以及领导层需要加强自身与时进步的意识，加强理解对于高校图书馆电子化发展方向的要求，并紧跟随图书馆服务转化升级的步伐，强化自身的管理能力，提升自身在图书馆行业中的竞争力，同时发挥自身的价值和优势。

3.2 强化高校电子图书馆的服务形式

高校电子图书馆是基于世界先进的技术，新兴技术所建设的图书馆服务空间，包括大数据以及智能化等技术在内的服务模式，高校电子图书馆服务的主要特征在很大程度上实现了智能化的服务，解放了人们的双手，减少了对人力资源的投入，同时在满足不同类型用户阅读需求的基础上，还能根据用户的不同偏好为用户提供更具针对性的服务方式和搜索方式，同时高校电子图书馆的服务类型，不仅可以通过人工服务的引导，还能为用户提供人机交互、智能语音交互的手段，根据语音提示选择适合自己的阅读方式[8，9]。

3.3 加强电子图书馆信息资源交互

交互功能的设计是电子图书馆中最需要的功能之一，利用交互功能可提升用户的阅读兴趣，图书馆就是知识传递和传播、保存信息交互的重要场所之一，构建交互功能，用户可以在交互平台上，发表自己对于阅读的理解，或者对于图书馆服务相关内容的满意与否，电子信息资源保管是否安全，图书馆中的交互功能能够更好的促进人与人之间的交流，并推动知识的传播，通过交互动能的密码强化，同时还能够对电子图书馆资源的建设起到促进作用[10]，从而进一步加强高校图书馆电

子信息资源的利用程度。

3.4 加强图书馆基于物联网的信息化建设

随着信息化的不断发展，图书馆的智慧转型升级势在必行。通过对图书馆智能化转型升级，借助物联网平台，可以推动图书馆数字化转型，并通过对图书馆已有的设备，添加智能化的服务平台与内容，以此来实现图书馆智慧化升级服务。如图 2 所示，展示的是基于互联网的智慧图书馆服务平台的框架。

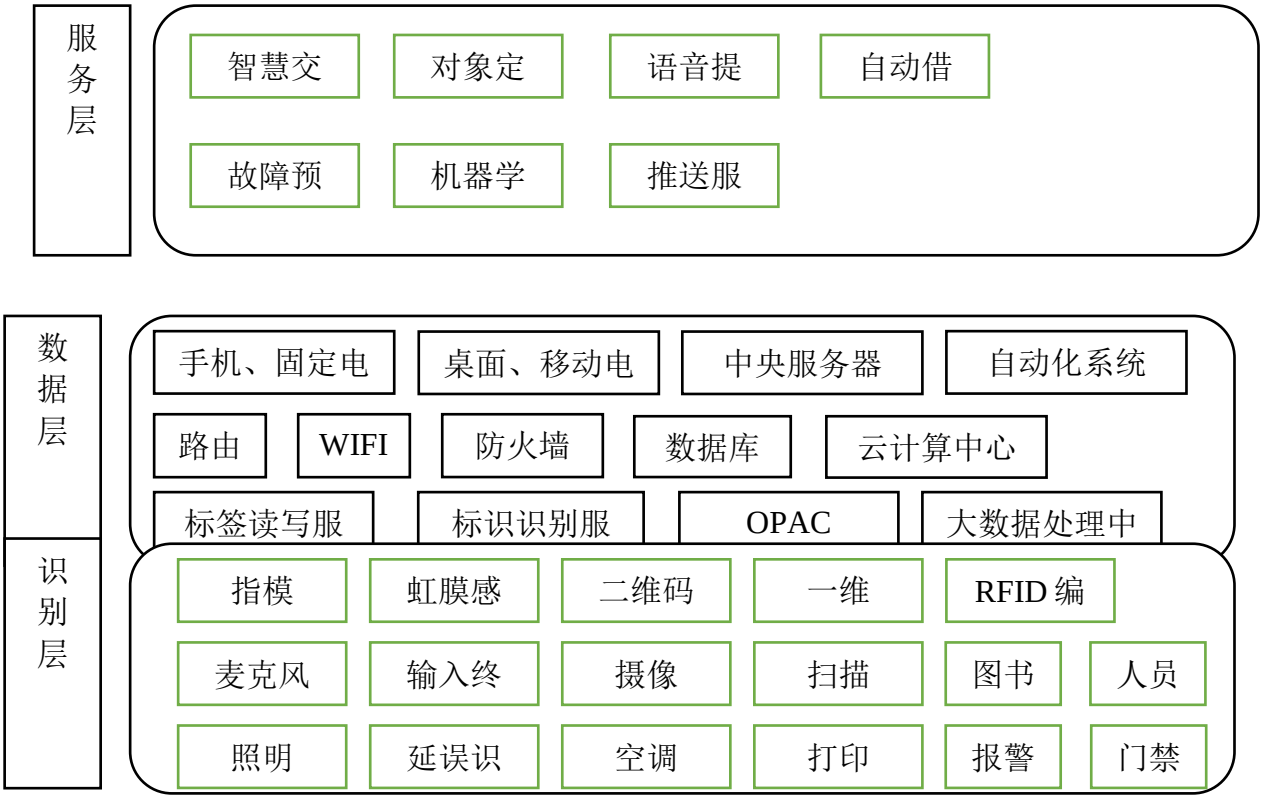


图 2：基于物联网表示体系的信息化图书馆建设模型

基于物联网构建的信息化图书馆建设模型，是以三个层次构成的，也就是由图书馆的支撑层面是，信息层、数据层以及服务层，对应到图书馆服务模型上，可以用数据层、识别层和服务层进行对应，下面分别介绍每个层级的主要作用。

（1）识别层，也就是信息化图书馆的信息收集阶段，在该阶段中，主要收集包括用户信息、身份等内容，识别层主要包含的功能包括了，指模、虹膜感应、二维码、一维、RFID 编码、麦克风、输入终端、摄像、扫描、图书、人员、照明、延误识别、空调、打印、报警、门禁等多种功能在内的信息数据收集层面，能够更好的为图书馆用户提供智能化的服务。

（2）数据层，该阶段主要是通过物联网收集到的数据，利用大数据的分析方式，进行数据分析，手机、固定电话、桌面、移动电脑、中央服务器、自动化系统、路由、WIFI、防火墙、数据库、云计算中心、标签读写服务、标识识别服务、OPAC、大数据处理中心等在内的多种数据服务功能，可以通过这些综合的数据处理功能，分析用户的偏好，深入的对用户的阅读需求进行挖掘，通过大数据与云计算中心，制定导向性的服务方向。在这个阶段中，用户可以利用搜索引擎功能进行想要阅读内容的搜索以及书籍的搜索，通过大数据与云计算的自动化系统构建，可以为图书馆管理人员减轻负担，能更好的服务于用户的阅读与学习。

（3）服务层，该阶段中主要包括智慧交流、对象定位、语音提示、自动借还、故障预测、机器学习、推送服务等内容，也就是说，最终为用户提供的服务类型包含以上几种，通过数据层的分析，可以实现用户偏好的分析，以及不同群体的阅读需求分析，定向的推送相关内容，吸引用户的

阅读；同时，信息化图书馆服务平台的智慧交流模块，可以实现用户的智慧交流，除了能与图书馆里员进行交流与沟通之外，还可以与图书产品进行信息交流与沟通；通过语音交互的方式，为用于提供更好的阅读服务，例如有使用障碍的用户等，可以利用该语音功能进行图书的阅读；信息化图书馆服务体系中，智能借还服务，可以减少用户的操作流程和等待时间，节约时间成本，同时也降低了图书管理员的工作量；信息化图书馆服务内容中的故障预测，可以对故障内容进行预测，更快地的定位到故障点，从而迅速进行响应，减少故障的发生[11]。

4 总结

综上所述，在我国图书馆信息化建设过程中，信息化图书馆建设的要求需要满足快速响应、协调、精度和改进等要求，通过加强对图书馆管理人员信息技术培训，加强图书馆信息化服务形式，加强信息资源交互，以及物联网状态下的信息化建设等，提升图书馆信息化建设水平。

PU-149

医学信息素养教育理论与实践

刘文峰
河南大学淮河医院

引言：

在信息爆炸的时代，医学信息素养已成为医学教育的重要组成部分。医学信息素养是指医学生或医师在临床实践、科研和学术交流中获取、理解、评估和使用信息的技能。它包括信息意识、信息技能和信息安全三个方面，其中信息技能是核心，信息安全是保障。医务人员信息素养不仅关乎医学方面的学术成就，也关乎工作中临床决策的准确性和效率，更与医疗从业人员合法权益、医患关系息息相关。因此，开展医学信息素养教育具有重要的现实意义。

医学信息素养教育的理论和实践中，对现状的相关研究：

首先通过查阅相关资料了解到：国内外学者已对医学信息素养教育进行了广泛的研究。研究表明，医学生和在院医师的信息学水平普遍偏低，存在信息获取、理解和评估能力不足等问题。相关研究还发现，信息素养水平与临床决策能力、医患沟通水平呈正相关，提高信息素养水平有助于提高临床决策的准确性和效率，有利于缓解医患矛盾，切实保护医务人员合法权益。此外，学者们还提出了多种提高医学信息素养的策略，如开设信息素养课程、开展信息素养培训、建立信息共享平台等。

本人也对河南大学淮河医院近三年新入职医务人员进行了一些简单的调研。调研采用问卷调查法，选取每年新入职医生和通过多方途径来院进行规范化培训的医师作为研究对象。问卷主要包括基础信息学水平调查、信息学在医疗工作中应用能力调查和信息安全意识三个部分。本研究共发放问卷 562 份，回收有效问卷 534 份，有效回收率为 95%。采用 SPSS 对数据进行统计分析，其中基础信息水平测试平均得分 72.73，信息学在医疗工作中应用能力得分 81.25，信息安全得分 69.7。

从以上数据不难看出针对医务工作者的信息素养教育还存在一些问题和不足。首先，信息素养教育尚未得到足够的重视，很多医学院校没有开设相关的课程和培训。其次，现有的医学信息素养教育内容和方法比较单一，缺乏针对性和实效性，不能满足医务工作者的需求。通过现状分析对发现的问题我们不能回避，需要在以后的教育工作中不断努力去改善和解决，但是我们又不能盲目的大刀阔斧做事情，首先需要梳理教育理论和实践中值得关注的重点和关键。

医学信息素养教育的理论和实践中，需要关注的重点:

1. 教育目标的明确与实践方法的选择: 教育目标应该与医务工作者的需求和实际工作情况相匹配。实践方法可以包括课堂教学、案例分析、模拟演练等，旨在激发医生学习兴趣和提高信息素养。

2.信息资源的充分利用与评估: 医务人员需要了解并熟练运用各种信息资源，包括医学电子文献数据库、临床指南、药物数据库等。同时，评估信息的质量和可靠性也是十分重要的，教育内容应该包括信息评估的方法和技巧。

3.信息技术的培养与应用: 随着医疗信息化的发展，医务工作者需要掌握更多的信息技术，如电子病历系统、临床决策支持系统等。教育应该注重培养医生的信息技术接受能力，使其对新型信息技术保持好奇心，更容易学会新技术新方法，并能够熟练应用信息技术解决临床问题。

4.信息行为与伦理的培养: 医生在信息获取和应用过程中，需要具备正确的信息行为和职业伦理，树立信息安全意识。教育中应该强调医生应当遵守信息保密、知识产权保护等相关法律法规和伦理准则，以保证信息的安全和医生的职业道德。

5.医务工作者信息素养提升需要与临床实际工作相结合，建立起完善的评估体系，以监测医务工作者信息素养的提升情况。同时，教育的效果也需要与患者的满意度、医疗质量和安全性等指标相关联，形成相互促进的良性循环。

总之，医学信息素养教育理论与实践是一项长期而复杂的工作，需要各方合作，不断探索和创新。只有通过有效的教育和实践，才能提高医务工作者的信息素养，提升医疗质量，为患者带来更好的医疗体验和健康效果。

有了工作的重心，我们需要采取一些措施和方法。加强医学信息素养教育的宣传和推广，提高医学院校和进修班对信息素养的重视程度。开设多样化的医学信息素养课程和培训，包括文献检索、医学文献综述、医学统计学等课程，以满足医务工作者的需求。同时，可以采用案例教学、实践教学等多种教学方法，以提高教育的实效性。建立医学信息素养评价体系，对医学生和医生的信息素养进行评估和考核，以便发现问题并及时进行改进。

医学信息素养教育理论与实践工作从以下几个方面展开:

1.学科知识与信息素养的融合: 医生需要掌握丰富的医学知识，但单纯的知识储备并不能完全满足临床实践的需求。因此，教育应该注重将学科知识与信息素养相融合，培养医生的综合素养和批判性思维能力，使其能够在临床实践中充分利用和应用所学的知识。

2.科学研究与信息素养的结合: 医务工作者需要从科学研究中获取最新的临床实践指南和最新的医学进展信息。因此，教育应该培养医生的科学素养，教导他们如何阅读并评估科学文献，以及如何运用科学研究成果提升自己的临床实践水平。

3.团队合作与信息共享: 医生在团队中与其他医务人员合作，共同提供优质医疗服务。在这个过程中，信息共享和交流至关重要。因此，教育应该培养医生的合作与交流能力，教导他们如何有效地与其他医务人员共享重要信息，以提高整个团队的工作效率和质量。

4.持续学习与自我提升: 医务工作者的信息素养教育应该是一个持续学习和自我提升的过程。医学和医疗技术都在不断发展，同样信息学也在突飞猛进的向前迈进，医务工作者需要不断更新知识，学习新的信息技术和信息管理方法，能够熟练使用各种信息载体和信息传播介质。因此，教育应该鼓励医生保持学习的热情，提供多样的教育途径和教育方式，提供相关的继续教育和培训机会，以帮助医生与时俱进。

综上所述，医学信息素养教育理论与实践是一个多层次、多方位的工作，需要教育者、医学院校、医疗机构以及医务工作者本身的共同努力。通过有效的教育和实践，可以提高医务工作者的信息素养水平，为患者提供更加安全、高质量的医疗服务，推动医疗领域的发展和进步。

医学信息素养教育的实践中，对未来教育方向进行深入探讨:

1.多渠道的信息素养教育: 除了传统的教室教学和面对面培训，可以利用信息技术手段，如在线学习平台、网络资源等，提供多样化的学习材料和教育内容，满足不同医务人员的学习需求。

2.推动信息素养教育的持续发展: 信息技术和医学知识的不断更新和发展，要求医务工作者不断学习和更新自己的信息素养。因此，应该建立起持续的信息素养教育体系，包括定期培训、继续教育等机制，使医生能够不断跟进最新的医学进展和信息技术应用。

3.建立团队合作与协同学习机制: 信息素养教育不仅仅是个体学习的过程，也需要通过团队合作和协同学习的方式进行。医生可以在团队中共同学习、分享经验和解决问题，提高信息素养的实际应用能力。

4.加强评估与反馈机制: 教育的效果需要通过评估和反馈机制来检验和改进。可以通过问卷调查、考试成绩、临床绩效等评估手段，及时了解医务工作者的信息素养水平，并提供针对性的反馈和改进意见。

总的来说，医学信息素养教育的理论和实践是一个综合性的工作，需要多方合作、多维度的思考和创新。通过有效的教育与实践，能够提高医务工作者的信息素养水平，提高信息安全意识，进而提升医疗质量和效率。

医学信息素养教育实践是将信息素养教育理论付诸实际行动，通过培训、课程设计、实践教学等方式，提升医务工作者的信息学水平。实践中可以结合实际临床案例，引导医生了解和运用各种信息资源和信息工具，培养信息素养相关的技能和能力。同时，也需要促进医务工作者的信息安全意识和信息伦理的培养，使其在信息化的医疗环境中具备正确的信息行为和职业道德。

医学信息素养教育理论与实践是一个不断探索和完善的过程，建立科学合理的培养体系，提供有效的教育资源和培训机会，促进医务工作者信息素养的全面提升，医学信息素养的教育任重而道远。

PU-150

Hotspots and research trends relating to infectious-disease-specific health literacy in the past decade: a bibliometric study

严波[1] 陈晃荣[1] 谢瑶[1] 刘靖榆[1] 赖雨婷[1] 黄丹蓝[1] 周旭毓[2]
1.广州新华学院
2.中山大学

目的 Objective: To explore the current hotspots and predict future research trends, providing reference for research in related fields.
方法 Methods: Relevant references (from May 1, 2014, to April 30, 2023) were retrieved from China National Knowledge Infrastructure (CNKI), and analyzed by CiteSpace 6.2.R2 software.
结果 Results: A total of 713 references were included. Cooperation between institutions was not common. Topics such as "students" etc. are current hot topics. Several topics, including "migrations" etc., may be hotspots in the future.
结论 Conclusion: The current research hotspots covered multiple populations (e.g., health personnel, middle-aged and elderly people, etc.). There is a relative lack of intervention studies and longitudinal studies. Topics such as "immigrant population" may become hotspots in future research.

PU-151

智慧医疗背景下医护人员信息素养探究

古林林 王琳凯
驻马店市中心医院

1 引言

随着互联网技术与医疗行业的结合，智慧医疗发展突飞猛进，云计算、人工智能等新兴技术为医院患者提供了方便快捷服务的同时也为医护人员带来了挑战，如何有效的利用新兴技术，涉及到医护人员自身的信息素养[1]。

2 智慧医疗

智慧医疗（Smart Healthcare）是指在信息技术的支持下，将医疗服务与先进的技术手段相结合，实现从医疗资源，医疗过程到健康管理各个环节的智能化和数字化。它旨在提高医疗服务的效率、质量和安全性，为患者提供个性化的医疗健康管理服务。

智慧医疗的核心是运用大数据、云计算、物联网、人工智能等新兴技术，实现医疗信息的集中管理、共享和分析，实时监测患者的健康状况，提供个性化的医疗方案和健康管理建议。具体来说，智慧医疗包括以下几个方面：

1. 互联网医疗（Internet Healthcare）：利用互联网技术实现医疗信息的传输、共享和交流，使医患之间的沟通更加便捷和高效。

2. 远程医疗（Telemedicine）：利用视频会议、远程监测等技术手段，实现医生与患者之间的远程诊疗和健康管理，节约时间和资源。

3. 智能医疗设备（Smart Medical Devices）：通过智能传感器和无线通信技术，实现对患者健康数据的实时监测和追踪，提供个性化的健康管理服务。

4. 电子病历（Electronic Health Records）：将患者的医疗档案数字化，实现电子病历的共享和管理，提高医疗信息的安全性和准确性。

5. 数据分析和预测性医疗（Predictive Analytics）：通过对大数据的分析，发现医疗模式和规律，为医生提供辅助决策，预测疾病风险和治疗效果。

智慧医疗的发展可以提高医疗资源的利用效率，改善医疗服务的质量和安全性，降低医疗成本，提升患者的医疗体验，满足更高效、更低失误的发展需求，符合医疗改革的发展需要，需要不断开拓创新[2]。

3 信息素养

3.1 概念

信息素养（Information Literacy）是指一个人在面对信息时具备获取、评估和应用信息的能力。它涵盖了信息检索、信息评估、信息使用和信息创造等方面。信息素养使个人能够有效获取所需信息，正确评估其可信度和质量，有效应用信息解决问题，同时也注重信息使用的伦理和法律问题。

信息素养起源于 20 世纪 70 年代的信息革命和信息社会的发展。随着信息技术的迅速发展和信息的大规模产生和传播，人们在面对庞大的信息量时往往感到困惑和无所适从。信息素养的概念应运而生，旨在帮助人们应对信息爆炸时代的挑战。

1989 年，美国国家教育研究院（National Council of Educational Research and Training，简称 ACRL）首次提出了信息素养的定义，将其定义为“独立地获取、评估和使用信息的能力，以及在产生信息时具备创新和批判性思维的能力。”随后，信息素养的概念逐渐被学术界和教育界广泛关注和研究，并且被纳入到教育和培训的内容中。国际上，很多国家和地区都将信息素养作为一项关键的学习目标，将其纳入到教育体系和教学计划中。

3.2 医疗信息素养

医疗信息素养是指医护人员在医疗和健康领域中获取、评估和有效应用医疗信息的能力，包括对医疗信息的理解、运用和管理，以及对医学知识、技能和技术的适当应用。

具体来说，医疗信息素养包括以下方面：

1. 信息获取能力：医护人员应具备获取医学信息和最新研究成果的能力，掌握有效的信息搜索和筛选方法，运用专业的数据库、在线资源和学术期刊等获取准确可靠的医学信息。

2. 信息评估能力：医护人员应具备分析和评估医疗信息的能力。通过判断信息的可靠性、权威性和适用性，了解不同研究方法，对医学研究结果进行批判性思考和判断。

3. 信息应用能力：医护人员应具备将医疗信息应用于临床实践和决策的能力。通过将医学知识和技术与患者的个体情况相结合，制定个性化的医疗方案和治疗计划，并及时更新和调整。

4. 信息管理能力：医护人员应具备有效管理医疗信息和电子病历的能力。遵守信息安全和隐私保护的规定，妥善管理和传输患者的医疗数据，确保数据的完整性、安全性和保密性。

5. 医患沟通能力：医护人员应具备与患者和家属进行有效医患沟通的能力，能够以简明易懂的方式解释医学知识和医疗信息，提供清晰的建议和指导，促进患者的理解和配合。

医疗信息素养的提高可以帮助医护人员更好地应对医学知识的多样性和快速变化，提升医疗服务的质量和效果，促进患者的安全和满意度。因此，在医学教育和培训中应重视医疗信息素养的培养，为医护人员提供相关的培训和支持，使他们能够不断适应医疗信息化发展的需求。

4 医护人员信息素养现状

医护人员信息素养在不同医疗机构和个体之间存在一定的差异。以下是一般医护人员信息素养的现状：

1.信息管理不规范：医护人员在信息管理和电子病历的使用方面存在差异。一些医护人员能够熟练操作电子病历系统，遵循信息安全和隐私保护的规定；但也有一些医护人员在信息管理方面存在不规范的问题，导致信息流通不畅或者存在信息泄露的风险。

2.缺乏信息应用能力：医护人员在将医疗信息应用于实践和决策方面的能力也存在差异。一些医护人员能够灵活运用医学知识和技术，结合患者的情况制定个性化的治疗方案；但有些医护人员缺乏更新和应用最新医学知识的动力。

3.缺乏医患沟通能力：医护人员在医患沟通方面的信息传递能力也各有不同。一些医护人员能够以简明易懂的方式解释医学知识和医疗信息，得到患者的理解和配合；但也有一些医护人员可能在沟通时过多使用技术术语或专业术语，难以被患者理解。

总体来说，医护信息素养整体处于不断提高的过程中，但仍然有一些待改进的地方。医疗机构和教育机构应该加强对医护人员信息素养的培养和教育，提供相关的培训和支持，鼓励医护人员积

极参与信息化医疗的发展，提升医护人员的信息应用能力和医患沟通能力。

5 医护人员信息素养培养

有效培养医护人员的信息素养，可以采取以下措施：

1. 提供专业培训：为医护人员提供系统的、针对性的信息技术培训，包括信息检索技巧、文献评估方法、疾病诊疗指南的应用等。培训内容应贴近实际工作需求，通过实践操作和案例分析等方式增强学习效果。

2. 制定规范和指南：制定医疗机构内部的信息管理规范和操作指南，并加强对其的宣传和培训。规范和指南可以明确信息使用的流程、质量要求和原则，帮助医护人员正确获取、评估和应用医学信息。

3. 创建资源平台：建立医学信息资源平台，提供全面、可靠的医学文献、临床指南、疾病诊疗方案等信息资源。平台应具备便捷的检索功能、信息更新和质量审核机制，为医护人员提供可靠的支持和参考。

4. 鼓励知识分享和交流：创建知识分享和交流的机制和平台，鼓励医护人员分享他们在信息获取和应用方面的经验和技巧。定期组织研讨会和经验交流会议，提供交流和学习的机会。

5. 强调持续学习：信息技术和医学知识都在不断更新和发展，医护人员需要不断学习和更新自己的知识和技能。医疗机构应提供学习资源和机会，鼓励医护人员积极参与学术研究、继续教育和专业培训，不断提升信息素养水平。

6 小结

智慧医疗建设是社会医疗卫生事业发展的新阶段，是医疗建设的一项大工程，需要政府、医疗机构、医护人员的多方参与[3]。在先进设施落地后医护人员只有通过加强信息技术等相关知识的学习，增强个人信息意识，提高信息工具的熟练度，才能更好的使用设施，医疗机构才能提供更高效、便捷的医疗服务[4]。

PU-152

医学院校图书馆多元信息素养教育服务体系建设研究

张帆

内蒙古医科大学图书馆（含档案馆）

1 国内外相关研究现状分析

在中国知网总库中，检索途径选择高级检索，用主题字段检索。检索条件为：(主题%='多元' or 题名%='多元') AND (主题%='信息素养' or 题名%='信息素养') 。时间范围：2003 年 1 月 2023 年 9 月，共找到 800 条检索结果。（数据来源：中国知网总库）

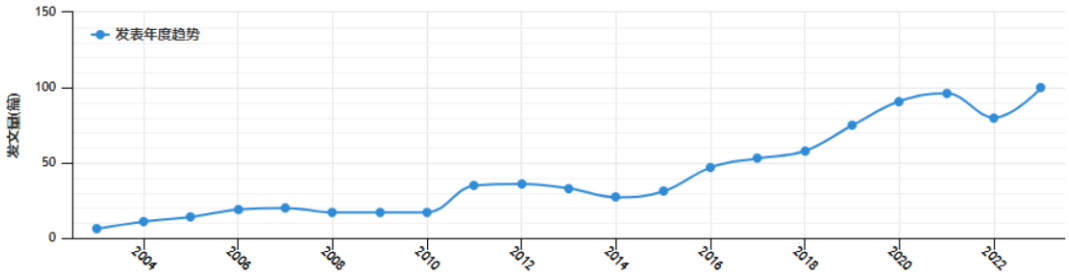


图 1 发表年度趋势

由上图可以看出，近 20 年关于“多元信息素养”研究数量日趋增多，由 2003 年的 6 篇增加到 2021 年 96 篇，2023 年论文数量仍在增加。

2“图书馆多元信息素养教育服务体系”建设内容

医学院图书馆开展“多元信息素养教育服务体系建设”，既符合图书馆发展的实际，与学校的

人才培养工作同向同行。图书馆多元信息素养教育服务体系建设主要包括以下四个方面：信息素养一流课程建设；“教学空间+平台”建设；多元信息素养教育服务体系建设；师资队伍建设。

2.1 信息素养一流课程建设。

加强信息素养一流课程建设，提升教育教学质量，将学生培养成为德智体美劳全面发展的社会主义事业的建设者和接班人。信息素养一流课程也要从教学理念、课程教学团队、课程目标、教学设计、课程内容、教学组织、课程管理与评价等 7 个方面开展建设。使高校使每一门课程发挥好最好的育人作用，将专业课程与弘扬真善美结合。

2.2“教学空间+平台”建设。

文献检索课程是信息素养教育的重要组成部分，是培养学生的信息意识、掌握获取知识和情报的一门科学方法课，是提高学生自学能力和独立研究问题能力的工具课，对学生的实习能力要求较高，因此加强教学空间建设和多元信息素养教育服务平台建设。

2.3 多元信息素养教育服务体系建设。

多元信息素养教育体系包括 4 个维度，11 个层次。

第一阶段开展大一的新生入馆教育（图书馆主页、图书馆微信公众号、图书馆宣讲团），第二阶段开展常规培训（数字资源使用培训、学知讲坛、学术资源推荐），第三阶段开展文献检索课教学（线下教学、学习通网课、线下多媒体机房实习），第四阶段鼓励学生参加各类信息素养比赛（“利用图书馆”知识竞赛、各类数据库检索比赛）。

其中文献检索课教学是多元信息素养教育体系的核心。创新文献检索课教学模式，从课程教学顶层设计，包括课程目标凝练、教学团队建设；课程教学实践设计，包括教学大纲编写、教学案例撰写以及课堂教学设计；课程案例成果库建设等多方面开展研究和实践，最终形成新时代医学文献检索课创新模式。实现大学阶段全过程育人。形成“三全育人”视域下的图书馆多元信息素养教育平台。依托“学习通”及“读者之星”平台，全方位、立体化开展多元信息素养教育。

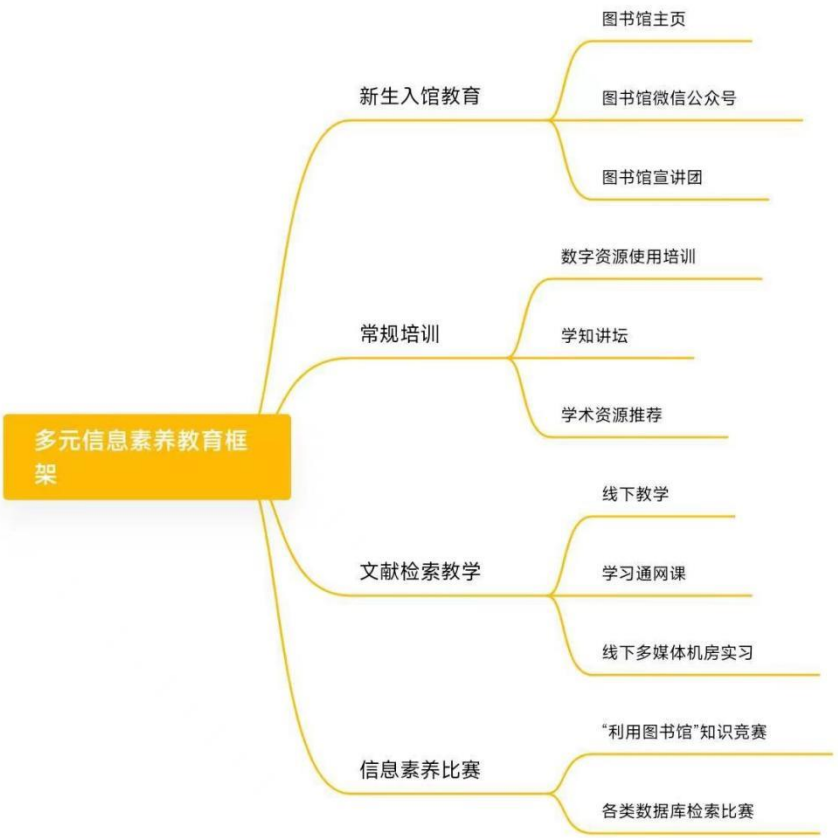


图 1 多元信息素养教育框架

2.4 师资队伍建设。

师资队伍建设一直是制约图书馆教学事业发展的瓶颈，注重师资队伍的引育建设，既要引进人才，更要培育已有人才。医学文献检索课创新教学人才队伍的建设，对于下一步图书馆教学事业的发展而言是至关重要的。只有医学文献检索课创新教学人才队伍的建设好了，才能培养具有历史使命以及良好科学素养的医学人才。

3“图书馆多元信息素养教育服务体系”实践成果

为检验图书馆多元信息素养教育服务体系的育人效果，培养学生信息素养实践能力，在 2017 级、2018 级、2019 级、2020 级的临床医学专业、中医学专业实施“以赛促教 以赛促学”模式。在《文献检索》课程教学过程中，通过文献检索与数据库检索的实践帮助医学生解决实际医学科研问题。并通过参与高校信息素养大赛，教师发现《文献检索》课程在教与学方面存在的问题，帮助教师找到教与学有效结合路径，实现“以赛促教 以赛促学”，全面提升学生信息素养实践能力。2019 年“incopat 专利检索大赛”5 名同学获奖。2020 年首届“桃李杯”内蒙古高校信息素养大赛 1 名同学获得一等奖，6 名同学获优秀奖。2021 年第二届“桃李杯”内蒙古高校信息素养大赛 4 名同学获优秀奖。2023 年“万方数据杯”高校知识答题竞赛（内蒙古赛区）1 名学生获得二等奖、1 名学生获得三等奖。

PU-153

医院信息化人才队伍建设研究及对策

王文峰 陈贝贝 李砚
三门峡市中心医院

一、我国人才队伍建设现状

根据“我国三级医院卫生信息人力资源建设情况研究”指出，三级医院的信息技术部门员工中80%以上是本科，且信息化高层次人才队伍大部分集中在东部地区的三级医院。在全国三级医院信息技术部门员工的职称构成中，全国三级医院信息技术部门员工中 62.66%是初级职称与无职称人员。专业背景构成中，73.41%的医院信息技术人员是计算机专业。研究表明信息化人才配置薄弱，数量与质量不能满足医疗信息化建设的发展需要。

二、医疗信息化人才队伍现状调查

本次研究人员采用的调查问卷方法进行研究，调研的主要内容是医疗信息化人才队伍建设现状、人才培养的方向及制约因素、人才培养方式等内容。项目组发出调查问卷共计 200 份，收到有效调查问卷共 198 份。

通过对调查问卷的统计分析，得到医院信息化人才现状为：人员年龄比例中，信息化后备力量准备不足；人员学历比例中本科占据多数；人员学习动力不足，阻碍人员能力提升的因素众多；调查人员对信息化工作的定位大部分认为是“技术支持”，但多数医院现行体制中仍被认定为“后勤保障”；调查人员中工作氛围和学习氛围均呈现较好氛围，这对于人才培养有正向的推动；调查人员对工资待遇的满意度一般；对需要提升的能力范围提出了多方位的要求；同时，也提出了多元化的能力提升方式。

三、医疗信息化人才队伍建设培养方案

根据对人才培养现状的调查数据进行研究，对比全国医院信息技术人员的数据，利用小组讨论、头脑风暴、鱼骨图、数据分析等技术，对人才匮乏原因进行分析。分析结果表明医院信息化人才队伍匮乏的原因包括：历史原因、定位原因、经济原因、政策原因、个人发展空间原因和其他原因等。根据“人本发展理论”、“马斯洛需求层次理论”进行研究，结合信息化人才队伍现状和智慧医院建设的要求，制定智慧医院医疗信息化人才队伍建设的人才培养方案，具体包括三个方面：人才培养方向和目标、人才培养方法、人才培养制度。

1、人才培养方向和目标

根据“人本发展理论”，人的发展受环境、资源的影响，通过资源合理配置可以全面提升人的能力。医疗信息化人才培养受到医疗政策和医院规划等外部环境的影响。因此，制定医疗信息化人才培养方向和目标，要做到以下几点：首先，人才培养要开拓思想、了解国家医疗信息化发展方向是智慧医院建设中人才培养的总方向和总指引，制定人才培养的总目标。其次，在智慧医院建设的大环境下，医疗信息化人才的培养、人才储备，要以医院总体人才建设规划为导向。最后，根据国家医疗信息化发展方向，以及智慧医院建设中对人才培养的需求作为总方向和总指引，以医院总体人才建设规划为导向，制定人才培养的总目标。

2、人才培养方法

根据“马斯洛需求层次理论”在职场中的应用，针对不同的人员制定不同的培养方案方法。

（1）信息化人才专业技能培养。医院信息技术能力提升范围包罗万象，包括网络技术、网络安全、大数据、数据库技术、人工智能、项目管理、信息安全、项目管理、政府采购招投标管理等等。人才培养方法主要包含：培训、职称考试、参观学习、项目带动人才建设、问题带动人才建设、竞赛、科研、会议等。

（2）临床医疗人员信息化技术培养。依托医院信息化系统建设与应用，培养临床信息化人才，在临床科室设立信息联络员，负责该科室的信息化问题反馈以及科室内部培训。

（3）专业技能以外的技术培养。医疗信息化涉及领域众多，信息部门支持鼓励技术人员不断拓宽知识领域，勇于探索相关知识。

（4）信息化人才心理疏导。信息技术人才培养周期长，见效慢，信息技术人员要保持长期的学习状态，在长期的持续坚持下，取得长足的进步。因此，对信息化人才的心理疏导不容忽视，通过定期组织团建、聚餐等集体活动，增进团队凝聚力、提高员工归属感。

3、人才培养制度

根据“人本发展理论”，通过制度可以创新开发人的发展潜能手段，医院信息化管理制度是医院人才培养的有力抓手，使得各项工作有制度可依，有制度必依，推进人才培养工作制度化、体系化、标准化。我院制定了一系列信息管理制度、信息安全管理制度、信息化人才培养制度、知识产权保护制度等。

四、总结

截至研究结束时，人才培养取得的成果：①刷新医院对信息技术的定位与认知。②科研攻关稳步进行。③科研能力提升带动经济效益。④学习氛围浓厚，职称晋升突破性增长。⑤完善各项信息化制度，使得医院信息化建设中的各项工作趋于流程化、标准化、规范化。

本项目研究仍存在一些问题需要解决，主要包括如下：多栖人才队伍的培养缺口仍然存在；薪酬激励措施少；人才培养周期长等。这些问题的研究涉及周期长，并受科学技术发展及社会进步的影响大，需要长期关注和研究。

因此，下一阶段的工作，应依据国家医疗信息化发展方向与智慧医院建设中对人才培养的需求，以医院总体人才建设规划为导向，不断完善医院信息化人才培养方案，培养多元化、全方位的信息化人才队伍，为实现智慧医院建设和医院高质量发展提供稳健的技术支撑和强有力的技术支持。

PU-154

医学信息素养教育理论实践

吴含
郑州大学第一附属医院

引言：近年来，随着互联网技术的飞速发展，信息时代已经全面来临。在医学领域，越来越多的医学资源、研究成果和临床经验以信息化的形式在互联网上广泛传播。然而，信息的爆炸性增长也给医学生们带来了崭新的挑战，他们需要具备信息获取、信息评估、信息应用和信息管理的能

力，以满足学习、实践和研究的需求。因此，医学信息素养的重要性日益凸显。

一、医学信息素养的概念与特点

1. 医学信息素养的概念：

医学信息素养是指医学专业人员具备的获取、评估、应用和管理医学信息的能力。其中，信息获取能力包括对医学文献、研究成果等信息的有效检索和筛选；信息评估能力指对所获取信息的准确性、可靠性和适用性进行判断；信息应用能力是将所获信息应用于临床实践、科研和管理中的能力；信息管理能力包括对获取的信息进行整理、归纳和保存的能力。

2. 医学信息素养的特点：

（1）数据量大、复杂性高：生物医学领域作为科技发展最迅速的领域之一，医学信息资源不断增加，信息数据量庞大，而且往往具有复杂性，包括疾病信息、药物研究、临床试验等

（2）资源多样、文种繁多：医学信息不仅包括传统的医药科技文献，如图书、期刊、特种文献，还包括电子出版物、网络文献，以及书写型和电子型病历、诊断报告等。这些资源的多样性要求医学人员能够灵活应对不同类型的信息

（3）学科交叉、内容分散：医学领域与其他领域交叉渗透，如信息科学、工程学、物理学、化学等。医学信息素养需要在多学科知识的

（4）半衰期短、更新迅速：医学信息的更新速度非常快，知识的半衰期很短，新的研究、技术和治疗方法不断涌现。因此，医学人员需要不断更新自己的知识和信息[2]

二、医学信息素养教育的现状和问题

1. 现状：目前，医学信息素养教育已经在理论和实践中取得初步进展。许多高校开始开设医学信息素养相关课程和活动，增强了对这一教育领域的关注。同时，在临床实践中，不少医疗机构引入了信息技术，加强了医学信息素养培训，提升了医护人员的信息素养水平。

2. 问题：尽管医学信息素养教育有了一些进展，但仍存在一些问题。首先，教育目标不够明确，缺乏明确的衡量标准，难以准确评估学生的信息素养水平。其次，教学方式和内容相对单一，缺乏灵活性和针对性。此外，教师的信息素养和教育能力不足，影响了教育质量[3]。

三、医学信息素养教育的基本原则与方法

1. 教师素质与培训：医学信息素养教育的成功与教师素质直接相关。教师需要掌握先进的信息技术知识和教育理念，以便更好地引导学生。然而，目前教师队伍中仍有部分人员缺乏相关知识和培训，因此需要加强教师培训，提升其信息素养和教学能力。

2. 课程设计与实施： 医学信息素养教育的课程设计应当贴近临床实际，结合学生的需求和未来的职业发展。课程内容应丰富多样，包括信息获取、评估、应用、管理等方面，同时要与医学领域的最新发展保持同步。

3. 评估体系建立：目前缺乏科学合理的医学信息素养评估体系，难以全面衡量学生的信息素养水平。建立完善的评估体系，包括知识、技能和实际应用的综合评价，有助于指导教学改进和学生能力培养。

4. 教育技术创新：随着信息技术的迅速发展，教育技术在医学信息素养教育中的作用日益凸显。应该积极探索新的教育技术应用，如虚拟实验、在线学习平台等，以提升教学效果。

5. 跨学科合作： 医学信息素养涉及医学、信息技术、教育学等多个领域的知识。建立跨学科合作的机制，促进不同领域专家的合作，可以为医学信息素养教育提供更丰富的资源和视角。

6. 教育宣传与推广： 医学信息素养教育的重要性需要被更多人认知。加强教育宣传和推广，提高学生、教师和社会大众对医学信息素养的认识，促进教育的普及。

7. 信息伦理与法律教育： 医学信息素养教育应该强调信息伦理和法律意识的培养，引导学生正确、合法地获取和应用医学信息，遵循信息使用的道德规范和法律法规。

8. 教育资源共享： 建立医学信息素养教育的资源共享平台，可以促进教育资源的交流和分享。教师可以从中获取教材、案例和教学方法，提升教育质量。

9. 实践环节强化： 医学信息素养教育需要强化实践环节，让学生在临床实际中应用所学知识，提高解决问题的能力和信心。

10. 终身学习观念培养： 医学信息素养教育应该培养学生具备终身学习的观念，鼓励他们持续关注医学和信息技术的发展，不断更新知识[4]。

四、医学信息素养教育的实际应用与效果

医学信息素养教育的实际应用和效果评估是确保教育成果的关键环节。通过实际案例的介绍以及对学生在临床实践中的表现的观察，可以更加具体地了解医学信息素养教育的影响和成果。

实验设计与方法

为了评估医学信息素养教育的实际应用与效果，可以采用以下方法：

1. 设计实验组和对照组： 将学生分成实验组和对照组，实验组接受医学信息素养教育，对照组采用传统教学方法。这样可以比较两组学生在信息素养方面的差异。

2. 教学内容设计： 在实验组中，设计涵盖信息获取、评估、应用等方面的丰富教学内容，使学生全面掌握相关知识和技能。

3. 教学方法： 实验组中可以采用问题解决、案例分析、小组讨论等互动教学方法，培养学生主动学习和实践能力。

4. 效果评估： 设计评估指标，包括学生信息素养水平的提升、实际应用能力的增强等，可以通过问卷调查、实际操作评估等方法进行量化评估。

实践案例介绍与效果评估

举一个实际案例，例如在临床医学课程中引入真实病例，要求学生通过信息检索工具获取相关文献，分析病情，提出诊疗方案，并在小组中展示。通过实际案例的应用，可以观察到以下效果：

1. 信息获取能力提升： 经过医学信息素养教育，学生能够更快速、准确地从信息中获取医学知识，提高了信息检索能力。

2. 信息评估能力增强： 学生学会了对医学文献的评估，能够判断信息的可靠性和适用性，降低了信息误导的风险。

3. 实际应用能力提高： 在实际案例中，学生需要将所学的信息应用于病情分析和治疗方案制定，培养了将理论知识转化为实际操作的能力。

4. 团队合作与交流：学生在小组讨论和展示中，增强了团队合作和交流能力，通过分享不同观点促进了知识的深入理解。

5. 自主学习能力培养：医学信息素养教育鼓励学生主动获取信息、解决问题，培养了他们的自主学习意识和能力。

五、医学信息素养教育面临的挑战和发展对策：

医学信息素养教育虽然取得了一些进展，但仍然面临一些挑战。了解并应对这些挑战，制定相应的发展对策，对于医学信息素养教育的全面发展至关重要。

挑战：

1. 教师素质不足： 医学信息素养教育需要教师具备医学知识和信息技术背景，但部分教师在信息素养和教育方法方面可能存在不足。

2. 教学方法单一： 目前的医学信息素养教育教学方法主要以理论课为主，缺乏实践和互动环节，难以激发学生的学习兴趣和动力。

3. 评估体系不完善： 缺乏科学有效的评估体系，难以准确地评价学生在信息素养方面的提升和成长。

4. 跨学科整合难度： 医学信息素养涉及医学、信息技术等多个学科领域，跨学科整合与合作可能面临难度。

发展对策：

1. 提升教师素质： 加强教师培训，提升他们的信息素养和教育能力，使其能够更好地引导学生培养信息素养。

2. 多元化教学方法： 探索多种教学方法，如案例教学、实践操作、小组讨论等，激发学生积极参与，提高教学效果。

3. 构建评估体系： 建立完善的评估指标体系，包括知识掌握、实际应用、创新能力等方面，为教学评估提供科学依据。

4. 跨学科合作： 建立跨学科合作机制，促进医学、信息技术等领域专家之间的合作，共同推动医学信息素养教育的发展。

5. 强化实践环节： 增加实践环节，如临床实习、实际案例分析等，使学生能够将所学知识应用于实际情境。

6. 推广宣传与资源共享： 加强医学信息素养教育的宣传力度，同时建立教育资源共享平台，促进教育经验和资源的交流与分享。

通过采取这些发展对策，医学信息素养教育将能够更好地应对挑战，不断提升教育质量，培养更具信息素养的医学专业人才[5]。

结论：医学信息素养教育是培养医学生信息获取、评估、应用和管理能力的重要途径。通过教育理论的借鉴和实践的应用，医学信息素养教育能够提高学生信息素养能力。然而，教育仍面临教育目标明确、教学方法多样化等挑战。未来，应创新教育技术、加强师资培训，推动医学信息素养教育的发展。

PU-155

慢性病用户健康信息搜寻行为的影响因素探索

赵梦圆[1]，应峻[1]
1. 复旦大学图书馆

1.引言

我国人口慢性病的发病率在逐渐增加。为了预防和控制慢性病的发作或恶化，患者与慢性病的抗衡是一场持久战，研究表明慢性病患者对健康信息的需求相对较高[1]。特别是当身体出现健康问题时，用户会更加积极地搜寻健康信息来解答自身的健康问题[2]。虽然大家都普遍认为专业的医务人员是最可靠的健康信息来源[3]，但这个问诊过程通常是耗时且昂贵的。尤其对于慢性病患者来说，慢性病具有的病因复杂、病程长、难治愈、易反复等特点让他们在日常生活中更加需要关注健康信息进行自我健康管理。而且互联网给用户提供了分享、交流、传播和咨询等健康信息的各种平台，许多研究都表明用户会经常在互联网上搜寻与健康相关的信息。

健康信息搜寻行为（Health Information Seeking Behaviour，HISB）是近年来信息科学、公共卫生、卫生传播等领域共同关注的话题，然而，专门探究慢性病用户的健康信息搜寻行为的研究还相对匮乏。已有一些研究讨论了年龄对用户健康信息行为的影响[4,5]，不同年龄段的用户在搜寻健康信息的过程中表现出不同的搜寻策略。同时，一些研究注意到患有特定疾病用户的健康信息搜寻行为，比如癌症[6,7]、心理疾病[8,9]和物理疾病[10,11]。伴随着生活节奏的逐渐加快以及工作或学习压力的不断增加，慢性病已经成为严重威胁身体健康的公共卫生问题，对于慢性病患者来说日常生活中的自我健康管理显得尤为重要。因此，慢性病患者在日常生活中如何利用互联网搜寻健康信息值得进一步的探索，通过充分掌握慢性病患者的健康信息搜寻行为的特征以及影响因素，将有助于后续一系列慢性病管理工作的科学有序开展。

2.文献综述

目前，国内外许多学者对用户健康搜寻行为的影响因素进行研究[12]，如 Chisolm 采用行为模型进行建模，通过分层逻辑回归等算法研究了健康信息搜寻行为及影响因素，指出健康信息搜寻行为不是单一的行为；性别、年龄大于 65 岁、当前健康危机、经常使用网络的习惯是对用户健康信息搜寻行为最稳定的预测因素[13]。总的来说，学者们普遍关注的影响因素主要包括四个方面：社会人口因素、认知因素、心理因素和情境因素。

从已有的相关研究来看，学者普遍考虑到的影响用户健康信息搜寻行为的社会人口学因素主要有年龄、性别、社会经济地位等，而且在不同的研究情境下的社会人口学特征对健康信息搜寻行为的作用不尽相同。随着年龄的增长，用户一方面要面临更多的健康问题，促使他们搜索更多的健康相关信息；另一方面，身体机能和自我效能的逐渐下降，使他们无法寻求健康信息。研究结果普遍显示，年龄越大的用户越不太可能通过互联网获取健康信息[14]。受教育程度对用户的健康信息搜寻行为有显著的积极影响[15-17]，通常教育水平高的用户的计算机技能和健康素养越高，这些都有利于他们从互联网上搜索健康信息[17,18]。而对于用户搜寻健康信息行为方面的性别差异在各个研究中的结果并不一致。

认知因素是影响用户健康信息搜寻行为的内在特征[19]，如健康素养、感知有用性和感知易用性。认知能力是用户在互联网上准确搜寻健康信息成功的关键。大多数研究都表明，健康素养高的用户更有可能定期或经常使用互联网阅读健康信息[20]。此外，许多学者利用技术接受模型来探究影响健康信息搜寻行为的信息通信技术因素，即感知有用性和感知易用性[16,21]。

用户的心理因素也会决定他们搜寻健康信息的行为，比如健康信息的信任程度、对健康的态度和信念等。首先，对健康信息网站的信任是影响用户使用互联网搜寻健康信息的决定性因素，通常对在线健康信息越信任的用户越愿意将互联网作为健康信息来源[22]。用户的健康意识在很大程度上反映了他们对自身健康的态度和努力，对自身健康的关注增加了他们获取健康信息的意愿，这促使他们更有可能去搜索健康信息[23]。此外，担心自己健康状况的用户也更有可能在网上搜索健康信息，从而带给想自己心理安慰[24]。

在不同情境下，用户的健康信息搜寻行为也是有所差别的。有研究发现用户所患慢性病的数量

与互联网使用呈正相关，但疾病引起的生理和心理障碍也会在一定程度上阻碍互联网的使用[25]。不同疾病类型的患者在治疗疾病和健康管理过程中，其健康信息行为存在差异。尤其对于患有尴尬或精神疾病的患者，他们更有可能访问在线健康信息资源，比如低收入家庭的抑郁症患者更倾向于使用互联网[25]。

3.研究设计

3.1 数据收集

（1）访谈对象

本研究是以慢性病患者为研究对象来探究其健康信息搜寻行为，因此参与访谈的志愿者需被诊断患有至少一种慢性病。笔者采取线上和线下两种方式来招募访谈对象。一是在线招募。笔者为了招募访谈对象先设计了前测问卷，包括性别、出生年份、联系方式、常住地、受教育水平等基本信息，将招募信息发放至慢性病相关的在线健康社区，并附有调查问卷，想要参与调研的志愿者可通过填写问卷等待联系，每位参与正式访谈的用户可获得 20 元的红包奖励。二是线下招募。通过滚雪球的方式从身边的亲朋好友招募符合条件的访谈对象。最终，参与正式访谈的样本量共有 23 人，其中，从线上挑选了 12 人参与调研，线下有 11 人接受访谈，受访者的基本信息如表 1 所示。

表 1 受访者的基本信息

受访者编号	性别	年龄	受教育水平	常住地	健康状况	慢性病类型	健康信息素养
P1		50	小学及以下	农村	很好	高血压和颈椎病	2.86
P2		69	高中	农村	较好	高血压	3.71
P3		65	初中	城镇	较好	高血压	3.71
P4		51	小学及以下	城镇	较好	高血压和糖尿病	2.00
P5		52	小学及以下	城镇	一般	慢性胃炎	2.29
P6		51	高中	城镇	较差	慢性胃炎、关节炎和高血脂	4.00
P7		26	研究生	城镇	一般	腰椎膨出	4.29
P8		51	本科	城镇	一般	高血压	3.43
P9		27	研究生	城镇	一般	鼻炎	2.14
P10		51	本科	城镇	一般	高血压	4.29
P11		24	研究生	城镇	较好	突发性耳聋和皮肤病	4.29
P12		22	大专	城镇	较好	慢性结膜炎	4.57
P13		27	本科	城镇	较好	鼻炎和皮肤病	4.86

P14		26	本科	城镇	较差	慢性胃炎和鼻炎	4.71
P15		23	本科	农村	一般	慢性胃炎、鼻炎和慢性结膜炎	4.29
P16		28	本科	农村	较差	慢性结膜炎和皮肤病	3.43
P17		41	本科	城镇	一般	糖尿病、鼻炎和慢性结膜炎	4.43
P18		55	小学及以下	农村	一般	高血压、慢性胃炎和高血脂	3.57
P19		43	中专	城镇	较差	慢性结膜炎	4.43
P20		46	中专	城镇	一般	慢性胃炎	4.14
P21		60	高中	农村	一般	高血压、风湿和慢性胃炎	3.71
P22		32	高中	农村	很好	慢性胃炎和腰椎膨出	4.00
P23		38	高中	农村	一般	慢性胃炎	4.43

（2）访谈过程

本研究的访谈工作从 2021 年 6 月开展至 2022 年 3 月，采用面对面或者在线语音的方式对 23 位受访对象进行半结构化访谈，每次访谈持续约 40-60 分钟。在正式访谈开始前，笔者会将访谈主题与目的等内容介绍给受访者，并征得他们同样的前提下对访谈过程进行录音。此外，为了保障访谈问题设置的合理性，先对 2 位关注健康信息的慢性病患者进行预访谈，通过参与者在访谈过程中的反馈以及意见，进一步改进和完善访谈大纲。最终，经过几轮修改和完善形成了正式访谈大纲，具体内容如下：（1）您平常会获取到哪些类型的健康相关信息？您对哪些方面的健康信息最感兴趣？（2）您通常会在哪些渠道搜索或阅读与健康相关的信息？您经常使用哪个渠来道获取健康信息？为什么？（3）您如何评价所获取到的健康相关信息？哪些因素会影响您对其可信性的判断？（4）您在什么情境下会想要主动搜索与健康相关的信息？您在什么情境下会主动浏览健康信息？您在什么情境下会让其他人帮助搜索健康相关信息？您在什么情境下会偶然看到与健康相关的信息？（5）您搜寻健康信息的过程是否顺利，遇到的障碍有哪些？（6）您平常如何应用自己所获取的健康信息？为了能够深入了解受访者的真实且全面的感受和经验，笔者在访谈过程中通过逻辑递进和思维发散的模式根据访谈对象的回答来即时调整访谈问题，尽可能地让访谈对象唤醒更多回忆并表达自己的观点和经验。

3.2 数据分析

笔者借鉴 Strauss 提出的扎根理论三阶段分析法对访谈过程中获取的原始资料按照开放式编码到主轴编码再到选择性编码的流程进行主题分析，逐步概念化、范畴化和主范畴化，进而归纳互联网环境下慢性病患者健康信息搜寻行为的影响因素。为了保证编码结果的客观性和可信性，本研究邀请了一位研究生和一位博士生对 18 份的访谈记录进行编码，其余 5 份用于理论饱和度检验。首先，两位编码者分别对 18 份访谈记录进行独立分析和编码；然后，将两位编码者的编码结果进行整合分析，并针对不同之处进行讨论和再次确认，形成最终的编码结果。

4.研究结果与讨论

通过对在线慢性病社区的患者以及可接触的慢性病患者进行访谈招募，系统围绕“您在日常生活中搜寻哪些类型的健康信息”“您搜寻健康信息过程中遇到哪些障碍”等问题，采用线上和线下相结合的方式对慢性病患者展开半结构化访谈，获取文本、录音等方式的访谈数据，进而按照开放式编

码到主轴编码再到选择性编码的流程进行主题编码。编码结果表明在慢性病患者搜寻健康信息的过程中，患者的健康信息搜寻行为主要受到其心理因素、情感感知、社会支持、信息因素和认知能力的共同影响。

4.1 心理因素

慢性病用户所提及的影响其健康信息搜寻行为的心理因素主要包括感知易感性、感知严重性、焦虑和态度这四个方面。感知易感性是指慢性病患者认为自己很容易患上健康信息中提到的健康相关问题[26]。许多受访者看到健康信息时会习惯性地对号入座，下意识将自身的情况带入，判断感染疾病的可能性，很容易产生恐慌，幻想着自己处于健康风险中。感知严重性在这里指慢性病患者认为接触到的健康信息中所描述的健康问题后果比较严重[26]。部分受访者提到百度的搜索结果，总是反馈一些负面信息，会激发他们想要获取更多关于健康问题的信息来了解并进一步确定病情。这同一些学者的研究结果相一致，即用户对疾病的感知易感性和感知严重性正向影响其使用网站获取健康信息的行为[27]。受访者的焦虑主要体现在两个方面，一是对自身健康状况的担忧，二是对互联网使用过程中遇到的问题而烦躁。尤其是当他们身体出现健康问题的时候，很容易陷入胡思乱想的状态，过度担心自己的健康状况。另外，年龄稍长且受教育水平较低的受访者在使用互联网过程中很容易处于焦虑状态，一旦手机界面不是往常所熟悉的操作，就无所适从。尤其对于老年人来说，信息技术使用的焦虑是阻碍其使用互联网搜寻健康信息的阻碍之一[21]。态度主要表现在患者对于互联网上健康信息的直观感受，部分受访者还是保持积极态度，对互联网上提供的健康信息很感兴趣，认为多了解健康知识可以有备无患。通常来说，当用户对互联网使用持积极态度时，他们从互联网上寻求健康信息的意愿更强[15]。由此可见，慢性病患者的心理因素是影响其产生健康信息搜寻行为的重要因素。

4.2 情感感知

情感感知是指慢性病患者对外部因素和内部因素的反应而引起的情感倾向所产生的健康信息搜寻行为的动机，主要涉及感知有帮助、感知障碍、感知健康风险和感知隐私风险这四个方面，能够直接影响慢性病患者的健康信息搜寻行为。感知有帮助主要体现在慢性病患者认为所获得的健康信息大多是有用的、有价值的，能够帮助自己解决健康问题。笔者发现他们的共同点是倾向于相信健康养生类的信息，即使亲身尝试健康信息中提到的方法，无论失败与否也都不会造成很大的损害。McKinley 和 Wright 的研究结果也表明大学生从互联网上获取的营养/健康信息的感知有帮助对其在线搜索健康信息的频率有积极作用[28]。感知障碍在这里指慢性病患者使用互联网获取所需健康信息的过程中会遇到一些困难，比如手机字体太小、界面设计复杂、注册过程繁琐等。尤其是年龄渐长的慢性病患者，会伴随着身体机能的下降，比如视力衰退、认知退化等症状，增加了他们搜寻健康信息的负担。有研究表明，更清晰、视觉效果更好的网站更能吸引用户来获取健康信息[29]。感知健康风险是指慢性病患者对自己可能面临的健康威胁的主观感受。由于慢性病具有病因复杂、病程长、难治愈、易反复等特点，患者经常处于疾病发作的恐慌中，尤其是当身体出现不舒服的情况，更会增加他们对自己健康状况的担忧。Ahadzadeh 等人的研究结果也表明感知健康风险能够影响用户使用互联网搜寻健康信息的行为[26]。感知隐私风险是指慢性病患者担心自己在互联网上的操作会泄露信息甚至造成财产损失。尤其是许多软件需要填写个人信息才能使用，但有些软件后台却泄露用户信息，有的受访者就提到平常会经常收到一些骚扰电话或短信，这可能会抑制患者搜寻在线健康信息的行为。用户从使用互联网感知到的隐私问题的风险也是阻碍其搜寻在线信息的一个重要因素[30]。整体来看，慢性病患者的情感感知对其健康信息搜寻行为有一定的影响作用。

4.3 社会支持

社会支持是影响慢性病用户健康信息搜寻行为的外部因素，通常是指来自个人之外的各种支持总称，在本研究中主要涉及信息支持、情感支持和家人支持三个方面。信息支持主要指互联网上的其他用户为慢性病患者提供信息和建议，并帮助他们发现原因，比如有类似情况的病友分享贴、专业医生科普、在线专家问诊等。部分受访者提到对一些细节性的经历描述很感兴趣，参考价值也相对较高。情感支持体现在慢性病患者从互联网上得到的安慰、关怀和同情，这在参与慢性病在线健康社区的活跃用户上表现比较明显。他们会在微信群、百度贴吧或豆瓣小组等病友交流平台积极分

享自己的日常生活以及困惑，社区成员通常会互帮互助并提供个人经验，自然而然地形成一种归属感。从已有研究来看，Li 和 Wang 从社会支持的视角调查了用户在社交媒体上搜寻健康信息的行为意图，其中，信息支持和情感支持可以显著影响用户对互联网健康服务的信任进而影响其健康信息搜寻意图[31]。此外，需要家人支持的患者可以归纳为两种类型，一是其家人从事与医疗卫生有关的工作，可以从家人那里很方便地获取专业的健康知识来解答疑问；二是年龄稍长且文化水平较低的患者，他们自己无法从互联网上主动获取所需的健康信息，倾向于通过替代搜索来满足其健康信息需求。家人支持对用户的健康信息搜寻行为的影响也是学者讨论的话题[32]。当慢性病患者感知所得到的社会支持能够帮助自己解决问题时，会产生良性循环促使他们不断获取健康信息。基于此，笔者认为从外部获取的社会支持对慢性病患者健康信息的搜寻行为有着重要的影响。

4.4 认知能力

慢性病患者的认知能力主要体现在健康信息素养、自我效能和搜索技能有限这三个方面。虽然慢性病患者心理因素、情感感知、社会支持三个因素对其健康信息搜寻行为有直接影响，但个人认知能力也间接影响着患者的健康信息搜寻行为。健康信息素养被定义为一种能力，即认识到对健康信息的需求、知道如何及在哪里找到健康相关的信息以及如何在日常生活中评估和使用这些信息。对于健康信息素养高的患者来说，他们擅于使用互联网并具有一定的健康知识积累，因此一旦产生健康信息需求，有能力直接去互联网上搜索来寻找正确的答案。Lam 和 Lam 探讨了澳洲老年人在互联网上搜寻健康信息的行为与健康信息素养之间的关系，结果发现对于受教育水平较高的用户群体，其健康信息素养和在线健康信息搜寻行为显著相关[33]。自我效能是慢性病患者对自己能否成功使用互联网搜寻到健康信息的主观判断。部分受访者表示在互联网上的搜索经验较丰富，知道如何构建检索式来描述自己的问题，对自己能够成功搜索到所需要的健康信息很有信心。但也有部分受访者表示搜索技能有限，对于互联网使用存在畏难心理，尤其是年龄稍长且文化水平较低的患者，不愿意去学习新技能，无法构建准确的检索式来描述自己的健康问题，导致搜索结果不理想，这阻碍了他们想要在互联网上搜寻健康信息的行为[34]。

4.5 信息因素

信息因素主要包括信息质量、信息过载、信息类型和信息获取这四个方面。部分受访者意识到互联网上健康信息的质量参差不齐，一是广告信息层出不穷，而且宣传方式越加智能和多样，很难判断出商品质量的真假；二是虚假信息遍布，互联网允许用户自由发表观点，却缺少管控信息质量的平台或机构，让用户充斥在海量的真假信息中，增加了他们获取所需健康信息的负担[35]。同时，这也让用户产生信息过载的感受，很难从繁杂的在线信息中筛选出适用于自己情况的病情信息；而且许多软件通过大数据分析对用户进行个性化推荐，但有些受访者提到会在多个软件上被推荐重复的信息，这会对他们浏览健康信息的过程产生干扰。Jensen 等学者的研究结果表明搜寻健康信息的用户相较于未搜寻者具有更低的信息过载得分[36]。笔者发现用户对信息类型的偏好存在年龄差异，年龄稍长的用户更喜欢通过视频形式来接收和理解信息，比如抖音、快手、小红书和微信“看一看”[37]，年轻人更愿意在知乎、豆瓣小组或百度贴吧以文字形式来交流和获取信息。虽然从互联网上获取健康信息有利有弊，但大部分受访者都愿意使用互联网进行主动搜索或者浏览健康信息，因为不仅允许用户随时随地搜索健康信息，而且使用互联网获取健康信息也很方便快捷，可以很快得到反馈解答自己的健康问题，省去预约、挂号、排队等一系列过程，同时有助于改善医患之间的信息不对称状态。基于上述分析，笔者认为信息因素可以影响慢性病用户的健康信息搜寻行为。

5.研究启示

本研究立足于慢性病用户的健康信息搜寻行为，在广泛阅读和梳理国内外相关研究的基础上，利用定性研究方法深入挖掘患者的健康信息搜寻特征及其影响因素，为进一步开展实证研究探索影响因素与健康信息搜寻行为具体关系提供理论基础和框架借鉴，从而有助于政府机构、互联网传播平台等多方力量为提升慢性病患者健康信息素养水平做出努力。

对于政府相关卫生部门，本研究为相关管理机构了解在线健康信息现状和问题提供一定的参考，有助于管理部门采取有效措施对平台进行监督、引导；助于卫生部门了解慢性病用户获取健康信息进行自我健康管理的情况和障碍，优化已有的慢性病综合防治工作机制和防治服务网络，积极

响应“健康中国”策略。从定性访谈中可以发现，慢性病患者对于互联网上健康信息质量无法做出准确的判断，而且还经常遇到广告信息以及虚假信息，这都对他们搜寻在线健康产生干扰。不过，多个受访者提到会比较信任官方认证的信息来源。因此，政府机构有必要完善互联网上信息传播秩序，营造良好的网络信息环境；同时相关专业机构有责任加强健康教育，提升慢性病患者甚至全民的健康素质，不给不良媒体和不良机构以可乘之机。因此，政府相关部门有必要提倡健康文明的生活方式，鼓励专业机构为不同群体开展个性化的健康指导，增强群众维护和促进自身健康的能力。

对于互联网健康信息传播平台，本研究为他们展现出慢性病用户的健康信息搜寻行为及其影响因素，有助于其构建慢性病患者健康信息搜寻行为的用户画像，为不同类型的患者群体提供个性化的服务，完善互联网上健康信息传播平台的建设并优化服务。从访谈结果来看，受访者提到的在线健康信息渠道主要包括社交平台（微信、微博、豆瓣）、短视频平台（抖音、快手）、社会问答平台（知乎、百度贴吧、百度知道）、医疗健康类 APP（好大夫在线、丁香医生、平安健康、蛋壳健康、小荷）、搜索引擎（百度、浏览器）、新闻客户端（今日头条、腾讯新闻），而且他们对不同健康信息来源的信任程度有差异，大多数受访者认为专业医疗平台的可信度最高以及百度搜索结果的可信度最差。同时他们也提到 APP 过度推荐信息以及注册过程繁琐等现象。因此，互联网平台可以进一步了解用户的需求适当推荐精准信息避免让用户反感的信息推荐，尽可能简化 APP 的注册过程，在设计过程中也要考虑老年人的需求，从而为用户提供更加个性化和人性化的服务。

6.结语

本研究仅通过访谈的方式初步探索慢性病患者健康信息搜寻行为的影响因素，但无法明确每个因素与健康信息搜寻行为之间的具体关系，还需借鉴成熟理论利用定量研究方法明晰深层次的影响机理。未来的研究也可以采用移动经验取样等方法，获取患者日常生活情境下信息行为的认知、心理、情感与体验等数据，更加真实揭示慢性病患者的健康信息搜寻行为特征，以进一步探究其健康信息搜寻行为的影响因素及变量关系，从而提升慢病患者健康信息搜寻质量，提升其生活满意度。

PU-156

高校医学图书馆与医学生信息素养教育的思考

李华
空军军医大学图书馆

当今时代，医学信息海量增长，医学知识更新周期不断加快，医务工作者面临着日益复杂的临床诊疗和科研等工作。医学生的信息素养主要指医学生能意识到信息需求，并进行检索、评估和有效利用信息，主要有信息意识、信息能力和信息道德三个方面，具有自主认识并运用信息提高学习水平的综合能力[1]。以医学信息获取、评价和利用等处理能力为核心的信息素养是医学生综合素养的核心，医学信息素养能力将成为医学生从事科研工作的必备条件。

一 医学模式的改变要求医学生有更高的信息素养

医学模式是人们对待和处理疾病、健康问题的态度或方式。早期是生物医学模式，基于特异性诊断和治疗的发展，有效保障人类生命健康。但这种医学模式忽视人的社会属性，没有考虑到人类心理和社会因素对健康的影响。随着工业化、城镇化、信息化快速推进，人们生活方式和思想观念发生了巨大变化，同时给人们身心健康带来很大挑战。在生物－心理－社会医学模式背景下，医学主要目的是鼓励普通大众参与到健康保健和疾病预防之中，通过增强自身健康意识，促进健康水平不断提升。随着“以疾病为中心”向“以健康为中心”转变，迫切需要进一步提高医学生的信息素养，满足现代社会需要医生由“治已病”向“治未病”转变。

二 加强医学生信息素养教育的必要性

1.1 医学生信息素养意识薄弱

医学生在学习过程普遍存在的问题是：许多本科生没有清晰的认识到信息的重要性和价值，缺乏对信息的敏锐度，信息意识淡薄，不能很好的运用信息资源和检索工具来解决问题。缺乏独立思考和判断能力，缺乏对信息来源的正确认知能力。高校图书馆员要对医学生进行引导，要培养他们的信息意识，积极运用信息技术，养成良好的信息习惯。

1.2 信息素养技能不足

许多医学院校开设医学文献检索课程，但由于教学课程内容枯燥、呆板，没有根据专业、学生类别进行个性化教学。很多医学生在学习过程中，不关注医学专业期刊文献和数据库，缺乏熟练使用检索工具和各种资源获取信息的能力。医学生要积极利用学校现有的期刊、工具书和数据库资源查找所需的资料，要学会分析、评价各种信息，提高自身信息获取能力。高校图书馆对医学生信息素养技能教育工作仍有很大空间。

三 高校图书馆在加强医学生信息素养教育中的对策

对医学生进行信息素养教育方面，高校图书馆具有信息资源优势、信息环境优势、信息技术优势、信息人才优势。因此，高校图书馆应积极利用自身丰富的文献资源和现代化的信息设施，为提高医学生的信息素养服务。培养具有信息素养、终身学习能力的人才，能够在信息时代具备获取信息与辨别利用信息的能力[2]。

3.1 加强高校图书馆信息资源建设

图书馆是高校的信息资源中心，拥有丰富的馆藏资源和完整系统的文献资源保障体系，是医学生学习知识、获取信息的重要场所。在文献信息资源建设上，应根据学校的发展目标和教学、科研的需要，制订文献信息资源建设方案，形成具有特色的馆藏体系。开展特色数字资源建设和网络虚拟资源建设，整合实体资源与虚拟资源，形成网上统一的馆藏体系。应科学合理地组织馆藏，既有利于文献信息的管理和保护，更要有利于文献信息的充分利用，为提升医学生信息素养提供物质保障。

3.2 创造泛在的信息素养教育环境

网络改变了用户的行为模式与学习习惯，医学生不一定到图书馆培训室才能接受信息素养教育。高校图书馆的信息素养教育必须适应这种改变，利用网络平台，创造泛在的信息素养教育环境。泛在的信息环境突破物理空间，不受时间限制，让医学生在何时何地都能很方便地接收到信息素养教育的内容，即将与信息素养教育有关的资源提供在网上供读医学生自行学习与浏览。除了将

馆内开设的信息素养教育课程、医学生培训讲座的相关课件放到图书馆网页上，还可将网络中的信息素养教育资源进行整理利用，让医学生能够方便、快捷地在图书馆网站上就查询到自己所需要的信息。

3.3 利用新媒体进行信息素养教育

网络视频因其直观生动、短小精悍的特点，用户可以观看、互动和分享的社交属性，适应现代社会生活快节奏，成为人们离不开的一种社交工具[3]。网络视频广泛应用，不仅给人们社会交往带来便利，也给知识生产、传播、获得带来新变化，使其成为人们获取信息的重要工具。根据中国互联网络信息中心（CNNIC）发布第 49 次《中国互联网络发展状况统计报告》显示，截至 2021 年 12 月，在网民中，即时通信、网络视频、短视频用户使用率分别为 97.5%、94.5%和 90.5%，用户规模分别达 10.07 亿、9.75 亿和 9.34 亿； 在线办公、在线医疗用户规模分别达 4.69 亿和 2.98 亿，同比分别增长 35.7%和 38.7% [4]。

因网络视频 APP 用户增长极快，社交平台迅速成为大众交流的重要工具和公众获取各种信息的重要渠道。特别是新冠疫情下，大众主要通过线上学习、交流信息。为适应这一变化，图书馆应积极制作短视频，帮助医学生学习如何利用图书馆的各种资源。

微信是当下主流社交媒体平台，也是图书馆主要的资源展示平台和服务宣传窗口[5]，在高校医学生群体中相当普及，不管在哪里，医学生都能够通过它直接获取信息。利用微信平台，高校图书馆可以主动把关于信息素养的知识传达给医学生，利用语音、文字、图片、视频等多种方式在线解决医学生在信息查找方面碰到的问题。通过微信平台渗透信息检索知识，让信息素养教育自然而然地融入到医学生的日常生活中，是普及信息素养教育的有效途径。

3.4 针对医学生不同阶段进行信息素养教育

新生对图书馆功能的认识和利用概念模糊，要重点加强对新生的入馆教育。例如：举办“图书馆基础知识与利用”、 “馆藏资源及查询”、“工具书使用”、“电子资源的检索与利用”等专题讲座。使新生基本了解图书馆馆藏资源和服务，掌握中文资源的查找方法，让新生具备能够基本利用图书馆各项资源的能力。大二、大三的医学生对文献资源有一定的需求，可进一步培养他们的信息检索能力，了解跨库检索原理与方法的同时，熟练掌握网络检索的技巧和方法，并能够利用图书馆、网络各种资源进行文献综述。大四、大五的医学生要开始准备完成期刊论文，在信息检索的基础上，培养他们的信息分析能力，加深对检索基础知识及外文摘要数据库检索方法的了解；同时，掌握一定的信息分析基础知识和方法，能够在获取信息的基础上进行文献分析、评价和利用。针对这个阶段的医学生设置“课题文献调研”、“信息伦理、知识产权与学术规范”等讲座，让医学生了解文献调研、信息伦理、知识产权相关知识，从而能够更好地独立完成期刊论文的文献调研，掌握和遵守撰写期刊论文的学术规范。研究生阶段的学习以做研究或做项目为主，要完成高质量的硕士论文或博士论文，自发性学习、碰到问题自己可以去查询相关资料的能力非常重要。这一时期的信息素养教育内容不仅要全面，而且要与专业相结合，要有一定的深度，可介绍一些专业外文数据库的使用。

3.5 提高高校图书馆员的信息素养

高校图书馆要成为大学生信息素养教育的重要场所，就必须不断提高图书馆员自身的信息素养。图书馆员的工作涉及到多个学科、专业，既要熟悉图书馆各项业务和馆藏资源情况，掌握各种检索工具的使用方法，还应加强信息素养的培训。信息素养与许多新技术紧密相连，各种新名词、新技术不断涌现，高校图书馆员对信息环境、读者需求要有敏锐的洞察力，对新技术、新内容要有善于自我学习的能力，与时俱进不断地完善自己，提高自己的信息储备与信息能力，才能提供真正高质量的信息素养教育内容。

综上所述，医学信息素养教育是医学学综合教育的基础，高校图书馆既是学校教学、科研的重要组成部分，也是信息资源中心。图书馆应充分利用微博、微信、网络短视频等途径开展健康信息素养教育,提升医学生的健康信息素养水平,遇到突发公共卫生事件时,医学生能够迅速理解、甄别信息真假,快速检索并获取健康信息,合理运用健康信息,并做出正确有效的健康决策[6]。

PU-157

数智时代医学生信息素养教育的内涵与培养路径初探

刘靓靓，聂小莺，胡波
国家儿童医学中心首都医科大学附属北京儿童医院

1 研究背景

信息素养（Information Literacy）又称信息素质，是个人进入社会进行终身学习的必备能力，也是众多其他相关素养形成的基础。大数据、云计算等新兴技术在健康医疗领域的广泛应用，以及包括科学文献在内的医学信息资源极速增长，对医学生的信息获取、信息利用和信息评价等能力提出了更高要求。信息素养教育是培养医学生信息获取与利用能力的重要途径，对提高医学生学习能力、创新能力、科研能力以及未来从业过程中解决问题的能力有深远影响。经过30多年的建设和发展，我国医学院校图书馆信息素养教育已取得一定成绩，但仍然存在医学生信息素养教育内容偏移、医学生信息素养教学方式单一等问题。随着健康医疗信息化建设的快速发展、医疗卫生行业职业岗位需求的变化，以及慕课、翻转课堂等新型教育模式的兴起，传统的医学院校图书馆信息素养教育面临极大挑战，需进一步总结新环境下医学生对信息素养教育的需求特点，在分析传统医学院校图书馆信息素养教育存在问题的基础上，探讨医学院校图书馆信息素养教育的发展方向。

1.1 新环境下医学生的信息获取与利用渠道更加多元

Web2.0时代，个人既是信息的接收者，也是信息的生产者。手机、自媒体等信息传播新兴渠道改变了人们的信息行为。医学生的信息获取渠道从传统的图书馆文献信息获取扩展为更加丰富的方式，包括网络搜索引擎、学术社区、门户网站、新媒体、手机APP等渠道成为了信息获取新的来源，对于医学图书馆的利用率造成极大冲击。医学生信息利用的方式也更加多元，除通过图书馆服务以外，还可以借助搜索引擎、学术社区、商业学术网站、自媒体等多种方式学习信息分析与利用知识、掌握信息处理工具，医学院校图书馆不再是医学生信息素养教育的单一主体。

1.2 新技术丰富健康信息素养教学方法

网络技术使得信息素养教学突破时间与空间的限制，实现了线上方式教学。各信息素养教育平台的模块功能、实现技术更加完备，互动性极大提升。加上教学思路不断优化和教学内容不断改革，促成Blackboard教学平台、MOOC、翻转课堂等新型教学方法出现，丰富了信息素养教学方法。

1.3 信息素养教育被赋予个人和社会发展的深层意义

信息社会背景与全球化视野下，人们对于信息素养重要性认识不断加深，意识到信息素养不仅关系到个人是否能够完成一次信息获取和利用过程，而是关乎终身教育与长期发展，能够提升个人对于社会的融入程度与自信心，帮助个人实现社会目标与社会价值。从这一层面来讲，信息素养教育不仅对于个人知识技能转换能力的培养与长期发展有着深远影响，甚至成为了影响组织机构、地区和国家竞争优势的重要因素。

2 医学生信息素养教育的内涵

2.1 信息素养概念梳理

信息素养概念最早由美国信息产业协会主席 Paul Zurkowski 于 1974 年提出，定义为“利用大量的信息工具及主要信息资源使问题得到解答的技术和技能”。早期较为经典的定义有 1989 年美国图书馆协会（American Library Association，ALA）和 2000 年美国大学与图书馆研究协会（Association of College & Research Libraries，ACRL）对于信息素养的定义，主要将信息素养聚焦于个人的信息检索、利用与评价能力。

随着网络技术的发展和社会信息环境的变化，人们的信息行为也发生了改变，信息获取渠道更加多元，个人不仅是信息的接收者，还可以是信息生产者。随之信息素养概念的边界更加模糊，开放度提升，与其他素养概念融合、衍生，联合国教科文组织（UNESCO）发布的《全球媒体与信息素养评估框架》（Media and Information Literacy Assessment Framework）就将媒体素养、信息素养和数字素养等概念融为一

体，定义为“媒体与信息素养”（Media and Information Literacy，MIL）。信息素养的内涵更加丰富。2015 年，美国大学与图书馆研究协会（ACRL）在《高等教育信息素养能力标准》的基础上修订形成了《高等教育信息素养框架》。《框架》对信息素养的定义是“包括对信息的反思性发现，对信息如何产生和评价的理解，以及利用信息创造新知识并合理参与学习团体的一组综合能力”。

国内对信息素养内涵的划分主要有两种方法，第一种依据信息素养的过程结构，将信息素养内涵分为运用信息工具、获取信息、处理信息、表达信息、使用信息和信息交流等能力；近年来对信息评价、信息创新等维度的研究增多。第二种是从解决信息问题所需心理技能的视角，即根据目标结构理论，将信息素养内涵分为信息意识、信息道德、信息知识、信息能力等若干要素。从数量上看，“要素说”主要有一要素说、二要素说、三要素说、四要素说、五要素说、六要素说等，发展上基本是对各要素进行拆分、重组和补充，因此从要素内容的角度进行介绍。同时，信息素养也是一个动态变化着的概念，在数智时代衍生出了更为丰富的素养概念，包括数据素养、元素养、健康信息素养等。

2.2 医学生信息素养教育内涵

2.2.1 信息素养通识教育

通识教育（General Education）一词来源于古希腊的“自由教育”，19 世纪 20 年代末被美国学者帕卡德(A.S.Parkard)首次使用到大学教育的情境中，这种教育理念目前已被普遍接受和应用。通识教育蕴含“全人”教育理念，从信息素养教育的角度来看，医学生不仅需要接受健康信息素养教育，还需要掌握最基本的信息素养，具备获取学习、工作、生活和个人健康相关信息的能力，因此需要接受信息素养通识教育。

信息素养通识教育是不区分特定人群的一般性、基础性信息素养教育，主要以培养个人掌握基本的信息意识、信息道德、信息知识、信息能力为目标，以信息素养通识课程和活动为渠道。医学生信息素养通识教育主要是通过高校公共选修课程和高校图书馆专题培训来实现，内容主要包括常规信息资源的检索、利用、分析、评价培训，新生入馆教育，计算机技能培训，文献管理软件培训等。通过开展信息素养通识教育，医学生具备了基本的信息素养，也为开展医学相关的信息素养教育奠定知识与能力基础。

2.2.2 医学文献检索与利用教育

与一般性的文献检索课程不同，医学文献检索课程的重要目标是为医学生的实验与临床检索需求提供指导与支持，进而提升他们的医学研究能力与临床思维，因此医学文献检索课程设计和实验课选题具有以下 4 个特点：

（1）与医学生早期科研紧密结合，为医学生论文选题与写作提供实用帮助。医学文献检索的课程或者实验作业要求医学生开展医学领域的课题查新、追踪领域研究热点及前沿，为医学生的研究选题检索策略提供指导。例如，授课教师会提出文献检索的选题围绕某个医学研究领域，医学生根据限定范围，选择适当的检索策略，探索领域内的经典文献与热点文献。在医学生通过查新选择到较为新颖的论文题目之后，结合文献管理软件教学、英文写作教学，对医学文献的管理与利用策略进行指导。

（2）与专类文献信息检索紧密结合。包括对医学会议信息、专利文献、药物信息、生物信息等专类文献信息的检索教学。这类专题文献信息具有先导性、动态性和时效性的特点，由于传统的论文和专利发表具有一定时滞期，因此医学图书馆会在讲授文献信息检索的一般性方法之外，从基金资助信息、智库报告、数据库工具等角度为医学生介绍发现领域前沿的办法。

（3）在文献信息检索语言教学方面，引入对医学相关的词表、分类法、术语集的介绍，包括国际疾病分类法、国际系统医学术语集、医学主题词表（Mesh 和 CMesh）等。

（4）在数据库检索与利用教学方面，引入对医学相关数据库和网络检索工具的教学培训。在使用医学信息资源、查阅医学文献时，离不开对医学相关数据库的使用，这些数据库在资源特点、检索功能、检索技巧等方面都与一般性数据库有很多不同，医学文献检索课通常会专题讲解这些医学数据库的检索和使用。

2.2.3 医学信息分析与评价教育

面对海量的医学文献资源，医学文献信息分析与评价能力能够指导医学生遴选优质信息资源，为科研与临床工作提供文献支持，是医学生信息素养教育的重要内容之一。根据医学生的医学信息分析与评价具体应用场景，医学院校图书馆通常从以下几个角度开展信息素养教育。

（1） 医学期刊投稿策略指导

针对医学生论文发表的普遍需求，医学院校图书馆一般会与期刊出版商、期刊编辑和医学专家展开合作，对医学期刊投稿策略进行指导，内容包括：①期刊选择。从科学引文索引（Scientific Citation Index，SCI）、Medline、中国科学引文数据库（Chinese Science Citation Database，CSCD）、北京大学《中文核心期刊要目总览》、中国科技核心期刊等多种期刊评价系统来了解医学核心期刊的评价标准，帮助医学生更高效地选择合适期刊进行投稿，提高论文的录用率；②期刊投稿流程介绍。包括对期刊投稿流程的具体介绍，Cover letter 等材料写作、与审稿人书面交流等方面的注意事项；③出版伦理。包括厘清作者权益与版权事项、科研诚信培训等。

（2） 医学文献信息分析与评价

对医学文献信息的科学分析评价有助于医学生快速、准确识别领域内的经典文献、热点文献和前沿文献，做好研究工作的文献基础。常用的医学文献信息分析与评价教育内容包括：①批判性阅读（Critical Reading），从研究开始时干预组和对照组的预后是否相同、研究过程中是否保持预后平衡、研究完成时各组在预后上是否平衡等维度来评价文献中研究结果的可靠性；②系统评价（Systematic Reviews）。针对某一具体临床问题，系统全面地收集临床研究文献，用统一的科学评价标准筛选出符合质量标准的文献，用统计方法进行综合（Meta 分析），得到定量结果，以最新的文献信息和决策依据改进临床实践；③基于内容的评价方法。经过自评或他评（如 F1000 等第三方文献评价），从选题新颖性、研究设计、研究论证、研究结果、表达规范性、结构逻辑等维度对文献信息进行评价；④基于外部特征的评价方法。参考依据包括文献的被引次数、文献所在期刊影响因子与期刊评价系统等指标。

2.2.4 信息道德教育

信息道德指个体在信息活动中应遵守的道德规范,它是调节信息生产者、加工者、传递者及使用者之间相互关系的行为规范的总和。信息道德不仅是个人信息活动遵循法律法规和公共秩序的内在要求，也是医学生在医学研究和临床中遵守医学伦理的必备素养。医学生在入学伊始就会接受信息道德教育，教育内容包括信息鉴别、知识产权、信息获取与引用、科研数据使用与共享的伦理准则、数据安全等。

通过信息道德价值揭示、价值纠偏、价值引导和价值增强，医学生能够提高信息鉴别意识，树立知识产权意识和正确的科研理念,科研行为规范化，从而促进医学科研和临床实践的良性发展。

2.2.5 循证医学教育

在医学领域，文献检索的目的一般包括开展科学研究和解决临床问题两种情况，前者需要广泛收集信息，侧重查全率，后者需要获取解决临床问题的最佳证据，侧重查准率，即循证医学范畴。循证医学(Evidence-Based Medicine，EBM) 指的是临床实践需结合临床医生个人经验、患者意愿和来自系统化评价和合成的研究证据，高质量的循证医学决策与文献检索、文献评价的结果密切相关，其主要实现步骤是：①确定临床实践问题；②检索相关文献；③严格评价文献④应用最佳证据，指导临床决策；⑤通过实践，提高临床学术水平和医疗质量。因此，循证医学教育也一般与医学文献检索与评价教学相结合。

医学文献检索课通过引入循证医学的思想, 引导医学生使用高质量的临床实践指南、系统评价和 Meta 分析，通过不同等级证据解决临床问题，提高医学生的循证决策能力和创新能力。

2.2.6 数据素养教育

医学生在科研与临床上接触大量的医学数据，对这些数据的采集、管理、分析与共享是有效开展相关研究或实践的前提，也是医学院校图书馆信息素养教育的重要内容。医学

生数据素养教育的内容主要包括数据意识、数据技能和数据伦理。数据意识指的是医学生对医学数据的感知、判断力，即知道自己何时需要数据、如何获取数据；数据技能指的是医学生获取、使用数据的能力，即知道自己如何利用数据；数据伦理是医学生使用医学数据的信息道德意识，即知道自己如何恰当地使用数据。

医学院校图书馆对医学生的数据素养教育通常以结合公共选修课、图书馆专题培训的形式，辅以数据获取、数据管理、数据分析相关工具的使用培训，从知识层面和能力层面提升医学生的数据素养。

3 医学生信息素养教育案例调查与分析

医学院校图书馆是医学生信息素养教育的前沿阵地。考虑到美国是较早开展信息素养教育的国家，积累了丰富的信息素养教育经验，因此选取了医学院排名靠前的医学院（校）图书馆作为调查对象，排除网页无法打开的调查对象，最终确定 9 所医学院（校）图书馆作为调查对象（见表 3.1）。在调查方法上，本文主要采用网站调查法与内容分析法，通过浏览图书馆网站各版块内容，提取、总结、分析与信息素养教育主题相关的内容。

表 3.1 美国医学院（校）图书馆信息素养教育调查对象名单

哈佛大学图书馆	https://library.harvard.edu/
约翰霍普金斯大学图书馆	https://www.library.jhu.edu/
纽约大学图书馆	http://library.nyu.edu/
斯坦福大学图书馆	https://library.stanford.edu/
加州大学旧金山分校图书馆	https://www.library.ucsf.edu/
宾夕法尼亚大学图书馆	https://www.library.upenn.edu/
加州大学洛杉矶分校图书馆	https://www.library.ucla.edu/
圣路易斯华盛顿大学图书馆	https://library.wustl.edu/
杜克大学图书馆	https://library.duke.edu/

3.1 信息素养教育资源网页组织方式

根据调查，只有圣路易斯华盛顿大学图书馆网站拥有信息素养教育二级专题页面（The Instruction and Information Literacy Program），其余 8 所医学院（校）图书馆都采用了分散式的信息素养教育网页信息组织方式，即图书馆网页没有以信息素养教育命名的专题栏目，而是分散在网页各个版块，将信息素养教育融入了图书馆资源与服务中。例如，哈佛大学图书馆网页的信息素养教育资源分布在用户指南（How to）、分馆链接（Library）、服务（Service）、工具（Tool）四个一级版块，内容分别是：①用户指南版块以简洁浅显的语言指导用户如何利用图书馆的资源与服务；②分馆链接版块以各个分馆链接为文本组织单元，信息素养教育内容蕴含在分馆服务中。例如，点击医学分馆（Francis A.Countway Library of Medicine）网站的链接，在分馆服务版块中可以看到分馆培训的相关信息，包括培训日期、培训内容和培训形式；③服务版块汇聚了哈佛大学图书馆提供的所有服务内容，如面向教职工的科研协助（Research Help for Harvard Offices）、研究咨询（Research Consultations）等；④工具版块发布了哈佛大学图书馆面向用户提供的相关工具，如用来创建和共享数据管理计划的在线工具 DMPTool，对工具功能、使用方法进行详细说明。除此之外，调查对象的信息素养教育网页文本内容组织方式还具有以下几种特征：

（1）根据研究主题组织相关信息

被调查的 9 所医学院（校）图书馆网站全部设有研究主题（Subject）索引，将图书馆

的信息素养教育资源依据主题进行分类组织。研究主题索引通常位于研究指南（Research Guide）目录下，按照字母顺序排列，主题内容包括学科专业、学习与科研周期的信息需求。以圣路易斯华盛顿大学图书馆网站的研究主题索引为例，共包含 83 个研究主题，覆盖人文社科、管理学、理学、工学、医学等领域的 71 个专业学科，以及论文写作、版权与出版、数据管理等 12 个学习与科研主题。每个研究主题含有二级资源链接、资源描述、主题标引馆员信息、标引日期、浏览量、标签等详细信息。

（2）根据用户类型细分信息层次

一些被调查图书馆还根据用户类型，细分信息素养教育网页信息的层次深度。例如，加州大学洛杉矶分校图书馆首页有研究指南（Research Guide）和研究与教学支持（Research &Teaching Support）版块，两者都含有信息素养教育相关内容，但是在文字表达、内容深度上都有所不同。研究指南版块以一站式检索和初级指南（Get Started:Information Research Tips）为主。用户在一站式检索框输入资源或服务的关键词，快速定位到相关指导。搜索框旁边列有搜索提示，引导用户使用简单词汇进行检索。初级指南以简单句介绍电子资源、纸本资源的检索方法、如何缩小检索范围等基本知识，满足图书馆新用户的信息需要。研究与教学支持版块则针对教职工和学生的不同信息需求，分别列举相应的图书馆服务内容。这样差异化的文本层次有助于不同信息素养水平的用户快速找到匹配自身信息能力与信息需求的资源和服务。

（3）根据总-分馆模式组织相关信息

被调查的国外医学院（校）图书馆都设有分馆，在信息素养教育资源的内容组织上也使用了总-分馆模式，具体可以划分为两种形式。第一种是总馆和分馆都具有独立页面。点击分馆网页链接可以查看该馆的资源与服务，能够充分揭示分馆的线下资源和特色资源，有利于用户针对性掌握图书馆讲座与培训信息。例如哈佛大学图书馆、约翰霍普金斯大学图书馆、杜克大学图书馆都设有医学分馆，医学生可以专门访问分馆网站，直接获取医学领域的信息素养教育资源。第二种是总馆页面下设分馆资源的指向链接。例如哈佛大学图书馆和纽约大学图书馆网站的研究指南版块可选择根据分馆进行分面查看，分馆资源与服务指导按照字母顺序排列，方便用户快速定位，见表 3.2。

表 3.2 国外调查对象信息素养教育资源网页组织方式

机构名称	链接目录层次	所属版块
哈佛大学图书馆	2 级	用户指南（How to）、分馆链接（Library）、服务（Service）、工具（Tool）
约翰霍普金斯大学图书馆	2 级	帮助（Help&Support）、服务（Services）
纽约大学图书馆	2 级	分馆链接（Locations）、服务（Service）、图书馆活动（Library Workshops & Events）、研究指南（Research Guides）
斯坦福大学图书馆	2 级	关于图书馆（About）、分馆链接（Libraries）、利用图书馆（Using the Libraries）、研究支持（Research Support）
加州大学旧金山分校图书馆	2 级	资源发现（Find Materials）、利用图书馆（Use the Library）、咨询服务（Ask an Expert）
宾夕法尼亚大学图书	2 级	帮助（Help with）、分馆链接

馆		（Libraries）、图书馆新闻与活动（News&Events）
加州大学洛杉矶分校图书馆	2 级	研究指南（Research Guide）、研究与教学支持（Research &Teaching Support）、分馆链接（Locations）、图书馆新闻与活动（News&Events）、利用图书馆（Using the Libraries）
圣路易斯华盛顿大学图书馆	2 级	分馆链接（Libraries）、研究与指导（Research &Instruction）、服务（Service）
杜克大学图书馆	2 级	利用图书馆（Use the Library）、研究支持（Research Support）、课程支持（Course Support）、分馆链接（Libraries）

3.2 信息素养教育内容

3.2.1 健康信息资源整合与利用指导

如表 3.3 所示，被调查的国外医学院（校）图书馆均提供了健康信息资源整合与利用指导，用户可以在网站的医学学科资源导航、资源常见问题汇总的位置获取健康信息资源及利用方法。每个医学院（校）图书馆的医学学科资源导航分面维度不同，例如约翰霍普金斯大学图书馆在研究指南中列出了医学、药学和生命科学研究学科相关的馆藏文献资源，每个学科下设二级学科专题资源页面，提供在线健康信息资源导航、相关医学数据库导航、药物指南、公众健康信息、论文撰写和馆藏指引等分面导航条目；加州大学旧金山分校图书馆的医学学科资源导航页面包含了获取印本资源、学科研究指导、远程资源访问、论文撰写等分面导航条目。杜克大学图书馆导航页面则是从资源入门、文献获取、数据获取、图书获取、查找国家研究背景、引用与研究工具、二级与三级学科分类等维度进行医学学科资源组织的。

表 3.3 国外调查对象网站健康信息资源整合与利用指导

机构名称	医学学科资源导航	常见问题汇总
哈佛大学图书馆	生物与化学科学，生物科学，生物学，人口统计学，营养学，健康学，医学科学，药学、口腔医学与公共卫生，心理学，公共卫生，	有
约翰霍普金斯大学图书馆	生物科学，医学与预防保健，心理学	有
纽约大学图书馆	生物学，健康、护理、保健与医学，神经科学，营养与食品健康，公众健康	有
斯坦福大学图书馆	细胞生物学，发展生物学，心理学，生物学（一般问题），实验室安全，药物和生物化学，微生物学，生物学移动应用，医疗器械设计与规范，神经生物学，神经科学	无
加州大学旧金山分校图书馆	基础科学，生物信息学和统计学，临床医学，公众健康信息素养，口腔医学，循证医学，公共卫生，医学史，LGBT 人群	有

	健康，医学（一般问题），护理学，药学，物理疗法，医学行政管理，系统评价	
宾夕法尼亚大学图书馆	生物学，生物医学，临床医学，口腔医学，健康与生命科学，健康与社会，卫生保健与管理，护理学，心理学，人口研究与人口统计，退伍军人用药	无
加州大学洛杉矶分校图书馆	生物学与生命科学，健康科学，心理学	无
圣路易斯华盛顿大学图书馆	生物学，医学史，生命科学，医学，哲学、神经科学与心理学，公共卫生，	有
杜克大学图书馆	健康与医学科学，	有

3.2.2 健康信息素养培训

经过调查，这 9 所国外医学院（校）图书馆的健康信息素养培训方式主要有新生教育和参考咨询。新生教育方面，哈佛大学图书馆设有新生馆员（Librarians for First-Year Students）岗位，馆员与新生结对，负责新生入学第一学年的图书馆利用指导、论文指导和研究指导。圣路易斯华盛顿大学图书馆在其实体馆内区域开辟了新生探索中心（Newman Exploration Center），供集中开展新生教育活动。

读者培训规模一般有 Workshops、Events、Classes 三类，主要区别在于：①Workshops 通常聚焦于某一培训主题，举办时间、次数比较随机，具有讲座的特征，如约翰霍普金斯大学图书馆的 Workshops 时长从 30 分钟-2 小时不等，一些专题会举办连续性讲座；②Events 和 Classes 更倾向于通常意义上的课程，授课时间更加固定，一些课程还是某些专业学生的必修课。读者培训内容主要包括了医学文献与医学数据库检索培训、医学论文写作与研究指导、医学数据管理培训、工具软件培训、信息道德与知识产权相关培训、教师信息素养培训和信息素养专题培训。

（1）医学文献与医学数据库检索培训

数据库培训方面，依据数据库类型，可以将培训内容分为医学文献资源数据库培训、循证医学数据库培训、临床信息资源数据库培训、药学资源数据库培训。例如，宾夕法尼亚大学图书馆提供生物医学学科的一对一的面对面指和网络研讨会，培训内容包括医学文献数据库 PubMed、临床信息资源数据库 BLAST、ClinVar、MedGen 等，用户通过培训后可以掌握文献信息的检索策略，使用和分析核苷酸、蛋白质序列等生物学数据。圣路易斯华盛顿大学图书馆为全校学生和教职工提供培训服务，特色之处是对美国国立卫生研究院（National Institutes of Health，NIH）的一系列医学信息产品（如 PubMed， My Bibliography，NIH Biosketch，NCBI SciENcv 及 NIH 产品的开放获取政策）开展定制化培训服务，用户填写培训申请表，即可以小组或班级形式接受医学文献与医学数据库检索定制培训。

（2）医学论文写作与研究指导

被调查的图书馆都开展了医学写作与研究指导。如斯坦福大学图书馆开设了科研专题培训班（Gear Up for Research），旨在为斯坦福大学的研究生、博士后、教职员工和研究人员增强研究经验。课程每年三次，冬季和春季课程侧重于提升科学和工程研究人员的技能，而秋季课程侧重于社会科学和人文科学中的数据管理。哈佛大学图书馆设有“朋辈研究员”（Peer Research Fellows）项目。该项目由本科生担任研究指导员，解答学生在论文写作与科研方面的问题，还可以把问题转交给专业学科馆员。纽约大学图书馆开设参考文献格式指导专题页面，整合汇聚了 AMA、APA、MLA 等各类参考文献格式的标注方法，配有参考书目、工具和相关网页链接。此外，斯坦福大学图书馆提供研究提醒（Research Alerts）插件，用户可以设置检索条件与数据库范围，当符合提醒条件的新文献出现时，会自动向用户发送电子邮件或更新的 RSS 提要，使用户及时了解相关研究进

展。加州大学洛杉矶分校图书馆开发了一款研究论文计划助手（Research Paper Planner），用户可以创建或者复制写作计划，工具会将写作任务分解为可管理的阶段性内容，可以随时跟踪进度并关联使用图书馆资源。杜克大学图书馆网页在列举常见参考文献引用格式的同时，还列出了杜克大学课程论文与毕业论文的参考文献要求与指南。

除了直接提供论文写作与研究指导，一些图书馆也从辩证角度指出用户需要在提出图书馆服务与指导的需求前，首先充分评估研究中的问题，以及思考图书馆指导是否适用于研究与写作的当前环节。哈佛大学图书馆针对用户文献检索结果不理想的情况，开设了检索问题诊断页面。该页面引导用户寻找符合自身检索问题的描述（检索结果过多、过少、结果不相关），查看原因分析与解决建议，参考提供的检索案例重新制定检索策略。系统综述是医学研究与论文撰写的重要方法之一，加州大学旧金山分校图书馆设有系统综述专题页面。值得注意的是，该页面没有直接介绍系统综述的写作方法，而是首先指出系统综述并非适合所有研究和所有环节，需要经过充分评估。许多用户不了解系统评估，往往在研究过程中才提交系统评估指导需求，例如咨询某个检索词是否需要调整，浪费了大量的时间和精力。因此，图书馆目前不接受研究过程中提交的系统综述指导申请，而是要求用户在研究开始前邀请图书馆员参与其中。图书馆员在系统综述中的角色包括：参与选题、根据 PICO 原则设计方案、制定检索策略、筛选文献、文献管理软件使用帮助、指导研究方法章节撰写等。该专题网页还详细阐述了系统综述的定义、适用范围、研究方法和参考资料，帮助用户加深对系统综述的了解，更明智地选择科研与写作方法。

（3）医学数据管理培训

被调查的所有医学院（校）图书馆全部设有医学数据管理的专题页面，页面内容包括：数据管理的内涵，如何进行数据管理，包括数据获取、数据管理、数据存储、数据可视化、数据出版、数据存档等方面，以及图书馆在此方面的培训与帮助。所有图书馆都会开展的一项工作是对于撰写数据管理计划（Data Managing Plan）和使用撰写工具的指导。在美国，申请大多数基金项目都需要提交一份数据管理计划，内容包括制定有竞争力的研究计划，能够对研究数据管理和基金管理方要求提出有效解决方案，常用的数据管理计划撰写工具是 DMPTool。哈佛大学图书馆 DMPS 样例（Sample DMPS）服务含有哈佛 Dataverse 样本数据管理计划、DMPTool 提供的公共数据管理计划以及收录 NIH 等常用机构的数据共享计划示例。约翰霍普金斯设有 8 个模块的线上培训，对查询、撰写和复制 DMP、下载、共享、审核 DMP 进行了指导。宾夕法尼亚大学图书馆对 DMP Tools 的数据安全问题进行了详细解答，包括使用工具创建数据管理计划的版权归属、使用政策、隐私政策等。

除了提供常规培训，许多图书馆也通过项目（projects）形式探索数据管理培训的新方向。斯坦福大学图书馆的 Bibliopedia 是一个可视化的文献管理平台，支持文献资源的上传、检索、共享和可视化展示。可视化网络揭示了文献之间的关联，帮助用户了解研究发展历程、研究合作等。该平台是一个开源平台，可供所有研究人员、图书馆和其他机构使用。通过使共享、分析和保存研究成果变得更加容易来增强数字学术可见度。图书馆的新兴技术团队（the Emerging Technologies Team ,ETT）由图书馆管理人员和技术部门人员组成，负责识别、测试、评估、推广新技术在图书馆内的使用，为图书馆提供新技术咨询和评估调研，他们撰写了 Stanford Box 等工具软件的使用测评，发布信息技术相关的博客、期刊资源，帮助用户更好地掌握工具软件的使用。R 语言培训是斯坦福大学图书馆面向全校师生的培训项目，图书馆拥有 R 语言实验室（R Open Lab）供用户咨询、培训。开设讲习班，咨询，线上教程包。宾夕法尼亚大学图书馆每年举办数字技能研讨会（Digital skills workshops），围绕信息素养主题开展数字出版、数字研究方法、沉浸式技术以及工具软件的研讨和实际操作。

（4）工具软件培训

调查发现，按照工具软件类型分类，培训内容可以分为参考文献管理软件培训、论文写作工具培训、文件与数据管理工具培训、数据分析与统计工具培训、多媒体工具培训和其他工具软件培训。为了帮助用户更好地依据使用场景判断所需工具，约翰霍普金斯大学

图书馆、纽约大学图书馆、斯坦福大学图书馆都设有文档管理解决方案网页。页面以表格形式体现，将各种用户需求与各类工具软件对应，用图标或文字表示图书馆推荐在此需求场景下使用该款工具软件，用户可以在表格中查找适合自身需求的工具软件。例如，斯坦福大学图书馆网页上，用户需求是与多人在线同时编辑同一个文档，图书馆推荐使用的工具软件是 Google Drive for Stanford、Google Shared drives for Stanford、OneDrive for Business。每个工具软件名称都附有链接，用户可以点击链接进一步学习该工具软件的使用方法。宾夕法尼亚大学图书馆在常规软件工具的培训基础上，还提供了网页制作、数字人文主题相关的工具软件培训。

此外，许多被调查对象结合图书馆空间服务、特色服务和优势学科开展培训。例如哈佛大学图书馆和斯坦福大学图书馆内设有多媒体工作室，为学生提供加州大学旧金山分校图书馆和加州大学洛杉矶分校图书馆拥有实体创客空间和丰富的 3D 打印与创客空间培训，因此在工具软件的使用培训上也突出了对于相关技术的指导。加州大学旧金山分校图书馆开展了 Bioconductor 、Bowtie、Galaxy、GATK、Samtools 等生物信息学软件的培训。

按照工具软件来源分类，被调查图书馆会对所购买、获取的第三方工具软件进行培训，还会对学校及图书馆自行研发的工具软件进行培训。Perma.cc 是哈佛大学图书馆开发的一款针对网页类型参考文献的引用工具，通过网页快照和链接能够防止网页丢失无法引用的情况。WashU ELN 是圣路易斯华盛顿大学图书馆开发的面向实验数据的电子笔记本，是基于云的研究数据协作和管理工具，图书馆用户可以在该平台上组织、存储和共享实验室数据，操作快速简单。

（5）信息道德与知识产权相关培训

被调查的 9 所国外医学院（校）图书馆均开展了信息道德与知识产权内容的培训（见表 3.4）。培训形式分为线上网页资源和线下讲座课程。培训内容方面可分为常规知识普及和专题项目。常规知识普及方面，如约翰霍普金斯大学图书馆介绍了科研诚信概念、开放获取资源概念与使用规范、区分抄袭和剽窃区别。加州大学旧金山分校图书馆对“孤儿作品”（Orphan Works）、文献属于公有领域，无明确知识产权说明（Public Domain）等特殊的知识产权情况进行了阐述。纽约大学图书馆网页重点讲解了文字、图像等不同类型资源的知识产权与许可申请。加州大学洛杉矶分校图书馆从知识产权使用方与版权拥有方的角度分别阐述了相关知识，为论文作者提供维与咨询工作。

专题项目方面，哈佛大学在学校图书馆、档案馆和博物馆都实施了版权第一响应者（Copyright First Responder）项目。项目专员具有版权法律知识专业背景，及时响应用户知识产权相关问题。 斯坦福大学图书馆主办了“数字时代伦理学”（Ethics in the Digital Age Series）线下研讨会项目，会议主题覆盖数字健康、数据科学与信息学、新闻学、社交媒体等，旨在搭建跨学科、跨行业的研讨平台，呼吁人们关注数字时代的学术道德问题。2018 年 7 月和 11 月举办的两次研讨会主题是医学健康信息伦理。与会专家探讨了数据时代个人隐私保护、健康信息共享的潜在风险等议题，形成了《斯坦福大学数字医疗道德指导原则声明》（Stanford Statement of Guiding Principles for Ethics in Digital Health）。圣路易斯华盛顿大学图书馆的版权知识专题（Learning about Copyright）设有免费的知识产权线上课程，每个课程由 3-10 分钟的视频及小测验组成，方便学生和教职员工开展知识产权自测。

表 3.4 国外调查对象信息道德与知识产权培训情况

机构名称	常规项目	专题项目名称	专题项目面向人群
哈佛大学图书馆	有	版权第一响应者（Copyright First Responder）	教职工、学生
约翰霍普金斯大学图书馆	有	无	无

纽约大学图书馆	有	无	无
斯坦福大学图书馆	有	数字时代伦理学（Ethics in the Digital Age Series）	教职工、学生、公众
加州大学旧金山分校图书馆	有	无	无
宾夕法尼亚大学图书馆	有	无	无
加州大学洛杉矶分校图书馆	有	无	无
圣路易斯华盛顿大学图书馆	有	版权知识专题（Learning about Copyright）	教职工、学生、公众
杜克大学图书馆	有	无	无

结果表明，被调查的美国医学院（校）图书馆信息素养教育的培训形式以线上和线下培训相结合为主，在开展信息素养通识教育的基础上，重视培养医学生的“数智时代”信息素养，如全文获取、数据管理、信息道德与知识产权、健康信息素养教育相关的培训，满足医学生学习、工作、科研方面的各类培训需求，围绕生物信息检索、药学检索、医学专利检索等主题开展了信息素养教育培养。这些案例经验为我国医学院校图书馆信息素养教育思路提供借鉴与参考。

对于医学生信息素养教育的研究是当下也是未来长期一段时间内的重点与热点，理论探讨与经验交流有助于提升我国信息素养教育的研究广度与研究深度，增强图情学科在信息素养教育中的话语权与领导力，提升公众对于图书馆信息素养教育的认可、重视与参与。医学院校图书馆应当充分把握这一契机，积极做好信息素养教育各项工作，为提升医学生终身学习发展的公民信息素养、促进实现联合国 2030 可持续发展目标而贡献力量。

PU-158

医工融合的创新创业教育生态研究——以重庆医科大学为例

王佳 赵文龙 贺向前
重庆医科大学医学信息学院

2019 年，教育部全面启动一流本科专业建设“双万计划”，建设新工科、新医科、新农科、新文科示范性本科专业，推动形成高水平人才培养体系，强调基于人工智能等技术，培养“医+X 交叉学科”等复合式创新人才。多学科交叉融合拓宽了医学类、工学类专业的创新创业通道，也对《创新创业》课程提出了更高的要求。

1 创新创业研究现状及问题

在创新创业的全球化进程中，国外学者研究了创新创业教育与国家政策和经济发展的关系以及创新创业教育课程设置、教学模式、评价体系、创新能力等[1-3]。国内学者对创业教育绩效[4]、体制机制[5]、制度保障、课程体系、师资建设、平台搭建、教学改革等方面提出对策建议[6-7]。各大高校纷纷针对创新创业采取了系列促进措施，湖南大学打造研究生创新创业在线课程学习平台；广东工业大学打造“多专业融合”人才的课程体系[8]，浙江大学、湖南大学等高校在协同创新模式引领下，引导师生科技创业[9]。

这些研究极大地丰富了创新创业课程建设的内涵，但是，当前创新创业课程建设中的培养模式、实践平台、教学内容等制约因素依然突出。

1.1 现有的院系培养模式与创新创业的跨学科需求不平衡

医工融合的创新创业活动中，往往需要医学、工学 、管理学、金融学等多学科交叉，多院系共同参与培养，需要师资团队具备较好的专业背景以及与时俱进的学习能力。 然而，在实践中，不同合作培养单位管理难以同质化，教师跨学科背景缺乏，学生涵盖面广，管理难度大，从而难以在学生的创新创业过程中发挥全面的指导作用。

1.2 当前的实践平台与创新创业实践需求不匹配

实践平台对学生的创新创业能力培养至关重要。创新创业实践平台能为学生模拟真实的创业环境，优化和提升实践能力和创新思维，但是大多数创新创业课程的建设都缺乏实践平台的支持，使得创新创业课程停留在理论层面，学生难以深入理解和掌握创业的真正挑战和机遇。

1.3 传统的课程教学内容与创新创业实践应用脱节

创新创业旨在市场、行业及社会中寻找痛点和需求，并结合专业领域建立技术壁垒，创新性地解决痛点。随着云计算、大数据、物联网、人工智能在医疗领域中的需求不断增大，然而目前的《创新创业》课程内容停留在创新创业理念、技巧等层面，与医科、工科等专业课程内容脱节，与学科专业的社会应用脱节。这直接导致学生在开展创新创业实践过程中，远离学科专业特点和现实需求，而盲目地追求社会热点、痛点，难以构建技术壁垒，从而无法在实践中走得更远。

2 重庆医科大学创新创业教育探索

2019 年开始，重庆医科大学开始建设智慧医学学科群。基于医工科融合特色，从专业论证及建构伊始，打破单一学科的培养模式，融合信息学与医学学科、计算机学科特色，面向医疗健康领域中信息的采集与转化、传输与存储、整理与序化以及解决问题，临床医院、信息学院、基础医学院、药学院、公共卫生学院等共同合作开展跨学科人才培养。建设数年来，不断拓展创新创业课程的内外延内容，加强医学与工程的深度融合，强调实践能力以及创新能力的培养，开展了不少有意义的探索。

2.1 建立医工“332”培养体系

率先在信息管理与信息系统、智能医学工程、医学信息工程三个专业中推行“332”培养体系：三个专业在 3 大基础模块（医学、计算机、数学）、3 个专业模块（智能感知、智能分析、智慧决策）中选择 2 个课程模块，领会医学、计算机科学、信息工程和数据科学的基本概念及其应用，实现多学科交叉。

2.2 打造创新创业第二课堂

以医学、人工智能和管理专业核心能力建设为基础，契合社会对医学信息人才需求，开展创新创业课程的外延活动--ICP（Innovation、Cooperation、Practice）特色实践活动，形成系列以创新性、合作性、实践性为核心的特色主题活动，包括 ICP 创新实践活动，主要包括创新实验、创新创业大赛等，提升学生的专业技能和实践能力，培养创新能力。

2.3 建设创新创业实践应用场景

打造了面积达 3500 平米的以“点亮智慧医学”为主题的现代化体验式实验教学中心，开放智慧医院的实践场景，并与企业合作，提供与健康信息领域相关的实际案例和研究；以便学生在学习中联系实际工作场景，更好地理解和掌握实际工作的应用技能。建设多功能的实验室，用于医学、信息和工科领域研究领域的跨学科合作。

2.4 开展创新创业混合式教学

紧密围绕专业发展方向以及就业前景，通过课程内容讲授、课堂互动、无领导小组、创业团队等形式，开展丰富的课堂教学活动；以 51 个知识点为基本单元，建设并开放了职业生涯规划及创新创业及创新创业网络课程，作为课堂学习的有效补充，满足学生充分利用碎片化时间学习的需求，强化学生对于技能的理解和掌握。

3 医工融合教育生态的创新创业教育建设

重庆医科大学智慧医学学科群在创新创业课程的建设探索方面取得了一定的经验，为医工融合的创新创业生态建设提供了建设基础和发展思路。

教育生态为各类教育主体生存发展的环境及其伴随的状态[10]，而“医工融合教育生态”则是共同对工科医科相关专业人才的培养产生推动或制约作用的环境及状态。在医工人才培养过程中，来自个体、校园和社会的各种能量、信息不断发生交换并相互作用。为达到医工融合教育生态平衡的

平衡，发挥创新创业建设的最优化，统筹各种生态因子，建立健康的生态系统势在必行。

3.1 提升师生素养，优化创新创业个体生态因子

优化创新创业个体生态因子是推动创业创新发展的关键任务之一。个体生态因子指的是影响个体创新创业能力的各种要素，包括师生创新创业知识结构体系、资源等。

①完善师生创新创业知识结构体系。建立“课程教学-创新实验-创新创业项目-毕业项目”的一体化教学路径，拓展创新创业深度；引导多学科师生建立创新创业团队，基于不同学科的视觉，共同完成创新创业项目的设计、打造和运营，提升医工交叉的知识和技能。

②拓展师生创新创业资源。开展跨院系、跨学科的创新实验以及创新创业沙龙等活动，引导师生与真实的创业者、投资人对话、交流和合作，获取真实的行业需求以及支持，挖掘和解决医学中的实际问题，培育创新创业项目。

3.2 加强校园建设，提升创新创业校园生态因子

创新创业校园生态因子对于创新创业的落地起着至关重要的作用。校园生态因子包括实践平台和创新创业氛围等。

①加强创新创业实践平台建设。依托众创空间等平台，建立创业孵化基地，为优秀的创业团队提供场地、资金、办公设施等支持以及更系统的创新创业教育和培训，形成市场调研、产品研发、市场营销、财务管理、成果转化等创新创业实践的完整产业链，有效推动推动创新创业实践。

②营造创新创业氛围。开展企业家进校园、创业大赛等活动，分享真实的创业经验和成功故事，激发师生创新创业热情。

3.3 社会协同育人，统筹创新创业社会生态因子

创新创业的实现需要社会生态因子的支持与促进。社会生态因子中包含校外课堂、社会实践等。

①拓展创新创业校外课堂。依托行业的实习基地、就业基地等，积极引导行业参与创新创业校外课堂教学，结合行业发展方向，深化创新创业教学实践内容，促进医工科创新创业教育的深入发展。

②推动行业参与创新创业实践指导。充分利用现有的医院、医疗卫生企业等行业资源，以寒暑假的社会实践等为载体，以创新创业团队双导师（校内导师+行业导师）为基础，充分挖掘行业痛点和需求，形成学校、社会协同育人的培养机制，为学生提供更接地气的创新创业指导。

PU-159

智能化时代背景下医学信息素养教育研究

崔宇鹏 张帆 赵宏亮
内蒙古医科大学

1 引言

智能化技术极大提高了社会生产力，对人类社会生产生活方式和社会结构都产生了深远影响，国内智能医疗以及医学信息技术的快速变革，推动智能化医学健康产业高质量发展。对于医学生群体而言，智能医疗的快速发展加速了学生知识结构的改变，对适应新技术下的信息素养能力提出了更高挑战，已有的信息素养教育已然无法适应这样的技术冲击。因此如何提高医学生信息素养能力，使其了解医学信息新技术新知识，通过强化医学信息素养解决医疗问题和技术难题，成为摆在当前智能化时代背景下信息素养教育的重要课题[1]。

2 医学信息素养教育的研究现状及特点变化

信息素养教育研究一直是国内外图书馆学研究的重点和热点，随着大数据、人工智能等智能化信息技术的不断发展，图书馆已从传统图书馆的 1.0 时代发展到数字图书馆 2.0 再到如今的智慧图书馆 3.0 时代，图书馆的每一次转型升级都带来了信息素养教育与之匹配的深刻变化[2]。以中国知网、万方、维普三大数据库作为样本，信息素养教育研究内容很多，智能化信息素养教育模式也逐渐成为图书馆界研究的方向。笔者以“智慧、智能、信息素养、医学”为主题，先用检索式为(SU %= '医学' OR SU %= '医科') AND SU %= '信息素养'进行检索，共检出结果 998 条，这也印

证了医学信息素养一直是业内研究的重点和热点，但是再以检索式为(SU %= '医学' OR SU %= '医科') AND (SU %= '智慧' OR SU %= '智能') AND SU %= '信息素养' 进行检索，相关结果仅 19 篇文献，最早文献为 2017 年。可见，智能化医学信息素养的研究正处于初级阶段，所检文献研究内容之间的关联度并不高，文献内容多以片面强调图书馆智慧化服务，而对与之相适应的信息素养教育方式改进上报道甚少。鉴于医学信息素养教育对医学生掌握新知识认识新技术的重要性，信息素养专业教师深入研究智能化医学信息素养教育模式，对提升医学生适应智能化医学发展的要求起到关键性作用。因此建设现代智能化大学，厘清智能化背景下医学信息素养教育产生的新变化，向师生提供智能化医学信息素养教育已经成为高校图书馆向前发展的必由之路。

2.1 内涵和外延发生变化

随着智能化的不断发展，信息素养教育有了新的内涵和外延。据大数据研究机构表明：大数据是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产[3]，大数据大量和多样的这些特点，导致人们所具有的信息素养水平的高低决定了大数据处理技术的能力以及其价值所能实现的程度。大数据的来源和承载方式有物联网、云计算、人工智能。因此，在医学信息素养教育课程设计中，对医学课程的数据信息进行深度挖掘整合分析，已经成为智能化时代下信息素养课的重要授课内容。智能化是互联网、大数据、云计算以及物联网技术发展的最终呈现形式，从生产、生活到学习、工作，最新的医学技术及手段纷繁复杂，医学生要具有智能化的信息素养能力去检索、获取和利用，让数据变知识、让知识变智慧，以便更利于医学教育的发展。

2.2 信息素养能力要求的变化

智能化时代之前，信息的产生和传播速度还不是太快，与之相适应的信息素养能力要求也不高，信息素养知识更新换代比较慢，大多数时候使用的课程内容好多年都不用更新也能使用。但是，智能化时代来临后，知识迭代的速率是以前无法想象的，与之相适应的对师生信息素养能力也提出了更高的要求。在原有信息素养能力的基础上，为适应智能化时代医学生学习能力提升应包括数据获取及信息分析、数据价值挖掘及数据安全防护能力。智能化时代信息迭代和传播速度比以往更快，这就要求医学生具有更高的信息素养能力，包括分辨有用信息、干扰信息、有害信息的能力，对有用信息利用智能化技术手段进行加工分类整合并加以综合利用，利用信息并最终应用到医疗活动的实践中。

3 智能化时代下医学信息素养提升路径研究

在智能化时代知识大爆炸的背景下，如何从海量信息中快速高效地获取自己所需信息，是摆在医学师生所面临的共同困境。笔者认为，医学高校信息素养教育可以从打造智能化信息素养教育理念、建设智能化基础设施、强化信息素养实践性三个方面下功夫，以提升医学生在智能化时代背景下的信息素养水平。

3.1 打造智能化信息素养教育理念

理念指引行动，先进的教育理念对一门课的重要性是显而易见的。智能化信息素养教育模式的教育理念应重点考虑以下三点：1）立足国内医学高校培养人才的总要求，将立德树人作为教育理念的第一理念。2）遵循教育学的一般规律，分阶段、分层次、分级别，针对性强地开展多形式、多内容的智能化信息素养教育模式。3）培养具有国际化大视野的信息素养培养人才体系，借鉴国内外先进教育理念的经验。智能化信息素养教育理念指导下的教学方法主要是融合线上线下优质的课程资源，并加入先进的智能化技术，打造成为智能化信息素养教育教学的云课堂，比如翻转课堂，就是这一理念下的成功应用。翻转课堂是利用现代智能化技术设备，比如物联网、虚拟现实等技术，学生不需要通过老师直接的面授，而是自己先把基本的知识通过技术设备自主完成学习后，再带着问题同老师交流，这样更有利于学生接受知识，体验感更强。

3.2 建设智能化基础设施

智能化基础设施是完成智能化信息素养教育的重要物质基础。首先是要建设相应的基础设施，如今，高校的信息化建设正处于方心未艾的发展时期，打造信息化教学环境的条件是具备的。具体需要软硬件的配套资源，硬件资源主要需要更高的网络宽带运行速度，只有带宽足够才能为人工智能及物联网的实现提供硬件保障，VR 这些翻转课堂的技术才能实现。软件主要包括更权威的课程建设平台，以及更加便利流畅的以文献检索和分类查阅的信息资源。在软件资源方面，要结合教育

部一流课程建设需求，加强与国际间其他医学院校的合作与交流，实现优质教学资源的共享。

3.3 强化信息素养实践性

学校先进的智能化和信息化基础设施，图书馆全面和完备的培养方案，都只是医学生信息素养提升的基础条件[4]。主动学习最新的医学领域智能化新技术，并在学校社团及举办的各类活动中尝试应用最新的理论成果，并在医疗机构实习中运用最先进的仪器设备去贯彻最新的治疗理念，都可以达到强化医学信息素养实践的目的。医学生信息素养的培养提升不应局限于通识课程，“三层次信息素养培养”的观点认为，信息素养应该包括通识信息素养、专业信息素养、实践和创新能力的信息素养[5]。其中，实践与创新能力的信息素养要结合具体的实践活动进行。

4 结语

在智能时代背景下，完善和创新高校医学生分级信息素养教育，以达到全面提升医学生信息素养水平的目的，这是高校图书馆步入智能化时代发展下的信息素养教育可持续发展要求[6]。运用先进的智能化技术，构建“智慧化学科馆员”的信息素养教学团队，不断完善智能化信息素养课程新体系，从而不断提升医学生的信息素养能力。

PU-160

大数据视角下国内医学信息人才培养研究热点及需求研究

周跃[1,2,3,4]，徐倩[1,2,3,4]，杨郑子衿[1,2,3,4]，胡超[3]，罗爱静[1,2,3,4*]

1.中南大学湘雅二医院

2.中南大学生命科学学院

3.医学信息研究湖南省普通高等学校重点实验室（中南大学）

4.湖南省心血管智能医疗临床医学研究中心

0 引言

伴随着大数据时代的到来，人们对信息的需求不断增长。大数据和互联网的新兴技术与社会各领域密切相关。为了满足这一时代的迫切需求，国家提出了“互联网+”发展战略。在 11 个专项计划之一的“互联网+健康医疗”推进下，智慧医疗时代已经到来，这也为国内医疗健康产业带来无限商机。智能与医疗健康相结合的智慧应用也在迅速发展,本文从国内医学信息人才培养现状出发，剖析当前高校医学信息学人才培养面临的机遇与挑战，特别关注于“互联网+”以及大数据相关视角对医学信息学的发展困境。同时，本文探讨如何在国内高校医学信息学人才培养中，培养出符合“互联网+”时代需求的高素质医学信息应用人才。

1 数据来源与研究方法

1.1 数据来源

分析一个研究领域的期刊论文是最常用的一种总结学科发展的方法,能够反映出该学科的最新研究成果和发展趋势[2]。为探究医学信息人才培养问题并提出对策，本文选用中国知网期刊数据库（CNKI）为数据源，首先依据医学信息人才培养研究热点，以“医学信息”和“人才培养”为主题词进行限定检索，2022 年 9 月 30 日为检索截止日期，在筛选后，从 359 篇学术期刊中，筛除了会议文献和学位论文，并选出 183 篇文献作为相关度较高的文献。

1.2 研究方法

本文通过可视化分析，从作者合作网络分析与关键词共现网络分析两方面对国内医学信息人才培养相关文献进行了引文网络汇总分析，旨在探究该领域的研究热点和现状。高频关键词的共现有助于提炼医学信息人才培养模式的热点关注，并指出在大数据时代背景下，如何改进医学信息人才培养模式。

2 医学信息人才培养研究热点文献计量分析

2.1 文献年代分布

为直观反映某领域的研究发展水平和发展趋势，通常选择展示分析该领域在一定时间范围内的文献发文数量，并将其与时间变化关联，以此可以在一定范围程度内反映该领域的学术研究发展速

度。对医学信息人才培养文献数量进行时间分布统计（见图 1），图示 2012-2022 年间中国知网筛选后，与医学信息人才培养相关的可视化研究论文的年代分布情况。该时间段内共挖掘到 183 篇研究论文，总参考数 99 篇，总被引次数 723，平均每篇被引次数 3.95。可以看出该领域的研究是在近几年才有所重视和突破，这也与相应的时代背景有关，随着国家政策的助推，教育与科技的发展也促进着该领域学者的发散性思考。

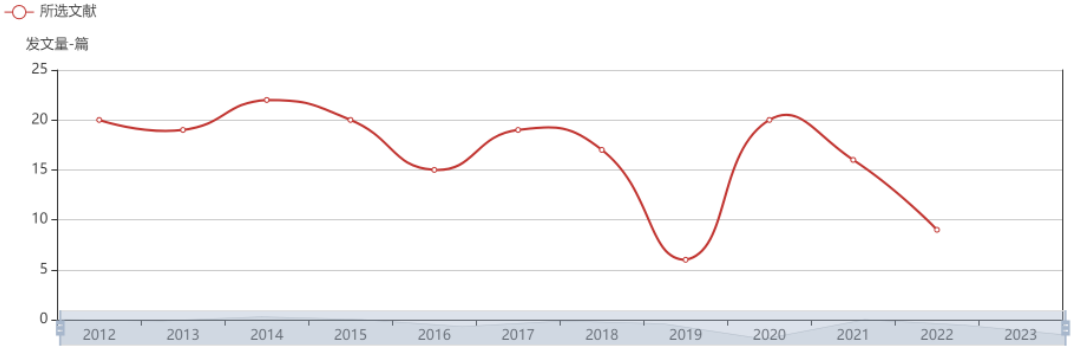


图 1 2012-2022 年医学信息人才培养发文量分析

该领域在近十年间总体的发文量波动较大，自 2012-2017 年有小幅度的波动下降趋势，2017-2019 年的发文量呈现骤降现象。由于疫情爆发，该学科领域在公共卫生突发事件中发挥着积极作用，因此在 2019-2020 年间发文量加速度大幅增长。虽然总体发文量有所下降，但并不代表医学信息学在该领域的研究停滞不前。由于医学信息学是近年来成为专业学科，其教育培养研究发展相对较慢。因此，该领域的人才培养研究正面临着新的突破。尽管该学科领域有专业性限制，但总体发文量相对稳定，迫切需要引入新的理论方法和进行实践研究，以适应新的大数据环境。

2.2 作者合作网络分析

本研究所选的 183 篇文献共有 102 位相关作者，根据其合作关系构建合作网络图，如图 2 所示。其中圆形节点代表作者，节点大小表示合作论文数量，连线表示作者间合作关系。根据图示，作者们呈现出分散的小聚类模式，网络密度则反映出他们之间的联系紧密程度[3]。图中整体网络密度偏低，结构相对松散，多数作者与本单位或所在高校的人合作，这与地理位置及单位影响有关，致使合作伙伴无法增多。吉林大学中，最大的子网络以王伟为核心，该高产作者位于网络中心位置，度数中心较高。该学科领域中还有以成都中医药大学的温川飙、山西医科大学的贺培风、中国医学科学院医学信息研究所的邱五七、中国疾病预防控制中心的马家奇、徐州医科大学附属医院的董瑞国以及湖北中医药大学的沈绍武等学者为中心的学术研究团体。

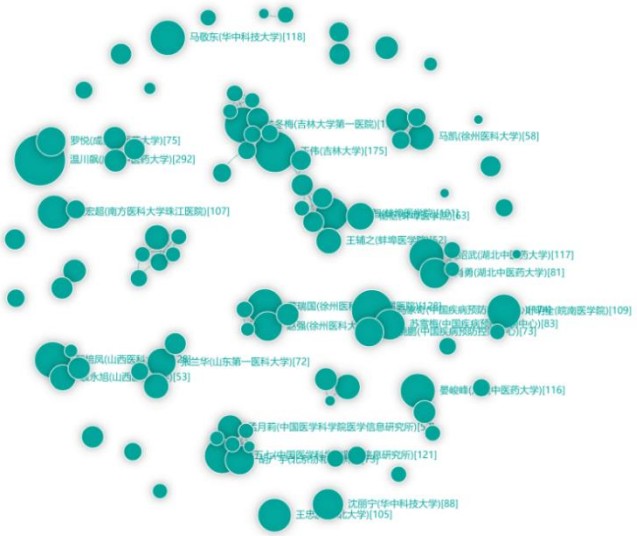
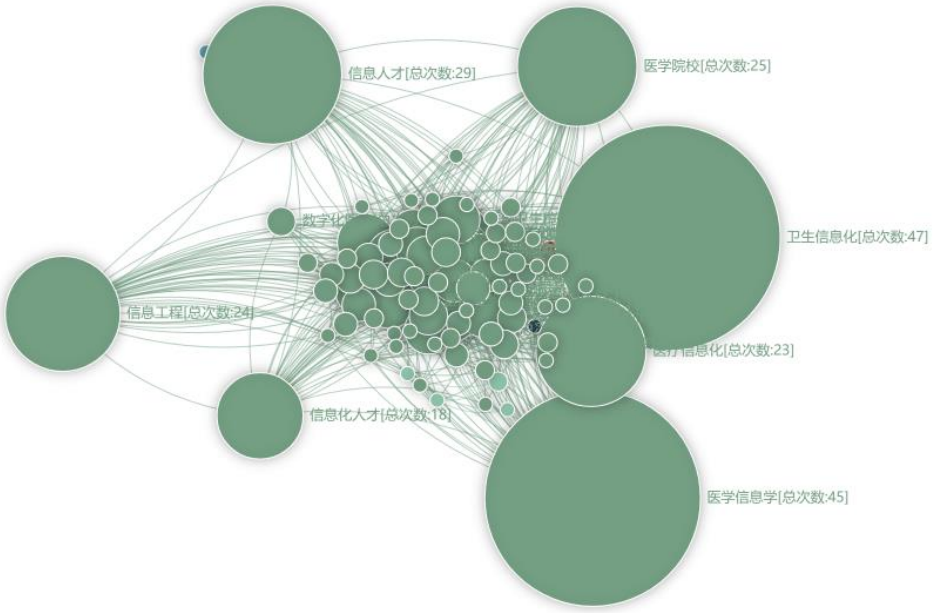


图 3 医学信息人才培养研究作者合作网络图

由图可见，目前该领域的研究存在着合作作者关联较窄，小团体较多且分散的问题，研究该领域的机构单位几乎只存在于各医学院校、医学信息专业的高校院系及医学信息研究所中，最为典型的则为吉林大学的医学信息系。该问题也侧面印证了，医学信息人才培养要想改革培养全面适应新时期的符合性人才还有很长远的路需要走下去，首先应考虑打破目前各院校不交互的困境，并将医学信息人才培育重心放在交叉学科的革新上，尝试选择与计算机相关专业的课程及师资相互合作，从而使得教育的学习成果收益最大化。

2.4 关键词共现网络图分析

研究发现，通过对 183 篇文献进行关键词共现的聚类分析，我们可以清晰地将学术研究的热点领域划分为不同的色块区域，根据图 3 的聚类结果，可以发现热点领域主要聚焦于卫生信息化、医学信息学、信息人才、医学院校、信息工程、医疗信息化、信息化人才等主题研究，同时也关注复合型人才、课程体系、教学方法、教育改革等议题，与大数据直接相关的主体关键词包括信息人才、医疗信息化、教学方法和课程设置，这证实了在大数据时代中，医学信息人才的培养应侧重于实现医疗信息化、探索教学方法和深化改革课程设置。研究者已经认识到在大数据研究中信息素养教育的重要性，要将信息处理挖掘与卫生医疗行业相结合，并实施培养新时代所需复合型人才的新



教育模式。

图 3 医学信息人才培养研究关键词共现网络图

3 大数据环境下国内医学信息人才培养分析

3.1 深化课程体系改革

由已有的 “大数据—医学信息人才培养”关键词共现关联可以看出，目前国内急需改革的应是相应的医学信息课程体系。在现有的医学信息人才培养环节中，课程体系多集中于管理学、信息学和统计学的知识学习。为适应大数据环境下的智慧医疗对医学数据知识发现的需求，还应加强医学信息数据挖掘、数据可视化及数据库应用等大数据时代医疗信息化人才必备的知识与技能，以满足“互联网+”医疗卫生行业在医疗数据分析、处理方面的需求[4]。由于客观环境与主观因素的影响，课程体系改革是一个漫长且复杂的过程。虽然深化课程体系改革从课程设置上无法快速更新普及，但可以从知识更新角度着手，要求任课教师结合大数据时代背景在授课时更新知识体系，将大数据与医学信息融合，为学生提供更全面有效的应用型知识。理论知识更新的同时，为应对大数据背景下的应用技术革新，还应在应用实践上进行课程更新，结合新时期需要，在医疗相关应用数据库、医院信息系统、医院及医用软件、远程医疗系统及临床辅助决策平台等设置相应课程实践学习。深化国内医学信息人才培养课程体系改革，应将理论与实践相结合，从双方面落实，既要保证学生的知

识储备及时更新，又应培养学生掌握实际应用以应对技术的更新。

3.2 培育复合型人才

新时期的医疗信息数据处理人才应与时代所需相适应，由于新时代的“互联网+”模式与多学科领域交叉，因此急需培育具有医学信息专业特色的复合型人才。医学信息学专业应以学科专业交叉为依托，明确专业培养方向，不仅可以保证人才培养的目标明确性，还可以提高医学信息学的专业认可度。由于大数据技术现已与医疗卫生行业紧密相融，因此为提升学生学习的有效性，应将医疗大数据、大数据技术医疗应用与计算机科学专业相结合，为保证医学信息人才的多方面提升，开设跨学科相关专业的交叉课程，研究并构建跨院系、跨学科、跨专业交叉培养医学信息复合型人才的新机制。明确将大数据与医学信息相融合，可以使培养出的人才与医疗卫生行业需求接轨，把握好医疗信息化的应用热点，落实培养满足 “互联网+医疗信息化”需求的人才。

3.3 加强融合医疗大数据构建特色专业

由于医学信息专业与医疗卫生行业密不可分的关联，因此医学信息教育多由医学院校授课，但为顺应时代保证医学信息学及时更新避免淘汰，国内也应尝试加强融合医疗大数据构建医学信息特色专业。回顾已有研究及过去的发展历史可见，国内的医学信息学教育源自过去的医学图书情报学，大部分院校将图书情报作为医学信息学的专业硕士学位培养方向，因此在专业构建上内容多与图书情报专业课程相似，真正结合医学、医疗大数据的专业特色鲜明的人才教育成果并不多见[5]。因此国内的医学信息人才培养应纵向深化培养层次，横向拓展培养范围，从多方面融合医疗大数据的应用教学。利用健康医疗大数据的相关平台，与教学平台有机融合，为科研的反哺作用促进教学。

4 结语

大数据时代为医学信息专业建设提出了新的挑战，也提供了发展机遇，在新时期背景下急需更符合先进医疗技术的复合型人才处理爆炸式的信息数据，但专业教育内容的改革仍是一个缓慢且严谨的过程，不仅需要时间的考验还需要实践的试错。于高校而言，应打破学校间的沟通障碍，适当探索多校联合建立医学信息学专业的联合培养平台，打造信息共享、经验互鉴、实践互助的人才联合培养模式。于专业而言，应开展多学科交叉互融的方式，提升学生的学习应用能力，尝试与数据统计学院、计算机科学学院进行交叉学科的课程革新。本研究的可视化分析也表明国内医学信息人才培养应重视与时代紧密关联，注意课程体系的改革和复合型专业人才的培养，建立更巨特色的医学信息专业学科教学。“互联网+健康医疗”行业带来的人才市场需求风向较为重要，高校应重点研究目前相关人才市场迫切需要的能力与技术，促进学生相应的学习能力培养，从而更好的建设医学信息学科并与大数据背景相适应。

PU-161

我国医药信息管理研究现状、热点与趋势——基于 SATI 的可视化分析

喻真[1,3,4,5]，谢文照[2,3,4,5]，罗爱静[1,3,4,5*]

1. 中南大学湘雅二医院

2.中南大学湘雅三医院

3.中南大学生命科学学院

4. 医学信息研究湖南省普通高等学校重点实验室（中南大学）

5.湖南省心血管智能医疗临床医学研究中心

1 前言

随着信息化和网络化的飞速发展，医疗卫生行业的信息化进程加快，医药信息管理作为一门医药卫生科学、计算机科学、信息科学、管理科学等相互交叉融合而成的综合性学科应运而生。医药信息管理学是信息管理学在医药卫生领域的应用，直接体现了卫生事业发展的需要，研究成果服务

于卫生事业管理，有助于合理配置医疗资源、降低医疗费用以及提高医疗质量。近年来精准医学、智慧医疗以及健康管理等领域迅速发展，为医药信息管理学带来了很大挑战，与大数据的无缝融合、智能分析挖掘、安全与隐私保护等问题成为医药信息管理学研究的新的前沿方向[1]，医药信息管理学日益受到众多学科研究者的青睐，逐渐成为国内外多个学科的研究热点。

医药信息的概念目前尚无定论，所以才有了众多的“名称”，如医学信息、生物信息、卫生信息、健康信息、临床信息等等。对我国医药信息管理专业教育发展回顾[2]，教育部 1988 年 9 月颁发了《普通高等学校本科专业目录》。在医药科中设置了“图书情报学 (医学.药学) ”专业和“中医文献学”试办专业，这是医药信息管理专业的萌芽。“八五"期间,国家卫生部组织开发的“中国医院信息系统(CHIS)"被国家列为“九五"规划。1995 年 5 月 19 日，由国家卫生部组织实施的"金卫工程"也正式启动。同时为适应医药信息处理发展的需要，很多学校都在积极开办医药信息管理专业，培养这方面的人才。

本研究基于 CNKI 数据采用共词分析，利用知识图谱，对我国医药信息管理教育研究的时间分布、高影响力机构分布、核心作者分布、来源期刊以及关键词分布进行可视化分析，为进一步预测我国医药信息管理教育研究领域的未来发展趋势提供一定的参考和借鉴依据。

2 数据来源与研究方法

2.1 数据来源

数据来源于中国知网(CNKI)平台下的中文数据库高级搜索，检索时间不限；为了更精确得到学科现状，采用关键词检索，构建检索式(关键词=医药信息管理) OR (关键词=医学信息学) OR (关键词=生物信息学) OR (关键词=医学情报学)OR (关键词=健康信息学) OR (关键词=卫生信息学) OR (关键词=临床信息学)，经人工筛查，得到医药信息管理有统计意义的国内文献总数 966 篇。

2.2 研究方法

基于 SATI 的可视化途径[3]，SATI 用来进行关键词的字段抽取、词频统计和矩阵生成，绘制可视化知识图谱，借助于图形化手段，统计关键词词频并生成关键词共现网络图谱。采用 SATI 在线软件来分析我国医药信息管理教育研究的文献数量、核心作者、高频关键词、研究热点等信息。

3 我国医药信息管理研究现状及热点分析

3.1 文献时间分布的可视化分析

学术论文的数量及其在一段时间内的数量变化趋势，是衡量研究发展状况的重要指标之一，可以用来分析该研究领域的发展趋势和动态[4]。本研究绘制了 2011-2020 年我国医药信息管理教育研究的论文数量曲线图，见图 1。

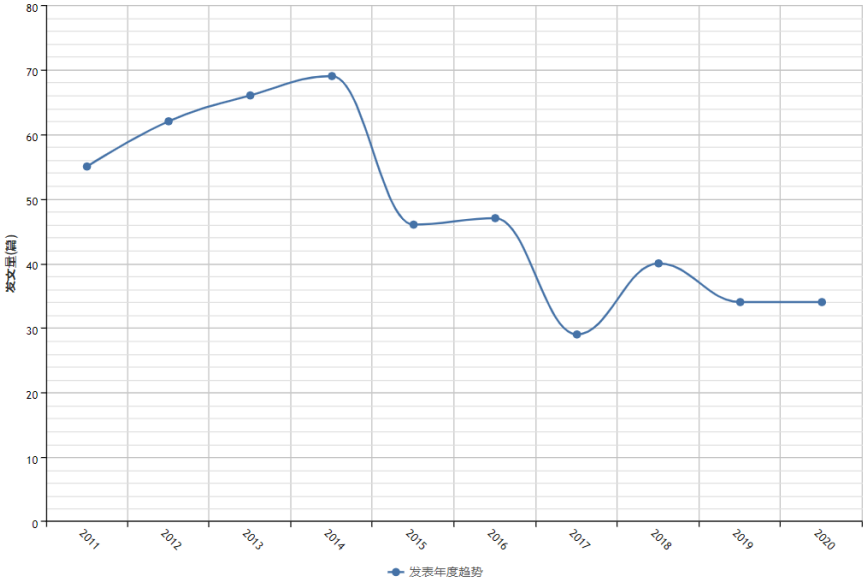


图 1 2011-2020 年我国医药信息管理教育研究相关期刊论文数量曲线图

由图 1 可知，2011 年至 2014 年，我国医药信息管理教育研究的期刊文献数量整体呈上升趋势，年平均文献数量保持在 50 篇以上，表明了我国医药信息管理学科不断发展壮大，成果不断涌

现；2014 年至 2017 年，我国医药信息管理教育研究的期刊文献数量开始呈现快速下降趋势，这可能是因为这段时间研究主题没有创新，研究人员研究主题热情下降；2017 年至 2020 年，我国医药信息管理教育研究的期刊文献数量处于平缓状态，这段时间该研究主题趋于饱和。结合近 3 年来医药信息管理教育研究相关期刊论文数量曲线图的走向趋势，可以预测我国该教育研究的态势没有受到很大的重视，没有相关政策的出台。

3.2 文献来源的可视化分析

对我国医要信息管理研究文献来源分布进行分析，文献来源分布主要反映了学者研究成果的层次，一般硕博士论文表明学者对该领域展开了较全面系统的研究，而文献来源的期刊级别越高，表明研究的越深入[5]。文献来源分布情况可以反映出哪些期刊是职业教育政策研究的主要阵地[6]。本研究绘制了我国医药信息管理教育研究的文献来源分布图，见图 2。

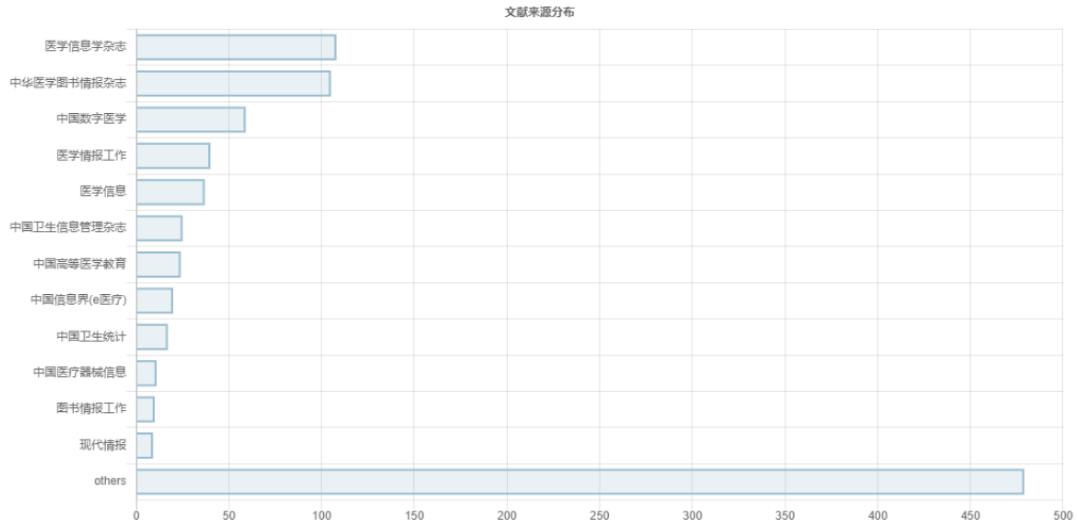


图 2 我国医药信息管理教育研究文献来源分布图

附注：频次 < 1% 的数据归类于 "Others"。

由图 2 可知，数量排名前 10 的杂志从高到低分别为：医学信息学杂志、中华医学图书情报杂志、中国数字医学、医学情报工作、医学信息、中国卫生信息管理杂志、中国高等医学教育、中国信息界(e 医疗)、中国卫生统计、中国医疗器械信息。

3.3 高影响力作者的频次分析

核心作者代表该研究领域的核心研究群体。通过将检索到的有关医药信息管理教育研究文献导入 SATI 软件中，选择'作者进行字段抽取以及频次统计。根据普赖斯定律，$M=0.749(N_{max})1/2$，其中，M 为文献篇数，N_{max} 为所统计年限中最高产作者发文量，发表文献数在 M 篇及 M 篇以上的人则视为该领域的核心作者。在统计的 966 篇文献中，王伟发文量最多，即 $N_{ma}=19$，将其代入公式计算出 M=7.116，按照取整原则，发表文献数量在 8 篇以上的作者可视为劳动教育研究领域的核心作者。经统计，发文量在 8 篇及以上的作者有 16 人，具体数据见表 1。

表 1 核心作者统计

作者	频次	%
王伟	19	0.9009 %
董建成	19	0.9009 %
刘岩	14	0.6638 %

蒋葵	14	0.6638 %
崔雷	13	0.6164 %
刘亚民	12	0.569 %
刘文君	11	0.5216 %
包含飞	11	0.5216 %
秦惠基	11	0.5216 %
牟冬梅	10	0.4742 %
刘燕	9	0.4267 %
周毅	9	0.4267 %
孙丽	9	0.4267 %
李华才	8	0.3793 %
李后卿	8	0.3793 %
钱庆	8	0.3793 %

3.4 来源期刊的可视化分析

文献的机构分布能够展现研究的平台情况，从而反映出研究的主要力量所在。图 3 展示了高频机构云图，我国医药信息管理研究的文献机构以中华医学会医学信息学分会为首，发文 16 篇；次之则为吉林大学公共卫生学院、河北省医学情报研究所和中国医学科学院医学信息研究所(见图 3)。可以看出医药信息管理研究的机构较为分散。

图 3 我国医药信息管理教育研究高频机构分布云图

3.5 关键词的可视化分析

3.5.1 高频关键词分析

笔者利用 SATI 对文献进行了关键词数量统计，得到了我国医药信息管理教育研究高频关键词频次时间序列图，如图 4，展示字段个数 30 个。从图 4 可以看出，在 2000 年前，医药信息管理主要是以“医学情报学”存在研究领域，而在 2009 年至 2015 年出现“医学信息学”概念，开始取代“医学情报学”，成为医学信息领域研究热点，并且在这段时间，研究主题也开始趋于多样化，但对医药信息管理没有一个确定的概念定义。2016 年至今，健康信息学以一个新的研究主题出现在医药信息管理领域，是除了医学信息学之外研究最多的主题，已经引起相关学者的重视和关注。

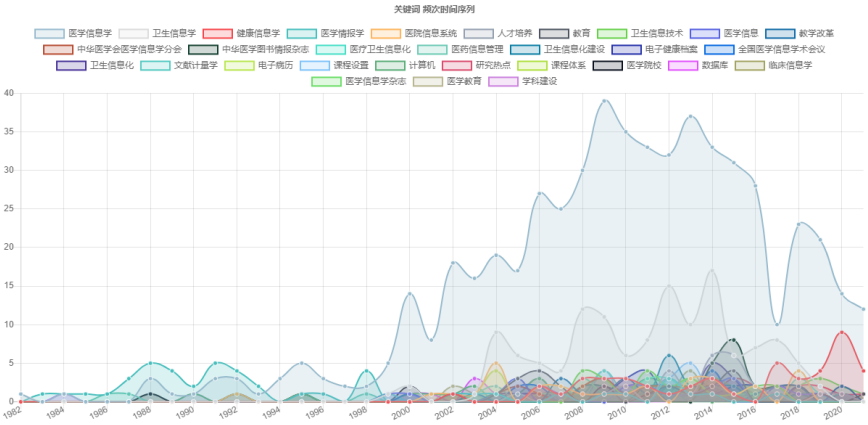


图 4 我国医药信息管理教育研究高频关键词频次时间序列图

附注：频率 = 频次 / 当年总和 * 100。这里的[频次]是该词条在某年出现的次数；[频率]是该词条在某年出现次数占总词条数的百分比。

3.5.2 关键词知识图谱分析

SATI 软件绘制高频关键词知识图谱，可以看出高频关键词在网络图谱中的影响力的大小，节点间的连线的粗细表示该关键词出现的频率高低，具体见图 5。“医学信息学”居于网络的中心，“卫生信息学”、“医药信息学”以及“健康信息学”等关键词也相对突出，说明这些关键词是当前研究的热点；“医学会”、“中华医学会中华医学信息分会”、“医学教育”等关键词的节点连线也较粗，说明医学会、中华医学中华医学信息分会旨在推进我国医学教育。但是从整个知识图谱可以看出，我国医药信息管理研究还是比较分散，概念定义以及研究主题都没有一个确定的大主题，除了对“医学信息学”概念研究的内容之外。

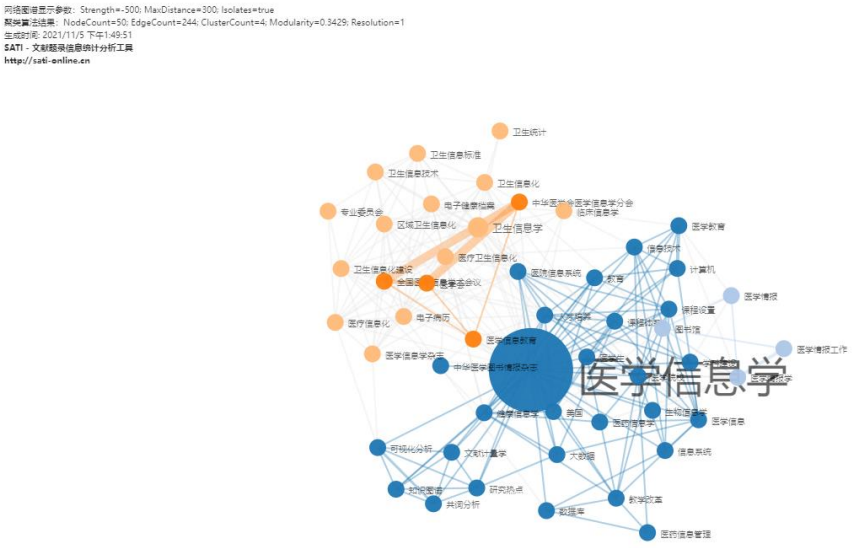


图 5 我国医药信息管理教育研究高频关键词知识图谱

4 我国医药信息管理研究的发展趋势

以 CNKI 数据库发表的医药信息管理学文献为研究对象，运用文献计量法，从文前发表年份、作者、高频引用文献、研究热点、研究领域和研究机构等方面分析医药信息管理学的研究现状以及发展趋势。研究结果表明：(1)从年发文量来看，医药信息学研究在过去十年中处于活跃状态。(2)从作者角度看，低产作者占比较大，研究人员合作关系较为封闭和松散。(3)从高频引用文献看，较早引入信息科学领域最新研究成果的文献备受关注。4)从研究领域看，卫生信息学、医学信息学、医学教育等领域研究成果较为丰硕。(5)从研究机构看，医学类科研院所和高校拥有较高的发文量，这些机构大多分布在我国中东部地区，西部地区依然落后。

综上所述，医药信息管理学的发展发展呈现出时代特点，但不可否认的是出现按的很大的局限性。在初始阶段，以概念性的探索为主，这个阶段以医学情报学为代表；中期，与不同领域交叉融合，出现了医学信息学研究热潮；后期，出现以健康信息学为代表的研究，基于新型的信息技术，学者们开始探索健康信息学这个发展方向。由此推测，未来医药信息管理学的发展仍会结合新出现的信息技术而出现新的变化，如人工智能、机器学习等技术的出现，将给医药信息管理学带来新的发展契机。不仅如此，健康信息学还将与不同的学科相融合，碰撞出新的思想火花。

PU-162

“新医科”背景下的医学生数字素养培育与创新思考

江帅卿[1]，秦选斌[1]，南京辉[2]，沈丽宁[1]，张韦[1]，马敬东[1]
1.华中科技大学同济医学院医药卫生管理学院
2.华中科技大学同济医学院附属协和医院

一、引言

培养有道德、有理想、有信念、有本领、有担当、有作为的卓越医学人才，是新时代医学教育提出的要求[1]。医学教育涉及教育和卫生两个重要领域，推进医学教育高质量发展是新时代卫生和教育系统高质量发展的共同要求。同时，人民群众对于卫生健康的高质量需求，推动医学教育供给侧结构性改革[2]。为加快建设高水平本科教育，全面提高人才培养能力。2018 年多部门联合发布《卓越医生教育培训计划 2.0》，对新医科建设展开全面部署，重构医学教育新体系、新模式，促进信息技术与医学教育深度融合[3]。“新医科”建设强调理念新、背景新和专业新。数字技术及其在卫生健康系统中的广泛应用将是“新医科”建设中无法回避的议题。

在社会数字化转型背景下，数字技术正在改变医学研究和医疗服务的传统模式，也正重新塑造着医学生未来的执业环境。这对医护人员的数字能力和数字素养提出了新的要求。西方国家的调查表明超过一半以上临床医学生或者护理学生无法将其计算机或信息技术能力有效转换成为执业能力。研究者以此提出在医学教育体系中增设数字素养培育模块的必要性。同时，培育医学生数字素养、提高医学生数字能力，是对接国家“四新”学科建设和“六卓越一拔尖”计划 2.0 的新要求，也是“新医科”建设面对社会和卫生健康服务系统数字化转型的必要举措，

二、基本内涵

2018 年我国印发新时代教育改革重要文件，正式提出了“新医科”概念。高教育部高教司司长吴岩就“新医科”的核心要义做出强调：加强“新医科”建设，一是理念新，二是背景新，三是专业新[4]。发展“新医科”是新时代党和国家对医学教育发展的最新要求，是将医学学科与理、工科等前沿科学进行有效整合，前瞻性布局新兴医学或医学相关专业建设。

“新医科”背景下的数字素养培育工作以“医为本质”为起点，以数字素养教育的关键理论为

核心内容，是培养学习者数字能力、弥合公民数字鸿沟、塑造符合时代发展人才的重要途径。数字素养强调应用数字技术，获取、理解、整合和共享数字信息，最终实现价值创造。这一概念与信息素养既有联系也有区别。笔者认为数字素养是信息素养的拓展和延伸，既包括了传统上的信息意识、信息能力和信息伦理等内容，同时更加强调数字环境中数字技术的应用能力，从而与数字时代的日常生活和工作有着更加紧密的联系[5]。

三、数字素养培育的重要性和必要性

1.数字素养培育的重要性

高校被誉为数字素养培育的土壤，提升医学生的数字素养势在必行。开展数字素养培育工作有利于提升医学生对数字技术的理解和认知能力，增强自身在未来执业场景中应用数字技术的能力，能够更好地适应数字化工作环境。对于医疗机构而言，数字技术的应用促进了医疗健康服务水平的提升，为健康产业创新发展提供了技术支持[6]。研究表明，具有较好数字素养的人往往对采用新技术有更积极的态度和行为[7]。因此，提高数字素养可以帮助医护人员增强其在临床实践中自主采纳数字工具和平台的意识和能力，有助于改善数字技术的使用情况，改变传统的临床服务提供方式，为患者提供更加优质、便捷的医疗服务[5]。

2.数字素养培育的必要性

随着以移动互联网、大数据、云计算、物联网、区块链等为代表的数字技术的发展，催生了移动健康（Mobile Health）、普适健康（Ubiquitous Health）以及数字健康（Digital Health）等新概念，推动卫生健康领域形成新业态[8]。具体表现在两个方面：一方面，数字技术有助于促进服务质量和患者安全，减少医护人员的工作负荷，提高服务系统运行效率；另一方面，大量应用数字技术增加了服务系统、业务活动的复杂性。无法有效使用数字技术或者错用误用基于数字技术的临床业务系统，将给医护人员和患者带来一系列风险。这对医护人员的数字能力和数字素养提出了新的要求。同时，如何对接国家数字人才发展战略需求，在现有实践基础上进一步提升医学人才的数字素养，成为当前卫生健康领域以及高校所面临的重要课题[5]。

四、“新医科”背景下医学领域数字素养培育研究

1.“新医科”医学教育现状

2020 年 9 月，国务院提出“以新医科统领医学教育创新，着力创新体制机制，分类培养研究型、复合型和应用型人才”[9]。“新医科”作为一种新型医学教育模式和人才培养体系，我国不少高校和研究者围绕“新医科”医学教育实践展开先期探索。南京中医药大学推动医学学科专业结构的横向拓展，实施“M+中医药”和“中医药+M”交叉学科人才培养，整合和增设医学信息工程、大数据管理与应用等新专业，开设以临床问题为基础的多学科融合课程[10]。李珊珊等对我国 24 所中医药院校的“新医科”建设现状进行分析，研究发现非医学类专业普遍存在建设基础薄弱、专业培养口径窄化等问题[11]。尚丽丽针对当前医学生数字素养普遍偏低等问题提出培养策略，并划定医学人才应当具备的数字素养范围[12]。

国内很多高校和研究者根据医疗卫生技术人员的职业定位，结合“新医科”建设背景和医学生未来执业场景，提出人才培养方案。医师培养方面，有学者提出“新医科”背景下的医师培育的教学体系，关注数字能力的培养。通过开发适用于医学学科的高级结构化课程和开展跨学科合作，使医生掌握基本数字能力[13]。护理人员方面，林敏等从职业教育出发提出完善护理学二级学科体系建设，引导专业型研究生教育向职业化转变[14]。药学技术人员培养方面，都丽萍从药学信息服务出发，提出符合我国药师工作需求的 DIS 服务模式流程[15]。北京市已将文献检索课作为住院药师规范化培训的公共必修课,成为数字素养教育的重要一环。

从综述内容我们可发现，现代医学教育注重从医学生未来执业场景出发，分职业培养医学生的职业能力，数字技术在医学生未来执业场景中所起到的作用愈发重要。在“新医科”背景下，我国高校开始尝试传统医学和各学科交叉融合,医学生的数字素养培育工作越来越受到重视。不可忽略的是，不同高校之间在学科交叉融合上存在一定的现实困境，普遍存在着学科结构单一、发展不平衡、多学科协作和创新能力不足等问题。同时，医科大学尤其是地方性医科大学还受制于学科壁垒，缺乏理科、工科等教学资源，医学生的数字素养普遍偏低，这在一定程度上制约了“新医科”建设的进程[16]。

2.数字素养培育现状

2.1 国外数字素养培育现状（除了项目之外可以提下卫生领域概念框架）

在医学领域，数字素养培育项目从关注医疗领域转变为关注特定的医疗角色（如护理人员、医师、药师以及医技人员）。当前，国际上已开展的数字素养培育项目大多是针对医学生、家庭医生以及全科医生，然而近几年培育的重点已经转向护理人员[17]。比如美国护理学会将计算机课程列为护理专业课程，并划定培育实践的主要内容，分为计算机技能、信息技术应用、信息管理能力三部分内容。[18]据美国医学信息学会网站统计，设置与培养医学生数字素养相关课程的高校达 71 所，涉及的专业包括医学信息学、临床信息学、护理信息学等。调查美国高校医学信息类课程设置情况发现，这类高校主要将医学与卫生信息学、统计学、计算机等学科相结合[19]。2013 年，英国实施技术强化学习项目，数字扫盲项目作为重要抓手，帮助卫生健康领域工作队伍应用新兴技术，培养一批数字化后备军。该项目共发布了 5 项行动指南，确保项目落地[7]。英国国家临床医学卓越研究所(National Institute for Clinical Excellence，NICE)与本土大学联合启动 NICE 学生冠军计划（NICE SCS），通过使用 NICE ES 搜索引擎向医学、药学和护理等专业的学生提供数字扫盲课程[20]，采用传统教学与同伴领导的教学方式培养学生的数字能力[5]。

2.2 国内数字素养培育现状

近年来，我国数字素养研究和培育实践工作不断受到业内重视。目前国内依托于基金资助项目的数字素养研究实践达到 141 项，近五年的研究占比高达 90%。但在医学领域，我国尚未在国家层面实施数字素养培育项目，开展数字素养培育实践相对不足，培养形式较为单一。数字素养培育内容仍停留在信息素养层面， 其内涵和外延还缺乏清晰的认知，相关研究更多是依附于信息素养而存在。国内研究者围绕培养策略、培养模式以及测评量表等方面开展了大量信息素养培育研究。比如朱妍昕[21]、彭骏[22]结合国内医务人员信息素养的继续教育形式，提出适合我国医务人员的继续教育策略。和欢等[23]围绕信息意识、信息能力、信息安全与道德 3 个维度编制了护理人员信息素养自评量表，具体细分为 35 个条目。我国有研究者为护理信息学的开设做出了很多努力，为护理信息学课程的开设与实践，做出一系列可行性研究[18]。在培育实践上，国内护理院校陆续开设护理信息学课程，如吉林大学护理学院开设护理信息学课程，杭州师范大学钱江学院、湖北中医药大学纷纷效仿，率先将自动识别技术、移动计算技术、无线网络技术引入教学和实验中[24]，课程的设置内容已经超出传统的信息素养范畴。如今，医学类院校中的信息素养课程越来越多涉及到数字化的内容，逐渐体现数字素养培育的特征。

在医学领域，单纯信息素养培育研究不能满足数字化执业环境的需求，仍需开展专门的数字素养培育工作。虽然我国数字素养培育取得一定成果，但总体来看还不完善[25]。数字素养与信息素养的差异除了传统的信息意识、信息能力和信息伦理之外，数字素养培育注重医护人员数字技术的应用能力，使其能更好地适应数字化执业环境[5]。

五、“新医科”背景下医学生数字素养培育的创新途径

1.关注新医科发展需求，明确医学生数字素养发展目标

在新医科背景下，培养医学生数字素养要顺应“新医科”学科建设的需求变化，在核心能力的培养上实现突破，超越原本属于信息素养的能力范畴。医学生数字素养培养策略应注重多学科的交叉融合，培养精通大数据、人工智能等新一代信息技术的“数字化医生”，真正能够利用数字技能参与到医学活动内核，发挥数字技术对医疗服务方式和服务模式所带来的颠覆性作用。一方面有利于医学生了解如何运用数字技术解决临床工作中遇到的问题，适应数字化的工作环境，为临床诊疗服务赋能。另一方面，依托数字技术有利于为医学研究寻找新突破口，提供全新的方法路径，推动“医学+数字技术”研究深入开展。

2 .以培养数字素养为主线，构建合理的课程体系

构建合理的课程体系是培养学生的数字素养和数字化思维方式的核心因素和基础性环节。作为医学生这一特殊群体，课程设置要紧紧围绕医学专业的学科特色，与基础医学、临床医学、预防医学等多学科开展合作，将数字素养提升重点与临床问题紧密结合，将数字素养的核心知识与教学资源嵌入到学科专业课程体系中，形成医学领域的数字素养课程体系。教学内容以培养医学生数字素养为主线，在教学方式上可侧重案例教学法并结合虚拟教研室进行教学。例如,将医院数字化转型典型应用场景和数据处理软件的使用引入到教学中，使学生身临其境地对医院数字化应用和数据处理技术有深入的了解。包括远程医疗、生物电信号处理、医学影像诊断技术、智能医疗辅助诊断系统、数据挖掘等各个领域的具体应用。

3.优化医学生数字素养培养模式，建立面向未来执业环境的人才培养方案

在医学生人才培养方案中，将原本培养医学生“信息素养”为核心的培养策略转变为培养医学生的“数字素养”，从强调培养信息管理、信息服务等核心能力，逐步过渡为培养基于临床工作环境下的“数据分析”“数据挖掘”能力以及数字工具的实践运用能力，以此更好的适应未来数字化的工作环境，能够利用自身的数字能力解决实际工作中遇到的临床问题。在未来职业规划上，医学生就业方向的选择应更加多样化，除了医疗机构之外，还可面向医学数据分析的企事业单位。

4.加强医学生科研与项目实践，多方渠道培养数字技能和数字素养

高校可与医疗机构、企业、政府等多方主体开展实践交流合作，提高医学生数字素养和数字技能。首先，高校可与企业合作举办数据分析类竞赛，利用平台资源优势，锻炼学生的知识运用能力。其次，高校与医疗机构可开展医院临床信息化应用项目合作,提升医学生临床诊疗数据挖掘与利用能力；政府和学校可组织类似于医疗健康首席数据官（CDO）培训项目，针对正式学院开展培训的同时择优选择部分优秀学生供其旁听学习，提升学生数字化思维和数据管理能力。

PU-163

数据科学在医疗保健发展中的作用

张超
郑州大学第一附属医院

背景

数字时代的发展导致了医疗保健和技术的混乱，同时带来了不断创新的数据相关应用程序的出现[1]。医疗保健部门的电子病历、处方、临床报告、购买药品的信息，医疗付费，调查，和实验室报告等数据，给分析和研究的新技术带来了巨大的机遇[2]。大量的数据可以汇集在一起，并使用机器学习算法进行单独的分析，分析细节和理解数据中的模式可以帮助进行更好的决策，从而提高病人护理质量，改善医疗保健结果，增加预期寿命，实现疾患的早期识别，进而以达到较好诊治效果的基础上降低医疗成本[3]。因此，医疗保健相关组织必须尝试获取所有可用的大数据，建立更好的医疗保健网络，在公立医院考核中脱颖而出，以获得更好的医疗质量[4, 5]。

随着社交媒体时代以及传感器和分析仪监测个人健康参数的智能应用程序的到来，大数据及其在医疗保健的应用已经变得更加重要[6, 7]。数据挖掘的在医疗中的作用提取全方面信息，以提供更好的治疗和护理。大数据挖掘与分析需要技术复杂的应用程序，使用高端计算资源和基于人工智能的算法来处理庞大的数据，通过应用程序的模糊逻辑和神经网络的自动决策将增加优势。大数据需要创新和有效的数据处理策略、基于数据的智能云的应用程序、有效的存储和用户友好的可视化，以获得实际的见解[8]。

1 医疗保健与大数据

1.1 医疗保健数据的来源及多样性

医疗保健的关键要素是卫生从业人员、卫生保健设施，以及为前者提供资金的一个供资机构。卫生保健从业者属于不同的健康领域，如牙科、药学、医学、护理、心理学、联合健康科学等。根据病例的严重程度，在许多层面提供保健。在所有这些阶段，卫生从业人员都需要不同形式的信息，如患者的病史、临床数据和其他个人或私人医疗数据、书面笔记或以打印报告[9]。

1.2 电子健康记录与大数据

临床病例记录保存了一个人体内疾病的发病率和结果[10]。随着电子系统及其能力的不断发展，医疗检查、健康记录和调查的数字化在当今已是常见。2003 年，美国国家科学院和工程院创造了“电子健康记录”一词，用来代表保存患者记录的电子门户。电子健康记录是与个人的身心健康相关的医疗记录，保存在电子系统中，用于记录、发送、接收、存储、检索和医疗服务连接等，保存于电子系统中的数据汇总成为大数据[11]。

2 数据科学

2.1 大数据平台

数据挖掘技术包括数据分组、数据聚类、数据相关性、顺序模式的挖掘、回归和数据存储。然而，数据的量通常非常大，以至于一定时间内完成工作就需要大量的计算机来分配和处理。即使最强大的计算机也不能直接处理巨量的数据，需先将其纳入存储中，可实现随时，有针对性地调取数据，完成分析，降低处理成本。因此，大数据存储平台作为接收、存储、下载及挖掘的意义在大数据分析中具有重要意义。

2.2 数据类型多样性

数据类型可以根据它们的性质、源和数据收集方法进行分类[12]。最常用的类型是由人类生成的数据，包括电子医疗记录、电子健康记录和电子患者记录。由于国家医保政策的不断改革，医疗部门数字化发展及人们对健康的重视，医疗数据的类型及体积还在不断。可穿戴的个人健康监测设备的突出地位和市场增长，为数据挖掘提供的信息也在不断增长[13]。

3 数据分析在医疗保健中的应用

3.1 疾病监测

它涉及到对疾病的感知、了解其状况、病因（疾病的病因形成方式）和预防。通过电子病历和互联网获得的信息在疾病分析方面具有巨大的前景。各种监测方法将有助于规划服务、评价治疗、确定优先事项以及制定卫生政策和措施。

3.2 图像处理

医疗保健数据的图像处理具有对解剖学和器官功能有价值的知识，并识别疾病和患者的健康状况。目前已被用于器官勾画、肿瘤识别、脊柱畸形诊断、动脉狭窄检测、动脉瘤检测等[14]。广泛应用的小波技术常用于图像分割、增强、降噪等处理技术。在图像处理中使用人工智能将增强卫生保健的各个方面，包括筛查、诊断和预后，并将医学图像与其他类型的数据和基因组数据相结合，将提高准确性，促进疾病的早期诊断[15, 16]。

3.3 可穿戴技术数据

不少的科技公司已研发出基于健康的应用程序和可穿戴健康设备。用以收集与生理指标（如心率、燃烧卡路里、血糖水平、皮质醇水平）相关的实时医疗数据（如情绪、饮食、运动和睡眠周期模式）与传统医疗大不相同，其旨在收集以患者为中心的连续数据，具有较好的时效性，同时也具备较强的隐私性。但是，医疗机构可在将来的研发更多的具有病种针对性的，简便的可穿戴设备，用以患者在诊疗过程中及诊疗后的评价研究。

3.4 数据云存储和云计算

数据云存储主要用于安全，结构化存储以患者为中心的电子数据[17, 18]，以提高医疗质量。除了医疗目的外，数据存储还用于研究、培训、教育和质量控制。决策者或医疗工作者可以通过专业的检索策略，有针对性从存储库中提取相关数据，用以分析，决策及研究诊疗新思路。

3.5 控制成本

电子医疗记录系统的迭代，大量数据的不断更新、合并和交换，数据类型的不断增加都会使得大数据有些杂乱无章，且易产生不必要的诊疗费用。结构化这些数据的异质性仍然构成了重大挑战。可以将不同来源，同一目标的数据转化为统一标准的数据类型，避免医生过度使用筛查和检测往往导致数据过剩和额外的成本[19]。基于大数据处理可构建更为科学的决策支持系统，尽量减少不必要的检测，以减少患者诊治的总费用。

3.6 公共健康

大数据及相关技术可以帮助研究人员理解目前正在产生的大量丰富和多样的健康数据。数据分析是必不可少的，因为当今社会生成的数据已经超过了我们在没有计算机支持的情况下从数据中获取知识的能力的程度[20]。为了进行这些分析，需要使用强大的计算和存储技术，并结合人工智能算法，以理解和了解大数据。数据分析也被应用于疫情爆发期间的疾病检测，对于疫情的早期发现，及时管控、疾病诊治及疫情预测具有重大意义。

3.7 药物安全

药物警戒需要在药物启动后跟踪和识别药物不良反应，以保证患者安全[21]。也有学者认为药物的不良事件可能不仅由单一药物引起，而且还可能由合成药物的混合物引起的理论[22]。通过联合多个数据源，可识别药品不良反应的精度[21]。识别药物和潜在不良反应时有可能获得错误信号，可利用大数据处理，建立可实时更新的预警模型，从而避免传统的假警报，对潜在的药物警戒活动有很大的帮助[15]。

4 面临挑战

与序列数据一样，处理和控制完整数据也将成为数据科学面临的一个主要挑战[22]。

1 收集的数据大多混乱且不准确。

2 确保这些数据的质量和可访问性之间的平衡是难以确定的。

3 数据标准化、维护隐私、高效的存储和传输需要大量的人力来持续监控和确保需求得到满足。

4 基因组数据集成到医学研究中是至关重要的。

5 处理数据时的计算机语言障碍。

5 展望

大数据的应用是医疗保健领域的一场革命。患者、医生和医疗保健提供者对医疗服务的态度才刚刚开始转变。所讨论的对大数据的使用只是冰山一角。随着数据科学的普及和各种数据驱动应用程序的出现，卫生部门仍然是数据驱动解决方案的引路人，为其客户提供更好的生活和量身定制的服务。数据科学家可以通过他们关于医疗保健领域的广泛数据，包括财务、临床、研发、管理和运营细节，获得对提高医疗服务质量有意义的见解。

PU-164

基于 CAS 理论的图书情报专业硕士创新实践能力提升体系构建

杨梦柯 沈丽宁*
华中科技大学

目的 通过对图书情报专业硕士创新实践能力培养现状的分析，总结目前培养困境，基于 CAS 理论构建图书情报专业硕士创新实践能力提升体系并提出具体实施建议，期望能够促进图书情报专业硕士创新思维的锻炼和实践能力的提升，培养符合卫生健康行业高质量发展需求的创新应用型信息专业人才，助力健康中国建设。

方法 通过文献梳理发现目前图书情报专业硕士创新实践能力培养存在信息流通闭塞、培养定位不清晰、教学方式单一、多方协同育人培养不成体系、过程管理不到位、实践教学效果评价难概全貌、校外导师不足、教师队伍专业水平有待提升、基础设施不足、资源建设有待加强等问题。图书情报专业硕士创新实践能力提升系统从基本特征上符合 CAS 理论的系统模式，结合培养困境、基于 CAS 理论、从探测器、规则集和效应器三部分构建了图书情报专业硕士创新实践能力提升体系。

结果 明确了专业硕士教学委员会、教务人员、教师、学生四大主体，形成了以案例教学、双师课堂、专题讲座、现场教学为核心的情景课堂和以专业实习、项目实践、创新创业大赛为核心的创新实践能力训练为图书情报创新实践能力提升体系的核心实现路径，为图书情报专业硕士创新实践能力培养提供指导。

结论 在困境梳理和体系构建的基础上，提出图书情报专业硕士创新实践能力提升的具体对策，包括以社会需求为导向、加强创新实践能力培养、丰富教学方式、构建协同育人的创新实践能力提升体系、优化过程管理、拓展实践教学考核手段、加强师资队伍建设、提高创新实践能力教育质量、强化资源建设、为创新实践能力提升提供基础保障。期望能够促进图书情报专业硕士创新思维的锻炼和实践能力的提升，培养符合卫生健康行业高质量发展需求的创新应用型信息专业人才，助力健康中国建设。

PU-165

基于数字科研培训平台本科生深度参与暑期科研项目的探索研究

梁宏锌[1,2] 周海榆[1,2,3,4] 刘盛波[2] 邱鸿锐[2] 唐子茗[3] 尤培蒙[4]
1.广东省人民医院(广东省医学科学院)广东省心血管研究所
2.广东省人民医院胸外科
3.南方医科大学
4.江西省肿瘤医院江西省肺癌研究所

目的 在“新医科”背景下为医学生本科科研教育与数字医学平台的结合探索提供一些参考，建立以数字医学科研培训平台为基础，多阶段、多维度、多层次融合创新的医学本科生暑期科研培训。积极探索一项整合各大医院与医科高校的项目实践资源，创新医学科研培训模式，线上、线下有机结合的智慧教培一体化方案，为我国医学本科教育添砖加瓦。

方法 分析新医科建设背景下数字医学科研培训平台特色，制定一套合理的学、思、用贯通，知、信、行统一的医学暑期科研项目运行模式，打造系统的医学本科生培养知识树。

结果 合理利用医学本科生暑期时间，早期深度参与各医学院校，医院课题组科研项目；高效利用数字医学科研培训平台，提高医学本科生在参与过程中的系统性认知。实时地了解医学本科生的培养方案和成长方式，更好地进行教学改革，在医学研究和临床教学上进一步探索，找到适合他们的培养方式。

结论 抓住新一代年轻医学生在学习成长过程中的痛点与难点，提高教学的质量和效率，实现临床医学本科生早期参与临床或基础医学研究，调动积极性，提前对医学科研有初步认知，同时为医院科研教育改革尝试，贡献实践经验。

PU-166

“智慧医疗”背景下医院信息化复合型人才培养探究

韩明泽① 彭晨辉②

①哈尔滨医科大学附属第六医院

②哈尔滨医科大学附属第六医院

1.引言

智慧医疗是指为实现患者与医务人员之间的数字化、智能化互动，打造健康档案区域医疗信息平台、利用最先进的互联网技术实现智慧医疗[1-2]。《“十四五”全民健康信息化规划》提出，到2025 年，我国将初步建设形成统一权威、互联互通的全民健康信息平台支撑保障体系，基本实现公立医疗卫生机构与全民健康信息平台联通全覆盖。随着医疗信息化的不断发展，在医院信息化建设中存在诸多问题。目前，我国医院信息化建设中存在医学与技术融合较为困难、懂“医”精“技”专业人才资源紧缺，科技研发卡脖子，创新技术应用难等相关问题，进而影响医疗服务从“信息化”向“智慧化”的发展[3-5]。因此，在智慧医疗背景下，如何培养可持续发展的、适应行业需求的医院信息化复合型人才培养，成为智慧医疗发展面临的重要问题之一。

2.医疗信息化复合型人才培养的必要性

《“十四五”全民健康信息化规划》提出：全民健康信息化建设创新引领卫生健康事业高质量发展，医疗信息化发展是推动全面建设智慧医疗的重要战略举措。人才问题极大限制了智慧医疗产业发展进程，医疗信息化复合型人才保障不健全，全民智慧健康落实不到位，导致医院难以实现真正的数字化的智慧医疗转型。所以，从医院发展来看，医院信息化的充分应用和合理规范都是其必然发展趋势。“十三五”期间我国全民健康信息化建设成效显著，但目前仍处在夯台垒基、爬坡过坎的关键时期，专业人才匮乏、“数字鸿沟”、“数据壁垒”依然存在。虽然国内开设培养医疗信息化人才专业的高校不在少数，但由于缺少顶层设计和系统规划，人才培养目标模糊，专业理论实践课程设置不规范，对应师资力量配备不足，专业课程设置不合理，实践能力缺乏等[6-7]，导致培养的医疗信息化人才不符合行业要求。目前，医院信息部门人员缺乏医学或计算机相关等专业知识，临床医务工作者缺乏信息化相关知识[8-9]，导致双方沟通存在障碍，不利于医院信息化建设，阻碍智慧医疗的发展，进而导致高校人才培养和社行业需求不能形成呼应。加之目前医院信息部门人员对于如何运用互联网、大数据等信息技术对业务流程进行优化一头雾水，缺乏数据挖掘、软件开发等能力[10]，导致医院信息化建设发展不能满足时代需求。为更好应对新时代数字化变革带来的机遇与挑战，必须进一步夯实医院信息化复合型人才培养计划，为我国智慧医疗事业发展注入源源不断的新动能。

3.培育交叉医院信息化复合型人才培养探究

《关于教育支持服务产业发展，提高紧缺人才培养培训质量的意见》提出联合高校建设卫生健康创新人才孵化基地，推动政府、科研院所、医疗机构、继续教育形式，完善多层次、多类型人才培养培训体系为企业共同培养人才，促进培养跨界复合型专业人才队伍建设。所以，医院信息化人才培养需要与时俱进，可以通过国家、医院、高校的资源整合，使这支队伍发挥最大效能。

3.1 明确国家发展，完善落实政策

中国智慧医疗将在未来五年迎来发展提升的黄金时代。《“十四五”全民健康信息化规划》中明确表明：要注重拓宽人才培养渠道，充分发挥科研院所、高等院校特别是国家健康医疗大数据研究院等机构，在全民健康信息化工作中的智力支撑作用，加快建立适应行业特点的新一代信息技术创新应用人才队伍培养体系。国家应研究制定更符合大数据时代要求的卫生健康信息化人员配备标准，加强完善数据分析、网络安全等技术人员配备，完善复合型人才培训机制，针对不同岗位的医

护工作人员对信息化、智能化医疗的需求制定系列个性化培训方案。《规划》中还对各地建立首席信息官制度的探索工作进行鼓励,进一步落实对医学信息技术、生物医学工程等新兴专业和复合型人才的建设研究工作，形成促进复合型人才与实用型人才相辅相承、齐头并进的良好环境，为智慧医疗高质量发展奠定有力基石。国家应该对数字健康管理服务模式进行重塑，丰富服务空间，拓宽数字健康应用场景。

3.2 规范人才培养，明确医院发展

3.2.1 扫除盲区提高认知，顺应自主研发趋势 医院需要因地制宜，开展智慧医疗建设相关培训、讲座。医院的管理者对于信息化建设的重视程度不同，致使各医院智慧医疗发展水平不尽相同。有些医院管理者十分重视智慧医疗的实践建设工作，能够顺应时代变更，跟进信息化发展，把互联网、信息化、大数据等与医院的医疗、服务、管理等各方面工作有机结合，优化患者诊治流程、提高就诊效率。反观某些医院及医疗机构，其管理者对互联网、大数据、信息化等概念一头雾水。尤其在当今的数字化时代，医院管理人员的智慧化思维水平不够，定然会阻碍医疗信息化复合型人才培养发展的进程。因此消除医院管理者对互联网医院建设，信息化医疗发展的认知盲区是开展医院层面信息化复合型人才培养的首要工作。依托健康中国战略及国家宏观政策，各级医院应重点增强对管理人员的互联网医疗实践培训，进而促进医院管理人员的医学计算机应用能力，形成智慧化管理思维。

目前医院信息化人员欠缺研发能力，大多数医院会采取外包的形式进行信息平台开发，而信息平台的开发和利用会根据医疗行业内政策的变化和各医院管理机制的调整而不断更新，所以后期投入较多，如果完全依赖第三方，维护成本会比自主研发高出很多。如果医院意识到自主研发的重要性，培养并组建自己的研发团队，不仅可以优化医院信息化的整理流程，也会让团队成员成长为更专业化的医疗信息化人才。

3.2.2 成立基金引进人才，合理设置激励机制 以引进高端人才为措施，加强高端人才培养为重点，形成引进人才带动本地人才培养模式和本地人才争做高端人才的良性竞争机制。积极引进医、教、研全面发展的高层次人才，及有一技之长的专业技术骨干。对于引进和培养的人才，医院将为其构建良好的工作平台，配备工作、科研所需的设施，提供科研启动资金，解决高层次人才的后顾之忧，创造良好工作环境和生活环境。

科室培养有用之才，让人才学成归来有用武之地，使科室建设与人才培养融为一体。医院应加大赴国内、外进修人员的支持力度，改善进修人员外出期间的福利待遇，减轻进修人员进修期间的生活压力，最大程度解决进修人员的后顾之忧。同时，完善人才培养的选拔机制，合理制定进修人员的进修计划，进修结束后通过学习分享的形式，实现前沿理念、前沿技术的共享。医院也将为进修人员回院开展新技术、新项目予以相应的政策支持。此外，医院应根据自身发展状况合理采取人才激励机制。可通过发放绩效奖金、项目薪酬奖励等方式提高高层次复合型人才待遇，完善医院的人才招聘准入机制及在职人员晋升、评奖评优机制，提高行业吸引力。

3.2.3 鼓励竞赛开展培训，善用以“赛”促“技”抓手 医院以提高信息化人员实做能力和故障处理能力为重点开展技能竞赛活动。以调查问卷、人物访谈的形式征集信息化建设中问题工作案例，了解信息化人员迫切需要学习的知识技能及在工作中遇到的技术难题，组织故障问题专项演练，定准实操工作中的薄弱点和关键点，坚持用问题训、有组织训，确保医院信息化人员胜任力提升。

人才培养取决于自身的努力，更离不开环境的塑造。医院应坚持创造技术学习的好氛围，打造一支业务过硬的医疗信息化人才队伍为目标，通过营造有利的环境，设立专项科研基金，激发信息化人员自我奋进的内在动力，不断提高技术素质。鼓励信息化人员参加竞赛，坚持以“院内比、院间比”的原则，提高专业技能。医院应坚持“以点带面、整体推进”的原则，医院联合筹资，重点对竞赛培训基地给与基金投入，在院内建立读书角，满足不同层次的学习需求；根据培训需求的针对性、内容的实用性和操作性目的，建设技能型、知识型、学习型、创新型、研发型为活动载体的专业队伍，采用“走出去学人之长，引进来补己之短”的做法，结合信息化建设不同时期的要求，安排医院信息化人员进行专业化技术学习，对院内职工分类进行专业培训，着力培养技术骨干人才，挖掘施教资源。为保证医院信息化建设的顺利展开，采用“请进来面对面授课，提问题实打实交流”的互动培养方式，邀请医院信息化建设成熟单位及负责医院信息化软件开发公司的专家及技

术人员进行专题讲座、案例分享、实操训练、解惑释疑，进而提高医院信息化人员对新政策、新技术、新规章的认知能力和岗位技能能力，保证了医院信息化建设的有效开展。

强化技能竞赛，提升职工职业素质是职工及医院的共同需要。针对技术业务能力不足和复合型人才短缺的实际情况，可结合行业特点，把提高职工技能水平和培养复合型人才作为抓手，建立学习论坛，组织技能拔尖人才进行讲座，建立一对一、面对面的传授，建立“师徒制”、“以老带新制”，把个人学习和集中学习进行整合，每周定期开展，形成传授知识的“小讲堂”，解决难题的“讨论会”，练习技术的“练兵场”。组织设擂比赛，以医院为活动单元，医院信息化技能比武为“擂台赛”，以优胜者为状元队伍，可设立技术骨干津贴作为奖励，面向合作单位新挑战，挑战成功者为新状元队伍，守擂成功的队伍，待遇不变，将准备下一年的擂台赛。通过这样的比赛形式，可以为全国网络安全技能大赛等权威赛事做准备，进而为国家培养实战高端人才。

3.3.适应时代发展，创新教育模式

3.3.1 接轨行业明晰需求，对应技能扩充培养 以推进“医学＋X”多学科背景的复合型创新拔尖人才培养为目的，高校在设置交叉学科培养，促进医文、医理、医工学科交叉融合的同时，利用竞赛促进人才的发展，利用多学科、多角度培训和参观企业等实践活动，促进学生的综合素质培养。高校可依托学科和医院优势，组建跨学科教学团队，开设医院信息化建设大数据分析、医用机器人技术等医工、医理交叉的课程，加深学生专业背景，充分应用专业技能，将本专业知识和其它专业知识学习和应用相互融合，以培养顺应时代的智慧医疗背景下医院信息化复合型创新人才。

3.3.2 试点开办联合培养，打造就业保障模板 在校企联合培养模式的基础上尝试开展试点实验校、院联合培养体系，把培养医院信息化建设相关专业学生的全面素质和综合能力与就业竞争能力目标，围绕以学校和医疗机构的教育资源和环境，以课堂教学与实践联系有机结合的方式，培养具有全面素质与创新能力人才，成为相关专业学生毕业后就业保障的一层“安全网”。同时，高校也应和医院保持良好的合作关系，及时了解医院信息化人才的市场需求，及时调整对于相关专业的学生的就业指导规划，让这些学生能更好的适应智慧医疗的人才需求，提高学生的就业成功率。在联合培养体系中的实践模块，高校与医疗机构合作支持学生参加机构实习，参与科研项目的研究以及医疗信息技术和互联网医疗企业的实践活动，提高复合型创新人才的综合素质。

4.结语

智慧健康背景下医院信息化复合型人才培养是当前医疗行业发展的必然趋势。医院信息化复合型人才的培养以知识为本、实践为手段、以创新为核心，提升素质为目标，需要充分发挥政府、医疗服务机构、高校三方的合力，才能实现加快落实国家智慧医疗、智慧健康的宏伟蓝图，攻克人才培养的难关。

PU-167

基于 OBE-CDIO 理念的医学信息人才实践教学体系构建

符凌嘉[1]，张治国[1,2]，沈丽宁[1,*]
1、华中科技大学同济医学院医药卫生管理学院
2、湖北省卫生技术评估研究中心

1 引言

随着国家“健康中国”战略的不断推进，以及我国医疗卫生领域信息化的快速发展。我国医疗卫生信息化需要大批专业能力扎实、实践能力强、综合素质高的医信交叉复合型人才，通过跨学科合作与融合，将当前医疗业务需求与信息管理融为一体，推动医疗领域信息化建设[1]。2021 年 4 月，习近平总书记考察清华大学并发表重要讲话，强调要用好学科交叉融合的“催化剂”，打破学科专业壁垒，对现有学科专业体系进行调整升级，推进新工科、新医科、新农科、新文科建设。这与我校的一流本科建设、一流人才培养理念是紧密相联的。当前，医学教育不断向医文、医工、医理等交叉学科支持的新模式转变。在坚持提升教育质量、推动教育公平的二十大精神的指引下，培

育适应人工智能、大数据等新科技产业革命、能利用交叉学科知识来解决医疗信息化领域难题的创新应用医学信息人才，已经成为了当前大学医学教育和人才培养的必然要求。

医学信息学旨在应用信息科学和技术来满足医疗卫生领域的数据管理、决策支持、研究和临床实践等需求，具有较强的专业性和交叉性[2]。目前，全国已有 40 余所高校开设了医学信息学相关专业的本科教育[3]。那些较早开始开设相关学科的学校，其课程和培养体系已经逐渐趋于成熟，但其教学体系在培养学术创新型和实践应用型的医学信息人才协同方面还存在一些难题。培养的人才与现实工作岗位的实际需求之间存在差距，实践教学通常只被视为理论教学的补充，忽略了实践教学的重要性。专业实践能力作为医学信息专业核心的培养目标之一，要求坚持以就业为导向、以学生为中心的教育原则，实时关注市场对于医学信息专业人才能力要求进行教学设计[4]。以培养具有良好的综合素质，能够适应现代医疗信息化发展高质量医学信息人才，确保其毕业后有一定的岗位胜任力和在该领域中进一步深造的能力。但从目前的教学实践上来看，医学信息专业普遍存在培养目标不够清晰，理论知识与实际应用相脱节的现象，导致部分毕业生无法满足工作岗位的需求[3]。已有研究对于医学信息专业的实践教学进行探索，通过分析医学信息实践教学存在问题，从实践教学课程体系、校内外实践教学基地、实践教学方法等方面提出了对策建议[5-7]。但较少有研究借鉴先进教学理论进行教学设计，亟需采取先进的教学理念，应用于医学信息人才实践教学设计中，以更好的适应时代发展。

随着教育的国际化和全球化，高等教育理念也不断的推陈出新。OBE（Outcome based education，OBE）教育理念基于教学成果来组织、实施和评价教学的模式，保证学生具备工作所需要的能力和素质，非常符合应用型人才的培养需求。OBE 理念在相关专业课程的教学体系构建、教学内容制定、考核评价改革中都得到了一定的应用[8]。CDIO 工程教育模式是近年来国际工程教育改革的最新成果，以项目驱动的教学方法为主，以构思（Conceive）、设计（Design）、实现（Implement）和运作（Operate）为教学实施流程，实现专业知识、实践技能的掌握，培养学生的实践应用和创新能力[9]。近年来，一些研究将 OBE 教育理念和 CDIO 工程教育模式相结合构建 OBE-CDIO 理念，部分大学探索开展 OBE-CDIO 工程教育改革，取得了不错的成效[10]。也有学者基于 OBE-CDIO 理念应用于遥感实践教学模式改革，确定教学目标后，建立课程与学习产出匹配矩阵，优化课程体系；再按 CDIO 组织课程教学，实现教学产出[11]。但目前将 OBE-CDIO 理念应用于医学信息实践教学的研究还尚不多见。

实践教学在创新应用型人才培养过程中具有举足轻重的地位，是培养学生基本技能、实践能力、创新精神的有效途径[12]。而 OBE-CDIO 教育模式与实践教学理念高度契合，将其运用于创新应用医学信息人才实践教学体系建设与创新具有现实意义。本研究强调以目标为导向、以学生为中心进行实践教学体系设计，包括构建多门学科间统筹规划分层实施的实践教学内容，并采用项目式的教学模式，丰富实践教学环节，同时优化多方参与、突出过程的实践考核方法。形成持续改进的实践教学体系，最终达成学生实践能力从基础到综合、创新的递进式培养。

2 医学信息人才实践教学体系存在的问题

2.1 实践教学培养目标不明确

医学信息学作为一个多学科交叉领域，涵盖了医学、信息科学、管理学等多个学科的内容。目前医学信息专业培养主要是以课程为导向，以专业基础知识积累为主。虽然在培养计划中设置了实践学时，但是没有明确的实践教学目标[13]。对于学生在实践教学结束后所需具备的能力类型、能力层次等方面都缺乏明确的指导。且医学信息学是一门交叉性很强的学科，不同高校医学信息学专业依托院系不同，培养目标也各有侧重[14]。有的院系依托于医学院，注重对医学基础课程的学习，有的院校更偏向于医院管理类课程，有的院校则重点是计算机编程类课程。培养目标各不相同，缺乏一致性和系统性，这也进一步导致实践教学的目的难以统一。因此为保证学生快速适应行业、企业等用人方的实际需求，应该从实际需求出发，明确实践教学培养的目标，以提高医学信息人才的实践能力以适应实际岗位要求。

2.2 实践教学内容难成体系

在 2020 年最新发布的《工程教育认证标准》中指出，高校的工科类专业教学学分中，实践和毕业实习至少占总学分的 20%，并要求进一步完善实践教学体系[15]。然而，由于专业课程实践教学培养目标不明晰以及交叉学科的特殊性，在实践课程设置上往往忽略了不同学科课程之间的联系，学科交叉融合不明显，实践教学内容难成体系，导致学生难以通过实践教学全面理解和应用所学知识[16]。尽管许多课程都要求开展实践教学，如临床类课程会有实践课，有些信息类课程如信息组织、病案管理等有上机要求，但实践内容未能与现实需求相结合。课堂实践、课外实习、毕业设计、大学生创新训练等环节之间的相关性不强[17]。因此，需改进实践课程的横向关联，以及提高学生不同层次的实践能力。为了应对这些挑战，本研究以实践教学目标为导向，建设信息科学、医学、管理科学等多学科交叉的实践课程体系，这将有助于学生更好地理解医学信息学的核心概念。并将数据、信息、知识和智慧有机地连接起来，提高学生的综合素质，使他们更好地适应未来医学信息学领域的挑战。

2.3 实践教学方式重知识轻能力、重传授轻实践

长期以来，我国的传统工科教育一直侧重理论教学，实践教育则通常被视为理论教育的附加环节[18]。有研究指出医学信息学本科教育更注重课堂教学，强调学生在了解掌握医学、信息学、管理学、计算机科学等专业基本知识的前提下，能从事一些简单的医学数据处理、信息系统设计维护等方面工作，但对实践能力要求不高[19]。在实践教学环节中，创新的教学方式如发现式、探究式等并未得到充分的运用，缺乏从实际应用情景出发进行设计，学生在学习过程中兴趣不高，这造成部分同学在课程结束后仍然对于自己专业工作的具体情境模糊不清[20]。

校外实践是实践教学的重要方式之一，但目前医学信息实践教学产教融合程度不够。虽然学院有较多的实践平台，提供了较好的资源。但现有培养模式没有很好地将高校、医院、卫生行政部门和医疗软件企业等多方有机协同，学生没有真正进入实践工作，无法在真正的实践环境中应用所学知识。因此，需要将“专业小课堂”和“社会大课堂”的结合，实现课本与现实相结合、理论与实践相统一，促使学生的自我学习能力、实践操作与应用能力、处理问题能力的全面提升，从而实现应用型人才实践能力的全面培养。

2.4 实践教学评价体系不完善

教学评价是促进教学体系持续改进的重要手段之一，而大部分实践教学的成绩评定主要是以课程为单位，较少有单独针对实践成绩评定及详细的评定标准[21]。与理论教学相比，实践教学的要求通常较低，缺乏淘汰机制。这也使学生对实践课程产生了较低的兴趣和动力。通常情况下，教师按照教学要求在课程中设置实践教学学时，并根据实践成果或者实践报告进行成绩评定。但这种评估方式的评价主体较为单一，不能真正对学生的实践学习过程和实践能力培养的进行衡量[22]。在实习环节中，也面临着对专业实践结果的多方考察和效果评价不足的问题，这使得评价难以发挥对教学的引导和监督作用。从而影响了实践教学的效果，使其无法达到培养方案中规定的培养目标。

3 基于 OBE-CDIO 理念下医学信息人才实践教学体系设计

医学信息专业实践教学的目的是培养学生能以交叉学科的视角，将知识更好的迁移至医学领域来解决相关问题。本文以学习产出为导向的 OBE 理念和以项目为主线的 CDIO 模式应用于医学信息人才实践教学体系的构建，强调从知识教育到目标教育的转变[23]，如图 1 所示。其中，遵循 OBE 理念明确实践教学目标，确定教学内容，通过 CDIO 模式改进实践教学内容，两者相辅相成[24]。首先，OBE 理念下实践目标的确认既要立足社会和市场需求，又要关注学生未来职业发展的能力要求，通过强化实践能力和技能，帮助他们更好的将学科知识与实际问题结合起来，以应对医学信息的复杂环境。并将不同能力分层并细化到各个阶段，形成实践能力培养的递增线。其次，依据预期学习产出，结合医学信息专业特色和行业需求反向设计实践课程内容，实现教学目标和教学内容的追溯式匹配。在教学内容的实施上，依托 CDIO 项目式教学模式，依据“构思-设计-实施-运行”精心设计教学环节，在课堂教学中采用先进的教学方法，以提高学生实践学习兴趣。在课堂之外构建多元化实践教学平台，打破大学封闭型人才训练与社会人才需求的隔离，借由学校与行业的

合作培养模式，让学生所学的理论知识能够真实应用在实际工作场景，营造开放式的交流以及协同创新的教学模式。最后，对实践教学考核体系进行优化，加强对实践过程的考核，全面考量医学信息人才实践教学效果，发现教学过程中的不足，促使实践教学的目标达成和效果提升，形成持续改进、一体化的 CDIO-OBE 相结合的医学信息人才实践教学体系。

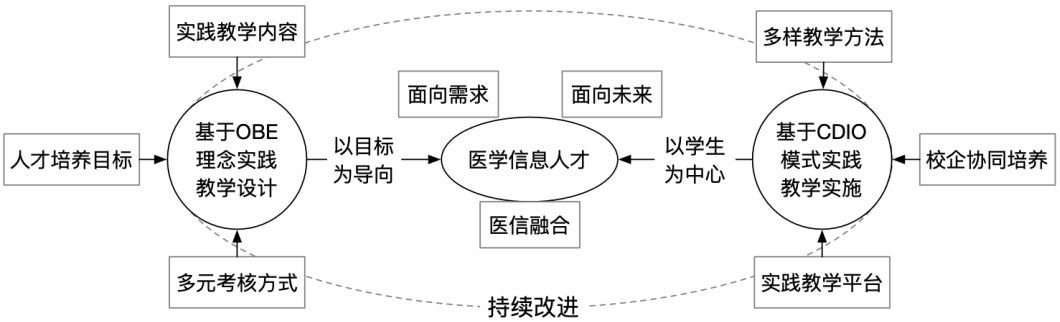


图 1 基于 OBE-CDIO 的医学信息人才实践教学体系框架

4 基于 OBE-CDIO 理念下医学信息人才实践教学体系建设

4.1 以能力为核心，优化人才培养目标

CDIO 培养大纲将学生的能力分为工程基础知识、个人能力、人际团队能力和工程系统能力四个层面[25]。从认知维度来看，能力对维度可以分为记忆、应用和理解、分析与创造[26]。通过实践教学，培养学生的问题识别和分析、理论综合和应用、创新思维和改进、人际交往和团队合作等能力以及在构思、设计、实施、运行项目实践过程中所表现出来的综合能力。我们借鉴以上原则，结合医学信息人才行业需求，将医学信息人才专业实践能力分解为专业知识能力、综合应用能力、创新思维和改进能力，形成人才培养目标的递增线。

其中基础实践能力包括对医学基础知识的了解、一定的计算机编程能力和良好的信息素养，能针对医疗信息领域相关问题，用现代信息技术进行信息收集、加工、处理、分析研究和开发的能力。综合应用能力是能融会贯通的使用不同专业技能，解决实际问题，体现医学信息人才的复合型、交叉性。进一步可分解为问题识别和分析能力、综合素养和跨学科能力、团队协作和沟通能力，培养的人才能在实践过程中敏锐的发现实际工作中存在问题，并能整合和运用不同领域的知识和理论，通过团队协作等方式，深入探讨医疗卫生服务模式的发展。创新思维和改进能力是实践教学的最终目的，培养一批综合实践能力强、科研意识好的高质量医学信息人才，能针对我国医疗信息化领域发展现存的主要问题，提出关键的创新和改进方向。

4.2 以目标为导向，反向设计实践教学内容

从实践教学内容构建上，依据培养目标将学生能力层次映射到实践教学课程设置上，将实践教学的内容体系设置为基础实践教学、综合实践教学、创新实践教学。基础实践教学主要以提高学生多领域的专业知识，包括医学知识如基础医学、临床医学、公共卫生等，信息科学知识如信息管理、信息技术等，管理科学知识如服务管理、项目管理等，并教学中设置相应实践教学内容，将实践教学融入课堂。专业老师的课堂教学的同时，聘请校外导师入校园，开设学习讲座，帮助学生拓宽行业视野，了解专业实践应用领域。综合实践教学以培养学习利用综合利用专业技能解决实际问题的能力，以及团队协作与组织、沟通协调、交流等，主要实践课程通过专业课程实验与实训、实践课程、小组汇报等，以项目形式、团队合作方式开展实践教学，在实践项目的设计上要结合医学信息专业的具体应用情境。除了校内实践，学生将有机会在校外实践平台进行实习，加强理论与实践的联系，扩宽学生视野，巩固和运用课堂教学所掌握的理论知识，提高专业技能，培养分析问题、解决问题等专业应用能力及初步进行科学研究的能力。创新实践教学主要对学生的创新思维进行训练，以实践活动为载体，以学科竞赛和科研项目为抓手，鼓励学生理论联系实际，反思自己的学习，提升创新应用能力。鼓励学生参与大学生创新创业比赛、挑战杯比赛、“互联网+”等赛事、参与科研课题以及撰写毕业论文，创新应用专业知识。多层次的实践教学内容拓展了实践教学的功

能，使学生的创新应用能力和综合素质在完整的实践教学过程中实现递进式的提升[27]。

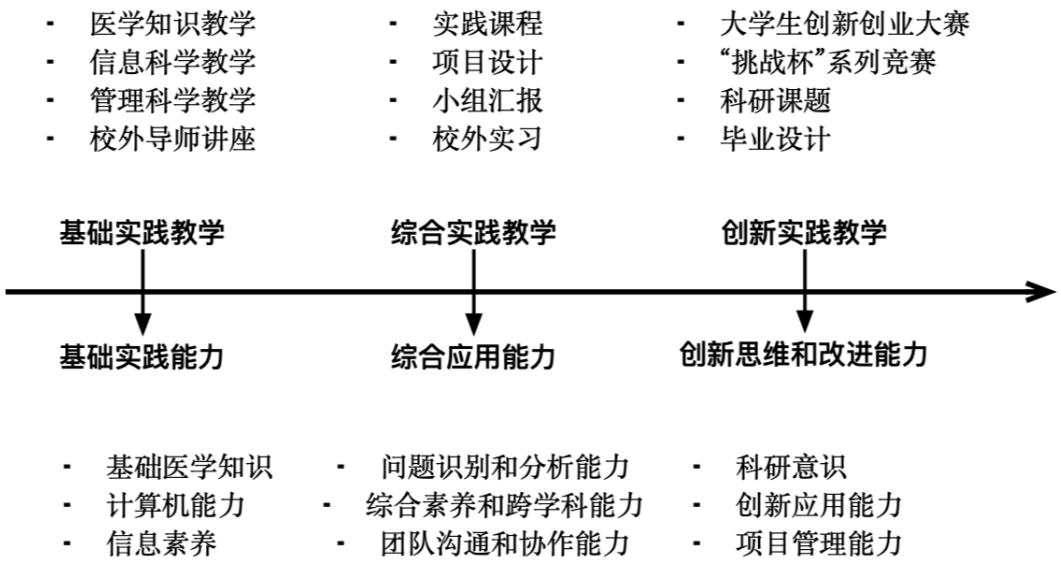


图 2 医学信息人才实践能力层次及实践教学内容

4.3 以学生为中心，丰富实践教学环节

实践教学环节应以学生为中心、以学习成果为导向，围绕教学内容、教学策略等方面进行改革优化。在教学内容设计上，将不同层次培养目标能力要落实到具体教学内容之中，这就要求教师进行层次化的内容设计，使每个学生都能得到最佳的培养。CDIO 理念强调以项目形式开展教学设计，以项目为媒介，激发学生积极参与教学过程的实践活动。因此在教学方法上，鼓励运用问题驱动教学法和项目驱动教学法，将大量实际案例融入课堂教学。以解决实际问题为出发点，设计贯穿课程核心知识点的个人或小组项目，由浅到深、由具体到抽象、由简单到复杂，按照构思、设计、实现、运作四个环节循序渐进组织教学。并将案例教学融入课堂，将“教、学、做”融为一体，实现教学过程与生产过程对接，激发学生的学习兴趣，引导学生主动思考、乐于实践。

4.4 以交叉学科为基础，建设多元化实践教学平台

建设多元化的实践平台，以帮助医学信息人才从不同视角了解医学信息相关领域工作人员的现实需求，也能满足不同学生未来职业发展规划的需要[28]。秉承对学生教育因材施教、分流培养的理念，将教学实践基地分别设置在各级医学图书馆、各级医院信息部门、健康医疗信息相关企业。拉近实践教学与企业、行业之间的距离，提升实践教学内容的专业性和实用性。将学生在学校里学习的理论知识真正应用到医学环境中，在熟悉其工作环境、基本流程和内容的基础上，鼓励学生应用所学知识创新医疗信息服务方式，提升学生的学习积极性和岗位胜任力。

不同的实践平台对于学生有不同的内容和要求。基于各级医学图书馆的实践平台，要求学生熟悉图书馆的全套操作流程以及自动化管理的具体情况，通过实地学习和实践，使他们能够理解中英文全文数据库的收录范围、检索路径和检索技巧，了解文献查找工作的评估标准，并有能力独立提供情报更新、情报分析等情报服务。基于各级医院信息部门的实践平台，要求学生理解各类医疗信息系统的基本架构和功能，熟识实习医院的组织架构和管理流程，包括电子病历和电子病历管理，使他们对医院的信息服务有更深刻的理解。基于健康医疗信息相关企业的实践平台，要求学生熟悉信息系统开发常用方法、常见工具、主要流程，熟悉卫生信息系统主要类型、基本结构、关键模块，以全生命周期为理念，全程参与信息系统设计、开发与实施，以项目管理的形式培养学生的综合实践能力。多元化的实践平台满足了医学信息人才复合型的实践需求，通过实践培养帮助医学信

息人才能够整合健康医疗大数据中的文献数据、调查报告、统计年鉴、电子病历等，以更精确的目标和更具灵活性的模式提供医学信息服务。

4.5 以持续改进为目的，优化实践教学考核体系

考核环节对于对于调动学生的学习积极性、促进教学体系持续改进具有重要意义[29]。实践教学考核要注重对学生整个学习过程的考核和评价，学生的实践教学考核要综合考虑学生的过程表现、实践能力、综合素质等。根据多层次递进式的实践教学目标，建立实践课程多维考核体系，如图 3 所示。对于专业知识能力的考核主要通过课堂考勤、课堂考试等形式进行考核，对学生对理论、技术和方法进行全面评估。综合应用能力考核主要通过项目设计报告、小组汇报、实践报告为主，并增加过程性考核的权重和分值，通过小组互评、教师点评等方式，培养学生发现问题、解决问题的能力。创新实践能力主要通过论文答辩形式展开，并将学生的创新实践成果和科研表现纳入实践考核内容，鼓励学习进行创新活动，并请任课教师、行业教师进行综合考核，从专业和实践角度对学生创新实践能力进行评估。该多元化考核体系，使任课教师、行业教师、学生都成为教学考核的主体，深度展示出学生的知识、技能以及整体素质。同时，实践教学考核不单仅要发挥其引领、检测并评估的作用，也要根据评级结果和建议反馈，推动实践教学体系的持续优化。

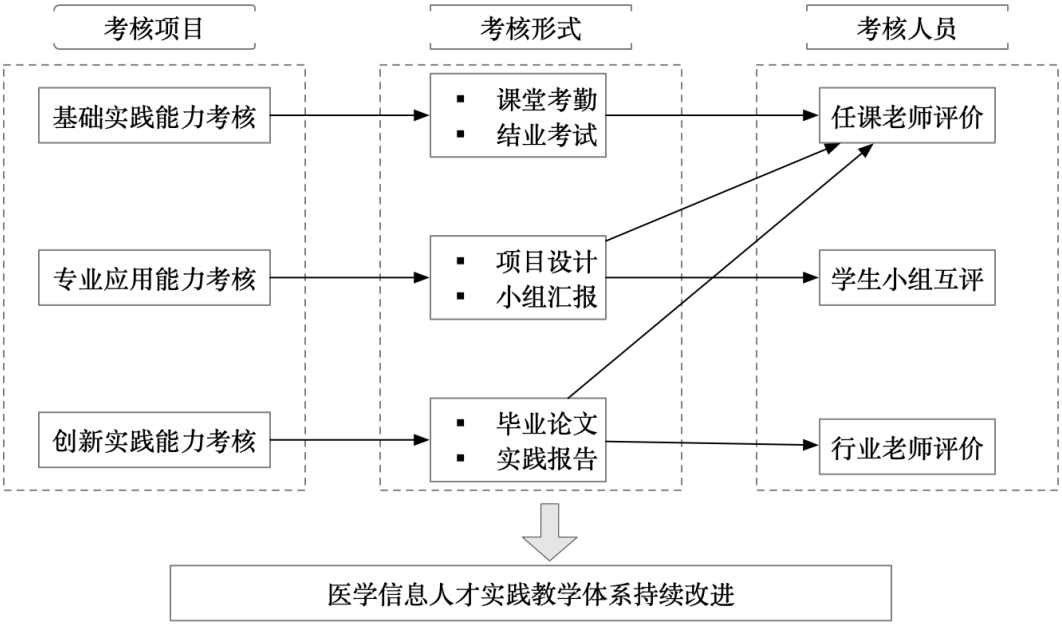


图 3 医学信息人才实践教学考核体系

PU-168

医学信息学未来的发展与展望

李新龙
山西医科大学

1 数据分析与人工智能

随着医疗技术的进步，医学数据的数量和复杂性不断增加。利用人工智能和机器学习技术，将能够更准确地分析和理解这些数据，以改善疾病诊断、治疗和预防。总之，医学信息学与人工智能的结合为医学领域带来了许多潜力和创新。它可以提供高效、准确的医疗服务，并促进个性化医学的发展，从而提高患者的生活质量和健康结果。

1.1 数据分析和预测

医学信息学利用人工智能技术，可以对大规模的医学数据进行分析和挖掘，以发现隐藏的模式、趋势和关联性。通过训练模型和算法，AI 可以预测疾病发展的可能性，辅助医生做出更准确的诊断和治疗决策。

1.2 临床决策支持

医学信息学中的人工智能工具可以为医生提供临床决策支持。基于患者的病历数据、实验室结果和影像学检查，AI 系统可以根据以往的经验和最新的医学研究，提供诊断和治疗建议。这有助于医生更好地理解患者的情况，并提供更加个性化和准确的医疗护理。

1.3 图像分析和识别

在医疗影像领域，人工智能技术可以帮助解析和识别各种医学图像，如 X 射线、CT 扫描和 MRI 图像。通过深度学习和计算机视觉等技术，AI 可以快速、准确地检测异常情况，辅助医生进行影像诊断。

1.4 自然语言处理

医学信息学与人工智能结合，也可以应用于医学文本和语言处理领域。通过自然语言处理技术，AI 系统能够理解和分析医学文献、病历记录和临床笔记等，为医生提供关键信息和知识检索服务。

1.5 虚拟助手和智能机器人

医学信息学中的人工智能还可以催生虚拟助手和智能机器人应用。这些虚拟助手可以与患者进行交流和咨询，回答常见问题，提供健康建议和自我管理指导。而智能机器人可以协助医生进行手术、提供康复训练等医疗操作。

1.6 研究与创新

医学信息学与人工智能的结合也推动了医学研究的进步。通过分析大规模的医疗数据和生物信息，AI 可以发现新的疾病模式、诊断标志和治疗靶点。这有助于加速新药的开发、药物安全性评估和临床试验的进行。

2 基因组学和精准医学

基因组学是研究个体基因组的学科，而精准医学是一种以个体基因组数据为基础，通过深入理解个体的基因变异、环境因素和生活方式等因素，制定个性化的预防、诊断和治疗方案的医疗模式。随着基因组学技术的突破，个体基因组数据的获取和分析将成为常态。医学信息学在整合基因组数据、临床数据和环境因素的基础上，可以实现精准医学，为患者提供更加个性化的治疗和预防方案。总体而言，基因组学的发展为精准医学提供了关键的科学基础和技术支持。通过深入理解个体的基因组特征，精准医学可以实现更加个体化的医疗和健康管理，提高治疗效果和患者的生活质量。

2.1 基因变异的理解

基因组学研究揭示了我们在基因层面的差异。通过基因组学技术，可以检测和解读个体的基因组序列，如单核苷酸多态性（SNP）和结构变异等。这些变异可能与疾病的发生、发展和治疗反应等相关。

2.2 疾病预测和风险评估

通过基因组学的方法，可以发现与特定疾病相关的遗传变异。这可以用于早期疾病风险评估和

预测，提前采取干预措施以预防疾病的发生。

2.3 个体化治疗策略

通过基因组学分析，可以了解个体对药物的反应和代谢特征。这有助于医生根据个体的基因组信息，选择最适合患者的药物和剂量，以提高治疗效果并减少不良反应。

2.4 新的靶向治疗方法

基因组学的发展为精准医学引入了新的治疗方法。通过识别某些疾病相关基因的异常，可以针对这些变异开发特定的靶向药物，以更精准地干预疾病进展。

2.5 疾病机制的研究

基因组学提供了研究疾病机制的重要工具。通过比较正常和疾病相关基因的表达差异，揭示疾病发生机理，有助于开发新的治疗方法和预防策略。

2.6 生活方式干预

基因组学的研究表明，基因与环境互动对健康的影响非常重要。通过基因组学的指导，人们可以了解自己的遗传风险，采取针对性的生活方式干预措施，如饮食、运动和环境调整，从而减少疾病风险。生活上的干预可以说是我们最简单有效的提高我们健康水平的方法。

3 云计算和数据安全

为了应对医疗数据不断增长的挑战，云计算技术将在医学信息学中发挥重要作用。通过云计算，医疗机构可以存储和共享大量的医疗数据，同时实现数据备份和安全性保护，保护医疗数据的隐私和安全。总的来说，云计算和数据安全在医学信息学中发挥着越来越重要的作用。它们提供了便利性、高效性和可靠性，使医疗领域能够更好地进行数据管理和分析。同时，我们也需要关注信息安全和隐私保护等挑战，并采取有效的安全措施来确保数据的安全性和合规性。在医学信息学教育和发展方面，云计算和数据安全技术也扮演着重要的角色。

3.1 数据存储和传输

在医学领域，大量的数据需要被存储和传输，例如医学图像、电子病历等。云计算技术使得这些数据的存储和传输变得更加方便和可靠。通过云存储，医疗机构可以轻松地备份和访问数据，无论是在本地还是远程。此外，云计算还提供了高效的数据传输机制，例如通过云服务器进行医患之间的数据交流，使得远程医疗变得更加便捷。对于医疗机构来说，也能够更方便地共享患者的信息，例如通过云平台实现不同医疗机构之间的信息共享，提高了医疗效率和诊断准确性。

3.2 计算能力

在生物信息学、药物研究和疾病预测等领域，需要进行大量的数据处理和分析。云计算技术提供了强大的计算能力，使得这些复杂的计算任务能够快速而准确地完成。通过云计算平台，医疗机构可以实时处理和分析大量数据，从而更好地预测疾病趋势、研究疾病病因和治疗方法等。此外，云计算技术还支持在虚拟环境中进行实验和研究，例如在药物研发方面，可以在虚拟环境中测试新药的效果，缩短了药物研发的时间和成本。

3.3 信息安全和隐私保护

在医学领域，患者的个人信息和敏感数据需要得到高度的保护。云计算技术提供了多种安全措施，如数据加密、访问控制和身份验证等，以保护患者的个人信息和敏感数据。这些安全措施可以确保只有授权人员能够访问和处理相关数据，从而避免数据泄露和滥用。此外，云计算平台也需要遵守相关的法规和标准，确保数据的合规性。例如，需要遵守 HIPAA（健康保险流通与责任法案）等法规，确保患者个人信息的保护。

然而，与此同时，也需要注意医学信息安全面临的挑战。例如，云计算平台可能面临来自外部的攻击或内部误操作等安全风险。因此，在医学信息学中应用云计算技术时，应充分考虑数据的安全性、隐私保护和合规性。这需要采取有效的安全措施，例如使用强密码、定期更新安全补丁、实施多层次的安全防护等。此外，也需要对云计算服务提供商进行仔细选择，选择具有高安全性和信誉的服务提供商。

3.4 医学信息学教育和发展

云计算和数据安全技术在医学信息学教育中也发挥了重要作用。通过云计算平台，学生可以方便地获取和处理大量医学数据，提高了实践能力和技能水平。此外，云计算平台也提供了在线教育和培训的机会，使得更多的人可以学习和掌握医学信息学的知识和技能。

4 医学伦理和隐私保护

医学伦理和隐私保护在医学信息学中具有重要关系。医学信息学涉及到医疗数据的收集、存储、处理和共享，而这些数据往往包含着患者的个人信息和敏感信息。因此，医学伦理和隐私保护在医学信息学中是不可忽视的重要问题。一方面，医学伦理要求医学信息学在处理患者数据时遵循道德和伦理原则，尊重患者的隐私权。这包括保护患者的个人信息，确保数据的保密性和完整性，避免数据泄露和滥用。医学信息学在设计和开发医疗信息系统时，需要考虑到这些伦理原则，确保系统的设计和实施都是符合伦理标准和患者隐私权益的。另一方面，隐私保护也是医学信息学中的重要问题。患者不愿意公开或透露自己的个人信息和病情，这涉及到隐私权的问题。因此，医学信息学需要采取技术措施和管理措施来保护患者的隐私。例如，采用数据加密和安全传输技术来确保数据的安全性，同时也需要制定相关的隐私政策和程序来确保患者数据的使用符合伦理标准和法规要求。

总之，医学伦理和隐私保护在医学信息学中具有密切关系。医学信息学需要在技术和设计上考虑到伦理原则和隐私保护，确保医疗数据的合法使用和患者隐私权益的保护。同时，医学伦理也要求医学信息学在处理患者数据时遵循道德和伦理原则，尊重患者的隐私权。

4.1 保障数据安全和隐私保护

医学信息学可以采取技术措施和管理措施来保障医疗数据的安全性和隐私保护。例如，采用数据加密和安全传输技术可以确保数据在传输过程中的安全性，防止数据泄露和被非法获取。同时，医学信息学也可以制定相关的隐私政策和程序，规定医疗机构和医护人员对患者的个人信息和敏感信息的处理和使用方式，确保数据的合法使用和隐私保护。

4.2 提高信息管理和决策能力

医学信息学可以通过数据管理和分析，提高医疗机构和医护人员的信息管理和决策能力。通过对医疗数据的收集、整理和分析，可以发现疾病的趋势和规律，为疾病预防和控制提供科学依据。同时，医学信息学也可以通过数据可视化技术，将复杂的数据转化为易于理解的图表和报告，帮助医护人员更好地理解患者病情和做出准确的诊断和治疗决策。

4.3 促进医疗质量和效率提升

医学信息学可以通过医疗信息共享和远程医疗等技术手段，促进医疗质量和效率的提升。通过医疗信息共享，医疗机构可以更好地共享患者的信息，避免重复检查和用药，降低医疗成本和提高医疗效率。同时，医学信息学也可以通过远程医疗技术，让患者在家中就能接受专业的医疗服务和健康管理，缓解医院拥挤和患者看病难的问题，提高医疗质量和患者满意度。

4.4 增强法律和伦理意识

医学信息学可以通过教育和宣传，增强医疗机构和医护人员的法律和伦理意识。医学信息学可以向医护人员普及相关的法律法规和伦理准则，让他们了解自己的权利和义务，遵守相关规定和制度，避免侵犯患者权益和发生不良事件。同时，医学信息学也可以通过案例分析和警示教育等方式，提醒医护人员时刻保持警觉，加强风险意识和责任意识，确保医疗活动的合法性和规范性。

5 讨论

医学信息学在未来将继续发挥重要作用，随着人工智能、医学伦理和数据安全等领域的不断发展，医学信息学也将面临新的挑战和机遇。

首先，人工智能将在医学信息学中发挥越来越重要的作用。医学信息学可以通过人工智能技术，提高医疗数据的质量和可靠性，实现更加精准的诊断和治疗。例如，通过自然语言处理技术，可以将医学文献和病例数据转化为机器可读的形式，为医学研究和临床实践提供更加全面的信息支持。同时，医学信息学也可以通过人工智能技术，实现个性化医疗和精准医学，根据患者的基因、生活习惯和健康状况，提供更加精准和个性化的治疗方案。

然而，人工智能在医学信息学中的应用也面临着一些挑战和问题。首先，人工智能算法的可靠性和准确性还需要不断提高，避免出现误诊和漏诊的情况。其次，人工智能在医疗决策中的透明度和可解释性也需要进一步加强，让医生和患者能够理解和信任人工智能的决策结果。此外，人工智能在医疗数据的安全和隐私保护方面也面临着巨大的挑战，需要采取更加严格的安全措施和技术手段，确保医疗数据的安全和隐私权益不受侵犯。

其次，医学伦理也是医学信息学未来发展的重要问题。医学信息学在收集、存储、处理和共享

医疗数据时，必须遵循医学伦理原则和相关法规，尊重患者的隐私权和合法权益。例如，在医疗数据的收集和处理过程中，需要征得患者的明确同意，确保数据使用的合法性和合规性。同时，医学信息学也需要加强对医疗数据的监管和保护，避免数据泄露和被滥用。在医学信息学的教育和培训中，也需要加强医学伦理的教育和宣传，提高医护人员的道德意识和伦理素养。

最后，数据安全也是医学信息学未来发展的关键问题。随着医疗信息的不断增加和数据价值的不断提升，医疗数据的安全和隐私保护也变得越来越重要。医学信息学需要采取更加严格的安全措施和技术手段，保障医疗数据的安全性和隐私权益。例如，可以采用数据加密、安全传输和数据备份等技术手段，确保数据在传输和存储过程中的安全性。同时，也需要制定更加严格的隐私政策和程序，加强对医疗数据的监管和保护，避免数据泄露和被滥用。

总之，医学信息学在未来将继续发挥重要作用，随着人工智能、医学伦理和数据安全等领域的不断发展，医学信息学也将面临新的挑战和机遇。医学信息学需要在技术和管理上不断创新和完善，同时也需要加强对医学伦理和数据安全等问题的关注和研究，为医疗事业的健康发展做出更大的贡献。

PU-169

智能医学发展与思考

卜庆录
哈尔滨医科大学附属第三医院

新冠肺炎疫情期间，人工智能技术得到快速应用，在辅助诊断等方面发 挥了积极作用，让人们对人工智能医学在医疗领域的应用有了更清晰的认识。 智能医学以现代医学为基础，融合人工智能、大数据、云计算、深度学习等相关领 域的工程技术，探索人体生命奥秘与疾病发 生机理，提升医疗与疾病防控的水平。智能 医学是现代医学发展的新方向，是全球医疗 卫生产业布局的重点领域，更是人类社会民 生健康与福祉的需求。智能医学的发展，需要兼具深厚的数理基础、完备的医学知识、前沿的智能工程技术手段，能够综合 运用所学知识进行临床转化创新的高素质复合型人才。

21 世纪以来，随着第四次工业革命的兴起，大数 据、云计算、人工智能以及机器人等新兴技术发展，促使医学模式面临又一次重大变革，工程医学最终进化 到“智能医学”的阶段。与传统的医学模式相比，智能 医学模式具有以下特点：一是以患者为中心。随着医 疗装备的微型化、网络化、家庭化和智能化，以及可穿 戴设备、互联网、5G 移动通信等技术的发展，医疗的 场景逐渐延展至患者家庭，形成以患者为中心的个性 化的精准医疗和健康管理。由智慧医院、区域医疗中心、家庭自助健康监护终端所构成的三级远程医疗网 络，将逐渐取代原有的以医院为中心的医疗模式，极 大缓解我国现阶段“看病难”“看病贵”等一系列社会 问题，并在疫情等重大灾难发生时，迅速形成有效的 防御和救助体系。二是由大数据驱动。通过可穿戴或 可植入设备，医生可全天候地实时获取患者的生理、 心理数据，再加上患者的各类基因组学、代谢组学、影 像组学数据，以及通过物联网获取的行为学数据等， 在理想情况下，医生可在数字空间形成一个完整的、全方位的患者的数字医疗映像。通过对数字映像的大 数据分析、追踪，可形成针对患者个体的精准医疗方案，通过对虚拟患者诊疗过程的模拟、推演以及药物 试错、疗效评估，我们可极大降低误诊的风险，提高疾 病的治愈率。同时，针对某一特定的疾病，可通过汇聚 海量的患者数据，利用数据挖掘的手段，从中找出隐 含的、以往未知的、具有潜在医学价值的信息，从而揭 示疾病的发生发展规律。尤为有趣的是，医疗大数据 的挖掘，有可能突破人类思维的局限，从意想不到的角度启迪新的医学发现。三是人机协同。在智能医学 阶段，人工智能将发挥越来越重要的作用，协助人类 医生完成海量医学数据的处理、手术机器人的操控 等，从而将人类医生逐步从低端的重复性、重体力劳 动中解放出来，转而从事更为高端的综合决策、复杂 病例研究等工作。因此，如果说工程医学的目标是应 用工程技术手段辅助诊疗的话，那么在智能医学阶 段，将形成一种新型的人机协作的形式，完成疾病的 智能诊断和精准治疗。

人工智能医学在医学中的应用场景众多，例 如：健康管理、辅助诊疗、风险预测、药物 挖掘、医院管理。智能与信息化医院，为患者提供预约、 分诊、结算、随访等线上一站式服务；构建 相匹配的网络安全、系统稳定、数据安全等 体系；医院内部信息系统集成整合，医疗数 据统一管理与共享互用；医院实施电子病历 系统，探索与居民电子健康、公共卫生服务数据互融互用。

互联网医院是指实体医院、智能与信息 化医院以及互联网融为一体的功能平台，它 涉及到区域与国家范围内院际之间的医疗信 息互动互联、检查结果的互享互认、线上诊断、 咨询与药物配送、远程服务、医保异地结算 以及分级诊疗健康的新模式。通过互联网医 院的建设，不断地提升医疗质量与效率，改 善医疗资源配置不平等、不平衡，降低医疗 成本，提供便捷优质的医疗资源和医疗服务， 增加老百姓的获得感，为老百姓带来切实的 健康利益。智慧医疗，是现代信息技术、人工智能、 大数据、互联网、机器人等与医学深度融合、 交叉发展的时代产物，是新型医疗模式与医 疗手段。它涉及到个人健康管理与医疗信息。

我国智能医学发展路径和可探索空间广阔，建议从以下方面着力建设：

一、着力自主研发创新，加速智慧医疗全面融合。 搭建一体化健康医疗大数据平台，助推分级诊疗体 系。健康医疗大数据是国家基础性战略资源，是智 能医学发展的基石，通过对医疗大数据进行集成分 析和归类，有助于为疾病诊断或治疗手段迅速提供 快速优化解决方案。

二、加速医疗信息互联互通，推广互联网医院建设。 以实体医院为主体，借助“互联网+医疗”创新成果， 建设智能医院，同时需要国家大力推动院际医疗信息 互联互通、检查结果共享互认、医保异地结算机制，助 力完善分级诊疗模式，有利于医疗质量与效率的提升，改善医疗资源分配不平衡，避免医疗资源过度浪 费。在此基础上，借助互联网、物联网、人工智能技术 和信息化手段，推广建立互联网医院，实现在线诊断 及药物配送并接入医保，降低医疗成本，切实提供便 捷、优质的医疗资源，增加老百姓获得感。同时建立 大众化与智能化的医疗服务体系，实现优质医疗资源 示范和中国医改模式转型，真正推动中国医疗改革。

三、加强学科支撑建设，打造人才培养高地。人才 资源是第一资源，是发展智能医学的关键。智能医 学及智能医疗产业的发展需要创新实践能力强、具 备国际竞争力的医工交叉复合型人才。在新时代新 医科发展布局下，需要国家推动智能医学学科发展 和专业人才队伍建设，逐步打通医工高端复合型人 才培养的壁垒，为智能医学科技创新及产业发展提 供人才保障，使医工复合型人才能够真正在临床上 发挥作用，推动智能医学发展。

四、完善法律规范及标准化建设，强化制度保障。 随着智能医学的发展以及智能医学新业态的不断涌 现，管理规范和法律制度亟待完善，为智能医学的在我国健康平稳 发展铺平道路，抢占国际智能医学发展的主导权。 智能医疗设备准入和安全性认证需要建立审批与评 判标准，划出法律和伦理道德的边界，打造兼顾科技 与人文的“中国智造”；健康医疗大数据需要从国 家层面加强科学布局和监管，保障数据安全性、有效 性和可控性。

五、 建设战略谋划载体，营造良好生态。智能医学 的发展需要战略性载体，以整合行业优质资源，有利 于促进智能医学科学技术的创新和发展、促进智能 医学科技的普及与推广、促进智能医学科学人才队 伍的发展壮大、促进国产化智能制造、促进国内外社 会组织及学术组织交流合作，促进智能医学与经济 建设相结合。有助于国家在智能医学领域的战略布局， 有组织地进行价值引领、道德规范、创新协同、健康 发展，推动我国在全球科技人文治理方面发挥主导 作用，增加主动权和国际话语权。

目前，我国在人工智能医学应用领域有比较 好的研发基础，为进一步推动该领域发展，充分利用我国在 5G 通信技术、医疗数据丰富等方面的优势，并补足短板，驱动智慧医疗新业态变革。加强人工智能医学应用所需的数据标准建设， 建立并完善医学数据收集、标识、分类、存储、训练、 验证、应用、评价、共享和互操作的标准体系；加 强监管科学的研究投入，针对人工智能医疗产品的 特点，明确相关产品、技术和服务的监管要求，制 定审批标准和指南，并根据技术和行业发展程度及 时更新相关标准和指南，促进人工智能医学应用的 可持续发展。

PU-170

医疗信息教育与人才培养

谢添
郑州大学第一附属医院

引言:

随着医疗信息的爆炸式增长，医疗信息教育在当今社会显得尤为重要。它不仅有助于提高医疗保健行业的效率和质量，还有助于推动医学研究的进程。医疗信息教育旨在培养具有医学知识、信息技术和数据科学技能的复合型人才，以满足社会对跨学科医疗信息人才的需求

1. 中国医疗信息教育与人才培养的现状

1.1 医学信息继续教育的重要性

继续教育是面向学校教育之后所有社会成员特别是成人的教育活动，是终身学习体系的重要组成部分[2]。 医学继续教育是针对在职的医务人员进行培训的教育方式。 医学教育是一种精英教育，它以人体为研究对象，以生命科学为主导，综合运用多种分析方法，最终目的是造福于人类社会。 而人体的深奥博大和现代信息技术的迅猛发展决定了当代医学教育必须在高等院校的一次性教育的基础之上，跟随时代的变化、知识的累计不断在探索中向前发展，是一种终身教育。 因此，继续教育已成为医学教育领域中不可缺少的组成部分。

1.2 医学信息继续教育的发展趋势

在保持医学教育本身内在具有的整体性和关联性情况下，医学与其他学科和领域的相互交叉融合已成为当前学科发展的一个重大趋势，医工、医理、医文等“医学+X”交叉学科支撑的“新医科”应运而生。在解决复杂生命问题的过程中，既要发挥各学科的优势，又要综合各学科的知识，往往能在破解重大技术难题、攻克疑难杂症方面取得重要进展和突破。

1.3 医学信息继续教育的挑战与问题

现代科学技术的迅猛发展,对生物医学起到了巨大的推动作用，电子计算机、生物工程、纤维光学、激光技术、超声技术、新材料技术等科学技术在医学领域的渗透和应用，使各个学科都取得了惊人进展，医学知识与信息激增,医学的知识体系随之扩增，医学教育的专业设置、各专业设置的课程门类也随之增多，高等医学教育内容处于高度的动态变化之中，这给医学生增加了学习难度，也给医务工作者带来了新的挑战。必须顺应时代变化，时刻学习，终身学习。

2.中国医疗信息教育与人才培养的发展方向

2.1 培养模式的创新与灵活性提升

中国救死扶伤的医德发展是几千年的传统医德，有坚定的医德基础。随着社会的快速发展，医学的定义不单单是从前的单一的定义，医学信息化，多学科的发展，医院向综合性的发展，我们也要进行革新。以新思想，新社会需求，以国家发展要求进行新的革新[3]。

——以新理念谋划医学发展。将医学发展理念从疾病诊疗提升拓展为预防、诊疗和康养，加快以疾病治疗为中心向以健康促进为中心转变，服务生命全周期、健康全过程。

——以新定位推进医学教育发展。以“大国计、大民生、大学科、大专业”的新定位推进医学教育改革创新发展，服务健康中国建设和教育强国建设。

——以新内涵强化医学生培养。加强救死扶伤的道术、心中有爱的仁术、知识扎实的学术、本领过硬的技术、方法科学的艺术的教育，培养医德高尚、医术精湛的人民健康守护者。

2.2 增加医学信息继续教育的投入与资源支持

依据现代化社会发展要求，以社会人们群众需求为基础，以未来发展趋势进行规划，以国家发展计划相结合进行。

一是加强组织领导。有关部门要进一步加强医学教育综合管理和统筹，协调解决医学教育创新发展有关问题。各地、各有关部门要把医学教育创新发展纳入本地区经济社会发展规划和本部门重点工作计划。

二是实施国家重大战略工程。统筹各方资金资源，加强对医学教育投入保障，支持国家及区域院校医学教育发展基地、一流医学院、高水平公共卫生学院、医药基础研究创新基地、国家住院医师培训示范基地、继续医学教育信息化等建设，中央预算内投资加大对医学院校支持力度。

三是保障经费投入。积极支持医学教育创新发展，根据财力、物价变动水平、培养成本等情况合理确定并适时调整医学门类专业生均定额拨款标准、住培补助标准，地方各级人民政府按照规定落实投入责任。

2.3 建设高水平的师资队伍

强化医学教育统筹管理。支持行业学（协）会参与学科专业设置、人才培养规划、标准制修订、考核评估等工作。深化综合性大学医学教育管理体制改革。遵循医学教育规律，完善大学、医学院（部）、附属医院医学教育管理运行机制，保障医学教育的完整性，加强对医学教育的组织领导。[4]建立健全组织机构，强化对医学教育的统筹管理，承担医学相关院系和附属医院教学、科研、人事、学生管理、教师队伍建设、国际交流等职能。教育部、国家卫生计生委要组织开展综合性大学医学教育管理体制改革试点，在国家改革建设重大项目上对试点高校予以倾斜支持。

3.案例分析与经验借鉴

3.1 某医院医学信息继续教育的成功经验

1.制定个性化的教育计划：医院了解到，不同医务人员的学习需求和能力是不同的，因此他们制定了个性化的教育计划。计划根据每个人的专业领域、职位和技能水平进行定制，确保每个人都能得到最适合自己的教育。

2.采用多种教育方式：医院采用了多种教育方式，包括在线课程、研讨会、实践操作训练[5]等。这种方式可以使医务人员根据自己的时间和学习方式进行选择，提高了学习的效率和效果。

3.注重实践操作：医院明白，医学不仅需要理论知识，更需要实践技能。因此，他们在继续教育中强调实践操作，让医务人员有机会在真实的环境中锻炼技能。

4.建立学术交流平台：医院搭建了一个学术交流平台，鼓励医务人员分享自己的经验和研究成果，从而促进了学术交流和知识传播。

5.结合临床实际：医院继续教育的课程和主题紧密结合临床实际，确保教育的内容对医务人员的日常工作有直接帮助。

6.定期评估效果：医院定期对医学信息继续教育项目进行评估，了解其效果和影响。这样可以帮助他们识别问题，不断改进并提高教育项目的质量。

7.提供技术支持：医院提供必要的技术支持，确保继续教育的顺利进行。他们也为医务人员提供相关软件和工具的培训，帮助他们更好地利用技术来促进学习和实践。

通过以上七个方面的实施，该医院成功地开展了医学信息继续教育，提升了医务人员的专业知识和技能，从而提高了医疗服务的质量。这也为其他医院提供了有益的参考和启示。

总结起来，该医院的成功经验在于他们个性化地考虑了每个医务人员的需求[5]，采用了多元化的教育方式，注重实践操作，提供了学术交流平台，紧密结合临床实际，定期评估效果以及提供必要的技术支持。这些措施共同构成了个全面、系统化的医学信息继续教育方案。

未来，该医院将继续优化他们的医学信息继续教育项目，探索更多的教育方式和手段，以满足医务人员在技术和信息方面的不断增长的需求。他们也希望通过分享自己的成功经验，帮助更多的医疗机构提升医学信息继续教育的质量和效果。

3.2 某高校医学信息教育与人才培养的模式创新

随着医疗技术的快速发展和信息化程度的不断提高，医学信息教育在人才培养中的地位日益重要。为了适应这一趋势，某高校积极探索医学信息教育与人才培养的新型模式，取得了显著成效。本文将从创新教育理念、优化课程设置、强化实践教学、深化产教融合、个性化人才培养、信息技术应用和综合素质提升等方面对这一模式进行介绍和分析。

1.创新教育理念

高校在医学信息教育方面积极推行创新教育理念，强调以学生为中心，以能力培养为导向，注重培养学生的创新精神和实践能力。通过开设跨学科课程、鼓励学生参与科研项目、提供丰富的课外学习资源等方式，为学生提供全方位的学习支持[6]。

2.优化课程设置

在课程设置方面，该高校紧密结合医学信息专业的特点，开设了一系列具有前瞻性和实用性的课程，包括医学信息理论、医学图像处理、医学数据挖掘、医学信息系统开发等[7]。同时，通过课程整合和内容更新，确保课程与行业发展的前沿技术紧密结合。

3、强化实践教学

为了提高学生的实践能力和解决问题的能力该高校加大了实践教学的力度。通过开设实验课程、组织学生参与项目开发、提供实习机会等方式，让学生在实践中学习，提高自己的技能。此外，学校还与多家医疗机构合作，为学生提供实习和就业机会。

4、深化产教融合

该高校积极推进产教融合，与多家医学信息领域的知名企业建立了战略合作关系，共同开展课程建设、师资培训、科研合作等工作。通过与企业合作，学校能够更好地了解行业需求，调整人才培养方案，为企业提供符合需求的高素质人才。

5、个性化人才培养，高校重视学生的个性化发展，根据学生的兴趣和能力进行个性化培养。通过开设选修课程、提供课外辅导、组织学术竞赛等方式，让学生在专业学习的同时，能够发挥自己的特长和兴趣，实现全面发展。

6、信息技术应用

在医学信息教育中，信息技术的重要性不言而喻。该高校注重培养学生的信息技术应用能力，通过开设相关课程、提供计算机实验室、组织编程竞赛等方式，提高学生的计算机水平和应用能力。此外，学校还积极引进先进的技术和工具，不断更新教学资源和技术设施，为学生提供更好的学习条件。

7、综合素质提升

医学信息专业不仅要求学生具备扎实的专业知识和技能，还需要具备较高的综合素质。该高校注重培养学生的综合素质，通过开设人文课程、组织社会实践、提供心理咨询等方式，提高学生的文化素养、社会责任感和心理素质。同时，学校还鼓励学生参与志愿服务、科技创新等活动，培养学生的领导力、团队协作能力和创新能力。

综上所述该高校通过创新教育理念、优化课程设置、强化实践教学、深化产教融合、个性化人才培养、信息技术应用和综合素质提升等方面的努力，构建了医学信息教育与人才培养的新型模式。这种模式能够有效提高人才培养质量，满足行业需求，为学生的职业发展和未来的职业生涯奠定了坚实的基础。

结论：

中国医疗信息教育与人才培养在医学信息继续教育方面的发展仍面临一些挑战与问题，但也存在着巨大的发展潜力。通过创新培养模式、加大资源投入、优化师资队伍、推动跨学科合作等策略，我们可以进一步推动中国医疗信息教育与人才培养在医学信息继续教育领域的发展，以满足医务人员不断更新知识的需求，提高医疗服务质量，推动医疗信息技术的创新与应用。

PU-171

浅析医学信息学发展和教育

孙楠
郑州大学第一附属医院

1 医学信息学在新时代的发展壮大

1.1 医学和信息学改变我们生活

进入 21 世纪，人类战胜了自然灾害、疾病、人口及落后的生产力等一系列困扰我们很长时间的世界性难题。

在这众多的难题中，疾病无疑是对人类危害最大影响最深远的，医学的不断进步帮助我们战胜疾病的困扰，而信息在这场漫长的战役中起到了关键性的作用，它在有形的和无形的医学领域帮助我们，帮我们分析和记录着我们生活中哪怕最不起眼的琐事，为临床医师治疗疾病，解除我们的痛苦提供数据支持。

1.2 医学信息学的专业领域和需求

医学信息学，这门传统的信息学学科，在新的时代也有了对其新的理解，它所涵盖的专业领域知识变的更加广泛：

韩国光州女子大学医学信息学系,学制为四年。教学内容按其课程内容分为 5 大类

①医学基础方面的课程有医学用语、解剖学和病理学、公共保健学、急救及处置等

②医院行政管理方面的课程有医院管理体制、保健行政学、病案管理、医疗财会、人事管理等

③医疗制度及相关法规方面的课程有医疗制度、医疗相关法规、医疗保险等

④信息处理技术方面的课程有情报学入门、电子病案、医院计算机系统、医学信息学、医药信息管理、疾病及手术分类、癌症登记、医院统计学、网络信息处理、医学信息通讯、信息处理应用、医学信息开发等

⑤计算机技术方面的课程有电算学基础、数据库、数据库开发、程序设计人门、程序设计语言C++、VB 程序设计、网络程序设计、系统分析及设计、主页制作等。就业去向主要有大学医院、药局、全民健康保险机构、医疗情报机[1]

从这份韩国大学医学信息学专业的课程内容我们不难看出，医学信息学已经演变成了一门集医学、管理学、公共卫生学、信息学和计算机学的复合型学科，学生在学校接受到领域如此广泛的课程教育，为的就是适应在新时代下对医学信息复合型人才的职业素养需求，不夸张的说，培养一个合格的医学信息学人才，就好比培养一个即会打篮球又会踢足球，同时还要精通羽毛球和网球运动员，他们所要掌握的知识和技巧，不是其中任何一个领域独自就能满足需求的。

接下来我们再来看看医学信息学人才在社会中的职业分布：

美国医学信息专业的就业情况包括首席信息执行官,首席医学信息学家、首席信息安全官、系统分析家、部门领导、卫生信息技术顾问、市场项目和执行经理、企业家等。由此,不难理解为什么许多院校对学生都要求有很强的交流能力,因为他们是管理者,而不是单纯的技术人员。当然,还有一些基本工作,如程序设计员、站点管理者、软件工程师、系统工程师、项目管理者、技术销售员等,分布在工厂、教育、卫生、政府、军队等部门。[2]

可见，由于专业领域知识广泛，社会对医学信息人才的需求也不局限于医院或者医疗单位，他们所能从事的职业分布于各个领域，这就是新时代对复合型人才的需求。

2 医学信息学的教育和人才培养

2.1 我国医学信息学教育现状与进展

和国外医学信息学发展相比，我国的医学信息学发展较晚，相对还比较落后，而医学产业却发展较早，随着医院信息化建设的发展，医学信息在医疗过程中的作用越来越明显，引起了医疗单位和企业的重视，为医学信息学的未来发展奠定了坚实基础。但学术上的落后和产业上的飞速发展，缺造成了医学信息学人才的严重缺口，其中掌握多领域知识的符合型医学信息人才就更是凤毛麟角了，我国医学信息学教育与国外相比差异较大,其主要根源在于国外的医学信息学的起源与我国不同,从而导致培养层次、课程设置、培养目标、师资水平等的一系列差异。由于历史原因，我国医学信息学的教学工作起步较晚，最初进入医学信息领域的计算机工程技术人员只从事科研与开发，没有从事医学信息学的教学工作，而在医学文献检索、医学文献信息处理领域工作的学者首先尝到了应用计算机带来的方便，他们也在医学文献信息教学的第一线，因此，医学文献检索教学率先进入了我国的医学信息学的教学领域。然而专业上的复合性和多元性，也造成了医学信息学人才的培养成本不断提高。很多大学即使建立了该专业，也只能从医学或者信息学为主的院系开设课程，而单一专业水平强的教师往往很难掌握多领域的授课知识，这就造成了培养出来的学生“专而不博”而要做到“即专又博”又要求学生分阶段攻读不同学科专业，时间成本过高。国内的大学毕业生面临着“毕业即失业”的压力，大部分很难静下心来再次攻读不同学科或者继续留校深造，所以这样的医学信息学教育现状，是很难培养出合格的复合型医学信息人才的。

经过近十年的发展，我国的医学信息学教育也有了长足的进步，在医学院校从事计算机教学的教师，慢慢地体会到计算机教学与医学相结合的重要性，他们也逐渐进入医学信息学领域编著了数量不少的“医学计算机与信息技术”方面的著作，他们人数众多，虽是这一领域的后来者，但有着后来居上的势头。各大院校和医疗单位也找到了解决培养医学信息人才的方法，在多次尝试和研究中发现，把医学和信息学整合到一起，形成一门新的学科，让学生从信息数据的角度了解和分析医学管理的知识，学会如何让信息为医疗服务。

2.2 加强医学信息人才培养的方法

2.2.1 加强医学信息学课程改革和教育理念

医学信息学发展的基础是我国卫生领域的信息化，而高校对医学信息学的课程设置又直接决定了学生的知识结构和就业方向，因此要以医疗信息化为根据，依照我国卫生领域信息化建设和发展的需要，紧跟国际医学信息学发展的脚步，同时结合自身的实际情况，设置并及时调整医学信息学专业课程，如开设相关的统计学和信息管理学课程，还应该随着计算机技术的发展对计算机信息方

面的相关课程进行内容的扩展和加深。教育理念要从单一学科向复合型学科转变，加强学生的多向思维能力，用不同学科知识去解决问题，把医学信息学发展成涵盖多领域教育的综合性学科。

2.2.2 发展医学信息学相关内容的继续教育

培养医学信息人才是一个漫长的过程，人才稀缺问题很难在短时间内得到解决，这就需要医疗单位根据自身的情况发展继续教育事业，为现有的职工进行相关的医学信息培训。随着医院信息化方面的继续教育不断发展，医学信息学的继续教育也分可分为三个方向：

第一个方向是计算机学人才，相对于医学，计算机学发展较快，近年来计算机学也成了热门学科，每年毕业的计算机类学科人才远高于其他学科，而且计算机是个应用广泛的学科，几乎没有哪个科学领域离得开计算机提供的分析和存储功能，所以医院在选择继续教育的对象的时候可以优先考虑计算机人才。

第二个是信息管理学，信息管理学也是个应用领域广泛的学科，它涵盖了计算机学和管理学的相关知识，在医学信息学继续教育方面有它的先天优势，培养出的人才可以胜任医院管理、人事和财务管理等职位。

第三个培养方向是医学人才，相对于前面两种人才培养方向，要把医学人才培养成医学信息人才，付出的成本会高一些，但通过这个方向培养出来的人才，医学知识扎实，又兼顾管理学，无论是成为临床医师还是医院管理人员，都有他坚实的专业知识基础做后盾。

在培养方式上，除了常规的讲授形式外，还可以结合医院自身条件，进行岗位循环等形式，让医学信息人才接触医院更多环节的工作，在适应工作的过程中达到完善知识结构、提高自身素质的目的。轮岗还可以使新近人才快速了解医院流程、医院部门建制，学习各类医疗信息化软件处理方法，并可以在与各部门医务人员交流中了解医院的需求，了解医疗人员对医学信息的的需求、培养沟通能力等。

结束语：医学信息学在我国还处在一个发展阶段，但必将成为我国未来医疗卫生信息化建设的重要组成部分。作为医学信息人才人才，要了解到医院每个部门对医学信息的需求形式和内容，解决医疗过程中出现的各种信息方面的问题，为医院的发展提供正确信息支持，也从一个侧面决定了医院的发展方向。我国的医学信息学还需要很长时间的探索和发展才能赶得上国外同行业学科取得的成绩，但只要我们明确了医学信息人才的培养方向，并且坚持不懈的走下去，有一天一定可以跟上国际的脚步。

PU-172

基于现状的基层全科医生医学信息继续教育发展对策研究

孙翔飞，于琦
山西医科大学管理学院

作为三级分级分类诊疗体系的底层基石，基层全科医生在基层医疗服务中扮演着极为重要的角色。他们直接接触病患，负责基础医疗服务，因此他们的专业水平和综合素养对于基层医疗体系的运转和卫生健康工作的推进至关重要。

1 基层全科医生的重要性

1.1 提供全方位的医疗服务

基层全科医生具备较为全面的医学知识和技能，能够为患者提供全方位的医疗服务，包括初步诊断、治疗、健康咨询等，为患者提供全程健康管理。

1.2 疾病筛查与早期干预

基层全科医生常常是疾病的最早接触者，通过及时的疾病筛查和早期干预，可以有效地控制疾病的发展，提高治愈率和生存率。

1.3 建立稳定的医疗关系

基层全科医生通常会长期服务于同一社区或者病患群体，建立了稳定的医患关系。这种关系有助于医生更好地了解患者的病情和需求，也使患者更有信心和依赖感。

1.4 减轻大医院压力

基层全科医生在基层医疗机构工作，可以处理一些常见病、多发病，减轻了大医院的就诊压力，使大医院资源得以更加集中用于处理疑难杂症和重症患者。

2 对基层全科医生加强医学信息继续教育的必要性

2.1 医学知识不断更新

医学领域的知识在不断更新，新的疾病、治疗方法和技术层出不穷。加强医学信息继续教育可以保证基层全科医生掌握最新的医学知识，提高他们的临床水平。

2.2 提升应对突发状况的能力

基层全科医生通常会面对各种各样的医疗问题，包括急诊情况。通过继续教育，他们可以学习到最新的急救和处理方法，提升在突发状况下的应对能力。

2.3 拓展诊疗技能和服务范围

继续教育可以使基层全科医生学习到更多的诊疗技能和服务方法，拓展其在基层医疗服务中的能力范围，提供更加全面的医疗服务。

2.4 促进医患沟通和健康教育

继续教育不仅仅包括医学知识，还包括医患沟通技巧和健康教育知识。这对于建立良好的医患关系，提高患者健康意识至关重要。

2.5 提升职业满意度和幸福感

通过继续教育，基层全科医生可以不断提升自己的专业水平，提高诊疗成功率，这将带来职业成就感和幸福感的提升。

3 全科医生医学信息继续教育发展现状

3.1 开展情况

近年来，随着《“健康中国 2030”规划纲要》的不断推进落实，医学信息继续教育稳步发展，已经取得了一定的成就。国家和地方政府、医学院校、医疗机构等都纷纷加大了对医学信息继续教育的投入，建立了一系列的培训机制和平台，为基层全科医生提供了学习的机会。

同时，许多医学院校和医疗机构也积极组织了各种形式的医学信息继续教育活动，包括学术讲座、病例讨论、临床实践等，为基层全科医生提供了丰富多样的学习资源。

然而，也不能忽视的是，目前我国医学信息继续教育存在一些亟待解决的问题，比如：

首先，医学信息继续教育的内容和形式还不够多样化，缺乏针对性和实用性。

其次，医学信息继续教育的培训机构和师资力量需要进一步加强，提升教育质量。

最重要的是，医学信息继续教育的管理和监督机制还不够健全，需要加强对培训活动的评估和监控。

3.2 基层全科医生对现有医学信息继续教育的看法

基于已经开展的医学继续教育，笔者对山西省基层全科医生的服务现状进行过系统地调查[1]，在调查对象对目前已开展课程总体较为满意的前提下，医学信息相关课程的评分相对较低，其中卫生信息管理部分为所有课程中总体评分最低的课程。

表 1 调查对象对于目前全科医生转岗培训体系中医学信息相关课程的看法评分（均数±标准差）

名称	低分组(n=201)	总体(n=723)
健康管理及慢性病管理	2.97±1.15	4.391±1.109
卫生信息管理	2.38±0.98	4.141±1.305

4 全科医生医学信息继续教育发展对策

医学信息继续教育是基层全科医生在获得基本医学知识和技能后，通过学习最新的医学知识、临床经验等手段，不断提升自己的医学水平，以适应医学科技的发展和医疗模式的更新。它对于全科医生的发展至关重要，也是保障患者健康和提升医疗服务质量的重要保证。下面将从多个方面详细展开医学信息继续教育的发展对策。

4.1 整合健全医学继续教育平台

目前，各省、区、市乃至地级市的医学继续教育平台虽都有不同程度的建立和发展，但课程规划等均不统一，整合统一建立全国规模的大型医学信息继续教育的学习平台是实现医学信息持续学习的重要基础。这一平台应包括以下几个方面：

电子学习平台： 创建一个线上的医学信息学习和继续教育平台，包括课程资源、视频讲座、病例讨论等，方便基层全科医生随时随地进行学习。

实时更新机制： 确保平台的内容能够及时更新，反映最新的医学研究成果、临床实践经验等，保证基层全科医生接触到最新的医学知识。

个性化学习推荐： 根据基层全科医生的兴趣和学习需求，为其推荐相应的医学信息内容，提升学习的针对性和效果。

4.2 丰富多样的学习形式

医学信息继续教育的学习形式应该丰富多样，包括：

线上课程和线下培训相结合： 组织线上课程和线下实践相结合的培训活动，使基层全科医生既能学到理论知识，又能实践应用。

临床病例分享： 组织基层全科医生分享实际临床病例，让大家从实践中学习，提升临床诊疗能力。

4.3 提供个性化的学习计划和指导

针对不同基层全科医生的个性化需求，可以采取以下措施：

制定个性化学习计划： 根据基层全科医生的职称、专业领域、兴趣等，为其制定个性化的学习计划，突出重点和补充短板。

专业导师指导： 为基层全科医生提供专业导师，指导其在特定领域的学习和实践，提升专业

能力。

4.4 推动学术交流和合作

学术交流和合作对于医学信息继续教育来说至关重要：

举办学术研讨会和讲座： 组织医学领域的专家学者举办研讨会和讲座，分享最新的医学研究成果。

建立医学交流平台： 创建医学交流社群或平台，让基层全科医生能够相互交流经验、讨论疑难病例等。

4.5 加强实践和培训

提供丰富的实践和培训机会，使基层全科医生能够在实践中不断提升自己的医学水平：

提供临床实践机会： 给基层全科医生提供在高级医疗机构别参与实际临床工作的机会，提升实践经验和应用能力。

举办技能培训班： 组织针对特定技能和手术的培训班，提高基层全科医生在实际操作方面的水平。

4.6 定期评估和反馈

为了保证医学信息继续教育的效果，可以采取以下措施：

定期考核和评估： 设立定期的考核机制，评估基层全科医生在医学信息继续教育方面的学习成果和综合能力。

提供个性化反馈：根据评估结果，为基层全科医生提供个性化的学习反馈和建议，指导其在学习方向上的调整和改进。

4.7 政策支持

制定明确的医学信息继续教育政策： 政府应该明确医学继续教育中医学信息相关内容的政策，明确培训内容、培训对象、培训时长等方面的要求。

设立医学继续教育基金：政府可以设立专项基金，用于资助医学信息继续教育的培训活动，提供资金支持。

优化医务人员职称评定机制：建立科学、合理的职称评定机制，将医学信息继续教育的学习成果纳入职称评定的考核范围。

建立奖惩机制：对积极参与医学信息继续教育的医务人员给予奖励，同时也要建立相应的惩戒机制，鼓励医务人员积极参与学习。

推动医学信息继续教育的法律法规完善：保障医学继续教育机构的合法性和权益，明确培训机构的资质要求、教学内容的合法性等。

4.8 资源保障

提供学习资源：政府可以向医疗机构、医学院校提供医学图书、期刊、数据库等资源，丰富医务人员的学习材料。

建立培训机构：政府可以引导和鼓励各类医学培训机构的发展，确保培训机构的合法性和教学质量。

提供培训经费支持：政府可以为医务人员参加医学信息继续教育提供一定的培训经费，用于报名费、学习材料等方面。

优化医疗资源配置：合理配置医疗资源，提高基层医疗机构的设施和技术水平，为医务人员提供更好的学习和实践条件。

建立专业指导团队：政府可以组建专业的医学信息继续教育指导团队，为医务人员提供专业的指导和培训。

推动企业参与培训：鼓励医疗机构和企业开展医学信息继续教育培训，提供更多的学习机会和资源支持。

4.9 加强监管和评估

建立医学信息继续教育质量评估体系：设立医学信息继续教育的质量评估机构，对培训机构和培训内容进行评估，确保培训质量。

加强对培训机构的监管：建立健全医学信息继续教育培训机构的审批和监管制度，严格审核培训机构的资质。

定期检查和考核： 对基层全科医生参与医学信息继续教育的学习情况进行定期检查和考核，督促基层全科医生积极参与学习。

5 总结

通过多管并用的“组合拳”措施，可以促进医学信息继续教育的健康发展，提升基层全科医生的综合素质和医疗服务质量，从而更好地满足社会对医疗服务的需求。同时，也有助于推动医学领域的进步和发展。医学信息继续教育是医疗行业不可或缺的一环，对于全民健康和医学科研的发展都具有积极的推动作用。

PU-173

人工智能 GPT 技术在智慧医院应用中的机遇与思考

马青 孟啸
南京市中医院

一、引言

自 ChatGPT 问世以来，其在中国迅速风靡，OpenAI 的 GPT-4 发布进一步提升了 ChatGPT 能力。截至 2023 年 9 月 8 日，CNKI 上有 3247 篇以“ChatGPT”为关键词的文献，可以说 GPT 等自然语言处理技术在医疗领域有着广阔的发展和应用空间。微软全球资深副总裁彼得·李等描述了结合医生、患者、机器的“三方模型”，预测 GPT-4 等大语言模型将颠覆医疗行业。2023 年 2 月 25 日，国家健康医疗大数据研究院（深圳）展示了名为“华佗 GPT”的人工智能产品。2023 年 5 月 25 日，国内互联网医院、慢病管理平台医联体发布自主研发的基于 Transformer 架构的国内首款医疗大语言模型“MedGPT”，旨在实现全流程智能化诊疗。谷歌医疗同步推出医疗大模型 Med-PaLM2，在对外发布前已在多家诊所内测试，并在 MedQA 测试集中达到“专家”水平的 AI 大模型。谷歌还在探索多模态功能，比如识别患者的 X 光来辅助诊断和护理。

本文将以 GPT 为例，分析通用大模型技术在智慧医院应用中带来的机遇与挑战，并展望该技术未来在医院不同环境下的应用场景。

二、GPT 简述

GPT（Generative Pre-train Transformer），是一种预训练语言模型，使用 Transformer 架构处理自然语言处理（NLP）任务。GPT 目前被认为是最先进和最具有代表性的自然语言处理模型之一，通过使用大规模的预训练数据素材和自主学习技术，能够捕捉自然语言中语法规则[1]、习惯表达式、上下文意思和语义含义等，并生成自然语言文本同人类进行交流，实现例如聊天、写作、发布新闻等思维活动。

据 OpenAI 官网介绍，ChatGPT 能在几分钟内形成一篇结构性完整的论文，完成家庭作业，甚至可以通过某些考试，最新的 GPT-4 在一些专业和学术领域表现出超出人类水平，约能以前 10%分数通过模拟律师考试。在一些政府门户网站、电商自助问答、服务型行业、咨询行业、信息服务行业等人机交互、智能问答已得到初步应用[2]。2023 年 3 月 Elondouetal 发布的研究结果表明，有 86 个行业的工作者的全部工作将被 GPT 降低 50%以上的工作时长。在交通方面,GPT 通过实时监测和分析车辆、道路、信号灯等信息，智能调度交通流量，预防交通拥堵。在医疗领域，除了前文提到的“华佗 GPT”、MedGPT 及医疗大模型 Med-PaLM2 外，浙江大学医学院附属第二医院

神经外科吴承瀚团队最新研究也证实，ChatGPT 能够顺利通过中国职业医师资格考试[3]。

三、 GPT 在医院建设发展中的应用场景

在医院信息化建设发展过程中，以 GPT 为代表的人工智能技术，利用自然语言处理、机器学习、深度学习等技术手段，彻底改变传统就诊模式，推动构建覆盖诊前、诊中、诊后的线上线下一体化、智能化、智慧化全流程医疗服务体系。

AI 预约导诊，患者在家即可在线交流，通过患者病情描述协助患者初步判断病情、提供就诊指引，告知患者应该挂哪个科室，推荐合适的看诊医生信息，在线预约就诊时间，并定时推送医院就诊的消息提醒、院内就诊路线、就医流程及看诊前注意事项等。到达医院后，通过 AI 移动端引导患者取号、候诊区签到等候，实时更新显示排队等候人数及等候时间。在等待就诊时间， AI 移动端看诊机器人会模拟医生展开预问诊，同时将患者病情信息转化为规范表达，帮助医生提前书写电子病历[4]。针对患者给出的症状描述和疾病史进行分析匹配，提供患者初步的诊断结果，也可以给出对应的检验检查建议，以供医生做进一步的诊疗判断依据。例如， GPT 依据医学知识库和医疗大数据，构建一个强大的病例分析与匹配系统[4]，例如，当患者在对话中表述这 2 天有发烧、头痛并伴有咳嗽、流鼻涕的症状时， GPT 可初步给出判断为感冒，具体是属于什么性质的感冒，建议开具血常规检查做进一步判断，还可以提供相应的治疗方案和建议，帮助医生在看诊时准确快速地进行诊断和治疗决策[4]。

待进入医生看诊环节，医生已通过预生成的电子病历基本了解该患者的现病史、既往病史、检查结果等相关就诊信息，节省了填写病历的时间，有更充足的精力放在对患者望闻问切的诊疗中。医生只需调整和完善病历，这极大提高了医疗服务的效率和准确性，也减轻了医生的工作负担。在开具处方时，GPT 通过对医药理论、医学文献学习，病例数据分析等提醒医生注意不同药物的适应症、副作用、禁忌症等相关信息[5]，辅助医生进行用药辅助决策，提高医疗效率和安全性。

AI 医生在患者看诊结束后会继续陪护、监测患者病情，提供个性化健康管理计划、合理饮食和运动方案、健康知识科普，以及通过医用级智能穿戴设备采集体征数据、整理分析健康变化数据，帮助医生改进治疗方案，建立对慢病患者的长期跟踪随访。

GPT 经过预训练处理深度学习技术实现对病历文本的分析与归纳，将病历信自动分类处理，集成到医疗数据管理的系统中[6]，医生可以通过自然语言进行快速检索病历信息。例如，医疗美容科医生想检索 2020 年有“瘢痕”诊断的门诊患者信息，进一步筛选这些患者中哪些人做过放射治疗，即可通过对话的方式，让 ChatGPT 返回精准快速的检索结果，帮助医生实现文献调研、数据分析等科研工作，还可以对大量的医疗大数据进行挖掘与分析，探究疾病的发生规律、分析患病的风险因素、不同的治疗方案的疗效等，为临床科研提供有力的支持和参考。

对 GPT 输入大量中医经典文献、名老中医口述的实践经验等，对名老中医的问诊开方经验进行分析编写成文案。利用 ChatGPT 可让学医者和医学爱好者随时随地学习中医知识和中医文化。还可以辅助医教研与各种培训，GPT 可以根据给定的临床案例和条件，模拟真实的临床环境，锻炼提高医生的临床推理和决策能力。

基于 GPT 深度学习技术与算法，实现对医疗数据的监测与管理，分析利用医疗数据推动医院各项工作的优化管理与提升。通过对医疗资源、医疗人员等数据的分析和处理，智能合理安排医护人员的工作计划和岗位调度，提升医务管理效率。实现对医疗治疗管理、医疗质量安全管理、药品采购管理、财务出纳管理等[7]。例如，通过数据筛查分析，采取有效管理措施降低住院患者围术期死亡率；通过每月开具的药品占比及库存情况，及时调整采购方案。GPT 还可以集成接入医疗设备，实现对医疗器械的智能监测、数据采集和问题发现，改进护理方案等。将医院数据库与 GPT 相结合，使用统一规范的数据治理标准，打通所有业务系统之间的信息壁垒，互联互通和信息的交互共享[8]。如此，医生通过语言描述出想要查询的数据，ChatGPT 即可自动制作提供相应的查询报表。

四、限制与挑战

GPT 虽然在医院发展中的应用前景广阔，但同时也面临着诸多限制与挑战，具体有以下几个方面：

（一）运营成本高。GPT 技术在医疗领域的应用需要大量的资金投入，包括技术研发、硬件设备采购、人员培训等方面。以 GPT 为代表的人工智能依赖大模型、大数据与大算力支撑，对服务器硬件配置要求非常高。大模型庞大的训练任务，导致传统的 CPU 根本无法支撑，需要由大量 GPU 服务器组成的算力集群来提供算力，这些服务器之间还需要性能稳定、高强度的网络连接，进行海量数据交换，这就意味着投入成本非常巨大，人工智能大规模算力的建立不能仅仅依靠某个医院实现，而是应该靠国家，投入大量的资源去解决。GPT 生产结果肯定存在一定程度的不确定性和错误，需要人为在实际应用中进行干预和优化调整。

（二）人力投入大。医疗领域的知识体系极其庞大复杂，医疗大语言模型需被 “投喂”大量的医学书籍、临床诊疗指南、医学论文专业知识、临床实际的患者病例信息数据既包括医生对该疾病的诊断记录，也包括患者身体特征、检测数据、家族史、环境信息等多维度的信息等。因为医疗 GPT 技术与生命健康息息相关，要求输出的信息精准度更高，在对其进行大量数据训练及深度学习过程中，需要专业的医护人员不断去测试随机抽取的医学问题，对 GPT 生产的结果校验，不断干预和优化调整学习内容及生产结果。这就需要医院抽调各专业各科室的具备执医资格的医生参与此活动，医护人员在繁忙的工作之余投入更多的时间和精力，可能会增加他们的负担和压力。

（三）数据安全面临巨大挑战。GPT 技术的应用涉及大量的医疗数据，如何保障数据的安全性和隐私性成为了一个重要的问题。首先，医院需要采取严格的数据安全措施，以防止数据泄露和保护患者隐私。其次，GPT 技术在医疗领域的应用需要符合相关的法律法规和伦理规范，以确保其合法性和道德性。此外，医院还需要考虑 GPT 技术在数据存储和传输过程中的安全性问题，以防止数据被篡改或损坏。

五、总结

以 GPT 为代表的人工智能技术在医院建设中具有广阔的应用前景，随着技术的不断发展，GPT 可以通过不断技术升级、预训练和优化，有望成为医生和医疗团队的得力助手，提高医疗服务的质量和效率，提升患者良好的就医体验。然而，在应用过程中也需要考虑信息安全、数据准确性可靠性、伦理问题等方面的挑战，以确保技术的可持续发展与安全应用。尽管面临一些挑战和限制，但相信随着科技不断进步，GPT 人工智能技术必将在促进医疗服务的提升和医院建设的发展中发挥巨大的作用。

PU-174

基于 5W1H 分析法的元宇宙智慧养老服务研究

阳杉[1,3] 王安莉[2,3]
1.中南大学生命科学学院
2.中南大学湘雅三医院
3.医学信息研究湖南省普通高等学校重点实验室（中南大学）

0 引言

随着社会老龄化程度的日益加剧，养老问题越来越突出。据国家统计局数据显示，截至 2021 年底，全国 60 周岁及以上老年人口约 2.67 亿人，占总人口的 18.90%，老年人口抚养比为 21.10%，比 2010 年提高 9.20 个百分点[1]。并且，老年群体的素质也在不断提高，在 60 周岁及以

上老年人口中，拥有高中及以上文化程度的人口比重为 13.90%，比 2010 年提高了 4.98 个百分点[1]。因此，如今的养老问题不再单单是数量上的挑战，更是质量上的挑战。养老需求呈现出个性化、多元化的特征，不仅仅需要注重老年群体的物质需求，更要注重其精神需求。

与此同时，元宇宙概念兴起，元宇宙的兴起顺应时代发展趋势，国家“十四五”规划中强调了医疗领域的创新发展需进一步加强，元宇宙医疗是借助元宇宙技术在传统医疗基础上发展起来的，若应用于养老服务领域，也将对助推传统养老逐步向智慧养老演进。

5W1H 分析法，即“对选定的项目、工序或操作，都要从原因(Why) 、对象(What) 、地点(Where) 、时间(When)、人员(Wh0)、方法(How)等 6 个方面提出问题并进行思考"[2]。目前，5W1H 分析法已被广泛应用于各行各业，如陈桂香和谭翠运用 5W1H 分析法，对公共图书馆开展游学阅读服务的原因、对象、地点、时间、人员、方法等基础方面进行分析[3]；杨智勇和贺奕静运用 5W1H 分析法，对数字档案馆智慧服务的产生背景、概念内涵、服务主体、服务空间、服务时点等一系列问题进行分析[4]；于兴尚等人运用 5W1H 分析法详实分析和论述了图书馆智慧服务兴起的缘由、模式类别、布局要则、流程定制、参与主体、改良路径，为图书馆实现资源建设创新与服务机制转型提供了一套图书馆智慧服务的理论依据与价值基础[5]；韩一平等人在全科医师培训中探索应用 5W1H 方法，将全科医师培训从原因(Why)、对象(What)、人员(Who)、时间/程序(When)、地点(Where)、方法(How)6 个方面分模块实施[6]；曾莉芬通过管理学当中 5W1H 分析方法就工业品的品牌建立和传播进行分析，通过对工业品牌传播的研究现状及问题的分析，进一步地看清工业品传播的本质所在[7]等。综合来看，5W1H 分析法在图书馆领域的应用最广，而在医疗服务、养老服务方面暂未有相关研究。

因此，本文将从元宇宙视角出发，结合 5W1H 分析法，深入分析智慧养老服务的整体态势，助力数智时代卫生健康事业高质量发展。

1 Why——为什么要将元宇宙与智慧养老服务结合

1.1 养老产品“适老性”不足

目前，智慧养老对于智能设备的开发与应用还处于初级阶段，涉老信息的应用、整合和处理等方面还无法实现数据的充分采集、整合处理、有效分析以及开放共享。而且大部分的智慧养老产品操作复杂、流程繁琐，产品设计人员仅关注到产品功能的叠加，而没有考虑到很好地贴合老年人群使用电子设备的实际能力和生活习惯，从而导致了养老产品“不养老”的问题。而元宇宙技术所构建的真实沉浸式体验感、强社交性、内容多元化、去中心化的虚实共生形态，不需要老年人去学习使用复杂的电子产品、系统，不局限于移动手机、平板电脑的便捷访问方式。元宇宙视域下的各种穿戴设备、智能家居都可以实现五官的延长，捕获到更多的外界信息，让老年人享受到社会全产业链带来的优质服务，实现“傻瓜式”操作。

1.2 养老服务忽视了老年人的精神需求

如今养老的工作重心一般放在提供医疗养护和基本生活保障上，较少重视老年人精神层面的需求。老年人缺乏情感慰藉以及接受教育的机会，更不用谈自身的价值再创造。智慧养老若只是重视技术精度，缺乏人文温度，反而会使老年生活变得更加封闭[8]。数字孪生技术背景下，元宇宙向现实世界的镜像发展，为大众带来沉浸式体验的同时，还不断扩展了人们生活维度、感官维度以及思维空间。虚拟现实的结合可以降低老年人参与社会活动的门槛和扩大其生活范围，满足老年群体的精神需求。

2 What——元宇宙视域下的智慧养老服务是什么

元宇宙是整合诸多独立工具、平台、基础设施及各主体间的协同协议等共同相互融合的新型虚拟形态，特别是虚拟现实、5G、云计算等技术的成熟和发展使得元宇宙从概念走向现实。元宇宙智慧养老服务需立足于老年人的根本需求，从学习教育、医疗康复保健、日常生活照料、人际交往、文化娱乐等多个方面来满足老年人需求，帮助老年人建立社会连接，提高老年人的生命生活质量。通过元宇宙赋能智慧养老服务模式，形成对物理空间健康养老服务的补充，使不同年龄、能力、地域的老年人均能“老有所养、老有所学、老有所乐”[9]。

3 Where——元宇宙结合智慧养老服务的场景

3.1 元宇宙助力康复医疗服务

元宇宙可提供一个沉浸式和交互式虚拟平台，通过模拟真实世界场景，使患者在安全可控的环境中进行功能性锻炼[10]。由于疾病、残疾等因素的影响导致部分老年人行动不便，一些老年人位于偏远地区，医疗资源不足，使老年人在就医、进行康复治疗时存在障碍。例如，使用 VR 游戏和模拟来改善中风和其他神经系统疾病患者的上肢功能，计算机向老年患者发布特定的运动任务让其跟随锻炼，以趣味性的方式提高患者的恢复动力和参与度。

3.2 元宇宙助力居家养老服务

老年人对精准化、个性化的医疗服务和护理服务的需求逐渐增加。元宇宙技术可以提供自动化辅助设备和个性化健康信息服务，也能为医疗诊断提供决策支持。通过智能家居监控和随身携带的传感器，获取数据，建立老年人养老数据库，这既可以用于子女和社区对老年人行为和心理的分析，以便进行适当的积极干预，也可以为服务提供者进行决策时提供依据。元宇宙智慧养老服务在老年人健康管理监测、紧急呼救、家庭护理等方面都可以发挥重要作用。

3.3 元宇宙助力精神需求服务

文化娱乐需求是老年人重要的精神文化需求，对于促进老年人身心健康，提高老年人生活品质具有重要意义。在养老社交元宇宙空间中，老年人能够以虚拟化身的形式进入社交元宇宙，虚拟化身可减少老年人对自身形象、身体机能的顾虑，同时打破了物理空间对老年交际圈层的限制，可以随时与亲朋好友一起聊天。除社交外，基于扩展现实技术打造的文化旅游、图书馆、博物馆元宇宙，也将成为丰富老年人文化娱乐生活的重要途径，老年人可足不出户即能前往世界各地旅行、在家博览各地博物馆，极大地满足老年人的精神需求。

4 When——元宇宙智慧养老服务的时间节点

相比传统养老服务时点的定期性（相关医护人员在工作日提供服务）、定点性（相关医护人员在工作时间点内提供服务）和延迟性（服务申请到服务完成存在时间间隔），元宇宙智慧养老服务则可以实现全天候、实时且及时服务。而且元宇宙智慧养老服务的时间节点将进一步延伸到前期和后续阶段，前期阶段将具有前瞻预测性，后续阶段将具有持续反馈性。例如，通过老年人随身携带的传感器监测老年人的身体健康数据，若监测出现异常数据，可及时为其提供就医服务，提前做出预测；而在提供了相关服务后，还可以持续对用户的基本信息、行为等数据进行采集、整理和组织，构建相应的群体用户画像库。

5 Who——元宇宙智慧养老服务的主客体

服务主体即元宇宙智慧养老服务的提供者，作为智慧养老服务的实施者和主导者，其服务主体包括参与智慧养老服务的一切组织、机构和个人，如医院诊所、医疗护理人员、元宇宙技术开发人员等。当今时代，新的信息服务环境、服务理念和服务方式对服务主体整体素质提出了更高的要求。服务客体即接受养老服务的老年群体，其需求特点也呈现出更加多元化、个性化、时效性和准确性、集成化与智能化的特点，而以上新变化都将推动传统养老服务向智慧养老服务的转变。

6 How——如何推动元宇宙智慧养老服务

6.1 加快 AR、VR 等技术在养老领域的应用

AR、VR 技术是元宇宙行业发展的基础，无论是从元宇宙发展的国家战略角度考虑，还是从元宇宙在智慧养老领域应用的角度考虑，都应推动 AR、VR 技术的发展，逐步实现智慧养老领域与元宇宙的深度接入与融合[11]。例如借助元宇宙 VR 虚拟现实技术营造 360 度实景三维场景，提供给老年人足不出户便可以全球旅游的体验等，这些都离不开 AR、VR 等技术的应用。

6.2 加大元宇宙智慧养老的宣传力度

元宇宙作为当下新兴的科技概念，人们对其的了解还不深入，许多医疗机构、养老机构、养老服务人员仍处于观望状态，不愿轻易尝试新兴事物，对于元宇宙智慧养老的模式、应用场景不熟悉，导致元宇宙技术目前的被接受程度还比较低。因此，我们需要加大对于元宇宙智慧养老的宣传力度，鼓励医疗机构、养老机构等积极尝试元宇宙相关技术。

6.3 加强元宇宙智慧养老政策引导

任何新技术、新领域的发展都离不开国家政策引导，只有跟随国家政策方向，才能在正确的道路上走的更快、更远。因此，需要从国家层面加强元宇宙智慧养老政策体系的建设，建立防范相应风险的政策体系，规范行业标准，根据企业与机构自身的实际情况，通过提供扶持资金、遴选优秀案例、开展协作交流等方式，鼓励、引导相关企业与机构加入元宇宙的研发与建设，促进健康养老元宇宙领域的发展。

7 结语

元宇宙是信息时代发展的新阶段，随着互联网、物联网等基础设施的成熟应用，区块链、数字孪生、扩展现实等技术的飞速发展，元宇宙在多个领域进行应用，呈现了广阔的发展前景，也将在智慧养老领域发挥重要作用。本文以 5W1H 分析法，对元宇宙智慧养老服务进行了深入的分析，助力传统养老服务向智慧养老服务进行转变，促进数智时代卫生健康事业的高质量发展。

PU-175

心房颤动导管消融手术患者管理中的机器学习：系统综述

陈炜[1,2,3,4] 罗爱静[1,2,3,4] 辛梓睿[1,3,4]

1.中南大学湘雅二医院
2.中南大学生命科学学院
3.医学信息研究湖南省普通高等学校重点实验室（中南大学）
4.湖南省心血管智能医疗临床医学研究中心

目的 导管消融（CA）是目前治疗心房颤动（AF）的主流手段之一，机器学习（ML）在帮助医生开展 AFCA 患者管理方面具备潜力，一些研究探讨了 ML 在 AFCA 患者管理中的应用，值得关注。本综述旨在系统地回顾在 AFCA 患者管理中使用 ML 的研究，强调 ML 在该领域凸显的优势以及暴露出的局限。

方法 在 PubMed、Web of Science、Embase、the Cochrane Library、ScienceDirect 上进行检索，根据纳排标准和人工审查确认纳入研究，使用调整后的诊断准确性研究质量评估工具（QUADAS-2）评估其偏倚风险和适用性。

结果 检索出的 246 项研究中有 25 项被最终纳入，确认了几个主要方向：（1）识别房颤导管消融靶点；（2）改进导管消融策略；（3）预测术后复发及再入院情况。纳入研究使用的患者数据除了一般人口统计变量、临床特征外，还涉及到影像检查（n=9，36%）、电生理信号（n=10，40%）等不同类型；进行模型开发时，以卷积神经网络为代表的深度学习被应用最多（n=15，60%）。由于缺乏外部验证等原因，大部分研究存在较高的偏倚风险（n=20，80%）。

结论 ML 凭借对复杂数据的强大学习能力，在 AFCA 患者管理方面展现了广阔前景，面对特定医学任务时甚至表现出与人类临床医生相当的优良性能，但模型是否可泛化、可解释仍需未来研究的进一步讨论。

PU-176

基于 ChatGPT 的医生诊疗 AI 助手系统的设计与实现

殷鸣
湘潭市中心医院 信息科

引言：

随着人工智能技术的飞速发展，其在医疗领域中的应用正以前所未有的速度和规模展开。在医疗诊疗过程中，医生常常面临着海量的病历数据、复杂的疾病诊断和治疗方案选择等各种挑战。这需要医生具备广泛而深入的医学知识和临床经验。然而，传统的医学教育和知识传递方式存在一定的限制，无法满足医学知识快速增长和不断演进的医疗需求。近年来，基于自然语言处理和机器学习的人工智能技术取得了巨大突破，其中 ChatGPT 作为一种强大的对话生成模型，呈现出了巨大的潜力。ChatGPT 是 OpenAI 团队基于 GPT-3.5 模型开发的一种大规模语言模型，能够根据输入的文本生成连贯、语义丰富的回复。这使得 ChatGPT 成为一种理想的工具，可以应用于开发医生诊疗 AI 助手系统，为医生提供实时的医疗建议和诊断支持。本论文旨在设计和实现一种基于 ChatGPT 的医生诊疗 AI 助手系统，利用 ChatGPT 的强大语言理解和生成能力，结合医疗领域的专业知识，为医生提供辅助决策和诊断的功能。该系统通过分析患者的主诉症状、检查检验结果等信息为临床医生提供诊断和治疗建议，并辅助制定个性化的诊疗方案。

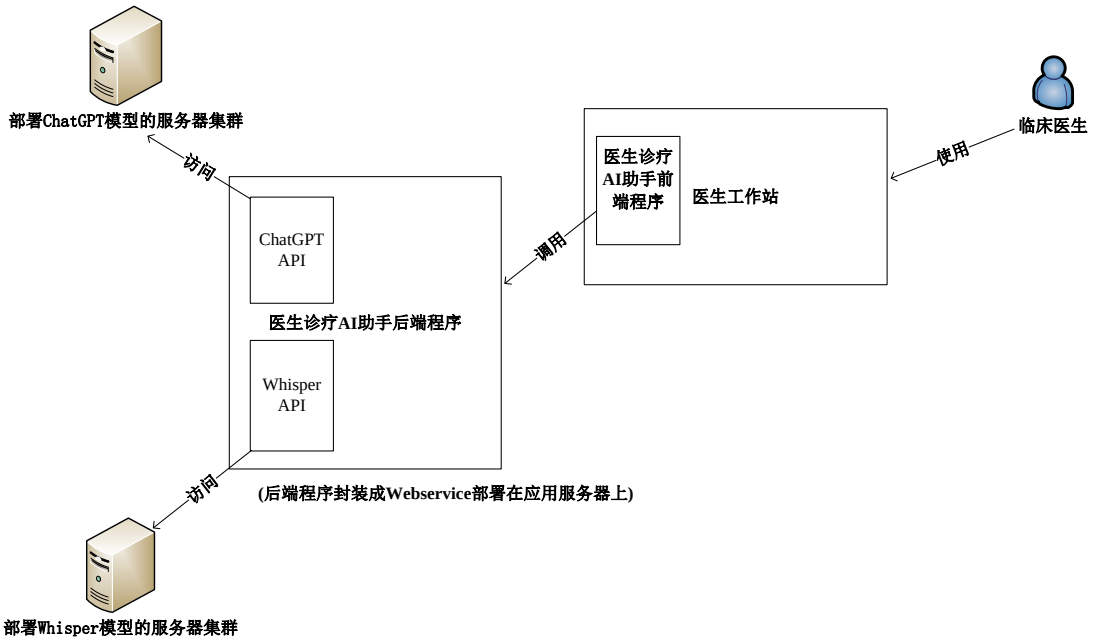


图 1 系统架构图

1. 系统设计

基于 ChatGPT 的医生诊疗 AI 助手系统由 ChatGPT API、Whisper API、医生诊疗 AI 助手后端程序和医生诊疗 AI 助手前端程序组成（系统架构图如图 1 所示）。ChatGPT API 为 OPEN AI 公司提供的 ChatGPT 模型的调用接口，通过这个接口可以调用 ChatGPT 的 AIGC（人工智能生成内容）能力。Whisper API 为 OPEN AI 公司提供的 Whisper 模型的调用接口，通过该接口可以实现将语音转成文字，大大地提高了用户与 ChatGPT 模型交互的便捷性。医生诊疗 AI 助手后台程序负责对接口进行集成，里面集成了 ChatGPT API 与 Whisper API 两个接口，医生诊疗 AI 助手后台程序被封装成 Webservice 供医生诊疗 AI 助手前端程序调用。医生诊疗 AI 助手前端程序作为一个客户端可以由医生工作站调用或嵌入集成到医生工作站中。医生诊疗 AI 助手系统的运行机制为：临床医生在医生工作站程序中打开医生诊疗 AI 助手系统，以打字或语音口述的方式（语音转文字的工作由 Whisper API 完成）将病人的主诉症状、检查检验结果输入到医生诊疗 AI 助手前端程序中，医生诊疗 AI 助手前端程序调用医生诊疗 AI 助手后端程序 Webservice 将病人的主诉症状、检查检验结果等信息通过 ChatGPT API 发送给 ChatGPT 模型，ChatGPT 模型根据输入的文本，利用其强大的语言理解和生成能力，结合医疗领域的专业知识，为临床医生提供诊断和治疗建议，并辅助制定个性化的诊疗方案。

2.系统实现

2.1 ChatGPT API 与 Whisper API 的实现

首先在 OpenAI 官网上进行注册并获取 API Key,根据官方接口文档写出接口代码，接口代码用 Powerbuilder 11.5 实现，ChatGPT API 的接口关键代码为：

```
curl.seturl("https://api.openai.com/v1/chat/completions","POST")  //配置 url
curl.setheader("Content-Type","application/Json")   //配置参数格式
curl.setHeader("Authorization","Bearer " + as_apikey)  //配置权限信息
//配置 ChatGPT 模型参数
ls_data =  '{"model":"' + as_model +'","messages": [' + ls_context + '{"role": "user", "content": "'+ as_prompt + '"}],"temperature":' + string(ai_temperature) + ',"max_tokens":'+ string(ai_max_tokens) }'
```

```
curl.setbody(ls_data)
curl.request()    //发送请求
```

在接口代码中依次配置好 ChatGPT 的接口地址 URL,参数格式、权限信息、模型参数等信息。ChatGPT 模型参数采用的 JSON 格式，其中参数 model 用于定义模型类型，参数 role 用于定于模型角色，参数 content 用于定义 prompt（提示词）即用户输入给 ChatGPT 的文字信息，参数 temperature 用于定义 ChatGPT 回答内容的多样性和创造性，参数 max_tokens 用来定义 ChatGPT 生成内容的最大长度。参数设置好以后采用 request 函数将参数发送给 URL 接口.ChatGPT 接口响应后，会将生成的内容以 JSON 的格式返回给接口调用者。

Whisper 模型是 OpenAI 公司研发一种基于自然语言处理技术的语音转文本模型。Whisper 模型使用了一种新型的转录算法，可以将人类语音转换为高质量的文本。该模型采用了一种称为“自适应语音识别”的技术，其能够自动适应说话人的语种语调和语言习惯。考虑临床医生工作繁忙，打字输入患者的主诉、检查检验报告效率太低。医生诊疗 AI 助手提供了语音输入功能，语音转文字的工作由 Whisper 模型完成，Whisper 接口代码用 Powerbuilder 11.5 实现，Whisper API 的接口关键代码为：

```
curl.seturl("https://api.openai.com/v1/audio/transcriptions","POST")  //配置 url
curl.setheader("Content-Type","multipart/form-data") //配置参数格式
curl.setHeader("Authorization","Bearer " + as_apikey) //配置权限信息
 //配置 Whisper 模型参数
curl.setform("model", "whisper-1")
curl.setform("file",as_filename,true)
curl.setform("response_format","text")
curl.request() // 发送请求
```

在接口代码中依次配置好 Whisper 的接口地址 URL,参数格式、权限信息、模型参数等信息。在 Whisper 的模型参数中,参数 model 用于定义模型类型，参数 file 用于指定需要转换成文字的语音文件，参数 response_format 用于定义语音转换成文字后返回来的格式。参数设置好以后采用 request 函数将参数发送给 URL 接口.Whisper 接口响应后，会将语音转成的文字内容以 text 的格式返回给接口调用者。

2.2 医生诊疗 AI 助手后端程序的实现

医生诊疗 AI 助手后端程序的作用是集成 ChatGPT API 与 Whisper API 两个接口，该程序用 Powerbuilder 11.5 实现，将两个接口分别定义为 gf_chatgpt_api()函数和 gf_whisper_api()函数。用 PB.NET 技术将两个函数封装成 Webservice 部署在应用服务器上供医生诊疗 AI 助手前端程序调用。

2.3 医生诊疗 AI 助手前端程序的实现

医生诊疗 AI 助手前端程序也是用 Powerbuilder 11.5 实现，该程序为临床医生提供了一个与 ChatGPT 模型进行交互的界面。医生可以在医生工作站里通过相关的菜单链接打开医生诊疗 AI 助手前端程序，医生可以用文字或者是语音的方式将患者的主诉症状、检查检验报告等临床病历信息输入给医生诊疗 AI 助手前端程序，医生诊疗 AI 助手前端程序通过调用医生诊疗 AI 助手后端 Webservice 与 ChatGPT 模型进行交互，ChatGPT 模型利用其强大的语言理解和生成能力，结合医疗领域的专业知识，对临床医生所提供的患者病历信息进行分析并生成诊断和治疗建议供医生参考。（医生诊疗 AI 助手前端程序如图 2 所示）

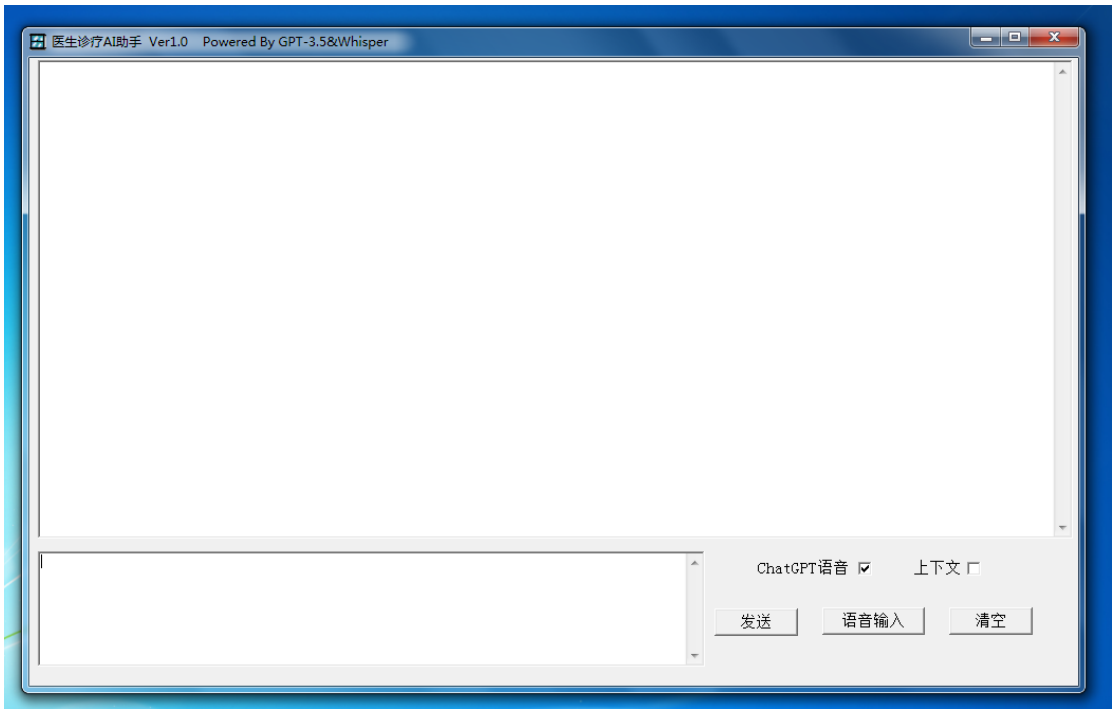


图 2 医生诊疗 AI 助手前端程序

3.应用情况

目前医生诊疗 AI 助手在我院还处在临床测试阶段，为了评估该系统的性能，我们邀请了一批专业医生参与了测试。我们将医生诊疗 AI 助手提供的诊断建议与人工医生的表现进行对比分析，以此来验证医生诊疗助手所给出诊断建议和回答问题的准确性。我们举一个测试例子，我们将一个经过脱敏的某患者病历数据输入给医生诊疗 AI 助手，我们往医生诊疗 AI 助手中输入某患者的主诉、体征、化验结果、影像所见、B 超所见、肌电图结果等患者的病历信息(如图 3 所示)，并要医生诊疗 AI 助手给出诊断意见和治疗方案，医生诊疗 AI 助手通过 ChatGPT 模型给出了诊断意见和治疗方案（如图 4 与图 5 所示）。ChatGPT 给出了“高血压”、“脑血管疾病”、“糖尿病”、“甲状腺功能亢进”、“颈椎病和腰椎间盘突出症”、“甲状腺结节”等六个诊断。我们又要人工医生在不参考 ChatGPT 所给出结果的情况下分析该患者的病历信息，人工医生给出的诊断为“甲状腺功能亢进性肌病”、“脑血管意外”、“糖尿病”、“甲状腺功能亢进症”、“2 型糖尿病”、“高血压”、“甲状腺结节”等七个诊断。通过比较分析，发现医生诊疗 AI 助手所给出的诊断意见与人工医生给出的诊断基本相符，且给出的治疗方案也具有一定的实用性和可参考性。经过对大量脱敏的真实病历数据进行测试，测试结果表明，该医生诊疗 AI 助手在诊断和治疗决策方面表现出了较高的准确性和实用性。医生们普遍认为该系统能够提供有价值的帮助，提高了他们的工作效率和患者的治疗质量。

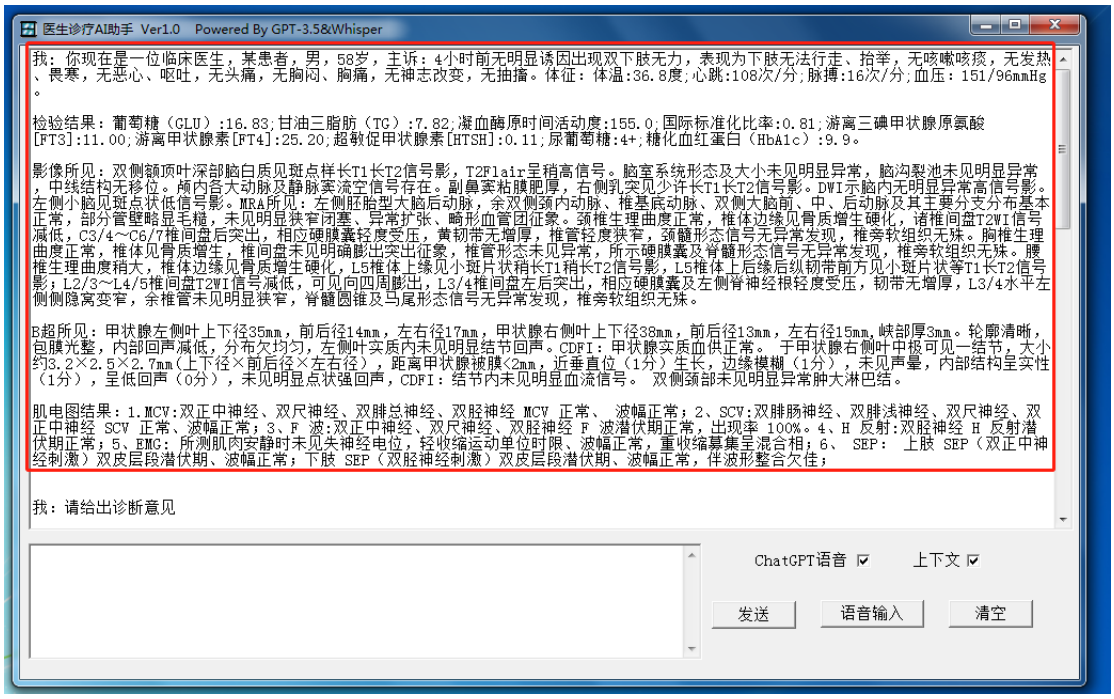


图 3 输入患者的主诉、体征、检查检验结果到医生诊疗 AI 助手

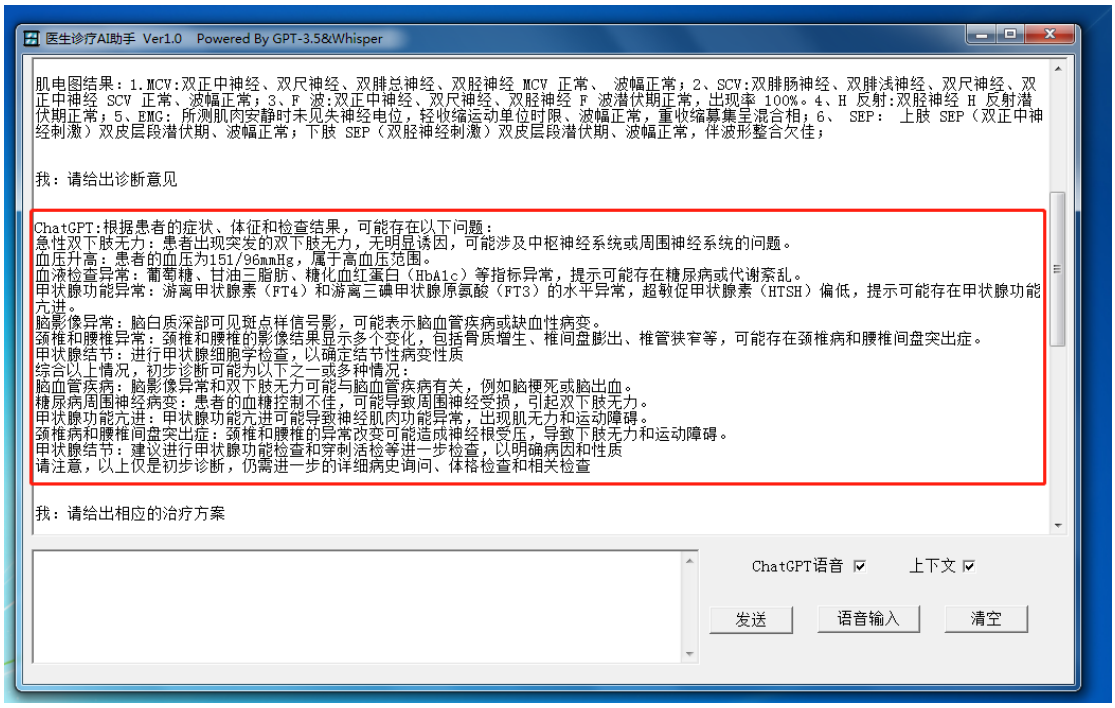


图 4 医生诊疗 AI 助手给出的诊断意见

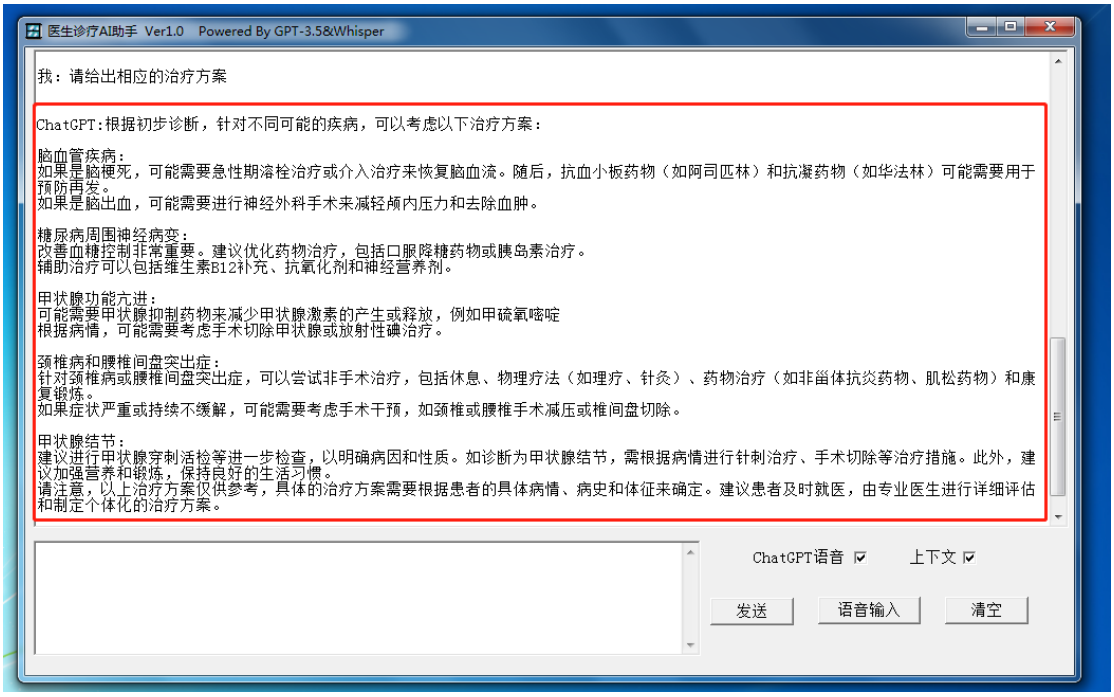


图 5 医生诊疗 AI 助手根据诊断意见给出的治疗方案

4. 结论

在本论文中，我们设计并实现了基于 ChatGPT 的医生诊疗 AI 助手，并对其进行了评估。

经过对大量脱敏的病历数据进行测试，我们发现基于 ChatGPT 的医生诊疗 AI 助手在提供医疗辅助服务方面具有巨大潜力。该系统通过自然语言处理和机器学习技术，能够理解医学文本并生成准确的回答，为医生提供了一个智能化的诊疗工具。然而，我们也要意识到该系统面临的挑战和局限性。数据隐私和安全性是重要考虑因素，需要采取适当的措施来保护患者的隐私和确保数据的安全。此外，模型的可解释性和责任问题也需要进一步研究，以增强医生和患者对系统的信任度和接受度。未来的研究方向包括使用专业的医疗数据集利用 Fine-tuning 技术对 ChatGPT 模型进行更个性化的训练，以提高它在诊疗辅助方面的服务质量和效率。以及将医生诊疗 AI 助手与其他医疗信息系统作更好的界面集成和数据集成让其能提供更全面、更高效的医疗服务。

PU-177

基于机器学习的药物组合协同预测方法综述

徐梦蝶

南京医科大学生物医学工程与信息学院医学信息工程学系

引言

近几十年来，药物的研究发展得到了长足的进步，在治疗肿瘤时也达到了较好的疗效。然而，单独使用一种药物对肿瘤进行治疗，可能会存在耐药性、毒性等问题。在许多情况下，耐药性往往是有效治疗的主要障碍。与单一治疗相反，联合治疗能够克服耐药性，减少用药剂量，降低毒性，实现更有效的疾病控制[1]。因此，联合治疗已成为许多复杂疾病的标准临床治疗策略[2]。

1 基于计算的药物预测方法

早期的基于计算的药物筛选方法是根据肿瘤的特性，采取大量的患者数据和计算方法，构建预测模型获得对大多数患者有效的最佳药物。尽管使用基于计算的单药筛选方法一直在发展，单药治

疗仍然存在耐药性等缺点，并且使用一种药物可能无法解决患者体内的问题。药物联合作用相比单一用药具有提高疗效，降低或克服耐药性等优点。在进行对联合治疗的药物的筛选时，虽然通过临床经验筛选结果准确，但是筛选费时费力，并不适合大规模的药物筛选[6]。随着技术的发展，大规模筛选技术开始被应用，运用比较广泛的方法包括高通量筛选(High throughput screening, HTS)方法[7]和相互作用化合物方法的多重筛选[8]，HTS 以分子水平和细胞水平的实验方法为基础，利用自动化系统控制实验的执行，具有快速、灵敏和精准的优点[7]。使用这两种方法大大加快了药物组合的发现，然而随着药物数量的指数型增长，为每一个药物筛选其可能的药物组合[9] [10] [11]变得不切实际，因此使用计算方法来减少药物组合的搜索空间成为目前的迫切需求。

基于计算的药物组合预测方法可以分为系统生物学方法、动力学模型方法和机器学习方法。系统生物学方法对专业的生物知识需求较高[12]，方法的重点在于生物网络中的参数控制和网络分析，然而生物网络具有复杂性，应用于大规模药物组合时会受到一定的限制；动力学建模方法[13]利用动力学方程模拟现实生物网络中节点的动态变化，然而大多数生物网络的动力学还不能被明确定义。与这些方法不同，机器学习方法可以在大规模数据中应用。在给定一定的训练数据后，机器学习方法可以学习输入的属性数据和相关输出（协同分数）之间的复杂关系，能够有效地预测药物的协同，大大降低了药物试验测试成本，因此机器学习被广泛的应用在药物组合预测中。

在机器学习方法中，药物组合预测问题通常被定义为多分类或回归任务。回归任务是预测联合用药的定量协同评分，分类则是判断是否具有协同作用。在分类任务中药物组合效应可分为协同效应、加性效应和拮抗效应。大多数研究将药物组合分为协同作用和非协同作用两类。非协同作用包括轻微的协同作用、相加效应和拮抗作用。

基于机器学习的药物组合协同预测范式为：首先从药物组合数据库和其他相关数据库中获取药物特征和标签，每个实例由特征向量表示，这些特征向量可能会是对相应实体的结构或者其他属性的反应。将特征向量进行连接或融合，得到每个样本的特征向量。最后将特征向量和标签一起送入机器学习模型或者深度学习模型中进行训练。模型网络中的参数可以在训练阶段中被优化，在训练完成后，可以通过各种指标对模型的性能加以评价，并选择最优的模型并将其应用于外部测试集，评价其泛化性，并通过生物实验或者文献检索来验证预测结果。该过程示意图如图 1 所示。

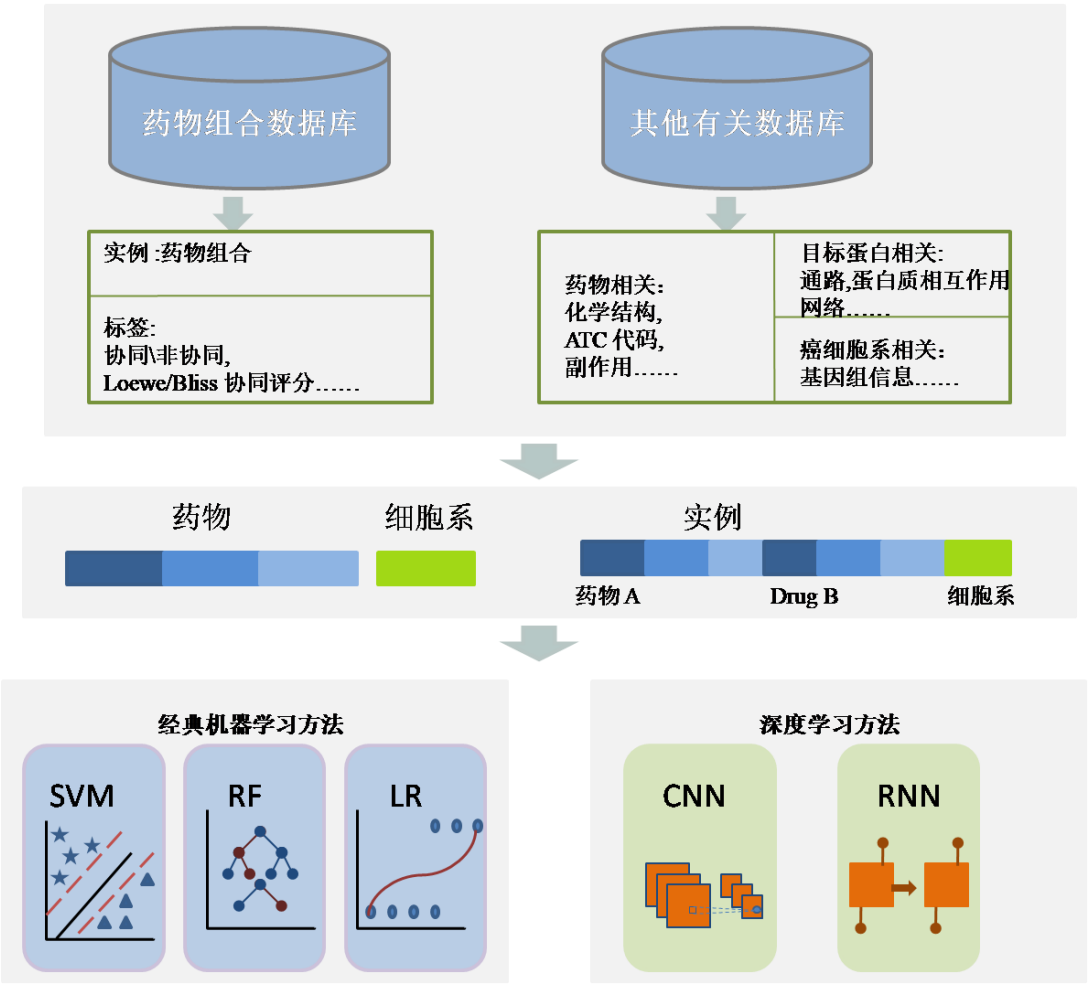


图 1 基于机器学习的药物组合协同预测示意图

1.1 药物组合协同效应

药物组合的协同作用、加性作用和拮抗作用分别对应两种药物联合作用效果大于、等于或小于单个药物效果之和，药物的作用或反应通过细胞增殖的抑制或活力衡量[14] [15]。如何衡量协同作用和拮抗作用的强弱以及量化协同作用仍然是目前面临的难题，也是需要面对的重要挑战之一。计算药物组合协同和拮抗作用最常用的两个模型就是 Loewe 加性模型和 Bliss 独立模型，其他方法大多是基于这两种方法的推导和演化。Loewe 加性模型将化合物和自身组合效应定义为相加效应；Bliss 独立模型采用概率的观点，认为两种药物的组合具有概率独立性。在 Loewe 和 Bliss 的比较中，Loewe 模型更符合两种作用于同一靶点或途径的药物的预期联合效应，Bliss 则是针对非交互作用的药物组合，并且 Loewe 模型需要单一药物的剂量-反应曲线。

1.2 药物组合协同预测数据

本文列出了 5 个数据库。在这些数据库中，Drugcomb、DrugcombDB、SYNERGxDB 等数据集主要关注抗癌药物组合。除了癌症外，Drugcomb 还提供了疟疾和 COVID-19 等其他疾病的数据。

表 1 部分药物协同组合预测数据库

数据库	数据量		网页	最新更新时间
	药物/化合物	组合实验		

Drugcomb_v1.5[16]	8397	739	https://drugcomb.fimm.fi/	03	2021.
	5350	664			
DrugcombDB[17]	1977	498	http://drugcombdb.denglab.org/	05	2019.
SYNERGxDB[18]	104	865			
NCI-ALMANAC[15]	38	536	https://www.synergxdb.ca/		2019
		596			2017.
O'Neil et al.study[19]		304	https://dtp.cancer.gov/ncialmanac	12	
		549			2016.
		227		03	
		37			

Drugcomb 是一个基于网络的数据库，发布于 2019 年[14]，在 2021 年 3 月更新[16]。Drugcomb 提供在 288 个细胞系中测试的 4268 种药物的 739 964 个药物组合实验数据。DrugCombDB 是 2019 年[17]发布的一个基于网络的药物组合数据库。DrugcombDB 包括在 104 种细胞系上测试的 5350 种药物的 498 865 个联合实验，同时在网站上也总结了组合药物的性质。SYNERGxDB 是 2020 年[18]发布的一个药物基因组数据库。它整合了 9 个药物组合数据集，并在在 151 个细胞系中对 1977 个化合物进行了 536,596 个联合实验。同时还提供癌细胞系的代谢组学、基因表达、拷贝数和突变谱等特征。NCI-ALMANAC[17]（抗肿瘤药物组合的大矩阵）数据库由美国国家癌症研究所统计局发布。它通过高通量筛选来自 NCI-60 面板的 104 种研究和批准的 60 种癌细胞系的 304,549 种药物组合。来自默克公司研究实验室的 O'Neil[19]等人进行了一次大规模的肿瘤学筛查实验，并于 2016 年发表了 O'Neil et al.study。它涉及 22737 种药物组合和 38 种癌细胞系及其所涉及的药物。

2 基于机器学习的药物组合协同预测

目前广泛使用的机器学习方法可分为经典机器学习方法和深度学习方法。经典机器学习方法包括支持向量机[21](Support Vector Machine,SVM)、决策树[22](Decision Tree,DT)、随机森林(Random Forest,RF)和极端梯度增强(XGBoost)[23]。深度学习模型包含前馈神经网络[24](Feedforward Neural Network,FNN)、深度信念网络[25](Deep Belief Network,DBN)、自动编码器[26](Autoencoder,AE)等。在小（数百个样本）和中等规模（千或万个样本）数据集的条件下，经典的机器方法可能比深度学习方法性能更好。随着数据集规模的增加，深度学习模型的性能将大大提高。

2.1 传统机器学习方法

在药物组合预测中应用的机器学习方法，可以在中小型数据集上实现较高的预测精度，包括人工神经网络、随机森林、随机梯度增强、XGBoost 等集成学习的方法。

基于传统机器学习的药物组合协同预测方法通过使用不同的输入来获取药物和细胞系的特征，细胞系特征最常见的是基因表达、突变和拷贝数变异。药物的特征主要可以分为结构特征、药物靶点信息以及其他属性等。表 1 总结了部分近年使用传统机器学习对药物组合协同作用进行预测的方法。Janizek 等人[29]提出了 TreeCombo 模型，此方法基于药物的化学和物理特征、癌症细胞系的基因表达特征，采用可扩展梯度提升树方法，对药物协同效果进行预测并给出解释。Jeon 等人[30]使用无偏高通量药物筛选数据，将线性回归，岭回归和随机森林算法用于预测药物协同效果，然而此方法过于依赖数据集，难以对未知的药物组合进行预测；Sidorov 等人[31]在 NCI-ALMANAC 数据集上，将 5000 多对药物组合、60 个细胞系、4 种模型和 5 种化学特征使用随机森林和 XGBoost 进行建模，预测药物组合是否具有协同作用；Zhang 等人[32]提出了 SyFFM 模型，能够在不同的疾病数据集上验证双药或者三药组合协同效果，主要基于药物与靶点、酶和适应症的关联构建药物对特征，通过特征的潜在向量空间上的域感知分解机，获得药物组合的协同分数。在 2020 年，Wang 等人[33]提出了 comboLTR 方法，以化学和组学特征为输入，使用潜在张量重建方式的多项式回归建模，描述治疗剂组合在不同剂量和癌细胞环境中的反应。2021 年 Jilkunen 等人提出了 comboFM,通过预测完整的剂量-反应矩阵来获得具有协同作用的药物对，comboFM[34]的输入特征包括两种药物的分子指纹、两种药物的浓度值和癌细胞的基因表达谱，应用因子分解

FM(Factorization machine)对回归任务进行实现，并通过剂量-反应矩阵对药物对的协同作用进行量化。此处所使用的的剂量-反应矩阵考虑了浓度信息，这也是一种预测药物对协同作用的有效策略。

表 2 基于传统机器学习的药物组合协同预测研究现状

时间	方法	药物特征	细胞系特征	建模任务
2018	XGBoost	化合物指纹、毒效团	外显子组序、RNA 序列	回归任务
2018	SVM、随机森林、逻辑回归	药物靶点信息	基因表达、突变、拷贝数变异和单药治疗信息	回归任务 分类任务
2019	XGBoost	Morgan 指纹、MACCS 键、ISIDA 片段、理化特性	-	回归任务
2019	域分解机模型	药物靶点信息、药物酶信息、解剖学治疗化学分类系统代码	-	回归任务 分类任务
2020	张量分解	分子指纹	基因表达	回归任务
2021	多项式回归	MACCS、蛋白组学数据	基因表达、突变、拷贝数变异等	回归任务

2.2 深度学习方法

传统的机器学习更多地依赖人工提取的特征。因此近年基于深度学习的药物协同预测受到了广泛的关注。深度学习是一种具有多个数据处理层基于人工神经络的机器学习方法。深度网络结构使其能够更好地捕获输入[35]和输出[36]之前的非线性或复杂关系。表 3 列出了近年来深度学习的药物组合协同预测研究现状。

表 3 基于深度学习的药物组合协同预测研究现状

时间	网络结构	药物特征	细胞系特征	建模任务
2018	前馈神经网络	分子指纹、理化属性	基因表达	回归任务
2018	深度信念网络	药物-靶点信息	本体指纹、基因表达	分类任务
2020	图卷积神经网络	药物-靶点信息	蛋白质-蛋白质相互作用网络	分类任务
2021	图卷积神经网络	蛋白-蛋白相互作用网络	蛋白-蛋白相互作用网络	分类任务
2022	图神经网络、注意力机制	药物分子结构	基因表达谱	分类任务
2022	深度森林	药物分子指纹、理化属性	特异性药物诱导基因表达谱	回归任务
2022	图卷积神经网络	药物分子指纹、理化属性	基因表达、突变、蛋白质互作用	回归任务

和传统机器学习相似，选择合适的药物和细胞系特征对预测的结果有决定性的影响，早期的深度学习预测方法使用分子指纹描述符来表示化合物的化学信息。Preuer 等人在 2018 年提出的基于前馈神经网络的药物组合协同预测模型 DeepSynergy[6]经常被用作预测药物组合的基线，具体来说，DeepSynergy 将两种药物的化学描述符和相应癌细胞系的基因表达值作为前馈神经网络的输入，和传统机器学习中的 RF,SVM,GBM(Gradient Boosting Machine)、弹性网络（Elastic Net）四种方法相比，DeepSynergy 在各种评分指标中都具有最好的性能, 然而，DeepSynergy 在应用于没有见过的细胞系和药物数据中时，很难对未学习过的药物组合协同作用进行预测。为了更好的提升预测性能，Lin 等人[38]提出将细胞系特异性诱导基因表达谱作为新的特征类型，使用深度森林作为特征学习模型，捕捉药物的细胞反应并阐明生物学协同作用的机制。Wang 等人[39]在 2022 年提出了一种新的模型 PRODeepSyn,同时将生物网络数据和基因组学数据纳入预测模型中，将蛋白质-蛋白质相互作用网络数据与组学数据相结合，通过图卷积网络提取拓扑结构，为细胞系构建低维密集嵌入。

在几十年研究的基础上，专家定义了多种分子表征，这些表征可以称为特征工程，最近的研究表明现有的分子描述符不足以充分描述化合物[40]。因此有研究人员从简化分子线性输入规范（SMILES）和分子图结构角度对药物进行表示，并进行药物协同预测。中南大学王建新教授[41]将图注意网络应用于药物组合筛选，提出了基于图神经网络和注意力机制的深度学习模型来识别出具有协同作用的药物组合。Kim 等人[42]提出融合 SMILES 序列和分子访问系统（MACCS）密钥指纹对药物进行表征。

目前深度学习预测方法在药物组合协同预测方面已取得了显著成果，但是并没有充分利用药物组合协同涉及到的丰富信息，因此设计更有效的、信息更丰富的药物特征表示将是进一步提高预测模型性能的关键。

3 问题及展望

为了加快对复杂疾病药物组合的治疗，许多人工智能驱动的生物制药公司已经致力于开发基于机器学习和网络药理学方法的药物组合。虽然许多学者和行业权威都已经对药物组合预测进行了研究，并且获得了高性能的模型，但是在这一领域仍然存在一些挑战。

3.1 有限的样本数据和方法缺乏普遍性

根据上述研究，使用合适的数据集是获得更好的预测性能的关键。虽然目前已经整理并公开了

一些大型的药物数据库，但是所提到的癌细胞系和药物的数量有限，这可能会影响模型预测新药物组合的泛化性，学术界所使用的大部分方法很难对新的药物和细胞系进行推广，在未来的研究中，研究者可能应该更加注重模型的泛化性，比如使用外部测试集来测试模型的通用性。

3.2 特定剂量的药物组合协同预测

药物的剂量对于协同作用的预测也有影响。药物组合可能在某一种[43]剂量范围中具有协同作用，在另外的剂量范围中则反而是拮抗效果，对于药物协同将来的预测，可能更精确到药物的剂量范围，而不是简单地分类一个药物组合是否具有协同效应，而是考虑剂量如何优化协同作用，使药物协同作用预测更加精准，这些研究有助于更接近于实现临床精准医疗[35]的目标，在未来的研究中可以进行进一步的探索。

3.3 深度学习模型的可解释性

可解释性指算法要对特定任务给出清晰概括，并与人类世界中已定义的原则或原理联结。使用深度学习对药物组合协同作用进行预测时，病人往往需要知晓算法给出结果的依据，大多数深度学习模型是由数据驱动的黑盒模型，人们很难理解深度学习中隐藏层数目、神经元个数和激活函数形式等会对结果产生怎样的影响，解释性较弱。因此，如何在未来的研究中透明化深度学习的黑盒，使其具有可解释性，具有重要意义。

4 结论

随着计算机技术的不断发展，机器学习尤其是深度学习在药物组合协同预测中展现出了巨大的潜力，并且有望在未来成为医生和临床药师制定药物治疗策略、监测药物治疗效果和优化药物治疗方案的理想工具。尽管目前仍然存在亟待解决的问题，未来随着高质量数据库的建立和新算法的开发，通过科研人员和临床工作者的有机协作，借助于优秀的机器学习算法和工具，可以有效克服挑战，提高预测模型在临床实践中的实用性和诊疗效率。

PU-178

人工智能技术在手术器械清点中的应用研究

靳超 孙传会

新乡医学院第一附属医院

1 背景

我国医疗行业的快速发展使得外科手术的开展从最初的粗放式操作迅速过渡到精细化、微创型阶段。各种类型的外科手术随之大大增加，手术中需要使用到各种不同形状和结构的器械。在手术前后各个阶段，按照手术安全管理规范，工作人员需要对手术器械进行多次、细致的清点。由于医疗需求的增长，各大医院手术量和手术复杂度急剧增加，高压之下必然导致人的疲乏，由此可能产生较多的失误和遗漏。为此，许多医院开始利用科学的精细化管理方法和制度来减少手术器械清点的差错[1]，并且希望借此提高器械清点效率，但都无法从根本上解决因为手术量剧增导致手术器械清点耗时多、效率低、失误多的难题。人工智能技术在一些领域的应用使得人们的生活更加便利，比如，汽车行业的自动驾驶技术解放了驾驶员的双手，智能安防更有效地维护人们的安全，AI 辅助医疗减轻了医务人员的工作压力，提高了疾病诊断的准确度，各种智能小家电使人们的生活更加丰富多彩。我国一些技术人员已经开始了相关研究，将人工智能技术应用于手术器械清点领域。

2 手术器械清点的信息化现状

从消毒供应中心的消毒处理流程到整个手术过程结束，器械包内部手术器械的清点贯穿于其中。消毒供应中心工作人员需要在器械分拣、清洗消毒、器械打包前对各类手术器械清点；手术室

护士需要在器械包拆包后、腔体关闭前、手术结束后这几个阶段分别对手术器械进行清点。一个完整的手术器械流转过程需要进行多次清点工作，每次清点耗费时间较长，效率低下且易出错，增加了清点人员的重复型机械劳动，非常容易引起清点人员的疲劳，降低其工作满意度。

现阶段，在手术器械消毒供应方面已有许多的信息化应用，一般多集中于医院的消毒供应中心和手术室，比如 CSSD（消毒供应中心）管理信息系统、移动式无菌包追溯系统、手术器械包 FRID 射频识别系统等。在此类系统中多集成有手术器械清点管理模块。

使用知网和万方两大知识检索平台搜索部分公开的医院手术器械清点信息化的应用案例发现，许多医院已经上线了消毒供应追溯管理信息系统[2]，并针对手术器械清点环节在器械包层面上采用了信息化的手段，对手术器械管理采取了全流程闭环管理，一定程度上保证了手术器械消毒质量安全。在应用技术上多采用条形码或二维码形式，部分医院购置了 PDA 移动终端[3]，方便了对器械包信息的识别和确认。也有一些医院购置 RFID 射频识别设备对手术器械包进行闭环管理[4]。

3 人工智能技术在手术器械清点中的应用

3.1 手术器械清点的智能化现状

随着机器视觉技术研究的深入，技术的不断革新给手术器械清点智能化指明了新的方向，技术人员将机器视觉图像识别、深度学习、神经网络等 AI 技术应用于手术器械清点上，使手术器械清点摆脱了完全依赖人工的窘况。

国外将图像识别技术应用在手术器械清点上较早，商业化也较为成熟，但其应用效果仍然不佳。主要方法为在手术器械上粘贴、激光雕刻微型条码或二维码，利用扫码枪或摄像头进行扫码识别[5]。手术器械较多使用金属材料，外表面光滑并需要经常进行消毒清洗，因此纸质条码易受污损；激光二维码由于受器械体积限制，形状不规则且面积较小，较容易受到光线影响，操作员扫码操作比较困难，识别效率低下。

3.2 手术器械专用数据集的构建

国内技术人员对手术器械进行图像检测时，较多使用机器视觉技术，并通过深度学习算法构建模型对手术器械进行识别，分辨器械类型，计算器械数量。受目前公开的训练数据集的限制，首先需要构建手术器械专用的数据集[6]，以便于使用其进行模型训练，构建过程如图 1。

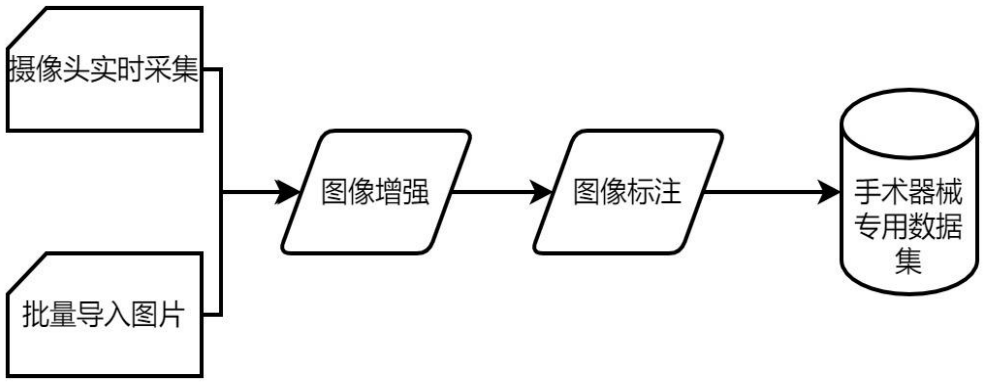


图 1 手术器械专用数据集构建示意图

优秀的算法模型并不能直接拿来使用，需要先行构建大量配套的数据集或利用大规模成熟数据集对其进行训练。在使用摄像头采集图像数据时，环境光源的影响易造成手术器械反光，图像过曝产生光斑给后面图像增强修正处理造成困扰。因此，使用稳定可靠的光源和合适的摄像角度对于手术器械的准确识别有很大帮助。同时手术器械的摆放背景对于图像识别尤其重要。背景复杂程度，颜色数量严重影响手术器械识别准确度，通常多使用手术室台面绿色无菌巾作为背景，既符合手术室器械操作习惯又增加了与金属器械的对比度。通过摄像头或图片批量导入得到的图像仍需要进行

增强修正处理，通常选择适合的图像增强算法对图片进行强化，以此来减弱环境光产生的器械表面反光影响。使用图像标注软件对修正后的图像进行手工标注，这一过程需要耗费大量的人工时间，为模型训练提供足够的训练数据。目前倾向于使用半自动或自动化的图像标注工具或算法来处理图像，以减少人工标注耗费的时间和体力。手术器械图像的标注完成后需要保存成合适的注释格式，通常的注释文件格式有 YOLO、PASCAL VOC、COCO、JSON 等。

3.3 手术器械清点目标检测模型构建

在目标检测模型选择上，更多人愿意使用开源的目标检测模型，此类模型有：YOLO、Faster RCNN 、SSD、Fast RCNN、R-CNN、SPPNet[7]。其中 YOLO 算法检测速度较快，但其对于复杂背景下，手术器械高重叠图像的识别准确率不佳。目前 YOLO 模型项目更新频繁，已经发展到了第五版。这些模型各有其应用特点，在确定最佳应用模型时，往往将多个模型同时进行训练，以比较各个模型的识别速度和识别准确率。使用目标检测模型前，需要安装该模型的配置环境，针对采集到的手术器械数据集做参数调整。深度学习目标检测模型算法主要分为两大类，一种基于锚框进行目标预测，一种无需锚框对目标进行预测。各类模型使用时有其不同的难点和特性，针对手术器械图像识别特点，多使用基于锚框的目标检测算法。此类算法有 YOLOv3，Fast RCNN、R-FCN 等。常见目标检测模型的锚框算法主要配合公开的通用数据集使用，手术器械图像外形与通用的数据集差距较大，因此，需要针对手术器械图像的特点对选定模型中的锚框算法进行调整。将准备好的手术器械专用数据采用调整后的锚框算法进行训练。模型训练好后，使用该模型进行测试，根据测试结果对模型参数，锚框算法进行调整，然后再进行模型训练。经过反复多次调整最后得到最优的目标检测模型，构建过程如图 2。以上过程仅仅是对目标检测模型大致的使用方法进行介绍，实际在对手术器械数据集进行训练时，将会遇到许多问题。图像增强修正算法、环境光源照射、摄像头的拍摄角度、手术器械摆放的重叠度等等每一项参数的变化都会对手术器械的识别准确度产生影响。在实际的应用中，如何找到性能、准确度、识别速度和图像采集便利性等因素的最佳平衡点非常关键，这有时候会决定着手术器械清点系统能否被用户所接受。

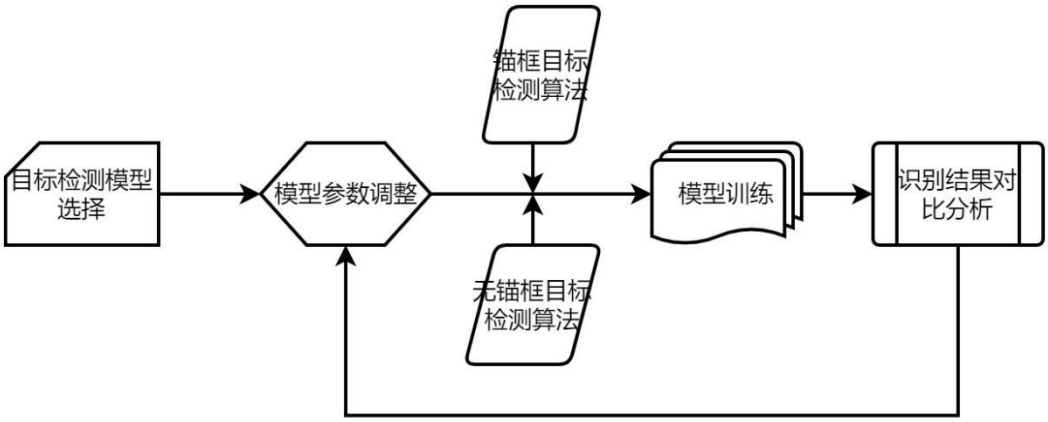


图 2 手术器械图像目标检测模型构建示意图

4 总结和展望

基于目标检测模型的手术器械清点系统目前仍处于起步阶段，虽然许多技术人员已经开发出了原始的应用程序，但成熟稳定的商业化方案仍然很少。手术器械清点的实际操作环境在每家医院的手术室都有部分差异，器械清点人员的操作习惯也有不同，在将 AI 辅助器械清点系统推广的道路上，有许多复杂的因素需要考虑。

此外，目前基于目标检测模型的手术器械清点系统应用时，器械识别与清点的效率和准确度仍无法保证完全脱离人工，完全由 AI 系统进行器械清点仍会出现漏认错认的情况，因此，许多方案的最终形式仍是系统清点加人工审核的模式。一旦人工审核时发现问题，仍需要操作人员对漏点错

点的器械进行重新清点。对 AI 系统的清点结果进行人工干预审核往往也需要花费操作人员的时间和精力。今后随着 AI 手术器械清点系统的完善，人工审核确认模块的人机交互会更加简单化和人性化，人工弥补性清点工作变得更加省时省力，甚至达到完全脱离人工确认的最佳状态。

人工智能手术器械清点系统因为使用了最前沿的深度学习技术，对数码摄像和计算机等硬件设备都有较高的硬件要求，目前的系统方案通常使用的是本地服务器，并且对服务器的配置要求较高。计算机 CPU、内存、GPU 等对人工智能手术器械清点系统运行效率和结果影响很大，普通的院内信息系统硬件要求一般都难以达到，这也会在一定程度上增加人工智能手术器械清点系统的应用成本。

5G 通信技术和云计算的迅速发展也将给人工智能手术器械清点系统的应用带来新的契机，通过 5G 网络将模型系统接入深度学习私有云中，高性能云服务器的大规模计算能力可以进一步缩短目标检测模型的训练时间，提高模型的识别效率和准确度，最终真正实现手术器械清点工作的智能化。

PU-179

移动医疗用于分娩镇痛管理的研究

刘新元 尚小平 杨林朋 吕倩 刘新奎
郑州大学第一附属医院 病案管理科

背景与目的 分娩痛可引起神经系统的过度刺激、血压升高；还可降低子宫流向胎盘的血流，引起胎儿窘迫；并可引起产后抑郁和产后创伤应激性疾病。因此减轻分娩痛是妇女分娩过程中的重要主题之一。我国的分娩镇痛率与发达国家还存在很大差距。本文通过对移动医疗用于孕妇分娩镇痛管理的研究，探讨移动医疗对于孕妇分娩镇痛管理的效果，为推行移动医疗提供理论依据。

方法 采用前瞻性队列研究方法，选取 2022 年 1 月 1 日至 2023 年 5 月 30 日就诊麻醉门诊的孕妇 580 例为研究对象。将研究对象随机分为两组：对照组 290 例，研究组 290 例。对照组孕妇就诊麻醉门诊，由麻醉师给予答疑解惑，并听取我院孕妇学校课程；研究组孕妇除了就诊麻醉门诊、由麻醉师给予答疑解惑、听取我院孕妇学校课程外，利用微信群进行移动医疗干预。

结果 ①研究组分娩镇痛率高于对照组，差异具有统计学意义（P<0.05），研究组剖宫产率与对照组无显著差异（P>0.05）；②研究组孕妇签知情同意前与麻醉师沟通时间短于对照组，差异具有统计学意义（P<0.05））；研究组孕妇签知情同意前向麻醉师咨询问题数量少于对照组，差异具有统计学意义（P<0.05）；③研究组满意率为 89.7%，显著高于对照组（82.4%），差异有统计学意义。

结论 移动医疗是现代信息技术和医学的结合，可以提高分娩镇痛率，缩短分娩镇痛前患者和麻醉师的沟通时间，可以提高孕妇满意度，具有良好的临床应用价值。

PU-180

虚拟科室在慢阻肺防治中的应用与展望

胡莉娟[1] 胡译心[1] 张静[1] 朱晓丹[1] 计海婴[1] 朱振华[1] 张捷青[1] 卞非[2] 梅园[2] 钱琨[1] 宋元林[1]

1.复旦大学附属中山医院

2.上海申挚医疗科技有限公司

慢性阻塞性肺疾病（简称：慢阻肺，英文简称：COPD）是我国常见的可预防可治疗的气道疾病。据“中国成人肺部健康研究”调查显示，我国患有慢阻肺疾病人数近 1 亿。由于公众对慢阻肺缺少深刻认知，重视程度不够，导致就医人群中重高风险和急性患者占愿意就医人群总数的81.6%；此外，二级医院与基层医疗机构的管理脱节、专业技术人员缺口、专业诊疗工具严重配备不足、基层医务人员诊疗规范亟待完善等问题，导致医疗机构长期面临巨大的慢阻肺诊治压力。

为此，复旦大学附属中山医院（以下简称：中山医院）为进一步改善就医服务，通过优化就医流程，探索全新就医模式，以数字孪生、机器学习、人工智能、大数据等关键技术为基础，构建覆盖慢阻肺专病诊治全流程的虚拟科室，实现智能临床辅助的领先探索与应用。

虚拟科室技术路线

慢阻肺专病虚拟科室建设目标

本研究主要围绕虚拟科室在慢阻肺专病诊疗全过程中的应用，通过数字化、虚拟化、AI 智能化等方式，从患者、医护人员、专病科室等角度考虑，优化虚拟科室功能，以非传统诊疗形式，实现慢阻肺患者、医生从诊断到治疗、预后管理的全流程智能辅助服务体系，以解决慢阻肺专病早筛难、普及难、诊疗不规范、预后管理不足等问题。

整体架构

结合中山医院慢阻肺专病诊室现有的信息基础和服务流程，根据数据采集、治理、存储、仿真建模、智能交互驱动的技术路线，分步实施。对专病诊治过程中的患者数据、诊疗数据进行统一采集、分类治理、统一存储和管理，为虚拟科室的算法模型提供数据基础；运用数字孪生、三维建模、物联网等技术手段，结合医务人员的专病经验、临床病例，以统一治理管理的数据为基础，将多维数据与专病诊室应用场景深度融合，实现对慢阻肺专病诊治全流程的仿真建模；以分类采集，统一治理的数据为基础，构建智能算法模型，运用人工智能、机器学习等技术，驱动虚拟科室场景，实现诊室的智能应用。

数据采集、治理、存储

基于中山院内慢阻肺专病诊室信息系统的底层架构与数据基础，通过高速数据采集、稳定数据治理和存储的孪生数据中台，根据专病诊室的信息系统、诊疗设备的开放接口，进行点位对接，治理实时数据，同时，利用数据接口实现专病诊室的数据采集。

主要采集数据范围包括专病病例、院外随访、常见呼吸疾病症状病例、多病种交叉病例、患者基础信息（例如：患者年龄、性别、生活环境、初次发病时间、是否有其他诱发性基础病、是否有家族病史等）、患者行为、医生诊断行为等相关数据，包括但不限于各类音频、视频、文字、数字等多元化数据结构。为提高数字建模的有效性与增益性，对数据进行标签、清洗、转换等处理流程。

根据不同数据性质进行多样化的存储方式，对于专病常见症状关联性更强、更具有规则的业务数据采用长期关系数据库存储；对于实时性要求高的动态数据采用即时性数据库；对于强交互性的业务数据采用独立存储库；对于结构性文本、非结构性音视频等数据采用文档云存储。

虚拟孪生仿真建模

仿真建模是针对患者就医、医生诊治整个业务流程的智能动态仿真模拟，是虚拟科室的主要业务内容。仿真建模不仅仅对于患者和医生的形象、行为、声音需要精细化、准确的孪生，更需要智能、精准的数据模型驱动。仿真建模主要分为标准化的患者孪生体建模、医生孪生体建模、专病科室几何场景建模。利用 3D MAX 专业建模软件，结合患者基本信息、专病病种特征、多病种交叉病例特征等信息创建标准化的患者孪生体；结合医生特征、就诊流程、专病病例、多病种交叉病例图谱等信息 1：1 高还原医生数字孪生体，创建数字医生；根据实际专病门诊业务流程进行数字场景模拟。

智能算法驱动

虚拟科室系统内部设置了深度学习卷积神经网络，通过智能算法对虚拟科室中的预问诊、辅助诊断等模型进行驱动训练，通过用户端的使用习惯及真实场景模拟训练，使智能模型越来越“聪明”。

当前虚拟科室已训练以下场景的智能模型：

1) 慢阻肺预诊模型，包含常见的慢阻肺的辅助检查和鉴别诊断，通过患者描述确定患者是否患有该病种。同时模型拥有呼吸内科多病种鉴别诊断的能力，根据患者的电子病历与该次预诊症状描述，推断患者在呼吸科领域存在引发多病种的概率和潜在风险；

2) 慢阻肺复查与随访模型，支持对慢阻肺患者的长期追踪、随访、量表分析、诊疗建议等功能；

3) 呼吸科患教模型，用于对患者教育，针对呼吸科常见病种的预防、判断、治疗、注意事项等知识科普。

4) 自然语言理解与生成模型和语音算法模型，通过自然语言的理解与生成，实现与患者的人机交互，通过学习现有医患对话训练语义理解模型，达到患者语言与医学术语的相互转换从而生成规范病历报告。

虚拟科室功能与应用

智能预诊

患者可在门诊等候区间，通过移动智能端在虚拟科室进行账号注册、身份验证、电子健康档案导入，根据患者全息信息创建个人的数字孪生档案。患者可通过智能设备端对数字医生发起预诊请求，数字医生在接收请求后唤起服务状态，进行数字问诊流程，患者通过 VR 设备与数字医生进行语音沟通，通过诊室内置的语音识别系统对患者描述转化成数字语言，数字医生对数字语言进行判断，从知识图谱中找到对应适应症进行辅助诊断，并进行下一步提示。患者可根据诊断结果和下一步提示提前准备面诊材料，系统将会在医生 PC 端同步该患者预诊结果，并且进行面诊准备工作的提示。

档案管理

在患者进行完一系列的信息注册、创建，以及预问诊服务等工作，该患者账号会留存归档历史信息，系统会对该条历史信息进行爬虫识别，判断知识图谱中是否有该病例症状指南，如果有，则系统将不再存储在长期数据库，仅归类存储为患者账号中的云数据；若果不存在，系统会对医生 PC 端发起请求，是否将该条病例症状描述添加至专病知识图谱中，根据医生的临床经验判断是否对现有知识图谱进行补充。

患者教育

慢阻肺的患病率高，易漏诊，预后差等问题很大原因在于患者对于该病种的认知不足，同时，由于知识能力的认知偏差，导致患者依从性差，医患矛盾大等问题，亟需解决，加强患者对该病种的教育十分有必要。

患者可在候诊期间通过几分钟短小动画，对患病潜在原因、患病常见症状、该病种的影响、一般治疗方式以及预后保健等进行沉浸式的科普教育，沉浸式的体验更便于患者加强专病认知，为后期诊治、预后工作奠定良好基础。

知识图谱

通过爬虫技术搜集医学书籍、文献中的专病症状、交叉多病症状、临床病历、专家共识以及医生临床诊断逻辑等补充专病数据，利用大数据分析、数理统计等技术对数据进行分析总结，归档梳理成专病知识图谱，同时使用深度学习卷积神经网络，在使用中不断的对知识图谱进行补充，使图谱内容越来越丰富。

虚拟科室的增益

本系统上线后为患者和医生等用户开放智能问诊、知识图谱等功能模块。基于人工智能、数字孪生、大数据、VR 等技术的虚拟科室，不仅极大地提升患者就医体验、优化门诊服务流程，还大大减轻医生繁琐复杂的工作，对慢阻肺的早筛、诊治提供技术保障，更科学、更高效地为医院专病事务管理提供示范。

在后疫情时代背景下，目前迫切需要一种非接触的新型诊治方式，满足公众日常对慢阻肺诊治的需求，缓解医院与患者的诊治压力。

思考

本研究借助人工智能、数字孪生、大数据、VR 等技术手段实现了智能辅助诊断、专病知识图谱等功能服务，取得了一些成果，但对于算法模型更仿真、更精准服务还有待进一步的提升，后期可不断地通过模型训练、智能算法、爬虫等技术对数据库进行优化，提高模型的精度与灵敏性，通过互联网、5G 等更强大的技术，对患者预后、随访进行联网跟踪，未来通过对技术的进一步探索，充分利用科技的进步与实践，提升虚拟科室的精准服务成效。

PU-181

5G 技术在智慧医疗领域中的应用探讨

张靓 王海星 汪伟*
山西医科大学第一医院

1 引言

随着计算机科学和互联网技术的迅猛发展，国民生活智能化程度的逐渐提升，我们正逐步进入了数字化、网络化、智能化的高速移动社会[1]。5G 作为新一代移动通信技术，因其传输速率高，能量消耗低，性能相对安全可靠等特点逐渐应用于智慧医疗领域[2]，为我国智慧医疗的发展提供了新的通信技术应用支持[3]。当前，我国 5G 医疗健康的发展尚处于起步阶段，在顶层架构、系统设计和落地模式上还需要不断完善，但是 5G 医疗健康前期探索已取得良好的应用示范作用，实现了 5G 在医疗健康领域包括远程会诊、远程超声、远程手术、应急救援、远程示教、智慧导诊、移动医护、智慧院区管理、AI 辅助诊断等众多场景的广泛应用。

2. 5G 智慧医疗的概念及定义

智慧医疗是当前医疗服务事业发展中所提出的新型医疗模式，是实现医疗资源整合与优化利用的重要基础[3]。目前，智慧医疗的概念已经拓展到医疗信息互联、共享协作、临床创新、诊断科学等领域。

5G 智慧医疗是指结合 5G 技术及互联网、物联网、云计算、大数据、人工智能等先进的信息通信技术，通过丰富的智能医疗应用、智能医疗器械、智能医疗平台等，将患者、医护人员、医疗设备和医疗机构等连接起来，实现在诊断、治疗、康复、支付、卫生管理等各环节的高度信息化、自动化、移动化和智能化。进一步为人们提供高质量的医疗服务，提升医疗效率和诊断水平，提高医院的运营效率，降低医院的运营成本，促进医疗资源共享下沉等一系列问题[4]。

3. 5G 技术在智慧医疗领域中的应用

5G 技术在智慧医疗领域的应用中，涉及到了多种医疗服务场景，包括了远程医疗、临床医疗、信息化建设、数据科研等应用场景，能够全面创新现代医疗服务模式，推动我国医疗服务事业发展，解决民生医疗的难题。其应用的场景主要包含以下方面[3]。

3.1 远程医疗应用场景

3.1.1 远程会诊

目前我国医疗资源不均衡，偏远落后地区还存在就医难的问题[2]。传统的远程会诊采用有线连接方式进行视频通信，建设和维护成本高、移动性差。在智慧医疗建设中，发挥智慧医疗平台的远程会诊也是推动我国医疗服务事业发展的重要内容。而 5G 网络高速率的特性，能够支持 4K/8K 的远程高清会诊和医学影像数据的高速传输与共享，实现患者数据信息的高速传输与同步调阅，进一步实现传统医疗服务中会诊模式的转变。同时，患者也可以利用 5G 技术的高清视频通信，与医生进行远程沟通，并通过监测设备，将自身的生命体征数据信息传递给医生，实现患者与医生之间的实时诊断，全面提升诊断工作的效率。并且远程会诊的应用，也能够促进我国传统医疗会诊模式的改革与发展，全面提升医疗诊断的质量，对于患者的医疗服务工作开展，有着积极的促进意义[3]。

3.1.2 远程超声

远程超声技术包括远程通信、信息学及超声医学三大技术[5]。由于基层医院往往缺乏优秀的超声医生，故需要建立能够实现高清无延迟的远程超声系统，充分发挥优质医院专家优质诊断能力，实现跨区域、跨医院之间的的业务指导、质量管控，保障下级医院进行超声工作时手法的规范性和合理性。在智慧医疗领域的诊断中，远程超声诊断是重要的诊断方式。在实际诊断中，通过 5G 通信技术，可保证数据传输的稳定性和高速性[3]，医生可以通过远程视频通信功能，通过机械臂操纵超声检测设备对患者进行检查诊断。

3.1.3 远程手术

远程外科手术也是当前智慧医疗建设中重要的内容之一，是解决偏远地区医院手术能力不足问题的重要方式。在远程外科手术中，医生通过 5G 通信技术的低延时、高带宽的技术保障，利用医工机器人和高清音视频交互系统，以遥感操纵机械臂对基层医疗机构的患者进行及时的远程手术救治[3]。利用 5G 网络切片技术，可快速建立上下级医院间的专属通讯通道，有效保障远程手术的稳定性、实时性和安全性，让专家随时随地掌控手术进程和病人情况，实现跨地域远程精准手术操控和指导，对降低患者就医成本、助力优质医疗资源下沉具有重要意义。

3.1.4 远程示教

医疗教育指面向医疗卫生技术人员进行的教育培训，用户包括医疗、护理、医技人员。医学继续教育主要分为会议讲座、病例讨论、技术操作示教、培训研讨、论文与成果发表等形式，可线下组织也可线上远程进行。5G 手术示教指通过对于医院手术相关病例进行直播、录播等形式进行教学培训，核心功能包括手术图像采集、手术转播、手术指导、手机等移动端应用等，主要面向医院普外科、麻醉科、心外科、神外科等外科相关科室医疗技术人员，旨在提高外科相关科室医护人员案例经验及实操水平。

3.2 临床医疗应用场景

3.2.1 健康数据监测

健康数据监测中 5G 技术的应用是以患者的电子健康档作为核心的，对于患者健康数据的监测与管理是智慧医疗体系的重要功能，更是现代医疗服务智慧化发展的重要体现。而智慧医疗的远程健康数据监测，需要对患者的生命体征信息进行实时监测与分析管理，这就需要通过 5G 技术来实现对患者生命体征监测数据的实时传输，以保证医生能够根据患者生命体征的监测数据，及时为患者提供远程医疗服务[3]。随着 5G 技术发展，将有更多远程监护医疗设备接入互联网，健康信息可传输到专业健康机构，通过大数据实时监控用户身体健康，及时提供健康指导和预警[6]

3.2.2 应急救援急救

智慧医疗领域应用 5G 网络通信技术，在急救的过程中实现医疗急救中院前急救与院中治疗的有机结合，实现医疗急救的一体化发展与急救衔接[3]。应急救援通过 5G 网络实时传输医疗设备监测患者生命体征和危急报警信息、车辆实时定位信息和车内外视频画面至远端专家侧，获得上级专家远程指导，同时远程监护可使医院在第一时间掌握病人情况，预先制定急救治疗方案并准备所需资源，做到院前急救与院内救治的无缝对接，大大缩短抢救响应时间，为病人争取更大生机。5G 技术可实现安全可靠医疗数据传输，进行信息资源共享和信息系统的互联互通，为院前急救、智慧医疗提供强大技术支撑[7]。

3.2.3 AI 辅助医疗

随着计算机技术和医学影像技术的不断进步，医学影像已逐渐由辅助检查手段发展成为现代医学最重要的临床诊断和鉴别诊断方法。5G 智慧医疗解决方案以 PACS 影像数据为依托，通过大数据+人工智能技术方案，构建 AI 辅助诊疗应用，对影像医学数据进行建模分析，对病情、病灶进行分析，为医生提供决策支撑，提升医疗效率和质量，能够很好的解决我国的医学影像领域存在诸多问题。

3.3 信息化建设应用场景

3.3.1 智慧导诊

随着医疗体制改革的不断深入，利用现代医疗信息化手段，优化就医流程，让广大患者有序、轻松就医已成为医院提高服务水平的迫切需求。医院通过部署 5G 智慧导诊机器人，利用 5G 边缘计算能力，提供基于自然语义分析的人工智能导诊服务，可以提高医院的服务效率，改善服务环境，减轻大厅导诊台护士的工作量，减少医患矛盾纠纷，提高导诊效率。

3.3.2 智慧院区管理

智慧医院管理利用 5G 海量连接的特性建设构建院内医疗物联网，使医院海量诊疗设备和非医疗类设备有机连接，能够实现医院后勤和人员管理、患者体征实时监测和院内急救调度、门禁安

防、院内导航等服务，提升医院管理效率和患者就医体验[7]。5G 网络技术所具有的网络传输、接入、共享功能为其在智慧医院建设中的应用和全面提升医院现代治理能力提供了相应的技术支撑，整合的智慧医疗体系除去了医疗服务当中各种重复环节，降低了医院运营成本的同时也提高了运营效率和监管效率[8]。

3.3.3 移动医护

移动医护将医生和护士的诊疗护理服务延伸至患者床边。在日常查房护理的基础上引入 5G 网络技术，可以实现病人与医生的无阻碍交流，医生也可以随时对病人的情况进行监测，进而可以更好的观察病人的实际情况，可以在最短的时间内找到最为合适的医疗方案。在 5G 网络技术的背景之下形成的临床护理机制，可以有效的对病人体温、血压等指标进行实时的监控，进而可以制定一份详细的生命体征表和测评单。除此之外，医生在进行查房工作时，还可以根据病人当时的实际情况修改病历，将要施行的医疗方案或是要服用的药物录入到系统中，可以清楚的对病人的情况进行实时登记，病人根据医嘱进行治疗，可以降低出现问题的风险，在一定程度上可以提升医院的医疗质量[9]。

3.4 数据科研应用场景

目前各医院的医疗服务是相对独立的，检查、检验、治疗等都分别在不同的医疗部门进行。虽然医院信息系统网络普遍可以支持院内信息共享，但在网络通信方面，由于受容量和传输能力所限，接诊医生还不能随时查阅患者在其他医疗机构就诊的历史病例资料，限制了区域医疗的进一步发展。5G 网络可以发挥大带宽、低时延、网络切片等优点，助力解决区域内不同医疗卫生机构之间的网络互联互通、信息共享和医疗资源合理分配等问题。此外，通过 5G 网络，数据高效流通，各医疗机构之间可形成数据中心[2]。这些数据看似庞大、杂乱，但其中又包含一定的医学规律，与数据挖掘、人工智能等技术结合，催生医学影像分析、健康管理和疾病预测等医疗应用场景。同时可对患者的数据与历史数据进行对比分析，找出其中的特征关系，建立疾病与患者临床表现之间的推理关系，在动态、有效预测病情发展的基础上，制定合理的医疗及护理计划，提高医疗质量，减少医护人员的工作量[10]。

4 5G 技术在智慧医疗领域中挑战及措施

随着 5G 和相关技术的发展，更多的医疗健康领域不断涉及，更多的用户场景成为现实，但是在 5G 技术的实际应用中，仍然面临着一些现实的挑战，给智慧医疗的建设与发展造成了不利影响。目前阶段 5G 技术在智慧医疗领域应用所面临的挑战及应对措施主要体现在以下几个方面：

4.1 面临的挑战

4.1.1 稳定性

5G 的网络覆盖面积以及网络的稳定性对医疗领域的众多应用场景要求极高。5G 网络的建设速度很快，但是因为智慧医疗的主要应用场景是在室内，很难完成 5G 信号的室内完全覆盖。这需要运营商根据医院的不同建筑结构，不同科室的功能分区，重新布设 5G 室分系统，实现医院室内、室外 5G 信号全覆盖[4]。同时，在基站的运行过程中，耗电量都较为昂贵，而 5G 基站的耗电量是 4G 基站的 4 倍左右，基站建设数量的大量增加，就造成了 5G 技术在智慧医疗领域应用中电费成本巨大，增加了 5G 技术在智慧医疗领域的应用难度[3]。

4.1.2 安全性

5G 技术在智慧医疗领域应用面临的信息安全困境是网络通信技术发展中所面临的关键问题之一。当 5G 技术在智慧医疗领域的实际应用中，涉及到人民群众的海量健康数据、诊疗数据、用药数据、大量医学影像数据以及个人隐私数据等，需要通过 5G 通信技术来实现医疗数据的传输与储存，这就对数据安全问题提出了更加严格的要求，然而现阶段的 5G 网络传输信息安全保障问题依然没有得到有效解决，这也给 5G 技术在智慧医疗领域的应用造成了不利影响[3]。加强智慧医院、智慧医疗领域的数据监管是保护患者隐私的重中之重，建立健全智慧医疗领域的安全监管体系，才能确保智慧医疗的可持续发展[4]。

4.1.3 标准化

当 5G 技术在新的医疗应用场景出现，与之对应的是新的终端设备和仪器的互联互通。目前缺乏统一的标准与评价体系，无线医疗应用场景众多，不同应用场景对于网络的需求差别较大，设备、网络、业务应用、数据互通互联、信息安全标准的都存在差异，目前尚无健全统一的 5G 智慧医疗健康标准体系[2]；要建立完善仪器设备的质量标准体系、技术标准体系、数据标准体系、接口标准体系等，需要通过逐步完善 5G 智慧医疗、5G 智慧医院相关的技术标准体系，才能有效地对新式医疗健康终端设备、仪器进行检测和质量把关，才能加快 5G 技术与医疗行业的相互融合[4]。

4.2 应对措施

4.2.1 加强信息化建设

加强机构内部 5G 基础网络建设，加强信息化人员培养，提高医院信息化水平；医疗机构及时跟进国家 5G 发展情况，遵循最新 5G 标准进行建设及优化、升级，实现可持续发展[6]。基于此，5G 技术在智慧医疗领域应用发展的实际过程中，就需要加强基站建设，为智慧医疗事业的发展提供基础设施保障[3]。

4.2.2 注重网络安全

加强信息安全建设，需要加强数据安全保障技术的研发，将之应用到智慧医疗信息服务平台之中，实现对 5G 技术下智慧医疗服务数据的安全保护，避免数据传输中信息安全问题的产生[3]。由于医疗行业对数据的安全性、可靠性及数据质量要求非常高，为确保医疗安全及患者隐私等问题得到保障，应创新安全管理模式，加强网络安全技术的研发，健全安全管理制度、操作规程和技术规范，以保证 5G 智慧医疗网络、数据、设备等安全可靠[2]。同时加强内部人员思想道德建设及职业技能培训，杜绝内部人员泄露信息；强化医院内部监管，对关键信息、设备进行监控，定期安排人员检查[6]。

4.2.3 建立健全标准体系

建立健全标准体系，结合医疗健康行业应用特点，接入设备多、接口标准不一、数据格式多样的现状，不断完善标准体系研究，推动面向医疗行业的 5G 标准体系的制定、实施和应用，在统一安全标准的基础上，加快信息的互通互联[2]。健全相关政策机制，为 5G 技术在智慧医疗领域应用提供保障随着我国 5G 技术体系以及智慧医疗服务事业的逐渐发展，5G 技术在智慧医疗领域的应用试验也愈加成熟。面对这一现实发展趋势，国家政府职能机构就需要根据实际情况，注重对相关政策以及法律机制的完善与补充，以此建立 5G 技术在智慧医疗领域应用的上下联动关系，发挥政府职能机构对 5G 技术在智慧医疗领域应用的引导与推动作用，解决 5G 技术在智慧医疗领域应用的现实困境，全面推动我国医疗服务事业的发展[3]。

5.总结

智慧医疗作为 5G 的重要应用领域，发展前景广阔。智慧医疗与 5G 的结合，将加速远程医疗的发展和落地，有利于优质医疗资源向地方医院、隔离区或交通不畅的地方输出，有效缓解地方医院或诊所尖端医疗人才和高端医疗仪器不足的现状。随着医疗行业融入更多的人工智慧，借助 5G 技术智能化，使智慧医疗走进寻常百姓生活[11]。随着运营商对 5G 网络建设的加快，智慧医疗、远程医疗等将充分发挥其移动、实时、多样化的特点，以临床需求为导向，解决具体问题，真正让 5G 与医疗健康融合在一起，为广大人民求医、问诊提供便利，为我国医疗健康卫生事业的快速发展做出贡献[11]。

PU-182

基于深度学习的心血管病风险预测模型研究

洪石陈[1]，王逸嘉[2]，丁祎[1]，关红伟[1]

1.上海大学附属南通医院（南通市第六人民医院）

2.南通科技职业学院

1 引言

近年来，心血管病在城乡居民总死亡原因排名中处于高位，长期影响着中老年患者的身体健康和生活质量。心血管病具有病程长、恢复慢等特点。因此，一旦患上心血管疾病，不仅病人需要承担身体和精神上的痛苦，社会也要付出大量的医疗资源[1]。而在患者确诊之前及时进行风险预测，结合预测结果积极的给予患者科学的指导，改变生活方式、改善饮食、合理用药，可极大的降低老年病的患病风险[2]。

虽然目前在临床上可以借助影像学技术较为准确的对心血管病进行诊断[3]，但一些影像技术不仅费用较为昂贵，而且对患者身体有一定的伤害（如血管造影术），同时也对于医院的医疗水平要求较高。

近些年，部分学着开始利用机器学习方法对患者电子病历中的风险因子进行研究分析，从而实现疾病的风险预测[4]。这类方法不仅对患者的身体没有伤害，同时大幅降低了成本，因此已逐渐成为当今具有重要现实意义和挑战性的研究方向。

本文基于深度学习技术，充分利用电子病历中大量有价值的患者历史数据和各类异构数据设计了一种基于多源异构数据的心血管病深度预测模型(Deep Prediction Model for Cardiovascular Disease Based On Multi-source Heterogeneous Data，DPM-CDMHD)，该模型在实验中相比基线模型能够更充分的从多种异构数据中获取信息，实现更准确的心血管病风险预测。

2 基于多源异构数据的心血管病深度预测模型

2.1 数据描述

本文的实验数据采集于南通市第六人民医院大数据集成平台中所积累的真实医疗数据，经过脱敏和数据处理制作成本文研究所需的医学数据集，其中包含 19365 条心血管病患者的记录，进行初步筛选后用于本文研究的数据特征如表 1 所示。

表 1 候选特征表

输入特征	特征名称
diag_code	诊断编码（多个）
tests_res	检验结果（多个）
exam_res	检查结果（多个）
age	年龄
gender	性别

patient_type	患者类型
dept	就诊科室
height	身高
weight	体重
sur_history	手术史
visit_num	就诊次数

2.2 数据处理

首先，设置时间段构成观测窗口和预测窗口，训练数据从观测窗口的患者电子病历数据中取出，分类标签则从预测窗口中获得。

其次，对于结构化数据，从患者的电子病历历史数据中聚合整理出每次就诊的记录（将时间间隔较小的就诊记录视为同一次），并根据时间顺序排序（为保证时序数据的质量，去除掉就诊次数较少的患者的数据）。

最后，对于非结构化的患者既往史、家族史等文本数据，先利用停用词词典删除了文本中的无关内容，再使用分词工具分词后将文本长度进行统一。

2.3 数据表示

首先，患者就诊序列中，每一次就诊记录以一个一维向量表示，向量的每个维度代表一种医学编码[5]（包含诊断编码和检验检查结果编码）。对于诊断编码，如患者有此诊断，则该维度值为 1，否则为 0。对于检验结果编码，如患者的检验指标在正常值之内，则该维度值为 1，否则为 2，未做的项目则记为 0。检查结果编码的处理方法与检验结果编码相同。单次就诊记录向量如图 2-1 所示。

诊断1	诊断2	诊断3	诊断4	...	检验1	检验2	检验3	检验4	...	检查1	检查2	检查3	检查4	...
1	0	1	0		1	0	2	0		1	1	2	2	

图 2-1 单次就诊记录向量

其次，对于电子病历中的患者人口学数据用 One-Hot 向量表示。具体来说，数值型数据经过分桶处理，映射成 One-Hot 向量表示。ID 类数据直接转化成 One-Hot 向量表示。

最后，患者既往史、家族史等文本数据在预处理阶段分词后，同样转化成 One-Hot 向量表示。

2.4 模型架构

模型由输入层、嵌入层、特征提取层和特征融合层组成，最后通过输出层得到模型的预测评分，如图 2-2 所示。

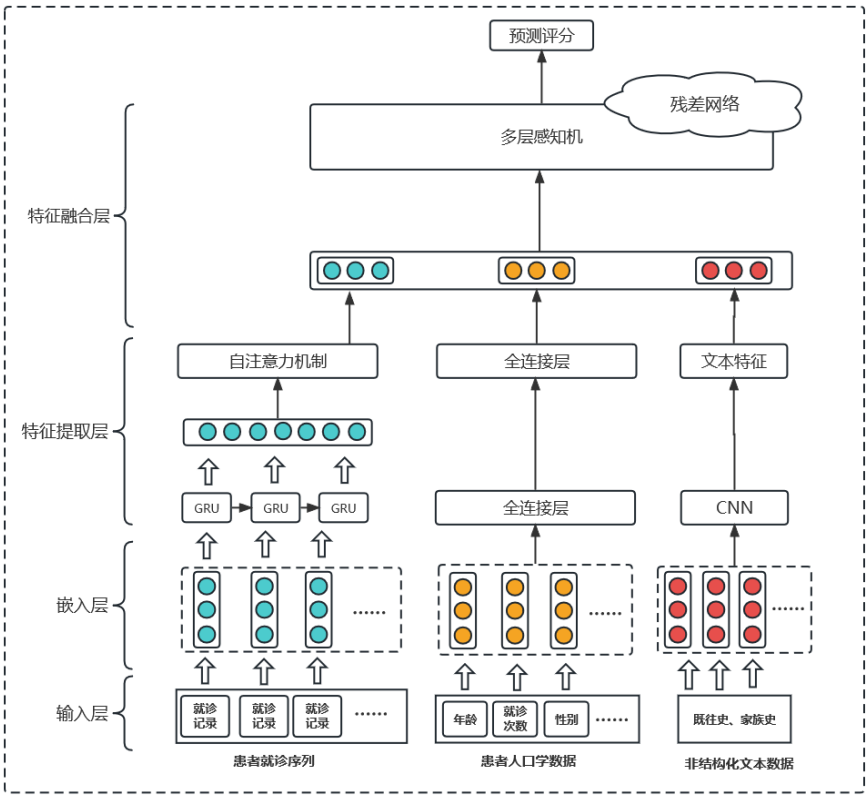


图 2-2 DPM-CDMHD 模型结构图

1） 输入层。

此部分的输入为患者就诊历史序列、患人口学数据以及非结构化文本数据，其中的患者人口学数据和既往史、家族史等文本数据均为 One-Hot 表示，患者就诊序列中每一次就诊记录由预处理完成的医学编码向量表示。

2） 嵌入层。

输入层中，大量 One-Hot 形式表示的特征导致了输入层向量是高维稀疏的，特别是在患者人口学数据和非结构化文本数据中，大量 One-Hot 形式表示向量使得输入维度非常高，会造成较大运算压力，使模型训练效率低下[6]。经过嵌入层的处理，可以将高维稀疏的 One-Hot 形式向量转化为低维稠密的嵌入向量，缓解了运算压力[7]。

以患者就诊序列为例，序列中某次就诊 y_j，式 2-1 展示了其嵌入过程

$$y_j^e = M^T y_j \qquad (2\text{-}1)$$

其中，M 为一个初始化后的嵌入词矩阵，维度是 $n \times d$，其中 n 表示所有特征数量，d 是经过嵌入层后，嵌入向量的维度。经过嵌入层之后，患者就诊序列的表示如式 2-2 所示。

$$Y^e = \{y_1^e, y_2^e, \cdots y_{t_1}^e\}$$

3）特征提取层。

患者就诊序列在经过嵌入层后，利用门控循环单元（Gate Recurrent Unit，GRU）对用户每次的就诊医学编码向量分别进行聚合，得到代表患者真实就诊信息的潜在关系序列

$X = [x_1, x_2 \cdots x_n]$。

将潜在关系序列作为自注意力机制的输入[8]，自注意力机制的计算如式 2-3、2-4、2-5 和 2-6 所示。

$$Q = W_q X \quad (2\text{-}3)$$

$$K = W_k X \quad (2\text{-}4)$$

$$V = W_v X \quad (2\text{-}5)$$

$$H = V\ \mathrm{softmax}\left(\frac{K^T Q}{\sqrt{D_k}}\right) \quad (2\text{-}6)$$

其中，W_q、W_k和W_v分别为查询向量、键向量和值向量的参数矩阵，分别与潜在关系序列计算得到查询向量矩阵Q、键向量矩阵K以及值向量矩阵V，进而通过打分函数得到输出序列H，D_k为K的维度。输出序列中向量拼接后得到患者就诊序列特征向量P_1

患者人口学数据经过嵌入层后，经过两层全连接层进行融合，得到更高阶的患者人口学数据特征向量P_2。

既往症、家族史等非结构化文本中的词向量经过嵌入层处理后，利用卷积神经网络完成文本特征提取，融合出完整的患者相关病历文本特征向量。首先，将嵌入层输出的词向量矩阵 M 输入卷积神经网络。矩阵 M 的维度是$m \times n$，其中 n 为词向量的数量，m 为词向量的维度。然后，通过多个不同大小的卷积核处理输入矩阵，计算过程如式 2-7 所示。

$$z_i = f\left(W_j * V_{(:,\ i:(i+h-1))} + b_c\right) \quad (2\text{-}7)$$

其中，$V_{(:,\ i:(i+h-1))}$为词向量矩阵 M 中第 i 到$i+h-1$个词向量组成的子矩阵，b_c为偏置项，$*$表示卷积运算，$f(\cdot)$为 Relu 激活函数。随后，用最大池化的方法从特征组每个区域中取最大值作为此区域的表示。式 2-8 为最大池化方法的数学表示。

$$f(x) = \max\{0,\ x\} \quad (2\text{-}8)$$

最后，化层的输出d_i，通过全连接层，生成维度固定的文本特征向量。过程如式 2-9 所示。

$$P_3 = \mathrm{Relu}(w_i d_i + b_i) \quad (2\text{-}9)$$

其中，P_3为经过全连接层后生成的职位文本特征向量，w_i为权重参数，b_i为偏置项。

3） 特征融合层。

电子病历中的三种异构数据，经过特征交互层后，在特征融合层进行拼接后，输入到以残差网络作为具体实现的多层感知机中以学习更高阶的患者特征，最后经过输出层得到最终的预测评分。

3 实验结果与分析

3.1 评价指标

本文所用数据集来源于公立医院的真实临床数据，数据集中正负样本比重不够平衡(约为1:4.8)，负样本的比重较高。因此，为了真实反应模型的预测性能，本文选择精确度、召回率和 F1 值作为评价指标。各指标的计算过程如式 3-1、3-2 和 3-3 所示。

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (3\text{-}1)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (3\text{-}2)$$

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \quad (3\text{-}3)$$

3.2 实验结果

本文提出的 DPM-CDMHD 预测模型旨在将多源异构的医疗数据融入模型，以提高心血管疾病的预测准确性。因此，我们将 DPM-CDMHD 模型去除患者人口学数据处理部分以及去除非结构化文本数据处理部分的两种简化模型作为两种基线模型（记为 METHOD-1 和 METHOD-2），加入到实验中，用以分析多元异构数据对模型性能的影响。同时，将支持向量机、逻辑回归和随机森林三种常用模型作为额外的三种常用基线模型[9]进行实验对比。实验结果如表 2 所示。

表 2 实验结果

模型(方法)	Precision	Recall	F1
DPM-CDMHD	**0.6578**	**0.7532**	**0.6881**
METHOD-1	0.6241	0.7158	0.6609
METHOD-2	0.6426	0.7442	0.6733
支持向量机	0.5993	0.6838	0.6355
逻辑回归	0.6022	0.6892	0.6420
随机森林	0.6327	0.7025	0.6652

实验结果表明，本文提出的 DPM-CDMHD 相比其他五种基线模型在精确度、召回率和 F1 值三种指标上均有最优表现。其中，在 DPM-CDMHD 基础上简化后的 METHOD-1 和 METHOD-2 在三种评价指标上均与 DPM-CDMHD 有所差距，说明在医疗场景中，多源异构数据的加入能够提升模型的性能。METHOD-2 在三种评价指标上的表现均略优于 METHOD-1，说明人口学数据对于模型性能的影响略大于非结构化文本数据。总体来看，基于深度学习方法和多源异构数据的 DPM-CDMHD 预测模型在本文实验中的表现是最优的。

4 结语

本文针对心血管病对患者健康深入且长期的影响，基于深度学习技术和医学多源异构数据，提出了一种心血管病风险预测模型 DPM-CDMHD。该模型融入电子病历中各类异构数据，利用循环神经网络、注意力机制、卷积神经网络等深度学习技术充分学习患者的全方位信息，使多源异构的医疗数据得到充分有效利用。通过实验验证，DPM-CDMHD 的预测性能均优于其他基线模型，能够更精确地预测患者是否患有心血管疾病或未来几年可能患上心血管疾病的风险，帮助患者预防心血管病，并帮助医疗机构提供个性化的医疗服务。

PU-183

基于 5G 技术的医院双域无线专网建设与应用

左锋、陈利佳、胡冰、金泠漪

陆军军医大学第一附属医院信息科

传统 WiFi 网络为移动医疗的普及推广发挥了巨大作用，也推动了在线教学的发展，但随着应用的深入和技术革新，其弊端也逐渐显现出来[1]。如 WiFi 信号强弱受距离、穿墙等环境制约，移动终端普遍存在信号弱、网速慢的问题。另外，WiFi 无线信号盲区和网络干扰现象普遍存在，严重影响了用户的使用体验。而且，因 WiFi 信号覆盖区域性限制，无线数据不出院区，无法满足当前院外至院内教学、医疗网的访问需求。而 5G 专网以其大带宽、低时延和高安全的特性，彻底解决了上述问题，是取代 WiFi 网络的有效方式[2-3]。重庆西南医院通过建设覆盖全院的智慧医疗和教学 5G 双域专网，为院内移动医护、智慧教学、远程示教等场景提供了高速便捷的无线连接以及院外到院内医疗、教学网络的稳定安全访问。

1 系统设计

我院对应用场景的实时性、可靠性与安全性均有严格要求，在现有的以云计算为核心的集中式数据处理模式下，业务流量需经由接入网和核心网进行交互，这种模式的突出特点是具备较强的计算和存储能力，但是在面对诸多新业务所提出的极致体验需求时则存在不足。一方面，所有的业务流均通过云计算中心进行处理，时延和拥塞将严重影响业务体验，无法满足超低时延的要求；另一方面，随着接入终端数的迅速增加，海量数据回传会对运营商接入网和核心网形成巨大挑战，进而降低网络的运行效率[4-5]。

1.1 系统架构

根据实际使用需求，建设两张专网：教学 5G 专网和医疗 5G 专网，院内在物理上实现两张网的隔离。

1.1.1 教学 5G 专网

2C 业务：利用大区共享 UPF，通过 100M 数据专线打通共享 UPF 和教育专网应用系统的传输链路，实现院内、院外的教职工及学生访问内网资源。2B 业务：建立一套规格为 5Gbps 的 UPF，下沉院内教育专网机房，满足远程手术示教的数据不出院区。

1.1.2 医疗 5G 专网

2B 业务：建立一套规格为 10Gbps 的 UPF，下沉我院教学楼医疗网机房，支撑院区内的 5G 医疗应用系统。2C 业务：利用 mini MEC 建设专网，下沉我院外科楼移动机房，支撑院区内的 5G 医疗应用系统。

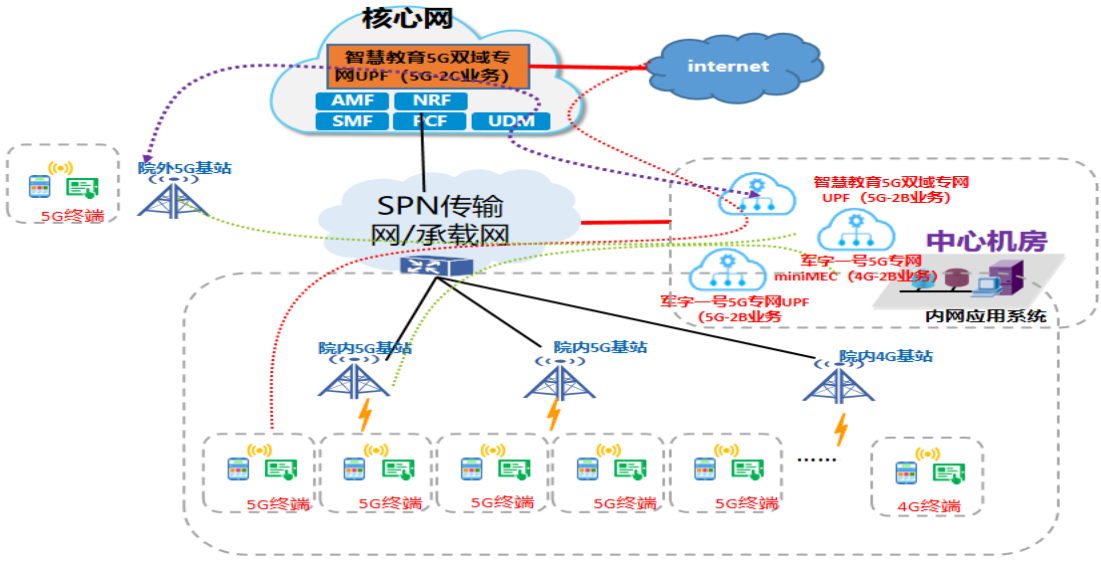


图 1 系统架构

系统使用移动边缘计算（MEC）技术，在提高网络资源利用效率的同时，有效满足各类业务迅猛发展对时延和带宽需求，而且为运营商网络与互联网、物联网的深度融合提供技术基础。医疗 5G 专网，满足医疗系统应用数据的安全性要求（5G-2B 业务），HIS、LIS 等业务数据不出院区[6]。

1.2 5G 专网总体设计

搭建院内 5G 传输专属通道，满足全院 5G 应用需求，实现全院 5G 网络连续覆盖并支持重要教学区域 5G 用户以及终端接入内网需求，实现覆盖区域 5G 网络信号强度 RSRP>-110dBm，信噪比 SINR>-3dB。 数字医疗 5G 双域专网平滑升级为同时具备 5G SA 和 NSA 双模组网架构，支持院内多场景多终端接入院内内网，满足大带宽、大链接、低时延的网络传输要求，上行传输网接入带宽直接平滑升级到不低于 10G，网络下行速率≥500Mbps[7]。统一 5G 边缘计算数据交换标准和接口规范，统一安全接入准入机制，统一数据加密、隔离、安全域划分等标准，统一业务数据和管理数据的协议转发标准。

制定统一的终端数据接入标准，自主适配全院多终端主流标准协议的 5G IOT 设备，统一设备管理，满足院内核心平台对前端 5G IOT 终端的统一调度，包含但不限于智慧教学平台、数据安全交换平台、智慧后勤管理系统、智慧会议平台、智慧营房管理系统、HIS、PACS、LIS 等系统[8]。支持升级改造后的教室、学习室、手术室内 5G+MEC+有线网络自适应传输接入，满足多媒体、物联网管控、精品录播课程建设、课堂内师生交互、远程互动、远程会诊、手术示教等应用的 5G 传输接入。

1.3 安全防护设计

院内 5G 网络涉及的安全内容包含接入安全、网络安全、网络管理安全、业务安全和数据安全，支持 NSA 和 SA 网络中物理环境下的设备组成独立的 5G 安全域，在不同安全域直接实施安全连接防护。同时，SA 系统设备不同安全需求以及与外部网元之间的关系，支持划分 5 个网络安全域：无线接入安全域、核心网安全域、互联网安全域、网管安全域、计费安全域；终端用户与接入网络加密保护，SA 支持分层密匙体系和密匙生成要求；支持用户向 5G 网络注册/去注册，空闲/连接不同状态转换时的安全上下文处理[9]。

1.3.1 客户终端安全

5G 终端在 5G 网络接入方面，认证协议上应遵循 EAP-AKA 以实现统一框架下的双向认证。防止终端接入非授权切片：在行业需求中，既需要防止非授权终端接入本行业的切片，也需要防止本行业终端误接入到其它切片；针对非授权接入切片：在终端完成接入认证后，应先由网络识别终端的切片标识信息，然后再建立会话，保障接入的合法性。对本地和远程登录的用户进行身份鉴别和

授权，对终端用户登录、终端管理系统登录及其他系统级应用登录进行身份鉴别，确保使用者身份的合法。

1.3.2 接入安全

网络接入方面，认证协议应使用规范的 EAP-AKA 认证机制，以实现统一框架下的双向认证，并使用 5G-AKA 增强归属网络控制，防止终端接入虚假网络。同时，在 USIM 卡增加设定的公钥，在首次附着网络使用公钥加密 IMSI，解决初始接入身份泄露问题[10]。mMTC 场景中存在大量的终端需要接入网络，这些终端通常为低功耗资源受限设备，则极容易产生信令风暴，造成网络拥塞。针对该问题，需对传统认证机制进行局部优化来满足海量终端接入网络的要求。

1.3.3 网络切片安全

5G 网络切片为不同业务提供差异化的安全服务，根据业务需求针对切片定制其安全保护机制，实现客户化的安全分级服务，同时网络切片也对安全提出了新的挑战，如切片之间的网络资源隔离、安全策略定制、切片内安全控制、业务端到端加密等；5G 网络切片的做到信息隐私保护，在安全上下文建立后，对切片选择信息必须使用机密性保护。采用认证机制，实现切片内 NF 与切片外公共 NF 间相互可信；切片管理设置相应的安全保护机制，包括切片管理服务使用双向认证、授权机制，切片管理系统及切片网络间通信需做完整性、机密性保护及防重放攻击[11]。

2 主要应用场景

通过该项目建设，建立覆盖全院的教学 5G 和医疗 5G 双域专网。实现全院覆盖，师生、医护人员、管理人员都可以通过智慧教学 5G 专网实时访问教学资源，开启视频会议，实现远程示范和教学，随时随地进行流畅的学习与交流；医护人员可以通过智慧医疗 5G 专网开展移动医护、远程诊断等医疗业务，将智慧医疗全面推广。主要应用场景如下：

2.1 移动医护

移动医护是院内医护人员利用移动终端进行医护协同工作与医疗差错管控工作的场景[12]。主要包括移动护理、移动查房、移动输液、护理质控、护理白板等多种业务场景。

图 2 移动医护场景

2.2 5G 医疗影像云

5G 医疗影像云业务包括影像存储、手机阅片等场景。通过前置机将影像数据传输到云端，实现影像数据的云端存储和云端调阅。支持医生手机阅片，终端通过院内 5G 专网接入院内 IDC，通过 IDC 访问院内数据。

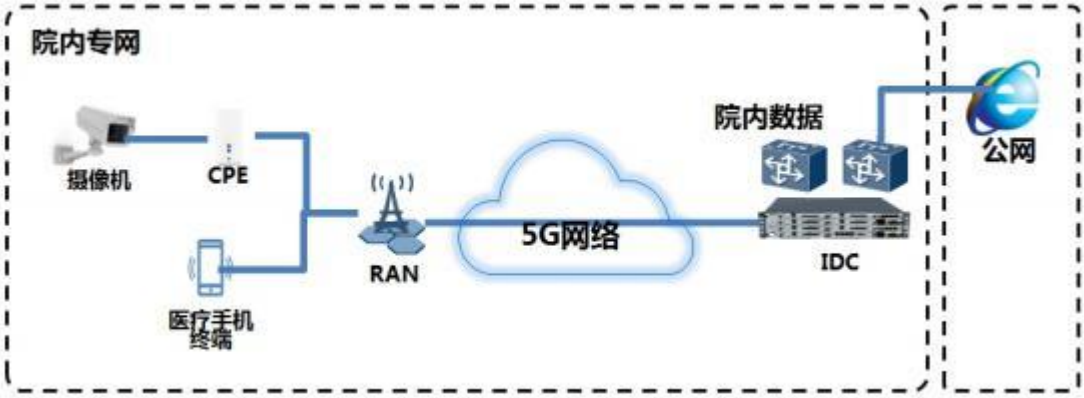


图 3 医疗影像云场景

2.3 5G 远程教学

我院 5G 智慧教育主要依托 5G 网络的连续广域覆盖、热点高容量、低时延高可靠等特点，打造特色智慧教学 5G 专网，实现超高速、超便捷、高安全的智慧教室网络体系建设，为远程教学、智慧教学提供可靠稳定且快速的网络环境，为师生提供优秀的智慧教学服务体验，解决智慧教室各设备间无缝交换连接。

2019 年我院打造了 5G 专网一期项目，针对教学楼重点区域、教室等进行了覆盖。2020 年打造了智慧教学能力提升项目、 教室智慧化升级、学习室智慧化升级多个智慧教育项目。2023 年我院又进行了 5G 专网全院覆盖，扩大 5G 专网覆盖区域，将 5G 专网一期项目平滑过渡，打造校园网智慧教学 5G 个人终端专网、校园网智慧教学 5G 业务终端专网、医疗网 5G 业务终端专网，针对智慧教学和智慧医疗的不同应用提供不同的网络保障，满足全院师生随时随地接入智慧教学网络的需求，提升智慧教学的效率[13]。

2.4 5G 床旁会诊

以医疗手推车为载体，可搭载视频会议系统，基于 5G 网络传输优势，实现现阶段省内院区与院区之间、上级医院与基层医院之间的高清视讯连接。5G 床旁会诊搭载院内信息系统，可为远端专家提供患者病历、影像信息，提高远程会诊效率。

图 4 5G 床旁会诊场景

2.5 远程超声

远程超声由远端专家对基层医院的患者开展超声检查，用于现阶段省内上下级医院，及偏远地区对口援助帮扶，提升基层医疗服务能力。通过项目的建设，构建“高速便捷、稳定安全”的新型智慧医疗 5G 专网，为我院智慧医院建设打好坚实基础，增强群众对健康服务的获得感、幸福感、安全感。

图 5 远程超声场景

3 应用效果

3.1 5G 双域专网打通了从院外至院内业务的大容量访问通道。

5G 双域专网为院内业务提供了大规模设备接入的通道。医院 5G 智慧教学 5G 专网容纳七千余部教职工及学生的 5G 手机（5G-2C 业务），可在院内院外无缝访问内部网络资源；同时支撑 20 个 5G 终端接入（5G-2B 业务），用于对远程教学的示范效应。医疗 5G 专网容纳七百余个 4G 移动医护的接入（4G-2B 业务）；同时保证上百个 5G 终端接入（5G-2B 业务），用于支撑远程心电等应用[14]。满足教职工及学生对内网资源的访问需求，市内不限区域访问（5G-2C 业务）；同时满足了院内各种 5G-2B 业务应用要求。

3.2 5G 双域专网为院内各种移动应用场景提供了高速安全的数据传输网络。

5G 网络具有大带宽、低时延和安全性高的特点，满足了对无线网络速度要求高的院内业务场景需求[15]。同时为各类移动医疗业务提供包括接入安全、网络管理安全、业务安全和数据安全等全方位的数据安全防护。

4 结论

5G 双域专网的建设与应用，是突破传统 WIFI 网络局限性的技术革新，具有良好的安全性和高速稳定的数据传输性能，可以很好地满足医院复杂的无线应用场景需求，具有广泛的应用和推广前景。

PU-184

基于物联网技术的智慧病房设计与实现

郝晓伟 卫新乐 赵治远 许锋波
河南科技大学第一附属医院

1 引言

随着新一代 5G 技术和物联网技术的快速发展，以智慧医院为核心的建设目标成为了各医院深化医疗改革的重要内容[1]。智慧医院的建设无疑是引领医疗机构从传统信息化建设转向数字化、智慧化转型的核心基础，其中智慧病房建设则是体现医院医疗服务水平和质量的核心环节[2-3]。如 2012 年以来，国家卫健委相继发布一系列智慧医院建设的纲要政策，作为智慧医院建设的三个主要板块（智慧医疗、智慧服务、智慧管理）之一的智慧医疗起了决定性作用，可见智慧病房的规划建设正在越来越凸显其重要性。然而，当前医院床护比仍然紧张的情况下，医护人员超负荷工作量与患者日益提高的住院体验要求之间的矛盾正在日益加剧，医院距离“以患者为中心的目标”还有一定距离[4-5]，如图 1 医院现阶段主要面临的问题所示。因此，如何通过更加有效的科技手段、管

理手段，构建智慧病房把“把医护人员还给病人”，从而显著增加医护人员的工作效率，减少低效流程的精力消耗，成为了各医院关心的热点。

图 1 医院信息化发展需求调查

在此背景下，我院在建设智慧病房的过程中，可遵循“打通最后十米”的总体原则。其一，尽可能实时性、远程性将临床数据应用于 5G 高速宽带数据传输技术、物联网传感采集技术、人工智能技术、大数据算法分析策略等技术手段；其二，融合智慧床旁护理、智慧床旁交互、实时生命体征监测与辅助决策等具有临床实用价值的应用开展；其三，实现多种业务的联动和闭环，在现有的标准和规范基础上实现标准化创新和流程再造。

2 我院病房存在的痛点

2.1 超高速宽带数据应用需求

随着我院病房护理能力的不断提升，各类先进医疗及护理设备越来越构筑于网络系统之上，依托于网络系统传输文字、语音、图像、视频等医疗数据，动辄要在秒级回传数百兆的图像文件到医生工作站，对于病房网络的传输速度要求越来越高。

2.2 床旁输液的精细化护理需求

根据世界卫生组织最新的公开调查显示，临床输液导致的错误包括药品种类错误 26.4%，药品时间错误 18.1%，药品剂量错误 36.1%。这其中，药品剂量错误更是 36.1%，可见人为的疏忽大意和过高的工作负荷是主要问题，精细化的临床输液管理已成为一个明显短板。

2.3 床旁信息的整合与交互需求

在面向医院的医技、就诊、运营管理等信息化领域，我院都得到了充分发展，衍生并发展出一系列应用系统。但在病房区域尚未真正转变为以以患者为中心、涵盖患者信息、护理信息、查房信息、缴费信息、医嘱与病历查询信息、餐饮、服务信息等综合一体的床旁智能交互式应用。

2.4 临床生命体征监测与辅助决策需求

在传统病区，医护人员每天需要对患者的体征进行测量，该操作的频繁程度对医护人员和患者都有可能造成影响。因此，需要通过先进的物联网传感器技术，通过远程采集患者各类体征，可在非接触的情况下远程采集患者的各项基础生命体征数据，包括心率、呼吸率、体动、体温、血氧等指标，再经信息平台对数据进行筛选和整合，从而形成医疗大数据，为临床诊疗和护理提供智能辅助决策，不仅能更好地保障医疗安全，还能为患者提供有品质的医疗服务。

3 智慧病房系统架构

3.1 系统逻辑架构

我院智慧病房系统架构设计在确保智能化、网络化、规范化的临床管理理念基础上，遵循了超高传输、超低延迟、低功耗长距离传感及临床统计算法融合的目标，采用 5G 网络、长距离低功耗物联网传输技术（LoRaWAN 网络）、人工智能算法分析模型等技术，构建了网络汇聚层、平台数据层、物联网应用层、可视化展现层等层次化智能病房总体架构。整个智慧病房形成全覆盖性物联网，充分利用 5G 技术混合型的网络架构、低功耗长距离传输技术的零布线模式，形成更高安全等级的医疗业务专网模式，并显著提高医院信息化平台数据超高速传输速度、容载能力、低延时能力。

如图 2 所示，该架构图由下及上，稳步部署，调动医院内部业务系统包括信息化部门、临床业务部门、设备与物资部门、后勤与保障部门等在内的高效率联动机制，定义清晰的分工与流程边界，兼顾技术可行性、临床落地可行性、医学法规的依从性、结论数据的可量化性、分析结论的对照性、工作流程改造的决策性等多个角度。

图 2 智慧病房逻辑架构图

4 智慧病房的应用

根据智慧病房系统规划与顶层设计的指导思想，将人工智能、5G 通信技术、物联网、大数据融合等技术进行临床应用融合，集中确定在三个重点发展领域，分别是智慧床旁护理应用、智慧床旁交互应用、临床生命体征监测及辅助决策应用，该三项领域分别对应了目前在卫生健康医护领域非常突出、亟待解决的典型应用难点和挑战性场景。

4.1 智慧床旁护理应用

作为智慧床旁护理场景中的智能输液管理，我院开创性地借鉴了 ICU 病房的输液泵技术思路，借助创新性的低功耗传感技术，以低成本、高普及率在更大量的、普通非 ICU 病房中开展，以“打通最后十米”的方式，填补了临床信息化的空白。

我院的智能输液管理系统，实现从医嘱下达、配液到执行的全流程闭环输液管理，有效保障输液安全。如图 3 所示。

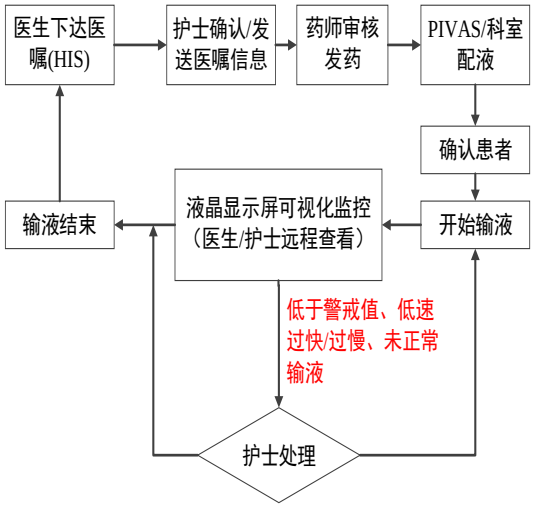


图 3 智能输液管理工作流程

通过 LoRa 长距离低功耗物联网传输技术的优势，一层病区仅需通过一台通信基站即可完成覆盖，实现多设备稳定连接，且能最大程度节省成本。智能输液设备采用高精度称重式传感技术，设备精度小于 5‰（通常为±1g），设备操作简单、便捷。具体应用中的工作界面如下图 4 所示。

图 4 智能输液管理工作界面

4.2 智慧床旁交互应用

智慧床旁交互应用的核心目标是构建以患者为中心的动态全集成系统，包括了患者信息、医嘱、护理、用药、诊断、缴费、病历、餐饮、调查、查询等为一体的综合式床旁智能化系统，患者既是信息的获取者，也是信息的产生者、参与者。

整个交互系统依托于医疗辅助服务信息系统为核心，通过与床旁主机、床头分机、卫生间分机、门口分机、走廊显示屏、医护主机、信息看板显示屏、台式电脑、值班室分机等硬件智能联动，组合打造出真正的智慧病房高效率的交互式管理。如图 5 所示。

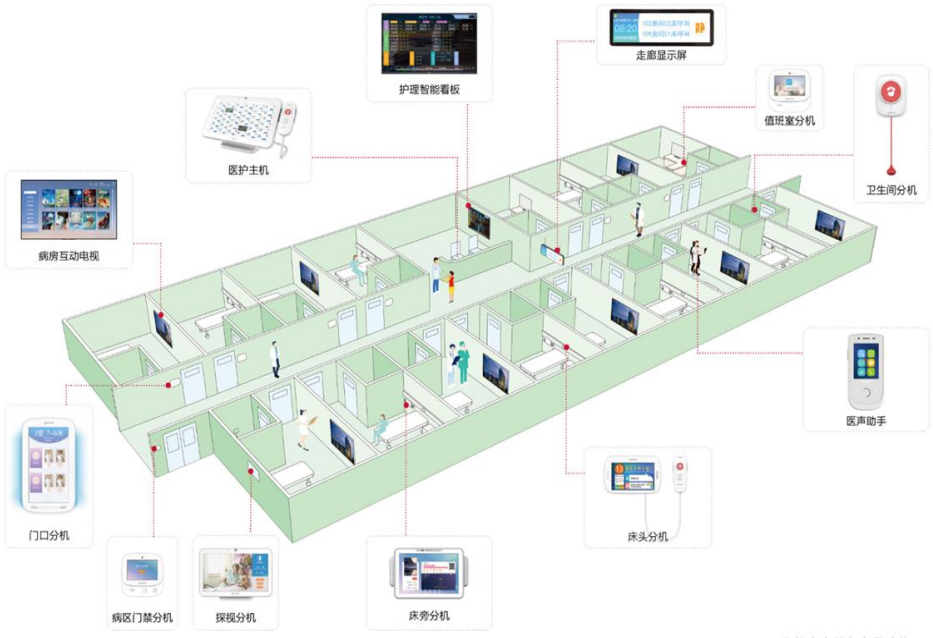


图 5 智慧床旁交互系统全景

4.3 临床生命体征监测及辅助决策应用

临床生命体征监测及辅助决策应用的发展是临床场景探索出一条全新的思路，能从 ICU 实时监测患者生命体征并应于临床决策的治疗与护理方式，全面移植或借鉴到普通非 ICU 病房，同时实现低成本、极简操作性、高普及率的目标。

围绕着这一目标而展开的体征监测，创新地将这些数据进行模型训练，与现行医嘱决策进行算法模型比对和融合，形成实时数据模型对现行临床护理及医嘱的重要辅助参考。

如下图 6 所示，采用柔性蓝牙体温贴采集临床体温数据，可连续 24 小时不间断

对患者体温进行实时监测，特别是针对某些病症或术后留观，从微弱的体温变化就可对患者恢复情况进行更细颗粒度的判断，并进而对临床医嘱提供更有价值的参考。

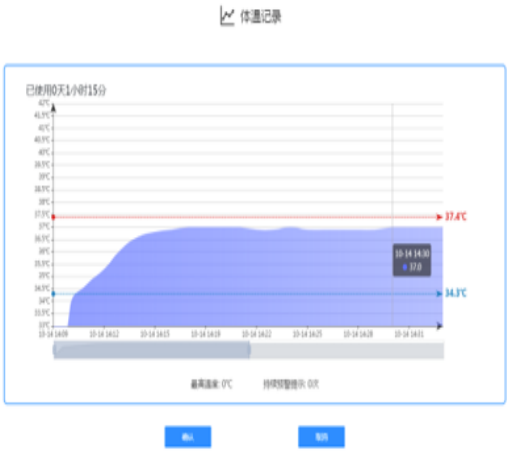


图 6 实时体温监测统计

如图 7 所示，通过采用智能床垫传感器有效跟踪监测患者心血管情况，采用非接触式技术，通过心冲击图（BCG）来进行 24 小时实时心率监测。

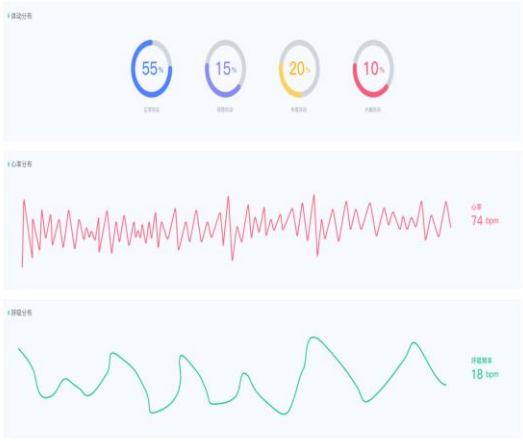


图 7 智能床垫产生的心冲击图与呼吸率实时统计

4.4 临床成效与分析

我院建成基于物联网技术的智慧病房后，在患者数量相近的情况下，分别对使用和未使用智慧病房系统的护士进出病房频率、病房响铃次数、患者满意度进行统计和分析，结果显示，护士进出频率由 66%降至 31%，医疗效率提升 35%；病房响铃次数由 25.64%下降至 2.4%，医疗质量提升 83%；患者及家属满意度从 90.6%提升至 96.5%，如表 1 所示。

表 1 智慧病房建设后医疗指标对比

医疗指标	建成前/%	建成后/%
护士进出病房频率	66%	31%
病房响铃次数	25.34%	1.4%
患者满意度	90.6%	96%

5 结束语

在基于物联网技术的智慧病房的临床实践中，产生了一系列全新的物联网数据和人工智能算法分析策略，由此产生的数据产生方法、挖掘方法、分析方法、整合方法、呈现方法等，均会带来一系列新的临床处置与作业思路。

这些全新的应用思路，从数据远程化、信息交互集中化的角度，创新地变革了现有的病房信息产生、传递、交互、呈现的模式，其带来的工作流程创新主要体现在从以医护为中心向以患者为中心的转变，从而也将改变临床效果的评价模式，更突显病人的临床体验，从更大的范围内契合了国家卫健委提出的关于智慧医院规划建设要从“供给侧”（即医院或医护人员）向“需求侧”（即病患人员）转变的指导方针，从院前、院中到院后的评价体系完全打通，以患者的综合体验、病症的全周期救治及护理效果评价为核心思想。

PU-185

面向儿科医学试题的参考解析推荐方法

[1]王娟，[1]侯丽，[1]孙月萍，[1]李佳明，[2]李云汉
[1]中国医学科学院医学信息研究所
[2]北京市测绘设计研究院

1 引言

问答（Question answering，QA）是自然语言处理（Natural Language Processing，NLP）中的一项基本任务，需要模型来回答特定问题。试题文本是一种常见的问答任务形式，用于考察学生对特定主题或领域知识的理解和应用能力。参考解析是回答试题的线索，能够帮助解读答案推理的整个过程，引导学习者思考问题、分析情景，促进深层次的学习和思考能力的培养，自动参考解析可以与学习者的答案进行对比，评估学生对医学知识的掌握程度和思维能力并能提高教学效率和一致性。基于以上现状，本文面向儿科医学试题，提出了一种基于隐性语义标引（latent semantic indexing，LSI）与 CoSENT（Cosine Sentence Bert）模型相结合的试题参考解析推荐方法。通过结合 LSI 和 CoSENT 模型，可以提高试题参考解析的准确性和推荐效果。LSI 可以帮助模型理解试题和参考解析之间的语义关系，而 CoSENT 模型可以衡量参考解析与试题的相似度。这样，模型可以根据相似度得分来推荐最相关的参考解析，提供更准确的答案线索。

2 相关研究

目前对于问题答案进行参考解析的推荐已经拥有不少成果：Trivedi 等人[错误!未找到引用源.]使用蕴含资源，包括标记的训练数据集（如 SNLI [错误!未找到引用源.]和 MultiNLI [错误!未找到引用源.]），来训练用于选择答案推理证明语句；Nishida 等人[错误!未找到引用源.]使用基于摘要模型的查询聚焦提取器，将可解释的多跳 QA 的证据提取视为以查询为中心的摘要任务，通过在问题句上使用具有注意机制的 RNN 顺序提取证据句，考虑证据句子之间的依赖关系并覆盖问句中的重要信息；对于没有提供黄金推理证明语句的数据集，使用结构化知识库（例如 ConceptNet、维基百科等）或远程监督以生成带标签语料，并从中获取与问题回答相关的推理证明语句，例如 Lin 等人[错误!未找到引用源.]使用 ConceptNet 构建模式图，并提出了一个结合卷积网络和 LSTM、具有基于路径的分层注意机制的关系图表示架构，以透明、可解释的方式有效地表示用于推理证明语句；Lv 等人[错误!未找到引用源.]从 ConceptNet 和维基百科纯文本中提取证据，构建图表以获得证据的关系结构，利用图形结构信息重新定义单词之间的距离，以学习更好的上下文单词表示，采用图卷积网络将邻居信息编码到节点的表示中，并使用图注意力机制聚合证据以探索答案的推理；Wang 等人[错误!未找到引用源.]应用远程监督来生成标签，然后用它们来训练推理证明语句提取器，应用深度概率逻辑学习框架对嘈杂的标签进行去噪，将句子级和跨句子语言指标结合起来进行间接监督；Zhao 等人[错误!未找到引用源.]使用维基百科中的文本片段构建证据图，然后应用训练好的基于 Transformer 的模型对这些证据片段进行推理，验证是否为证明语句。以上这些方法旨在提供模型训练所需的样本数据，以帮助模型学习选择 QA 理由句子的能力。Yadav [错误!未找到引用源.]和 Bethard 等人[错误!未找到引用源.]依靠基于余弦相似度算法的对齐信息检索方法检索给定当前查询的最明显的证明语句，并根据先前检索的理由动态地将新术语添加到查询中，以推动多跳检索。

3 面向儿科医学试题的参考解析检索方法

3.1 数据

3.1.1 任务制定

该任务包含三种数据成分：

试题题干：在本文中既可以为询问某类知识的句子，又可以为一段描述病人情况的段落。

答案选项：每道试题包含 4~5 个答案选项，其中包含一个或多个正确答案。

文档：来自合作单位提供的组织成段落的半结构化文本材料《儿科学》教材，从其中找出适用于试题题干和答案选项的参考解析语句。

本文的任务是根据已有试题与答案选项，从相关文档中检索得到每个答案选项匹配的参考解析，即每个答案选项的理由句，例如表 1 所示。在该题中，已经将题目按照《儿科学》的章节进行所属目录的标注，试题所属目录表示它属于《儿科学》中的“内分泌疾病”章节中“性早熟”这一节内容中的“诊断与鉴别诊断”小节，而由参考解析可以得出支持该题的正确答案选项的充分依据。

表 1 儿科医学试题参考解析推荐示例

试题所属目录	题干	正确答案选项	参考解析
内分泌疾病 性早熟 诊断和鉴别诊断	由于下丘脑对性激素的负反馈作用下降，使促性腺激素释放激素过早分泌所致的性早熟为?	A.特发性性早熟	第十五章 内分泌疾病 第四节 性早熟 （1）特发性性早熟（idiopathic precocious puberty）：又称体质性性早熟，是由于下丘脑对性激素的负反馈的敏感性下降、促性腺素释放激素过早增加分泌所致。

3.1.2 数据基本情况

试题主要用于医学生的各类考试，本文主要以儿科为切入点。试题可以评估医学生应用知识、概念和原则的能力，以及展示以患者为中心的基本技能的能力。

本文对试题进行了处理：删除重复问题、打乱顺序答案选项、清楚脏数据等。试题共分为 6 种类型：A1、A2、A3、A4、B1 和 X 型，其中 A1 为单句型最佳选择题；A2 型题为病例摘要型最佳选择题；A3 型题为病例组型最佳选择题；A4 型题为病例串型最佳选择题；B1 型题为标准配伍题。其中案例类型题目数量为 4131，非案例类试题数量为 6108。

用于参考解析查找的文档为电子版《儿科学》教科书，格式为半结构化的 XML。本文对文档按照章节进行分割，共计分为 135 小节。

3.1.3 数据分析

根据对目前已有的数据进行相关分析，发现数据有以下特点：

1. 专业知识：数据集中的每个问题都蕴含丰富的医学领域专业知识，而儿科学作为一门相对综合性学科，涉及到广泛而复杂的领域，这使得对于参考解析的检索需要对上下文有一定的理解。

2. 试题具有多样性：临床医学领域的问题类型如 3.1.2 中所述，通常可以分为两类。第一类是询问单一知识点的问题，例如“小儿用于预防佝偻病的维生素 D 剂量为？”；第二类问题首先描述患者的病情，然后询问最可能的诊断、最合适的治疗、需要的检查或者某种治疗的结果等。通常第一类问题需要进一步推理，第二类问题可能需要多跳推理，因此比第一类复杂得多，对文本的检索提出挑战。在本文中，主要对第一类问题进行处理，后续工作将研究对第二类问题的参考解析推荐。

3. 参考解析检索嘈杂：教科书中的段落通常无法直接提供试题的答案，并且广泛采用的信息检索方法所检索到的许多段落可能是嘈杂的并带有干扰的甚至是不相关的，因此需要对参考解析的检索范围做限定，在本文中用到了试题与教材对应的章节目录信息。

总之，应对这些挑战需要模型具备对医学领域的专业知识理解和推理能力，并能够从大量文本中筛选出相关的证据进行准确综合。

3.2 方法设计

为探索面向儿科医学在问答任务中的参考解析，本文提出了一种基于 LSI 与 CoSENT 模型相结合的试题参考解析推荐方法。首先通过基于隐性语义标引并结合电子书资源中的章节目录信息和试题所包含的章节目录信息进行参考解析的抽取，获取多个可供选择的参考解析，然后通过训练调优后的 CoSENT 模型对已有的参考解析与试题题干和答案选项进行相似度计算，得到相似度最高的参考解析即为最终的参考解析。方法流程如图 1 所示。

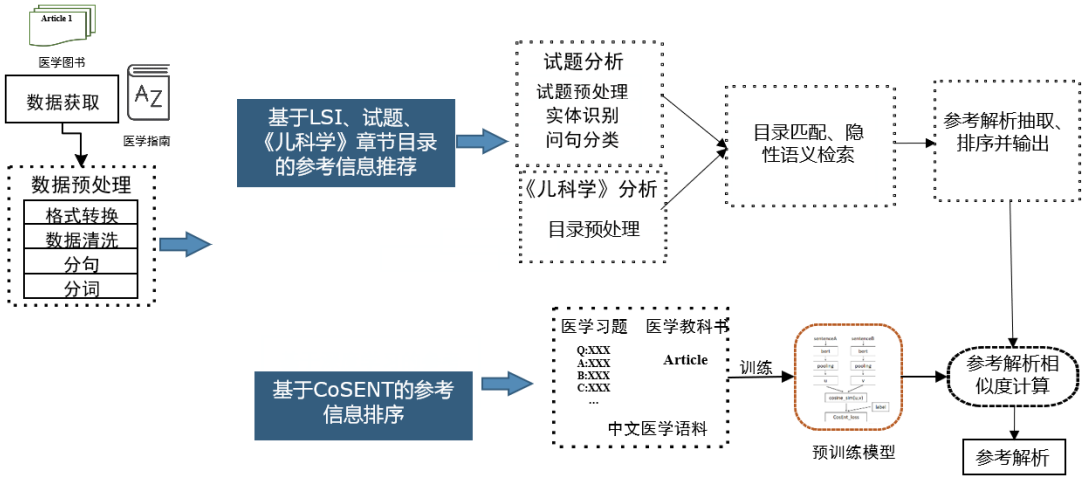


图 1 参考解析推荐方法流程

首先将试题题干与答案选项相拼接，形成一个 Query 查询，然后分别对试题题干、选项以及按照章节分割后的《儿科学》电子书进行分词、去停用词等；利用 TF-IDF 算法提取各篇文档关键词并构造关键词库；分别统计、分析试题题干与选项中的关键词；将该试题题干、选项分别与关键词库中其他包含相同关键词或部分相同关键词的文档进行比对，以确定不同文档间的语义相关性以及文档与特定关键词间的相关性，计算关键词在文档中出现的频率，取其中出现频率最高的前 3 个文档；利用余弦相似度计算文档中的每个句子与 Query 的相似度，然后排序输出匹配结果。

利用电子书资源中的章节目录信息和试题所包含的章节目录信息，将二者相结合，得到参考解析所在的精确章节位置，然后再利用相应规则提取具体的解析内容。最后将基于 LSI 的参考解析模型和基于章节目录信息得到的参考解析模型相结合，得出多个参考解析内容供下一步 CoSENT 模型进行选择。

得到多个参考解析内容后，采用 CoSENT 模型对参考解析内容与试题进行相似度计算。CoSENT 模型是一种有监督句向量模型，解决了 BERT 在语义相似度检索的巨大时间开销和其句子表征不适用于非监督任务如聚类，句子相似度计算等任务的问题。CoSENT 和 Sentence-BERT 在结构上相似，同样使用孪生网络结构。但和 Sentence-BERT 不同的是，CoSENT 的训练目标和推理任务的目标是相一致的，其本质是一种新的损失函数，统计了所有正样本对余弦值与负样本对余弦值的差值同样本标签差值的损失。本文使用儿科学的部分试题答案和参考解析语句构建相应的模型训练数据集，标注人选择拥有医学背景的学生四人，并选择第五人对标注数据进行一致性检验，最终形成数据训练集 8.2K、验证集 1.9K 和测试集 1.9K。预训练模型为 text2vec-base-

chinese [错误!未找到引用源。]，该模型是 chinese-macbert-base 在中文 STS-B 数据训练得到，并在中文 STS-B 测试集评估达到较好效果，最终该模型在本文的测试集上 person 相关系数评价表现为 66.7。然后利用该模型对基于 LSI 算法和章节目录信息获得的参考解析进行相似度计算并排序，选取相似度最高的作为最终的参考解析。

3.3 结果展示

以儿科学主题试题题库和儿科学主题电子书为核心资源，基于传统机器学习、深度学习等技术，开发为答案选项推荐参考信息并定位到参考信息出处信息的医学问答辅助应用系统，如图 2 所示，并开展参考信息评测工作，参考信息精确度为 72.6%。医学问答辅助应用系统为开展基于医学问答系统的慕课知识应用，便利学生掌握知识点及其关联，提升传统图书的知识服务能力等提供技术支撑，有助于提升传统图书的知识服务能力。

图 2 医学问答辅助应用系统

4 总结

本文利用《儿科学》电子书资源，面向儿科医学试题提出了一种参考解析自动推荐的方法，通过基于 LSI 和基于试题、电子书章节目录信息对参考解析进行自动提取，并通过训练后的 CoSENT 模型对所得到的多个参考解析语句进行排序，选取与试题题干与答案内容相似度最高的作为最终的参考解析。基于以上方法，本文还开发了医药问答辅助应用系统，为便利医学生掌握知识、教师教

学模式创新等提供技术便利，针对案例性试题的研究不足，后续将开展针对性研究，不断完善参考解析自动推荐方法，为后续覆盖全学科的自动参考解析研究打下基础。

PU-186

基于新一代 5G 技术的智慧医院建设研究与应用

许锋波 卫新乐 郝晓伟 赵治远
河南科技大学第一附属医院

1 引言

近年来，在新时代医疗体制改革的背景下，国家开始重点关注人民健康问题，然而，以“医院效率不高，患者就医效果不佳”的就医体验确尤为突出。因此，国家开始尝试新一轮的医疗体制深化改革，对医疗领域信息化建设、数据平台建设等工作都开始了积极推进。同时，随着当前互联网技术和通信技术的飞速发展，互联网技术与通信技术得到了融合性的提升，开始涌现了以 5G 技术为代表的业务和应用程序，形成了“5G+智慧医院”的模式，并取得了一定的成果。目前来说，在世界范围内很多国家的不同领域都在推动 5G 技术的应用，如 2013 年，日本、韩国及美国等多个国家开始研究 5G 技术；2016 年，美国某个运营商开始正式进行 5G 网络的试用；2017 年，美国部分区域实现 5G 技术的商业用途。

目前，国内的医疗信息化建设区域性分布式不均也是构建智慧医院的关键性问题，而 5G 技术的出现，使得数据传输效率和效果大大的提升。如在医疗领域利用高网速的视频通信能力，可以实现切片技术、移动边缘等技术，这些应用技术在医疗服务的实现打破了国内医疗机构信息化建设不均的现状，为国内各医疗机构的智慧医院建设提供了可行性，如使得“5G+智慧医院”真正进入实质性阶段。

基于此，本文分析当前国内医疗机构存在信息化建设分布不均的问题，导致医生和患者之间存在医疗资源利用效率不高及患者就医满度低的情况，并引入了以 5G 为核心的“5G+智慧医院”建设的解决方案，探讨以我院为主的“5G+智慧医院”具体实施方式，最后对我院 5G 技术实际应用过程出现的难点进行了分析和解决。

2 5G 技术的性能优势

2.1 5G 技术概述

所谓 5G 技术，就是第五代的移动通信技术，其传输速度快于 4G。5G 技术可以将其应用于增强型移动宽带、大连接物联网、低时延高可靠通信三种。

对于 5G 移动网络通信技术来说，互联网和物联网是研究和设计的重要方向，高清视频、虚拟现实等移动互联网技术的迅速发展，也催生了 5G 技术新的研究方向，从传统的人与人之间通信延伸到物与物、人与物之间的通信互联，并逐渐实现“万物互联”，使移动通信网络可以与更多的行业和领域进行深度融合，自动驾驶、智能家居、共享单车、智慧交通、工业机器人、环境监测等物联网服务和应用也呈现出爆发式增长的态势，进而推动移动通信网络以及新兴产业的繁荣发展。

2.2 5G 技术的性能优势

5G 属于最新一代通信技术，无论是网络信号传输速度，还是功耗、容量，都比 4G 更胜一筹，技术优势更多，如表 1 5G 与 4G 关键指标参数对比所示，5G 技术可以适应更多的医疗活动场景，对医院的技术管理的模式和服务产生革新。同时，5G 不只是网络性能方面的提升．其主要建设目标是解决“物物互联”和“人物互联”问题．将人与人之间的通信扩展到万物连接，提升移动网络所覆盖的深度和广度，打造全连接的智能化社会。

表 1 5G 与 4G 关键指标参数对比

技术指标	4G	5G	提升倍数(倍)
峰值速率(bps)	1G	20G	20
连接数密度（km2）	105	106	10
流量密度（Tbps/km2）	0.1	10	100
移动性(km/h)	350	500	1.43
端到端延迟(ms)	10	1	0.1
用户体验速率(bps)	10M	0.1～1G	10～100

3 5G 技术在我院信息化建设中的应用

3.1 远程会诊和诊断

5G 技术的发展无疑会为医院提高服务做出巨大贡献，并且其在稳定性、高质量以及实时性等方面都能够很大程度上超过 4G 技术，延伸大型医院的服务水平。

在此背景下，我院更是率先成立了“互联网医院”，实现 5G 网络全覆盖，建立远程医疗协作网，把市、县、乡、村医疗机构“网”起来，建立分级诊疗体系，进一步放大优势医疗资源效应，实现远程诊断、远程会诊、远程监护、远程教育。我院“互联网医院”试点先以栾川县为试点单位，在其县医院成立网络医院，在 15 个乡镇卫生院成立网络分中心、35 个村卫生室成立网络诊所，初步建设市、县、乡、村四级分级诊疗体系，实现医疗资源下沉，精准健康扶贫之目标。

3.2 智能导诊

我院的智能导诊系统目的是构建就诊病人的智能语音助手，患者可实时和智能导诊机器人就行语音交流，用于帮助患者提前预测疾病。同时根据预测的疾病智能为患者推荐相关科室和坐诊医生，相比于医院的人工全科医生，智能导诊大大为病人和医生节省了时间和提供了便利，提供了实时性的交流平台。例如某地的早期食管癌患者到我院就诊，患者偶尔出现进食困难、滞留不适感、咽喉部干燥不适感等症状、经智能导诊接收分析后，预测患者为早期食管癌，并向患者推荐相关科室和门诊医生，会极大的增加患者的生存、治愈几率。

3.3 患者体征监测

在患者体征测量方面应用 5G 技术能够实时对患者体征的监测、处理，从而实现检测等功能。5G 技术具有细小型化发展的特点，可以使得许多具有低消耗、小体积的设备能够更好地成型的佩戴在患者身上。对于 5G 高容量这一特点来说，其可以实现医疗互联网的最大程度接入，支持对患者实时数据在医疗设备上的上传。如我院智慧病房下的智慧腕带，基于 5G+LoraWan 实现信息传输和互享。

4 5G 技术应用过程中的难点与解决措施

4.1 技术难点

（1）云计算软硬件

一般情况下，大型服务器来支撑的 5G 网络，需要组建云计算平台的方式，利用基站、交换网络等设备与具有转换功能的路由器进行对接，实现系统智能化的相关服务与操作。然而，我院现阶段面临问题，如存储系统、中间件、云计算服务器等软硬件方面的技术难点，因此应该不断继续在

相关方面进行自主研究，提升开发力度，尽快摆脱一些关键技术层面制约，提升我院数字化和智慧化的发展速度。

（2）毫米波传输

基站本身的低、中、高频率的信号也会影响 5G 技术的应用场景，而 5G 技术更加依赖高频信号，这是因为高频信号的衰落速度较快，在具有崎岖地形或密集人口的大城市中，必须及时建立密集化组网，不断建设基站，满足城市的需求。由于我院的应用环境较为复杂、建设规模也比传统规模要大，因此在技术落实方面的难度更大，尤其是在信号传输方面有着更高的要求。此外，在系统设计方面，高频关键器件的特性与成熟程度也会影响整个系统的稳定性，尤其是在大宽带、低噪声、高效率、多功能性高频器件的设计、生产在产业化层面还存在着许多技术性难题需要时间来攻破。

（3）网络安全

5G 技术支撑平台，一般采用的是通用硬件平台，但是其却不具有较好的网络安全可靠性，为了能够更好地应用 5G 网络技术，提高其应用水平，在支持用户海量访问时，必须采用服务终端，并使得用户通过其安全认证。因此，如果需要对海量数据进行访问，系统可采取相应的防护措施，比如对服务器和用户方法进行加密等方法。

4.2 解决措施

（1）提高覆盖性

覆盖性增强技术采用密集型异构组成网络。通过缩小覆盖半径，以增强频谱资源空间的复用效率，从而提升 5G 网络覆盖度。对于 5G 超密集异构网络，我们采用宏站和低功率小型基站来实现全方位覆盖，通过提高站点密度来减少节点之间的距离，其网络节点距离终端设备更近，实现频谱效率和系统容量大幅度的提升。

（2）提高传输性

毫米波通信的方法能够提升技术的稳定性，对频谱资源紧缺程度有所缓解，同时提升系统的通信容量。对于毫米波通信方法来说，其具有波束集中、波束窄以及较强的抗干扰能力等优点，稳定网络传输过程的稳定性，使得系统具有更高的抗干扰能力。

（3）提升安全性

物联网、移动通信等领域的新型技术中的可见光通信，具备很多应用优势，比如较低的成本、宽频谱、较高的速率高等，此外，5G 技术的保密性特点也直接奠定了其在 5G 技术通信中的地位。

5 结论

当前，对于我院“5G 智慧医院建设”来说，依托于 5G 技术的强连接性和高质量性，实现了面向医务人员的智慧医疗及面向患者的智慧服务，为医务人员和患者之间的搭建更为便捷信息服务平台。未来我院考虑引入人工智能、边缘计算等一些前沿的技术，用于提升和改进信息化水平，使信息业务平台功能更加强大和完善。

PU-187

基因预测多发性硬化症的发生会增加患严重 COVID-19 的风险

胡景喜 王茜 郝佳炜 陈淼然 赵宇慧 贺培凤*
山西医科大学

介绍

多发性硬化症(MS)不仅是一种中枢神经系统自身免疫性疾病，也是最常见的慢性炎症性疾病[1]。它通常以行走障碍和认知过程减慢为特征，是导致残疾的主要非创伤性原因。全基因组关联研究(GWAS)鉴定出的大多数多发性硬化症风险等位基因都与免疫通路基因相关。

当前的 2019 冠状病毒病（COVID-19）疫情因严重急性呼吸综合征冠状病毒 2（SARS-CoV2）感染而对全球健康构成严重威胁[2]。截至 2023 年 9 月 6 日，累计 COVID-19 病例数和死亡人数分别超过 7 亿和 600 万（https://covid19.who.int/）。COVID-19 的易感性和严重程度具有高度异质性。呼吸衰竭和多器官功能障碍，以及免疫反应受损和不受控制的炎症过程，是导致感染的主要原因。COVID-19 患者死亡。严重的 COVID-19 与多种合并症相关，包括高龄、肥胖和糖尿病。了解 COVID-19 与其他健康状况之间的关系是治疗和管理 COVID-19 的关键。识别风险因素，尤其是可改变的风险因素，可以帮助制定疾病预防策略。呼吸衰竭和多器官功能障碍，以及免疫反应受损和炎症过程失控，是导致 COVID-19 患者死亡的主要原因。了解 COVID-19 与其他健康状况之间的关系是治疗和管理 COVID-19 的关键。识别风险因素，尤其是可改变的风险因素，可以帮助制定疾病预防策略。

自身免疫性风湿性疾病与 COVID-19 的易感性和严重程度之间的关联已得到广泛讨论。2020 年春季，瑞典国家卫生和福利委员会针对 COVID-19 的风险群体进行了讨论，并指出患有与多发性硬化症相关的严重运动障碍的人属于 COVID-19 的风险群体。接受抗 CD20 治疗的多发性硬化症患者患严重 COVID-19 的风险增加[3]。最近的病例对照研究进一步发现，多发性硬化症患者因 2019 年冠状病毒(COVID-19)大流行而出现并发症的风险可能增加。然而，一些系统分析表明，多发性硬化症并不会显着增加感染 COVID-19 的风险，因为多发性硬化症患者更有可能是女性和年轻人。目前尚不清楚 MS 与 COVID-19 发展之间的正相关性是否是由于混杂因素或偏见造成的。此外，由于可能的混杂因素和反向因果关系，观察性研究常常存在偏差。因此，迫切需要使用更基础的证据来确定它们的关联并阐明 COVID-19 和 MS 的潜在机制。

孟德尔随机化(MR)是一种流行病学方法，它使用单核苷酸多态性(SNP)作为工具变量(IV)来推断暴露与结果之间的因果关系[4]。根据孟德尔定律，MR 是随机对照试验的自然类似物，因为遗传变异是随机的分配于减数分裂。因此，MR 研究不太可能受到可能的混杂或反向因果关系的影响。

使用根据 MS GWAS 摘要统计估计的 MS 遗传决定因素进行了两个样本的 MR 分析，以评估遗传介导的 MS 风险与 COVID-19 严重程度之间是否存在因果关系。

材料与方法

研究设计

我们进行了一项双样本 MR 研究，以检验 MS 与 COVID-19 的因果关系。该分析模拟了后代 SNP 随机分配的随机对照试验（独立于性别和年龄等混杂因素）。此外，该 MR 设计满足三个假设：（i）遗传变异作为 IV 应与 MS 显着相关。(ii)IV 不应与任何混杂因素相关联。(iii)IV 应仅通过 MS 影响 COVID-19 的风险[1]（图 1）。本研究的数据来自先前经道德和机构审查委员会批准的公共数据库，因此不需要进一步的道德批准。

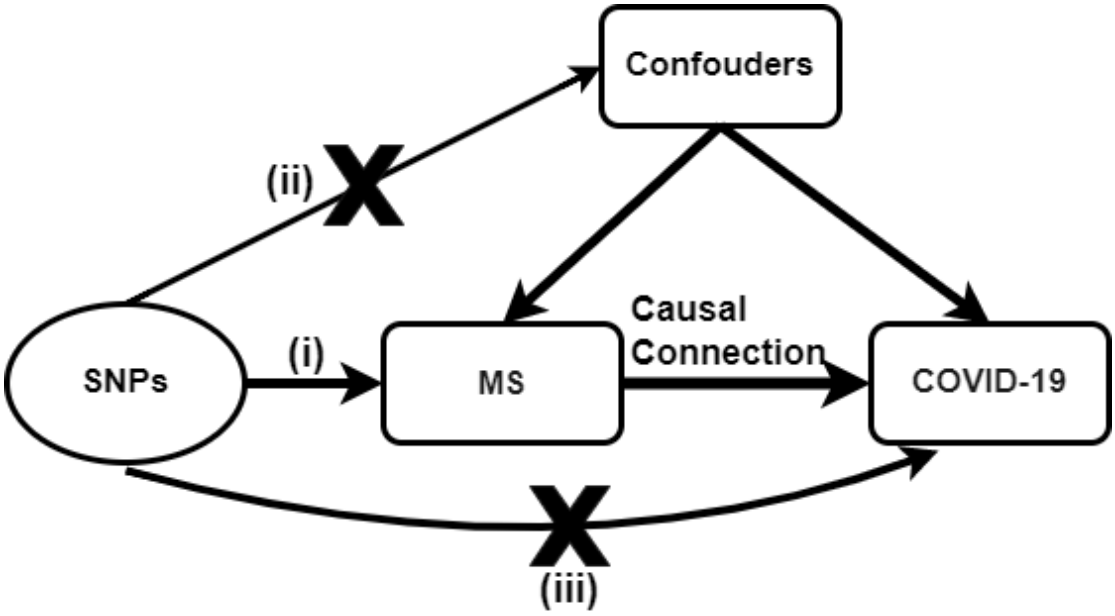


图 1：有向无环图(DAG)，代表 MS 和 COVID19 风险之间的因果关联。箭头代表直接因果效应。SNP，单核苷酸多态性；MS：多发性硬化症。

数据源

MS GWAS 数据由国际 MS 遗传学联盟(IMSGC)基于 IMSGC 免疫芯片获得，IMSGC 是一项大型 MS 遗传学研究，其中包括 14498 例 MS 病例和 24091 名健康对照，所有参与者都有欧洲血统[5]。COVID-19 表型的汇总统计数据来自最新版本的 COVID-19 宿主遗传学倡议(HGI) GWAS 荟萃分析（2022 年 4 月第 7 轮），这项名为 A2 研究的研究将诊断为“非常严重的呼吸道疾病”COVID-19 的患者(N=18,152)与普通人群的样本(N=1145,546)进行了比较[6]。

IV 的选择

在 MR 分析之前，执行严格的过滤步骤来控制 SNP 质量。首先，我们提取了与 MS 全基因组显着性阈值相关的 SNP（P<5×10-8）。其次，我们在连锁不平衡中聚合了 SNPS（LD、r2<0.001 且距离>1Mb））。第三，为了避免原始 GWSA 可能存在的统计偏差，我们排除了次要等位基因频率(MAF)<0.01 的 SNP，因为它们通常具有较低的置信水平。第四，F 统计量小于 10 的 SNP 被排除在量化遗传工具强度之外。我们在 phenoscanner 平台上进行测试，排除了任何与 COVID-19 的潜在危险因素相关 SNP。

孟德尔随机化分析

为了评估 MS 对 COVID-19 的因果影响，通过逆方差加权(IVW)、MR-Egger 回归和加权中位数(WM)进行 MR 分析[7]。在这些方法中，IVW 是估计暴露与结果之间因果关系的最重要方法。它计算了每个 SNP 对 COVID19 结果的 Wald 比率，并获得了汇总因果估计。所有分析均在 RV4.0.5 中使用 TwoSampleMR 包（版本 0.5.6）和 MRPRESSO 包（版本 1.0）进行。我们将估计效果值表示为优势比和相应的 95%置信区间(CI)表示，可以记录为 MS 风险每增加 1-SD 所导致的相对 COVID-19 风险的估计值。

敏感性分析

为了评估这些结果的稳健性并防止潜在的多效性和异质性，进行了一系列敏感性分析，包括 Cochran's Q 检验、漏斗发检验、留一法(LOO)分析、MR-Egger 截距检验和 MR 多效性残差和进行了异常值（MR-PRESSO）测试[8]。Cochran Q 检验中 P 值<0.05 被认为具有统计显着性。MR-egger 用于评估潜在的水平多效性。MR-Egger 回归的截距项用于评估水平多效性是否影响 MR 分析的结果，其中 P<0.05 表示水平多效性。通过漏斗图的目视检查，不对称性表明水平多效性。我们进行了留一法分析，以确定是否有任何单个 SNP 驱动了因果估计。我们通过依次丢弃每个与暴露相关的 SNP 来重复 IVW 分析。在 MR-PRESSO 测试中，消除了与异质性相关的 SNP，以减少估计因果效应时的异常值。

结果

SNP 的选择

总共提取了 49 个 SNP 作为 MS 的显着预测 SNP（P<5×10-8，LDR2<0.001）。计算 IV 和表型关系强度后，在 PhenoScanner 中验证统计 F 大于 10 的 SNP（p=1e-5，R2=0.8，build=37）：SNPrs8070345 与酒精摄入相关，需要删除。其余 48 条静脉显示出与主流 COVID-19 危险因素（吸烟、饮酒、肥胖）无关的疾病和特征。总共删除了 COVID-19GWAS 中不存在的 6 个 SNP（rs1131265、rs12296430、rs1359062、rs212405、rs9736016、rs9989735）。最终，42 个 SNP 被鉴定为与暴露相关的独立 SNP（图 2）。

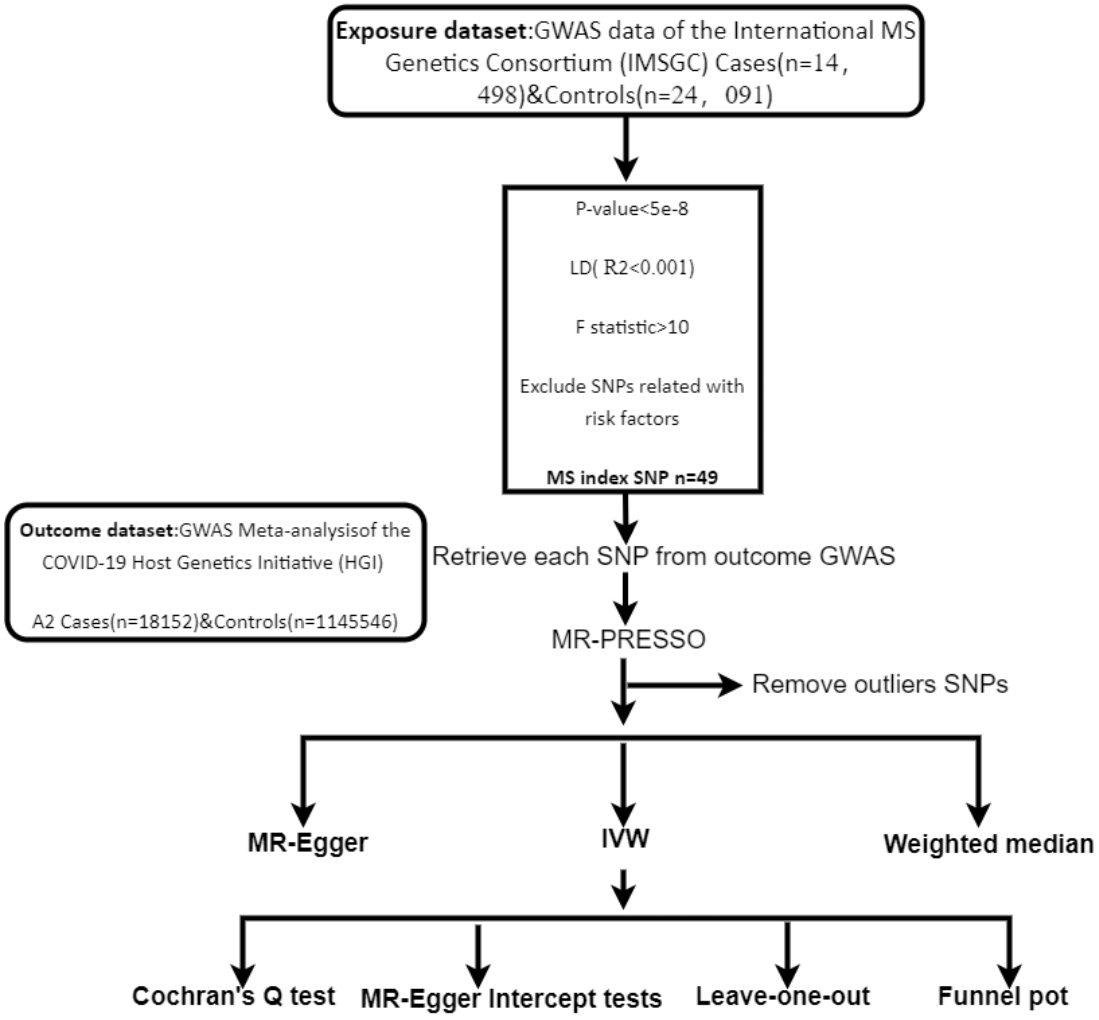


图 2：研究中的流程图和数据选择过程。对暴露数据进行一系列质量控制，并对暴露数据和结果数据进行协调，筛选出独立的 SNP，用于后续的 MR 分析和敏感性分析。

MS 与 COVID-19 之间的因果关系

在两个样本的 MR 分析中，MS 与诊断为严重呼吸道疾病的 COVID-19 风险之间存在显着异质性(Q=93.701，P=5.282E-06)。使用 MR-PRESSO 将一个 SNP(RS3129727)识别为异常值，然后重新计算 MR 估计值。各种 SNP 信息显示在补充表中。初步证实 MS 与 COVID-19 诊断为严重呼吸道疾病的风险之间存在因果关系（OR：1.063，95％CI：1.020-1.110，P=0.004），并使用加权中位数 MR-Egger 来解释阳性辅助结果（MR-Egger：OR：1.099，95％CI：1.009-1.198，P=0.038；加权中位数：OR：1.090，95％CI：1.037-1.146，P=0.0007）（（图 3）。尽管仍然存在异质性，比以前少得多(Q=73.227,P=0.001)。未发现水平多效性(Egger 回归截距=-0.006,P=0.394)。SNP 对暴露(MS)和结果的估计效应大小（确认患有非常严重的呼吸道疾病的 COVID-19 感染）以散点图显示（图 4）。在缺失的分析中，没有单个 SNP 强烈影响 MS 与 COVID-19 之间的因果关系。（图 5）。漏斗图的对称性表明，我们的研究严格遵循 IV 假设。（补充图）

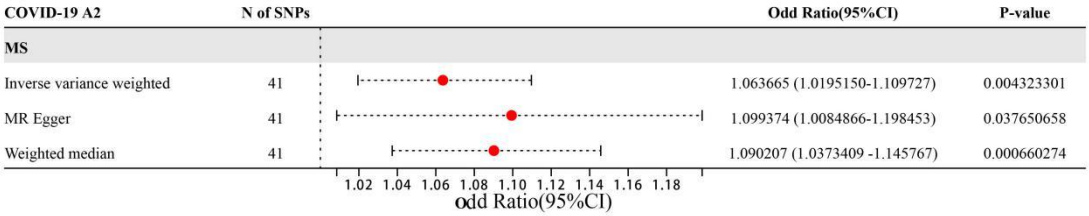


图 3：通过 IVW、MREgger 和加权中位数的比值比(OR)值评估 MS 与 COVID-19 之间的因果关系。

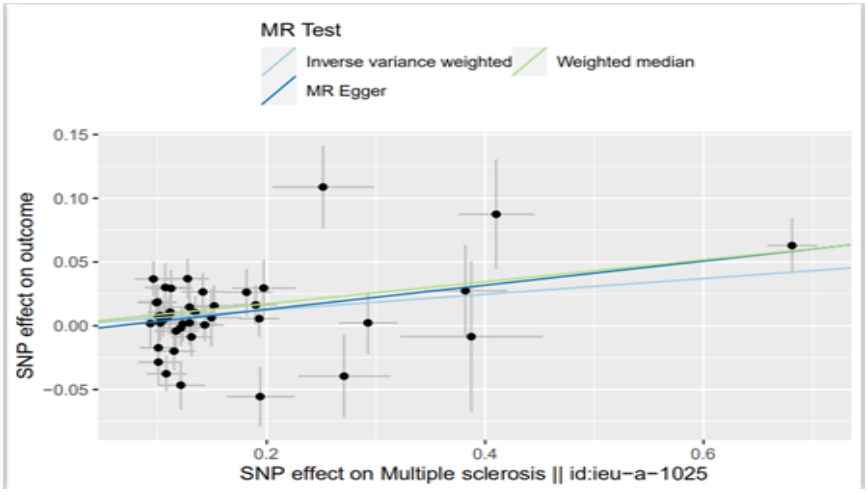


图 4：多效性检验散点图。每个点代表一个 IV，直线代表 95%置信区间，横坐标是暴露，纵坐标是结果，彩色线代表 MR 拟合的效果。

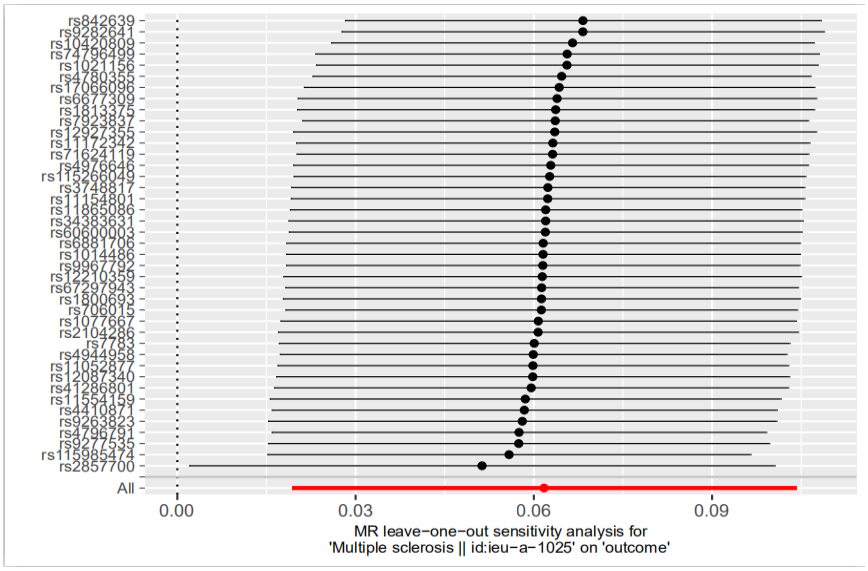
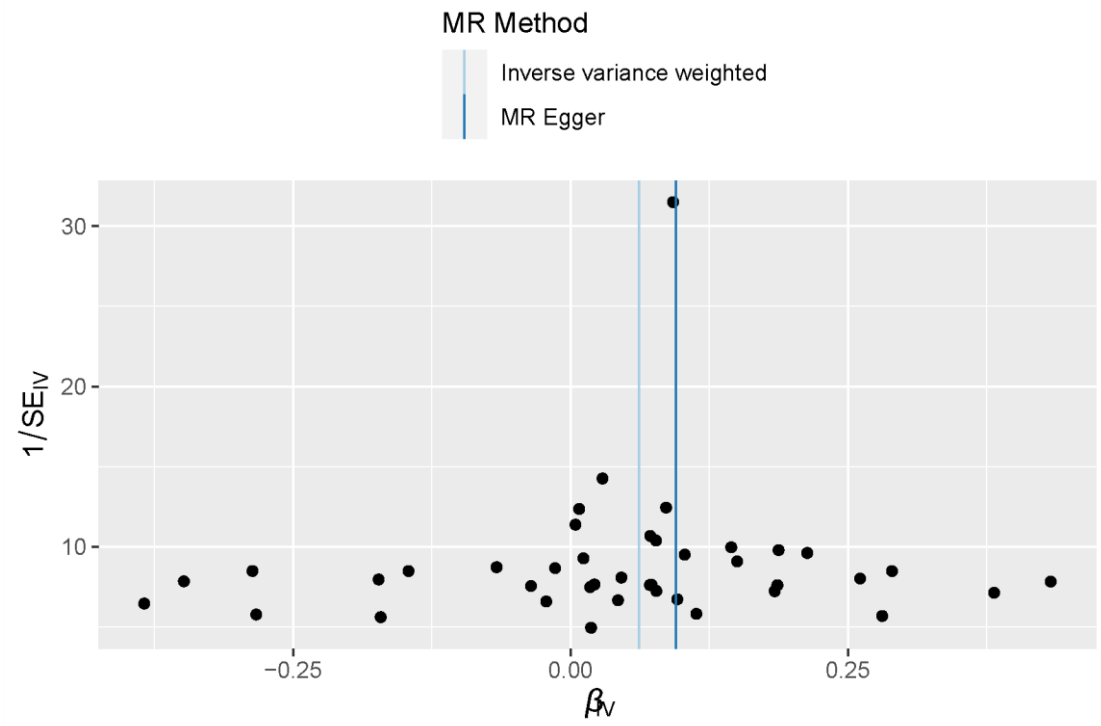


图 5：留一法在 MR 敏感性分析中的结果



补充图：漏斗图显示了 MS 对 COVID-19 风险影响的孟德尔随机化(MR)分析的异质性，表明不存在显着的多效性和实验偏差。

讨论

利用国际 MS 遗传学数据和 COVID-19 GWAS 数据，我们使用多种 MR 方法评估了 MS 和 COVID-19 与已确诊的严重呼吸道疾病之间的因果关系。总体而言，我们的结果表明，多发性硬化症的发病会导致 COVID-19 的严重程度增加。它对于探索哪些疾病可能诱发患者感染 COVID-19 是一个很好的指南。

作为一种自身免疫性疾病，多发性硬化症（MS）患者的免疫反应调节异常，导致某些细菌或病毒的感染频率高于普通人群。因此，推断自身免疫性疾病之间的因果关系是合理的例如多发性硬化症和 COVID-19。最近的研究表明，接受抗 CD20 治疗的多发性硬化症患者患严重 COVID-19 疾病的风险增加。然而，这些回顾性和前瞻性观察性研究缺乏证明 MS 与 COVID-19 之间因果关系的直接证据。与这些需要长期观察的大规模前瞻性临床试验相比，MR 研究以时间和成本效益的方式揭示了 MS 和新冠病毒之间的因果关系。

这项研究既有优点也有局限性。本研究的三个优势较为突出。首先，与传统的观察性流行病学研究设计相比，我们的 MR 分析可以最大限度地减少反向因果关系和混杂因素的影响。第二个是使用 COVID-19 患者的全面 MS 和最新 GWAS 数据集来系统地评估其因果关系。第三，我们选择的数据集都是来自大规模遗传联盟的汇总水平数据集，这些数据集具有强大的统计能力来检测 MS 和 COVID-19 之间的因果关系。

然而，这项研究存在一些局限性：首先，我们的暴露和结果的参与者是欧洲血统，因此研究结果不能立即推广到具有其他文化背景和生活方式的种族群体。其次，由于 MR 通过遵循等位基因随机分配的原则来推断暴露与结果之间的因果关系，因此通过 MR 方法完全区分多效性和中介作用具有挑战性。最后，虽然 MR-PRESSO 去除异常 SNP（RS3129727）后异质性显着降低，但异质性仍然存在。这是由于人群中的测序或抽样方法存在差异，而对于这些人群，可以获得有关暴露和结果的汇总水平统计数据。

总结

总之，我们提供了强有力的证据支持 MS 的发生对 COVID-19 风险的因果影响。我们的这项研究将促进未来对 COVID-19 相关免疫系统影响的研究，从而改进在当前大流行中监测多发性硬化症患者的风险评估工具。然而，未来的研究需要探索因果关系的潜在机制，这对于制定相关的临床建议至关重要。

PU-188

基于人工智能的 PI-RADS 自动评分在前列腺癌 MRI 诊断中作用与展望

康亚西[1] 王忠尧[1] 刘阳[2]
1.海南省人民医院 2.海口市人民医院

前言：前列腺癌（Prostate Cancer, PCa）是男性最常见的恶性肿瘤之一，其发病率在全球范围内呈逐年上升的趋势。根据统计数据显示，全球每年有超过 100 万人被诊断出患有前列腺癌，其中约有 30 万人死于该疾病。世界范围内，前列腺癌发病率在男性所有恶性肿瘤中位居第二。根据我国国家癌症中心统计，前列腺癌自 2008 年起在男性泌尿系统肿瘤中发病率第一。前列腺癌不仅对男性的身体健康造成威胁，而且对社会经济文化也产生了重要影响。如何早期诊断前列腺癌，成为一个亟待解决的问题。

1. 前列腺癌的早期诊断

目前常用的方法包括前列腺特异性抗原（PSA）检测、前列腺穿刺活检和前列腺 MRI 等，以下是它们各自的优缺点。

PSA 检测：PSA 是一种由前列腺细胞分泌的蛋白质，可以用来判断前列腺是否发生异常。PSA 检测简单、快速、无痛，但其准确率有限，容易出现假阳性或假阴性结果。

前列腺穿刺活检：是通过穿刺前列腺腺体，进行的活组织病理检查（活检）可以确定前列腺癌的确诊。但穿刺活检阳性率低（30%左右），且是一种侵入性操作，容易引起感染、出血等并发症，并且对患者的生活质量造成不良影响。

前列腺 MRI：前列腺 MRI 是一种通过磁共振成像技术对前列腺进行成像的方法，它可以提供更准确的前列腺癌定位和评估。多参数磁共振成像 (MP-MRI)及其 PI-RADS 评分，是当前最通用的无创检查，可以大幅降低不必要的前列腺穿刺活检，可显著提高 Pa 的早期诊断率，受到医生、病人和社会的高度关注。

PI-RADS 是 Prostate Imaging Reporting and Data System 的缩写，是一种基于前列腺 MRI 图像并用于前列腺癌筛查和诊断的标准化评分系统，可以评估前列腺癌的概率和病变的严重程度。该系统由欧洲泌尿放射学会(Europen Society of Urogenital Radiology, ESUR)制定[1]，2016 年推出 V1.0 版，2019 年更新为 V2.0 版，2021 年更新为 V2.1 版，通过评分使前列腺影像报告标准化、规范化，减少模糊的影像描述和诊断结果，旨在促进全球标准化，减少前列腺 mpMRI 检查的获取、解释和报告方面的差异，提高前列腺癌的早期诊断准确性和可靠性。

PI-RADS 评分是根据前列腺 T2WI、DWI、ADC 及 DCE 等的综合表现，对出现有临床意义前列腺癌的可能性给出评分方法。前列腺解剖上主要分为外周带和移行带、中央带，由于在 MRI 影像上中央带和移行带很难区分，合称中央腺体。在 PI-RADS 评分中，前列腺外周带（PZ）疾病以 DWI 结果为主，以 T2WI 来修正。前列腺中央腺体以 T2WI 结果为主，由 DWI 评分进行修正。

前列腺 PZ 评分标准	
评 分（分）	标准
1	呈均匀高信号；
2	线状、楔形或弥漫性轻度低信号，边界不清；
3	信号强度不均匀或界限不清，呈圆形、中等低信号，包括其他不符合 2、4 或 5 分标准者；
4	局限于前列腺内，边界清楚，均匀中等低信号病灶或肿块，最大径 <1.5cm；
5	与 4 分影像表现相同，但具有侵犯或突破前列腺包膜（局部隆起或宽基底相接处≥1.5cm）；

表 1 前列腺外周带 T2WI 评分标准

前列腺 TZ 评分标准	
评 分（分）	标准
1	均匀中等信号强度，边缘清晰（正常）；
2	局限性低信号或不均匀有包膜的结节，边缘仍清晰（前列腺增生）；
3	边缘模糊，信号强度不均匀，包括其他不符合 2、4 或 5 分标准者；
4	呈透镜状或边界不清，均匀中度低信号，最大径<1.5 4cm；
5	影像表现同 4 分，但同时累及前纤维基质或外周带前角，最大径≥ 1.5cm；

表 2 前列腺中央腺体 T2WI 评分标准

DWI 应用评分标准	
评 分（分）	标准
1	与正常组织 ADC 值未下降，DWI 上信号增高（b=800）；
2	ADC 值降低，呈模糊低信号，DWI 上弥漫性信号增高；
3	在 ADC 图上呈局灶轻、中度低信号，在高 b 值图像上呈等、轻度高信号，包括其他不符合 2、4 或 5 分标准者；
4	在 ADC 图上呈局灶明显低信号，在高 b 值图像上呈明显高信号，轴面最大径<1.5cm；
5	影像表现同 4 分，但最大径≥1.5cm。

表 3 前列腺 DWI+ADC 评分标准

PI-RADS 评分系统最终将前列腺 MRI 图像分为 5 个等级，即 PI-RADS 1 到 5 级。具体评分标准如下：

PI-RADS 1 分	前列腺没有可疑病变。
PI-RADS 2 分	前列腺有一些非癌性病变，不需要进一步检查。
PI-RADS 3 分	前列腺有一些非癌性病变，建议进行随访检查。
PI-RADS 4 分	前列腺有可疑的癌症病变，建议进行穿刺活检以确诊。
PI-RADS 5 分	前列腺有高度可疑的癌症病变，建议进行穿刺活检以确诊。

表 4 PI-RADS 评分标准

2. PI-RADS 评分系统的意义

通过统一的评分标准和报告形式，可以提高前列腺癌的早期诊断准确性和可靠性，降低漏诊率和误诊率。此外，该评分系统还可以减少不必要的穿刺活检和影像学检查，降低医疗成本，提高患者的生活质量。因此，PI-RADS 评分系统已经被广泛应用于临床前列腺癌的筛查和诊断。

但目前 PI-RADS 评分，仍是人工进行，存在以下缺点：

（1）主观性和人为失误的可能，不同医生之间的评分可能存在差异。

（2）需要医生进行大量的手动操作，耗费时间和精力，并且可能存在疲劳和注意力不集中等人为因素，影响评分的准确性。

（3）结果可能会受到医生的个人经验和知识背景的影响，如果医生的经验和知识不足，评分的准确性可能会受到影响。

人工智能（AI）可能有助于完成良性和恶性实体的癌症检测和分类，它可能有助于在磁共振成像（MRI）扫描上分割可疑病灶和正常解剖。AI 还可以帮助对前列腺多参数 MRI（mpMRI）病例进行初步评估或分类（即选择/识别具有更非典型图像特征的前列腺 MRI 检查）和图像质量（即将 mpMRI 扫描分为诊断和非诊断）[2-4]。

3. 用人工智能（AI）代替人工

用人工智能（AI）代替人工，进行 PI-RADS 评分的科研价值突显，主要包括：

（1）提高评分准确性。计算机可以准确识别和分割前列腺中的异常信号，从而减少评分误差，提高评分准确性。

（2）提高评分速度。计算机评分速度快，能够在短时间内完成大量的评分工作，大大节省了人工评分的时间和成本。

（3）促进病情诊断。通过计算机自动评分，医生可以更快地了解病人的病情，从而更准确地

制定治疗方案。

（4）提升医疗水平。计算机评分可以帮助医生准确地确定病人的病情，从而更好地指导治疗和手术决策，提升医疗水平。

（5）推动医疗科技发展。计算机评分的应用促进了医疗科技的发展，将人工智能技术应用于医学领域，为医疗行业带来更多的创新和进步。

因此，使用人工智能（Artificial Intelligence, AI）帮助医生准确评估 PI-RADS 分数，消除读片过程中主观因素带来的不一致性，就具有特别重要的科研价值，可以提高前列腺癌的检测率和治疗水平，促进医疗科技的发展。

图 1 AI 代替人工

4. 国内外现状

在阅读近三年国内外文献的过程中，将人工智能（AI）、深度学习、机器学习、前列腺癌、前列腺恶性肿瘤进行组合检索，发现相关文献并不多，外文具有参考意义和价值的不足 20 篇。Stabile A 等人研究发现 mpMRI 的性能通常会根据不同级别放射科医生的经验而有所不同[5]，对 MRI 解读和理解也不同。卷积神经网络（CNN）方法可以在自然图像分析中超越人类的表现，有望增强前列腺 MRI 的计算机辅助诊断[6.7]。扩散加权成像（DWI）、表观扩散系数（ADC）和 T2 加权成像（T2WI）序列可能是临床前列腺 MRI 检查的最重要和最实用的组成部分[8, 9]。基于 CNN 的深度学习（DL）方法彻底改变了现有的前列腺 MRI 阅片的工作模式。DL 研究涉及仅使用扩散加权成像（DWI）、表观扩散系数（ADC）和 T2 加权成像（T2WI）序列中的一个或两个即可进行 PCa 诊断[10, 11]。对于基于人工智能的检测，已经有几项验证研究在调查这些人工智能是否真的对放射科医生的工作产生影响。在最近的一项研究中，Gaur 等人[12]表明，基于人工智能的检测与 PI-RADS v2 [12]分类相结合，提高了特异性。基于人工智能的检测也略微提高了放射科医生的能力，并发现 PI-RADS v2≥3 的病变敏感性指数为 78% [12]。Litjens 等人[13]证明了结合基于 AI 的预测和 PI-RADS v2，改善了对癌症的检测和 csPCa 与 cisPCa 的鉴别。

国内文献，有 5 篇博士、硕士毕业论文以及 6 篇期刊论文。在《基于深度学习与影像组学的前列腺癌检测与 PI-RADS 评分》一文中，应用专家的阅片逻辑，包括修改输入图像的拼接方式，来模拟专家对 MRI 的三个主要参数（T2 和 DWI/ADC 图像）分别解读，使用独特的多任务模块，来分别给 TZ（移行带）和 PZ（外周带）区域癌灶打分。得到的 PI-RADS 人工智能 AI 模型 Kappa 系数达到了 0.716，与人工阅读比较，该模型超越了大多数普通医生。它的高敏感性，有助于缺乏经验的医生校准自己的评分[14]。

以上国内外文献显示，目前虽已有人工智能可以对前列腺影像的 PI-RADS 评分产生影响的研究，可以证明人工智能可以提高诊断的特异性和敏感性，但没有真正实现基于人工智能（AI）的 PI-RADS 自动评分。

5. 拟采用的研究方法

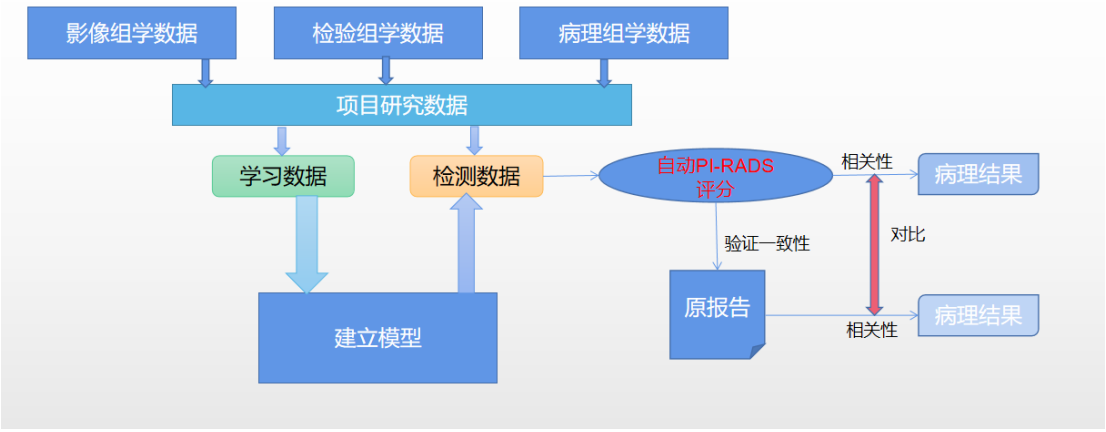


图 2 研究方法

（1）数据收集：收集大量的前列腺 MRI 图像数据，包括前列腺癌和非癌变病变的图像。这些数据应涵盖不同的病例和病变类型， 以确保模型的泛化能力和准确性。

（2）数据标注：对采集的前列腺 MRI 图像进行标注，标注出前列腺癌的位置和严重程度，并将其与 PI-RADS 评分进行对应。这是训练机器学习模型的基础。

（3）特征提取：使用计算机视觉技术对标注后的前列腺 MRI 图像进行特征提取，提取出与 PI-RADS 评分相关的特征。常用的特征提取方法包括卷积神经网络 (CNN) 、支持向量机 (SVM) 等。

（4）模型训练：选择合适的机器学习算法，将特征输入到模型中进行训练。常用的机器学习算法包括神经网络、决策树、随机森林等。在训练过程中，可以 采用交叉验证的方法来评估模型的性能。

（5）模型优化：根据训练结果和交叉验证结果对模型进行优化和改进，进一步 提高模型的准确性和可靠性。

（7）模型验证：使用验证数据集对模型进行测试和验证，评估模型的准确性和泛化能力。如果模型的准确性和泛化能力达到要求，则可以部署到临床实践中进行应用。

（8）系统实现：利用训练好的模型，实现 PI-RADS 评分自动化。构建 PI-RADS 评分 AI 模型，开发出一个准确性高、可靠性好的 PI-RADS 评分自动化诊断系统。

6.结论

综上所述，用人工智能（AI）代替人工进行 PI-RADS 评分具有重要的科研价值，可以提高前列腺癌的检测率和治疗水平，促进医疗科技的发展。将人工智能技术，应用于医学领域，为医疗行业带来更多的创新和进步。

PU-189

人工智能在麻醉学中的应用

闫丽璇

山西医科大学

目前，麻醉学科中面临着许多挑战。 根据国家统计局数据显示，2009—2018 年我国老年人口比重不断上升，人口老龄化趋势日益明显 [1]。此外，我国年度住院患者手术人次平均年增长速度远远超过麻醉科医生年增长速度。同时按照人口比例计算，目前我国麻醉科医师的缺口约为 30 万人。 社会人口老龄化、迅速增长的手术量和麻醉需求以及麻醉科医师的严重缺乏等因素导致我国麻醉科医师高负荷高强度工作，从而引起麻醉质量下降，职业倦怠，甚至患者围术期的风险增加。

而随着人工智能（artificial intelligence， AI）的发展，麻醉学科很有可能从中获益。

AI 可被定义为研究算法，使机器具有推理和执行诸如解决问题、识别物体和单词、推断世界状态和决策等功能的能力 [2]。在麻醉领域，医疗系统中的电子病历、围术期监测设备或患者穿戴的便携电子产品中可产生大量的医疗健康数据。同时它涉及到临床护理的多个元素，包括围手术期和重症监护、疼痛管理以及药物传递和发现。 AI 能在大量数据中识别并提取有用的数据，在如今的临床麻醉中发挥着越来越重要的作用。临床中 AI 应用最广的形式为机器学习，又可分为监督学习、无监督学习和强化学习，它们均可以用于解决分类和回归问题。例如研究者使用模糊逻辑来根据心率(HR)、血压和脉搏量的归一化值来评估轻度、中度和重度低血容量的存在，从而开发了检测低血容量的麻醉监测器 [3]。另外有研究者根据病人的人口统计、生命体征和病史的各个方面、手术类型和自控镇痛(PCA)剂量等特征使用决策树来预测病人 PCA 的总消耗量，优化 PCA 给药方案 [4]。今天机器学习中最流行的方法之一是使用神经网络，在麻醉领域应用于监控麻醉深度及麻醉药物输注控制。另外贝叶斯算法，自然语言处理和计算机视觉等也是常被应用在临床工作中的机器学习方法。

AI 在麻醉学科中的应用

评估麻醉深度 AI 具有强大的数据处理能力和自我学习能力。有多项研究证实，AI 可应用于麻醉管理中，改善麻醉深度的监测。一些研究者使用人工智能技术和频谱分析更直接地分析脑电图信号来估计麻醉深度。Shalbaf 等人 [5] 利用脑电图的多种特征对清醒和麻醉患者进行分类(四种可能的清醒、轻度、全麻或深度麻醉状态)，在使用七氟醚时准确率为 92.91%，而反应熵指数的准确率为 77.5%。同样的算法被推广到丙泊酚和挥发性麻醉患者，准确率为 93%，而 BIS 指数的准确率为 87%。Zhang 等人 [6] 记录了患者的中潜伏期听觉诱发电位，并利用神经网络评估这些信号的准确性，以确定患者何时清醒(准确率 96.8%)，何时接受充分麻醉(准确率 86%)，何时从全身麻醉中苏醒(准确率 86.6%)。Hever 等 [7] 开发的基于专家系统的规则算法及随机森林模型能够有效识别监护仪中的警报与伪象，明显降低机器的误报率。Sadrawi 等 [8] 综合分析受试者的 EEG、肌电图、心率、血压、脉搏以及信号质量指数等多个参数构建的人工神经网络模型能更精确地反映麻醉深度。有研究者 [9] 使用机器学习的控制系统调控麻醉药物的自动化输送，同时纳入了药物药代动力学指标进行预测，以达到最佳麻醉深度。

预测模型的构建 在围手术期护理风险预测方面，AI 各种技术应用于早期预测围术期不良反应如低血压、低氧血症保，保障患者围术期的安全。例如，神经网络被用来预测诱导大剂量异丙酚的催眠效果(BIS 测量)(敏感性为 82.35%，特异性为 64.38%，曲线下面积为 0.755)，并发现超过了执业麻醉师的平均估计水平(敏感性:20.64%，特异性 92.51%，曲线下面积 0.5605) [10]。神经网络也被用于预测神经肌肉阻滞恢复时间 [11] 和诱导后或脊髓麻醉期间低血压发作的恢复率 [12]，而其他机器学习方法已被测试用于自动分类术前患者 ASA 分级 [13]，定义喉镜检查发现困难 [14]，定义意识镇静时呼吸抑制 [15]，和协助决策最佳的小儿手术麻醉方法 [16]。Hatib 等 [17] 使用了两个大型数据库(共 1334 例患者)的有创动脉波形数据，开发和测试了一个可以提前 15 分钟预测低血压的 logistic 回归模型 [17]。其敏感度为 88%,特异度为 87%。 这将给临床医师提供充足的时间纠正潜在的生命机能紊乱,防止严重低血压的发生。在一项单中心随机对照试验中，比较了机器学习警报系统(使用 6 个生命体征参数作为特征)与基于监护仪等警报系统预测和评估脓毒血症，前者的表现更优。并可使平均住院时间缩短 20.6%，住院死亡率降低了 58%[18]。Kendale 等 [19] 利用机器学习方法,分析了 13323 例患者诱导后低血压的发生情况,以患者术前合并症、术前用药情况、诱导用药、术中生命征为临床特征,采用 Logistic 回归、随机森林、支持向量机、朴素贝叶斯、k-近邻、神经网络等方法进行建模,并对性能最优化的模型进行优化,结果证实,机器学习算法可成功预测全身麻醉后低血压的发生。

超声引导 借助 AI 在图像识别和处理方面的优势，可帮助麻醉医生高效快速地识别超声图像，高效安全地实施超声引导下的麻醉操作。Pesteie 等 [20] 使用卷积神经网络自动识别椎板的前基底，

而 Hetherington 等 [21] 使用卷积神经网络从超声图像中自动识别骶骨和 L1-L5 椎骨以及椎间隙等解剖定位，协助麻醉科医生进行硬膜外穿刺置管等操作，其准确率高达 95%。此外，AI 构建的机器学习算法能够自动测量心脏的射血分数等参数，快速评估心脏功能，其结果的准确性相当于心脏病学专家，优于超声科医师 [22]。

围术期危机事件管理 科学的预警和决策，积极处理及预防，能化解围术期危机事件。Amaral 等采用 k 近邻、随机森林、AdaBoost 等算法,研究哮喘发作时气道梗阻表现的诊断,结果表明,所有的算法均可较准确的诊断出哮喘发作 [23]。Olsen 等 [24] 采用卷积神经网络算法训练并建立了一个基于 44959 例患者的心电图和经胸心脏超声数据判断心衰与否的系统,其 AUC 可达 0. 93, 准确性达 85. 7%。

疼痛管理 借助机器学习能更有效地识别疼痛，但目前应用不是特别广泛，有待于提高。为了识别更客观的疼痛生物标志物，Gram 等人 [25] 使用机器学习分析了 81 名患者的脑电图信号，试图预测哪些患者会对急性疼痛的阿片类药物治疗有反应;结果表明术前脑电图对术后阿片类药物治疗有反应的患者的评估只有 65%的准确性。此外，Olesen 等人 [26] 进行了一项大数据研究，寻找 1237 例癌症患者的单核苷酸多态性，以预测这些患者的阿片类药物剂量;然而，他们的研究没有发现这一人群中任何单核苷酸多态性与阿片类药物剂量相关。

小结：

既往围术期对麻醉下患者的监测依赖于麻醉医生的经验来调整麻醉剂、神经肌肉阻滞和心血管药物，随着医疗技术的进步，现代麻醉学要求麻醉医生拥有分析处理及解释数据的能力，从而更安全地管理病人的麻醉。尽管目前可用设备如 BIS 及脑电图等监测器可提供一些数字化监控，但这些设备所依赖的算法有限，在极端情况或年龄时可能并不可靠、有时受所用麻醉药物类型的干扰、或运动及电刀干扰。而 AI 强大的计算及分析能力则可以规避这些缺点。

但 AI 也存在一定的局限性。AI 尤其是神经网络可能产生黑盒效应，即算法可以将预测通知给临床医生或研究人员，但不能解释背后的原因。相较于决策树，深度学习是通过归纳的方式进行评估，更不易于解释。因此在可解释的人工智能方面，人们正在努力提高算法的透明度。另外数据收集及隐私保护方面仍值得关注。

一些研究表明，在测量临床终点(如麻醉深度)方面，人工智能的方法和技术比现有的方法和技术有了改进，而另一些研究则表明，与目前可用的和广泛使用的技术相比，人工智能的性能没有区别，甚至更差。不同于有明确规则的棋类运动，临床医学具有显著的不确定性，许多临床数据只是潜在的生理或病理过程的表征，而不是直接测量。机器学习的优势之一是它从数据中学习的能力，这与现有的静态算法方法相比具有明显的优势。随着我们构建更好、更完整的数据集，我们可能会看到 AI 在医学上更复杂、更有临床意义的应用。

AI 为麻醉学科的发展开辟了新的可能。在 AI 的帮助下能显著提升麻醉科医生的诊疗效率和准确度，降低疲劳工作，提升围术期安全度。但是大多数 AI 在麻醉学上的应用仍处于研发阶段;因此，目前麻醉学领域人工智能的重点不是取代临床医生的判断或技能，而是研究如何增强这些判断或技能。未来还需要更多的研究来探索 AI 在麻醉学科中的应用。

PU-190

人工智能在医疗领域的应用与思考

周枫
郑州大学第一附属医院

人工智能（Artificial Intelligence, AI）浪潮席卷全球，人工智能技术为社会发展带来巨大变

革，是当前最热门的科学技术话题之一。人工智能技术的发展推动政府、企业、个人都涌进人工智能大潮。相较于传统医疗手段，人工智能技术在医学领域应用获得许多颠覆性的改进，它有利于帮助人类解决医学难题，甚至在未来实现人类改变生老病死的梦想，掀起新一轮的革命。目前我国人工智能在医学应用发展迅猛，科学界和各行业关注人工智能，我国制定的人工智能的发展计划旨在促进各领域与人工智能迸发出火花，通过研发不同的人工智能应用，以提升生产和生活效率。

1.人工智能的概念

人工智能的定义最早是被麻省理工学院的约翰麦卡锡在 1956 年的达特矛斯会议上被提出来的。早些年的人工智能的定义只是为了让机器表现出来人的一些行为。当今对人工智能又有了新的定义，是指让机器可以像人一样理性的思考和分析问题，并做出相关的行动。目前，人工智能可以跟人的大脑一样进行事物的推理和一些数据的运算，但人工智能通常还可以完成人脑无法完成的数据运算。从 21 世纪开始人工智能技术逐渐具备深度学习能力，深度学习就是通过算法，来对数据进行处理和分析。对于人工智能来讲，算法是它的灵魂跟核心，而大数据跟算力则是其基础。

2.人工智能在医疗领域的应用

当前，在医疗卫生领域中，人工智能应用于诸多场景：健康管理、临床决策支持、医疗虚拟助手、医疗诊断、疾病筛查和早期疾病预测、病历分析、医院管理和药物研发等，这些应用关乎人类的生命健康。

2.1 医疗虚拟助手

自 2022 年 11 月，ChatGPT 迅速成为现象级焦点，引发市场广泛关注，掀起新一轮人工智能浪潮。推出仅两个多月，ChatGPT 全球用户已突破 1 亿，成为历史上用户增长速度最快的应用程序。在医疗卫生领域，ChatGPT 可以成为辅助医生诊断和选择治疗 方法的智力支持。在慢性病研究方面，可以借鉴 ChatG PT 中 PPO 强化学习模型，应用马尔科夫链决策过程，动态学习前后病因影响大小、不同地域病因的影响关系，强化模型值与人工值差距研究，可以建立有成效的分析和决策系统，为慢性病等病因分析和治疗提供人工智能支持。在医院管理和卫生系统管理方面，ChatGPT 可应用于医院内部的自动化客服和咨询系统，为患者提供在线预约、挂号、查询检查结果等服务，减轻人工接待压力，提高工作效率；医院知识库管理和知识查询，帮助医护人员快速获取医学知识和临床指南，支持决策和诊断过程，提升医疗质量和安全性；卫生系统公众健康教育，通过智能问答和健康咨询，为大众提供个性化健康建议和预防措施，促进公众健康管理和健康意识。这些应用可以提高服务效率、增强医学知识的传播和应用、促进公众健康意识，为医疗行业的改进和发展作出积极贡献。

2.2 疾病筛查和预测

人工智能可以应用在医疗领域进行疾病的筛查，比如精神疾病诊断、自闭症筛查、通过病灶筛查糖尿病视网膜病变等。此外，人工智能也可以协助脑疝预测、慢性肾病预测、心脏病病人死亡预测、骨关节炎发展预测和流行病风险预测等疾病预测。遗传疾病方面，脸部图像分析框架 DeepGestalt 利用计算机视觉和深度学习算法识别遗传疾病的面部表型，它在 502 个不同图像上识别正确综合症的准确率达到 91%。

2.3 医疗诊断

人工智能可以在临床环境中提供算法解决方案，帮助诊断、治疗、预后，在放射学领域的视觉模式识别能够帮助解释成像。计算机可以存储正常人类的生理特征、已诊断疾病的病理模型和历代专家医生所总结的医疗知识，并根据患者不同的疾病特征，进行演绎推理，诊 断疾病，给出诊疗意见和方案。20 世纪 50 年代，就有把疾病症状与已有诊断病例作逻辑匹配的逻辑方法， 1958 年运用计算机结合数据统计和逻辑技术进行医疗诊断，60 年代引进贝斯判定、决策力以及模式识别，70 年代相关学者绘制了人工智能医疗诊断系统简图，人工智能逐步应用于医疗诊断。目前人工智能诊疗已得到迅速发展与广泛应用，Liu Wenjie 等用 JAVA 设计了人工智能蛋白分析程序，对具有代码的蛋白成分可能存在状态进行逻辑分析，对多种生物标记物基础分析通过远程医疗接口在

人工智能辅助医学诊断中提供潜在的应用。江涛通过患者数据收集以及病情医学数据分析构建模型，建立人工智能肺癌诊疗技术， 并以肺癌 CT 诊断为例进行实验验证分析，发现人工智能技术不会受到干扰序列影响，实现肺癌人工辅助诊断，对临床诊断具有实际应用价值。

2.4 临床决策系统

临床决策支持系统（CDSS）旨在为制定临床决策的医疗卫生专业人员提供专家支持计算机程序，其通过人机对话，根据病人数据并做出有关诊断、预防和治疗的决策，提高决策有效性。目前影响范围最广的医疗辅助诊断系统是“沃森”系统，该系统是由 IBM 与美国肿瘤医院共同开发完成的。沃森医生平台在医疗领域的运用，主要是辅助诊断人类疾病的一大难题——癌症，该系统通过自然语言处理和深度算法学习，整理并分析上万份的临床病例、治疗方案、近百本期刊和论文，从中可以快速的提取有效的信息，从而为医生的诊疗过程中提出合理的方案，从而进一步提高了诊疗的效率和精准度。

2.5 病历分析

人工智能技术能够对电子病历进行结构化处理，提升医护人员检索病案信息的效率，并从中挖掘多元数据。在医院信息系统（HIS）、实验室信息管理系统（LIS）、放射信息管理系统（RIS）、医学影像存档与通讯系统（PACS）和电子病历（EMR）等院内系统中，医学人工智能应用可以将病人信息和病案集合处理，通过时间顺序展示每位病人的门诊、住院病历信息，满足日益升级的医疗需求，解决医院信息互联问题，方便医生诊治和病人就诊。

2.6 健康管理

传统的健康管理通常在线下进行,效率比较低下，且很难切合用户个性化的真实需求，无法观测和激励用户的健康自控行为，难以达到预期效果。人工智能健康管理应用的工作机制是：人类通过穿戴可穿戴设备记录穿戴人的生命体征数据，人工智能健康管理应用对人体进行监控评估，并提供干预方案和建议。运用运用人工智能技术，不仅可以从门诊和住院电子病历中采集个体化的数据，还可通过可穿戴设备和智能家居，对用户健康数据信息和行为习惯进行监测和采集，并同步传输到医疗健康数据平台进行筛选、提炼和分析，进而预测个体的疾病易感性、药物敏感性等，自动匹配健康管理知识库，进行有针对性的干预。目前，主要应用于疾病风险预测与干预、慢病管理、运动管理、睡眠监测、母婴健康管理和老年人护理等。

2.7 药物研发

药物研发具有周期长、困难、低效、高成本的特点，并且研发药物的疗效跟副作用无法预测，只能后期在临床应用中进行观察。利用人工智能技术不仅可以减少研发周期，降低成本，还可以运用计算机算法计算能力和数据分析的能力，对该药物的众多分子结构进行分析 和排列重组，最后进行药物的合成模拟，可以有效的减少实验次数降低成本，预测新药物的效果跟副作用。目前通过人工智能技术辅助药物研发最多的就是心血管疾病相关药物、肿瘤相关药物、预防传染病的药物等，人工智能技术还在研究新冠疫情特效药和疫苗中发挥了 重要作用，通过对病毒的蛋白、基因组 RNA 结构进行推算，加快了新冠疫情药物和疫苗的研发进程。

3.挑战与思考

网络技术、基础设施、理论知识的进步使得信息技术和工程突飞猛进的发展，高精尖、技术先进、计算快速的方案变得越来越实用，极大的促进了人工智能等相关领域的发展壮大，人工智能机器人和人工智能模型在疾病诊疗管理中的崛起具有深远的临床意义。1989 年的一项研究显示，在门诊患者中，美国医生无法根据患者症状，得出准确诊断结果的高达 90%，许多学者认为，目前临床数据收集和医疗诊断过程受到人类分析能力限制，人工智能可以从根本上改变医疗决策中的不确定性。对于医生来说，通过人工智能可以辅助诊断，减少筛选对比病例的时间，为患者制订准确的治疗方案；对于患者来说，可以更快速的完成健康检查，获得更为精准的诊断建议，节省大量的时间、金钱成本。

人工智能的发展离不开人才的挖掘和培养。现阶段国内交叉学科人才普遍 匮乏，多数 AI 技术

人员没有医学背景，医学专业教学单一，缺乏系统性交叉学科的培养体系和方案。 医学信息学与生物工程学近几年发展趋势良好，但与医学素养、AI 技术相脱离，无法很好地培养跨学科人才。

人工智能医学应用在我国面临若干伦理挑战。人工智能医学应用在实践过程中可能产生安全性问题。当它对人造成伤害时，面临严峻的责任归属问题。人工智能医学应用的基本要素之一是医疗健康数据。这些数据因有意、无意泄露、或遭受网络攻击，最终导致隐私泄露，而人工智能医学应用的算法偏倚会增强对个人和群体的歧视。一些人工智能医学应用的治疗价格高于传统医疗手段，能让公众享受得到且负担得起相关服务是一个巨大挑战。此外，人工智能医学应用面临可解释性和透明性问题，它们可能导致公众的不信任。当医生建议或决定利用人工智能医学应用时，这种人机之间的不信任会使公众和病人怀疑医生是否为病人的利益做出最佳的判断，从而影响医患关系。我国亟需采取有力的伦理措施解决人工智能医学应用的伦理问题，提高各方利益相关者的伦理意识、建立健全人工智能医学应用的法律法规、监管标准、行业规范和公司准则，鼓励多方参与制定有关政策，加强相关培训和国际合作。

PU-191

医疗云应用的前景及优势

詹汉强

郑州大学第一附属医院

前言

随着云计算的发展，医疗云在国内外得到政府的大力支持和推广，医疗云存储安全问题也受到了广泛的重视。接下来，本文针对医疗云存储的特点和国内外研究进展，进行详细的描述和分析。

随着云计算技术的出现，使用户有能力利用可扩展的分布式计算环境。在云中，包括存储在内的计算资源由第三方服务商提供[1]。由于云中的数据通常存储在共享的环境中，因此用户不知道他们的数据的确切位置，也不知道与他们共同存储的数据的其他来源[2]。云存储系统的本质是以提供海量存储服务为核心任务的云计算系统[3]，云存储系统是云计算服务的重要基础，承担着收集、存储和处理数据等最底层的任务。云存储系统的主要组成部分为：存储硬件设备、应用软件程序、访问接口等。从存储结构上划分可以分为存储层、基础管理层、应用接口层以及访问层。（1）数据存储层：云存储系统利用软件通过网络将各类分散的存储设备统一管理、协同工作，提供超大存储空间和超强处理能力。数据存储层的核心是通过应用软件管理接入网络的各类存储设备，将分散的存储设备集中管理实现海量数据存储池，并可以对硬件设备的进行实时监控和容量动态扩展,是云存储系统的基础。（2）数据管理层：数据存储层之上则是数据管理层，主要为不同的存储服务向上一层提供统一管理界面。将数据存储层与数据服务层之间通过统一的用户、安全、副本及策略等管理方式进行无缝连接；（3）数据服务层：是云存储系统中面向用户的一层，为用户提供可应用服务。可以按照用户的不同需求，开发特定的应用接口，提供个性化的存储服务，也可以根据存储的需求量进行灵活扩展。（4）用户访问层：用户可以通过访问层直接访问云存储系统，随时随地通过互联网使用云存储服务，这是云存储服务的显著特点之一[4]。目前，在多种云存储的访问方法中最常用的是 Web 服务 API（Application Programming Interface）。 云服务提供商在存储技术和管理能力方面具有明显的优势，在专业技术服务能力、数据的容灾备份等方面，远胜于个人用户[5]。

云存储服务与个人存储管理

云存储服务与个人存储管理存在明显不同，主要表现在以下几个方面：一、存储功能方面。云存储提供在线存储服务，面向互联网，可应用于不同类型网络，个人存储是指面向单一用户，提供单一的存储、处理等应用；二、存储性能方面。云存储系统面向互联网用户，提供庞大的存储能力和丰富的服务类型，云存储在复杂的网络环境下对数据安全和存储效率等方面具有较高要求，需要更强大的技术支撑，个人存储则存储能力有限、类型单一、技术水平相对有限；三、数据管理方面。云存储是提供简单的数据存储、管理服务，同时需要实现复杂大数据的管理、分析、处理等功能，为其他数据服务提供支撑[6]。

云存储服务类型主要是面向个人用户的网络存储空间和在线访问等数据服务，面向企业用户的数据存储、管理、处理、备份与容灾等服务。目前，很多大型 IT 公司将云计算和云存储服务作为核心业务积极开展推进，促使云计算和云存储的相关技术突飞猛进。国际上大型的云服务提供商主要有：Amazon、Google、Rackspace 等公司，分别推出他们的云服务平台：Simple Storage Service (S3)、Google Cloud Storage、Cloud Files 等。国内的云存储服务也快速崛起蓬勃发展，代表性的公司主要有：阿里巴巴、金山、华为、腾讯等公司，分别推出了：“阿里云”、“金山快盘”、“DBank ”、“网盘”、“QQ 网盘”等云服务平台。同时，云存储服务的发展也面临着安全挑战。趋势科技率先推出“云安全”概念及技术后，瑞星也启动了“云安全”战略。此外，卡巴斯基、Mc Afee、赛门铁克、绿盟科技、山石网科、安全宝、瀚海源、梆梆安全、金山、360 等安全企业也都推出了各自的云安全解决方案[7]。由于云存储服务发展呈现规模化、平台化、网络化等特点，与原有的私有存储相比，若云存储出现安全问题，所造成的经济损失和社会危害会更大。因此，云存储安全的重要性更加不容忽视。

医疗云及发展现状

医疗云主要是指将医院的预约挂号、收费、电子病历、医生的电子处方以及检查报告和医疗影像文档等临床检验信息整合为完整的医疗电子健康档案(Electronic Health Record-EHR)系统，并将 EHR 档案存储到云服务器便于患者查询或者其他医生的诊断依据以及保险公司的评价依据和远程医疗、医疗研究数据信息的来源等[8]。依靠医疗云可以建立“多对多”的健康信息交流平台，患者可以更快的与医生、医院、保险公司等机构进行沟通，医疗云平台也可作为医疗信息系统、远程诊断系统、远程监护系统以及远程医疗教育系统的基础平台。医疗云与传统医院或卫生机构的信息管理系统相比具有如下优势：一、数据安全。医疗云采用多种网络安全措施，降低医疗信息被窃取的风险；利用云存储安全措施，定期备份医疗数据，提高医疗数据的安全性。二、信息共享。各医院、城市卫生系统以及全省全国的卫生系统信息可以整合，有利于各医院、研究机构、相关企业与政府部门的信息共享。三、动态扩展。医疗云平台可使各医疗单位的医疗信息系统的存储性能、访问性能、灾备性能等进行动态扩展升级。四、无限布局。医疗云的兼容性和远程可操控性，使得医疗云可形成无限覆盖的医疗健康信息平台，在云平台内可共享医疗信息。五、经济实惠。各医疗机构内部不需要再部署相关硬件设施和管理技术人员，可以按需购买云服务以支持医疗云平台的建设。

随着云计算技术的成熟，医疗云发展建设规模增长快速，国内外各大互联网公司先后推出相应的医疗云平台，主要有：亚马逊医疗云、阿里医疗云、英特尔医疗云、腾讯医疗云、华为医疗云、百度医疗云、中国移动医疗云平台、金山医疗云等。目前的云存储、云计算服务的相关技术已经能够满足医疗机构的各类需求，云服务正在快速的融入医疗行业，医疗云已经完成了市场培育期，进入快速发展阶段。为了便于管理、控制和确保安全，目前各大医疗机构的医疗云多数是以私有云为主，统计数据表明 46.5%的医疗云用户优先选择部署私有医疗云。而且，在保持现有的数据中心的同时，独立建设医疗云是医疗机构部署云计算的主要方式。此外，也有很大一部分医疗机构首先进行原有数据的虚拟化，然后逐步过渡到云计算平台。由此可见，医疗行业用户在部署医疗云时一般选择比较谨慎的方式循序渐进，各医疗机构目前对公有云的安全性和隐私性还有较大的忧虑。医疗云系统中的信息安全是复杂的综合性技术问题，基本的安全因素涉及到用户终端、医疗健康信息系

统和云平台网络环境等各个环节，一般 IT 架构的防护措施和安全技术已无法满足目前医疗大数据的安全需求[9]。因此在医疗云深入发展的过程中，安全保障技术能力和措施是实现“可信医疗云”的关键。目前多数医疗机构都有自己的数据库来管理电子病历等医疗信息，在同一医疗信息系统内交换数据更容易。但是，不同的医疗信息系统之间的医疗数据交换速度会变慢[10]。

医疗云数据共享

一些小规模医院没有相关的服务器来管理医疗信息，也没有提供医疗信息交换的能力。云存储平台可以为所有医院和诊所提供信息交换平台，可以作为医疗信息存储中心，这不仅可以简化不同医疗系统间复杂的数据交换流程，而且可以节省大量设备购置和管理的费用。另外，患者可以在医疗云平台的的一个界面中，查询到完整医疗信息，不必在不同医院的医疗信息系统中进行分散查找就诊信息。因此，云服务商提供的云计算已成为医疗机构部署医疗信息系统合适的平台[11]。 虽然众多医疗机构纷纷采用云平台部署医疗信息系统，共享医疗数据，但是数据的隐私和安全问题阻碍了云平台在医疗领域的应用[12]。关于医疗信息的隐私保护，常见的保护措施是通过匿名和化名的方法来隐藏患者与就诊信息之间的不可链接性[13]。目前，围绕医疗数据的各种安全和隐私问题已经开展了广泛的研究。 许多研究者或机构提出了互联网环境下处理医疗数据的安全和隐私问题的报告，其中健康保险可移植性和责任法案（Health Insurance Portability and Accountability，Act-HIPAA）和欧洲数据保护指令 95/46/EC[14]被广泛采用。医疗数据在传输过程和存储过程中的隐私保护是医疗数据共享存在两个基本安全问题，前者已经被广泛研究，可以采用安全套接层协议（Secure Socket Layer-SSL）和传输层安全（Transport Layer Security-TSL）协议解决。后者在云计算范例中的存储服务的研究较少，而且与云存储在云服务提供商站点的外包数据相关性更高[15]。云存储的特点是灵活、可扩展、经济、便捷等，可根据用户的不同需求制定不同类型的云存储服务系统，根据硬件基础的部署模式不同分为四种：公有云[16]、私有云[17]、社区云[18]、混合云[19]。不同的部署模式主要考虑的是成本经济性、风险大小、管理模式和所提供的服务形式等非技术层面因素，在技术层面上集中部署模式基本大同小异差别不大。同样，医疗机构的医疗云可以是自建的，也可以购买云服务提供商的云服务。通常医疗机构根据自身使用需求和经济实力，选择医疗云的建设模式，医疗云根据不同的硬件部署模式也可以分为以下四种模式：(1)公有医疗云：医疗机构根据实际需求购买第三方云服务商提供的云服务，云服务商搭建医疗云平台，提供计算处理、数据存储、软件应用等医疗信息服务，按需计费。医疗机构通过互联网访问云服务平台，实现对医疗信息系统的部署、使用和管理。如前面提到的亚马逊医疗云、阿里医疗云、英特尔医疗云、腾讯医疗云、华为医疗云、百度医疗云、中国移动医疗云平台、金山医疗云等[20]。(2)私有医疗云：医疗机构自行建设和使用的医疗云平台，只为指定单位提供特定的云服务，在系统管理和数据安全方面可控性强。私有医疗云由医疗机构自行进行基础设施、平台及软件服务建设，安全性高、服务质量好，但是需要较大的资金和人力投入。(3)社区医疗云：具有同样的需求和目标的多家机构共享医疗云基础设施，合作建设，共同使用，利益与风险共担。(4)混合医疗云：公有医疗云与私有医疗云结合利用，这种部署模式一般是由私有云承担机密性较强的数据处理，公有云承担私密性低却需要消耗大量计算或存储资源的任务，二者结合可以充分利用云计算的处理能力和数据安全保密。

医疗大数据的特点与应用分析

随着互联网和云计算等信息技术的发展，现在社会中的各行各业充满越来越多的数据信息，大量的各类数据成为一种新的社会资源需要分析、处理以为人们的生活生产服务，充分高效的利用数据资源将会为人类社会创造更巨大的财富。

大数据(Big Data)[21]：指快速增长、类型复杂、容量巨大、多样化的信息数据。传统的数据处理方式已经无法处理具有这类特点的数据，因此需要采用新的特殊处理方式才能对大数据进行处理。根据国际数据公司和 Gartne 对大数据的定义，可将大数据概括出 4 个特征[22]：超大容量

(Volume)、超多种类(Variety)、超快增长(Velocity)、超大价值(Value)，简称为 4V 特性，这一定义基本涵盖了大数据的基本特征和一般属性。显示了大数据的“4V”特性的具体描述： 在医疗领域，医疗数据和健康信息具有多样性，且因各医疗机构的信息系统之间的异构性，信息之间重复少，难以共享，这造成医疗信息容量比较庞大[23]。同时，随着医疗机构信息化的全球普及，各类医疗信息系统正在迅速的增长，随之而来的是所产生的医疗数据急剧增长。要发挥医疗大数据的作用，需要更多的资源，更强大的处理能力来处理分析。医疗大数据的处理分析有助于医学研究、医药开发、疾病诊断、卫生监控等。随着云计算与云存储的快速发展，医疗大数据正在悄然快速的产生，云计算催生医疗大数据的同时，也提供了高效的医疗大数据的存储与处理方法。这里的医疗大数据是指医疗领域内产生的对相关机构及研究人员有价值的大数据，主要有以下几个方面的来源[21]：（1）医院的临床医疗数据；（2）医学研究机构的实验室数据；（3）卫生管理部门的监控统计数据；（4）个人的健康信息数据；（5）医药企业的研发数据；医疗大数据的基本类型：（1）医疗机构管理信息系统数据：医疗机构的管理信息系统是指医疗机构的人、财、物等行政管理和就诊等业务信息融合的综合管理信息系统，在医疗机构中的各类活动的不同阶段生成的数据，主要包括门诊、病房、人事、财务等数据信息。（2）检验分析信息系统数据：检验分析系统的主要功能是对检验仪器输出数据进行分析生成检验报告，以检验报告的形式将检查数据存储在服务器中，医生可以随时访问服务器查看或下载被检查人员的检验报告。（3）医学影像信息系统数据：是指医疗机构的影像科室产生的医学影像资料数据，主要包括核磁、CT、超声、各种 X 片、心电图以及其他检查设备生成的的图像数据资料。（4）个人健康信息数据：个人健康信息数据不仅仅是就诊人员的电子病历，是指诊断和检验数据以及个人信息、财务信息等数据组成的全部信息，以统一的格式组织起来。

医疗大数据的多元优势

医疗大数据的特性除了普通大数据的“4V”特性以外，还具备以下几个特点：

（1）多态性：医疗大数据中包含数值、图像、文档等多类型数据，同时也包含动画、语音、视频信息等多种形态的数据，这些特征与其他领域大数据有明显区别。（2）不完整性：医疗大数据的搜集经常与处理过程存在非连续性，因此会造成医疗大数据的不全面，尤其是来源于人工录入的大量数据，可能存在严重的偏差和残缺。由于电子病历和诊断记录等医疗数据的自身表达和记录的不确定性，造成了医疗大数据的不完整性。（3）时序性：大量医疗数据具有明显的时序性，如诊断信息，疾病的发生、诊断、检验等过程的时间顺序，医疗数据的采集时间顺序等。同时，诸多医学影像函数也与时间相关，因此医疗大数据具有明显的时序性。（4）冗余性：医疗机构的各类信息系统每天会生成大量的就诊信息，其中很大一部分是重复的甚至是不必要的记录，这造成了医疗大数据冗余。对于医疗大数据的处理分析方式可概括为三个原则[24]：一、处理分析的对象是全体数据集，而不是部分随机数据；二、处理分析的数据具有混杂性特点，而不是单一的精确性数据；三、处理分析的数据是相关关系数据集，不是因果关系数据集。大数据的“4V”特性对于医疗大数据同样适用，因此，不能采用传统的医疗信息处理方法。随着医疗信息化的不断深入发展，医疗大数据急剧膨胀，但是，因为医疗大数据多样性和复杂性很难被直接用于医学研究和医疗服务，也给医疗数据的处理带来更大的技术挑战。如今，医疗工作者和科学研究人员可以充分利用云计算系统充分发掘利用医疗大数据的价值，且已发展为医学信息研究的热点。目前，医疗大数据的分析处理方法通常采用基于集体智慧(Collective Intelligence)[25]的个性化推荐系统。主要通过对集合整体的行为和数据信息的分析，获得有关集合整体的统计结果。获得的统计结论一般是表现集体的共性部分或某种趋势，这样的结果通常不能通过对个体行为的分析获得。对医疗大数据进行分析可以帮助构建各种个性化推荐系统，如临床决策支持系统等。关于医疗云中的大数据处理模式和共享方法已经有诸多研究工作开展，Yang 等人[26]基于医疗记录属性的分类，提出用医疗数据集的垂直分区来实现对不同隐私关注点的医疗数据不同部分的考虑，实现了大规模医疗数据访问和共享的原型系统。

Li 等人[27]提出了改进的患者与电子病历之间的不可链接机制，进而实现医疗大数据的安全共享。上述方案针对不同的医疗云环境和安全需求，分别提出相应的医疗大数据处理方案。

医疗大数据的应用前景

医学信息化有助于提升医院医疗服务水平和提高医院医疗服务质量，是以“患者为中心”理念解决“看病难、看病贵”问题的有效途径。因此，医学信息化已得到越来越多的重视，逐渐成为我国医疗改革进程中的工作重点。但其成长、完善、成熟还需要经历长期的发展过程，现有医学信息化还存在着广阔的发展空间。并随着计算机技术及信息技术的发展，医学信息化要结合自身的特点与需求快速应变以利用新的技术提升自己的水平和能力。

目前，利用云计算核心技术及理念解决医院信息化自身建设存在的问题及医样大数据的二次利用的难题已经得到越来越多的重视，逐渐成为我国推进医疗改革过程中的必然发展趋势。但由于医疗大数据的复杂性和云计算范畴本身包含的技术及概念的多样性，将云计算运用到医院信息化建设及医疗大数据的二次利用等问题中来，还需要长期的实践及探索过程，医疗云与医疗大数据还存在着广阔的发展空间。

PU-192

基于 RSM 框架的生存预测模型

唐朝飞
南京医科大学附属无锡市人民医院

1 引言

生存分析是机器学习中一个定义良好的问题，用于估计随时间推移感兴趣的事件发生的概率。此类事件的一个例子是器官移植后受者的器官衰竭，或入住重症监护病房(ICU)的病人死亡。深度神经网络(DNN)的出现及其在生存分析领域优于传统的基于 COX 的和浅层机器学习模型(如逻辑回归和随机生存森林(RF))的表现激发了本文的研究。

随着深度神经网络的新进展在生存分析应用中越来越普遍，确保其运行可靠性变得至关重要。DNN 模型通过学习观测值和事件之间复杂的相互关联关系，在估计事件的概率密度函数(pdf)方面通常优于传统的生存模型。Cox 和 RSF 等传统生存模型的决策过程简单，易于研究人员解释。另一方面，DNN 可以学习特征和目标之间复杂且相互关联的关系。然而，深层神经网络的内部工作过程看起来像一个黑盒子，因此，不容易被研究人员解释。

因此，DNN 的预测，特别是在医疗保健环境中，可能不容易解释或信任。已经有一些尝试在 DNN 预测上获得医疗专业人员的信任。Ribeiro 等人开发了一种名为 LIME 的算法，通过添加解释器模型(如决策树)来识别与预测相关的一系列重要特征，从而使分类器或回归模型具有可解释性。虽然该算法应用于文本处理和计算机视觉，但尚未应用于生存分析。将 LIME 应用于生存分析将会非常困难，因为生存函数没有基本的真理。LIME 算法需要同时训练至少两个模型，这使得训练过程非常复杂。LIME 的另一个缺点是，对特定问题的结果是主观的，其准确性无法通过统计来衡量。Yousefi 等引入了一种使用深度神经网络解释生存分析结果的管道。他们使用 Cox 偏似然作为成本函数，并将计算出的风险反向传播到网络的第一层，以确定每个输入特征对应于患者预后的风险因子。计算出的每个输入的风险为整个模型提供了一个解释因子。尽管该模型提供了一种使 DNN 模型可解释的方法，但当电子健康记录(EHR)队列具有纵向测量、缺失值、审查记录或竞争风险时，它不能使用。虽然提议的管道提供了最相关特征的排名列表，但它不能根据结果的相似性或

EHR 记录作为预测的信任来源来识别感兴趣的患者。

在医疗保健领域，确定具有与正在调查的病例相似的临床测量和结果的患者被认为是信任的来源。Gallego 等人开发了一种流程，可以根据过去的临床决策和临床判决，从数据库中提供相似的患者给单个患者。Sun 等使用广义马氏距离[18]根据医生的反馈得出相似的患者。他们的方法是一种监督学习方法，用于基于关键临床指标的 EHR 患者聚类。这种方法不适用于生存分析，因为生存分析的输出是一个没有基础真理的生存函数。Li 等人[19]讨论了几种在训练用于基因组研究的 DNN 时对特征进行排序的方法，其中所有这些方法都在模型的代价函数中添加一个正则化项来对输入特征进行排序。虽然该技术对输入特征进行排序简单有效，但它需要 DNN 目标的基本事实，以便能够对与每个类相关的重要特征进行分类。

2 模型与方法

生存分析的目的是根据观察结果估计一个或多个事件的发生时间。在医疗保健中，这些事件可以包括死亡时间、疾病开始时间、器官衰竭时间[26]等。在生存分析中，危险函数或相当于危险率显示了随着时间的推移，诸如死亡或器官衰竭等医疗事件发生的概率。开发可靠的生存模式面临着几个挑战。一些电子测量是纵向的，即记录每位患者随时间的每次临床测量的多个观察结果，这往往是高度稀疏的。因此，任何可靠的生存模型都必须有效处理缺失值，即使缺失率极高。另一个挑战是如何处理正确的审查数据[30]。正确的审查数据通常发生在医疗中心在一段时间后失去对患者的跟踪时，这段时间被称为审查时间。开发生存分析模型的另一个主要挑战是缺乏风险事件概率分布的基本真值。这些挑战使得开发用于生存分析的统计模型或机器学习模型的监督训练变得困难。最近的研究表明，与传统模型相比，深度学习模型在预测危险函数方面具有优越的性能。例如，Dynamic-DeepHit (DDH)是一种非参数深度生存模型，能够处理纵向测量。深度生存机器(DSM)是一种基于机器学习的参数模型的示例，该模型假设生存函数的多个威布尔和对数正态分布的组合。

随机生存森林(RSFs)方法是随机森林的扩展，支持对右截尾数据进行分析。DeepCox 提出了用于生存分析的 DeepCox 混合物(DeepCox mixture, dcm)，它通过一个混合模型推广了比例风险假设，假设存在潜在群体，并且在每个群体中，比例风险假设成立。最后一个例子是 Survival Seq2Seq，这是一个非参数多事件深度模型，能够处理具有非常高缺失率的纵向测量。生存 Seq2Seq 性能优于本文提到的其他模型，如 DDH。我们正在考虑将生存 Seq2Seq 作为下一节中描述的反向生存模型(RSM)的推导的一部分。

3 实验

本文在两个数据集上评估了 RSM 的性能:一个是部分类似于医疗数据集的合成数据集，另一个是著名的 ICU 数据集 MIMIC-IV。以下小节提供了每个数据集的详细描述。

右审查是医疗数据集的一个共同特征。患者经常在随访中丢失。因此，他们的医疗记录在审查时间之后没有被收集。为了考虑这种现实世界的情况，我们随机选择了一半的数据，即 10000 个数据样本，进行右审查。因此，每个数据样本用(x÷i, si, ki)表示，其中 si 表示给定数据样本的事件时间是否被右截(si = 1)或不被右截(si = 0)。ki 表示未被截的数据样本的事件时间和被截的数据样本的丢失时间。在纵向数据集中，如 MIMIC-IV，在给定时间内只记录少量临床测量值，其余临床测量值未记录。有人指出，缺失值及其缺失模式通常与目标标签相关，即信息缺失，这导致纵向数据集的高缺失率。我们通过为不同程度上与标签相关的协变量创建缺失模式，将这种随机值引入合成数据集。我们在数据集中引入了 45%的随机值。我们还在临床测量中引入了高达 5%的随机缺失值。综上所述，合成数据集的总体缺失率为 50%。

MIMIC-IV 是一个大型免费数据库，包含与 20 多万名患者相关的去识别健康相关数据，这些患者分为三个模块:核心、医院和 ICU。MIMIC-IV 的文件可在其网站上找到。在本研究中，我们使用 ICU 模块，该模块包含 ICU 患者的临床测量和结局事件，可用于生存分析(详见附录 A 和 B 部分)。

4 结果

在本节中，评估了 RSM 识别与事件时间最相关的变量的能力，并对 MIMIC-IV 和 Synthetic 数据集的相似患者进行排名。

	评价指标	百分比			
		25%	50%	75%	100%
MIMIC IV	MAE	35.63±4.3	38.06±4.5	39.44±4.2	62.75±3.3
MIMIC IV	CI	0.886±0.03	0.892±0.03	0.886±0.03	0.917±0.03
Synthetic	MAE	11.95±0.7	12.48±1.3	14.11±1.5	15.64±1.9
Synthetic	CI	0.884±0.10	0.787±0.04	0.773±0.06	0.818±0.09

为了评估 RSM 的性能，本文将其应用于中描述的 Survival Seq2Seq 模型。生存 Seq2Seq 最近被提出作为一种最先进的生存分析模型。Survival Seq2Seq 在 MIMIC-IV 和 Synthetic 数据集上的表现。

Survival Seq2Seq 预测数据集中每个数据样本和事件的 PDF。显示了一组 7 名模拟患者的生存 Seq2Seq 结果示例。通过测量距离平方的平均值来确定最优簇数，如图 4C 所示。如前所述，我们有意将非信息性临床测量添加到合成数据集中。RSM 可以成功识别那些非信息性临床测量，并通过分配较小的权重将其从其他特征中排除。我们认为 id 为 16 到 20 的特征是非信息性的，其中 RSM 成功地识别了其中 4 个最不重要的特征。

尽管在深度生存模型中添加了解释功能，但一些限制可能会影响 RSM 的性能。RSM 使用深度模型进行生存分析。因此，深度生存模型的精度直接影响到 RSM 的性能。RSM 不能正确地解释深度生存的结果，该模型的预测也很差。RSM 的另一个限制是，它的整体性能受到几个解释单元(如相似性度量、聚类模型、L1 显著特征选择和统计测试)的单个性能的限制。可以用更高级的算法替换 RSM 的上述每个组件，以实现更高的 RSM 整体性能。例如，为了获得更好的聚类性能，可以使用基于深度无监督学习的聚类算法。升级 RSM 中使用的单个或所有单元是我们未来研究的主题。此外，识别 RSM 中的重要特征是基于训练队列的。随着医疗保健行业追求个性化服务，针对特定病例的重要特征识别可能更加可取。

5 结论

在这项研究中，我们提出了 RSM，这是一个框架，用于识别与感兴趣的患者具有相似临床测量和结果的患者，从而验证深度生存模型的预测。我们将 RSM 应用于在合成数据集上训练的 survival Seq2Seq 模型的预测生存 pdf，以验证 RSM 识别重要临床测量和识别与给定数据样本最相似的数据样本的能力。在验证了这些功能之后，我们在 MIMIC-IV 上训练的生存 Seq2Seq 上测试了 RSM，以证明生存 Seq2Seq 的预测是正确的。据我们所知，这是第一次专门设计一个框架来解释经过训练的深度生存模型的输出。

PU-193

基于边缘计算架构的医疗设备智慧物联管理平台的构建及应用

陆靓亮，金伟
南京医科大学附属无锡人民医院

1.引言

伴随着医疗服务需求的扩展与医学装备行业的飞速发展，大型医疗设备的配置使用量显著增加，并成为各级医疗机构必不可少的临床诊断与治疗工具[1]。大型医疗设备配置使用成本高、技术专业性强且复杂度高、运维管理难度大，是医疗成本控制与医疗风险防范的重要环节。此前，医疗机构对大型医疗设备的管理主要停留在资产管理，在维修和维保的方面也仅仅停留在人工记录的层面，缺乏信息化手段实现设备精细化管理[2]。但是，对于国内绝大多数医疗机构而言，大型医疗设备资产占据了医院总固定资产的大部分比率，并且其通常都是高科技产品，每一个类别的设备往往涉及到多学科融合，综合着多样的新型技术成果[3]。因此，为获取大型医疗设备的运行、使用的实时状态信息，为购置设备、绩效考核等提供决策依据，实现数据的有效共享，助力医院精细化发展[4]，探索建立医学装备互联互通的数字化平台，利用真实世界数据支持医学装备的专业化管理，是行业发展新趋势。

近年来，随着万物互联时代的快速到来，网络边缘的设备数量和产生的数据都在快速增长。在此情形下，以云计算模型为核心的集中式处理模式将无法高效处理边缘设备产生的数据。尤其对于医院的医学装备而言，传统云计算有 4 个不足：实时性不够、数据安全和隐私无保障、能耗较大、带宽不足。尤其医学装备产生的数据与患者隐私和医疗行为极为紧密，设备运行数据传输到云端，会涉及到复杂的医疗信息安全监管风险。因此，需要引入面向边缘设备所产生数据计算的边缘计算模型与医学装备物联网平台的架构设计，确保数据不出院的前提下，实现医学装备实时运行数据的采集解析和数字化应用[5]。

基于以上原因，我们设计并构建了基于边缘计算架构的大型医疗设备智慧物联管理平台，利用物联网与数字化的技术手段[6]，实现了实时、客观、可持续的各节点数据支撑，可量化可追溯的使用质量评价体系，并为建立归一化的医疗大数据平台提供系统支撑。

平台架构及其特性

1.1. 边缘计算技术特性

边缘计算是在网络边缘执行计算的一种新型计算模型，边缘计算操作的对象包括来自于云服务的下行数据和来自于万物互联服务的上行数据，而边缘计算的边缘是指从数据源到云计算中心路径之间的任意计算和网络资源[5]。边缘计算模型和云计算模型二者相互协同融合，二者融合也是行业数字化转型发展的必然趋势[7]。

边缘计算模型具备三个明显的优点[5]：第一，可以避免全部上传云端，直接在网络边缘处理临时数据，减轻网络带宽和数据中心功耗的压力；第二，不通过网络请求云计算中心的响应，直接在靠近数据生产者处处理数据，整体上提升了服务响应效率；第三，边缘计算将用户数据存储在网络边缘设备上，不仅确保了用户数据的安全性和隐私性，同时也兼顾了云端核心算法的快速迭代和知识产权保护。

1.2. 平台架构设计

平台利用先进的边缘计算架构设计[8]，在确保医院数据不出院的前提下，算法模型迭代更新实现设备数据的动态采集和解析，构建医院自主可控的医疗设备数字化平台。平台架构图如图 1 所示，其工作流程如下：

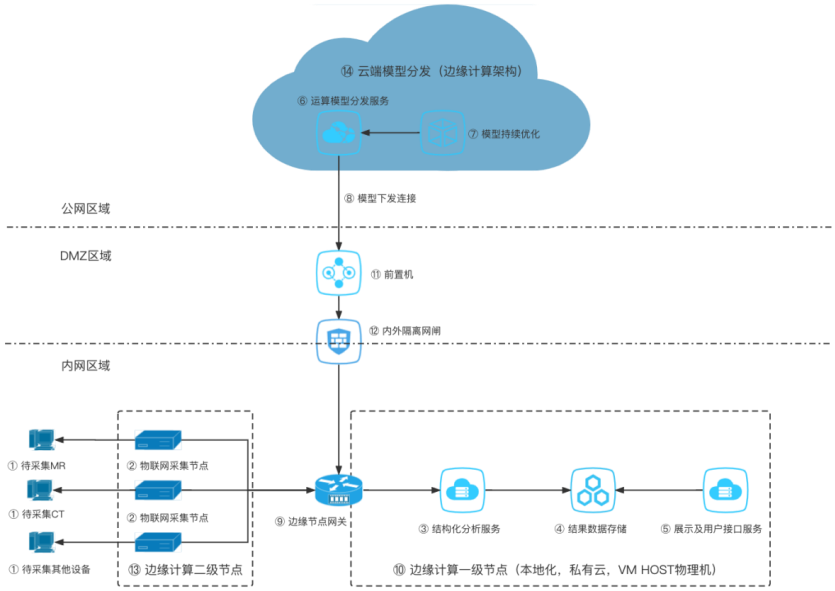


图 1 医疗设备智慧物联管理平台架构

物联网采集节点②，从待采集设备①获取设备日志，发送到结构化分析服务③，分析后数据推送到结果存储④，由用户接口服务⑤，向终端用户提供展示。

结构化分析服务在分析前，通过医院预设安全网闸⑫及架设于医院机房 DMZ 区域的前置机⑪发出模型请求连接⑧，从云端获取运算模型更新。从安全设计角度上讲，该模型下发属边缘端 pull 方式，不进行主动 push，即边缘端不监听公网 IP 及任何端口。

模型持续优化单元⑦，通过公有计算服务的统计结果，持续的对运算模型进行优化；当下一次⑧建立连接，本地结构化分析服务③将更新运算模型，提供更加精确和可靠的分析结果。系统前置机需求的公网连接，仅用于运算模型下载+更新，以及服务器/物联网采集节点的工况心跳（在线/离线，正常/异常）。

边缘计算一级节点⑩基于 INTEL/AMD X86-64 架构的机架式服务器，搭建服务器集群部署于内网，部署医疗设备的日志解析核心算法、数据库及应用展示系统。系统架构设计搭配基于 Linux KVM 稳定可靠的硬件虚拟化解决方案，为整套系统多个服务间逻辑隔离且稳定交互打下坚实基础，实现构建不超过 64 台物理服务器的集群结构，以及相应的 cephFS 共享存储服务。系统架构支持基于 iptables 的 IPv4 和 IPv6 防火墙设定，支持 Linux 磁盘分区 256bit AES 加密，有效保障业务数据不受窃取或篡改。云端运算模型迭代更新，边缘计算一级节点⑩从公网自动下载算法包更新。

1.3. 边缘计算架构特性

原本由云端运算集群完成的计算任务，现今由本地化部署的数据结构化分析节点来完成。

计算、存储、应用、通信等组件全部本地化部署，提供更快的网络响应，全程数据不出院，确保信息安全。

计算所需的数据分析模型由本地节点向云端请求，云端响应请求后，针对分析节点的工作目标进行动态下发。

由于云端对数千台设备提供 7×24 小时公网物联网服务，因而针对特定品牌和机型的被分析设备，其数据计算模型可以不断动态更新，迭代改进。院内版边缘计算架构的核心服务——结构化分析服务，也可在运行过程中从云端动态更新运算模型，从而做到在不间断服务过程中，让系统服务能力、准确性、及时性持续提升。

2. 平台关键技术设计

2.1. 大型设备物联网采集节点特性

该平台主要设计思路是，基于大型医疗设备均已接入院内网络系统的前提下，设计医疗设备日志采集器作为物联网采集节点，基于网口的日志采集与解析方式，通过大型医疗设备的网络接口进行大型医疗设备的实时运行日志数据的提取，并将采集到的日志数据通过有线网络传输至中心服务器中进行深度解析，结果数据及应用服务通过终端显示。

国内外大型医疗设备的数据采集尚无统一方式，目前主要用到的技术有电源采集，截屏采集，日志采集。借助于设备电源信号进行采集或设备显示器图像信号进行采集的方式均存在局限性，这些方式技术成熟度较高，但采集信号单一片面，使得数据采集后对大型医疗设备的实时监测与评价管理的应用意义有限；且在设备端直接接入采集器模块，对设备的常规运行使用及维护产生干扰风险。日志采集通过设备网口直接采集各品类品牌大型医疗设备运行日志，以实现对设备运行动态的全数字化还原，具有较高的技术领先性与广泛的应用前景。

日志采集器采用嵌入式芯片设计，ARM 处理器，定制 Linux 系统，在物联网采集节点实现日志源文件的结构化预处理。结构化预处理后的原始日志数据通过握手协议向边缘端核心服务器传输，再将日志数据进行解析、翻译与逻辑运算，最终得到结果数据并存入数据库。

2.2. 医疗设备数据中台设计

采用 MySQL 结构数据库实现对所有接入品类设备的结果数据存储、分发和用户接口服务，只监听必要网卡，SSL 通信，使得集群外无法对数据库进行嗅探和攻击；多个数据库节点采用主从结构，确保备份机制可靠运行。该服务同样支持存储加密、BFDP 和 iptables 防火墙保护。多个数据库节点采用主从结构，确保备份机制可靠运行。根据医院信息化建设需求，提供基于 HTTP (application/json)的第三方服务接口，赋能医疗机构设备大数据中心成为真正的全院数据中台。

2.3. 医疗设备管理平台设计

基于可靠且可扩展的前后端分离式设计，在 HTML5 页面基础上，架设业界领先的 js 封装数据可视化组件，将设备数据解析结果动态且实时的展示在任何用户终端浏览器界面（包含但不限于工作站、一体机、平板电脑、手机等），并将提取的目标设备风险故障预警和运营总结等关键信息以确定时效主动推送到用户终端。

3. 应用情况

3.1. 医疗设备管理平台主界面

平台主要有涵盖全院大型医疗设备的资产台账、运维预警、运行动态、调拨管理、使用效能评价、质控评价等功能，多维度自定义呈现设备实时动态的运行使用情况，全院近半资产动态清晰呈现，医疗设备管理核心指标智能分析，基于真实世界数据的量化分析指标有效指导医院管理决策，实现医疗设备智慧化管理升级。如图 2 所示。

图 2 医疗设备管理系统主界面

3.2. 医疗设备管理平台驾驶舱模块

该模块在一体化的界面上实现设备使用动态界面、运维动态展示界面的实时呈现和交互，根据医院设备管理的经验和需求，利用可视化数据聚合技术，定制化医疗设备动态运行管理中央平台。如图 3 所示。

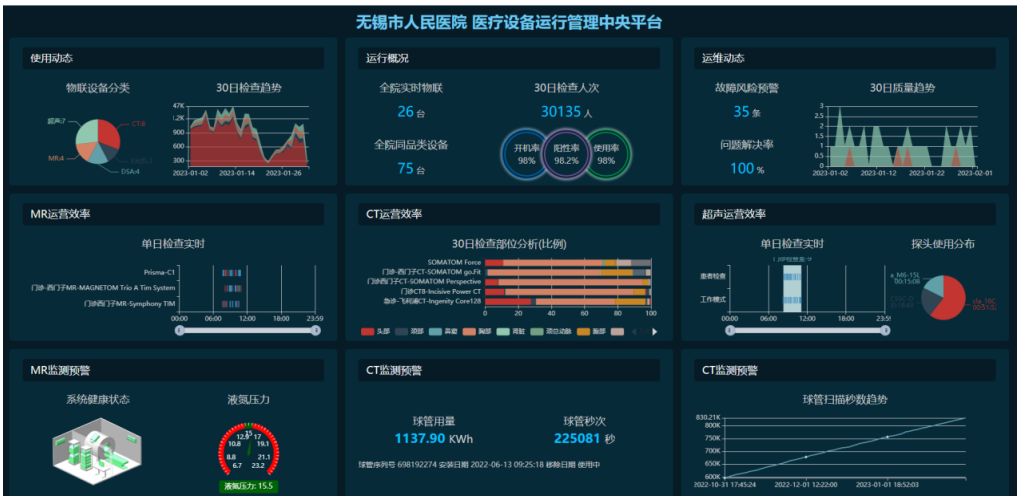


图 3 医疗设备管理平台驾驶舱模块

3.3. 实现各类设备使用全流程的数据闭环

该平台基于真实世界数据还原设备运行使用动态，并实现设备使用全流程的数据闭环，极大的提升了大型医用设备量化使用评价与精细管理能力，有效解决了不同品牌、型号大型医用设备的日志数据接口标准化问题，尽可能的利用设备厂家的通用端口及协议，设计物联网采集节点的硬件与软件技术，采集并解析设备运行日志。有效解决了大型医用设备实时巡检、质量安全、绩效管理、资产效益评价等问题。

平台正式运行 6 个月中，共接了 MR4 台，CT8 台，DSA4 台，X 光机 3 台，超声 7 台，数据采集频率每 15 分钟一次，数据准确率高达 97%。

3.4. 提高了大型医用设备运维效率

该平台大大提高了大型医用设备运维效率。以磁共振设备为例，原有的磁共振运维依赖于人工巡检，对核心部件参数如液氦压力、液氦水平等依赖人工记录，耗费人工且时效性低，风险故障的排除存在不及时的问题。通过实时监测与评价管理应用平台，对各台磁共振设备的开关机状态、运行状态、风险故障、核心部件参数均一目了然，随时调阅，超过阈值或产生高风险故障预警实时推送，大大缩短了维修响应时间，准确的风险预警消息对故障处理措施、维修前的准备工作均有显著的指导意义，显著提高故障维修效率。

平台正式运行 6 个月中，磁共振设备的平均故障维修响应时间由 52.1min 缩短至 23.8min，设备开机率从 94.3%提高至 96.8%，大型维修量由 33 次减少到 21 次，减少了 36%，保养校准监测各项指标均通过，同时临床科室反馈显著减少了设备宕机时间。

4. 结论

基于边缘计算架构建立的医疗设备管理系统，实现了不同品牌不同型号大型医疗设备系统底层互联互通，建立了标准化的日志解析数据库，并实现了对大型医疗设备核心参数实时监测、风险故障预警、使用动态分析、使用效能评价、远程协同等应用功能，驱动大型医疗设备的运维使用降本增效，实现了智慧医院管理的升级。

PU-194

通过转化基因组学确定多发性硬化易感性的蛋白靶标

王茜 王学琴 吕毅康 韩子怡 贺培凤*
山西医科大学

引言

多发性硬化症（MS）是中枢神经系统（CNS）的一种慢性炎症性疾[1]，其特征是中枢神经系统的白质和灰质均出现大面积融合性脱髓鞘病变[2]。多发性硬化症的发病机制被认为是对自体蛋白免疫耐受的丧失，这可能是遗传易感性和环境因素共同作用的结果。这可能会导致自身反应性 T 细胞和 B 细胞的产生，从而导致中枢神经系统的慢性炎症和脱髓鞘病变。[3]尽管过去我们对 MS 发病机理的认识取得了进展，并开发出了各种治疗药物，但目前的治疗方法并不能完全有效地控制疾病。因此，有必要确定多发性硬化症的关键分子特征，并探索基因变异在治疗中的作用。确定新的药物靶点，为患有这种疾病的患者开发出更加个性化的治疗方法。本研究我们旨在通过整合来自血液和基因数据的高通量蛋白质组学数据，进行全蛋白质组关联研究（PWAS）。目的是确定与基因组结构相关的蛋白质水平，为该疾病的潜在治疗靶点提供新的见解。

方法

MS GWAS 数据

本研究使用的 MS 汇总数据集来自国际 MS 遗传学联合会（International MS Genetics Consortium）的最新出版物，其中包括 47,429 例 MS 病例和 68,374 例对照，仅限于欧洲参与者[4]。

2.2 人类血浆 pQTL 数据

人类血浆蛋白质组是通过对社区人群动脉粥样硬化风险（ARIC）研究中的欧洲裔美国人（EA）进行 cis-pQTL 分析而生成的，总样本量为 n = 7,213 [5]。分析了顺式调节区域中欧洲血统的 6,181,856 个常见单核苷酸多态性 (SNPs)（次等位基因频率 (MAF) > 1%），并根据性别、年龄、研究地点和十个遗传主成分 (PCs) 进行了调整。

2.3 全蛋白质组关联研究

该研究基于 FUSION 来进行全蛋白质组关联分析（PWAS），其中涉及血浆蛋白质表达水平中的 cis-pQTL 和 GWAS 结果中的 SNPs 分析[6]。该方法利用连锁不平衡（LD）参考面板来减轻 LD 对计算的测试统计数据的影响。然后，估算每个基因基于 SNP 的遗传率，利用 FUSION 计算 SNP 对具有显著遗传率（即遗传率 p 值小于 0.01）的蛋白质丰度的影响。此外，还使用 Benjamini-Hochberg (BH) 方法计算了调整 FDR 的 P 值。

2.4 MR 分析

为了进一步验证 PWAS 确定的顺式调控血浆蛋白是否与 MS 存在因果关系，我们采用了 MR 分析，将与蛋白表达（cis-pQTL）密切相关的变异作为工具变量（IV）。研究人员选择了全基因组 $p<1\times10^{-5}$ SNPs 作为 IV，并将连锁不平衡聚类阈值规定为 $r^2<0.05$ 时，蛋白质水平被认为介导了 SNP 对表型的影响。

2.5 贝叶斯共定位分析

贝叶斯共定位分析可用于评估工具变量假说的有效性，该假说考虑了两个（或多个）因果变异体之间的重叠特征。在本研究中，贝叶斯共定位分析用于评估通过 MR 和 MS 鉴定的蛋白质共享相同单核苷酸变异的概率[7]。我们使用了 COLOC 先验概率，$p_1 = 1 \times 10^{-4}$、$p_2 = 1 \times 10^{-4}$和 $p_{12} = 1 \times 10^{-5}$，其中 p_1表示与 MS 相关的 SNP，p_2表示与重要 cis-pQTL 相关的 SNP，p_{12}表示

与两种性状相关的 SNP。这一过程的结果主要通过后验概率（PPH4）量化支持，PPH4 在 0.75 或以上被定义为共定位的有力证据[8]。

2.6 药物基因相互作用分析

为了深入研究通过贝叶斯共定位分析鉴定的基因编辑血浆蛋白在疾病药物治疗中的潜力，我们探索了基因 - 药物相互作用。通过将编码风险血浆蛋白的基因输入药物-基因相互作用数据库（DGIdb），我们搜索了可用的药物或化合物以及潜在的可药用基因。

结果

PWAS 使用血浆 cis-pQTL 鉴定了 25 种与 MS 相关的候选血浆蛋白

该研究利用 FUSION 开展了 MS 的 PWAS，旨在通过整合 cis-pQTL 和 GWAS 结果，确定其丰度与 MS 遗传风险相关的蛋白质。 PWAS 鉴定出 25 个统计显着的顺式调节血浆蛋白丰度与 MS 相关，错误发现率为 P<0.05 ，与 MS 最显着相关的前 10 个顺式调节血浆蛋白是顺式调节血浆蛋白是激活转录因子 6 β (ATF6B)，补体 C2 (C2)、同种异体移植物炎症因子 1 (AIF1)、主要组织相容性复合物 II 类、DQ α 2 (HLA-DQA2)、TNXB、补体因子 B (CFB)、天然细胞毒性触发受体 3 (NCR3) ）、MHC I 类多肽相关序列 B（MICB）、胸苷磷酸化酶（TYMP）、丝裂原激活蛋白激酶 3（MAPK3）（图 1 和表 1）。

MR 揭示了 7 种与 MS 相关的血浆蛋白

应用血浆 cis-pQTL 数据集和 MS GWAS 的 MR 分析，进一步评估 MS PWAS 显着的顺式调节血浆蛋白丰度是否与 MS 存在因果相关，并进一步鉴定候选靶向锚定血浆蛋白，为 PWAS 确定的血浆蛋白提供验证。 MR 结果表明七种血浆蛋白 [PLEK、TNXB、CASP3、CD59 、CR1、TAPBPL、ATXN3] 与 MS 有很强的因果关系，错误发现率为 P<0.05（图 2 图 3）。基因预测较高的 TNXB（OR：0.7150，95% CI：0.5939 至 0.8607，P = 0.0004）和 CD59（OR：0.8705，95% CI：0.8009 - 0.9461，P = 0.0011） 与较低的 MS 风险相关。较高的 PLEK（OR：1.4037，95% CI：1.0752-1.8325，P = 0.0127）、CASP3（OR：1.1191，95% CI：1.0491 - 1.1939，P = 0.0006）、CR1（OR：1.1353，95% CI：1.0649 - 1. 2102，P = 0.0001）、TAPBPL（OR：1.0561，95% CI：1.0223 - 1.0910，P = 0.0010）和 ATXN3（OR：1.0803，95% CI：1.0322 - 1.1304，P = 0.0009）水平与较高的 MS 风险。

MS 风险血浆蛋白与血液中 cis-pQTL 之间的共定位

如果一个变异体与蛋白质表达（即 cis-pQTL）和 MS 相关，则可能会导致 MS 的 PWAS 中存在关联，这被解释为与因果关系一致。在先前的 MR 中，通过统计共定位分析对符合 FDR 阈值的基因进行了研究，该分析报告了 GWAS 和 cis-pQTL 共享一个因果变异体的概率。该分析显示，7 个基因种有 3 个（PLEK、CR1、CD59）提供了共定位证据，表明这 3 个蛋白在 MS 的病理生理学中起着重要作用（PLEK：PPH4= 0.817；CR1：PPH4= 0.837；CD59：PPH4=0.817）（图 4）。

药物与基因的相互作用

我们试图确定针对风险相关蛋白质的药物再利用的潜在机会，并寻找潜在的可药用基因。为了实现这一目标，我们通过 DGIdb 查询了贝叶斯共定位分析 (n=3) 检测到的所有 MS 显着基因与处方药物的相互作用。从 DGIdb 中仅获得了一种药物，即依库珠单抗，它被批准为靶向 CR1 的免疫

抑制剂。此外，PLEK、CR1、CD59 可以作为可成药的靶基因，表明这三个基因具有很强的开发潜力，适合作为药物靶点，但目前尚无针对它们的药物。

讨论

我们研究的目的是鉴定导致 MS 发病机制的血浆蛋白，从而发现治疗该疾病的潜在新靶点。通过分析，我们确定了 25 种可通过调节血浆蛋白丰度而增加 MS 发病风险的蛋白质。这些蛋白质是开发 MS 新疗法的潜在靶点。其中 7 种蛋白（PLEK、TNXB、 CASP3、CD59、CR1、TAPBPL、ATXN3 ）在 MR 验证分析中得到验证。最后通过贝叶斯共定位分析对 PLEK、CD59 和 CR1 进行鉴定，并通过 DGIdb 进行分析，发现了与 CR1 相互作用的靶标药物依库珠单抗。识别疾病的治疗靶点并寻找靶向药物是人类遗传学研究的关键目标。

此前，Jia 等人研究了脑蛋白，发现 SHMT1、FAM120B 和 ICA1L 的失调可能会导致 MS 风险[10]。与其不同的是我们整合了外周血蛋白质组学和基因数据，为研究新型 MS 靶向药物拓展了新的蛋白质位点。血浆蛋白可以捕获活跃于不同生物过程的各种蛋白质，更有利于发现新的药物靶点。并且可以克服脑蛋白相关药物穿越血脑屏障的问题，因此更具临床转化价值。

研究发现了与 MS 相关的蛋白质，包括 PLEK、CD59、CR1、CASP3、ATXN3 和 TAPBPL。除了这些已知的 MS 蛋白质，该研究还发现了新的候选蛋白，例如 TNXB，它们可能在 MS 的发生和进展中发挥作用。这些发现可能对理解 MS 的潜在机制和确定潜在的治疗靶点具有重要意义。PLEK 是血小板、单核细胞中蛋白激酶 C (PKC) 家族的重要底物。 PKC 的胞内酶在单核吞噬细胞的各种细胞反应的信号传导中发挥着至关重要的作用。 PLEK 已被证明是其他炎症性疾病的易感位点[11]。此外，研究发现 CR1 和 CD59 均通过补体途径对 MS 产生影响。它在 MS 患者灰质和白质中的失调可能在疾病的不同阶段发挥作[12]。 CD59 是 MS 的保护因子，而 CR1 是 MS 的危险因子。 CR1 作为宿主细胞细胞膜上的补体调节因子取代经典途径中 C3 转化酶的一种成分，CR1 与多种自身免疫性疾病的发病和发展有关联。这些发现凸显了 CR1 的多面性及其作为一系列自身免疫疾病治疗靶点的潜力。通过 DGIdb 分析确定了与 CR1 相互作用的靶向药物依库珠单抗 。依库珠单抗目前用于治疗溶血性阵发性睡眠性血红蛋白尿症。依库珠单抗反应变异性的决定因素可能是可调节补体激活的因子水平，特别是 CR1 的多态性等位基。依库珠单抗可以在药理学上重新定位为治疗 MS 的新药。

新基因 TNXB 的鉴定可能会导致对 MS 新蛋白质靶点的探索。目前，TNXB 已被确定为 Ehlers-Danlos 综合征和膀胱输尿管反流病的多效应基因。在一些进行性 MS 患者中，可以观察到显着的突触丧失和中枢神经系统萎[14]。因此，TNXB 可能通过突触途径介导 MS 的发病机制。

我们的研究有几个优点。首先，这是第一项使用最大、最全面的人类血浆蛋白质组和最新的 MS GWAS 汇总数据进行 PWAS 分析的研究。为了尽量减少假阳性结果并提高结果的可信度，我们采用了独立的 MR 和基于贝叶斯的共定位方法。然而，本研究存在一些局限性。我们的研究专门调查了血液中的候选蛋白质，因此未来的研究也可能受益于探索其他组织。其次，我们的分析仅限于欧洲血统。我们所描述的结果是否与其他人群一致还有待研究。

结论

根据我们目前的蛋白质组学研究结果，PLEK、CR1 和 CD59 有希望成为治疗 MS 的新型治疗靶点。然而，需要进一步的研究来确定这些蛋白质表达的内源调节因子。这发现这些调节因子的药理诱导剂或抑制剂有可能开发出治疗 MS 的新方法。此类研究可能对未来治疗这种复杂而又使人衰弱的疾病产生重要影响。

PU-195

代谢相关基因作为预后生物标志物的泛癌分析

黄泰[1,2]，刘格良[1,2]，莫状[1,2]，陈浩然[1,3]，王灿[1,3]，于琦[1,3]，贺培风[1,3*]

1. 山西医科大学 临床决策研究大数据山西省重点实验室

2.山西医科大学 基础医学院

3. 山西医科大学 管理学院

引言

癌症作为一种极具复杂性的疾病，在全球范围内的不断增加，成为医学界的重要挑战。它的发生和发展受到多种遗传、环境和代谢因素的影响。代谢作为维持生命活动的基础过程之一，与癌症密切相关。越来越多的研究表明，代谢紊乱是癌症的一个重要特征，且与患者的预后密切相关[1]。代谢特征因此而成为癌症研究的备受关注的热点领域。

然而，代谢通路的复杂性和互相交错，传统的代谢分析方法难以全面解析代谢通路的作用机制[2]。此外，不同类型的癌症研究面临着样本量小、数据维度高、信息获取难等问题[3]。因此，如何从高通量代谢组学数据中准确鉴别与癌症相关的代谢特征，并构建有效的癌症预后模型，是当前代谢组学领域研究的重要问题。

在这一背景下，借助生物信息学技术与代谢特征的综合运用，我们可以进行泛癌预后分析，为临床治疗提供重要的参考和指导。通过分析癌症患者的代谢组学数据，鉴别出与预后相关的代谢特征，并利用机器学习方法构建泛癌预后模型，为临床医生提供更准确的癌症患者预后评估和治疗建议，如 Huang 等应用优化的铁粒子辅助激光解吸/电离质谱（LDI MS）方法对来自早期肺腺癌患者、健康对照、其他肺癌以及肺良性疾病的人血清样本进行代谢组学分析，通过机器学习模型算法，确定了可以区分早期 LA 和健康对照的 7 种代谢物[4]。

泛癌研究已成为癌症领域的关键方向，对于改进癌症治疗策略和揭示癌症机制至关重要[5]。同时，癌症预后评估在制定个性化治疗方案和提高患者生存质量方面具备重大意义。然而，现行的癌症预后评估方法存在一定准确性和可靠性方面的挑战[6]。泛癌分析代表了一种全面的生物信息学方法，为改进癌症预后评估提供了新途径，但构建基于多代谢特征组合的预后模型具有一定复杂性。

因此，本研究旨在应用泛癌分析方法，构建多代谢特征组合的癌症预后模型，旨在提高预后评估的准确性和可靠性，从而为个性化治疗方案的制定提供更为可靠的依据。这项研究具有重要的临床应用前景，有望在改善癌症预后评估和治疗方案制定方面取得新的重要进展。同时，它也将有助于更深入地理解癌症的分子机制和生物标志物的发现，为未来研究提供新的思路和方法。

方法

1 癌症 RNA-seq 数据集和处理

本研究的全癌 RNA 测序（RNA-seq）数据源自癌症基因组图谱（The Cancer Genome Atlas, TCGA）数据库。该数据集包括来自 33 种不同癌症类型的共计 10551 个样本。癌症的类型和样本数量详见表 1。

基于 KEGGREST R 包获取 KEGG 数据库全部的代谢相关基因列表。

表 1 TCGA 数据库获取 RNA-seq 数据样本数量

Table1 Number of RNA seq data samples obtained from TCGA database

TCGA	Detail	Tumor	Normal
ACC	Adrenocortical carcinoma	79	-
BLCA	Bladder Urothelial Carcinoma	403	19
BRCA	Breast invasive carcinoma	1081	99
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma	296	3
CHOL	Cholangio carcinoma	35	9
COAD	Colon adenocarcinoma	455	41
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma	47	-
ESCA	Esophageal carcinoma	173	13
GBM	Glioblastoma multiforme	145	5
HNSC	Head and Neck squamous cell carcinoma	515	44
KICH	Kidney Chromophobe	66	25
KIRC	Kidney renal clear cell carcinoma	529	72
KIRP	Kidney renal papillary cell carcinoma	289	32
LAML	Acute Myeloid Leukemia	137	-
LGG	Brain Lower Grade Glioma	505	-
LIHC	Liver hepatocellular carcinoma	369	50
LUAD	Lung adenocarcinoma	513	58
LUSC	Lung squamous cell carcinoma	496	51
MESO	Mesothelioma	82	-
OV	Ovarian serous cystadenocarcinoma	401	-
PAAD	Pancreatic adenocarcinoma	178	4
PCPG	Pheochromocytoma and Paraganglioma	176	3
PRAD	Prostate adenocarcinoma	483	51
READ	Rectum adenocarcinoma	163	10
SARC	Sarcoma	258	2
SKCM	Skin Cutaneous Melanoma	103	1
STAD	Stomach adenocarcinoma	410	36
TGCT	Testicular Germ Cell Tumors	149	-
THCA	Thyroid carcinoma	500	57
THYM	Thymoma	120	2
UCEC	Uterine Corpus Endometrial Carcinoma	539	35
UCS	Uterine Carcinosarcoma	57	-
UVM	Uveal Melanoma	77	-

注：-代表无数据

Note: - represents no data

2 差异分析

基于下载的表达谱数据，本研究使用 R 语言 DESeq2 包[16]对所有数据集进行筛选。对于 TCGA-LIHC 数据，DEGs 筛选条件设定为差异倍数大于或等于 2，校正后的 *P* 值小于 0.05。

3 富集分析

基于对应的 DEGs，本研究使用 clusterProfiler、org.Hs.eg.db、GseaVis 等 R 包[7]对 DEGs 进行京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes，KEGG）通路富集分析和基因集富集分析（Gene Set Enrichment Analysis, GSEA），并对聚类结果进行可视化处理。

4 生存相关基因的筛选

Cox 比例风险模型（Cox Proportional-Hazards Model），通常简称为 Cox 模型，是一种用于生存分析的统计方法。在本研究中，生存相关基因的筛选通过 survminer、glmnet、randomForestSRC R 包[8]完成。为了进一步验证随机生存森林模型的有效性，研究者采用了 survminer R 包中的 surv_cutpoint 函数。用于确定在生存数据中的最佳分割点，以便将癌症患者样本分成高代谢特征组和低代谢特征组，从而更好地评估模型的性能。Kaplan-Meier（KM）生存分析[9]用于分析基因与患者预后的关系。为了可视化相关的预测模型，使用 forestplot R 包绘制了森林图。

5 生存相关基因的评估

生存相关基因的评估通过 pysurvival[10] 、sklearn[11] python 包完成。本研究采用 timeROC[12] R 包对验证集进行了受试者工作特征（Receiver Operator Characteristic，ROC）曲线的绘制。ROC 曲线是一种用于评估关键基因的诊断性能的工具。它通过在不同的敏感性和特异性组合下绘制曲线，帮助我们理解在不同阈值下模型的表现。在 ROC 曲线下方的面积（Area Under Curve，AUC）是一个重要的指标，它代表了模型诊断性能的总体评估。AUC 值越接近 1，表示模型的性能越好，具有更高的诊断准确性。

统计分析

所有统计分析都是使用 R v4.2.1 进行的。使用 Kaplan-Meier 曲线和对数排名测试评估了两组之间的生存差异。使用单变量和多变量 Cox 回归分析来确定预后因素。对于相关性分析，使用 Pearson 计算正态分布数据，使用 Spearman 计算非正态分布数据的相关系数。为了分析两组数据之间的差异，分别对正常分布变量和非正常分布变量使用了未配对的 Student's t-test 和 Mann-Whitney U-test。调整后的 p 值（FDR）是通过 Benjamini-Hochberg 校正法计算的。误差条表示均值的标准差。除非另有说明，否则 $P< 0.05$ 被认为具有统计学意义。

结果

1 肝癌中代谢相关基因的表达分析

本研究首先获得了 84 条通路 1726 个与代谢相关的基因列表，这些基因来源于 KEGG 数据库。然后，研究人员将这些基因与肝癌样本中的基因表达数据进行交叉分析，得到了在肝癌中表达的代谢相关基因列表总 1680 个。在这项研究中，我们选取了 369 个肝癌肿瘤样本和 50 个相应的癌旁组织样本作为研究对象，以了解这些基因的表达情况。

分析结果表明，在这些代谢相关基因中，有 0.74%（1244/1670）的基因在肝癌中表现出显著的差异表达（调整 *P* 值小于 0.05），如图 1A 所示。这意味着在肝癌样本中，大多数代谢相关基因

的表达水平与正常组织相比存在明显的变化。此外，我们也观察到一小部分代谢相关基因在肝癌中发生了显著的表达变化，这可能表明这些基因在肝癌的发生和发展过程中具有重要的功能和调节作用。

2 差异基因富集分析

基于差异分析的结果，我们筛选出了 308 个上调基因和 164 个下调基因，然后进行了 KEGG 富集分析。如图 1C 所示，我们展示了富集通路的前 20 位结果，而图 1D 则呈现了与代谢相关的 KEGG 富集通路的前 20 位结果。这些富集通路结果为我们提供了有关代谢在肝癌中的重要性和功能的初步见解。

根据富集分析结果，一些显著的代谢通路凸显出在肝癌中的重要性，其中包括视黄醇代谢、花生四烯酸代谢、药物代谢-细胞色素 P450、脂肪酸降解、细胞色素 P450 对外源物质的代谢、缬氨酸、亮氨酸和异亮氨酸降解、色氨酸代谢、药物代谢-其他酶、甘油磷脂代谢以及糖酵解/葡萄糖异生等。这些结果更进一步强调了代谢异常在肝癌的发生和发展中的关键作用。

为了深入分析这些通路的生物学功能，我们进行了基因集富集分析（Gene Set Enrichment Analysis，GSEA），并确定了具有最高标准化富集分数（Normalized Enrichment Score，NES）绝对值的代谢相关通路。在这些通路中，酸分解代谢过程、氨基酸分解代谢过程、分子分解代谢过程、外源分解代谢过程、α-氨基酸代谢过程以及一元羧酸分解代谢过程被认为是在肝癌中最为重要的代谢通路。这些通路的激活或抑制可能在肝癌的发生、发展和代谢重塑中扮演着重要的角色。

Figure 1

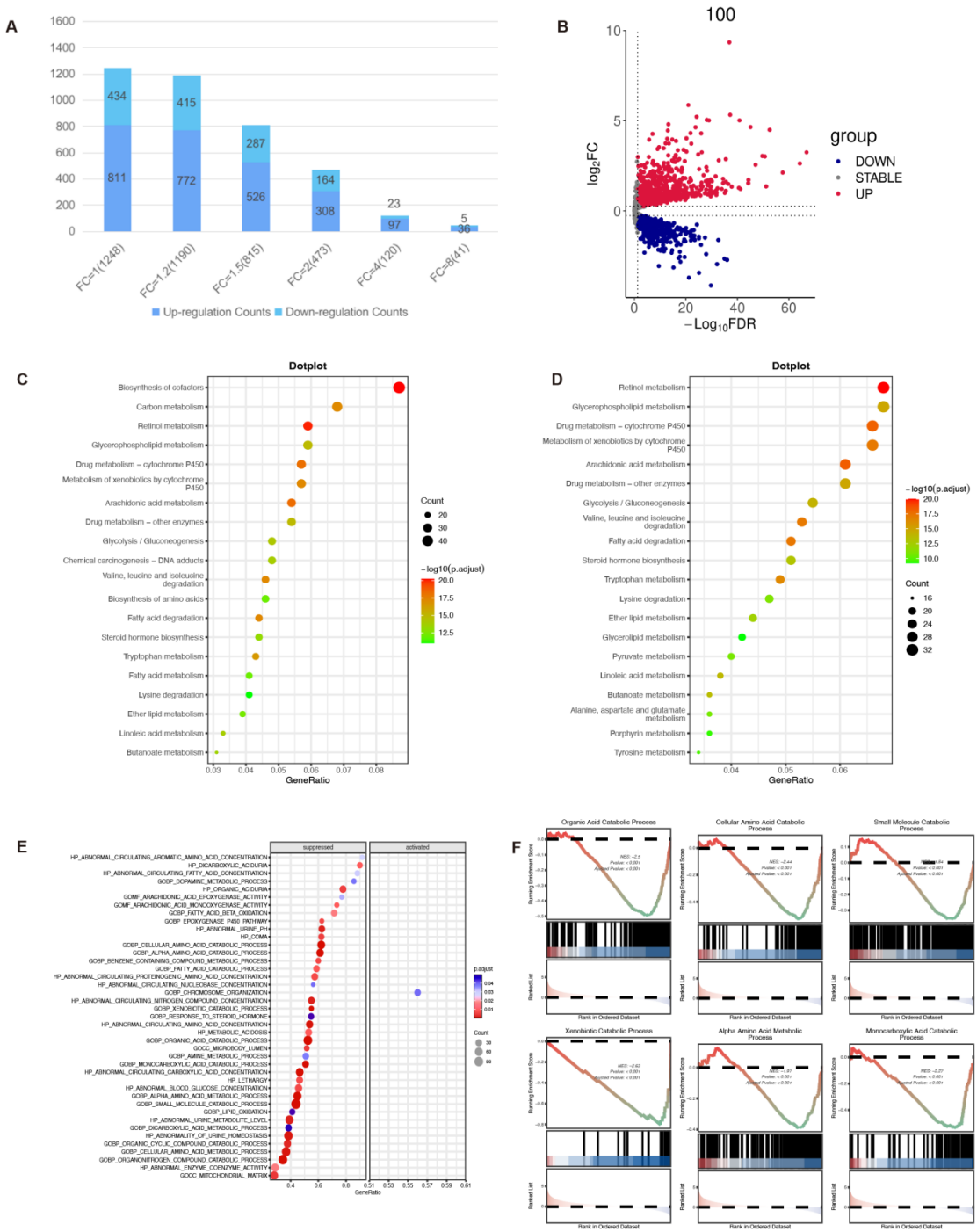

A
1600
1400
1200
1000
800
600
400
200
0
434
811
415
772
287
526
164
308
23
97
5
36
FC=1(1248)
FC=1.2(1190)
FC=1.5(615)
FC=2(473)
FC=4(120)
FC=8(41)
Up-regulation Counts
Down-regulation Counts
B
100
10
5
0
log2FC
0
20
40
60
−Log10FDR
group
DOWN
STABLE
UP
C
Dotplot
Biosynthesis of cofactors
Carbon metabolism
Retinol metabolism
Glycerophospholipid metabolism
Drug metabolism – cytochrome P450
Metabolism of xenobiotics by cytochrome P450
Arachidonic acid metabolism
Drug metabolism – other enzymes
Glycolysis / Gluconeogenesis
Chemical carcinogenesis – DNA adducts
Valine, leucine and isoleucine degradation
Biosynthesis of amino acids
Fatty acid degradation
Steroid hormone biosynthesis
Tryptophan metabolism
Fatty acid metabolism
Lysine degradation
Ether lipid metabolism
Linoleic acid metabolism
Butanoate metabolism
0.03
0.04
0.05
0.06
0.07
0.08
GeneRatio
Count
20
30
40
−log10(p.adjust)
20.0
17.5
15.0
12.5
D
Dotplot
Retinol metabolism
Glycerophospholipid metabolism
Drug metabolism – cytochrome P450
Metabolism of xenobiotics by cytochrome P450
Arachidonic acid metabolism
Drug metabolism – other enzymes
Glycolysis / Gluconeogenesis
Valine, leucine and isoleucine degradation
Fatty acid degradation
Steroid hormone biosynthesis
Tryptophan metabolism
Lysine degradation
Ether lipid metabolism
Glycerolipid metabolism
Pyruvate metabolism
Linoleic acid metabolism
Butanoate metabolism
Alanine, aspartate and glutamate metabolism
Porphyrin metabolism
Tyrosine metabolism
0.04
0.05
0.06
GeneRatio
−log10(p.adjust)
20.0
17.5
15.0
12.5
10.0
Count
16
20
24
28
32
E
suppressed
activated
HP_ABNORMAL_CIRCULATING_AROMATIC_AMINO_ACID_CONCENTRATION
HP_DICARBOXYLIC_ACIDURIA
HP_ABNORMAL_CIRCULATING_FATTY_ACID_CONCENTRATION
GOBP_DOPAMINE_METABOLIC_PROCESS
HP_ORGANIC_ACIDURIA
GOMF_ARACHIDONIC_ACID_EPOXYGENASE_ACTIVITY
GOMF_ARACHIDONIC_ACID_MONOOXYGENASE_ACTIVITY
GOBP_FATTY_ACID_BETA_OXIDATION
GOBP_EPOXYGENASE_P450_PATHWAY
HP_ABNORMAL_URINE_PH
HP_COMA
GOBP_CELLULAR_AMINO_ACID_CATABOLIC_PROCESS
GOBP_ALPHA_AMINO_ACID_CATABOLIC_PROCESS
GOBP_BENZENE_CONTAINING_COMPOUND_METABOLIC_PROCESS
GOBP_FATTY_ACID_CATABOLIC_PROCESS
HP_ABNORMAL_CIRCULATING_PROTEINOGENIC_AMINO_ACID_CONCENTRATION
HP_ABNORMAL_CIRCULATING_NUCLEOBASE_CONCENTRATION
GOBP_CHROMOSOME_ORGANIZATION
HP_ABNORMAL_CIRCULATING_NITROGEN_COMPOUND_CONCENTRATION
GOBP_XENOBIOTIC_CATABOLIC_PROCESS
GOBP_RESPONSE_TO_STEROID_HORMONE
HP_ABNORMAL_CIRCULATING_AMINO_ACID_CONCENTRATION
HP_METABOLIC_ACIDOSIS
GOBP_ORGANIC_ACID_CATABOLIC_PROCESS
GOCC_MICROBODY_LUMEN
GOBP_AMINE_METABOLIC_PROCESS
GOBP_MONOCARBOXYLIC_ACID_CATABOLIC_PROCESS
HP_ABNORMAL_CIRCULATING_CARBOXYLIC_ACID_CONCENTRATION
HP_LETHARGY
HP_ABNORMAL_BLOOD_GLUCOSE_CONCENTRATION
GOBP_ALPHA_AMINO_ACID_METABOLIC_PROCESS
GOBP_SMALL_MOLECULE_CATABOLIC_PROCESS
GOBP_LIPID_OXIDATION
HP_ABNORMAL_URINE_METABOLITE_LEVEL
GOBP_DICARBOXYLIC_ACID_METABOLIC_PROCESS
HP_ABNORMALITY_OF_URINE_HOMEOSTASIS
GOBP_ORGANIC_CYCLIC_COMPOUND_CATABOLIC_PROCESS
GOBP_CELLULAR_AMINO_ACID_METABOLIC_PROCESS
GOBP_ORGANONITROGEN_COMPOUND_CATABOLIC_PROCESS
HP_ABNORMAL_ENZYME_COENZYME_ACTIVITY
GOCC_MITOCHONDRIAL_MATRIX
0.4
0.6
0.8
0.51
0.53
0.55
0.57
0.59
0.61
GeneRatio
p.adjust
0.04
0.03
0.02
0.01
Count
30
60
90
F
Organic Acid Catabolic Process
Cellular Amino Acid Catabolic Process
Small Molecule Catabolic Process
Xenobiotic Catabolic Process
Alpha Amino Acid Metabolic Process
Monocarboxylic Acid Catabolic Process
Running Enrichment Score
Ranked List
Rank in Ordered Dataset

图 1

图 1 （A）肝癌中代谢相关基因的表达差异情况。（B）TCGA 肝癌数据火山图。（C）kegg 富集通路前 20 位。（D）代谢相关 kegg 富集通路前 20 位。（E）GSEA 富集通路前 20 位。（F）前 6 位 GSEA 富集通路点阵图

Figure 1 (A) Expression Discrepancy of Metabolism-Related Genes in Hepatocellular Carcinoma. (B) Volcano Plot of TCGA Hepatocellular Carcinoma Data. (C) The Top 20 KEGG Enriched Pathways. (D) The Top 20 KEGG Enriched Pathways Related to Metabolism. (E) Top 20 GSEA Enriched Pathways. (F) Top 6 GSEA Enriched Pathway Heatmaps.

3 单因素 Cox 分析和机器学习筛选

研究者对差异表达的代谢相关基因进行了预后相关性分析，采用了单因素 Cox 回归模型，获得了 561 个与预后相关的差异基因。同时，研究者运用 Lasso 回归和随机生存森林机器学习算法对 561 个差异基因进行筛选。随机生存森林算法中，获取了重要特征前 30 位（图 2A），并与 Lasso 模型的结果进行交集运算，最终得到了 7 个具有重要预测能力的代谢特征。建立 LASSO 模型时，通过交叉验证选择了平均误差最小的 λ 值（图 2B），得到了 15 个最重要的特征。

基于这 7 个代谢特征，研究者重新构建了随机生存森林模型，并使用该模型预测了 1-5 年生存状态的 AUC 值和 ROC 曲线（图 2CD）。结果显示，这些代谢特征在预测 1-5 年生存状态时表现出较高的 AUC 值，均超过了 0.89，说明模型具有良好的预测性能。此外，研究者还使用了 7 个代谢特征构建了多因素 Cox 模型，得到了这些特征的重要程度并以列线图形式表达（图 2E）。

为了进一步验证随机生存森林模型的有效性，研究者使用 survminer R 软件包中的 surv_cutpoint 函数确定了最优分割点，并将癌症患者样本分为高低代谢特征组，进行了 KM 生存分析。图 2F 展示了生存曲线，结果表明，基于这 7 个代谢特征构建的随机生存森林模型可以有效预测肝癌患者的生存风险（P<0.0001）。

Figure 2

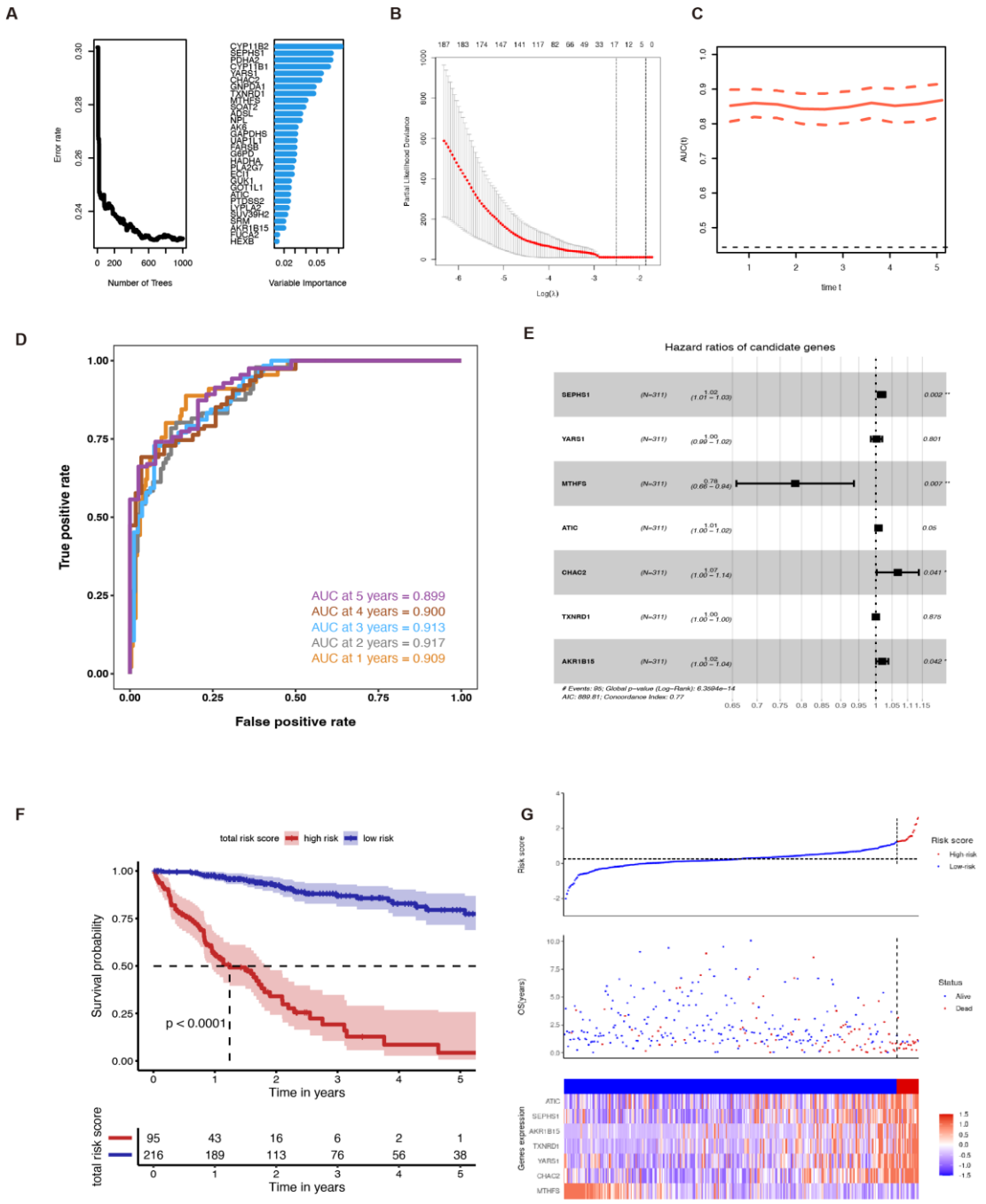


图 2 （A）随机生存森林前 30 位重要特征图 （B）LASSO 筛选重要特征图 （C）随机生存森林预测 1-5 年生存状态的 AUC 值 （D）随机生存森林预测 1-5 年生存状态的 ROC 曲线 （E）多因素 COX 回归模型森林图 （F）高低代谢特征组 KM 曲线 （G）风险曲线

Figure 2 (A) Top 30 Important Features in the Random Survival Forest. (B) LASSO Feature Selection. (C) AUC Values for Predicting 1-5 Year Survival Using the Random Survival Forest. (D) ROC Curves for Predicting 1-5 Year Survival Using the Random Survival Forest. (E) Forest Plot of the Multivariate COX Regression Model. (F) Kaplan-Meier Curves for High and Low Metabolism Feature Groups. (G) Risk Curves.

4 多机器学习算法验证 7 代谢特征价值

在本研究中，针对这 7 个代谢特征，我们采用了多种生存分析方法进行全面评估，包括标准的 Cox 比例风险模型（CoxPH）、非线性 CoxPH 模型（例如 DeepSurv/Nonlinear CoxPH）、线性多任务生存回归（MTLR）、神经多任务生存回归（神经 MTLR）、参数模型、条件生存森林、额外的生存树、随机生存森林、线性生存支持向量机（SVM）以及核生存支持向量机（SVM）。我们的分析结果见表 2，所有模型的 C-index 均超过 0.75，IBS 值均低于 0.25，这一结果表明这 7 个代谢特征在肝癌预后方面具备显著的预测能力。

表 2 7 代谢特征多机器学习算法检验

Table 2 Evaluation of the 7 Metabolic Features Using Multiple Machine Learning Algorithms

Model	C-index	IBS
Standard CoxPH	0.86	0.18
NonLinear CoxPH	0.87	0.18
Multi-Task Logistic Regression (MTLR)	0.76	0.12
Neural Multi-Task Logistic Regression model (N-MTLR)	0.78	0.12
Parametric models	0.86	0.17
Conditional Survival Forest	0.81	0.21
Extra Survival Trees	0.83	0.25
Random Survival Forest	0.83	0.19
Kernel SVM Model	0.83	0.13

5 代谢特征在泛癌中的预测价值

研究者首先对这 7 个代谢特征在多种癌症中的表达情况进行了 t-检验。检验结果显示，如图 3 所示，在大多数癌症中，这 7 个代谢特征的表达都呈现出显著的差异。其中，大部分代谢特征属于上调表达，只有极少数代谢特征呈现下调表达。接着，研究者采用了在泛癌中表现良好的随机生存森林算法来评估这 7 个代谢特征在泛癌中的预测价值。具体结果如表 3 所示。

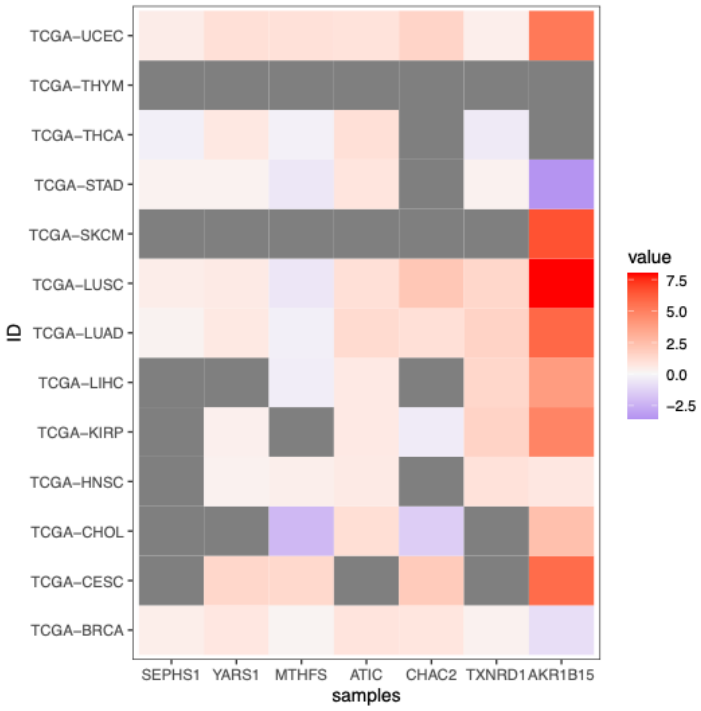


图 3 7 代谢特征在多种癌症表达情况和 t-检验结果

注：颜色越偏向红色代表上调表达越明显，越偏向紫色代表下调表达越明显，灰色代表未通过 t-检验

Figure 3 Expression and t-test Results of 7 Metabolic Features in Multiple Cancers

Note: The more the color tends towards red, the more significant the upregulation in expression; the more the color tends towards purple, the more significant the downregulation in expression; grey represents no significant difference based on the t-test

表 3 7 代谢特征在泛癌中的预测价值

Table 3 Predictive Value of 7 Metabolic Features in Pan-Cancer

project	C-index	IBS
TCGA-LAML	0.678276792	0.172243208
TCGA-LUAD	0.745822384	0.166253129
TCGA-KIRC	0.754246411	0.15347127
TCGA-LUSC	0.701354925	0.177928965
TCGA-LIHC	0.767348443	0.146933256
TCGA-MESO	0.5	0.183518171
TCGA-DLBC	0.5	0.129314857
TCGA-LGG	0.752793213	0.13524963
TCGA-ACC	0.5	0.076891376
TCGA-KICH	0.5	0.074923516
TCGA-PAAD	0.722592198	0.144884904
TCGA-BLCA	0.726143427	0.183862972
TCGA-STAD	0.740121217	0.194407285
TCGA-KIRP	0.840471128	0.094962475
TCGA-SKCM	0.621598317	0.169388035
TCGA-BRCA	0.786687343	0.095294196
TCGA-HNSC	0.746481146	0.178619454
TCGA-PRAD	0.998845777	0.027560805
TCGA-THYM	0.875474641	0.032383053
TCGA-PCPG	-1	-1
TCGA-UCEC	0.837883985	0.125960629
TCGA-ESCA	0.751144244	0.161051778
TCGA-THCA	0.994864187	0.007659777
TCGA-UVM	0.850790919	0.121165467
TCGA-CHOL	0.5	0.093973678
TCGA-OV	0.702578872	0.143858605
TCGA-CESC	0.836279354	0.131301303
TCGA-UCS	0.5	0.17259187
TCGA-GBM	0.664478812	0.07379792
TCGA-TGCT	1	4.42E-31
TCGA-COAD	0.793888126	0.158656581

讨论

基于肝癌中代谢基因的表达情况以及这些基因在其他癌症中的表达情况，可以得到如下一些思考。在肝癌和其他多种癌症类型中，代谢相关基因的表达存在明显的变化。这提示代谢的异常在不同类型的癌症中都可能扮演着重要角色。这种代谢的异常通常包括能量代谢、葡萄糖代谢、脂质代谢等方面的改变。因此，代谢的异常可能是癌症的一个普遍特征，而不仅仅局限于某一种癌症类型。同时，不同癌症类型中的代谢基因表达差异明显，这可能反映了不同亚型的存在。这些亚型可能在代谢特征上有所不同，这对于预测疾病亚型和选择合适的治疗方案至关重要。因此，代谢特征可能成为区分癌症亚型的重要生物标志物。

基于代谢通路在肝癌中富集通路情况，显示了在肝癌中代谢异常的重要性。这些通路涉及到视黄醇、脂肪酸、氨基酸、药物等的代谢，这些是维持正常细胞功能所必需的基本组成部分[13]。因此，这些通路的异常可能在肝癌的发生和发展中发挥关键作用。这些代谢通路中的一些可能成为治疗肝癌的潜在靶点。例如，细胞色素 P450（cytochrome P450, CYP450）参与内源性物质和包括药物、环境化合物在内的外源性物质的代谢，是涉及药物代谢与生物激活作用的主要酶类。因此，

可能是药物治疗的有效靶点。

了解这些通路的异常可以帮助开发针对特定通路的治疗策略。这些通路的富集情况可能有助于识别不同的肝癌亚型。不同的肝癌亚型可能在代谢特征上有所不同，这对于预测患者的预后和选择治疗方法至关重要。另外，在这些通路中，可能存在与肝癌相关的代谢标志物。这些标志物可以用于肝癌的早期诊断、疾病监测以及治疗效果的评估。这些代谢通路的异常表明代谢调控在肝癌中具有重要性。了解这些调控机制可能有助于揭示肝癌的病理生理机制，从而为研发新的治疗方法提供依据。基于患者肝癌样本中这些代谢通路的异常，可以考虑个体化治疗策略。例如，可以根据患者的代谢特征来选择合适的治疗方法，以提高治疗效果并减少不必要的治疗。总的来说，这些代谢通路的富集提供了深入了解肝癌代谢异常的机会，有助于我们更好地理解肝癌的生物学机制并开发相关治疗策略。这些思考也可以为未来的肝癌研究提供方向和灵感。

代谢相关基因在多种生物学过程中发挥着关键作用。这些基因的异常表达与多种疾病的发病机制密切相关，同时也参与调控多个生物通路。在本研究中，我们选择了七个代谢相关基因，SEPHS1、YARS1、MTHFS、ATIC、CHAC2、TXNRD1、AKR1B15，深入探讨它们与不同疾病和生物通路之间的联系。

SEPHS1 编码的蛋白质在硒代谢中起着关键作用。这种蛋白质的功能是从硒化物和 ATP 中合成硒磷酸。硒磷酸是用于合成硒半胱氨酸的硒供体，硒半胱氨酸是在帧内 UGA 密码子处共转录地嵌入硒蛋白中的。SEPHS1 的异常表达已经与一些疾病相关联。例如，Shu 等[14]研究 SEPHS1 在肝细胞癌（HCC）中的表达上调。SEPHS1 基因敲除导致 SMAD2/3/4 和间充质标记物（包括 snail、slug 和 N-cadherin）在 HCC 细胞中的表达下调。此外，SEPHS1 基因敲除导致 HCC 细胞迁移和侵袭的减少，并抑制了 TGF-β对 HCC 细胞迁移和侵袭的刺激。最后，SEPHS1 的高表达与 HCC 患者的预后不良相关，表现为总生存期和无病生存期的降低。

YARS1 编码一种酪氨酸 tRNA 合成酶，并与一些疾病有关，包括 Charcot-Marie-Tooth 疾病、Dominant Intermediate C 和神经、内分泌和胰脏疾病。此外，YARS1 还参与 tRNA 氨基酸化和蛋白质代谢等途径。MTHFS 编码一种与甲基化反应相关的酶。异常表达可能导致神经管畸形和癌症。此基因参与甲基化通路。ATIC 编码的酶在腺苷酸代谢中起着关键作用。这种酶可以生成肌苷酸，这是嘌呤环系统的完全形成的第一个化合物。ATIC 的异常表达已经在一些疾病中被观察到，包括肿瘤和免疫相关性疾病。例如，Li 等[15]研究 ATIC 在肝细胞癌的发展中起着重要作用，通过调节 AMPK-mTOR-S6 K1 信号通路来促进肿瘤的生存、增殖和迁移。此外，ATIC 的异常表达也与一些免疫相关性疾病有关。CHAC2 编码一种谷胱甘肽代谢酶。它与谷胱甘肽代谢通路相关，直接作用于 GSH 并调节 ChaC1 以持续控制 GSH 代谢。异常表达与神经退行性疾病、癌症和炎症有关。例如，CHAC2 已被证明抑制胃癌和结直肠癌的进展[16]。此外，CHAC2 还能抑制 ROS 和炎症细胞因子的表达，从而促进 A. pleuropneumoniae 在 PAM 中的存活。TXNRD1 编码一种酶，它与氧化还原反应和抗氧化防御有关。它的异常表达与多种癌症、心血管疾病和神经退行性疾病相关。TXNRD1 涉及氧化还原通路。例如，TXNRD1 能够还原并激活硫氧还蛋白，一种含有二硫-二硫键活性位点的氧化还原酶，从而减少细胞蛋白中的氧化半胱氨酸残基[17]。此外，TXNRD1 也参与硒代谢和抵御氧化应激。AKR1B15 是一种新发现的酶，与 AKR1B10 有 92%的氨基酸序列一致性[18]。AKR1B10 是一种具有高视黄醛还原酶活性的酶，已知与多种癌症的发展有关。然而，AKR1B15 的酶活性和生理作用仍然知之甚少。AKR1B15 能够催化多种羰基底物（如芳香醛、烯醛、酮和α-二羰基化合物）的 NADPH 依赖性还原。与 AKR1B10 不同，后者强烈偏好全反-视黄醛，而 AKR1B15 对 9-顺-视黄醛显示出更高的催化效率。这可能表明 AKR1B15 在视黄醛代谢中具有生理作用。

七个代谢相关基因的多重功能和其与不同疾病的关联揭示了代谢网络在生物学中的复杂性。这些基因不仅可以作为潜在的治疗靶点，还为探索相关疾病的发病机制提供了新的线索。此外，它们参与的生物通路也受到了广泛的调控，可能对多个生理和病理过程产生影响。深入研究这些基因和通路的关联将有助于我们更好地理解生物学系统的复杂性，为未来的疾病治疗和药物研发提供更多可能性。

结论

在本研究中，通过对公共数据库中的代谢组学数据和临床信息进行综合分析，我们成功构建了一个基于代谢相关基因跨多种癌症类型的泛癌预后模型。这一模型具备了良好的预测能力，可以为癌症患者的预后状况提供有力支持。这一发现对于癌症研究和临床实践都具有重要的意义。同时，我们发现若干代谢特征在不同癌症类型中均表现出差异，这为预测分子亚型和对不同治疗方案的反应提供了新的线索。本研究的成果为癌症患者的个体化治疗决策和精准医疗提供了参考，为代谢组学领域的发展提供了新的思路和方法。

PU-196

医疗信息化新需求下的桌面虚拟化

王旭
郑州大学第一附属医院 病案科

1 桌面虚拟化概述

传统的医院 PC 办公环境在实际使用中主要存在以下几个问题：（1）PC 机硬件的多样化，使得难以对企业内部进行统一的集中化管理。（2）用户数据分散存储，终端设备丢失会引起数据泄露等重大安全隐患。（3）办公环境固定，对于台式机，只能在固定场所解决用户的办公需求，对于便携机，只能在固定设备上解决用户的办公需求。

为了解决这些问题，使用虚拟桌面构建办公环境的方案正被越来越多的用户所接受。桌面虚拟化建设通过构建一种全新的桌面交付模型，将原有 PC 桌面工作负载(包括操作系统、应用程序、用户数据等)集中托管在数据中心的集群计算资源上执行，终端用户使用支持远程桌面协议（Remote Desktop Protocol，RDP）的客户端设备（如 PC 机、笔记本、平板电脑等）随时随地与数据中心内虚拟桌面进行通信[1]。

2 桌面虚拟化系统原理及相关技术

2.1 桌面虚拟化系统架构

桌面虚拟化系统通常都可以简化为如图 1 所示的系统架构，其中包括终端接入层、权限控制层、虚拟化业务层、基础设施层和维护管理层。

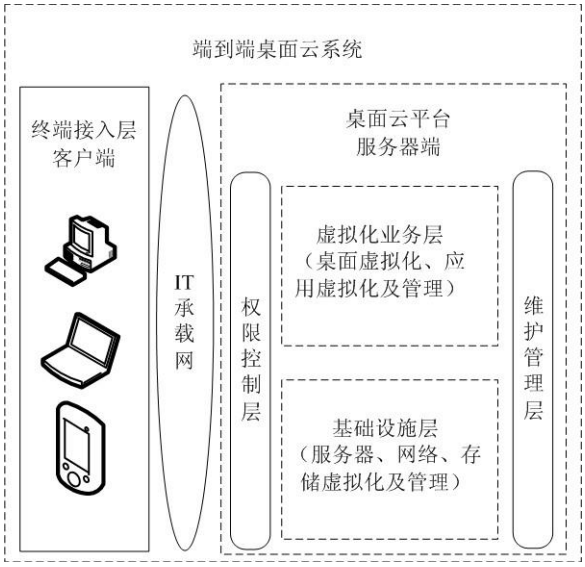


图 1 桌面虚拟化系统架构图

（1）终端接入层主要指负责接入桌面虚拟化系统的终端设备，按照访问类型可分为局域网用户和广域网用户两类[2]。局域网用户指通过局域网来进行访问桌面虚拟化系统的用户，访问方式属于内部网络访问，可以直接使用终端设备访问桌面虚拟化系统。广域网用户又称为远程用户，指通过广域网进行访问桌面虚拟化系统的用户，访问方式属于外部网络访问，基于数据安全的考虑，通常需要配置防火墙和接入网关等安全设备。

（2）权限控制层对接入桌面虚拟化系统的终端设备进行控制，根据访问方式或数据安全性等级确定是否需接入网关、防火墙等安全设备，同时也负责对访问桌面虚拟化系统的终端用户进行身份验证和识别。

（3）虚拟化业务层主要包括桌面虚拟化、应用虚拟化及会话的管理。桌面虚拟化组件负责实现虚拟桌面的创建、分配和交付；应用虚拟化组件负责实现虚拟应用的创建、分配和交付；桌面会话管理负责桌面云的接入控制和会话管理。

（4）基础设施层。该层主要包括虚拟化平台和资源的管理和调度，虚拟化平台完成对各种物理资源的虚拟化，为虚拟机提供各种虚拟资源[3]；资源管理及调度主要是对各种云物理资源进行管理，在创建虚拟机时，为虚拟机分配相应的虚拟资源，同时完成各种业务的计费功能。

（5）维护管理层。运维管理系统包括桌面云的业务运营部分和系统的维护管理部分。桌面云的业务运营管理可直接由运营人员管理。系统的维护管理主要由云管理平台负责完成，主要指桌面云基础设施的维护管理功能，包括业务节点的安装部署，节点性能监控，配置管理等功能[4]。

2.2 桌面虚拟化的影响因素

（1）存储因素。桌面虚拟化系统性能的优劣受很多因素的影响，其中最重要的一项就是存储因素。判断桌面虚拟化系统是否成功的关键是用户体验，而存储因素在最大程度上决定了用户体验。为了缓解对存储系统的压力，可以采用增加高速硬盘分担读写操作、升级存储网络链路、增大存储容量等方法。

（2）I/O 负载因素。I/O 高峰时桌面虚拟化系统常遇到的另外一个问题。I/O 高峰就是大量用户在同一时间内从服务器下载系统或在同一时间注销系统并向服务器保存所做修改时对服务器 I/O 设备的考验。如果不能顺利解决或很好应对 I/O 高峰，用户就会出现不能进入系统或无法保存数据等问题，从而影响用户的使用[6]。

（3）网络带宽因素。网络带宽对用户体验的影响很大，如果网络带宽足够，丢包和延时就会很少，用户体验度就会大大提高。但是在广域网中，网络情况比较复杂，有延时的不可预料性，这样虚拟桌面的传输就会受到很大影响，所以要尽量在网络带宽好的环境中使用桌面虚拟化，带给用户良好的体验。

3 医院桌面虚拟化系统的建设

桌面虚拟化系统在医院的应用场景主要有临床医疗、移动医疗和综合管理，根据桌面虚拟化系统应用场景的不同，其设计方案可以存在适当差异，以满足特定场景的特殊需要。以 VMware 系列产品为例，简要介绍前两种应用场景下桌面虚拟化系统的建设。

3.1 临床医疗桌面虚拟化系统建设

3.1.1 现状描述

医疗行业现有的临床诊疗应用中，基本上是通过 PC、笔记本电脑作为访问 PACS、LIS、HIS、EMRS 等应用系统的客户端。随着医疗信息化系统的建设，终端设备上需要承载的应用越来越多，终端设备老化也会导致各种系统问题。处于人力资源不足的医院 IT 机构，每天需要奔波与各个业务科室、甚至于分支机构，去处理应用软件、操作系统等软硬件问题。这不仅增加了 IT 人员的维护困难，更重要的是， 对于需要全天候运作的医生工作站、护士工作站，每一次计划外的停机，都有可能带来无法估量的损失。作为院内最重要的访问业务系统客户端的医生工作站、护士工作站存在多医师、护士共用一台主机的情况，如何既能保证多人使用同台终端主机的安全问题，又能给使用者带来便捷的使用感受，是一个急需解决的难题。

PACS 系统保存的病患影像资料，CT/MRI 单幅图像大小一般在 0.5～1M 之间；CR/DR 一般 10M 左右，每例的图片数量都在数十幅左右，因此每例总量比较大；需要通过局域网网络传输到医生工作站，再通过相应的 DICOM 软件打开。在这个过程中由于图像较大，导致传输时间长，特别是在网络应用高峰期时，更是会因为图像传输时间的长，影响医生工作效率。

3.1.2 方案设计

为解决上述临床医疗应用过程中的问题，可以采用 VMware 的 Horizon View 桌面云和 vCops for View 监管方案，资源池可自动浮动或固定。其基本架构如图 2 所示，对临床诊疗场景中的医生、护士工作站采用虚拟化解决方案，每个医护人员都可共同用一台终端设备，但是每个人都拥有独立的桌面。对拥有的桌面的任何操作都不会影响到他人。可以对不同的桌面进行不同的安全配置，这样不仅保证了终端主机的使用安全，又能给医护人员带来良好的使用感受。

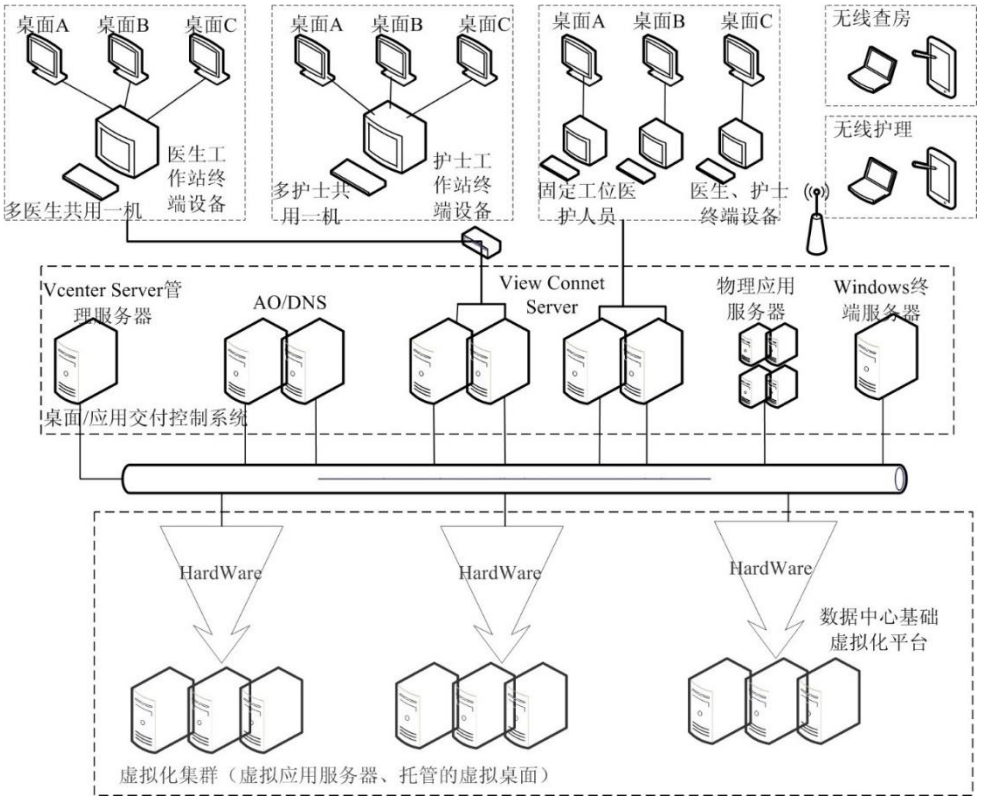


图 2 临床医疗桌面虚拟化系统建设方案

对于管理人员，可以通过监控平台监控整个虚拟桌面的运行情况，并且能及时对出现问题的虚拟桌面进行问题排查；对于需要进行移动查房、移送护理的医护人员同样采用虚拟化方案，通过移

动终端设备，连接托管在后台服务器上的虚拟桌面，进行常规医护操作；对于终端上的数据使用安全和传输过程中的安全，通过该方案，所有的数据都只是在数据中心内部流动，传输到终端设备上的只是屏幕和键盘的变化数据，对于数据的安全起到了很好的保障作用。通过该方案还可以很好的满足各种不同规格、不同型号的多外设使用；由于所有的信息流都只是在数据中心内流通，PACS 影像资料不需要从数据中心到医生工作站桌面的传输，在传输时间上比未采用该方案之前快数倍，很好地改善了用户体验。

3.2 移动医疗桌面虚拟化系统建设

移动医疗主要包括三种形态，分别为院内的移动查房业务、院内的移动护理业务和院外的社区、农村移动医疗业务。

3.2.1 现状描述

医护人员和患者之间需要频繁地在院内移动检查。传统模式下，护士在对患者进行护理中，需要在各式各样的表格手工记录众多的患者体征信息，然后录入到系统中，形成相应的患者信息库。需要查看住院医生对患者最下下达的医嘱，并记录在相应的纸质病历上，带着众多的病历文件，到各个病房为患者执行新医嘱，并登记患者的医嘱执行状态以及护理记录；对于需要到社区、农村进行移动医疗的医护人员来说，当前模式下，医护人员对患者的诊疗记录、医嘱都是通过纸质文档先行记录，回到医院内再录入到信息系统中。对于患者的过往病史，需要患者口述，医护人员无法实时调阅。由于患者与医生间的地区性差异和文化水平差异，对病史的描述和理解会存在一定的差异，这可能导致医生对患者的诊疗存在误差。

3.2.2 方案设计

为解决上述移动医疗应用过程中的问题，可以采用 VMware 的 Horizon Suite 进行方案设计，资源池可自动浮动或固定。其基本架构如图 3 所示。在方案中在客户端与数据中心之间传输的只有屏幕信息及键盘鼠标消息，这不仅能有效降低网络带宽需求，还能有效提高数据安全，保证敏感数据只在数据中心内部传输。为医护人员提供完整的、实时的、随时随地的患者信息访问，有助于提高医疗质量，保证各类手持设备问医疗信息系统体验上的一致性。

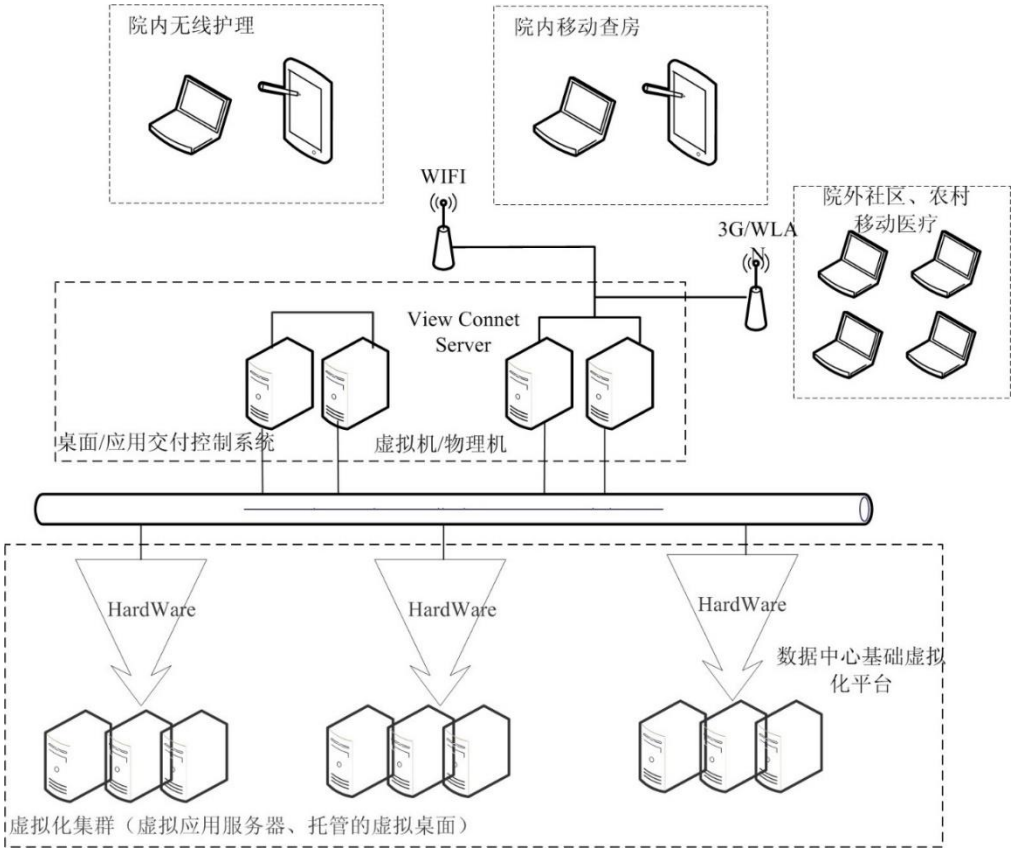


图 7-22 移动医疗桌面虚拟化建设方案

为了让所有用户都能有较好的使用体验，除了考虑每个分支机构的网络接入方式及带宽外，还需考虑数据中心虚拟化基础平台的对外接入带宽。这样用户可以通过移动设备高效的访问所需信

息。

4 总结与展望：

桌面虚拟化承载了移动医疗与远程医疗的应用，可保障医生出差时仍然能连接到服务器，它完成了医疗模式的新型转变。虚拟桌面以其集中管理、灵活部署、多设备接入、高效运维的特性和客户终端易部署、易管理、易维护的特性越来越多地受到医院关注。特别是桌面虚拟化的广泛应用将促进移动医疗、智慧医疗、远程医疗、区域医疗及医疗大数据的发展。

PU-197

5G 物联网下生命体征智能采集系统设计及实现

彭洋

湖南医药学院总医院信息部

一、背景

随着计算机技术和信息技术的高速发展，移动互联网和物联网技术正加速向各行各业渗透。尤其近几年来，随着智能平板电脑的普及和 5G 网络技术的成熟，移动医疗已在医院逐渐成为一种应用趋势。借助移动医疗平台，医院逐步实现了无纸化查房和床旁检验、检查项目的信息网络传输，体温、脉搏、血压、血氧饱和度等生命体征信息的采集——这项重要的临床护理工作，也由传统的手抄再转录，发展到手持无线数据终端（简称：PDA）床旁录入方式，从基础信息采集环节提高了医护人员的工作效率，降低了差错概率。临床生命体征信息采集方式分为手工录入和自动录入两种，目前，只有极少数移动应用产品能够做到自动采集数据，大多数基于移动医疗的生命体征采集还是采用护士床旁手工录入的方式。受 PDA 尺寸的限制，人工床旁录入效率相对低下且容易出错。因此，如何进一步提高采集效率和数据的准确性，借助物联网技术，实现生命体征智能自动采集，成为了医护人员、医疗信息技术人员和医院管理者关注的难题。

二、系统方案设计选型

物联网系统一般由感知层，网络层，应用层三层结构组成。目前国内患者生命体征智能采集系统感知层一般采用的方案有：

1. 设备采集：通过各种生理监测设备，如心电监护仪、血压计、血糖仪、血氧仪等，实时或定期采集生命体征数据。这些设备可以是传统的医疗器械，也可以是如智能手环、智能手表等可穿戴设备。2. 无线传感网络（WSN）：采用无线传感器节点部署在人体周围或内部，通过无线通信技术实现生命体征数据的实时采集和传输。这些传感器可以检测心率、呼吸率、体温等生命体征参数。3. 智能手机应用：利用智能手机内置的传感器（如加速度计、光学传感器等）以及外接的生理监测设备，通过各种生理信号处理算法实现生命体征的采集。用户可以通过安装相应的应用程序在手机上实时查看和分析生命体征数据。4. 图像识别技术：通过分析摄像头采集的人体图像或视频，利用图像处理技术和生物特征识别算法实现对生命体征的无接触式采集。例如，利用摄像头捕捉面部皮肤微弱的颜色变化，实现心率的无接触式检测。5. 嵌入式生物传感器：将生物传感器嵌入到人体内部，例如植入式心脏起搏器、可穿戴胰岛素泵等。这些传感器能够实时监测生命体征参数，并通过无线通信技术将数据传输至外部设备或医疗专业人员。以上几种方案可以相互结合使用，以满足不同场景和应用需求。

生命体征智能采集物联网系统中网络连接有多种方式，经过研究分析，各连接方式及其优缺点如下图所示：

连接方式	有线连接	Wi-Fi连接	蓝牙连接	ZigBee连接	LoRa连接	NarrowBand-IoT (NB-IoT)连接（窄带物联网）	2G/3G/4G/5G模式连接（宽带物联网）
优点	较高的速度、稳定性和安全性	无线连接，部署方便；传输速度较快；设备普及度高，易于实现互联互通	无线连接，部署方便；功耗较低；成本较低；适用于短距离设备之间的通信	低功耗，适合长时间工作的物联网设备；传输距离较远，可达数百米；自组网能力，可扩展性强	低功耗，适合长时间工作的物联网设备；超远距离传输，可达数公里；抗干扰能力强	低功耗；传输距离较远，可达数公里；抗干扰能力强；基于蜂窝网络，覆盖率广	可利用现有的移动网络，在终端接入上与现有的手机卡接入方式一样，远距离传输，传输速率高，抗干扰能力强
缺点	布线成本较高，连接物体受限于线缆长度，不适用于大范围部署	受制于无线路由器的信号范围，不适合长距离传输；信号受干扰影响较大；功耗较高，不适合低功耗设备	传输距离较短，通常在10米以内；传输速度较低，受信号干扰影响较大	传输速率较低；应用领域相对更专用，设备互通性较差	传输速率较低；成本相对较高；需要建立专用网关	传输速率较低；设备成本相对较高；需要运营商支持	耗电，网络覆盖穿透不足，基站难以支撑高度密集的联网数

图一 物联网网络连接方式及其优缺点分析图

多种生命体征智能采集方案和网络连接模式各有优缺点，根据护理人员需求，结合本院实际：①住院部只有病房内部实现了 WIFI 覆盖，配置室、治疗室、走廊等位置并未完全覆盖 WIFI；②采集设备数量有限，科室患者较多，如使用蓝牙频繁切换连接会大大降低采集效率；③生命体征实时采集要求传输效率高；④前期移动护理系统已使用移动运营商 5G 物联网模式，效果良好。故患者生命体征智能采集系统决定采用支持 5G 联网的多参数（支持同时采集体温、脉搏、血压、血氧饱和度）生命体征监测仪 aVS2000 设备采集方案。

三、系统需求及架构

1. 系统需求分析

通过深入临床调研分析，收集到的系统需求如下：

（1）可支持采集体温、脉搏、血压、血氧饱和度等生命体征；

（2）可对接移动护理系统，采集生命体征时不需要切换程序；

（3）可根据医嘱提取需要采集生命体征的患者；

（4）采集不同患者生命体征时不需要切换界面；

（5）采集设备轻便、便携，不增加护理人员负担；

（6）生命体征采集后可直接上传至护理系统，不需要多余操作；

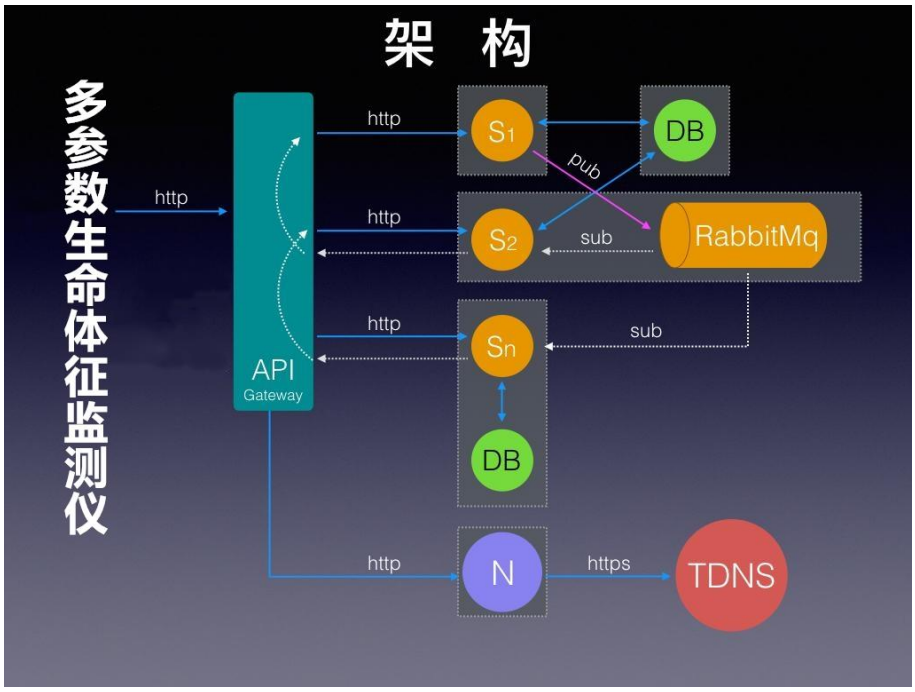
（7）可在一个界面操作采集多种生命体征。

2. 系统需求实现设计方案及架构

物联网架构模式是指在物联网技术中，为了实现设备之间互联互通以及数据的传输和处理，所采用的一种设计模式和技术架构。以下是几种常见的物联网架构模式：1. 集中式架构模式：所有物联网设备直接与云端服务器进行通信，所有数据处理和存储也在云端进行，设备之间的通信和数据传输都通过云端中转。这种架构模式适用于设备数量较少、数据处理量较小的场景。2. 分布式架构模式：在物联网系统中部署多个智能节点，每个节点都有自己的处理能力和存储能力，设备之间的通信和数据传输可以通过节点之间的互联来实现。这种架构模式适用于设备数量较多、数据处理量较大的场景。3. 边缘计算架构模式：在物联网系统中引入边缘计算技术，将数据处理和存储功能下移到设备端，通过在设备端运行算法和模型来实现数据处理和分析，减少了数据传输和处理的延迟，提高了系统的响应速度和效率。4. 微服务架构：这种架构将物联网系统划分为许多独立的、可扩展的、松耦合的服务。每个服务都有特定的功能，可以独立地开发、部署和更新。这有助于提高系统的灵活性、可维护性和可扩展性。5. 混合式架构模式：将集中式架构、分布式架构和边缘计算架构结合起来，根据不同的应用场景和数据处理需求，采用最优的架构模式来实现数据处理和传输。

由于多参数生命体征监测仪为 Android 系统，支持直接安装移动护理系统（移动端），相当于

一台支持多参数生命体征智能采集的 PDA。移动护理业务终端多，数据处理量大，要求响应速度快，传输效率高，单点故障不能影响全院整体业务，根据我院实际情况及需求，决定采用最先进的微服务框架。生命体征智能采集系统依托医院已建成的移动护理系统和多参数生命体征监测仪 aVS2000，使用移动运营商 5G 物联网络，前后台系统之间的交互采用 REST 协议，实现医护人员在病床边实时录入、查询、修改病人的基本信息、医嘱信息和智能采集生命体征等功能。将多参数生命体征监测仪和移动护理系统对接后，移动护理系统点击测量体征可直接调用多参数生命体征监测仪测量界面，体征检测仪通过传感器测量出数据后，将数据回写至移动护理系统，数据不再需要护理人员操作，直接填入相应表单。



图二 系统架构图

四、系统设计及实现

1.多参数生命体征监测仪接入运营商 5G 物联网

前期医院上线移动护理系统时已建成基于运营商 5G 物联网络的全院无线物联网络。多参数生命体征监测仪插入运营商物联网卡后接入运营商 5G 物联网，然后采用 APN 技术接入医院内部网络，即可通过移动护理系统获取患者和医嘱信息。

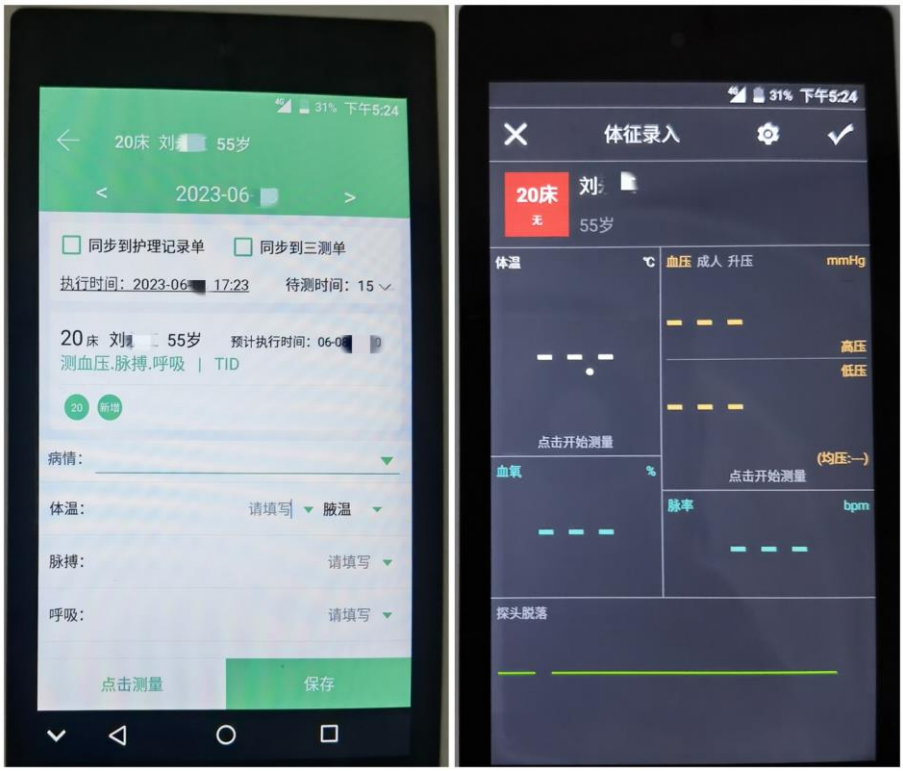
2.移动护理系统移动端 APP 功能实现

使用 Android Studio 编程软件实现对移动护理系统 APP 生命体征采集模块的设计。首先需要改造移动护理系统移动端，在点击相应医嘱任务测量按钮时调用多参数生命体征监测仪已封装好的测量接口方法类（startActivityForResult）直接跳转测量主界面。

```
try {
    Intent intent = new Intent( action: "android.intent.action.MAIN");
    ComponentName name = new ComponentName( pkg: "com.xicoo.vitalsignsmonitor",  cls: "com.xicoo.vitalsignsmor
    intent.setComponent(name);
    Bundle bundle = new Bundle();
    bundle.putBoolean(THIRD_PARTY_FLAG, true);              //第三方APP调用
    PatientInfo patientInfo =new PatientInfo();
    patientInfo.title="查房";
    patientInfo.name="王小明";
    patientInfo.info="测量 血压、血氧、体温、脉搏";
    patientInfo.bedNo="+11";
    patientInfo.levelOfCare="二级护理";

    bundle.putString("PatientInfo", JsonUtil.toJson(patientInfo));  //病人腕带ID

    intent.putExtras(bundle);
    startActivityForResult(intent, REQUEST_CODE_1);
} catch (Exception e) {
    Toast.makeText( context: this,  text: "请先安装体征测量APP", Toast.LENGTH_SHORT).show();
    e.printStackTrace();
}
}
```

图三 代码实现界面

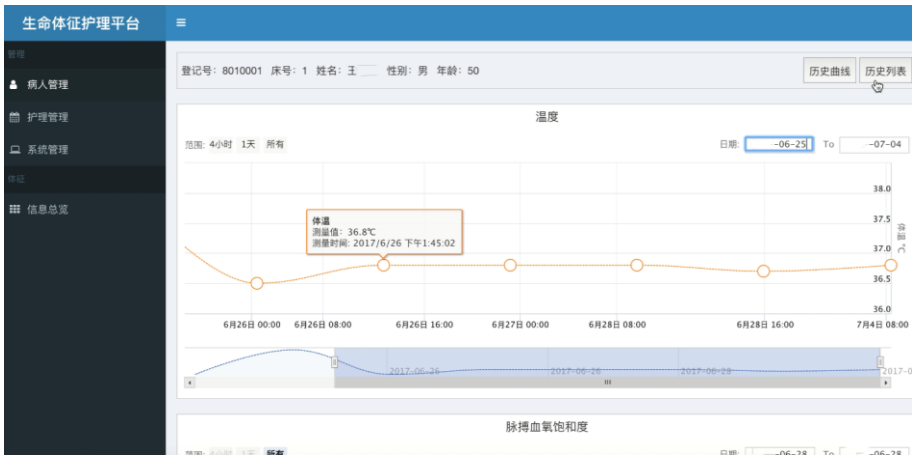
测量主界面可直接测量采集体温、脉搏、血压、血氧饱和度等多种生命体征，采集完体征后点击确定可直接将相应测量值填入移动护理系统，并可直接将测量值同步至护理记录单或三测单中，实现生命体征的智能自动采集。



图四 体征数据采集界面

3. 移动护理系统处理存储数据

数据采集保存后，经过校验、过滤、转换后存入移动护理系统数据库，可根据任务类型在移动护理系统三测单、护理记录单等单据中展示数据，并可将体征数据传送给集成平台，供医院其他系统获取调用。体征数据展示界面统一，来源一致，满足医疗数据采集规范要求，并可根据需求在生命体征采集分析平台进行二次分析，为科研人员提供更加准确、高效的数据支持。



图五 生命体征采集分析平台

五、系统应用体会

自 2022 年底上线使用以来，以有 60 张病床的科室为例，每天每个病床需要采集 1 次血氧、2 次血压、2 次体温、1 次脉搏，生命体征智能采集系统的使用能够节省护士誊抄和录入时间约为 2－4 个小时。将护理人员从繁杂的重复登记誊写劳动中解放出来，真正体现了医院还时间予护理人员，护理人员还时间予病人的理念，让护理人员能有更多的时间更好地为患者服务，提升了患者的就医感受，推进了我院卓越护理服务的建设。

生命体征智能采集系统初步实现了患者体征监护数据的自动采集、处理和存储，为移动医疗系统平台实时提供体征数据。医院其他系统通过统一的集成平台获取患者的体征数据，实现了体征数据一次采集、全程共享的效果。生命体征智能采集系统一次测量采集医嘱的所有体征，数据自动上传系统，节省了多仪器切换测量及护士手动录入体征的时间，大大提高了护士效率；基于信息化工具之间的互联互通，将录入出错率降低为 0，同时也有助于医院整体信息化水平的提升。

下一步规划将血糖等数据采集也纳入该系统中，优化采集速度，增加患者床位分组等功能，进一步提升采集效率，不断推广应用，充分利用信息化手段，解决临床痛点，服务广大群众，在建设智慧医院的道路上不断前进。

PU-198

生成式人工智能在智慧医疗方面的实现方法研究

张萍 乔一鸣*
大同市第五人民医院

一、目的：

通过应用生成式模型,开发智能问诊和病历系统,以提高医患交流效率与病历质量。方法采用知识图谱增强的 XLNet 进行问诊理解,LLM 模型生成答复;使用 LSTM 语音识别,BERT 提取关键信息,BART 生成病历。预计相比传统方法,提问的匹配率提高 20%,病历规范性提高 10%。结论表明生成式模型能显著改进智能问诊和病历系统,临床应用前景广阔。

1、智能问诊

智能问诊的目的是为患者提供更加快速、准确的初级医疗咨询和预诊服务，缓解医务人员的工作压力。提升医务人员的工作效率，通过减少冗余的工作和提高问诊的精准度来实现。同时通过提供基于医学知识库的有针对性的咨询和建议，增强患者的问诊体验。

2、智能病历

智能病历旨在通过数字化和智能化的方式，提高病历的管理效率和利用价值，进而优化医疗服务的质量和安全性。

二、方法：

1、智能问诊

自然语言处理：利用大语言模型技术来开发智能问诊系统，该系统能够理解和处理自然语言输入，提供基于医学知识库的咨询和建议。

知识图谱构建：创建一个包含大量医学知识的知识图谱，来支持智能问诊系统的决策过程。

实施自然语言处理和知识图谱构建的任务是一个系统和多阶段的项目。根据国际和国内的先进经验，初步可以设计这两个核心组件的具体实施方法：

1.1 数据采集

医学文献：利用公开可用的医学数据库收集相关文献。

电子病历：利用医院的历史电子病历数据。

1.2 数据预处理

数据清洗：移除冗余和错误的数据。

标注：对数据进行标注以训练模型更好地理解医学术语。

1.3 模型训练和开发

选择模型：选用先进的 NLP 模型如 BERT 或 GPT-3 作为基础模型。

训练：使用大量的医学数据来训练模型，使其能够理解和处理医学相关的查询。

1.4 系统集成

API 开发：开发 API 来将 NLP 模型集成到智能问诊系统中。

测试和优化：进行多轮的测试和优化来确保系统的稳定性和准确性。

1.5 知识提取

数据源：确定和整合多源数据，包括病历、医学研究报告和公开数据库。

实体识别：通过 NLP 技术来识别和抽取关键实体和关系。

1.6 图谱构建

节点定义：确定知识图谱中的节点，如病症、药物、治疗方案等。

关系定义：定义节点之间的关系，以建立一个结构化的知识体系。

1.7 知识库的维护和更新

持续更新：确保知识库能够反映最新的医学研究和发现。

反馈循环：建立一个反馈机制来从用户那里收集反馈，并用这些反馈来不断优化知识库。

2、智能病历

2.1 电子病历系统：开发一个可以自动化生成和维护电子病历的系统，该系统将包括智能录入、检索和数据分析功能。

2.2 智能录入：使用语音识别技术,识别医患对话并转换为文本。使用自然语言理解技术,从文本中提取症状、检查结果等结构化医学信息。应用知识图谱,自动将信息匹配到标准术语。

2.3 智能检索：构建基于医学知识图谱的智能问答系统,可以解析医生的查询意图,从病历数据库中检索需要的信息。

2.4 智能生成：使用自然语言生成技术,根据提取的结构化信息,自动生成叙述流畅、语法正确的病历文本。采用医学领域的语料进行模型预训练,提高生成文本的标准性和适合性。

2.5 数据分析：对病历文本进行命名实体识别、关系提取,构建结构化知识库。应用数据挖掘技术,分析病历数据,获得疾病信息统计、药物史分析等知识。

2.6 系统框架：基于深度学习和知识图谱技术,设计端到端的病历管理系统。使用区块链技术,保证病历数据的安全性、可追溯性和一致性。

2.7 模型迭代：利用真实的临床病历数据,持续优化模型,提高系统的智能化程度。评估系统在不

同医院、不同科室的适用性,进行部署调整。

三、结果：

1、智能问诊

提高预诊效率：减少医务人员的工作量，提高预诊的速度和准确度。

改善患者体验：通过提供快速、个性化的服务来改善患者的体验。

2、智能病历

提高数据准确性和完整性：减少人为错误，提高病历数据的准确性和完整性。

优化病历管理：通过数字化和智能化的方式来优化病历的管理和利用。

四、结论：

通过应用生成式人工智能模型开发智能问诊和病历系统，将极大地推进智慧医疗信息化建设，不仅将提高医疗服务的效率和准确性，而且还将通过高效的数据分析工具为医院提供数据驱动的决策支持，从而促进更有根据和安全的医疗决策。同时，它也将为患者提供更好的服务体验，提高他们的满意度。在未来的发展中，系统可以继续优化和更新，确保其始终保持先进性和适用性。通过实施这一创新项目，医疗机构可以成为国内外智慧医疗的典范，为医院赢得更多的认可和影响力，同时为医院开创更多的发展机遇。

PU-199

浅谈智慧医院管理中检验信息化

赵旭涛

汾阳医院

目的 检验科作为医院的一大分支，检验科信息化是未来数字化医院的基础，检验科大量的数据、信息构成了诊疗过程和决策的重要依据。检验科信息化大大提高了医学检验技术的更新速度和检验效率，同时也为众多的患者带来了方便。检验科传统的工作方式以及工作流程已经不再适应于当下的发展,该学科的专业性还在不断细化中，检验手段和项目也在不断增加，使得检验系统变得越来越复杂，此时检验科信息化尤为重要。

方法 1 HIS 和 LIS 数据共享

HIS 与 LIS 系统主要是两个独立的服务器，具有不同的数据库，不能进行信息共享。我院通过系统间的接口，实现 HIS 和 LIS 系统信息共享。(1)住院病人：医生通过住院医生站开立检验医嘱，传至护理站，护士审核执行后，HIS 将患者基本信息及医嘱信息传至 LIS 系统，LIS 生成条码,检验科签收条码同时将费用计入 HIS,检验科老师按照条码提示信息将标本按检验小组分好，由机器人送至各检验小组，检验科老师上机检验，检验完毕后 ，结果传送 LIS 中，检验科老师审核无误后发布报告，检验结果传至 HIS,供临床医生调阅。（2）门诊病人：门诊医生开具检验申请单，病人缴完费后，检验科老师在 LIS 系统中扫描申请单上二维码从 HIS 系统中调取患者基本信息及检验信息并将其保存至 LIS 系统中，同时打印条码及回执单，待检验结束后 ,患者持回执单扫码取检验报告，同时 LIS 将检验结果传至 HIS，供医生调阅。

2 条形码信息化管理

条形码信息与医院 HIS 系统患者信息相关联，引入既方便患者又方便护理人员，更能避免标本采集过程的种种差错。临床医师开具检验申请单，选择检验项目时，会自动筛选出检验标本；若医生开单时未选择标本，标本采集时自动默认标本，护士床旁 PDA 采集标本和门诊检验科扫描识别条码信息时，提醒标本种类，检验标本及采集时注意事项如试管颜色、采血量、是否需要抗凝血等。待

标本运送至检验窗口，工作人员扫描容器上的条形码，完成对样本的签收，对住院标本进行收费确认，防止了漏费，不合格的标本拒收，有“急”字的标本走急诊通道。随后将标本分发到各个实验室，快速完成样本的签收，节约了时间，加快了整个检验进程，减少了人工干预造成的差错。标本上机后，仪器自动扫描条形码，将条形码信息传到LIS 工作站对条形码进行分析，从LIS 系统中获取项目信息传送到仪器，仪器接收到信息后进行分析检验。

3 检验结果自动审核

传统的人工审核模式在高强度、高压力下已逐渐显现弊端，同时由于工作经验、知识背景和专业素质差异，存在对基本信息、极值数据、历史波动、逻辑错误等制约影响检验报告的及时、准确和质量。自动审核不仅在高效处理海量数据发挥了优势，同时规范了报告审核标准，避免人工审核可能出现的差错。使用自动审核的仪器通过接口程序，完成与 LIS 系统之间的数据交换。实现样本申请发送，自动编号、检验结果、审核提示信息、样本追踪信息接收、前处理、样本血清图片接收等功能。

结果 4.1 检验结果历史回顾

检验系统中审核报告时，能够自动显示病人同项目的历史检验结果和对应趋势图进行比较分析，非常直观明了。同时医生站也可以调阅患者历史检验结果及对应趋势图。

4.2 质量控制

全院各个检验系统都会进行质控管理，通过仪器和项目配置，质控规则设置确定检验科对每个实验室仪器质控的要求和条件。系统根据质控规则和项目对应对各个检验环节进行质量控制。系统也支持指定质控计划及质控记录报表统计。从而掌握仪器的可靠性和检验结果的准确性，以便适时采取措施。

4.3 报告查询

患者主索引可以将同一患者所有检验报告进行关联，在 360 视图中统一展示，医生站、护士站可点击患者 360 视图查看患者各项检验报告。检验系统中的检验结果可细化到各检验指标，检验项目中的每一个指标，均显示参考值范围及对应测量值、异常标志，并自动判断是否异常。

不合格标本管理

检验科可在 LIS 系统中对不合格标本进行登记，并退回不合格标本时填写退回原因，同时可以在检验系统中统计查询不合格标本记录，也可以查看护士处理结果，基本形成不合格标本的闭环管理。结合医院不同临床科室的诊疗特点和患者自身个体的动态变化，设置了检验结果动态变化对应的危急值，有助于临床及时发现病情变化并进行处置。系统实现了检验危急值的自动甄别、实时提醒的功能。当检验结果出现危急值时,检验科电脑立即出现弹窗提醒，工作人员审核报告后，临床科室医生和护士的电脑均出现弹窗提醒。同时，危急值的发布与接收之间的 “接收监控反馈回路” 有效地对危急值的接收进行监控。

设计了检验信息提醒大屏，包括样本运行状态、危急值接收情况、失控提醒及 TAT 提醒，非常直观明了 。

结论 信息化管理使检验科的工作更加规范化、制度化、现代化、专业化，为检验科工作的质量认证打下了坚实的基础，大大提高了医学检验技术的更新速度和检验效率，同时也为众多的患者带来了方便。

PU-200

信息化在检验结果自动审核中的应用

赵旭涛
汾阳医院

目的 自动审核是将临床实验室中各种独立的自动化仪器以特殊的物流传送设备串联起来，在信息流的主导控制下，构成流水线作业的组合，形成大规模的实验室常规检验过程的自动化。自动审核是计算机系统按照临床实验室设置的已通过验证的规则、标准和逻辑，自动对检测结果进行审核并发布检验报告的行为。利用条形码技术，实现从样本采集、标本登记、前处理、轨道流转、仪器检测、检后跟踪全流程管理，实现结果的自动审核，同时，在全流程中记录日志，便于后期跟踪、维护、纠正。通过对 LIS 的改造及流水线的应用，一方面提高检验科管理水平、检验质量，另一方面为临床、患者提供快速、精准报告，缩短了 TAT。

方法 近年来，随着先进检验技术及高精尖仪器在检验医学中的不断应用，检验医学得到迅猛发展，在临床诊疗工作中发挥着越来越重要的作用。临床实验室报告呈指数增长，传统的人工审核模式在高强度、高压力下已逐渐显现弊端，同时由于工作经验、知识背景和专业素质差异，存在对基本信息、极值数据、历史波动、逻辑错误等制约影响着检验报告的及时、准确和质量。自动审核不仅在高效处理海量数据发挥了优势，同时规范了报告审核标准，避免人工审核可能出现的差错。自动审核是目前提高临床检验工作效率并保证标准化审核最有效的辅助手段之一。
自动审核是计算机系统按照临床实验室设置的已通过验证的规则、标准和逻辑，自动对检测结果进行审核并发布检验报告的行为。自动审核的一般流程包括：室内质控、检验结果数据类型、仪器警告符号、危急值识别、自动审核范围比较、历史结果及极值范围比较。自动审核规则由流水线和 LIS 系统共同制定，最终通过 LIS 系统实现。待检测结果完成后，流水线会根据审核规则进行数据分析处理，LIS 系统通过通讯获取流水线处理信息，通过 LIS 系统设定反馈信息对应关系进行校验，对于存在报警信息的样本将不予自动审核，需人工干预审核。对于自动审核，LIS 系统需要做相应的改造，具体改造如下：自动审核维护功能：①自动审核开启与关闭；②自总审核样本范围如门诊、住院、体检等；③自动审核人员设置；我院自动审核通过 LIS 与流水线仪器的接口程序实现自动审核，流水线会根据内部规则设定审核范围，样本报警及审核信息均会通过共享文件传递给 LIS，LIS 建立相应的表存储这些内容。在完成自动审核前要获取信息，判断是否可以通过审核。图 1 为自动审核流程图
血清质量直接影响着检验结果的准确性，因此接收和显示该项目结果显得非常重要。我院罗氏流水线可以通过中间体软件根据血清图片转换为血清质量结果供实验室参考。罗氏该项目的通道为 SerQuality，结果为文字类型。检验系统只需要接收该项目并显示在报告界面，但是不能作为报告打印项目显示在患者报告单上，只是供检验科老师审核报告辅助作用。审核老师可根据该项目的结果判断该样本的检测结果是否准确，是否需要复查处理还是直接可以审核报告，因此该功能对检验报告审核的意义十分重大。图 2 仪器抓取到的血清图片传送至检验系统

结果 通过基于流水线的自动审核的应用不仅减少人工审核环节的工作量，缩短标本周转时间(TAT)，更有意义的是可以将报告审核工作做到规范化、标准化，提高实验室信息化智能化管理水平，同时可以让检验科的老师更专注于针对不同的科室 、病种及不同年龄组人群极值及危急值的研究，以辅助医患之间产生最佳的诊疗决策。也将加速实验室产生更多的检验数据资料，对疾病预防控制 、筛查 、诊疗带来巨大的应用价值。

结论 自动审核也存在潜在的风险，系统本身参数规则设置、分析前因素以及仪器检测、管理制度缺失以及信息数据安全等，因此利用 LIS 系统对检验结果进行自动审核，不仅要看到其优点，还要考虑其不足和不可控因素。建立抽查回顾制度，定期验证自动审核规则，定期考核授权人员，只有不断验证和学习训练，掌握理论和实际操作并结合不断细化的行业标准才能积极面对实验室自动审核所带来的风险。

PU-201

Immunotherapeutic response and prognosis prediction in lung adenocarcinoma based on lncRNA signatures of activated CD4+ memory T cells

刘盛波[1] 周海榆[1] 李洽轩[1] 尤培蒙[2,3]

1.广东省人民医院

2.江西省肿瘤医院

3.江西省肺癌研究所

目的 CD4 memory T cells are an important component of the tumor immune microenvironment (TIME), influencing survival and therapeutic strategies of patients. However, the mechanisms of activated CD4+ memory T cells (CD4amT) and associated lncRNAs in lung adenocarcinoma (LUAD) remain unknown.

方法 The LUAD dataset acquired by Public Genetic Database was used as training cohort (TCGA-LUAD) and validation cohort (GEO-LUAD). Univariate Cox analysis was used to infiltrate the prognostic CD4amT related lncRNAs (CD4amTRL). LASSO analysis and multivariate Cox analysis were performed to construct model by CD4amTRL. Patients were divided into high- and low-risk groups based on CD4amTRL signature. Tumor Immune Dysfunction and Exclusion (TIDE) was used to predict the response to immune checkpoint blockade and immunotherapy cohort was used to validate the immune responses. We also explored the heterogeneity of gene expression patterns regarding to TIME. Cellular localization of lncRNAs was assessed at the single cell level.

结果：In LUAD cohorts, compared to those with the high abundance of CD4amT, patients with low abundance had better OS (TCGA-LUAD: HR = 0.678; p = 0.028; GEO-LUAD: HR = 0.435; p < 0.0001). Then, we developed the CD4amTRL signature based on six lncRNAs, with good prognostic performance to stratify high and low group (1-year survival AUC of TCGA-LUAD and TCGA-LUAD: 0.737 and 0.719, respectively). TIDE indicated that low-risk group has better immunotherapy response (p < 0.01). In immunotherapy cohorts, compared with the high-risk of CD4amT signature, low- risk group also had better OS and PFS (p < 0.001). Significant differences were identified in molecular expression patterns relevant to immune process. We found that LINC00996 was localized to plasma cell-like DC cells.

结论 Our novel CD4amTRL signatures can help predict prognosis and guide treatment in patients with LUAD. These six lncRNAs served as participants in the tumor immune memory response process.

PU-202

智慧医院 5G 专网模型构建与应用

任晓强[1] 杨振宙[1] 范晓乐[1] 路东方[1] 崔昇[1] 李春晖[1] 陈超[1] 方志远[2]
[1]山西省人民医院
[2]中国移动通信集团山西有限公司

1 引言

医疗系统网络的“5G 化”是提升 5G+智慧医疗创新的关键，从 2020 年开始，随着国家工信部、国家卫健委先后下发了《关于推动 5G 加快发展的通知》、《关于下达 2020 年度卫生健康标准项目计划的通知》、《关于进一步加强远程医疗网络建设的通知》、《“十四五”全民健康信息化规划》等重要政策文件，将开展 5G+医疗健康应用试点，促进 5G 在卫生健康行业的重点应用创新。

智慧医院的建设旨在利用信息技术提升医疗服务的质量和效率，改善患者体验。而 5G 专网作为一种创新的通信技术，具备高速、低延迟等优势，为智慧医院的发展提供了良好的基础。对于智慧医院 5G 专网模型构建及应用的研究，国内外学者已取得了一定的成果。

国内方面，文献[1]中的研究聚焦于基于 5G 本地区域网络的智慧医院定位与导航系统，提出了相关的设计和实现方案。文献[2]则探讨了融合边缘计算的 5G 专网支撑下的远程麻醉机器人系统，该研究关注了远程医疗领域的创新应用。而文献[3]的研究内容为融合 5G 与边缘计算的智能医疗系统架构，这一研究在智慧医疗领域探索了结合 5G 和边缘计算技术的系统架构。这些文献提出了相应的系统设计和实现方案，并通过实验室环境或局部应用场景进行了初步验证。然而，目前国内的研究还相对较少，尚处于探索阶段。

国外方面，文献[4]主要描述了一个基于 5G 的智能医疗系统框架，并提到了相关的驱动场景以及安全挑战。文献[5]和[6]分别关注了 5G 与软件定义网络以及物联网在智慧医疗领域的应用。这两篇文献都提到了通过 5G 和相关技术实现智能医疗系统的潜力，以及边缘计算在智慧医疗中的作用。这些研究大多还处于理论探索和概念验证的阶段，需要进一步的研究和实践来实现实际落地和商业化应用。

目前国内外相关医疗机构等[7-11]，都重点围绕 5G+智慧医疗应用开展创新拓展，并取得一定的成绩，例如：5G+远程会诊、5G+手术示教、5G+VR 探视等。但所有的应用创新都是远程辅助类，非核心业务的创新，同时，也产生了对于 5G 网络的安全性、可靠性、以及接入方式的疑问，主要体现在：接入方式非 5G（所有的 5G+医疗应用主要采用的都是 5G CPE（路由器）转 WIFI 方式，医疗终端实际仍采用 WIFI 方式接入）、安全性存疑（绝大部分的 5G+智慧医疗应用都在互联网环境下，而不敢真正实现 5G 网络与医院内网的互通）、可靠性存疑（对于 5G 网络连接的可靠性缺乏理论与实践支持，可否保持 7*24 的稳定运行）。随着 5G 技术的不断成熟，特别是网络切片、SPN 传输、边缘计算、人工智能等应用在其他行业的广泛应用，解决医院终端“安全、可靠”的接入 5G 网络，“高效、便捷”的访问医院核心业务，是真正实现 5G“赋能”医院的关键，对于推动全国的 5G+智慧医疗创新有着重要的意义，且引领作用重大。

解决医疗终端设备“高效、快捷”的依托 5G 网络“安全、可靠”的接入医院的核心业务系统，是 5G 技术赋能智慧医疗的关键。实现智慧医院的全面升级需要强大的网络基础设施支持。这就需要 5G 专网的构建和应用，以满足大规模数据传输、高清视频传输、远程诊疗等医疗需求。因此，本研究旨在探讨智慧医院与 5G 专网的融合应用，解决智慧医院建设过程中面临的关键问题，提高医

疗服务质量和效率。

2 5G 医疗专网架构体系

2.1 5G 专网技术概述

与 4G 项目，5G 技术提供了更大的带宽，可以支持更高速的数据传输，从而实现更快的下载和上传速度。5G 网络具有更低的传输延迟，通信时延可以降低到毫秒级别，这对实时应用如远程手术、虚拟现实和自动驾驶等具有重要意义。5G 技术能够同时连接更多的设备，不仅涵盖移动终端设备，还可连接物联网设备和传感器，为智能城市和物联网提供了强大的支持。5G 具备更好的网络可靠性和稳定性，通过使用多种技术手段来保证通信的高质量和连续性，以满足医疗等关键应用的需求。

5G 专网采用了一种灵活的网络架构，以满足不同应用场景的需求，主要包括以下组成要素：

核心网（Core Network）：负责处理用户数据和控制信令，并提供服务管理功能。核心网包括核心网节点（Core Network Node）和数据网络（Data Network）。

网络切片（Network Slicing）：通过网络切片技术，5G 专网可以将网络资源按需划分为多个逻辑网络，以满足不同应用对网络性能和服务质量的要求。

边缘计算（Edge Computing）：边缘计算在 5G 专网中扮演重要角色，将计算和存储资源靠近终端设备，提供更低的延迟和更快的响应速度。

Massive MIMO：5G 专网利用 Massive MIMO（大规模多输入多输出）技术来提供更好的信号覆盖和更高的频谱效率，支持大量同时连接的设备。

虚拟化技术：通过网络功能虚拟化（NFV）和软件定义网络（SDN），5G 专网实现了网络资源的灵活分配和管理，提供更高的可扩展性和灵活性。

通过 5G 专网的架构和组成要素，智慧医院可以实现更快速、可靠和安全的数据传输，提供远程诊疗、远程手术、医疗设备的远程监控等创新医疗服务。此外，5G 专网还可以支持医疗设备和物联网设备的智能连接和管理，进一步提高医疗服务的效率和质量。

2.2 智慧医院的 5G 专网模型构建

2.2.1 智慧医院 5G 专网系统需求分析

智慧医院对于 5G 专网的需求主要包括以下几个方面：

（1）网络安全：智慧医院需要 5G 专网具备高级网络安全功能，以保护医疗数据和病人隐私的安全。这包括数据加密、身份认证、访问控制等安全机制，防止未经授权的访问和数据泄露。

（2）数据传输速率：智慧医院的医疗应用需要高速的数据传输能力，以支持实时的医疗数据传输和处理。5G 专网应提供低延迟和高带宽的网络连接，确保医疗数据的即时性和准确性。

（3）设备连接密度：智慧医院内部有大量医疗设备和传感器需要连接到 5G 专网进行数据采集和交互。因此，5G 专网需要具备高设备连接密度的能力，以支持同时连接大量设备，满足医疗设备的正常运行和数据采集的需求。

（4）应用场景需求：智慧医院涵盖多种应用场景，如远程手术、医疗设备的远程监控、医疗图像的传输等。对于不同的应用场景，5G 专网需要满足不同的服务质量需求，提供低延迟、高可靠性和大带宽等特性，以支持实时数据传输和处理。

综上所述，智慧医院对于 5G 专网的需求包括网络安全、数据传输速率、设备连接密度等方面的要求。为了满足这些需求，5G 专网需要提供高级的网络安全功能、低延迟高带宽的数据传输能力，并具备高设备连接密度的特性，以支持智慧医院的各类应用场景和服务需求。

2.2.2 智慧医院 5G 专网的设计和实施

基于 5G 专网的基础上，通过超级 SIM 安全网关产品为我院提供医疗终端的身份认证和业务访问能力，确保在 5G 医疗专网中实现让正确的人，使用正确的终端，使用正确的权限，访问正确的业务，获得正确的数据，确保我院 5G 专网服务的高效、稳定、可靠（如图 1 所示）。

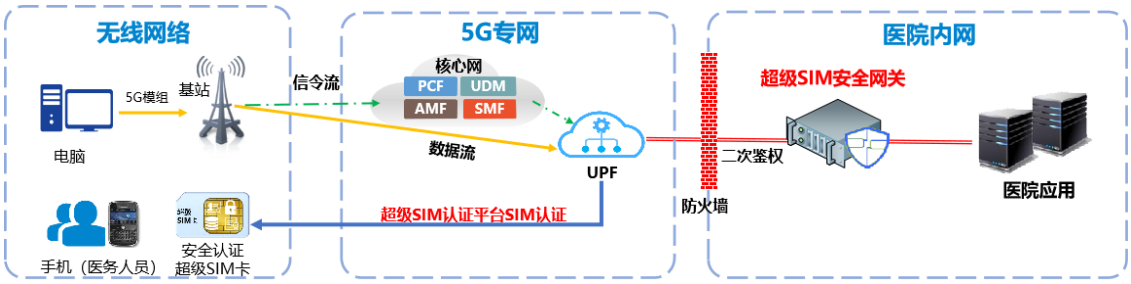


图 1 5G 专网+超级 SIM 安全网关方案示意图

5G 医疗专网的核心网与面向大网用户的 SA 网络相对独立，采用“无线基站共用，核心网独立部署，传输资源隔离”的方式组网，为医院客户提供一张隔离的、端到端高性能的专用接入网络（如图 2 所示）。为了保障用户业务安全，通过定制数据网络名称（data network name，DNN）来区分数据网络与路由隔离，保障用户业务体验，以及医院用户数据的绝对安全。

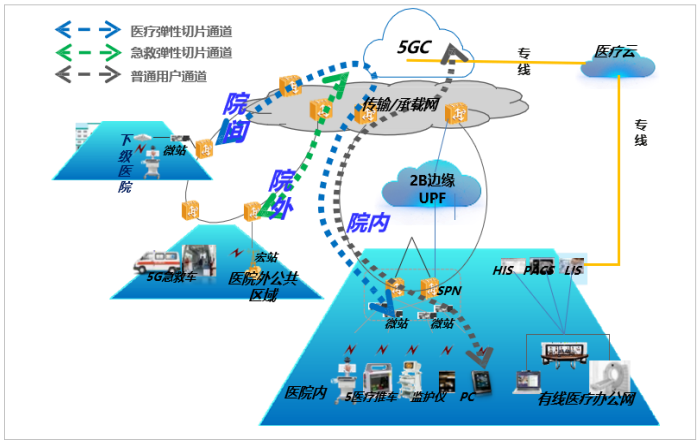


图 2 5G 医疗专网总体组网

3 智慧医院 5G 专网应用案例

3.1 医院 5G 专网实施方案

山西省人民医院移动终端 600 部，在医院机房内网区与内网前置区交界的防火墙旁挂部署超级 SIM 安全网关硬件，并在外网前置区部署一台代理服务器作请求转发到互联网，与超级 SIM 身份认证平台完成对接。

在完成设备网络对接后，方案计划整体分为 3 个阶段进行部署和测试。第一阶段为“连通测试”阶段；第二阶段为“功能测试验证”阶段；第三阶段为“生产试用”阶段。各阶段主要任务如下：

第一阶段——“连通测试”阶段，主要完成超级 SIM 安全网关和上下游系统（如 5G 专网客户端侧、内网业务系统侧、互联网出口等）网络联通、设备参数配置以及测试账号录入。

第二阶段——“功能测试验证”阶段，完成零信任安全网关各项功能场景（如 SIM 认证一键登录、连接稳定性、应用访问、SPA 端口隐藏、访问控制、审计等）进行测试验证。

第三阶段——“生产试用”阶段，主要将医院正式用户分批、逐步接入安全网关，正式开始分批次进行接入试用。

通过以上三个阶段，逐步推进完善，在尽可能的减少对现有生产环境的影响的条件下，完成网关部署和支撑试用工作。

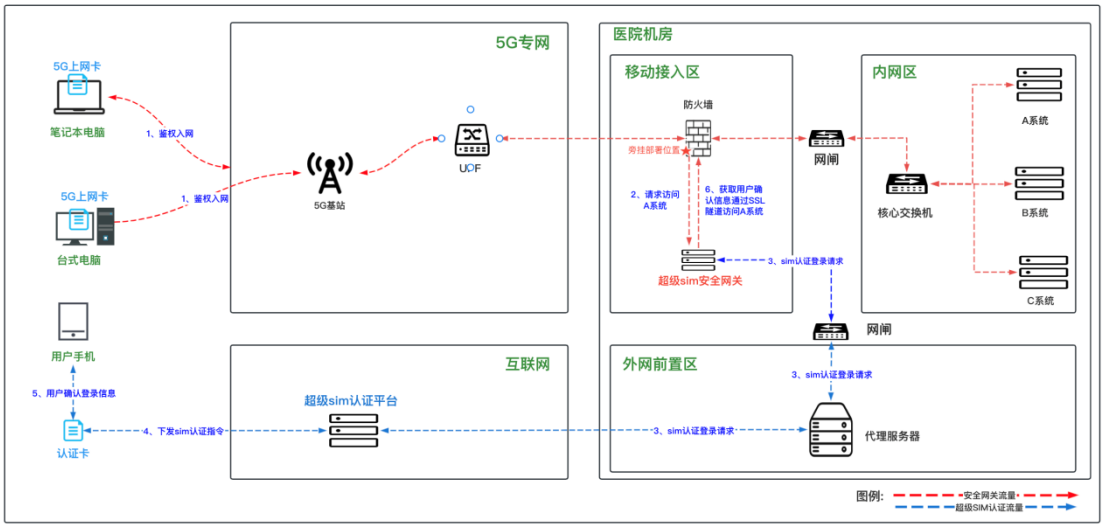


图 3 系统整体设计实施方案图

3.2 应用效果

3.2.1 性能对比

医院传统的网络建设方案主要采取综合布线+WLAN 方式，将各个终端设备连接至医院的数据中心，以满足终端访问内网的需求。由于包括查线、布置线缆、调试交换机、线路连接等过程，不仅整体建设耗时较长，而且后续维护工作大。此外，随着医院移动设备的增多、覆盖范围的增大，内网 WLAN 暴露出覆盖范围不足、补点难度大、单 AP 接入能力限制、易被网络攻击等众多问题。工程实施难度大：施工穿墙布线复杂，施工难度大；覆盖范围不足：随着医院业务的不断拓展，对于无线覆盖范围，不断提出要求，受限建设之初 AP 点位设计的影响，对于零星需求，基本无法满足；单 AP 接入能力限制：单 AP 的接入能力有限，随着医院移动业务的不断增减，高峰期对于无线网络需求不断增加，易造成网络拥塞；存在安全风险：WLAN 受限于其技术的广播特性，易通过接入侧发起攻击，存在接入的安全风险；兼容性差：不同厂家设备不兼容，更换难度大。（如表 1 所示），本文从速率/时延、安全、频段、移动性、终端、部署模式、维护方式等方面对 5G 专网与传统“WLAN”方式进行对比，5G 专网方式均优于传统“WLAN”模式。

表 1 5G 专网与传统“WLAN”模式性能对比

指标	5G (5G 及 LTE)	WiFi (802.11ax 及 802.11ac)
速率/时延	速率：10-20Gbps，时延 1-10ms	速率：9.6Gpbs，时延 20-200ms
安全	5G 支持终端与网络的双向认证，信令和数据都加密，256 位秘钥，安全性高	WiFi 仅支持单向认证，安全性差，加密算法简单，易破解，过去几年安全问题频发
频段	频谱为运营商独有，统一规划，抗干扰能力强，提供更好的速率保障 5G 的单用户体验更好，有保障	Wifi 频谱是公共的，易被侵入，AP 间干扰严重速率和接入用户数没有保证，难以达到大带宽接入，QOS 差
移动性	无缝切换，不影响业务	没有切换保证机制，经常源 AP 掉线后新 AP 重新连接。
终端	支持 5G CPE（室内及室外），5G 上网卡，智能 PDA，5G 行业用模组的终端设备	智能手机，终端等支持 WiFi 标准的终端设备
部署模式	按需求，按 SLA 切片部署，端到端保障	企业部署 WiFi 专网，无法满足高可靠、低时延需求

维护	运营商负责维护	客户自己维护

3.2.2 经济效益

以中型三甲医院为例，采用全部建设 WLAN 无线模式（含三年维保费），再外加三年维保费总计六年，以六年为一个周期计算，六年期总支出约 400 万元；采用 5G 专网服务方式，按照医院 600 个无线终端计算，每终端租赁费为 500 元/年，且可按需调整，年服务费约为 30 万元，六年期总费用约为 180 万元；单个中型三甲医院综合考虑六年期可节省 220 余万元左右，且不需考虑设备老化维保及报废等因素，如广泛应用经济效益会更好。

3.2.3 社会效益

虽然 5G 赋能智慧医疗提出多年，但始终未能形成好的解决方案，本次利用 5G 技术解决医院移动终端内网接入，是我院与中国移动的一种创新尝试，具备安全性、可靠性双提升，具有极高的创新性，改变了原有 wifi 模式，解决了原有 5G 应用仅在医疗互联网端的应用，对于赋能医卫行业，既有非常重要的社会意义。

4 总结

通过本次项目的实施，山西省人民医院及医联体单位完成建设试点部署，构建一个完全依托 5G 网络的医疗移动专网，实现智慧医院 5G 专网无线基站共用，核心网独立部署，传输资源隔离的新型 5G 专网模式。同时，为全国打造了具有可复制的 5G+医院场景技术方案及模式，利用 5G 技术，丰富医疗应用的场景，实现 5G 与医疗业务的深度融合，此模式进一步创新了 5G 在卫生健康行业的重点应用，并可借鉴全国推广应用。

智慧医院 5G 专网模型的未来发展将聚焦于更广阔的网络覆盖和容量、边缘计算与人工智能的结合、虚拟和增强现实应用、无线医疗设备和传感器的发展、远程医疗与移动医疗以及数据安全和隐私保护等方面。这些趋势将促进智慧医院的数字化转型，提供更高质量和便捷的医疗服务。

PU-203

5G+WiFi6 医院物联网安全的探索与研究

史乐萌 郝晓伟 许锋波 王腾飞 宫晨曦 刘轲

河南科技大学第一附属医院

1 引言

中国医疗行业近年来发展很快，医疗服务水平有了明显提高，但需要解决的问题还很多。到医院就诊，在一些医疗条件相对较差的偏远地区，会遇到挂号排队时间长、看病等候时间长等问题比较严重。看病效率低、医疗资料收集困难等问题，在医院也同样存在。不过这些问题都将随着 5G+WiFi6 和物联网技术的发展和应用而得到有效的解决，医疗服务水平也将得到迅速的改善和提高。

物联网对医疗智能化的促进作用体现在以下几个方面：

1.优化咨询流程。智慧医疗利用物联网技术，让患者轻松就诊，患者只需动动手指，就能轻松完成挂号、交费、查报告等工作，免去了排队等候和重复交费的辛苦。同时，通过视频远程会诊、信息化转诊平台等方式，让偏远地区也能享受到优质的医疗资源。

2.实现管理智能化。物联网技术可以帮助医院实现智能管理医疗设备的工作，通过智能改造医院医疗设备、车辆、基础设施等资源，使医疗设备无线、物管可视化，帮助医院实现智能管理。

3.快速收集有效数据。各种医疗设备通过物联网卡接入智能医疗系统，能够对患者数据进行有效识别，并将信息反馈给处理中心，及时对信息进行智能分析和处理，快速将患者的精准数据提供给医生，为医生减轻工作负担。

4.提升医疗服务水平。IOT（Internet of Things 物联网）技术不仅可以帮助医院提高工作效率，还可以在潜力巨大的医疗领域帮助减少医疗开支和开支；还能为患者提供更加便捷的服务，让更多的患者能够快速就医，享受到健康关爱。这些都产生了积极的促进作用，促进了医疗服务水平的提高。

随着物联网系统数量和数据量的增加，物联网安全问题越发显得突出和重要，针对安全问题，医院在建设过程中运用了多种安全技术措施。以下具体介绍关于 5G+WiFi6 医院物联网安全探索与研究。

2 物联网和 5G+WiFi6 介绍

2.1 物联网基本概念

物联网就是通过不同的通信网络，将不同类型的终端设备进行连接。经过二十多年的发展，物联网技术已经渗透到社会的各行各业。物联网未来发展的方向是在物联网的基础上，由终端、移动网络、互联网融合发展而形成的移动互联网。移动物联网的主要特点是网络无处不在，覆盖无缝隙，通过 5G、WiFi6 等无线传输新技术实现万物互联[1]。

2.2 5G+WiFi6 技术安全分析

(1)5G 是第五代移动通信网。该技术旨在大幅提高无线网络的速率和响应能力。通过 5G 无线宽带连接传输的数据潜在峰值速度高达每秒 20 千兆(Gbps)。

(2)WiFi6 是由无线网网络标准 WiFi 联盟提出的命名规则，将 802.11ax 改为 WiFi6[2]，于 2019 年发布。WiFi6 相比传统的 WiFi 速度更快，性能更强，采用了与众不同的技术。WiFi6 最高速率可达 9.6Gbps，并优化了信号上行覆盖，完美契合了 NB-IOT(NB-IOT 窄带物联网)物联网。

(3)5G 在实际应用中引入虚拟化技术[3]、网络切片、边缘计算、网络能力开放等创新技术，采用全新架构和新安全协议解决安全威胁。

(4)WiFi6 使用了新的安全协议 WPA3，使用户在访问公共网络时，能够通过个性化的数据加密来增强用户隐私的安全性， 同时将密码算法提高到 CNSA 等级算法的 192 位，这使得黑客无法窥探用户的流量，很难获取隐私信息[4]。

3 物联网安全分析

3.1 物联网体系结构

IoT 架构一般以 IoT 属性为基础，分为感知层、传输层、网络层、应用层[5]。如下图所示的架构。

物联网的体系结构图

3.2 物联网安全分析

(1)感知层的感知设备种类很多，为了保证其安全性，目前主要是加密和认证工作，利用认证机制避免非法访问标签和节点。针对感知层加密，目前已有一定的技术手段实现，但仍需提高安全等级，以应对更高的安全需求。

(2)传输层主要对节点到节点的机密性进行研究，利用节点与节点之间的严格认证，确保端到端机密性；支持数据的安全传输，使用与密钥相关的安全协议。

(3)对感知层采集数据的计算、处理、知识挖掘，实现实时控制、精确管理、科学决策的是物联网三层结构中最顶层应用层。IOT 主要威胁是攻击者可以向终端发送假触发消息，触发终端误操作的假终端触发威胁。

4 物联网安全研究探索

4.1 加密认证措施

物联网感知层具有可认证性，医院物联网采用了非对称密码加密算法,既加密法使用一组成对的私钥和公钥对消息进行加密和解密,确保机密性和身份验证性且具有比较好的通信能力和计算能力。

4.2 物联网设备协议安全性

医院物联网设备选用 RFID+蓝牙+LoRa 网关形式进行组网。RFID 是通过逻辑方法增加标签安全机制的隐私增强技术，其逻辑机制是通过安全认证协议(基于 HASH 功能)来实现的。由于医院物联网设备分布在各个建筑物中，数据交互采用通讯距离更远、耗电量更少、互连能力更强的 LoRa 网关，而 LoRaWaN 是一种基于 LoRa 技术的通讯协议，其特点是长距离传输技术能够完成灵活弹性的网络建网，并且成本较低。LoRaWaN 标准提供了两个加密层，网络安全层和应用安全层，LoRaWaN 协议提供端到端强大的安全[6]，通过对网络层和应用安全层进行加密。

4.3 网络安全接入环境

医院部署了 5G+WiFi6 专网，WiFi6 主内、5G 主外，医院物大部分物联网设备联网依托于支持 WiFi6 协议的内网建立，物联网设备接入内网需手动搜寻内网无线 SSID，内网无线设置为隐藏模式，防止人为非法连接；内网 连接时需进行身份认证，输入无线密码，基于 WPA3 安全加密协议; 物联网设备接入内网后还需进行内网认证管理，将设备的 SN 码加入内网网管认证白名单且在交换机端口上绑定设备 MAC 地址，这样才能真正接入网内。而像智能机器人和智能医废垃圾车则是在设备内嵌入 5G 模块，并使用 5G 专网进行连接，5G+WiFi6 两者是高度互补，同时满足不同的网络需求，使得整个网络系统性能能更好的发挥。更重要的是医院内外网还部署了防火墙、网闸、数据库审计、漏洞扫描、态势感知平台、内网探针、外网探针、日志审计、杀毒系统等安全设备，保证了网内所有设备的安全。后期，医院还会继续部署更多安全设备尤其是针对物联网安全方面做更精细的防护。

4.4 5G+WiFi6 及物联网系统维护

医院安排专业工程师对 5G 设备、信号、WiFi6 网络及物联网平台服务、设备维护、系统信息安全维护等进行日常巡检，尤其在服务器、操作系统、防火墙、网络设备等安全方面会定期进行安全扫描，并在安全方面对系统进行巡检；人员权限的设置遵循最小授权原则；强化密码策略，输入 3 次密码出错会被锁定，需管理员确认解锁，用户密码需符合每 3 个月修改 8 位(含)以上的用户密码，以及复杂的字母数字组合；系统最新补丁的安装和重要文件的安全测试，都是在安装前定期安装的。每月对操作系统安全漏洞进行一次扫描，及时发现最新的安全问题，并通过升级、补丁、加固等方式予以解决。

5 结论

通过 5G+WiFi6 高传输能力，医院已经建立了高可用、高效率、高效益的物联网应用平台，并成为省内人工智能医疗废物平台的引领者，全省的人工智能医疗废物系统需要与医院平台对接。并且随着医疗技术服务及物联网技术的进步，就诊患者更容易获得好的就医体验和花费更少的费用，但在提供这些医疗服务的同时，必须优先考虑其安全性，如若出现物联网安全事件，那必将是重大医疗事故！物联网安全技术目前还不完全成熟，使用过程中会面临更多安全问题和挑战，因此还需在实践中不断探索和完善，保证医院物联网安全可靠运行。本文针对医院 5G+WiFi 物理网安全的探索与研究，希望对医疗行业提供一定帮助。

PU-204

5G+WiFi6 医院物联网安全的探索与研究

史乐萌
河南科技大学第一附属医院

目的 随着 5G、WiFi6 及物联网技术快速发展，医疗行业的数字化转型也全面提速，通过物联网技术来提升医疗服务水平、实现智能化管理。当然，5G、WiFi6 和物联网在广泛的应用中都存在着安全威胁，本文着重介绍了医院在物联网安全方面的一些探索和研究，希望能够在物联网安全方面为相关行业提供帮助。
方法 加密认证、更安全的协议、网络安全防护。
结果 保证了网内所有设备的安全，后期，医院还会继续部署更多安全设备尤其是针对物联网安全方面做更精细的防护。
结论 物联网安全技术目前还不完全成熟，使用过程中会面临更多安全问题和挑战，因此还需在实践中不断探索和完善，保证医院物联网安全可靠运行。

PU-205

智慧医院物联网建设的应用

付宇 郭伟
太和县人民医院

引言

随着物联网技术的迅猛发展，医院逐渐认识到在医疗保健领域应用物联网的广泛潜力。智慧医院物联网建设能够显著改善医疗服务质量、提高效率、降低成本，并为患者和医务人员提供更加优质的体验。智慧医院物联网的建设可以实现实时监测和远程监护。通过物联网设备收集患者的生命体征和医疗设备状智慧医院物联网的基本概念和关键技术并通过平台实时传输给医务人员，他们可以随时远程监测患者的状况。智慧医院物联网建设有助于资产管理和设备维护的优化。通过将医疗设备与物联网连接，可以实现对设备状态的实时监测、预测性维护和远程管理。现代医疗系统中，医疗服务从传统的基于医生手动操作及人工判断的方式转变为以信息化技术为支撑的新型医疗模式。智慧医院物联网建设是国家卫生健康委所倡导的提高医疗卫生服务质量、推动医疗健康良性发展的重要举措之一。它将通过智能化管理和信息互通，提高医疗机构的运行效率和服务水平，推动医疗行业的创新和发展。

1、智慧医院物联网的基本概念和关键技术

智慧医院物联网的基本概念是将医疗设备、信息系统和人员通过传感器、网络和云计算等技术手段连接起来，实现数据的采集、传输和分析。通过这种方式，医院可以实现设备之间的互联互通，提高医疗服务的效率和质量。智慧医院物联网的关键技术：传感技术：包括各种传感器的应用，如温度传感器、压力传感器、心电图传感器等，用于采集医疗设备产生的各种数据。无线通信技术：使用无线网络技术（如 Wi-Fi、蓝牙、Zigbee 等）将传感器采集到的数据传输到网络中央服务器或云平台，实现数据的实时监测和交互。数据处理技术：对从传感器获取的大量数据进行处理和分析，包括数据清洗、过滤、整合和挖掘等，以提取有价值的信息，并支持医院决策和临床治疗。这些关键技术共同促进了智慧医院物联网的发展，使其能够实现设备互联、数据共享和智能化管理，提升医疗服务水平和患者体验。

2、智慧医院物联网建设的优势

2.1 提升医疗服务质量

信息化技术中，物联网（IoT）是一种能够将多种设备、传感器、智能设备和其他互联设备无缝连接在一起并共享数据的技术。智慧医院物联网建设可以提供更高效的心电及电生理诊断、实时监测和快速反应机制，从而进一步提升医疗服务的质量和水平。通过物联网设备，医院可以实时监测患者的生理指标和健康情况 2。例如，心电图、血氧饱和度、动态血压、体温等数据可以通过传感器实时传输到医护人员的监控系统中 3。这使得医生能够及时发现异常情况并采取必要的措施，避免潜在的风险。物联网将医院内部各个环节之间的快速信息传递和响应机制。比如，当患者的生

命体征出现异常时，监测设备可以自动发出警报，并将信息发送给相关医护人员。物联网技术还能帮助医院优化资源配置、提高工作效率，从而提升整体医疗服务的质量和水平。

2.2 降低医疗成本

智慧医院物联网建设可以实现医院管理的数字化 4。通过物联网技术，可以对医院内部的各种设施和设备进行监测和管理。通过物联网技术可以实现自动化管理和监控系统，减少人工干预和管理的需要 5。例如，设备的远程监控和维护可以减少维修人员的巡检频率，降低设备故障带来的管理成本。此外，物联网系统还能实时收集和分析数据，提供决策支持，帮助医院优化资源配置和流程管理，进一步降低管理成本。物联网技术可以代替部分人工操作，从而减少医院在人力资源方面的开支。例如，在药房管理中，物联网系统可以自动监测药品库存并进行智能补货，无需人工手动盘点，节省了人工成本。物联网建设可以提高医疗设备的利用率和效率 6。通过实时监测设备状态和预测性维护，可以减少设备故障导致的停机时间，避免生产损失和额外维修费用。此外，物联网系统还能帮助医院实施能源管理，优化用电和供暖系统，降低能源消耗成本。物联网技术可以实现资源利用优化设备共享和资源协同利用 7。医院内的各种设备可以通过互联互通，根据需求进行智能调度和共享使用，减少资产闲置和浪费。通过数据分析和预测模型，医院可以更好地规划和利用人力、设备和物资资源，提高效率并降低成本。

2.3 提高医患关系

智慧医院物联网建设可以让患者更好地了解自己的健康状况，增加患者对医院和医生的信任感和归属感。建立个性化慢性病健康管理计划，基于物联网技术收集的健康数据，医院可以为患者制定个性化的健康管理计划。这些计划可以根据患者的特定情况和需求进行定制，包括饮食、运动、用药等方面的指导。通过提供量身定制的慢病健康管理方案，医院能够更好地满足患者的需求，增加患者对医院的归属感。物联网技术使得远程医疗服务成为可能，患者可以通过互联网与医生进行视频咨询和诊断。这不仅提供了更便捷的医疗服务，同时也增加了患者对医生的信任感。患者可以在家中或其他舒适的环境中接受医疗咨询，减少因出行而带来的不便与压力。通过移动应用或在线平台，患者可以获取相关的健康信息、预防措施和治疗建议。这使得患者能够更好地了解自己的疾病和治疗方案，并能够主动参与到自己的慢病健康管理中去。这种积极参与的过程增强了患者对医院和医生的信任感和归属感。

2.4 提升患者体验和服务

利用物联网技术，可以实现患者服务的智能化，带来许多便利和改进。以下是一些具体应用：智能导航系统：医院内安装智能导航系统，利用感知设备和定位技术，为患者提供准确的导航和指引，使他们能够更快地找到目的地，如诊室、检查室等。这有助于简化就医流程，减少迷路和不便。（2）移动终端应用：通过移动终端设备，患者可以方便地获取医疗信息、预约挂号、在线问诊等服务。他们可以使用手机或平板电脑浏览医院的官方网站或专用应用程序，查询医生的资料、排班信息、就诊指南等。同时，患者还可以使用移动终端进行线上挂号，避免排队和耗费时间。（3）远程监护和健康管理：物联网技术可以实现患者远程监护和健康管理，通过连接医疗设备和传感器，将患者的健康数据传输到医院或医生的系统中进行监测和分析。这样，医生可以随时查看患者的健康状况，提供远程诊断和指导，减少不必要的门诊就诊。（4）智能医疗设备：物联网技术还可以将智能化应用于医疗设备，如药物管理系统、智能床位等。通过与网络连接，这些设备可以自动记录和传输数据，提醒护士或医生进行相应操作，提高医疗过程的效率和准确性。

3 智慧医院物联网建设面临的挑战

3.1 数据安全与隐私保护问题

智慧医院物联网建设需要处理大量的个人健康信息和敏感数据，因此数据的安全性和隐私保护是一个重要的问题。数据加密和传输安全，为确保数据在传输过程中的安全性，智慧医院物联网建设应采用数据加密技术来保护患者的个人健康信息。通过使用安全协议和加密算法，可以有效防止未经授权的访问和数据泄露。加强身份验证和权限管理，对于使用智慧医院物联网系统的人员，应实施严格的身份验证和权限管理措施。根据不同角色和职责，设置适当的权限级别，最小化数据被滥用或误用的风险。智慧医院物联网建设需要遵循相关的法律法规和隐私保护规定。敏感数据的匿名化和脱敏处理：为最大程度保护患者隐私，智慧医院物联网建设可以采用数据匿名化和脱敏处理技术。通过对个人身份信息和其他敏感数据进行去标识化或加密处理，将数据与特定个体分离，从而降低数据被关联到具体个人的风险。

3.2 技术标准问题

由于医疗设备种类繁多，不同厂商的设备之间存在兼容性问题，目前智慧医院物联网建设缺乏

统一的技术标准和规范，这会导致各种设备和系统之间的互操作性问题。互操作性问题导致数据传输错误、信息延迟、设备故障等一系列问题。例如，不同设备厂家的通信协议，使得它们无法进行有效的数据交换。设备的接口和数据格式也可能不一致，从而造成数据解析和解释的困难。缺乏统一的技术标准和规范还会给智慧医院物联网的安全性带来风险。如果设备和系统之间的通信未经适当的加密或认证，那么就可能存在数据泄露和篡改的风险。而且，由于缺乏规范，可能存在漏洞和弱点，使得恶意攻击者有机可乘。智慧医院物联网需要制定统一的技术标准和规范。这将有助于确保设备和系统之间的互操作性，并提高整个系统的安全性。此外，制定统一标准还能促进不同供应商之间的合作和竞争，推动技术创新和发展。在制定统一标准和规范时，需要广泛征求各方的意见和建议，包括医疗专业人员、技术专家、设备制造商等。只有通过共同努力，才能建立一个稳定、可靠、安全的智慧医院物联网系统，为患者提供更好的医疗服务。

3.3 技术更新和维护

定期更新是必要的。物联网系统中使用的软件、固件和操作系统等都需要及时更新，以修复已知漏洞、增加新功能或改进性能。这可以防止潜在的安全问题，并使系统保持与最新技术的兼容性。系统维护也非常重要，物联网设备和传感器需要进行定期检查和维护，以确保其正常运行。这包括更换电池、清洁设备、校准传感器等工作 8。此外，还需要对系统进行性能监测和故障排除，及时发现并解决潜在问题。

3.4 人员培训问题

智慧医院物联网建设需要有专业的技术团队来设计、部署和维护。但是，目前医院中缺乏相关的人才，很难找到具有相关技能和知识的人员。加强培养内部人才，通过提供培训课程、培训项目来实现。此外，还可以鼓励员工参加相关的行业研讨会和培训活动，以增加他们的知识和技能。医院可以与技术公司、专业服务提供商或物联网解决方案供应商建立合作关系。这些合作伙伴通常具备丰富的经验和专业知识，在物联网领域有着深入的了解。通过与他们合作，医院可以利用他们的专业团队来支持建设和维护物联网系统。

4 医院物联网未来发展趋势

4.1 边缘计算和边缘智能

随着物联网设备数量的增加和数据量的增长，边缘计算将在医院物联网中发挥更重要的作用。通过在设备端进行数据处理和决策，可以降低数据传输和处理的延迟，提高响应速度和效率。

4.2 人工智能和机器学习

医院物联网将与人工智能和机器学习等技术相结合，实现更智能化的应用。例如，利用机器学习算法对大量患者数据进行分析，可以提供更准确的诊断和治疗方案。

4.3 可穿戴设备和生物传感器

可穿戴设备和生物传感器将成为医院物联网的重要组成部分。这些设备可以实时监测患者的生理参数，并将数据传输到医疗系统中，实现个性化的健康监护和病情管理。

4.4 云计算和大数据分析

云计算将提供更大存储和计算能力，支持医院物联网系统处理海量的数据。大数据分析可以从庞大的数据中提取有价值的信息，帮助医院进行决策和改进医疗服务。

4.5 无线通信技术的发展

无线通信技术将不断提升，例如 5G 技术的普及，将为医院物联网提供更稳定、高速的网络连接，支持实时数据传输和远程监控。

5 结论

医院物联网建设推动医疗卫生行业的创新和发展。通过物联网技术的应用，可以实现医疗数据的互通共享，促进医疗资源的合理配置和利用，推动医疗卫生服务的整合和优化，提高医疗质量和效率。医院物联网建设是国家卫生健康委所倡导的提高医疗卫生服务质量、推动医疗健康良性发展的重要举措之一。它将通过智能化管理和信息互通，提高医疗机构的运行效率和服务水平，推动医疗行业的创新和发展。

PU-206

基于机器学习的乳腺癌治疗靶标生物活性预测

唐圣晟，周毅*
中山大学中山医学院

乳腺癌是全球发病率最高的癌症之一，对人们的生命健康构成严重威胁，每年夺走成千上万的生命。研究表明，雌激素受体 α 亚型（ERα）在乳腺癌的发生、侵袭、转移、化疗和内分泌治疗耐药中起着关键的调控作用[错误!未找到引用源。]。目前，内分泌疗法通过抑制 ERα 介导的雌激素信号途径来治疗乳腺癌，其中三苯氧胺是常用的药物[错误!未找到引用源。]。然而，抗药性的存在限制了该方法的有效性，因此需要研发新型药物来治疗乳腺癌，其中抑制 ERα 活性的化合物具有潜在价值。

计算机辅助药物研发已经成为药物化学家的重要工具，它不仅可以计算化学分子结构，还可以针对特定疾病靶标设计有效的化合物[错误!未找到引用源。错误!未找到引用源。]，从而减少了寻找有效化合物的时间并降低了动物实验和临床试验的次数[错误!未找到引用源。]。其中，定量结构-活性关系（QSAR）模型是目前最常用的药物设计方法。QSAR 模型通过数理统计的方法，量化分析某一特定靶标的生理活性与一组化合物的物理化学特性之间的关系。

为了开发治疗乳腺癌的药物，需要建立具有高预测能力和逻辑解释能力的 QSAR 模型，用于预测具有更好拮抗 ERα 活性的化合物分子结构。

在构建 QSAR 模型方面，多重线性回归(Multiple Linear Regression, MLR)、偏最小二乘法(Partial Least Squares Regression, PLS)、随机森林（Random Forest,RF）和支持向量机（Support Vector Machines ,SVM）等算法已经展现出良好的效果[6-10]。例如，Ahamed 等人使用 RF、SVM 和 MLR 等方法，开发了苯醌衍生物与 5-LOX 抑制活性的关系模型，其中 RF 模型表现最佳[6][错误!未找到引用源。]。李姜等人计算了药物之间的 Tanimoto 系数，并使用 SVM、朴素贝叶斯和逻辑回归构建了组合药物预测模型，其中 SVM 表现最佳[7]。

因此，建立高质量的 QSAR 模型对于预测具有更好拮抗 ERα 活性的化合物分子结构，对开发治疗乳腺癌的药物具有重要意义。

1 资料与方法

1.1 数据来源

从 DrugBank 药物分子数据库收集到不同化合物对乳腺癌治疗靶点 ERα 的生物活性数据，包括了 1974 种化合物的 729 个分子描述符和及其对应的 ERα 生物活性值。其中，化合物对 ERα 的生物活性值是通过实验值 IC50 记录的，其单位为 nM。该值越小，生物活性越大，抑制 ERα 活性越有效。但在建模中，常常使用它的负对数 pIC50 来表示生物活性值。其与生物活性呈正相关，即 pIC50 值越高，生物活性越高。

1.2 特征选择方法

1.2.1 Boruta 算法

Boruta 算法的核心思想是将数据中的真实特性变量和被随机加入的变量（也就是所谓的阴影变量）进行统计。具体操作流程如下：

(1) 建立阴影变量：对真实特征矩阵 R 进行随机的排序，获得一个阴影特征矩阵 S，并将其嵌入到特征矩阵 R，形成一个新的特征矩阵 N=[R, S];

(2) 通过引入新的矩阵 N， 利用如 RF 等能够输出特征重要性的模型对其进行训练，从而获得各个特征的重要性；

(3) 将阴影特征的重要性最大值作为阈值，而将真实特征中低于该值的则视为无关紧要的特征；

(4) 去除无关紧要的特性，并重复上述步骤，直到满足要求。

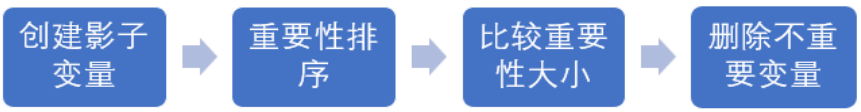


图 11 Boruta 算法流程图

1.2.2 基于学习模型的特征排序法

特征排序法是利用机器学习方法，通过对单个特征和因变量的分析，建立相应的预测模型，从中比较各变量重要性。如在随机森林中做预测，具体方法如下：

(1) 对于每棵决策树，使用袋外数据计算误差，记为$Err1$；

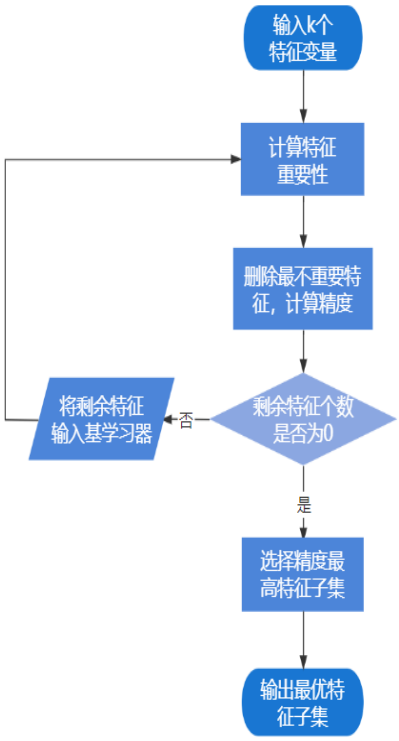
(2) 对袋外数据中全部数据样本的特征i随机加入噪声，重新计算误差，记为$Err2$；

(3) 假设有N棵树，则特征i的重要性为$\frac{\sum Err2 - Err1}{N}$。

如果当随机噪音增加时，袋外数据的准确度显著降低，表明此特性会对预测的效果有很大的影响，从而显示出它的重要性。

1.2.3 递归特征消除法

递归特征消除（RFE）是利用基模型进行多次训练，每次训练结束后，选出最差的特征，然后对其余的特性进行反复的训练，直至全部的特性都被迭代。



(1) 将待筛选的 k 个特征作为初值子集输入到基学习器中，对各个特征的重要性进行分析，并计算分类/回归精度；

(2) 从现有特征子集中删除特征值最小的特征子集，创建新的特征子集。 将其重新传回基学习器并计算新特征子集的重要性和分类/回归精度。

(3) 递归的重复步骤 2，直到特征子集为空，取 k 个不同特征个数的特征子集，选择准确率最高的特征子集作为最佳特征组合。

1.4 模型评价与模型解释

本研究采用可决系数和均方误差对模型的预测性能进行评估。可决系数（R Squared,R^2）也称拟合优度，它通过比值关系反映自变量对因变量变动的解释能力，其取值与 1 越相近，说明模型的效果越好（公式 1）。均方误差(MSE，Mean Squared Error)是一种常见的测量数值之间差异的度量，反映真实值与预测值之间的误差，故其取值越小说明两者的差距就越小，从而模型效果越好（公式 2）。

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=0}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=0}^{n}(y_i - \bar{y}_i)^2} \quad (1)$$

$$MSE(y,\hat{y}) = \frac{1}{n}\sum_{i=0}^{n}\|y_i - \hat{y}_i\|_2^2 \quad (2)$$

由于机器学习模型的欠解释性，采用 SHAP 方法对模型进行解释。SHAP 算法根据各特征加入模型时的边际贡献，计算各特征的重要性。取一个特征在所有特征组合下的不同边际贡献的均值，作为这个特征的重要性度量，即 SHAP 值。

2.建模

2.1 数据预处理

2.1.1 方差过滤

在对数据集查验的过程中，发现存在一些分子描述符的样本取值均为 0。从数理统计的角度来看，这些分子描述符的方差值较小，说明各个样本在该特征上彼此相似，未能体现差异信息，也就是说这些分子描述符对 ERα 的生物活性值的预测不能提供有用信息。因此，需要通过方差过滤法将这些变量进行过滤，减少数据维度，从而降低计算复杂度，提升模型的效能性。计算 729 个分子描述符的方差，根据阈值筛选出 225 个方差为 0 的分子描述符，将其剔除，得到剩下的 504 个变量。

2.1.2 相关性过滤

考虑到数据集中的分子描述符的类别，存在多个描述符共同描述同一性质的现象，其中各变量间可能存在高度的相关性。本文使用皮尔逊相关系数计算描述符之间的相关性，在相关性高于 0.9 的两个描述符中，保留与生物活性值相关性更高的描述符，删除另一个描述符，从而消除冗余和相关的特征。通过相关性过滤，数据集的 504 个分子描述符进一步减少到 237 个分子描述符。

2.1.3 数据归一化

在多维数据中的分子描述符计数单位各异，为了避免由于数据之间差异过大影响数据分析的结果，需要对数据进行归一化，清除各分子描述符之间的量纲，实现数据的一致性。

现假设x'_{ij}为标准化数据矩阵中第i个化合物的第j个分子描述符变量数据，其取值范围为[-1,1]，变量归一化公式如下：

$$x'_{ij}=\frac{x_{ij}-\min\limits_{1<i<n}(x_{ij})}{\max\limits_{1<i<n}(x_{ij})-\min\limits_{1<i<n}(x_{ij})},i=1,\cdots,1974;j=1,\cdots,729 \quad (3)$$

2.2 特征选择

变量的选择对机器学习模型至关重要，特征选择算法的应用有利于提升模型的预测性能，降低其计算复杂度。在药物分子学领域，许多的特征选择算法已被应用于选择最优分子描述符，本文将综合应用 Boruta 算法，基于学习模型的特征排序法，递归特征消除法（RFE）三种特征选择方法现实数据降维处理。

针对预处理后的数据集，利用由 Python 中的开源机器学习库 Scikit-learn 实现以下特征选择算法。先基于 XGBoost，ET，RF 模型分别应用 Boruta 算法和特征排序法进行变量筛选，得到各方法下的最优子集。而后，为同时发挥两种方法的优势，减少特征变量数量，以实现更少变量更高预测效果的目的，取上述两种方法的共同变量作为最终特征变量。

2.3 建立模型

本文主要评估了 XGBoost、LightGBM、RF 和极限树（Extra Tree, ET）、拉索回归(Lasso)、SVM、K 近邻算法(K-Nearest Neighbor, KNN)。

按照 7:3 的比例将已有数据随机分配为训练集和测试集。单次实验往往受样本随机性的影响，为确保模型的稳健性，使用 XGBoost 等七种机器学习算法的按照默认值进行 100 次训练，并采用 k 折交叉验证（k-fold cross validation）进行内部验证（k=10），计算并记录各评价指标。

3.结果

3.1 特征选择结果

应用 Boruta 算法进行筛选时，为避免单一算法带来的特征选择系统偏差，选取 XGBoost，ET，RF 共有的描述符作为该方法下的最优子集，得到 28 个分子描述符。而在特征排序法中，则是将 XGBoost，ET，RF 得到的三种特征变量重要性评分进行归一化处理后，再等权相加，得到综合评分。并根据最终的综合评分，选择 21 个分子描述符作为最终自变量。综合应用递归特征消除和交叉验证技术，验证了 21 个最终特征变量的选择合理性。经训练 ET，RF 模型来选择高排名特征，递归消除低重要性特征，发现该 21 个变量均为高排名特征，故作为最终特征变量。21 个最终特征变量名称，综合得分及所属类型如下（见表 3）：

表 3 最终特征变量

分子描述符名称	综合评分	所属类型
MDEC-23	2.74897	分子距边矢量
minsOH	1.22408	原子型电拓扑态
C1SP2	0.91764	碳元素类型

minHsOH	0.81373	原子型电拓扑态
LipoaffinityIndex	0.79412	原子型电拓扑态
maxsssN	0.65374	原子型电拓扑态
ndssC	0.45019	原子型电拓扑态
minHBint5	0.42150	原子型电拓扑态
nHBAcc	0.38806	受体数
SHsOH	0.23597	原子型电拓扑态
minssO	0.22870	原子型电拓扑态
TopoPSA	0.17456	原子型电拓扑态
MLFER_A	0.15884	分子线性自由能关系
MDEO-12	0.15256	分子距边矢量
VC-5	0.14759	气集群
maxHBint5	0.14252	原子型电拓扑态
MDEC-22	0.13783	分子距边矢量
maxsssCH	0.12735	原子型电拓扑态
SHBint10	0.12224	原子型电拓扑态
MDEC-33	0.11859	分子距边矢量
MDEC-34	0.11397	分子距边矢量

3.2 模型比较

各模型评价指标的平均值整理如下：

表 4 模型评价结果列表

模型	R2（测试集）	R2（训练集）	MSE（测试集）	MSE（10折）
XGBoost	0.71074	0.97874	0.56733	1.00265
LightGBM	0.73297	0.93957	0.51540	0.85948
ET	0.74051	0.98339	0.50366	0.86290
RF	0.73470	0.94993	0.51782	0.87796
Lasso	0.56188	0.57495	0.86679	1.03390
SVM	0.61986	0.65386	0.74978	0.95535
KNN	0.65610	0.77808	0.67207	1.16319

从评估指标来看，ET、LightGBM、RF 模型的拟合度更高，预测误差更低，具有较强的稳健性。相对来说，XGBoost 算法在模型外部拟合度却差强人意。而拉索回归，SVM ，KNN 的预测效果远不及其他模型。

3.3 模型优化与最优线性组合预测模型的建立

为了使模型能够更加准确可靠地预测新化合物的生物活性，利用网格搜索法对极限树、LightGBM、随机森林三种模型进行调参优化。在训练集上完成训练后，极限树的拆分内部节点所需的最小样本数为 3, 森林中的树木数量为 200；LightGBM 的学习率为 0.1，森林中的树木数量为 200；随机森林的最大深度为 19，森林中的树木数量为 600。再在训练集上利用改进后的三种模型进行 100 次训练，各评价指标平均值如下：

表 5 改进模型评价结果列表

模型	R2（测试集）	R2（训练集）	MSE（测试集）	MSE（10折）
极限树	0.75239	0.97815	0.48077	0.85058
LightGBM	0.73015	0.96627	0.52531	0.87322
随机森林	0.73747	0.95093	0.50775	0.86960

对比表 4 与表 5，可以发现三种模型的评价指标并没有明显改善，也就是说调整参数并不能较大程度地提高模型的拟合度和预测能力。考虑到单个模型的预测性能往往亚于组合预测模型，为进一步改进模型，决定将三种模型进行组合，构建最优线性组合预测模型，以实现更高的精准度。

分别将 3 种单项预测方法在训练集和测试集上的预测值进行合并，构成包含 3 种方法对 1974 个样本预测值的新数据集。设 ERα 生物活性值 pIC50 的序列为$\{x_t, t=1,2,\cdots,1974\}$，记第$^{\mathbf{i}}$种预测方法对第$^{\mathbf{t}}$个样本的预测结果为$x_{it}, i=1,2,3; t=1,2,\cdots,1974$。称$e_{it}=(x_t-x_{it})$为第$^{\mathbf{i}}$种单项预测方法的第$^{\mathbf{t}}$个数值的预测误差。

设$\hat{x}_t=l_1x_{1t}+l_2x_{2t}+l_3x_{3t}$为$x_t$的组合预测值，$e_t$为组合预测在第$^{\mathbf{t}}$个数值的预测误差，则有

$$e_t=x_t-\hat{x}_t=\sum_{i=1}^{3}l_ie_{it} \tag{4}$$

设J为线性组合预测模型中的预测误差平方和：

$$J=\sum_{t=1}^{1974}e_t^2=\sum_{t=1}^{1974}\left(\sum_{i=1}^{3}l_ie_{it}\right)^2=\sum_{t=1}^{1974}\sum_{i=1}^{3}\sum_{j=1}^{3}l_il_je_{it}e_{jt} \tag{5}$$

由此可得到以预测误差平方和为准则的最优化模型：

$$\begin{aligned}&\min J=\sum_{t=1}^{1974}e_t^2=\sum_{t=1}^{1974}\sum_{i=1}^{3}\sum_{j=1}^{3}l_il_je_{it}e_{jt}\\&s.t.\ \sum_{i=1}^{3}l_i=1\end{aligned} \tag{6}$$

设极限树预测值为x_{1t}、LightGBM 预测值为x_{2t}、随机森林预测值为x_{3t}，将各预测值代入上式求解得：

$$\hat{x}_t=0.932x_{1t}+0.191x_{2t}-0.125x_{3t}, t=1,2,\cdots,1974 \tag{7}$$

各回归系数均通过显著性检验，利用该回归模型在进行预测，模型在训练集上的 R^2=0.95，MSE=0.096，在测试集上的 R^2=0.90，MSE=0.179。根据线性组合模型在训练集上的预测结果绘制残差图和预测对比图如下：

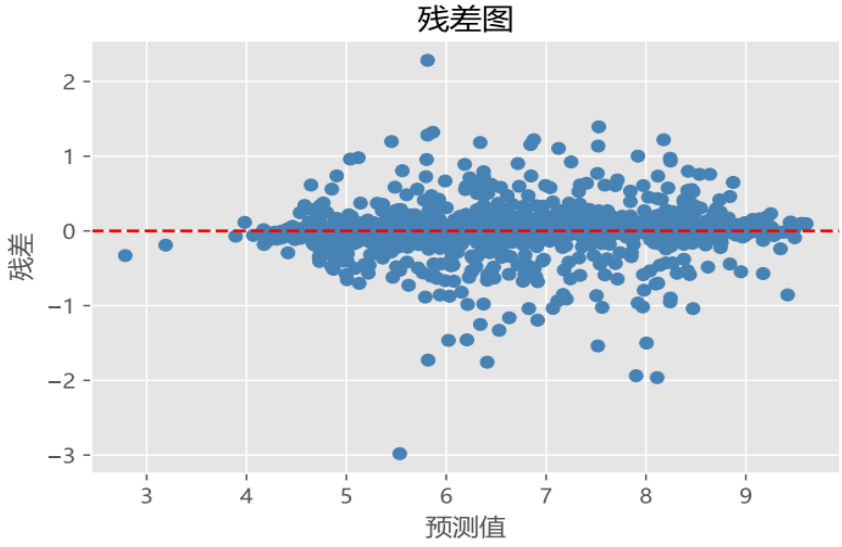


图 3 残差图

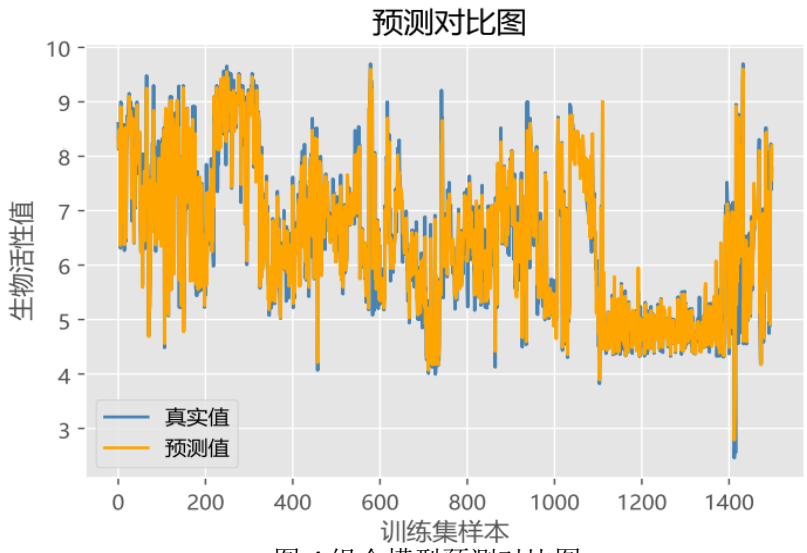


图 4 组合模型预测对比图

由图 3、图 4 可看出，在训练集上残差均匀的分布在 0 的附近，预测值贴近真实值，说明该模型有较强的预测能力。

3.4 模型解释

通过上述实证，不难发现无论是拟合优度，预测精度，还是从线性组合模型前的权重大小来看，极限树模型在各方面的性能更好，具有更强的稳定性。为更好的理解各分子描述符在极限树模型中预测生物活性的作用，利用 SHAP 算法对极限树模型进行可视化分析，见图 5，图 6。

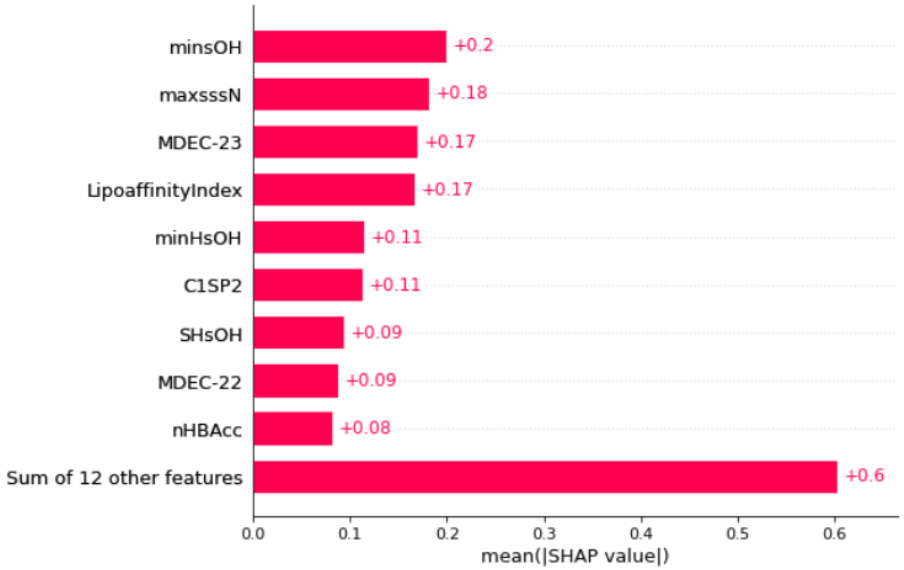


图 5 特征重要性 SHAP 值

由图 5 可以看出，在极限树模型中，minsOH,maxsssN、MDEC-23、LipoaffinityIndex 四种分子描述符对生物活性值的影响程度较大。其中 minsOH 的 SHAP 值最高，也就是说该特征在所有不同的特征序列情况下边际贡献最高。

结合图 6，图中横坐标是 SHAP 值，纵坐标是不同的分子描述符。可以看到，图中每一行都排布着蓝色或红色的点，这些点代表着不同的样本。其颜色越蓝表示该点的实际值越小，相反地，颜色越红表示该点的实际值越大。

特征的排序是按照图的特征重要性排列，其中宽的地方表示有大量的样本聚集。

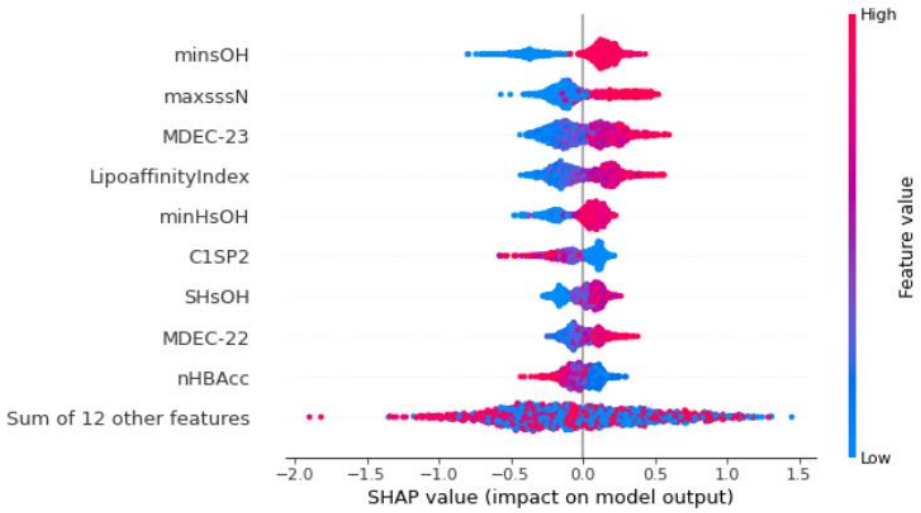


图 6 特征密度散点图

横向来看，minsOH 其高分红点值的化合物，SHAP 值大于 0，对 ERα 生物活性值产生正向影响；而低分蓝点的化合物，SHAP 值小于 0，则产生负向影响；排名第二位的 maxsssN 也会明显影响生物活性值，蓝色点主要集中在 SHAP 值小于 0 的区域，可见 maxsssN 值减小会降低生物活性。另外，比如排名稍靠后的特征 SHsOH，大多数的点弥漫在 SHAP 值为 0 的附近,表示该特征在大部分化合物中影响效果不明显，只对小部分化合物有影响。

纵向来看，特征排名在前四位的 minsOH,maxsssN,MDEC-23,LipoaffinityInde-x 四种分子描述符的样本分布较为分散，说明特征的影响程度较大。这些特征都是样本值越大，SHAP 值越大，样本值越小，SHAP 值越小，说明其取值与生物活性值 pIC50 成正比关系。

4.结论

本文针对乳腺癌治疗靶标 ERα，寻找到作用于该靶标的化合物分子描述符及其对应的生物活性数据集。利用 XGBoost， LightGBM，ET，RF，SVM，Lasso，KNN 七种机器学习模型将化合物的分子描述符作为自变量，ERα 的生物活性值作为因变量构建了 QSAR 模型和最优线性组合预测模型。根据高 R^2 低 MSE 原则比较不同的模型，发现相较于组合前的最优单预测模型 ET 的 R^2=0.75，MSE =0.48 而言，组合模型的 R^2=0.90，MSE =0.179 展现出较强预测能力，为后续的拮抗 ERα 生物活性的化合物预测提供基础。此外，在 SHAP 算法的帮助下，发现 minsOH、maxsssN、MDEC-23 对生物活性值 pIC50 产生正向影响，并在分子描述符中更大程度地影响预测结果。

综上，通过 QSAR 模型和最优线性组合预测模型的构建和 SHAP 算法下各分子描述符的影响程度分析，为乳腺癌药物研发中预测具有更好拮抗 ERα 生物活性的新化合物分子提供参考。

PU-207

基于深度学习的外展神经自动分割成像初步研究

孙钰茗 杨汉丰*
川北医学院附属医院

前言

外展神经是第六对脑神经，属运动神经，支配同侧外直肌控制眼球外展[1]。磁共振神经成像（magnetic resonance neurography，MRN）能提供高分辨率的神经图像，在周围神经疾病的诊断中发挥重要作用。但是由于外展神经自身直径纤细、周围毗邻关系复杂，对放射科医生的解剖学知识要求较高，而且人工重建是一个费时、费力的过程，重复性较差，不利于基于外展神经成像的进一步深度研究。因此，探究一个自动分割全段外展神经的方法，简化复杂、繁琐的重建过程，对于进一步进行外展神经相关疾病的研究十分重要。

目前人工智能辅助诊断技术已经在医学领域得到了广泛的应用。机器学习（machine learning，ML），尤其是深度学习（deep learning，DL）在医学图像分析领域发挥着越来越重要

的作用。因此，本研究拟通过深度学习方法建立外展神经自动分割模型，实现对外展神经进行全段自动分割，并与高年资放射科医师手动分割结果对比，验证模型效能，以期为未来进一步大批量研究奠定方法基础。

1 材料与方法

1.1 一般资料

本研究为回顾性研究，已通过本院伦理委员会的审查，并豁免知情同意。回顾性收集自 2021 年 5 月至 2022 年 12 月于我院行 MRN 检查的受试者。

纳入标准：（1）接受外展神经平扫+增强检查，且外展神经全程成像评分>3 分以上[2]（表 1）；（2）扫描范围从脑干至鼻尖。

排除标准：（1）外展神经走行路径周围病灶导致神经走行严重变化或显示不清者。（2）有外展神经相关疾病临床表现、阳性体征及病史者。

表 1 外展神经可视化评分标准（采用 5 分法[2]）

评分	评分标准
4 分	神经近段、远段均可识别，且连续性良好
3 分	神经近段、远段均可识别，但不连续
2 分	只有神经近段可识别，且连续
1 分	神经近段可识别，但不连续
0 分	神经无法识别

1.2 MRI 扫描方法

本研究所有受试者均使用 uMR 790 3.0T 和 Siemens Skyra 3.0T 进行扫描。扫描序列包括常规头颅磁共振扫描序列（轴位 T1WI、T2WI、T2-FLAIR、DWI 等）及磁共振神经成像序列（轴位 3D CISS、冠状位增强 3D SPACE STIR），后者具体扫描参数见表 2。

表 2 不同设备 3D CISS 及 3D SPACE STIR 序列参数

设备	序列	TR/TE (ms)	FOV(mm)	层厚（mm）	矩阵	码方向	励次数	翻转角（°）	带宽（HZ）
Siemens Skyra	3D CISS	5.16/2.2	180×180	0.7	320×307	F/L		30	460
	3D SPACE STIR	4000/225	256×256	1	256×218	S/I		120	55.8
uMR 790	3D CISS	1300/257.6	200×180	0.76	384×100	F/L		110	365
	3D SPACE STIR	4000/184	180×180	1	240×240	F/L		78	75

1.3 图像预处理

1.3.1 手动分割图像处理

本研究以手动分割为基础，由两位神经放射科医生独立进行外展神经全段分割。图像的分割在 3D Slicer 4.11（https://www.slicer.org/）软件上进行，绘制每位受试者外展神经从脑干到眼眶的示意图（图 1）。（1）从影像归档和通信系统（picture archiving and communication system，PACS）中导入并加载增强 3D SPACE STIR 及 3D CISS 图像。（2）将所有受试者的 3D SPACE STIR 及 3D CISS 图像进行配准，然后调整图像角度尽可能清晰地显示外展神经。（3）利用 3D Slicer 软件的分割工具，分别在轴位和冠状位对外展神经颅内段及颅外段进行分割。外展神经分割的准确性由两位放射科高年资医师以 3 分法进行评估：0，差（在临床实践中不可用，不纳入研究）；1，可接受（需要轻微修订，修订后纳入研究）；2，良好（几乎不需要修订，直接纳入研究）。需修订者由两位医生共同修改并达成一致后纳入研究。

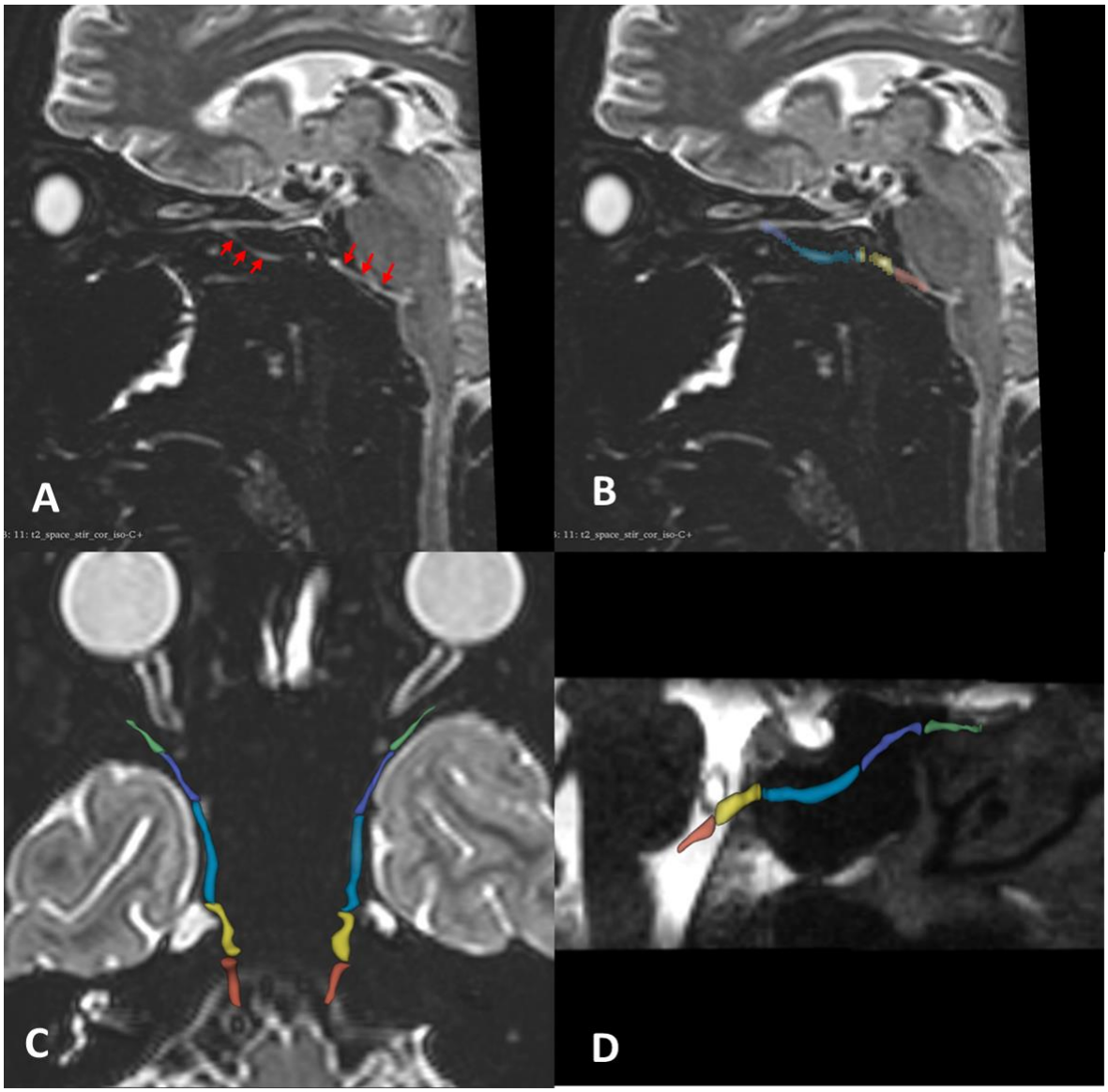


图 1 放射科医师在 3D Slicer 软件手动勾画双侧外展神经

（A）增强 3D SPACE STIR 序列上呈高信号的外展神经脑池段及海绵窦段；（B、C、D）在 3D Slicer 软件进行外展神经全程标注，脑池段（红色）、岩斜段（黄色）、海绵窦段（蓝色）、眶上裂段（紫色）及眶内段（绿色）。

1.3.2 深度学习图像预处理

首先，对所有图像应用 N4ITK 偏置场校正算法，减少噪声带来的偏差。然后对放射科医生手动分割的外展神经进行基于连通性的拆分，将原来的双侧外展神经分割为左右两条外展神经，不仅可以使训练数据量翻倍，而且减少了训练过程中对 GPU 的内存消耗。在训练过程中，采用数据增广的正则化方式，减少在模型训练的过程中由于样本数据的原因陷入过拟合的问题，有助于提高模型的泛化能力及稳健性。最后，对数据进行归一化处理。

1.4 研究方法

1.4.1 外展神经分割模型总体框架

本研究采用基于语义分割的端到端网络模型，这是一种三维分割方案。以 V-Net 网络为基线，引入 BottleNeck 和 SE-Net 来优化 V-Net 网络结构，并将这种结构命名为 SEVB-Net（SE 代表 SE-Net，B 代表 BottleNeck）。（详见图 2）

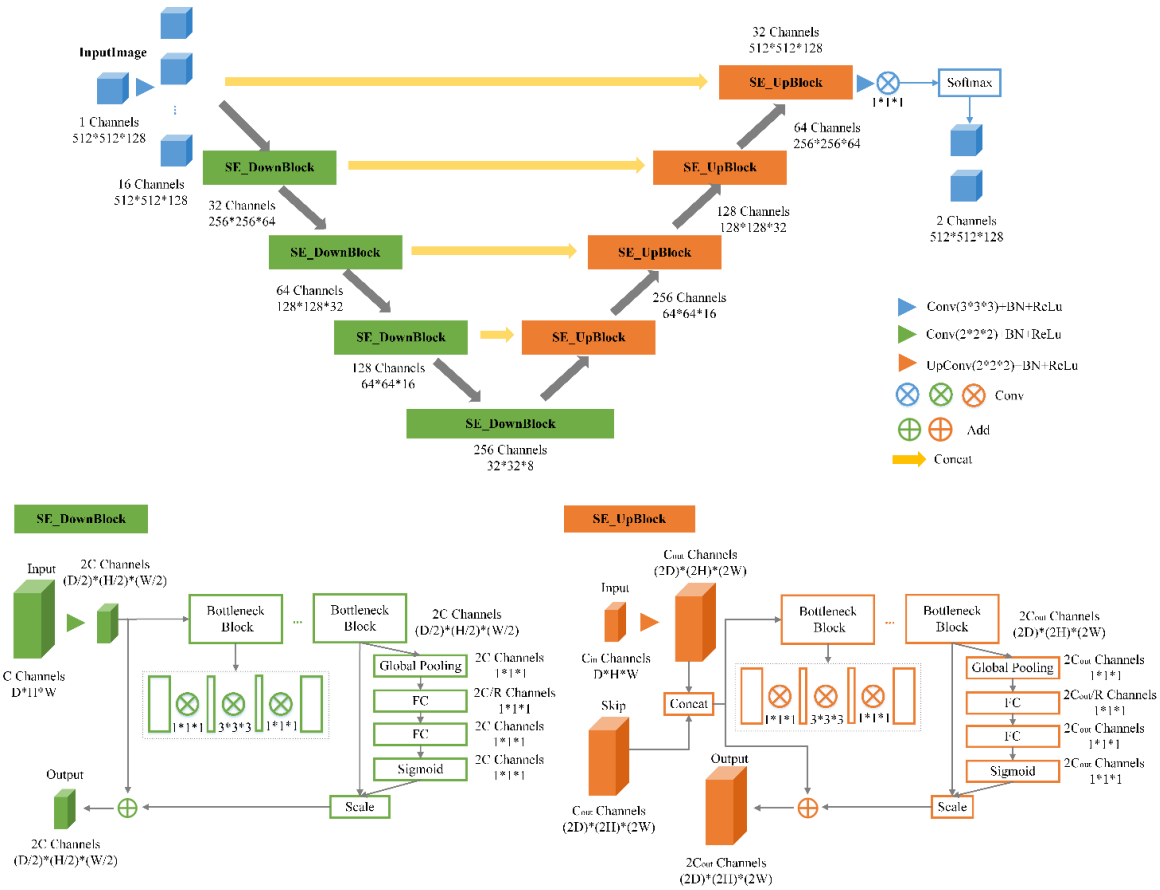


图 2 SEVB-Net 网络体系结构

BottleNeck 由三个卷积层组成，中间 3×3×3 的卷积层首先在一个降维 1×1×1 卷积层下减少了计算，然后在另一个 1×1×1 的卷积层下做了还原，在保持精度的同时大幅降低了模型的计算量。SE-Net 选择性地强调有用的信息特征而抑制不太有用的特征，提高了准确率。

1.4.2 训练过程

由于三维医学图像尺寸较大，为了减少内存大小，我们采用多尺度级联结构进行训练，可以有效减少存储空间。首先，通过粗略分割阶段生成不同解剖结构的初步置信图；其次，提出细化分割阶段，对粗略分割图进行细化和优化，以准确地分割出外展神经的边界。（详见图 3）

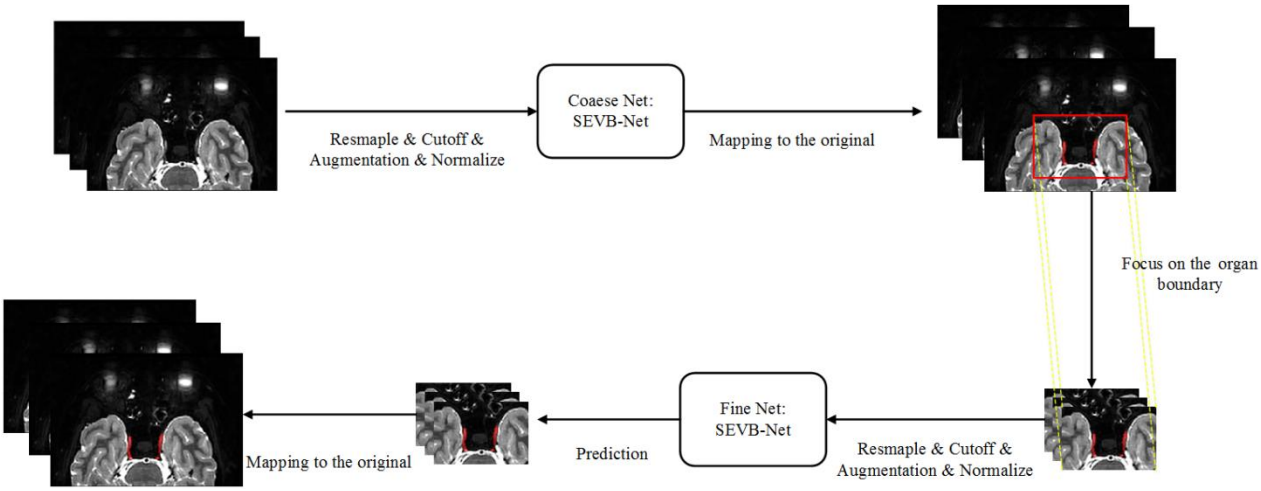


图 3 多尺度级联结构网络流程图

1.4.3 损失函数

在本研究中，我们使用一种混合损失函数来评估模型的预测值和真实标签的匹配程度，这种模型由骰子系数损失函数（L_{dice}）和焦点损失函数（L_{focal}）构成。L_{dice}[3]主要用来处理数据分类不均衡的情况，这些不均衡在很多情况下是由于背景和待分割区域之间的面积对比不均衡导致的。L_{focal}[4]从样本难易分类角度出发，解决样本非平衡带来的模型训练问题，降低了大量简单负样本在训练中所占的权重，从微观角度对像素逐个绘制，与 L_{dice} 互补。该混合损失函数，通过调整平衡参数的取值，在训练时尽量平衡敏感度和准确率。

1.4.4 评价指标

本研究采用 Dice 相似系数、精确度及灵敏度作为评价指标来评价外展神经的分割性能。上述指标的计算公式如下：

$$\mathrm{DSC（F1Score）} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \quad （1）$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad （2）$$

$$\mathrm{Sensitivity(Recall)} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad （3）$$

1.4.5 统计学分析

本研究采用 SPSS 25.0 软件进行数据分析。使用 Shapiro-Wilk 检验评估数据的正态性。符合正态分布的计量资料以均数±标准差（ $\bar{x} \pm s$ ）表示。使用配对样本 *t* 检验比较在自动分割模型辅助下两位医师对每例双侧外展神经进行手动校正所需的平均时间与手动标注所需的平均时间。*P*<0.05 为差异有统计学意义。

2 结果

2.1 参与者的人口资料

本研究共纳入 107 例符合要求的受试者，平均年龄 53.65±9.56 岁；年龄范围 28-83 岁；47 名男性和 60 名女性。

2.2 模型迭代后分割能力对比

我们的模型经过一次版本迭代，分别在训练集为 68 例及 86 例时建立自动分割模型。以高年资放射科医师的手动分割作为金标准，使用 68 例训练集建立的模型在测试集数量为 21 例时，Dice 相似系数、精确度及灵敏度分别为 0.678±0.088、0.724±0.086、0.657±0.138；在增加训练样本量至 86 例后，得到的模型在相同测试集下 Dice 相似系数、精确度及灵敏度分别为 0.692±0.076、0.728±0.082、0.680±0.129，较前均有提升。（详见表 3）

表 3 外展神经 AI 模型分割性能参数

模型（训练集：测试集）	DSC	精确度	灵敏度
A（68:21）	0.678±0.088	0.724±0.086	0.657±0.138
B（86:21）	0.692±0.076	0.728±0.082	0.680±0.129

在图 4 中，我们展示了 1 例随机样本的可视化结果，包括放射科医师手动分割和模型自动分割之间的可视化结果对比。如图所示，在随机抽样预测中，我们的模型得到的结果与手动分割的结果在整体上相似，该模型可以在一定程度上在 MRI 图像中对外展神经进行初步自动识别和分割，辅助医疗诊断。

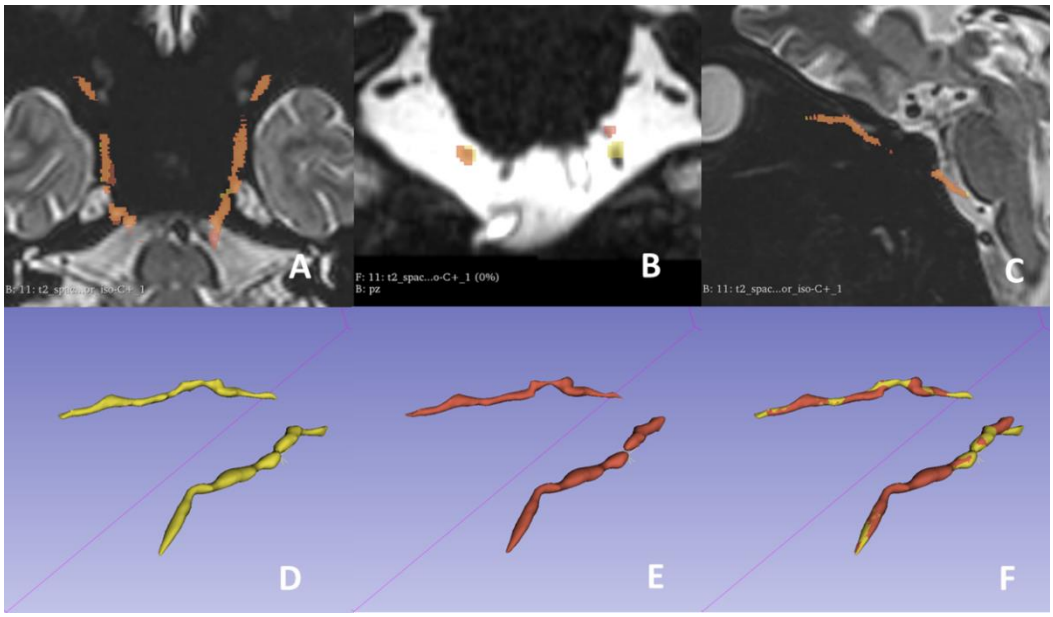


图 4 一个随机样本的分割结果

标准手动分割（黄色）、模型自动分割（红色）（A）轴位；（B）冠状位；（C）矢状位；（D）、（E）、（F）均为三维重建图。

2.3 手动分割与在自动分割模型辅助下进行手动校正所需时间分析

在验证集里随机抽取 10 例，利用该模型对双侧外展神经进行自动分割，然后上述两名医师对其进行手动校正并记录所需时间。两位放射科医师手动标注每例双侧外展神经所需的平均时间为 24.8±5.4 分钟；在自动分割模型辅助下进行手动校正所需的平均时间为 9.6±2.9 分钟；在自动分割模型辅助下进行手动校正所需的平均时间较放射科医师手动标注的平均时间明显减少，标注效率有所提高，差异有统计学意义（*P*<0.05）。（表 4）

表 4 手动分割及在自动分割模型辅助下进行手动校正所需的平均时间

	手动分割	自动分割模型辅助下手工校正	*P* 值
时间（min）	24.8±5.4	9.6±2.9	*P*<0.001

3 讨论

外展神经是第六对脑神经，根据其走行将外展神经离开脑干后分为五个节段，分别为脑池段、岩斜段、海绵窦段、眶上裂段和眶内段[5]，由于不同节段其周围微环境的差异较大，故采用不同的 MRI 扫描序列和成像方案。在脑池段，外展神经被脑脊液所包围，神经在重 T2 加权序列上能得到较好地显示，其中 3D CISS 是最为常用的序列。 增强 3D SPACE STIR 序列获得了高组织对比度的颅神经外周段 MRI 图像，为颅神经外周段的成像提供了方向[6]。

外展神经麻痹是外展神经最常见的病变，导致外展神经麻痹的原因复杂多样，微血管病变、脑血管疾病、肿瘤、创伤、炎症等都是其常见病因[7]。在外展神经全段走行上，任何部位的病变都可以导致外展神经麻痹的发生，对于病变的定位诊断至关重要。但由于外展神经直径纤细、走行迂曲，邻近血管、神经错综复杂，放射科医生必须仔细观察数百张图像，并对外展神经进行后处理重建以评估它们，这在很大程度上依赖于放射科医生的临床经验和熟练程度，因为它涉及后处理技术，是一个费时、费力的过程，并伴随着误诊或漏诊的可能。

目前人工智能辅助诊断技术已经在医学领域得到了广泛的应用。医学图像分割主要用于分割器官或病变，以定量分析相关的临床参数，为后续诊断和治疗提供进一步的指导。实现准确的器官分割是从患者的医学图像中找到病变区域的前提，器官分割的准确性将直接关系到病变区域的检测，而病变的划分对于疾病术前导航或肿瘤放疗指导等至关重要[8, 9]。

V-Net 是目前广泛应用于医学图像分割的一种三维图像分割方法，它吸收利用了残差的思想，在网络拓展路径和收缩路径的每个阶段中都引入残差学习的机制，加快收敛速度[10]。根据应用场景的不同特点，研究者们往往会有针对性地对网络进行改进，提高目标器官分割的精度和效率。

我们在 V-Net 的基础上，引入 BottleNeck 和 SE-Net 来优化 V-Net 网络结构。引入 BottleNeck 主要是为了降低计算消耗。该模块实际上是一种特殊的残差结构，将原通道数先减少，后增大到原来的大小，在保持精度的同时大幅降低了模型的计算量[11]。SE-Net 通过学习的方式自动获取每个特征通道的重要程度，在此基础上对重要特征进行强化，同时抑制对当前任务用处不大的非重要特征，该方法在不引入过大的参数量和计算量的情况下，可以很轻松的嵌入到现有网络中显著提升模型性能[12]。有研究在 U-Net 的基础上引入 SE-Net 后，发现模型在分割微血管时表现出优异的分割性能，尤其是血管的连续性有显著改善[13]。

大多数颅神经都纤细，且与周围软组织对比度不高，这使得颅神经的分割困难。Lin[14]提出了一种实现三叉神经及其周围脑血管的自动分割方法，并在三维磁共振血管成像中对这些解剖结构进行自动分割和重建，对三叉神经痛患者血管减压术前计划的制定具有重要指导意义。该方法包含两个阶段：粗分割和精细分割，前者用于获取三叉神经、脑血管系统和脑干的置信图，后者则根据置信图细化组织/器官的边界，该方法在模型的准确度上得到较好的验证。这与本研究所使用的多尺度级联结构类似。

一项多中心研究利用一种支持主动学习的阶梯式 3D nnU-Net 方法进行 CT 图像上的下颌管分割，以获得下牙槽神经走行，并通过 Dice 相似系数评价其准确性。该研究发现分割模型经过学习后 DSC 逐渐增加，该模型在有限数据的情况下可以提高训练效率、减少标记工作[15]。

一项研究提出一种自动检测和分割内耳道及其神经（如面神经、耳蜗神经、前庭下神经和前庭上神经）的方法，通过一种由 U-Net 与 Mask R-CNN 联合的方法来检测和分割内耳道及其神经。该方法在内耳道及其神经的定位和分割上取得了较好的效果[16]。

上述研究表明，基于深度学习的自动分割可以提高器官/病灶分割的准确性和分割效率，可以辅助医生诊断和制定患者的治疗计划，也是引导手术和临床治疗评价的主要方法。

在本研究中，结合随机抽取的样本的可视化结果及 Dice 相似系数等评价参数可以得知，我们的模型得到的结果与手动分割的结果在整体上相似，该模型可以在一定程度上实现基于 MRI 图像的外展神经自动识别和分割，并提高分割效率。

本研究也有一定的局限性。首先，作为深度学习的训练数据样本量较少，得到的模型准确性有待提高，在增加训数据集的数量后模型的性能可以得到进一步提升。其次，本研究为一项单中心研究，由于在深度学习过程中没有利用多中心数据集，因此无法实现人工智能学习阶段的多样性。通过比较模型在对不同公司和不同医院的设备获取的数据的分割效能，可以进一步确定需要改进的方向。

PU-208

基于深度学习面神经磁共振图像自动分割

曾祥鑫 杨汉丰*
川北医学院附属医院

1 目的

面神经是包含运动、感觉及副交感神经的混合神经，主要支配面部表情肌、舌前 2/3 的味觉及舌下腺、下颌下腺和泪腺的分泌。

面神经在骨管中走行较长，而且在颞骨内走行迂曲、血供比较脆弱，是颅神经中最容易受损的神经之一，所以对面神经的观察和定位较为困难且十分重要[2]。现在高分辨率 MRI 提供了神经精细的解剖细节，并可以观察面神经从脑池段至颌骨外段主干，由于面神经纤细、走行复杂，这对放射科医生的解剖知识和操作要求较高，人工重建不仅繁琐、耗时，而且重复性差，受主观判断的影响。因此，探究一个自动分割面神经方法，简化复杂、繁琐的重建过程，提高医生工作效率，以及对于进一步针对 pHFS 相关的研究十分重要。基于此，本研究

拟采用深度学习方法，构建面神经磁共振图像自动分割模型，实现面神经自动、高效分割，并与临床医生作对比，验证模型效能。

2 材料与方法

2.1 受试对象

本研究为回顾性研究，已通过本院伦理委员会的审查，免除受试者知情同意，收集 2018 年 1 月至 2022 年 11 月于川北医学院附属医院行面神经平扫+增强的检查者，符合下述纳入排除标准。

纳入标准：根据吴绯红学者在颅神经外周段磁共振神经成像研究中提出面神经可视化评分标准（采用 5 分法)（详见表 1），将面神经进入腮腺处定义为“中间点”，以此分为近段与远段，面神经磁共振图像可视化评分≥3 分。

排除标准：（1）面神经走形区域肿瘤等患者；（2）严重中耳乳突炎症干扰神经成像；（3）图像伪影严重。根据以上排除标准，最终纳入本研究的检查者图像共 100 例。

表 1.面神经可视化评分标准

评分	定义
0 分	面神经无法识别
1 分	只有面神经近段可识别，但不连续
2 分	只有面神经近段可识别，且连续
3 分	面神经近段、远段均可识别，但不连续
4 分	面神经近段、远段均可识别，且连续性好

2.2 检查方法

本研究纳入的检查者均由以下两台 3.0T 磁共振扫描仪(联影 uMR790、GE Discovery MR750) 进行扫描，仰卧位，头先进，扫描参数如下：

表 2.面神经磁共振成像扫描参数表

机型	序列	TR（ms）	TE（ms）	层厚（mm）	FOV（mm）	层数	扫描时间
联影 uMR790	增强 STIR-MX3D	4000	184	1	180×180	180	8 分 23 秒
GE Discovery MR750	增强 T2WI 3D CUBE fs	2500	Maximum	1	160×160	256	8 分 46 秒

2.3 图像预处理

将原始图像 DICOM 数据导入 3D slicer（版本 4.11.20210226）软件中，对图像进行重新格式化，以调整角度，为了提高人工分割的准确性，使分割图像尽可能垂直面神经走形方向，由经验丰富的放射科医师沿面神经走形逐层手动标注出面神经脑池段、内耳道段、迷路段、鼓室段、乳突段、颌骨外段主干（详见图 1），并记录分割每例检查者双侧面神经所需时间，分割结果经 2 名资深影像诊断医师审核后纳入，如对分割结果存在分歧，则由两名医师协商后共同决定。将 100 例分割图像按照 8:1:1 随机分为训练集、测试集及验证集。训练集及验证集用

于模型训练，测试集用于评估模型的分割性能，并将测试集自动分割图像重新导入 3D slicer 软件中，由标注医师进行图像修正，记录测试集每例检查者双侧面神经图像修正所需时间。

2.4 深度学习算法的应用

本研究采用上海联影医疗科技股份有限公司自主研发的一种基于 SE-VBNet 网络结构，在保持准确度的同时降低了模型的计算量，并且通过选择性地强化重要特征抑制非重要特征的，提高分割的准确率。首先通过粗分割网络实现面神经的粗略分割与定位。在粗略分割与定位的基础上，基于 SE-VBNet 与滑窗技巧实现面神经的精确分割，其中涉及到的主要技巧如下：（1）预处理：N4 偏置场校正，数据增广，包含随机旋转、缩放、平移、镜像、弹性变形、伽马校正等。（2）网络结构：SEVB-Net 在网络中引入通道注意力机制，对通道间的依赖关系进行建模，自适应调整各通道的权重，在增加很小的计算消耗的情况下可以极大地提升网络性能。（3）推理策略：用训练时的窗口尺寸遍历整张图像，每次只预测窗口内的图像，窗口滑动有重叠部分，边界重叠处的像素进行加权融合，这样可以在测试时最大限度地保持与训练过程的一致，进一步提升网络预测精度。（4）后处理：主要包含连通域分析（详见图 2）。

2.5 分割结果评价指标及统计学方法

本研究使用 Dice 相似系数来评价自动分割与人工分割之间的相似性。评价指标取值范围为 0-1，值越大，分割效果越好，计算公式如下，其中真阳性（True Positives，TP）、假阳性（False Positives，FP）、假阴性（False Negatives，FN）。本研究采用 SPSS 26.0 进行数据分析，采用独立样本 t 检验比较人工分割与自动分割后人工修正所需时间的差异。*P*<0.05 为有统计学意义。

$$\mathrm{DICE} = \frac{2\mathrm{TP}}{2\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \quad (1)$$

3 结果

3.1 检查者一般资料

本研究共纳入符合纳入排除标准检查者 100 例，男性 39 例，女性 61 例，平均年龄 53.12±10.80 岁，年龄范围为 28-75 岁，其中联影 uMR790 磁共振扫描仪检查者 62 例，GE Discovery MR750 磁共振扫描仪检查者 38 例。

3.2 深度学习模型准确性评价

本研究在评价深度学习模型时，以放射科医师人工分割作为金标准，基于 SE-VBNet 自动分割模型对面神经磁共振图像自动分割训练集、测试集及验证集 Dice 相似系数分别为 0.761、0.684、0.718（详见图 3）。

3.3 深度学习模型自动分割后人工修正与人工分割时间对比结果。

本研究随机选择 10 例检查者图像作为测试集，人工分割的平均时间为 34.4±3.66 分钟，明显大于深度学习模型自动分割后人工修正平均时间 9.90±2.28 分钟（*P*＜0.05）。

4 讨论

面神经是最复杂的颅神经之一，从脑干到颅外软组织走行复杂，特别是颞内段曲折而长，各种先天性的、炎性的、血管性的和肿瘤性病变可能影响每个节段，所以掌握面神经的复杂解剖对疾病定位是至关重要的。磁共振高分辨率各向同性成像序列是颅底面神经成像的最佳选择，各向同性的数据可以沿着面神经的走向进行多平面重建，有助于多方位观察面神经颞骨内段病变[30, 31]。本研究选择的增强 STIR-MX3D 序列及增强 T2WI 3D CUBE fs 序列，使富含钆对比剂的组织（如肌肉、血管丛、唾液腺等）在重 T2 加权的图像上呈明显低信号，而含有丰富结合水的外周神经纤维束在该序列上呈现的高信号，图像有更高对比度，神经周围背景抑制效果明显，对面神经脑池段至颌骨外段主干均能有较好的显示[32]。其次本研究使用不同厂家磁共振扫描仪及不同序列进行数据采集，所采集图像灰度具有差异，模型学习阶段数据多样性，这使结果更具有普遍性。

本研究基于 SE-VBNet 构建了面神经磁共振图像自动分割模型，该模型的分割结果与专家标注相符，显著缩短重建面神经所需时间，并且减少重建时主观因素的影响。与吴亚平等学者研究类似，该研究使用基于级联 VB-Net 分割模型对急性期脑缺血梗死磁共振 DWI 图像进行分割，Dice 相似系数为 0.76，该模型将 V-Net 中卷积层替换为瓶颈结构，其由三个卷积层组成，该方法可以大大减少参数的数量，从而减少计算量和模型的规模；降维后，可以更有效、直观地进行数据训练和特征提取，增维后的特征对当前任务更加具体；与传统的卷积结构相比，可以传输低层次的信息，使训练更深入、更容易[33]。本研究深度学习模型自动分割后人工修正时间明显小于人工分割，这与 Lim 等学者研究结果一致，该研究使用基于 3D nnU-Net 对三叉神经的外周分支下牙槽神经锥形束计算机断层扫描成像（cone-beam computed tomography，CBCT）进行分割，Dice 相似系数为 0.58，同时相对于人工分割，基于深度学习的自动分割减少了约 2/3 的分割时间，减轻了医师工作负担，提高了工作效率。

本研究仍具有局限性：本研究纳入的数据集较少，模型的准确性有待提高。

5 结论

本研究基于 SE-VBNet 构建的面神经磁共振图像自动分割模型能够实现自动分割面神经脑池段至颌骨外段主干，其分割结果与经验丰富的放射科医师手动分割一致性较好，有助于简化复杂、繁琐的分割重建过程，减轻医生劳动强度、提高工作效率，为面神经形态学的进一步高效、大量研究奠定基础。

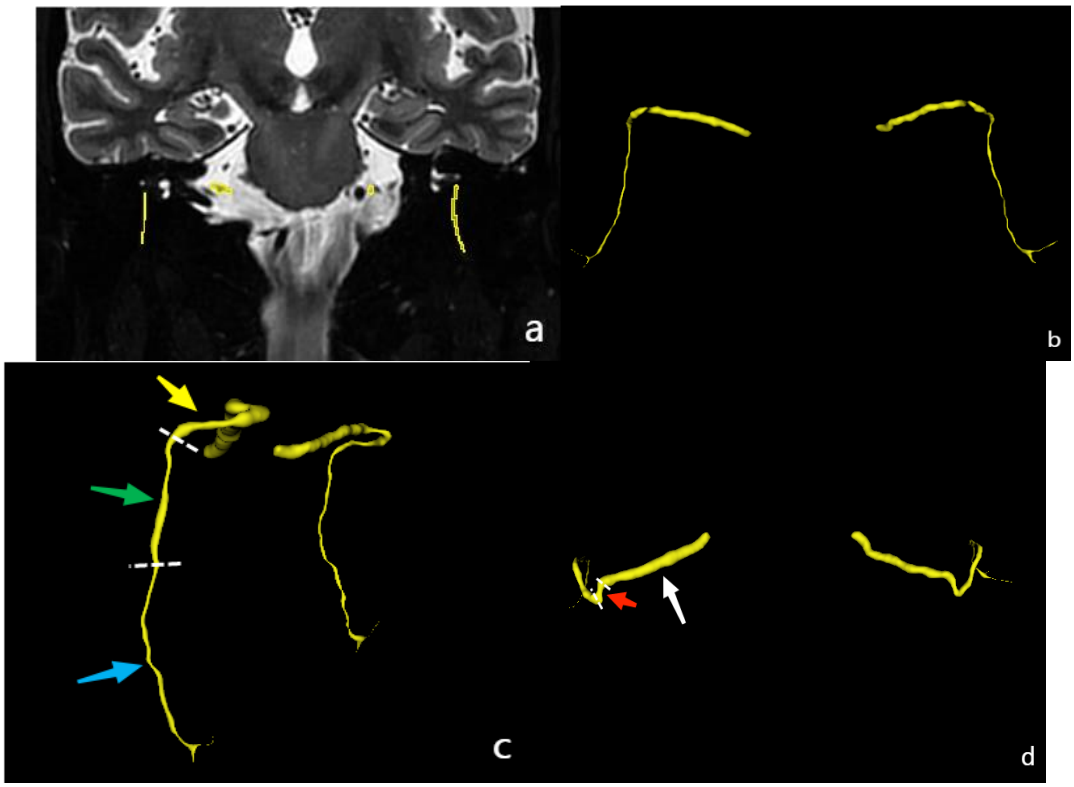


图 1 面神经脑池段至颌骨外段主干 MRI 图像人工标注。a 为 ROI 标准；b、c、d 为面神经标注后三维重建，c 图中黄色箭头为鼓室段，绿色箭头为乳突段，蓝色箭头为颌骨外段；d 图白色箭头为脑池段与内耳道段，红色箭头为迷路段。

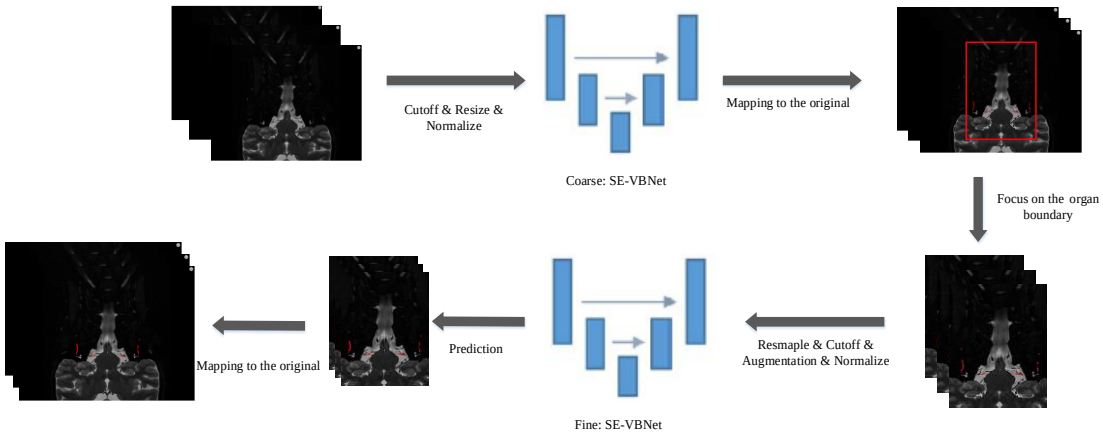


图 2 SE-VBNet 流程图。

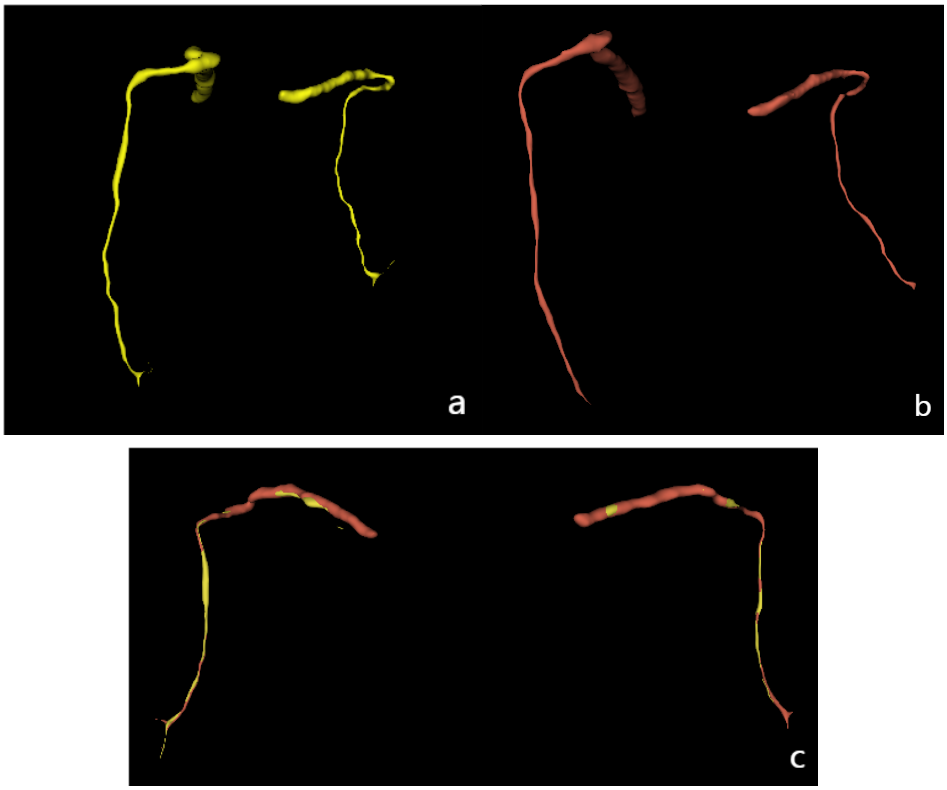


图 3 a 为手动分割 3D 重建，b 为深度学习模型自动分割 3D 重建，c 为手动分割与自动分割融合。黄色为手动分割，红色为自动分割，脑池段至鼓室段分割效能较好，颌骨外段至面神经远端颞面干及颈面干分割效能较差。

PU-209

健康医疗大数据支持下的脑胶质瘤诊治科研初探

李旭① 李挺②

①贵阳市第二人民医院肿瘤科 ②贵阳市卫生健康信息中心

引言

现代医学对脑胶质瘤的个体化诊断及治疗标准提示，需综合性地对患者病情、症状及体征等进行检查，具有一定的复杂性，且该病术后并发症发生率较高，临床工作中单体医疗机构实施的普通的管理系统无法满足脑胶质瘤患者信息的收集与整理，包括后续病人的全生命周期健康医疗

情况的追踪和信息分析。而在贵阳市健康医疗大数据平台的支持下，这一管理和研究模式得到了实现的可能。

一、贵阳市健康医疗大数据现状

贵阳市建立了以贵阳市人口健康云平台为核心，健康医疗大数据治理体系为支撑的全市区域范围内的健康医疗大数据平台体系。健康云平台是按照国家卫健委信息化建设的“46312”体系框架的要求进行的区域医疗大数据平台。 实现了全市公立医疗机构和基层社区卫生服务中心（乡镇卫生院）等三百家医疗机构的信息化覆盖。打通了市县乡三级公立医疗机构诊疗数据通路，支撑起市县乡村四级基层公共卫生体系。在“健康云”平台基础上，贵阳市健康医疗大数据治理平台更加深度和全面地整合全市医疗机构临床诊疗数据，对贵阳市健康医疗大数据进行收集、清洗、分析、应用、呈现，支撑相关监管部门决策及医疗大数据的综合应用。并形成优良的健康医疗数据为更多需求场景提供坚实的基础。

通过平台的数据汇集和科研分析功能的应用，笔者能方便和准确地在全市范围内对脑胶质瘤病人进行筛查和病情跟踪，通过大数据分析和联系，可以对病症在整个区域内的诊断，治疗，用药等各个方面进行跟踪分析。病人所有的健康医疗数据都能够得到 360°的展示，对病人的病情发展和延续有了良好的跟踪和管理手段。

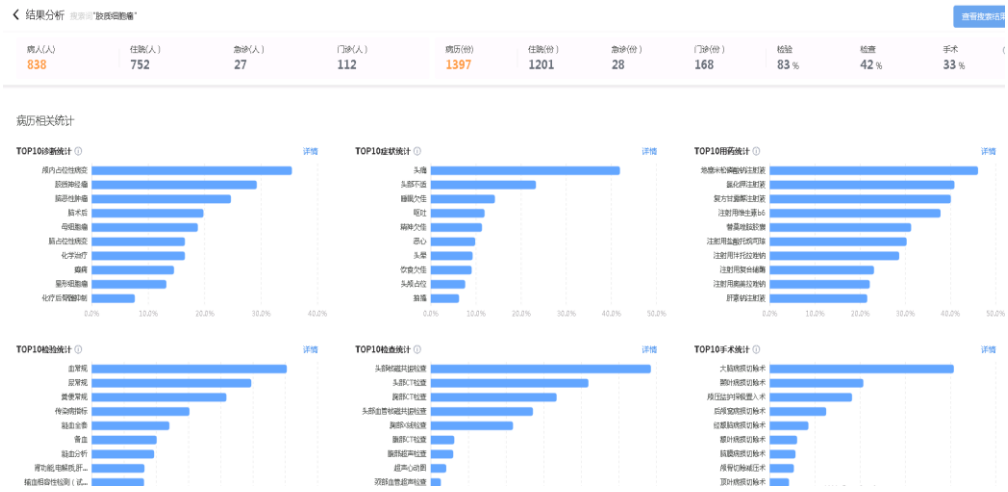
二、大数据带来胶质瘤研究观念转变

脑胶质瘤是最常见的原发性颅内肿瘤，传统上将弥漫性胶质瘤分为高级别胶质瘤 (WHO Ⅲ和Ⅳ级）和低级别胶质瘤 (WHO Ⅱ级)，前者约占所有脑胶质瘤的 70%，预后差 ，5 年生存率仅为 13%[1]，治疗以手术切除为主，并结合放疗和化疗等其他疗法。在神经胶质瘤的诊疗领域，无论是肿瘤的流行病学调查、临床诊断与治疗、预防与监测，还是临床试验研究都贯穿着对数据的收集、管理和分析[2]。

目前医疗大数据的特点是复杂化、碎片化、不兼容和非完整性，这就导致了临床医生及科研人员很难访问和使用，所以医疗数据标准化、规范化的收集成了一大难题。胶质瘤患者的医疗数据可以来自于不同的医院和科室、报告和记录，这些数据往往是非结构化的。与此同时，来源于不同地区、研究机构、人员、方法、仪器的测序数据，都包含不同程度的系统性差异。所有的精准模型都是建立在海量病人数据之上。越早建立起成熟的大数据质控机制，不断用更新的病人信息修正预测模型，成为目前大数据应用于胶质瘤诊疗的当务之急。

三、利用大数据平台的胶质瘤研究

笔者通过大数据平台，查询 2017 年 1 月 1 日到 2022 年 3 月 22 日的贵阳市所有市县乡三级公立医疗机构的主题词为“胶质细胞瘤”的病历，供检索到“胶质细胞瘤”病人 838 人，病历 1397 份病历。通过数据结果分析，快速获得前十位医院分布、诊断、症状、用药、检查、检验、手术方式统计结果。结合数据可以分析得出以下结论，胶质瘤这一肿瘤疾病在贵阳市县乡三级公立医疗机构形成了专病专治局面。我院在北京天坛医院多年帮助中已形成治疗脑部肿瘤疾病的专科医院，其他医院收治该类疾病基本为后续康复，支持治疗。结合健康医疗大数据，对胶质瘤诊治及科研进行一些初步探索。



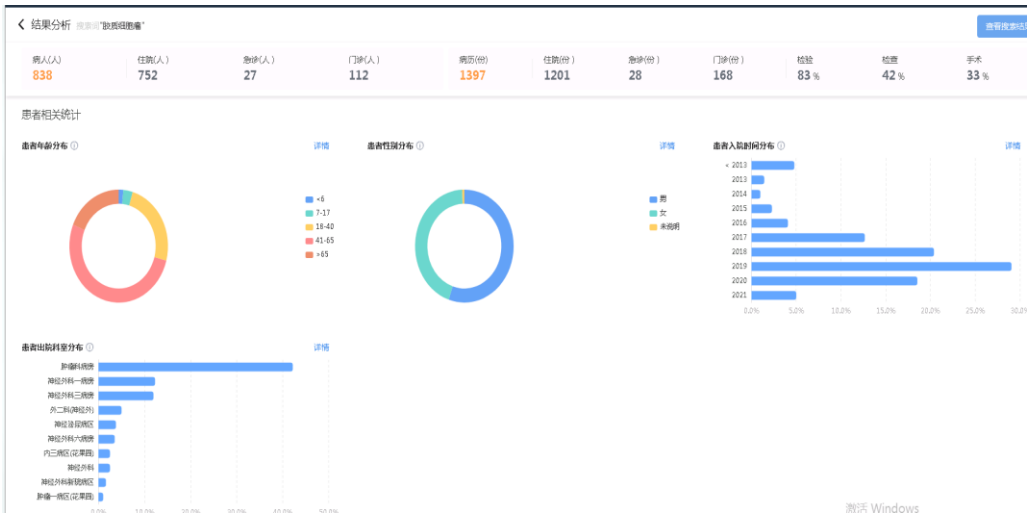


图 1：病历样本分析统计表

1. 临床诊断：

胶质瘤在诊断上，高度依赖高清晰的 MRI 影像，目前 MRI 可提供平扫、增强、flair、PWI、MRS 等多种信号，由此带来大量的图像数据，需要影像科医生，综合评定诊断。从数据上也可看出，胶质瘤影像诊断首选头颅 MRI 检查，CT 仅作为术后排除出血的急诊检查或特殊情况下 MRI 的重要补充。尽管 MRI 是目前临床上诊断及评估脑胶质瘤最重要的方法，但有时胶质瘤的影像表现不典型会使得诊断及其与一些非肿瘤组织（例如瘤周水肿和治疗后改变）的鉴别变得困难，人工诊断及评价胶质瘤需要耗费大量的时间和精力。大数据技术应用的飞速发展极大地推动了恶性肿瘤的个体化治疗。而利用现代医学平台全面地汇集患者的数据资料（基本临床信息、多模态影像数据和多组学数据 ），分析并利用这些数据，能够帮助临床医生针对患者个体制定出最佳治疗方案以及进行最优的药物选择，从而实现精准医疗。在胶质瘤的精准治疗方面中，基于大数据挖掘基础上个体化治疗为脑胶质瘤治疗开辟了全新的治疗天地[2]。

图 2：疾病关系图谱研究

图 3：研究病历样本

2. 治疗：

胶质瘤的治疗以手术为主，当影像诊断提示胶质瘤后，神经外科医生需要根据影像资料，决定患者是否能够手术，这其中包括患者年龄，有无并发症，肿瘤所在位置。大数据能够根据患者现有资料，快速选择相应的既往病历资料，及参考文献，为神经外科医生选择最合理的术前准备和手术入路，避免损伤那些重要功能区，结合脑胶质瘤患者的自身状况，制定科学、合理的个性化管理措施。术后，根据患者各项生理指标，评估患者状况，设立危及值报警，及时发现患者手术并发症，提供最佳处理方案。指导围手术期处理，保障患者平安。Robin [5]等对脑胶质瘤患者进行深入研究，以患者选择和循证管理为基础的手术决策能够有效且最大限度地降低患者的手术风险。

胶质瘤的重要补充治疗手段是放射治疗。由放疗科完成。根据术后复查影像资料，决定放疗靶区，每个医生勾画的靶区或多或少存在差别。大数据可减少认识误差，其中①放疗时间的选择，李凯新等[6]回归性分析 180 例高级别胶质瘤患者，从手术至放疗间隔时间以及生存关系中证实，放疗等待时间延长 1d，患者死亡风险增加 2%。大数据提示，可以让神经外科医生把握术后恢复时间，尽快安排患者转入放射治疗科，开始放疗。②患者完成放疗定位，影像数据传输工作站后，随后的放疗计划由物理师和放疗医师共同完成，需要计算机快速计算出治疗计划图，给予肿瘤最大剂量，周围正常组织最小剂量。临床实际操作方中，靶区制定仍然存在较多问题，如

肿瘤实际浸润范围和影像学检查是否一致，浸润程度与胶质瘤病理分级、肿瘤位置、邻近器官关系以及肿瘤水肿范围都相关。因此，需要根据实际病情选择放疗范围。目前高性能的直线加速器，带有人工智能模式的设备，可以借助大数据，完成诸如一键勾画正常组织，针对靶区要求，工作站自动完成计划设计，然后物理师和医师根据需要，作出修改。

药物治疗：目前在胶质瘤治疗上，替莫唑胺是唯一一个可通过血－脑脊液屏障并使用至今的一线化疗药物，目前尚无其他药物能取代替莫唑胺[7]。靶向免疫治疗也仅有贝伐单抗类抗血管生成药物取得一定疗效，但生存获益不多。

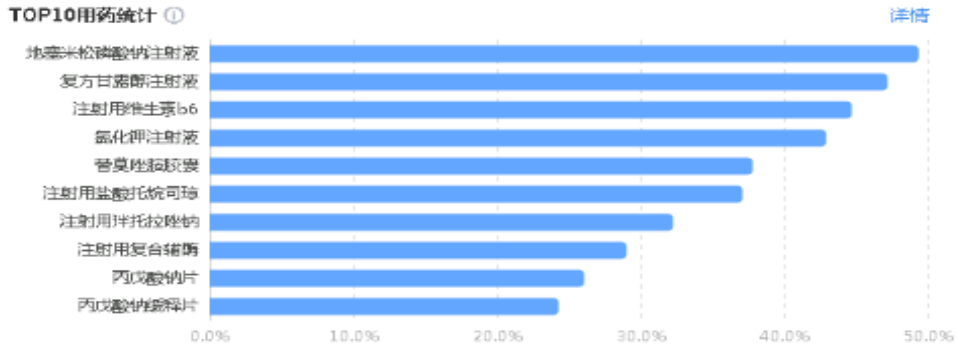


图 4：病历用药参考

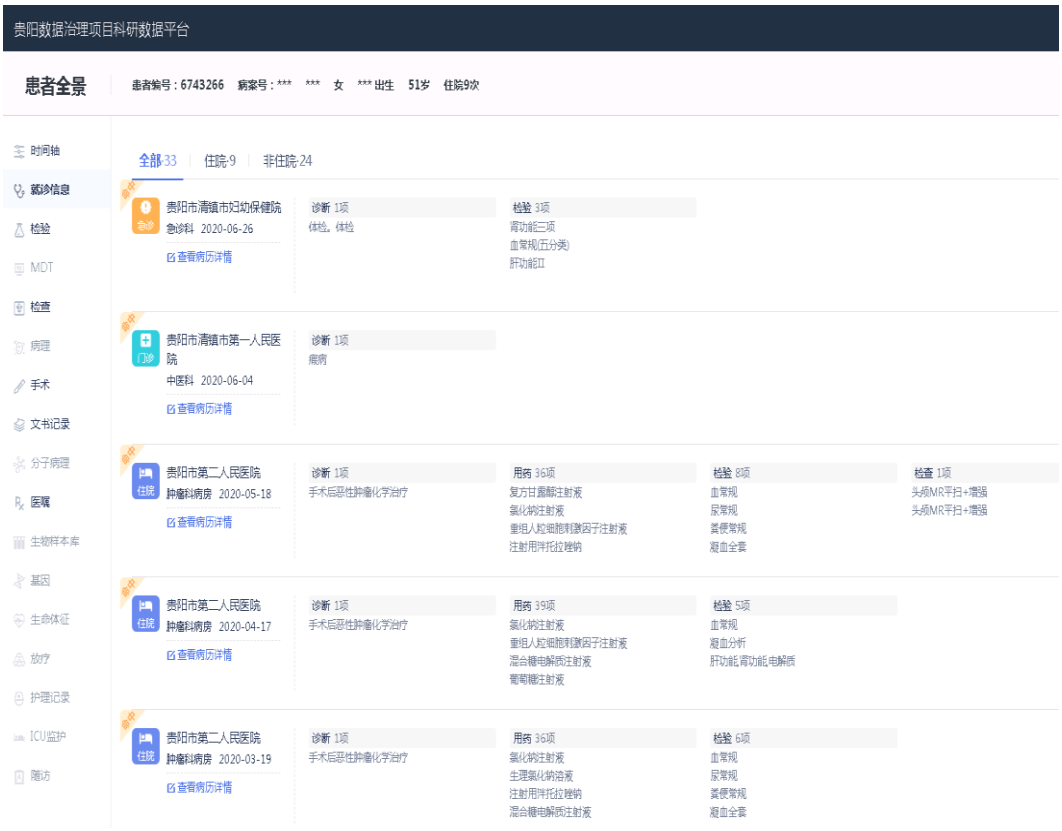
从前十位用药统计可以看出，市县乡三级公立医疗机构用药主要是改善脑水肿的激素和甘露醇、改善代谢和电解质平衡的维生素和氯化钾、化疗药替莫唑胺以及止吐护胃的对症治疗药物、胶质瘤易发癫痫，丙戊酸钠是首选药物。均是围绕病人治疗必不可少的基本药物。由此可对各医院收治相关病人后，提出用药指导。避免过度医疗发生。

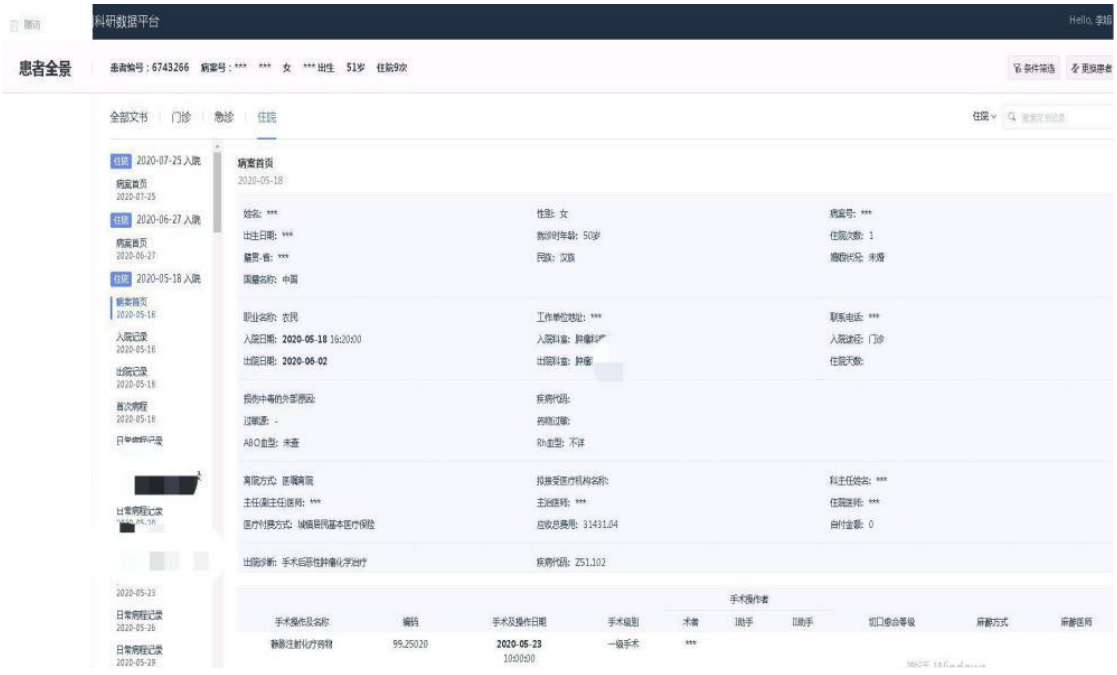
3. 预后：

目前胶质瘤总体治疗疗效，尤其是胶质母细胞瘤还不能令人满意，各种预后因素分析的文章[1]很多，如何将不同医院的资料进行整合，提供高质量临床指导意见，还需要更多高质量循证医学证据形成大数据。精准医学，通过手术病理组织标本进行基因分析，查找有意义基因，并通过大量数据分析，预测患者预后，目前的 MGMT,1p-19q,IDH1 等已经开始指导临床药物疗效、预后判断，此项工作在一些地区已经开展。

4. 病人病情全景跟踪

肿瘤治疗后，一个重要的工作就是定期随访，关注患者生活质量，判断是否复发，决定是否开始新的治疗。通过区域化的健康医疗大数据平台，能准确和大范围的帮助使用者随时随地获得与患者相关的各类诊疗信息，同时打破医疗边界，360°的掌握到病人的医疗健康情况。


贵阳数据治理项目科研数据平台
患者全景
患者编号：6743266 病案号：*** *** 女 *** 出生 51岁 住院9次
时间轴
就诊信息
检验
MDT
检查
病理
手术
文书记录
分子病理
医嘱
生物样本库
基因
生命体征
放疗
护理记录
ICU监护
随访
全部33 住院:9 非住院:24
贵阳市清镇市妇幼保健院
急诊科 2020-06-26
查看病历详情
诊断 1项
体检。体检
检验 3项
肾功能三项
血常规(五分类)
肝功能II
贵阳市清镇市第一人民医院
中医科 2020-06-04
查看病历详情
诊断 1项
痢病
贵阳市第二人民医院
肿瘤科病房 2020-05-18
查看病历详情
诊断 1项
手术后恶性肿瘤化学治疗
用药 36项
复方甘露醇注射液
氯化钠注射液
重组人粒细胞刺激因子注射液
注射用泮托拉唑钠
检验 8项
血常规
尿常规
粪便常规
凝血全套
检查 1项
头颅MR平扫+增强
头颅MR平扫+增强
贵阳市第二人民医院
肿瘤科病房 2020-04-17
查看病历详情
诊断 1项
手术后恶性肿瘤化学治疗
用药 39项
氯化钠注射液
重组人粒细胞刺激因子注射液
混合糖电解质注射液
葡萄糖注射液
检验 5项
血常规
凝血分析
肝功能,肾功能,电解质
贵阳市第二人民医院
肿瘤科病房 2020-03-19
查看病历详情
诊断 1项
手术后恶性肿瘤化学治疗
用药 36项
氯化钠注射液
生理氯化钠溶液
注射用泮托拉唑钠
混合糖电解质注射液
检验 6项
血常规
尿常规
粪便常规
凝血全套


科研数据平台
患者全景
全部文书 门诊 急诊 住院
病案首页
2020-05-18
手术操作者

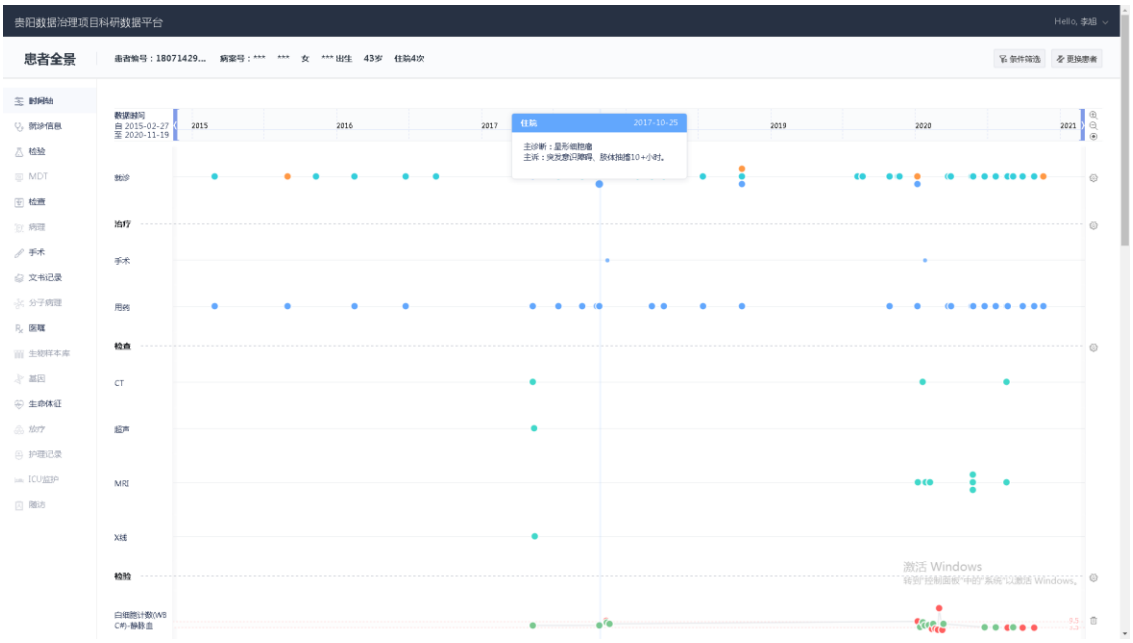


图 5：患者全景病历追踪

四、总结

综上，健康医疗大数据时代对传统的胶质瘤全程管理模式提出了挑战，在面临数据碎片化、欠规范化、隐私保护等问题的同时，也为胶质瘤的管理带来了重大机遇。它将影响到我们现行医疗模式，优化医院临床管理和科研方向。

PU-210

基于深度学习的眼底图像辅助诊断

邵家辉 [1] 吴辉群 [2]
[1] 南通市第二人民医院信息科
[2] 南通大学医学院

1 引言

智能医学是一个新兴的跨学科领域，旨在利用人工智能和机器学习等技术，改善医学领域的诊断和治疗过程[1]。在智能医学中，眼底图像识别是一个重要的研究方向，因为眼底图像包含了丰富的医学信息，如血管、视网膜和黄斑等部位的病变信息[2]。然而，由于眼底图像的复杂性和多样性，传统的图像处理方法无法完全满足实际应用的需求。因此，研究者使用深度学习技术来进行眼底图像识别，以提高识别的效率和准确性。

智能医学的发展得益于机器学习和深度学习等技术的不断进步。机器学习是一种基于数据驱动的技术，通过训练大量数据来提高算法的效率和准确性。深度学习是机器学习的一个分支，其通过构建深度神经网络来实现更加复杂的特征学习和分类[3]。在智能医学中，深度学习被广泛应用于医学图像处理、疾病诊断和医疗决策方案制定等方面。眼底图像深度学习模型的训练过程如图 1-1 所示。

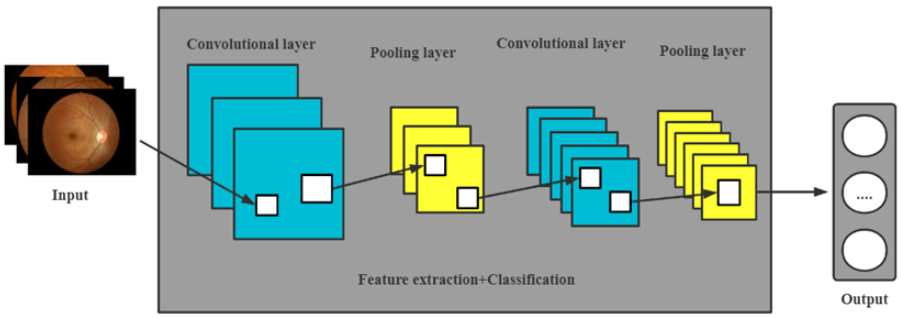


图 1-1 眼底图像深度学习模型的训练过程

眼底图像识别是智能医学应用的一个重要领域，具有较高的实际应用价值。在眼底图像识别中，图像处理和特征提取是两个关键步骤。图像处理主要包括图像预处理、分割和增强等操作，目的是去除图像中的噪声、突出关键区域，以便于后续的特征提取和分类。特征提取则是从处理后的图像中提取出有用的特征，如纹理、形状和颜色等，以供分类器进行训练和分类。

在眼底图像识别中，深度学习具有显著的优势。循环神经网络（RNN）和卷积神经网络（CNN）是两种常用的深度学习算法，分别适用于不同的眼底图像识别任务。循环神经网络是一种适用于序列数据的神经网络，其通过循环的方式对输入序列进行处理，适用于语音识别和自然语言处理等任务。卷积神经网络是一种专门针对图像处理的神经网络，其通过多层的卷积层和池化层来提取图像的特征，适用于图像分类和目标检测等任务。在眼底图像识别中，卷积神经网络可以用于疾病预测和病程进展分析等任务。

2 人工智能在眼底辅助诊断的优势

医学图像辅助诊断领域非常适合人工智能集成，因为该领域利用广泛的数字成像模式和结构化指标，如视网膜中央凹厚度、视敏度等。电子医疗数据集与多模态数字图像相结合，为人工智能分析提供了大量患者眼底图像信息。同时，进行眼科手术时所需的高精度仪器可以获取大量的眼底图像数据，通过辅助诊断工具的部署，能够促进智能医学技术的应用。

随着人们对眼部健康的重视，集成了人工智能算法的远程眼科解决方案能够通过批量筛查眼底图像文件来减少临床医生的审查时间，深度学习模型辅助眼底图像识别如图 2-1 所示。当下，由于城乡之间的资源分配不平衡，医疗保健的公平性一直是人们广泛关注的问题。在基层部署人工智能辅助诊断工具，提供完善的网络安全硬件设备，可以有效解决基层医疗资源不足的问题。通过人工智能的辅助诊断，基层眼科医生可以获得更多的诊疗支持，从而更好地服务患者，改善医疗资源分配。

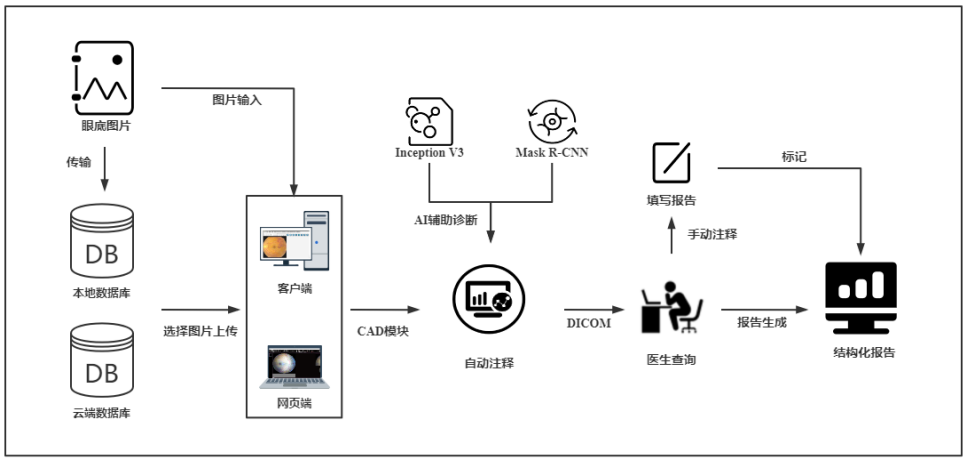


图 2-1 深度学习模型辅助眼底图像识别

为了验证人工智能在眼底图像识别中的优势，我们进行了一系列实验。首先，收集了一批眼底图像数据，包括健康人群和患者的眼底图像，并使用传统的图像处理方法对这些图像进行处理和特征提取。然后，我们使用 Inception-V3 和 Mask-Rcnn 深度学习模型对处理后的图像进行分类和预测，并将结果与传统的图像处理方法进行比较。如图 2-2 所示，训练深度学习模型对正常眼底和糖尿病视网膜病变眼底图像进行分类。实验结果表明，人工智能能够辅助识别出眼底病变区域和判断病情进展。

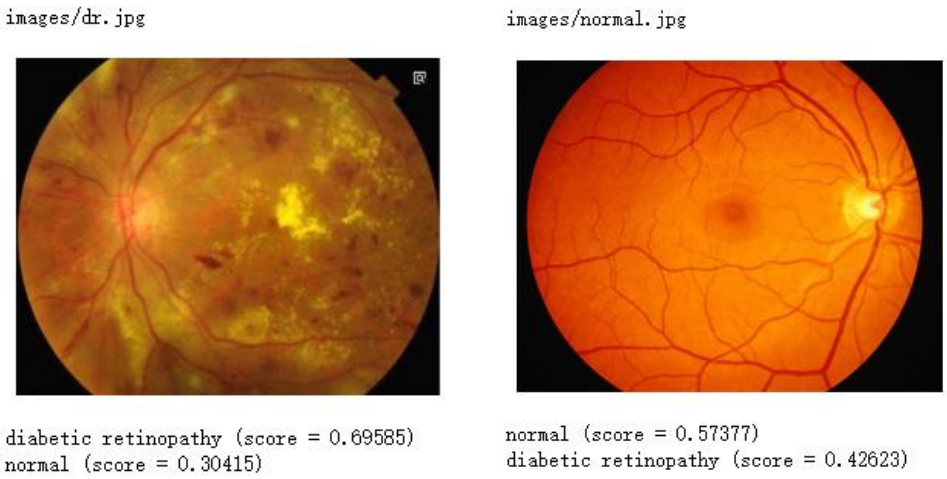


图 2-2 正常眼底和糖尿病视网膜病变眼底图像进行分类

人工智能辅助诊断通过深度学习算法对眼底图像进行分析和诊断，能够快速、准确地识别眼底病变，包括黄斑变性、青光眼、糖尿病视网膜病变等，从而提高诊断效率和准确性，并最大限度地降低对患者造成误诊漏诊的风险。可以通过建立图像数据库以提供足够的训练和验证样本，以开发强大的深度学习模型，并使用内部和外部多元数据进行独立测试，以评估其可行性。要为每种类型的应用程序建立性能标准，验收测试和质量保证程序，以确保深度学习模型的性能能够满足医院眼底图像分析的要求，并随着时间的推移不断更新迭代。对于模型研究人员和开发人员来说，重要的是要了解临床医生对眼底图像分析的需求，设计有效的辅助诊断工具，并通过考虑临床环境中的实际问题来提供智能化模块。如果开发、验证和实施得当，来自智能工具的高效数据分析可以辅助临床医生的诊断，从而整合临床工作流程，提高患者的生活质量和健康水平。

3 人工智能在眼底图像辅助诊断的应用

眼科医生对各种眼部疾病的诊断、监测和治疗决策都是基于其临床特征，在眼底镜、视野仪等眼科设备上进行识别，判断健康眼底与病变眼底不同阶段之间的差异。智能医学作为一门正在崛起的学科，它有可能将眼科医生从这些单调的筛查工作中解放出来。目前，智能医学的主要重点是使用人工智能进行眼底图像分析，以对眼部疾病进行分类、分级、分割和预测。眼底图像人工智能辅助诊断的应用流程如图 3-1 所示。

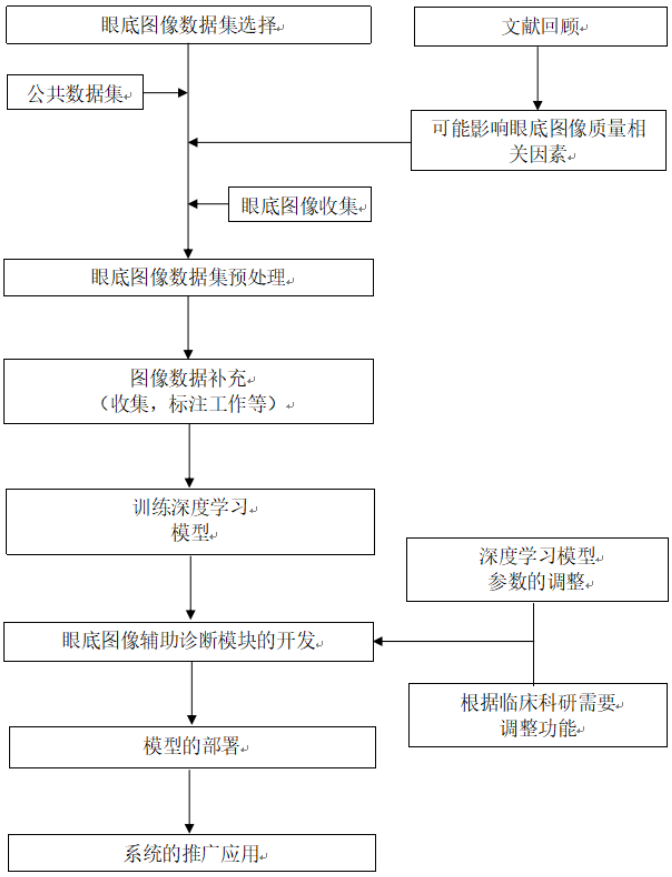


图 3-1 眼底疾病辅助诊断系统的技术路线图

3.1 糖尿病视网膜病变

据估计到 2040 年全球约有 6 亿人将患有糖尿病，其中三分之一预计将患有糖尿病性视网膜病变，这是全球成年人失明的主要原因[4]。中国目前是世界上糖尿病患者最多的国家[5]。2016 年，国务院发布了“健康中国 2030”规划纲要，为中国卫生改革的未来方向提供了进一步的指导。“健康中国 2030”概述了到 2030 年所有糖尿病患者将接受疾病管理和干预的目标。在中国，大约有 4 万名眼科医生，与糖尿病患者的比例为 1:3000。作为一种具有成本效益的预防措施，鼓励在社区一级进行定期视网膜筛查。

糖尿病视网膜病变的筛查至关重要，因为它有助于早期检测和治疗，从而防止视力下降[6]。在高危人群中，糖尿病视网膜病变的发病率高达 40%，而且有很大一部分患者没有接受筛查[7]。人工智能在糖尿病视网膜病变的筛查中能够发挥重要作用，它可以帮助克服筛查障碍，有效防止失明，提供个性化的诊断和治疗建议。

人工智能辅助糖尿病视网膜病变检测的早期研究集中在病变检测上，并且已经发展为将病变眼底分类，主要集中在标准彩色眼底照相上。2016 年，Abràmoff 等人[8]和 Gulshan 等人[9]设计了使用卷积神经网络的眼底图像算法，其能够检测糖尿病性视网膜病变，曲线下面积 AUCs 分别为 0.98 和 0.99。随后 Ting 等人[10]的研究使用了更大的数据集，证明了对相关糖尿病视网膜病变的良好检测，AUCs 分别为 0.97 和 0.94。

除了标准的眼底图像拍摄，研究人员还使用光学相干断层扫描图像、超宽视野成像，甚至智能手机捕捉的视网膜图像来研究视网膜病变的智能检测。学相干断层扫描识别的视网膜内液，卷积神经网络可以准确识别了，例如 Lee 等人[11]使用手动分割的黄斑光学相干断层扫描图像来开发能够检测黄斑水肿的卷积神经网络。超宽视野成像能够识别可视化角度为 200°的眼底，捕捉眼底图像上除糖尿病相关的潜在周围疾病，Nagasawa 等人[12]发现卷积神经网络在超宽视野成像图像上检测增生性视网膜病变具有较高的灵敏性性(94.7%)和特异性(97.2%)。同样，Wang 等人[13]发现使用超宽视野成像图像对可参考的糖尿病视网膜病变具有较高的敏感性(91.7%)。除了通过专业的眼底成像设备，Natarajan 等人[14]还开发了一个基于智能手机的离线人工智能系统，该系统对检测可参考的糖尿病视网膜病变具有高灵敏度。

3.2 年龄相关性黄斑变性

年龄相关性黄斑变性是导致视力丧失的常见原因之一，全球估计有 1.96 亿患者受到影响[15]。早期发现和治疗黄斑变性可以最大限度地减少视力下降。考虑到疾病的早发现早治疗，辅助眼底图像的大规模筛查。

Burlina 等人[16]使用来自 4613 名患者的超过 130，000 张图像的数据库来开发用于自动检测年龄相关性黄斑变性的深度学习算法。他们的辅助诊断系统能够识别中度和晚期年龄相关性黄斑变性患者，并且准确率达到 92%。Vaghefi 等人通过研究[17]表明，与单独使用光学相干断层扫描相比，在年龄相关性黄斑变性中结合使用深度学习模式可将检测的准确性从 91%提高到 96%。Keenan 等人[18]开发了一种眼底图像人工智能算法，该算法可以准确量化新生血管黄斑变性患者的液体体积。人工智能还被用于量化与年龄相关性黄斑变性相关的其他关键特征，包括视网膜内液、视网膜下液、色素上皮脱离、椭球区损失、玻璃疣、纤维化和视网膜下高反射物质[19]。

3.3 青光眼

青光眼是全球不可逆失明的第二大常见原因[20]，青光眼的早期检测已被证明可以减少视力下降[21]。眼底图像筛查是青光眼的常用检测方法，并被作为许多远视青光眼评估的一部分[22]。可疑青光眼的综合评估可能包括光谱域光学相干断层扫描(SD-OCT)、视野检查、眼压测量、角膜厚度测量和房角镜检查。人工智能算法已被开发用于通过视盘照片和 SD-OCT 识别视神经变化，从而预测青光眼视野变化。

Li 等[23]建立了一个大规模的基于注意力的青光眼数据集，包含 5824 张眼底图像和由眼科医生提供的注意力图。他们还提出了一种卷积神经网络来诊断青光眼。首先，引入了一个注意力预测子网来生成注意力图。在该子网中，使用了多尺度和通道注意方法。然后，设计了一个病理区域定位子网来定位病理区域，在每个阶段将注意力图嵌入到特征图中。最后，定位的病理区域和预测的注意力图被连接在一起，并被馈送到青光眼子网中，以预测青光眼的二类标签。Liao 等人[24]介绍了一种用于青光眼诊断的临床可解释的 ConvNet 架构 EAMNet。他们首先使用卷积神经网络作为骨干网络，使用多个残差块来提取有用的特征。然后提出了一种称为多层平均池 M-LAP 的方法来弥补低级别定位信息和高级别语义信息之间的差距。此外，通过特征图的加权求和获得了证据激活图。

3.4 视网膜中的其他应用

主要涉及多种疾病的诊断，例如 De Fauw 等人[25]设计了一种深度学习架构，通过光学相干断层扫描图像识别可参考的视网膜疾病，实现了与视网膜专科医生相当的性能，该系统能够识别新生血管黄斑变性、地理萎缩、玻璃膜疣、黄斑水肿、黄斑裂孔、中心性浆液性视网膜病、玻璃体黄斑牵引和视网膜前膜。深度学习能够基于疾病患者眼底图像的结构评估，在微视野计上预测视网膜功能[26]。这有助于评估患有遗传性视网膜疾病的患者，同时在临床试验中监测进展或治疗效果。在其他眼底图像智能辅助诊断系统中，实现了识别中心性浆液性视网膜病、脉络膜血管病变、镰状细胞病和黄斑毛细血管扩张症[27]。

4 讨论：挑战和机遇

本文通过对智能医学、眼底图像识别和深度学习的研究，介绍了深度学习在眼底图像识别中的优势。深度学习能够提高眼底图像识别的准确性和效率，为医学诊断和治疗提供了更好的支持。与传统的图像处理方法相比，深度学习具有更强的特征学习和分类能力，能够更好地处理复杂的医学图像数据。但在未来的应用中依然面临以下几个挑战：

4.1 数据集的整合与共享

原始数据的整合，主要包括多模态图像和来自不同地区的多个数据集的结构化临床信息，如何处理图像数据，并建立共享系统，使数据驱动的人工智能技术在眼科实践中发挥作用，需要进一步探索。同事要考虑网络安全问题，在网络安全框架内建立数据存储和共享系统，以便更安全广泛地使用。

4.2 数据隐私与信息透明

虽然眼底图像拥有可以通过人工智能技术重新识别的生物信息，但在收集和处理图像时仍需保存好个人信息。一些新的学习任务可以模糊生物识别信息，甚至为小规模数据集提供无监督模型。如何通用化图像数据个人信息的问题将是扩展人工智能技术的可扩展性和可推广性的关键因素之一。人工智能的复杂性和不可理解性包含在术语“黑盒算法”中[28]，黑盒算法有可能导致人工智能在医疗保健生态系统中被滥用。因此，算法的透明性是眼底图像数据收集的一个关键点。在人工智能技术的应用中，保持数据隐私和信息透明之间的适当平衡是一个重要问题。

4.3 人工智能在临床实践中的应用

人工智能技术在现实世界中的作用不是取代眼科医生，而是辅助他们诊疗，将人工智能模型和医生经验结合起来，以做出更高效、更准确的决策。这种提高效率的能力可以简化临床工作流，让临床医生有更多的时间与病人沟通。如果人工智能算法仅用于改善工作流程而非患者护理，人工智能的不当实施可能对医患关系有害，并可能影响患者的信任。

4.4 道德问题和监管问题

在临床实践中部署人工智能技术要考虑一个关键问题是如何避免误诊漏诊。随着人工智能应用越来越受欢迎，法律责任应该得到很好的界定。需要建立更多的立法和治理系统来完善责任规则和监管环境。不仅要为人工智能模型的验证，还要为数据安全和法律责任建立政策和具体的权威机构。跨部门和跨学科的合作对于确保人工智能医疗保健生态系统的完整性将非常重要。

4.5 专业人才的培养

继续教育、实际应用的推广以及面向医疗行业的用户友好、可理解的界面对于简化工作流程和拓宽人工智能系统的适用性同样重要。跨职能团队之间的沟通和协作，包括眼科医生、患者和软件开发工程师，可以通过人工智能技术对眼部疾病的诊疗产生积极影响。

5 结论

随着智能医学的不断发展，眼底图像辅助诊断将在医学领域发挥越来越重要的作用。未来，我们可以进一步探索智能医学在医学图像处理、疾病诊断和治疗方案制定等方面的应用，为医学发展提供更强大的技术支持。同时，我们也需要关注深度学习的局限性和挑战，如数据质量和标注问题、模型可解释性等问题，并积极寻求解决方案，以实现智能医学的更好发展。

PU-211

基于 5G 的全院心电网络建设与应用效果

高标 张旸
安庆市第一人民医院信息中心

0 引言

随着医疗信息化水平不断提高，医院息化建设的完善对实现“健康中国 2030”起着至关重要的作用。在医院的信息化建设中，心电图的全程数字化管理是医院建立完整电子病历管理的必要组成部分，全院心电网络的建设也是整个医院信息化重要的组成部分，与影像数据的数字化管理和化验数据的数字化管理具有同等重要的地位，同时心电图检查也是临床诊断上最常用的方法之一，因此心电网络系统也得到了快速发展，从之前的单机心电工作站通过打印的方式打印纸质心电报告到现在全院心电网络的普及，患者无需单独去心电图室即可床边完成心电检查，真正践行了“信息多跑路，患者少跑腿”的理念[1]。但现有的无线 WIFI 组网心电网络系统难以克服网络抗干扰能力差、带宽低、延迟高、稳定性差的缺点，加上心电图诊断在临床诊疗中的必要性，患者多、工作量大，一些常规心电诊断带来大量重复工作无疑给心电诊断工作带来了压力及增加了诊断的失误率，因此一个基于稳定网络环境具备高带宽、低时延的心电系统十分必要。5G 网络在医疗信息化中的应用大大改善了医疗信息承载方式，5G 不仅仅是 4G 网络的一个简单延伸，而是一个真正意义上的融合网络，将 5G 网络应用于医院多院区及医联体临床信息化系统建设，能有效提高诊断与医疗水平、降低医疗开支、减少看病花费的时间[2]。基于此在 5G+心电一种新模式下，医生可以更快调取心电数据、开展心电远程会诊，大大提高了心电业务处理效率。

1 需求分析

传统的心电检查需要患者单独前往心电图室取号、排队、做检查、等报告等流程对于特殊的、行动不便以及急诊的患者带来了极大的不便。由于传统心电图机单机无法接入医院信息系统，心电信息孤岛给临床诊疗带来了阻碍，导致增加患者就诊时间，加大了临床诊疗成本。一种智能化、高效化的全院心电网络才能够面向未来满足人民健康的发展需求。

早期心电图机无数字输出接口，只能采用扫描的方式将心电图波形以图片的方式保存，单机心电工作站，也只能通过打印机打印心电报告，因此需要更智能的数据传输方式来提升数据传输效率，利用该技术，在心电业务信息化管理平台上，心电检查可完全实现在网上申请、预约和登记，实时在线诊断，网上传输报告及远程会诊、示教等，基于 5G 网络技术，可使全院多院区、医联体

的心电图数据报告实现数字化、网络化、无纸化集中管理[3]。

心电图资料无法有效保存是长期以来困扰医院的难题，而实现数据的集中存储、统一管理，不仅使医生通过和历史心电图对比，更好的观察病情和提升治疗效果，也可以提高医院和基层医院的临床和科研工作水平，未来面向医院可持续发展的一种智能化、轻便化、可靠的全院心电网络系统是必不可少的。

2 系统设计与实现

2.1 5G 网络的优势

网络覆盖面广：传统的无线路由布网方式，网络抗干扰能力差、带宽低、延迟高、稳定性差的缺点，导致在同一楼层不同病房信号覆盖不全，相较于传统组网方式采用 5G 网络能够更好的扩大网络覆盖面积，在临床应用过程中，医生采集完心电图数据即可发送[4]。

超高可靠低时延通信：在许多的临床应用过程中，一份急性的心电数据需要请求快速会诊，在院内带宽高负荷使用情况下，采用 5G 网络更高效，急性的心电图数据传输更快更稳定，临床医生更快的获取的患者的心电图报告。

远程医疗：在 120 急救车、医联体单位远程心电图通过 5G 网络技术更好保证院外数据的稳定快速。

2.2 系统构成

中心服务器 服务器操作系统采用 Windows Server 2012，提供心电程序和数据库服务控制和管理硬件资源。数据库选择了 SQL server 进行数据管理，通过数据库组件统一设定后台数据库。文件存储采用了 NATAPP 方案，相对 FTP 或者直接存到数据库，提升了数据传输效率，同时也保证了数据安全。

5G-Advanced 设备 系统传输网络采用“5G 无线接入网+传输网+核心网”三部分构成，本系统采用的 5G-Advanced 超级时频折叠技术相较于传统 5G 网络在时延、带宽及抗干扰能力上有明显的优势。

移动心电采集仪 采集仪有导联线、肢体夹、吸球、智能终端、便携小推车等部分组成，主要功能是完成患者心电数据采集、存储和传输[5]。

心电诊断中心 由心电报告终端应用及计算机组成，心电采集传输至中心服务器，通过算法分配到不同的诊断终端，医生完成诊断后回传至中心服务器，供临床医生实时调阅。

2.3 系统架构

本文以安庆某医院为例，针对医院一地三院区的特殊情况，全院心电网络的建设依托 5G 网络技术的优势，将三个院区心电监护数据整合在同一平台包括门诊心电检查、急诊心电检查、住院心电检查等数据传送至院内心电中心服务器，通过接口与医院 HIS 信息系统、CIS 医生站、EMR 系统等信息交互，可实时获取患者基本信息、医嘱信息以及心电报告的发布和归档，解决心电类检查的信息孤岛，同时针对院内外业务流程进行优化，提高心电的工作效率和管理水平[6]。基于 5G 网络的移动心电系统设计方案如图 3 所示：

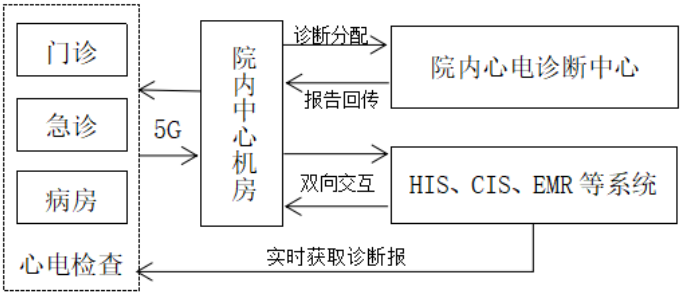


图 3 基于 5G 网络的移动心电系统设计方案示意图

全院多院区心电网络系统依托底层服务器架构，与 HIS 信息系统、排队分诊叫号系统对接实时

获取患者遗嘱及门诊处方信息，将多院区患者心电检查数据通过 5G 网络传输给中心服务器，院内心电诊断中心与中心服务器数据交换对患者心电诊疗数据进行人工干预审核辅助决策诊断报告发布，实现数据存储、调阅与共享。通过与 HIS 信息系统、EMR 系统、PACS 系统等融合，实现信息获取、心电检查、心电处理、数据管理等业务[7]。系统技术架构示意图如图 4 所示。

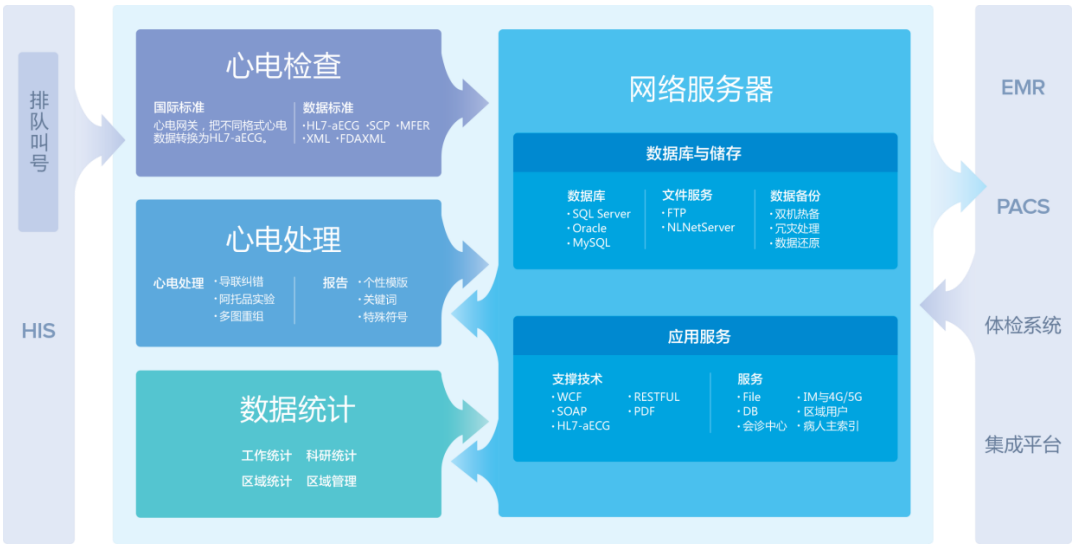


图 4 5G 网络的移动心电系统技术架构示意图

2.4 应用场景

安庆某医院 2022 年 12 月上线的基于 5G 网络的移动心电系统覆盖一院三区近 50 个临床科室病区，服务患者床边心电诊疗的需要，提高了临床心电诊疗效率。基于 5G 网络的移动全院心电系统在运营成本、诊疗效率、患者满意度等方面取得了很好的收益。

(1)门诊体检诊疗应用场景

1、心电网络系统通过与院内系统对接，获取患者的检查信息。

2、根据检查项目并通过排队叫号功能使患者有序完成相应检查。

3、检查数据上传至院内服务器。

4、诊断医生通过心电分析软件或智能终端完成诊断，并将诊断结论回传至院内服务器。

5、临床医生可通过系统接口获取患者诊断信息。

6、患者可通过院内自助打印机获取纸质报告。

(2)病区诊疗应用场景

1、临床医生开具心电检查医嘱。

2、病区医务人员通过数字式心电图机上的 Worklist 功能，一键获取待检者信息。

3、心电检查数据通过 5G 方式上传至院内诊断中心服务器。

4、诊断医生通过心电分析软件或智能终端完成诊断，并将诊断结论回传至院内服务器。

5、病区医务人员可在检查端获取患者诊断信息。

（3）诊断中心应用场景

院内心电诊断中心依托 5G 网络实时获取患者心电数据，心电采集完毕后，通过 AI 辅助功能分析，可将危急数据以危急预警信息和语音提示等形式告知检查医生。通过人工智能辅助决策系统快速将心电数据完成诊断报告，工作人员审核无误后完成心电报告发布及归档。同时根据接收的心

电检查信息危重程度不同，可自动进行诊断分级并通过弹窗报警提示，避免危急的数据淹没在常规的心电列表中，使危急的数据得到及时处理[8]。人工智能辅助决策诊断级别包含阳性危急、阳性、阳性可疑、正常、不可识别五种类型。

3 应用效果

安庆某医院床边移动心电系统自 2022 年 12 月上线，系统采用 5G 网络传输，依托 5G 网络优势实现了患者基本信息、医嘱信息的获取，心电图的数据的发送、信息归档、报告发布，网络稳定性相较于传统无线路由组网故障率大大减少，自 2023 年 1 至 3 月与去年同比因网络原因（包括断网、卡顿、信号差、网速慢等）问题报修平均减少 115.4%如图 5 所示。

通过对数据库中 2023 年 1 至 3 月期间门诊和体检患者的处方开具到心电报告发布完成的数据进行筛选与去年同期比较，患者心电检查平均时间同比较传统心电检查平均减少 26.7 分钟如图 6 所示。

基于 5G 网络的移动心电系统依托 5G 网络高效、准确的优势，工作人员只需审核发布心电报告，处理异常报警心电检查数据，相较于传统心电检查人工核对信息、书写报告、审核报告、发布报告、归档报告等流程，工作人员效率同比提高 67.1%如图 7 所示。

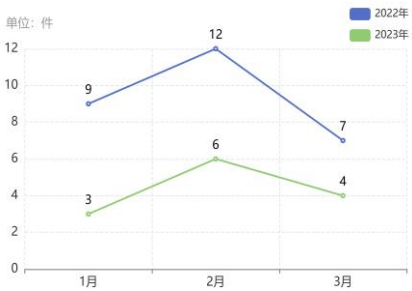


图 5 故障报修

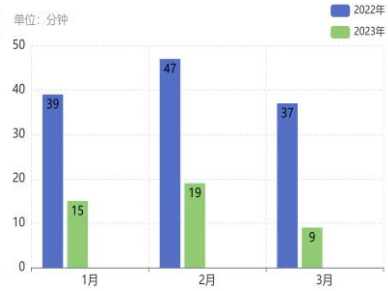


图 6 患者心电检查时间

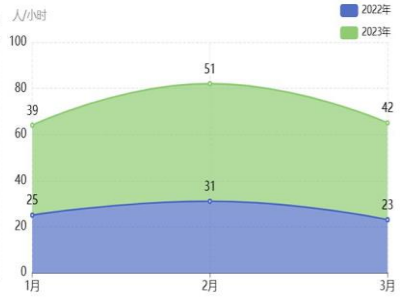


图 7 诊断效率

4 结论

基于 5G 网络的移动心电系统依托 5G 网络超低时延、超高带宽的优势，带来了高效化、智能化的全院心电网络解决方案，解决了传统单机心电图机无法长期保存只能采用扫描以图片保存的难题，实现了多院区心电数据的集中储存、统一管理，方便了医生比对历史心电图数据更好的观察病情和提升医疗水平，同时降低了心电系统故障报修率，减少了患者排队等待时间，提高了诊疗效率。

PU-212

人工智能在肺癌研究中的应用

王昕玥[1,2,3]，渠鸿竹[1,2,3]，方向东[1,2,3]
1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室
2. 中国科学院大学
3. 基因组与精准医学检测技术北京市重点实验室

1 背景

肺癌是一类原发于肺部支气管黏膜以及腺体的恶性肿瘤，是全世界癌症死亡的主要原因，约占世界上所有癌症相关死亡的 30%，其发病率和死亡率几十年来高居不下。肺癌分为小细胞肺癌（SCLC）和非小细胞肺癌（NSCLC）两种亚型，非小细胞肺癌又可凭细胞学形态分为肺腺癌（LUAD）和肺鳞癌（LUSC）。根据 2021 年全球癌症统计报告[1]显示，肺癌发病率在男性和女性中分别位列第 1 和第 2 位，死亡率则均位列第 1 位。其五年生存率也是所有癌症类型中最低之一。

现阶肺癌可以从胸部 X 光片和计算机断层扫描（Computed Tomography，CT）检查中被发现，但诊断和治疗主要还是依靠于传统的组织病理学[2]。肺癌具有早期症状不明显的特点，许多患者在被诊断时已经进入了晚期，早期肺癌患者可以通过手术治疗显著改善患者生存期，而中晚期患

者，即使使用了化疗或放疗，死亡率依旧很高，同时还伴随局部复发和远处转移，这也是肺癌患者死亡的主要原因。

随着大数据和计算机科技的迅猛发展，产生了基因组、转录组、代谢组、影像组和临床信息等多类型医学数据。肺癌的复杂性和异质性使得其在各个环节所产生的各种数据都具有研究价值。近十几年来，人工智能（Artificial Intelligence，AI）技术在肿瘤精准治疗方面的应用越来越广泛，诊断和预测能力丝毫不逊色于专业医生，助力肿瘤患者实现个性化精准治疗。总之，大数据和 AI 在肺癌领域中发挥着重要的作用，两者结合可以帮助医生提高肺癌筛查和诊断准确率，辅助患者确定最合适的治疗靶点和个性化治疗策略，预测生存期，降低病发率死亡率，显著改善患者治疗效果和预后情况。本文旨在介绍近年来 AI 在肺癌各个环节中的最新应用。

2 AI 用于肺癌风险预测和诊断

AI 在肺癌进程中应用的方面有很多，最早期的首先就是肺癌的高危人群筛查和肺癌风险因素的鉴定，提供早期预警和建议来帮助更多的潜在肺癌患者。Petousis [3]这篇论文设计了一组动态贝叶斯网络（Dynamic Bayesian Network，DBN）模型，可用于辨别肺癌病例以及预测高危人群未来发病几率。Ahmad 等人[4]使用决策树（Decision Tree，DT）和随机森林（Random Forest，RF）算法构建了一种基于危险因素的肺癌预测模型，该模型可以快速计算出危险因素对肺癌影响的百分比，敏感性和准确性都与专业医生研判结果相当。Chen 等人[5]在老年人群中使用了深度神经网络（Deep Neural Network，DNN）的方法来定量分析肺癌的行为、疾病史、环境等风险因素，其方法准确率范围高达为 0.927~0.962，并且发现了戒烟时间和吸烟量是 65 岁及以上女性患肺癌的主要危险因素，为防止老年肺癌提供了干预指标。

在肺癌整个进程中，AI 应用得最为广泛的是利用各种各样的数据对肺癌进行预测和诊断。血液检测操作简单快捷，且相对于传统的检测办法给患者带来的伤害更小，所以找寻血液中的肺癌生物标志物就成了此类研究的重中之重。Wang 等人[6]采用无监督去噪自动编码器（Denoising Automatic Encoder，DAE）鉴定到 35 个基因特征可作为肺癌的诊断和预后生物标志物。Xie 等人[7]使用朴素贝叶斯（Naive Bayes，NB）算法发现的六种血浆代谢物。而 Guan 等人[8]则使用代谢物数据利用极端梯度提升（eXtreme Gradient Boosting，XGBoost）方法鉴定到的血液中鸟氨酸和棕榈酰肉碱都可作为诊断依据，为肺癌血液筛查的可行性提供了有力支持。和血液检测一样无创的检查还有临床问诊，例如可通过个人病史，既往症状和合并症来预测肺癌风险。Lee 等人[10]选择了肺炎、急/慢性支气管炎和糖尿病等 13 个疾病特征，利用 DNN 开发了一个 10 年肺癌风险预测模型，准确率为 85.4%。Baidaa Al-Bander 等人[11]使用层次分析过程（Analytical Hierarchy Process，AHP）给网络调查收集到的每个临床数据特征分配权重，再用加权特征训练人工神经网络（Artificial Neural Network，ANN）来构建肺癌预测模型。评估结果表明，该模型在受试者中表现出良好的预测性能。

图像处理技术在肺癌的早期检测和治疗预防中发挥着重要的作用，可通过 AI 检索并训练图像中的特征，进行肺癌预测或诊断。及时准确地鉴别肺结节良恶性是肺癌早筛中至关重要的一环。Guobin 等人[12]开发了一种基于混合特征的梯度增强机（Gradient Boosting Machine，GBM）算法的卷积神经网络，可从 CT 图像中将判断肺结节的良恶性。Li 等人[13]结合 LDCT 肺癌筛查和 3D DNN 预测肺结节恶性程度。Zhang, G 等人[14]使用了一种名为 3D DenseNet 的网络结构来进行肺结节分类，比传统的 CNN 相比具有提高特征利用效率和缓解梯度消失等优点，该方法具有 92.4% 的分类准确度，可与计算机辅助诊断系统相辅相成。

常见的肺癌影像检查有 X 光、CT、正电子发射计算机断层显像(Positron Emission Tomography- Computed Tomography，PET-CT)和低剂量 CT（Low Dose CT，LDCT）等。Jayaraj 等人[15]实现了一种将逻辑回归（Logistic Regression，LR）和 Adaboost 分类器相结合的多阶段混合预测方法，应用于 CT 肺部图像，预测和检索肺癌的发生。相比 LR，神经网络更好地应对数据维度大和非线性的特点。Shimazaki 等人[16]基于卷积神经网络（Convolutional Neural Networks，CNN）架构开发了一种深度学习的模型，用于胸部 X 光照片上肺癌的检测和分割。Shakeel 等人[17]首先使用改进的 DNN 对肺部 CT 图像进行分割，准确定位肺癌区域，再利用集成分类器来综合判定肺癌，降低了误诊率，实现了肺癌的高效预测。Anita[18]等人则在 UCI 机器学习数据库中发现高斯朴素贝叶斯（Gaussian Naive Bayes，GNB）比 SVM、RF、NB 和人工神经网络 ANN 等机器学习算法，在 CT 图像预测患者的肺癌风险上表现更优，准确率达到了 97.5%。为了使模型具有更强的泛用性，Kalaivani 等人[19]的研究提出一种基于 ISBSSA 的机器学习方法，该

模型可自适应选择恰当的算法，例如 NB、CNN、SVMs、K 最近邻算法（K-Nearest Neighbor，KNN）和 RFs，来预测 CT 和 PET-CT DICOM 图像中肺癌细胞的存在和类型。此方法可有效地预测和分类肺癌细胞。

除开使用上述的转录组、代谢组和影像组数据，基因组数据同样也可与 AI 结合实现肺癌预测。Liu 等人[20]对 TCGA 和 ICGC 数据提出了基于 Kullback-Leibler（KL）散度基因选择的深度神经网络来预测肺癌，该方法极大地增强了模型的泛化能力，在验证集上的 AUC 为 0.99。Li 等人[21]开发了一种名为 Deep-LC 的深度学习方法，该方法首先用图卷积网络（Graph Convolutional Network，GCN）提取基因对之间的相互作用，再利用 CNN 模块作为解码器判断基因是否与终点事件肺癌相关，极大地提升了常见深度学习方法在该领域的准确率。

3 AI 应用于肺癌分期和亚型分类

确定肺癌分型和分期是患者获得治疗的前提和依据，有助于提高治疗的针对性和有效性[22]。Wang Z 等人[23]使用变分自动编码器（Variational AutoEncoder，VAE）来挖掘 LUAD 和 LUSC 两种肺癌亚型表观遗传数据中的差异甲基化模式，从而达到识别分类的效果。癌症的特点是无限增殖，扩散和转移，所以分期可用于表示病情的发展和转移程度，目前常用的主要是 TNM 分期标准。Yu L 等人[24]运用 RF 和 SMOTE 算法，在 IA 至 IV 期 NSCLC 患者的 CT 图像中挖掘出了 NSCLC 诊断和病理分期的成像标志物，在 TCGA 验证集中表现优异，尤其是对于 LUAD。Li W 等人[25]在微生物组、转录组、微生物和转录组结合组中探索 RF 算法作为肺癌分期预测方法的准确性，结果表明，结合组的预测准确率以 0.809 位居最高。

4 AI 用于肺癌治疗

AI 可以根据临床数据、病理、分子数据和影像学数据（如肿瘤大小、转移）构建模型预测患者的生存期和治疗反应，帮助医生和患者选择最佳的治疗策略，制定个性化医疗方案，避免浪费医疗资源和宝贵的时间，减少不必要的药物暴露。Buizza 等人[26]能够将 PET/CT 放射组学特征应用于机器学习之上，用于评估化放疗期间的早期肿瘤反应，区分放化疗的早期反应者和无反应者。Guo 等人[27]通过 LR 模型和 ANN 模型建立肺癌患者化疗期间感染的预测模型，研究发现年龄、住院时间、手术史、联合化疗、骨髓抑制和糖尿病等是肺癌患者化疗期间感染的危险因素。Ahmed 等人[28]首先使用基因共表达网络来识别药物反应预测的代表性特征，再集成到 GNN 中预测肺癌的药物敏感性。Dercle 等人[29]提取 CT 放射组学特征预测肺癌患者对纳武单抗、多西紫杉醇和吉非替尼的敏感性，发现肿瘤空间异质性、边界和体积与总体生存率相关。Moran 等人[30]使用 LR 评估放射肺损伤的严重程度，以及区分无/轻度肺损伤和中度/重度肺损伤，最后实现预测 SABR 后 3、6 和 9 个月时肺损伤的严重程度。

5 AI 用于肺癌复发、转移和预后

肺癌复发和转移也是肺癌的重要病理特征，通常与恶性肿瘤的生长和扩散过程密切相关。并对患者的预后产生重大影响，准确评估复发和转移情况对制定合适的治疗计划至关重要。为了提高肺癌转移的准确性，研究者进行了很多努力，如利用影像学技术、分子生物学标志物等进行预测和诊断。Abdu-Aljabar RD 等人[31]使用 XGBoost 模型来分析基因表达和转录组数据增强肺癌检测能力，并预测手术后的复发情况。Akinici 等人[32]通过提取肿瘤和瘤周放射组学来评估 NSCLC 患者手术后的复发风险，与 TNM 分期系统的结合使用后，大大改善了单独 TNM 分期复发风险预估能力。Zhao 等人[33]开发了一种基于 CT 图像和临床知识的跨模态 3D 神经网络，用于预测临床 T1 期肺腺癌的淋巴结转移。

肺癌预后通常指的是预测肺癌的未来可能病程和结局，可以给患者和医者提供重要时间线索，以此辅助调整临床决策。Hosny 等人[34]通过 3D CNN 对 NSCLC 患者 CT 数据进行预后评估。其能够在放射治疗和手术数据集中将低风险和高风险死亡患者群体明显区分开。Maleki 等人[35]对包含临床信息的数据集使用 k-NN 算法，并应用遗传算法进行特征选择，此方法可有效支持肺癌分期诊断和预后预测运算。Oh S 等人[36]提出多模态深度学习模型，针对 PET 图像使用 ResNet3D 模型，针对临床数据使用 Cox 比例风险模型，二者结合能够改善 NSCLC 患者的生存预测性能。Lai YH 等人[37]开发了一种结合基因表达和临床数据的异构数据 DNN，可以用来准确预测非小细胞肺癌患者的总体生存率，AUC 为 0.816。Lee 等人[38]进行的研究结合使用了 CT 和临床特征，旨在替代基因组特征来预测早期非小细胞肺癌患者总体生存率。Guo H 等人[39]用 CNNs 提取的结构图像特征和冠状动脉钙化、肌肉质量、脂肪衰减度和肺气肿严重程度四个临床指标，构建 KAMP-Net 方法对肺癌筛查对象进行全因素死亡风险的预测。

6 总结

虽然 AI 已经在肺癌领域大放异彩，但是应用到临床仍然需要一段时间，不仅因为肺癌的高复杂性和高异质性，也因为成熟的 AI 需要大量数据去进行训练，而这是一个费时费力的过程。但不可否认的是，现阶段爆发增长的肺癌 AI 应用，可以辅助医生实现了高效率肺癌预测，获得更高的诊断准确率，推动肺癌精准医学治疗的发展。

PU-213

基于人工智能的免疫检查点抑制剂反应预测模型：现状与前景

陈永森[1,2,3]，渠鸿竹[1,2,3]，方向东[1,2,3]
1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室
2. 中国科学院大学
3. 基因组与精准医学检测技术北京市重点实验室

1.引言

肿瘤免疫疗法是一种革命性的治疗方法，而免疫检查点抑制剂（ICIs）则是该领域的重要组成部分。ICIs 通过解除肿瘤细胞对免疫系统的抑制，激活患者自身的免疫应答来对抗肿瘤[1]。然而，ICIs 并非对所有患者都有效，且其疗效和不良反应的预测仍然是一个具有挑战性的问题。随着人工智能技术的快速发展，利用人工智能在预测 ICIs 反应方面已经展现出巨大的潜力。人工智能可以利用大规模的临床数据、基因组学数据、转录组学数据和蛋白质组学数据等多种数据源，通过机器学习和深度学习算法构建预测模型，以预测患者对 ICIs 的疗效和不良反应，推动肿瘤患者的个性化精准治疗[2]。

2.人工智能预测模型的构建与优化

在预测免疫检查点抑制剂反应方面，人工智能模型的构建和优化起着关键的作用。传统机器学习算法在预测 ICIs 反应方面已经得到广泛应用。这些算法包括支持向量机（SVM）、随机森林（Random Forest）、逻辑回归（Logistic Regression）等。在使用这些算法时，首先需要将数据集划分为训练集和测试集，然后根据训练集进行模型训练，最后在测试集上进行模型评估。传统机器学习算法的优势在于具有解释性强的特点，可以帮助研究人员理解预测模型所依赖的特征和规律。此外，通过调整模型的超参数，如正则化参数、核函数选择等，可以进一步优化模型性能。

深度学习算法在近年来的预测 ICIs 反应方面表现出了巨大的潜力。卷积神经网络（CNN）[3]、循环神经网络（RNN）[4]和自注意力机制（Attention）[5]等技术已经成功地应用于图像、序列和文本等数据类型的建模。深度学习算法的优势在于能够自动学习特征表示，无需人工设计特征，并且能够处理复杂的非线性关系。针对 ICIs 反应预测问题，可以设计深度学习模型，将临床特征、基因组学数据和免疫特征等多种数据类型进行融合，以提高模型性能。

在构建预测模型时，特征选择和数据降维技术对于提高模型的准确性和泛化能力非常重要。特征选择技术可以帮助从大量特征中筛选出与 ICIs 反应相关的重要特征。常用的特征选择方法包括相关系数分析、方差分析、LASSO 等。数据降维技术可以减少数据集的维度，提高模型训练和测试的效率，同时防止过拟合。常见的数据降维方法包括主成分分析（PCA）、线性判别分析（LDA）等。

为了评估预测模型的性能并提高模型的泛化能力，模型交叉验证和超参数调优是必不可少的步骤。模型交叉验证可以通过将数据集划分为多个子集，在训练和测试中多次使用，以评估模型的稳定性和性能。常用的交叉验证方法包括 k 折交叉验证和留一交叉验证。超参数调优是指通过对模型的超参数进行搜索和调整，以最大化模型的性能。常用的超参数调优方法包括网格搜索、随机搜索和贝叶斯优化等。

通过以上方法和技术的应用，研究人员可以构建和优化高性能的预测模型，以预测患者对免疫检查点抑制剂的反应。在模型构建和优化过程中，需要注意训练集和测试集的划分、模型性能评估和泛化能力的验证等，以确保模型的准确性和可靠性。

3.生物标志物的发现

生物标志物在预测免疫检查点抑制剂反应中具有重要的作用，它们是可以测量的分子、细胞或影像特征，可以反映患者免疫状态和治疗反应的变化。以下是用于发现和验证与 ICIs 反应相关的生物标志物的常见方法和策略[6]。

生物标志物的发现需要综合考虑多种数据源，如基因组学、转录组学、蛋白质组学以及免疫细胞浸润等。一种常用的策略是通过对比 ICIs 反应患者与非反应患者的差异，识别与 ICIs 疗效相关

的特征，该策略可以通过统计分析、机器学习和深度学习等方法来实现。另一种常见的方法是借助生物信息学工具，如基因集富集分析、通路分析和网络分析等，来发现与 ICIs 反应相关的生物功能和通路。这些方法可以帮助我们理解 ICIs 治疗机制，并寻找与 ICIs 反应相关的潜在生物标志物。例如有研究者采用一种利用网络分析和机器学习框架来预测免疫检查点抑制剂（ICI）治疗反应的新方法。研究表明，基于该方法的预测结果比传统标志物更准确，为精准肿瘤学提供了新的生物标志物选择方案[7]。

成功发现和验证的生物标志物可以在临床实践中发挥至关重要的作用。这些标志物可以用于筛选患者，预测患者对 ICIs 的疗效和治疗反应，并为个体化治疗方案的制定提供依据。例如，标志物可以用于选择患者是否接受 ICIs 治疗，或者调整 ICIs 治疗方案的剂量和时间。此外，生物标志物的发现和验证也为未来的前瞻性研究提供了方向。研究人员可以进一步评估和优化已有标志物，探索更准确、可靠和易用的标志物。此外，随着技术的不断发展，新的生物标志物可能会被发现，并进一步加强我们对 ICIs 反应的预测和理解。在发现和验证生物标志物时，仍然面临一些挑战，如数据质量和规模的限制、标准化的问题，以及多中心合作和外部验证的需求。未来的研究应致力于克服这些挑战，以实现生物标志物在 ICIs 反应预测中的临床应用。

4.个体化治疗方案的优化与实施

个体化治疗方案的优化与实施是预测免疫检查点抑制剂反应的关键环节。通过预测模型可以根据患者的临床特征、生物标志物和其他相关信息，预测患者对免疫检查点抑制剂的疗效和不良反应。基于这些预测结果，可以制定个体化的治疗方案。例如，对于预测为对 ICIs 具有良好疗效的患者，可以选择 ICIs 作为一线治疗。对于预测为对 ICIs 疗效不佳的患者，可以考虑其他治疗方案，如化疗、靶向治疗或免疫联合治疗等。此外，对于不能耐受 ICIs 的患者，可以及早进行调整和管理，以减少不良反应的发生[8]。

通过定期检查和评估患者的临床状况、生物标志物和影像学特征等，可以及时调整治疗方案。如果患者在治疗初期表现出不良反应或疗效不佳，可以针对性地调整剂量、联合其他治疗或转换为其他治疗策略。此外，根据治疗后的疗效和不良反应，可以及时进行疗效评估和副作用管理，以确保患者获得最佳的治疗效果。

多学科团队的合作和交流,包括临床医生、病理学家、遗传学家、生物信息学专家等不同领域的专业团队的合作可以促进数据共享、信息整合和交流。通过共同讨论和解读患者的临床特征和生物信息学数据，团队成员可以共同制定个体化治疗方案，并协调监测和调整治疗效果。为了进一步优化个体化治疗方案，临床试验和后续研究是必不可少的。临床试验可以用于验证和优化预测模型的性能，并评估新的治疗策略的有效性和安全性。此外，后续研究可以关注长期治疗效果、治疗耐药性和生存结果等[9]。这些研究结果可以为个体化治疗方案的优化和实施提供更准确、全面和长期的指导。通过基于预测模型的治疗决策、治疗方案的监测和调整、多学科团队的合作和交流，以及临床试验和后续研究的推进，我们能够不断提高个体化治疗的水平和效果，为患者提供更优质的医疗服务。

5.挑战与展望

在免疫检查点抑制剂反应的预测和个体化治疗方案的优化与实施过程中，还存在一些挑战和需要解决的问题。免疫检查点抑制剂反应的预测涉及多种数据类型，如临床数据、基因组学数据、转录组学数据等。然而，这些数据的质量和标准化仍然是一个挑战。数据的准确性、一致性和可靠性对于建立可靠的预测模型至关重要。因此，在数据收集和处理的过程中，需要注意数据的质量控制和标准化处理，以确保数据的准确性和可比性。

在个体化治疗方案的优化和实施中，模型的解释性与可解释性也是一个重要的考虑因素。传统的机器学习算法通常具有较强的解释性，可以帮助研究人员理解模型依赖的特征和规律[10]。然而，随着深度学习技术的应用，深度神经网络等复杂模型的解释性较差。因此，如何在保持模型预测性能的同时提高其解释性，仍然是一个挑战。为了提高预测模型的泛化能力和可靠性，多中心合作和外部验证是不可或缺的。通过多个医疗机构和研究团队的合作，可以收集更大规模的数据集，并验证预测模型在不同患者群体和疾病类型中的适用性。然而，多中心合作和数据共享面临着数据隐私和安全性的挑战，需要制定相应的隐私保护措施和法规。尽管个体化治疗方案在理论上具有很大的潜力，但其实施和可行性仍然面临一些挑战。一方面，个体化治疗方案的制定需要多学科团队的合作和交流，以确保各专业领域的知识和技术可以得到充分利用。另一方面，个体化治疗方案的实施需要充分考虑临床实践的可行性和经济可承受性。因此，如何在实践中平衡个体化治疗与标准化治疗的效果和成本，仍然需要深入研究和探索。

尽管面临着各种挑战，未来免疫检查点抑制剂反应的预测和个体化治疗仍然具有广阔的展望。随着技术的不断发展，我们可以期待以下方面的进展：

数据资源的增加和丰富：随着越来越多的医疗机构和研究团队参与数据收集和共享，我们可以期待更大规模、更全面的数据资源，从而提高预测模型的准确性和可靠性。

新的生物标志物和数据类型的发现：随着技术的进步，我们可以期待发现更多新的生物标志物，并利用更多种类的数据类型，如代谢组学[11]、微生物组学[12]等，来预测和优化个体化治疗方案。

智能化和自动化的工具和平台：随着人工智能技术的发展，可以期待出现更智能化、自动化的工具和平台，用于数据分析、模型构建和治疗决策，以支持个体化治疗的实施和优化。总之，免疫检查点抑制剂反应的预测和个体化治疗方案的优化与实施是一个正在快速发展的领域。通过克服挑战和不断追求创新，我们可以为患者提供更准确、有效和个体化的治疗方案，为免疫治疗的未来发展开辟新的可能性。

PU-214

智能医学在推进社区医院建设中的应用

时文睿[1,2,3]，渠鸿竹[1,2,3]，方向东[1,2,3]

1. 中国科学院北京基因组研究所（国家生物信息中心），中国科学院基因组科学与信息重点实验室

2. 中国科学院大学

3. 基因组与精准医学检测技术北京市重点实验室

1 引言

由于人口数量引起的就医压力大一直是我国医疗体系中不可忽略的问题，造成了居民就医难，医生压力大的局面。针对这一问题，国家颁布《关于推进分级诊疗制度建设的指导意见》(国发办[2015]70 号)、《关于全面推进社区医院建设工作的通知》、《关于加快推进社区医院建设的通知》等一系列文件指出要加快推进社区医院建设，解决看病难、看病贵，实现小病在社区、大病进医院、康复回社区，体现了社区医院建设的重要性[1]。为推行社区医院，实施分级诊疗制度，各地医疗机构积极探索，但是仍有一系列问题阻碍着社区医院的发展建设，如医疗资源不足，专业人才缺乏，转诊流程不畅，居民认可度不高等[2]。

目前利用人工智能（AI）、大数据和可穿戴设备等技术结合医学，在医疗领域中进行数据分析、模式识别、决策支持等功能的智能医学正在兴起，以改进医疗服务、提高诊断治疗效果、促进健康管理和研究[3, 4]。已有研究提出构建互联网医院体系助力基层医疗服务，而在此基础上利用智能医学可以助力社区居民初次诊断的下沉，有助于解决目前社区医院遇到的困境，推动实现分级诊疗落地，真正实现社区首诊，提高社区卫生服务能力，解决医疗卫生资源供给不足，形成合理的就医格局，更好地建设社区医院[5-7]。

2 人工智能辅助社区医院治疗

人工智能利用机器学习、深度学习、自然语言处理等技术，对医学数据、影像、文献等进行分析和挖掘，从中发现规律、建立模型，并运用这些模型进行辅助诊断、治疗决策、健康管理等方面的工作，在医疗领域正在发挥越来越重要的作用[8, 9]。

2.1 智能辅助诊断模型提高诊断质量

人工智能可以利用患者的临床数据，影像等信息构建智能辅助诊断模型，提供信息支持并辅助医生进行疾病的诊断和治疗，已经在神经、循环和消化系统疾病中显示出良好应用前景。大型医院拥有海量的病人病历描述和疾病案例，这为智能辅助诊断模型的建立提供了数据支持，通过对疾病特征的选择和训练，构建智能辅助诊断模型[10]。而目前社区医院通常面临专科医生资源紧缺，医生水平不足的问题，导致居民对社区医院诊断结果的不信任，无法实现在社区医院实现首诊，这样的现实情况为利用大型医院建立的诊断模型反哺基层医疗诊断提供了可能性。智能辅助诊断的建立有助于大型医院的诊断成果共享至基层社区医院，可以帮助社区医生更快速、准确地诊断疾病，提高诊断效率和质量，有效提振居民对社区诊断结果可信性的信心，形成小病先去社区医院进行首诊的良性循环。已有研究分别使用智能辅助诊断模型对儿童社区获得性肺炎、白内障、社区糖尿病性视网膜病变进行检测，证明了智能辅助诊断模型在社区筛查中的准确性和可行性，为使用智能辅助诊断模型在社区医院进行辅助诊断提供了依据[11-13]。

同时，智能辅助诊断模型的应用也离不开社区医生的帮助。患者往往存在无法准确表述自己症状的情况，需要社区医生进行引导式询问，对病人进行专业的体格检查和问诊，确保病人症状的准确描述，提高智能辅助诊断系统信息输入的准确性，预测结果的准确性。智能辅助诊断模型与社区医生的组合有助于解决社区医院人才不足的问题，提升居民对诊断结果的认可度，提高居民进入社区医院的诊疗意愿。

2.2 智能治疗模型提供个性化治疗方案

人工智能还可以通过学习医生对病人开具的治疗及用药方案构建智能治疗模型，提供个体化的治疗指导。大型医院医生接触患者多，治疗方案先进，用药更加个体化，积累了大量的治疗和用药案例，而社区医院医生接触病例相对较少，治疗方案往往较为单一固定。针对这一情况，通过利用大型医院的大量不同病人治疗方案数据构建的智能诊疗模型，社区医生可以对患者开具针对性的治疗方案，可以实现治疗方案的共享，提高了治疗的科学性和有效性。通过不断更新智能治疗模型，保证用药和治疗方案的先进性。已有研究使用人工智能进行药物推荐以及术后辅助治疗方案推荐，均发现与真实临床决策方案一致性较高，有助于提升科学治疗，合理用药[14, 15]。

此外，由于患者个体特征，病情数据的不同，使用智能治疗模型医还可以预测不同药物对患者的疗效和副作用，提供个性化的药物推荐，帮助社区医院医生针对不同个体使用不同药物。此外，社区医院还发挥着方便长期服药患者取药的作用，智能诊疗系统有助于针对患者取药时的健康状态评估，及时调整药物类型以及药物剂量，方便对患者进行治疗后的追踪和预后评估。

2.3 疾病风险预测模型助力居民预防保健

预防是保证居民健康的重要环节，通过提前预防可以帮助居民降低疾病的发生风险，保持良好生活方式，减轻医疗费用和社会经济负担。利用大量罹患某种疾病的病人临床信息，可以提取到该疾病相关的危险因素，利用深度学习技术建立多种疾病的风险预测模型。例如近期使用年龄、收缩压、舒张压、降压治疗、当前吸烟、糖尿病和腰围作为危险因素，构建了中国心血管疾病风险预测模型，有助于检测心血管疾病发病风险[16]。

社区医院作为与居民接触最便捷的医疗服务站点，数量众多分布广泛，可以最大程度地发挥医疗宣传作用，对社区高龄居民进行普及性的疾病风险预测，帮助居民评估自身罹患某种疾病的可能性，督促居民保持良好生活方式，降低疾病发生风险。已有研究针对社区居民进行心血管疾病，高血压，老年认知衰弱进行了风险预测模型的验证，表明在社区预防保健中风险预测模型可以发挥重要作用[17-19]。对预防保健不仅关乎个体的健康，也对整个社会的福祉和可持续发展具有重要意义。在社区医院利用疾病风险预测模型进行居民疾病风险预防，可以降低社会的疾病负担和劳动力损失，促进人口健康水平的提升，从而推动社会的经济繁荣和稳定。

3 可穿戴设备助力社区居民健康

可穿戴设备是一种融合芯片技术、交互技术和通信技术的便携式电子设备，能感知和收集佩戴者活动情况和的用户即时状态信息，在医疗健康领域发挥着重要作用。健康监测、远程康复和疾病治疗是可穿戴设备在医疗中主要的应用[20, 21]。

3.1 健康监测设备保障居民健康

健康监测设备可以获得佩戴者的即时生理指标，如心率，体温等，使个人成为自身身体健康情况的第一知情人，帮助佩戴者及时发现身体状态异常，及时到医院进行诊断和治疗。大型医院以诊断治疗为重点，医生精力不足以兼顾大量病人治疗后的长久的健康情况，无法进行持久的追踪和监测。社区医院作为与居民最贴近，居民最方便的医疗服务站点，利用健康监测设备有助于居民健康的追踪监测，掌握居民健康情况，增强社区医生与社区居民的联系。

健康监测设备可以在社区医院中发挥健康提醒和辅助诊断、康复辅助、健康教育和行为改变的作用。通过健康监测设备，社区医生可以向患者发送健康提醒和警示，例如服药提醒、定期检查提醒等。健康监测设备还可以辅助医生进行初步的疾病诊断，通过收集和分析患者的生理数据，提供线索和参考，帮助医生作出准确的判断。此外，健康监测设备在康复中的应用也很广泛。例如，针对运动损伤的康复，可穿戴设备可以记录患者的运动轨迹、姿势和力度等信息，帮助社区医生评估康复进展并提供相应的康复指导[22]。最后，通过健康监测设备得到的生理参数、运动情况和睡眠质量等数据，医生可以向患者提供健康教育和行为改变的建议。例如，设置健康目标、监测锻炼情况、跟踪饮食习惯等，激励患者积极参与健康管理和自我调整，提高治疗效果和生活质量。已有研究验证了可穿戴设备可以改善脑卒中患者社区康复效果，展示了健康监测设备在改善居民健康中的潜力[23]。

3.2 远程康复设备助推居民健康管理

远程康复设备如血压检测仪、血糖检测仪、心电图检测仪和脉搏血氧仪等可以收集慢性病患者的相关信息，这对于监测患者身体状态和药物效果有重要作用。社区医院在慢性病的管理方面起到重要作用[24]。慢性病如高血压、糖尿病、心脑血管疾病等需要长期治疗和定期随访，社区医院可以对患者进行管理和监测。通过定期复诊、调整药物剂量、提供生活方式指导等方式，帮助患者保持良好的病情控制。患者通过远程康复设备得到的生理数据可以实时传输给社区医院信息平台上，使社区医生可以远程监测患者的健康状况，及时发现异常变化，并采取相应的措施。有研究表明可穿戴设备可以帮助管理慢性病患者群体，改善慢性病高危人群的不良生活习惯和行为方式，对预防慢性病发生和发展具有极其重要的意义[25]。

目前，社区医院居民电子健康档案建设存在着居民建档意识不强，对自身疾病存在回避等问题[26]。而通过远程康复设备获得的血压，血糖，血氧数据可以直观主动地反应居民身体指标，对健康数据进行监测和管理，帮助社区医院推动居民健康档案和电子健康记录系统建设。已有研究分别使用电子健康档案改善社区罹患糖尿病、冠心病合并高血压以及血液病居民的健康情况，实现社区精准化健康管理模式[27-29]。此外，健康档案信息有助于上级医院医生了解对患者过往身体情况，方便对患者身体健康信息进行评估，选择更科学有效的药物和手术方案，减小治疗副作用和手术并发症。利用好社区医院便捷以及数量多分布广的特点，建立去中心化的居民健康档案储存，减小大型医院信息存储压力，同时建立好居民电子健康档案与医院的共享，确保居民健康信息的及时传递，提高电子档案利用率，提升医疗服务质量。

4 结语

4.1 存在问题

智能医学可以助推社区医院发展，提高医事服务质量，但仍有许多问题亟待解决。许多社区医院仍缺乏影像设备，阻碍了智能辅助诊断系统在图像识别鉴定方面的应用[30]。社区医院仍需加强人才培养及培训，医生在使用人工智能辅助决策时仍然扮演着关键角色。人工智能算法提供的是参考意见和推荐，最终的决策权仍在医生手中，需要医生结合临床经验和患者情况进行综合判断，只有提升社区医生水平才能更好地发挥智能医学的作用。人工智能在医疗领域的应用仍处于不断发展阶段，需要不断验证和改进，医生与人工智能共同发挥作用，确保其诊断治疗的安全性和准确性。

4.2 展望

智能医学已经展示出其在未来帮助社区医院进一步建设的可行性和应用场景。随着智能医学的普及和发展，有助于解决社区医院现有发展困境，为社区医院的建设和发展带来更多机遇，加快实现我国分级诊疗落地。同时社区医院需要及时掌握和采纳新兴技术，加强人才培训和合作，积极推动医疗信息化和智能化的发展，提高社区卫生服务能力，从而为患者提供更加优质、便捷、高效的医疗服务。

PU-215

基于病人诊断与检验数据缺失的智慧诊疗

张海翔[1]，陈香梓[1]，李培培[2]
1 合肥市第二人民医院
2 合肥工业大学

1 引言

数字化时代医疗健康从原始的挂号问诊向精确智慧医疗预防方向发展[1]，且我国丰富的医疗数据资源给人工智能为首的大数据模型[2]建立提供基础，如何从这些历史沉淀的医疗数据中提取出有用的信息便于智慧医疗的开展成为新的研究方向[3]。

已有工作从一维到二维、三维医疗数据借助数据挖掘技术对这些数据进行分析，训练后的模型用来进行支持病人诊断预测、未来病种变化及病人病情追踪[4]。医疗数据预处理后在人工智能加持下，给出的预测结果为一线临床医生提供精准辅助决策，有效避免一些误诊带来的医疗纠纷。例如使用深度卷积神经网络对异常心脏数据[5]进行分析；采用 ConvNeXt 网络与 BioBert 编码的语义向量[6]对 X 光图像结果智能分析；还有学者将架构 EfficientNetB4 与转移学习相结合对胸部 X 光图像分类[7]。给患者带来及时有效的病情确认，有助于医生更好的对症下药，实现疾病早发现早治疗。

而在现实生活中并非每次都可以获取到所有的历史数据，患者诊疗维度较长数据丢失，同一病种不同医生的检查检验处方不同，患者基础疾病不全等情况都会对人工智能神经网络模型训练带来误差。此时直接利用这些医疗数据进行训练、预测结果辅助分析[8]，给出的结果不仅没有达到辅助决策效果有时还会误导临床医生病情判断，因而对于缺失问题研究是我们合理使用医疗数据进行智慧医疗研究中必须考虑的内容之一。本文主要研究一维数据下的缺失问题，即借助抽血、体液检查等化验操作得到的患者属性体征，对患者病情做出基础疾病诊断预测。

处理缺失问题在数据层面通常分为三种问题：仅诊断数据缺失，仅体征数据缺失，诊断和体征数据都缺失。解决仅诊断数据缺失问题[9]常见有预处理策略恢复缺失的诊断矩阵然后训练模型或假设已知缺失诊断的位置在训练模型时忽略这些数据；解决仅体征数据缺失问题[10]有调整不同体征的重要性以适应当前数据内容。但是这两种方法并非每次都可以获取的另一维度的完整数据，给缺失数据处理带来误差。

基于此需要进一步研究和开发更有效的方法，利用诊断特定的体征信息来解决缺失体征和诊断的问题。本文提出基于缺失体征和诊断的诊断特定体征智慧诊疗方法，该方法通过学习体征和诊断相关矩阵来恢复缺失的数据矩阵。为降低模型的复杂性利用范数对系数矩阵约束。利用体征相关性相关的体征具有相似系数向量，为了更好体现出诊断间的关系对预测结果借助诊断相关性进行约束。

2 缺失体征和诊断的体征特定诊断预测模型

2.1 问题定义

本节给出缺失数据预测问题定义：给定数据集 $U=\{(X,Y)\}$，矩阵 $X=[x_1,x_2,...,x_n]^T\in R^{n\times d}$，$X$ 为 d 维度的患者体征数据集，$Y=[y_1,y_2,\cdots,y_n]^T\in\{-1,1\}^{n\times l}$，有 l 个诊断的诊断集，且当患者满足其中第 i 个诊断时 $y_i=1$，不满足等于 0。其中缺失环境下矩阵 X 和矩阵 Y 并非完整，$x_{ij}=0$ 表示患者样本 x_i 的第 j 个体征信息缺失，$y_{ij}=0$ 表示对应的诊断是缺失的或患者没有改诊断信息。

2.2 缺失数据恢复与诊断特定体征征学习

在模型训练过程中如果使用的患者沉淀医疗数据是不完整，得到的模型会降低新预测患者的诊疗预测结果[11]。因此为克服这种现实问题，本文提出补全缺失数据矩阵的方法，即学习患者体征相关性矩阵和诊断相关矩阵补全缺失的原始矩阵，新的补全矩阵较缺失矩阵内容更加完整、丰富。我们设置体征相关者矩阵中 A_{ij} 表示第 i 个体征信息和第 j 个体征间相关性，诊断相关性矩阵中 C_{ij} 表示第 i 个诊断和第 j 个诊断之间的相关性。满足相关性将 A_{ii} 和 C_{ii} 设置为 1，利用缺失数据借助强相关信息来恢复选择线性回归损失函数。补全问题可描述为：

$$\min_{W,A,C}\frac{1}{2}\left\|XAW-YC\right\|_F^2+\frac{\lambda_1}{2}\left\|XA-X\right\|_F^2+\frac{\lambda_2}{2}\left\|YC-Y\right\|_F^2+\lambda_3\left\|W\right\|_1 \quad (1)$$

其中 $\lambda_1,\lambda_2,\lambda_3$ 为权重参数，$W=[w^1,w^2,\cdots,w^l]\in R^{d\times l}$ 表示系数值，在补全原始缺失矩阵 $X(Y)$ 时引入正则化项 $\left\|XA-X\right\|_F^2$ 和 $\left\|YC-Y\right\|_F^2$ 控制新的补全后补全 $XA(YC)$ 与 $X(Y)$ 之间的差异。为提取诊断的特定体征信息，本文使用 l_1 范数正则化约束[12,13]去除每个诊断的不相关体征提高医疗分类性能，同时降低模型的复杂度。

2.3 学习患者体征相关性与诊断相关性

在预测分类过程中探索体征和诊断之间的相关性可以提高分类性能。而在处理缺失数据时直接使用不完整信息计算向量间的欧氏距离结果必然存在许多误差，因而学习一个体征相关性矩阵 A，在分类构建过程中引导模型参数 W 的学习[14]，如果两个体征数据 f_i，f_j 之间存在相关性，则相应参数向量 w_i 和 w_j 也相似即：

$$\sum_{i,j}^{d}A_{ij}\left\|w_i-w_j\right\|_2^2=tr(W^TL_1W) \quad (2)$$

其中 $L_1 = D - A \in R^{d \times d}$ 是体征相关性矩阵 A 的图拉普拉斯矩阵，$D_{ii} = \sum_{j=1}^{d} A_{ji}$ 是对角矩阵。同理探索诊断间的相关性也可提高分类性能。因而假设诊断 y_i 和 y_j 相关，XAw_i 和 XAw_j 输出结果相似[15]，有：

$$\sum_{i,j}^{l} C_{ij} \left\| XAw_i - XAw_j \right\|_2^2 = tr(XAWL_2(XAW)^T) \quad (3)$$

其中 $L_2 = D - C \in R^{l \times l}$ 是诊断矩阵相关 C 的图拉普拉斯矩阵，$D_{ii} = \sum_{i=1}^{l} C_{ii}$ 是对角矩阵。由于原始矩阵信息的缺失无法直接计算出相关性矩阵结果。因而需要计算 W 时同步迭代 A, C，最终优化问题：

$$\min_{W,A,C} \frac{1}{2}\left\| XAW - YC \right\|_F^2 + \frac{\lambda_1}{2}\left\| XA - X \right\|_F^2 + \frac{\lambda_2}{2}\left\| YC - Y \right\|_F^2 + \lambda_3 \left\| W \right\|_1 + \frac{\lambda_4}{2} tr(W^T L_1 W) + \frac{\lambda_5}{2} tr(XAWL_2(XAW)^T) \quad (4)$$

λ_4, λ_5 为权重参数。

2.4 分类模型优化

因为数据缺失我们无法直接从原始的矩阵中计算出相关性矩阵值，采用交替迭代算法进行优化。在每个迭代步骤中更新单个参数同时固定其他参数步骤如下：

1）更新体征相关性 A：固定模型参数 W 和诊断相关性矩阵 C 式（4）重写为：

$$\min_{A} \frac{1}{2}\left\| XAW - YC \right\|_F^2 + \frac{\lambda_1}{2}\left\| XA - X \right\|_F^2 + \frac{\lambda_5}{2} tr(XAWL_2(XAW)^T) \quad (5)$$

通过目标函数梯度下降算法优化 A：

$$A := A - \alpha \nabla f(A), s.t. A \geq 0; \quad (6)$$

其中 $\nabla f(A)$ 是 A 的梯度，α 是下降梯度步长。此时 A 的迭代公式可写为：

$$\nabla f(A) = X^T XA(W^T W + \lambda_1 I + \lambda_5 W L_2 W^T) - X^T YCW^T - \lambda_1 X^T X \quad (7)$$

2）更新诊断相关性 C：固定模型参数 W 和 A，式（4）重写为：

$$\min_{C} \frac{1}{2}\left\| XAW - YC \right\|_F^2 + \frac{\lambda_2}{2}\left\| YC - Y \right\|_F^2 \quad (8)$$

目标优化方法 C 与 A 一致，$\nabla f(C)$ 是 C 的梯度，δ 是梯度下降步长。

$$C := C - \delta \nabla f(C), s.t. C \geq 0; \nabla f(C) = (I + \lambda_2) Y^T YC - Y^T XAW - \lambda_2 Y^T Y \quad (9)$$

3）更新系数参数 W：固定矩阵 A 和 C，式（4）重写为：

$$\min_{W} \frac{1}{2}\left\| XAW - YC \right\|_F^2 + \lambda_3 \left\| W \right\|_1 + \frac{\lambda_4}{2} tr(W^T L_1 W) + \frac{\lambda_5}{2} tr(XAWL_2(XAW)^T) \quad (10)$$

系数矩阵 W 受 l_1 范数正则化约束因而采用加速近端梯度法 $\min F(W) = f(W) + g(W)$ [16]解决。，根据公式（4） $f(W)$ 和 $g(W)$ 定义如下：

$$f(W) \min_{W} \frac{1}{2}\left\| XAW - YC \right\|_F^2 + \frac{\lambda_4}{2} tr(W^T L_1 W) + \frac{\lambda_5}{2} tr(XAWL_2(XAW)^T) \quad (11)$$

$$g(W) = \lambda_3 \left\| W \right\|_1$$

其中 $f(W)$ 是利普希茨连续梯度，表示成 $\|\nabla f(W_1)-\nabla f(W_2)\|_F^2 \le L_f \|\Delta W\|$，$L_f$ 是利普希茨常数，$\Delta W = W_1 - W_2$，由近端梯度算法最小化方法，$F(W)$ 二次近似表示为：

$$Q(W,W^{(t)}) = f(W^{(t)}) + <\nabla f(W^{(t)}), W - W^{(t)}> + \frac{L_f}{2}\left\|W - W^{(t)}\right\|_F^2 + g(W) \tag{12}$$

定义 $G^{(t)} = W^{(t)} - \frac{1}{L_f}\nabla f(W^{(t)})$，此时有：

$$\begin{aligned} W &= \arg\min_W Q(W,W^{(t)}) = \arg\min_W g(W) + \frac{L_f}{2}\left\|W - G^{(t)}\right\|_F^2 \\ &= \arg\min_W \frac{1}{2}\left\|W - G^{(t)}\right\|_F^2 + \frac{\lambda_3}{L_f}\|W\|_1 \end{aligned} \tag{13}$$

当迭代至 ψ_t 时满足 $\psi_{t+1}^2 - \psi_{t+1} \le \psi_t^2$，$W^{(t)} = W_t - \frac{\psi_t - 1}{\psi_t}(W_t - W_{t-1})$，系数矩阵 W 优化过程：

$$W_{t+1} = S_{\frac{\lambda_3}{L_f}}[G^{(t)}] = \arg\min \frac{\lambda_3}{L_f}\|W\|_1 + \frac{1}{2}\left\|W - G^{(t)}\right\|_F^2 \tag{14}$$

$\nabla f(W)$ 可表示为：

$$\nabla f(W) = A^T X^T (XAW - YC) + \lambda_4 L_1 W + \lambda_5 A^T X^T XAWL_2 \tag{15}$$

给定 W_1，W_2，$\Delta W = W_1 - W_2$ 可得到：

$$\begin{aligned} &\|\nabla f(W_1) - \nabla f(W_2)\|_F^2 = \left\|A^T X^T XA\Delta W + \lambda_4 L_1 \Delta W + \lambda_5 A^T X^T XA\Delta W L_2\right\|_F^2 \\ &\le 3\left\|A^T X^T XA\Delta W\right\|_F^2 + 3\left\|\lambda_4 L_1 \Delta W\right\|_F^2 + 3\left\|\lambda_5 A^T X^T XA\Delta W L_2\right\|_F^2 \\ &\le 3(\left\|A^T X^T XA\right\|_2^2 + \left\|\lambda_4 L_1\right\|_2^2 + \left\|\lambda_5 A^T X^T XA\right\|_2^2 \cdot \left\|L_2\right\|_2^2)\|\Delta W\|_F^2 \end{aligned} \tag{16}$$

利普希茨常数表达为：

$$L_f = \sqrt{3(\left\|A^T X^T XA\right\|_2^2 + \left\|\lambda_4 L_1\right\|_2^2 + \left\|\lambda_5 A^T X^T XA\right\|_2^2 \cdot \left\|L_2\right\|_2^2)} \tag{17}$$

3 实验结果与分析

3.1 实验数据集与评价指标选取

本文选取多个公共开源数据集 Genbase，该数据集共有 645 条数据，体征类别 1186 维，诊断类别 27 个。为了模拟缺失数据情况我们在预处理时随机选择 80%的数据量用于训练，且按照缺失比例 10%递增缺失情况，对于缺失位置的值设为 0。

评价指标选取分类任务中常见的 Average Precision、Ranking Loss、Coverage、One-error 和 Hamming Loss。

$$AvePrec = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{|Y_i^l|}\sum_{y_j \in Y_i^l}\frac{\left|\left\{y_q \in Y_i^l : R_i(y_q \le R_i(y_j))\right\}\right|}{R_i(y_j)} \tag{18}$$

$$RankingLoss = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|(y_a, y_b) : R_i(y_a) > R_i(y_b), (y_a, y_b) \in Y_i^l \times Y_i^l\right|}{\left|Y_i^l\right|\left|Y_i^l\right|} \tag{19}$$

$$Coverage = \frac{1}{n}\sum_{i=1}^{n}\max_{y\in y_i} rank_f\left(x_i, y\right) - 1 \quad (20)$$

$$One-error = \frac{1}{n}\sum_{i=1}^{n}\left\{\left[\arg_{y\in Y_i}\max f\left(x_i, y\right)\right] \notin Y_i\right\} \quad (21)$$

$$hLoss(h) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{l}\left|h(x_i)\Delta Y_i\right| \quad (22)$$

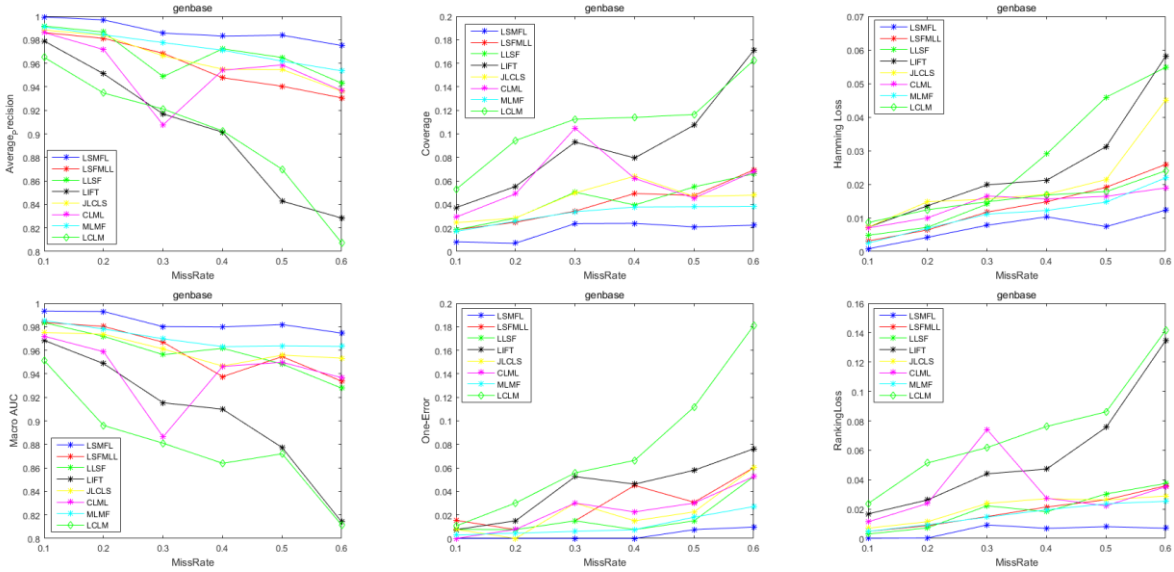
3.2 对比方法介绍

为了验证所提方法的有效性，我们选择当前主流预测分类方法，各方法介绍如下：LSFMLL, LLSF, LIFT, JLCLS, CLML, MLMF, LCLM。LLSF[13]利用线性回归方法进行标签特征学习时，通过最小二乘法来最小化预测标签与真实标签之间的误差，以求得最优的标签特定特征。LSFMLL[3]通过学习到的特征表示和度量矩阵来提高分类或回归任务的性能。LIFT[17]通过对正例和负例的聚类分析学习标签特定的特征。JLCLS[18]结合标签补全和标签特定的特征学习能够充分利用存在的标签信息，并通过学习到的特征表示提高多标签分类任务的准确性和泛化能力。CLML[14]借助相关性在分类过程学习公共特征和标签特定特征。MLMF[19]解决缺失特征问题、利用相关性信息构建分类器。LCLM[20]在缺失标签环境下引入公共特征和标签特定特征学习模型。

3.3 实验结果分析

通过实验结果可以明显看出，所提出的 LSMFL 方法在分类过程中处理缺失数据方面具有明显的优势。尤其在 Average Precision 和 Ranking Loss 等指标上显示出了出色的表现。随着缺失率的逐渐增加，所有算法的处理性能普遍下降，这是一个普遍存在的趋势。然而，LSMFL 方法通过利用体征间的相关性和诊断间的相关性来恢复缺失数据，并通过对系数矩阵施加范数限制来提取诊断特定的体征。实验结果表明，即使在面对缺失数据的挑战时，LSMFL 方法仍然能够有效地保持良好的性能。

需要注意的是，当缺失率从 20%增加到 30%时，与 LSMFL 相比，其他对比算法如 LLSF、CLML 和 LIFT 的性能立即下降。与之相比，LSMFL 方法展现出相对稳定的性能，差异微乎其微。这一观察结果表明，LSMFL 方法具有更好的稳健性，即使在存在缺失数据的情况下也能保持其有效性。此外随着缺失率的增加，每个对比算法的性能都会迅速下降。而 LSMFL 性能下降的速度相对较慢，尤其在 Average Precision、One-Error 和 Macro AUC 等指标上表现出较好的结果。因此，综合来看，LSMFL 方法处理缺失数据的能力以及对于高缺失率情况下的鲁棒性表现出了显著的优势。



4 结语

本文的提出的算法 LSMFL 能够有效处理存在缺失体征和诊断的智慧医疗分类问题。在这个过程中，它通过学习体征相关性矩阵和诊断相关性矩阵来恢复缺失的数据矩阵，并训练分类器。同时

为了降低模型复杂度对系数矩阵进行范数限制，以提取与诊断相关的患者体征信息。此外，LSMFL 根据体征相关性的假设，即相关体征具有高相似度的系数向量，来更好地利用体征之间的关系。最后，为了更有效地捕捉诊断之间的关系，LSMFL 还对输出的诊断相关性进行了限制。实验结果在多个数据集上验证了我们方法的优越性。

PU-216

医学知识大模型语料数据集构建研究

吕婷钰[1]，李晓瑛[1]，刘宇炀[1]，杜晋华[2]，李心怡[2]，罗妍[1]，唐小利[1]，
任慧玲[1]，刘辉[1*]，尹浩[2*]
1. 中国医学科学院医学信息研究所 100005
2. 清华大学网络大数据研究中心 100084

1 引言

大语言模型（Large Language Models，LLMs）采用深度学习技术，通过预训练学习大量自然语言文本，发挥自注意力机制和 Transformer 结构优势实现自然语言处理任务[1]。生成式预训练转换器（Generative Pre-trained Transformer，GPT）等大语言模型在自然语言处理中取得了巨大成功，其以生成式的自监督学习为基础，从 TB 级训练数据中学习隐含的语言规律和模式，训练出千亿级别参数量的大规模语言模型，使机器能够理解、回答和生成人类语言[2]。而这种颠覆性技术发展的一个关键目标是与医学健康领域有关的智能应用，因此医学知识大模型备受公众瞩目。

医学领域 LLMs 一定程度上能改变人们与在线医学信息系统的交互方式，如利用 BioGPT[3]能在医学信息检索中丰富和改进搜索结果、辅助临床决策、总结医学文献等。虽然 LLMs 能够分析大量数据，但由于医学健康领域的专业性和高度复杂性，它们无法完全学习医学专家对信息的判断决策和长期累积的知识经验[4]。利用大规模的医学论文数据进行预训练的医学大语言模型，通过增强其专业知识覆盖度和理解能力，能够提升大模型在多种医学自然语言处理任务上的性能[5]，提高医学知识问答能力。此外，在积极研发医学知识大模型、提升其专业知识理解能力、生成能力之余，对这类大模型的定量评估亦是当前的一个关键。本研究尝试构建中文医学知识语料数据集，建立中文医疗问答能力评测基准，促进中文医学知识大模型不断发展和应用创新。

2 国内外相关研究进展

目前，国内外正处于医学语料数据集建设和医学知识大模型研发训练、评测应用的关键阶段。医学数据资源类型和规模体量十分庞大，且应用范围不断扩大，为构建医学领域的语料数据集奠定了基础[6]。

2.1 国际研究进展

通过开展文献调研和网络调研，总结国际上广泛使用的医学语料数据集包括 MedQA、MedMCQA、MultiMedQA、PubMedQA 等。其中，MedQA 是第一个公开可用的针对医疗问题的大规模开放领域的多项选择数据集[7]，主要收集了美国医师执照考试题，支持大模型利用广泛和先进的医学领域知识来回答问题；同时，该数据集研发团队还收集并发布了一个大规模的医学教科书语料库。MedMCQA 包含来自两个印度医学院入学考试的模拟考试和历年发布的考试多项选择题，主题多样丰富[8]，涵盖 2000 多个医疗保健主题和 21 个医学科目。MultiMedQA 是 Google Research 团队为评估 LLMs 编码临床知识、回答医学问题的能力而建立的语料数据[9]，用于回答医学考试、医学研究等相关的问题，作为评估医学领域 LLMs 临床知识和问答能力的多样化基准。PubMedQA 来源于 PubMed 论文题目和摘要结论句[10]，也是第一个针对生物医学科研文本进行推理来回答医学问题的问答数据集。

表 2-1 国际医学语料数据集建设

序	数据名称	数据概况

号		
1	MedQA	收集自美国、中国大陆和中国台湾的医学委员会考试题，涵盖英语、简体中文和繁体中文的 12,723、34,251 和 14,123 个问题。
2	MedMCQA	整合超过 19.4 万条高质量印度医学院入学考试的多项选择题，涵盖 2,000 多个医疗保健主题和 21 个医学科目。
3	MultiMedQA	包含 HealthSearchQA 的 3,173 个在线搜索医学问题与六个现有的开源数据集。
4	PubMedQA	拥有 1,000 个专家注释、60,000 多个未标记和 21,000 多个人工生成的医学论文问答实例。

2.2 国内研究进展

在国际医学大语言模型和语料数据集启发下，中文医学语料数据集建设取得一些显著成果，典型代表如 MLEC-QA、CMExam、CMB 等。MLEC-QA 是目前规模最大的中国执业医师考试题，涵盖临床、口腔医学、公共卫生、中医、中西医结合等多个学科[11]。CMExam 为第一个提供医学注释的中国临床医学检查数据集，包含医学考试的多项选择题和医学专家的注释[12]。中文医疗模型评估基准 CMB 包括了不同临床职业、不同职业阶段考试中的多项选择题 CMB-Exam 和基于真实病例的复杂临床诊断问题 CMB-Clin[13]。CMB 数据集能更好地基于中国本土文化环境提供一个更加情境化的基准，支持评估医学领域 LLMs 的综合知识问答和推理能力。

表 2-2 国内语料数据集建设

序号	数据名称	数据概况
1	MLEC-QA	由五个子集构成，包含 136,236 个生物医学领域的多项选择题，并配有医学专家注释的图像或表格。
2	CMExam	包含来自中国执业医师考试的 60,000 多道多项选择题，医学专家标注包括疾病群组、临床科室、医学学科、能力领域和问题难度级别五个问题的注释。
3	CMB	包括中国执业医师考试中的多项选择题和基于实际病例研究的复杂临床诊断问题。临床诊断问题则基于实际的病例进行设置，由医学专家达成共识后确定正确答案。

3 语料数据集构建研究

自然语言理解的长期目标为构建通过自然语言进行推理的智能系统，而自动问答任务是评估智能系统推理能力的通用方法之一。评估医学领域 LLMs 需要标准化、规范化和统一的数据集。目前国际相关机构建设医学语料数据集大多基于美国执业医师考试题、英文医学论文摘要和临床医学检查等数据，而针对中文医学论文的语料数据集构建实践较少。因此，本研究通过分析中文医学论文题目和摘要特征，总结归纳了多种中文医学论文摘要的语料数据集构建模式，形成多样化的中文医学论文摘要问答数据集。再者，基于长期的中文医学专业名词建设成果，构建了医学名词解释问答数据集。此外，整理了近年来中国执业医师考试真题，作为 CMExam、CMB 的补充数据；并且收集了 Chinese-medical-dialogue-data、MedDialog 等开源问答数据集，提升医学知识大模型的多轮问答能力。最后，这些数据集均转化为大模型调优与评测可直接使用的规范格式，以期作为评测医学健康领域 LLMs 的基准。

3.1 中文医学论文摘要问答数据集

基于自主加工的中文医学论文题目、摘要和摘要结论句，构建中文医学论文摘要数据集，形成由问题、上下文和长答案组成的结构化问答对。其中，“问题”衍生自现有的医学论文标题，“上下文”是论文的摘要部分，“长答案”为摘要的结论部分。鉴于中文医学论文很少直接以疑问句作为论文

标题，故需要对论文标题经过特定的改造作为问题提出，并将论文摘要结论句作为问题的答案（即长答案）。目前已初步整理出六类摘要处理模式 pattern（见表 3-1），后续将进一步补充完善。

表 3-1 中文医学论文问答数据集构建模式

序号	模式	问题	回答
1	A 对 B 的敏感性	miR-503 下调 Bcl-2 增强 BEL-7402/5-FU 细胞对 5-氟尿嘧啶的敏感性有影响吗？	miR-503 可能通过下调 Bcl-2 的表达,进而增强 BEL-7402/5-FU 细胞对 5-氟尿嘧啶的药物敏感性,抑制细胞增殖。
2	A 对 B 的干预作用	维康颗粒对亚健康疲劳小鼠不同组织氧化应激损伤的干预作用是什么？	健康疲劳状态下小鼠存在显著地氧化应激损伤,维康颗粒可通过降低模型小鼠的氧化应激损伤,促进亚健康疲劳状态向正常状态转化。
3	A 的药理作用	菊芋多糖的药理作用是什么	菊芋多糖具有降血糖和抗病毒的作用,药物安全性好,有较高的开发价值。
4	A 对 B 的影响	Ad-HIF-1α 双突变体对大鼠缺血心肌及 Na~+-K~+ATP 酶的影响是什么？	AdHIF-1α-564/402 能减轻心肌细胞梗死程度,保护心脏功能,对 Na~+-K~+ATP 酶水平无明显影响。
5	A 与 B 的关系探讨	HCG 日孕酮水平与体外受精-胚胎移植治疗结局是什么关系？	HCG 日血清孕酮浓度水平在 3.0~4.77 nmol/L 之间能获得更成熟的卵子,同时不增加因多卵泡发育引起卵巢过度刺激综合征(OHSS)导致的周期取消率。
6	A 与 B 的比较	MRA 与颈部血管彩超诊断颈动脉狭窄的比较结果是什么？	MRA 对脑梗死患者颈部血管检查准确率优于颈部血管彩超,当患者颈部血管病变越严重时两者差异越大,且两种检查方法均与 DSA 有一定的差距,应仔细斟酌两者的结果。
7	A 与 B 的关联性	血脂相关基因 DNA 甲基化水平与冠心病有什么关联性？	两组间脂代谢相关基因的 DNA 甲基化水平存在统计学差异,提示脂代谢相关基因的 DNA 甲基化在冠心病中起一定的调控作用。
8	A 的临床疗效	阿替普酶静脉溶栓对穿支梗死患者的临床疗效如何？	采用 rt-PA 静脉溶栓治疗穿支梗死患者,有利于降低炎症因子水平及神经功能缺损程度,能提升临床疗效,临床应用安全可靠。
9	A 的治疗策略	颅颈交界区髓周动静脉瘘患者血管构筑特点及治疗策略是什么？	颅颈交界区 PAVF 患者多伴脊髓动脉参与供血和动脉瘤样结构,多以出血起病,更多使用介入栓塞联合显微手术治疗。PAVF 患者预后与非 PAVF 组类似,大多患者预后良好。
10	A 对 B 的治疗效果	氯吡格雷联合丁苯酞注射液对急性脑梗死的治疗效果如何？	氯吡格雷联合丁苯酞注射液对急性脑梗死患者治疗的效果较好,能促使患者的神经功能、抗氧化功能得到改善,值得在临床上推广。

3.2 医学名词解释问答数据集

领域专业词汇对提升 LLMs 的学科知识理解能力亦非常重要[14]。研究所构建的中文医学语料数据集，涵盖 2.7 万个来自《中文医学主题词表》（Chinese Medical Subject Headings, CMeSH）的医学名词及其定义注释信息。具体而言，从词表中获取“名词”和“定义注释”部分，以“什么是{名词}？”的格式作为问题提出，并以“{定义注释}”作为问题的答案。值得提出的是，该数据集收录了疾病、解剖学、有机体、化学物质和药物、诊疗技术与设备等多种类别的医学知识，有助于提升医学知识覆盖度。表 3-2 列举了其中的肿瘤、病毒、身体部位共三类的问答数据。

表 3-2 医学名词解释

序号	名词	提问 Q	回答 A
1	炎性乳腺肿瘤	什么是炎性乳腺肿瘤？	转移性乳腺癌，其特点为由淋巴转移和最终癌细胞阻塞淋巴管所致的受累乳腺的水肿和红斑。
2	肝炎病毒	什么是肝炎病毒？	任何引起肝脏炎症的病毒。包括来源于人类和动物的 DNA 病毒和 RNA 病毒。
3	消化系统	什么是消化系统？	一组从口腔延伸至肛门的器官，用于分解食物，吸收营养和排泄废物。人类消化系统包括胃肠道和附腺（肝、胆道、胰腺）。

3.3 中国执业医师考试真题（补充）

语料数据集是否能够及时更新，是衡量其内容质量和生命力的一个重要指标。为此，本研究收集了近年来中国临床执业医师考试真题，作为已有 CMExam、CMB 语料数据集的补充。下例为 2022 年一道单选题，经格式转化后的数据。

QU 提问：急性肾损伤后，患者出现肾小管原尿回漏的原因是？
IT 选项：A 肾小球滤过率降低
B 肾间质水肿，挤压肾小管
C 输尿管梗阻
D 肾小管上皮细胞坏死、脱落
E 肾小管的水钠重吸收减弱
AN 答案：D

3.4 开源问答数据集

开源数据集含有大量真实和丰富的医学知识，为构建和扩展中文医学语料数据集提供了重要的资源。具体而言，Chinese-medical-dialogue-data 数据集[15]涵盖了内科、妇产科、肿瘤科等六个医学领域，包含了共计 792,099 个问答对。MedDialog 中文数据集[16]则收录了来自医生和患者之间的 110 万条中文对话，数据总量达到 400 万。此外，Huatuo Encyclopedia QA 数据集[17]不仅收集了中文维基百科中的疾病和药物百科条目，而且覆盖来自钱问健康网站上的医学文章，共计 364,420 条医学问答数据。

综上所述，本研究构建了一套包含中文医学论文问答数据集、医学名词解释问答数据集及中国执业医师考试真题（补充）的中文医学语料数据集库，并收集了若干开源医学问答数据。这些数据集可用于医学大模型对中文医学基础知识、临床问题等相关问答任务的定量评测，并将不断扩充、更新完善，努力成为中文医学大模型的专业知识和理解能力、生成能力的评估基准。今后，将持续开展相关研究，进一步把电子病历、在线健康社区等真实世界数据纳入中文医学语料数据集构建中，提升医学知识大模型在临床智能诊疗、病历自动书写等任务上的应用效果，拓展新一代人工智能技术的应用范围。

4 结语

自主构建中文医学语料数据集，为研发医学知识大模型提供坚实的基础数据支持，有助于实现标准化和客观全面的定量评估，在语料数据层面推动中文医学知识大模型的创新发展和应用落地，从而提升医学垂域大模型的知识覆盖度、问答能力和行业认可度，促进医学科技自立自强。未来将进一步利用电子病历、在线健康社区等真实世界数据，不断扩充丰富中文医学语料数据集，以提高中文医学大模型的智能问诊能力，为健康中国战略实施提供更广泛、更坚实的 AI 支持。

PU-217

基于机器学习的成人严重细菌感染临床时序预警模型：一项回顾性研究

刘剑 裘云庆*
浙江大学医学院附属第一医院

研究背景：

感染性疾病占发热待查的首位[1]。严重的感染性疾病可迅速发展为脓毒症，感染性休克，甚至是死亡。在目前的数字时代，电子健康记录是一个庞大的电子数据存储库，代表海量的的临床信息[2]。机器学习的技术有独特的能力去分析庞大的数据，因为他们可以以灵活和可训练的方式处理大型数据集，并理解变量之间的复杂关系[3]。有关人工智能机器学习的方法在成人严重细菌感染中的报道较少。在本篇论文中，回顾性分析了发热待查患者的临床数据，建立了两个序贯预测模型去帮助临床医生做出合理的临床抉择，并进一步在感染患者中早期发现严重细菌感染的患者。

方法：

研究设计

这是一项回顾性的观察性队列研究。入选人群是 2010 年 1 月到 2020 年 12 月期间就诊于一家拥有 2200 张床位的教学医院的患者。

入组标准

患者入院原因为发热，基于发热待查专病库，具体入组标准如下述：（1）诊断为“发热待查”或者主诉为发热若干天；（2）年龄大于 18 岁；（3）体温＞37.5℃被认为是发热。

排除标准

患者如果满足以下标准将会被排除：（1）年龄＜18 岁；（2）怀孕；（3）明确诊断为血液疾病，包括白血病、免疫性血小板减少症、地中海贫血和多发性骨髓瘤等；（4）被确诊患有自身免疫性疾病，如血管炎，强直性脊柱炎，或系统性红斑狼疮等；（5）被确诊为肿瘤，如肺癌、肝癌、胃癌、卵巢癌等；（6）缺少必须的数据。

数据变量收集

住院记录中提取相关的感兴趣变量，包括性别、年龄、身体质量指数(BMI)、既往疾病（包括高血压，糖尿病，心脏病，肾脏病，乙肝及恶性肿瘤）、24 小时入院内的实验室检查结果。

数据特征分布

非正态分布的连续变量用四分位差表示，非正态分布的二分类变量用频率(比例)表示。采用 Wilcoxon 秩和检验和卡方检验对严重细菌感染组与非严重细菌感染组的差异进行统计学评估。变量之间的相关性使用 Pearson 相关系数进行测量，并使用热图进行可视化。所有统计分析和算法均使用 R -4.0.2 版统计软件。

特征选择

分别采用 3 种算法：Boruta、Lasso、RFE 对特征进行筛选。

结果

患者纳入和描述性数据

在本研究中，945 例临床诊断为感染性发热的患者被纳入分析。数据集以 7:3 的比例随机分为训练组和测试组。共有 661 例患者用于训练集，284 例用于测试集。患者的中位年龄为 54 岁(IQR: 38-66 岁)，47.4%的患者为女性。最常见的共病为高血压(27.9%; 264/945)，而其他合并症，如糖尿病、心脏病、肾病、乙型肝炎和恶性肿瘤分别 12.1% (114/945), 6.3% (60/945), 5.5% (52/945) , 4.3% (41/945), 和 3.7% (35/945)。

感染的特点

从严重感染发作部位来看，肺部感染是最常见的三种临床感染类型，约占 24%（224//945），尿路感染约占 9%（83/945），腹腔感染约占 8%（78/945）。根据临床病程和微生物学资料，717 例（76%）被诊断为细菌感染，185 例（20%）被诊断为病毒感染，17 例（2%）被诊断为真菌感染，11 例（1%）为寄生虫感染，15 例（2%）为未知病原体。62 例感染患者(7%)的微生物样本呈阳性，分离出细菌病原体 42 例，真菌病原体 13 例。

模型结果比较

在早期预测模型中，随机森林模型表现出最高的准确性，在训练集和测试集分别为 0.731（95%CI=0.695-0.764）和 0.729 （95%CI=0.637-0.780）。在入院 24 小时内的风险评估模型中，逻辑回归模型的准确率最高，在训练集和测试集分别为 0.777（95%CI=0.742-0.808）和 0.754 （95%CI=0.698-0.804），见表 1。

结论

早期预测成人严重细菌感染是一项重要的临床需求。在这项研究中，我们开发了两个模型(早期入院评估模型和入院 24 小时内的风险评估模型)。这些模型可以帮助临床医生在成人发病后早期识别严重的细菌感染，从而实现早期临床监测和治疗，避免不必要的抗生素使用。未来需要根据临床应用场景进行前瞻性临床试验，以确定两种模型的临床效用。

表 1 训练集和测试集的模型特点

早期预测模型

列	模型	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	LR(+)	LR(-)
练	LR	0.756(0.715-796)	0.684(0.647-719)	0.638(0.592-682)	0.785(0.723-840)	0.869(0.828-903)	0.494(0.438-550)	2.973(2.268-898)	0.461(0.400-531)
	RF	0.723(0.679-766)	0.731(0.695-764)	0.811(0.772-846)	0.551(0.480-621)	0.801(0.761-836)	0.568(0.496-638)	1.808(1.544-118)	0.342(0.273-429)
	XGboost	0.796(0.758-833)	0.770(0.736-802)	0.805(0.765-840)	0.693(0.625-755)	0.853(0.816-886)	0.615(0.549-678)	2.619(2.122-232)	0.282(0.229-347)
试	LR	0.686(0.617-754)	0.634(0.575-690)	0.605(0.534-672)	0.709(0.596-806)	0.844(0.775-898)	0.409(0.326-496)	2.078(1.447-982)	0.557(0.447-695)
	RF	0.685(0.612-759)	0.729(0.673-780)	0.790(0.728-844)	0.570(0.453-681)	0.827(0.766-877)	0.511(0.402-619)	1.836(1.411-389)	0.368(0.265-511)
	XGboost	0.708(0.641-775)	0.676(0.618-730)	0.663(0.594-728)	0.709(0.596-806)	0.855(0.791-906)	0.448(0.359-540)	2.279(1.594-258)	0.475(0.374-603)

入院 24 小时预测模型

列	模型	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	LR(+)	LR(-)
练	LR	0.807(0.771-844)	0.777(0.742-808)	0.839(0.801-872)	0.640(0.569-706)	0.837(0.799-870)	0.643(0.572-710)	2.330(1.928-815)	0.252(0.199-320)
	RF	0.809(0.772-846)	0.767(0.732-799)	0.775(0.733-813)	0.750(0.684-808)	0.872(0.835-904)	0.602(0.539-664)	3.100(2.426-962)	0.300(0.248-363)
	XGboost	0.846(0.812-879)	0.748(0.713-782)	0.698(0.652-740)	0.860(0.804-905)	0.916(0.881-944)	0.564(0.506-620)	4.984(3.516-065)	0.351(0.302-409)
试	LR	0.780(0.720-841)	0.754(0.698-804)	0.776(0.711-832)	0.697(0.581-798)	0.869(0.809-915)	0.546(0.442-648)	2.563(1.807-635)	0.322(0.239-434)
	RF	0.759(0.691-826)	0.706(0.648-759)	0.694(0.624-758)	0.737(0.623-831)	0.872(0.809-920)	0.483(0.389-577)	2.637(1.790-885)	0.415(0.324-533)
	XGboost	0.780(0.717-843)	0.717(0.659-770)	0.689(0.619-753)	0.789(0.681-875)	0.894(0.834-938)	0.496(0.404-588)	3.272(2.096-108)	0.394(0.311-500)

PU-218

基于混合云架构的医院信息系统安全外联及管控实践

沈亮[1] 周敏[1]
1.浙江大学医学院附属第一医院，浙江杭州 310003

为推进实施健康中国战略，国务院办公厅于 2018 年正式发布《关于促进“互联网+医疗健康”发展意见》，坚定不移地推动互联网与医疗健康深度融合，同时强调行业监管和数据安全保障[1],并且新冠肺炎爆发以来，患者对基于互联网的远程医疗服务需求也呈现爆发式增长。在国家政策及患者需求的双重加持下，医院面向互联网的应用服务与日俱增，信息系统的服务边界不断延伸，而基于物理隔离的医院传统网络架构已很难适应当前需求发展，一方面传统网络架构较为封闭，无法充分满足“互联网+”时代下深度数据融合的需求，另一方面构建于传统网络架构下的外联信息系统已成为导致医院信息安全风险的主要因素[2]。鉴于云计算技术在医疗行业逐渐得到应用并且在安全防护和安全治理层面所展现的独特优势[3]，我院于 2017 年开始引入云计算技术，构建混合云平台来统一管控信息系统的安全外联。

一. 医院信息系统外联的安全风险分析

为实现高效的医院运营以及提供优质的患者远程服务，医院信息系统服务边界不断从医院业务内网向互联网延伸。医疗健康领域由于其数据的高敏感性和隐私性，数据安全及患者隐私保护一直是普遍关注的焦点[4]。而医疗机构信息系统从传统封闭性网络快速延伸至互联网的过程中，由于其技术滞后、管理缺失等原因，仅满足医院信息系统的外联需求并未从根本上改进医院网络安全架构而导致面临较大的安全风险[5]。技术层面的安全风险包括网络安全防护体系不完善、主机安全策略不全面、数据和应用防护策略单一、安全监测与应急响应不及时等，管理层面的安全风险包括缺少统一安全运营中心、缺乏全面的审计能力、信息安全专业人员不足等。

二. 浙大一院基于医疗混合云的安全管理框架设计

基于医院传统网络安全架构下信息系统外联的安全风险分析，存在不可避免的安全管理缺陷，同时考虑医疗数据本身的敏感性和隐私性，避免数据直接存储于公有云而面临由于失去数据控制权及云平台管控权限发生数据泄露的风险。最终确定在医院内部网络建设拥有独立管控权限的专有云平台，同时结合具备动态、全栈安全防护能力的公有云平台，构建我院医疗混合云平台。

（一） 医疗混合云的网络基础架构

我院建设的医疗混合云平台主要是基于医院数据中心本地化部署专有云平台，同时结合院外的公有云专有网络(Virtual Private Cloud, VPC)，两侧通过合理网络规划实现物理专线互联[6]。我院医疗混合云平台模型如图 1 所示。

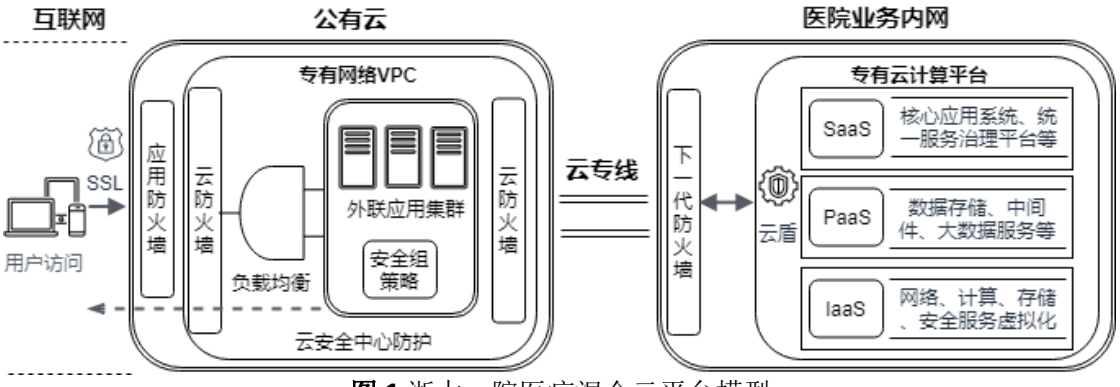


图 1 浙大一院医疗混合云平台模型

（二） 信息系统外联的安全管理框架

面向混合云平台动态的虚拟化网络，安全管理服务基于软件定义网络（Software Defined Network, SDN）/网络功能虚拟化（Network Function Virtualization, NFV）等技术，形成安全资源服务化交付、弹性伸缩、安全服务定制化的安全管理框架，同时不同安全资源可以相互作用，形成系统化的整体防护能力和动态化的适应能力[7]。图 2 为我院混合云模式下信息系统外联的安全管理

架构，基于等级保护 2.0 标准指导及测评[8]，云安全架构包括网络通信安全、主机安全、应用安全、数据安全 4 个层面，同时构建安全运营管理。

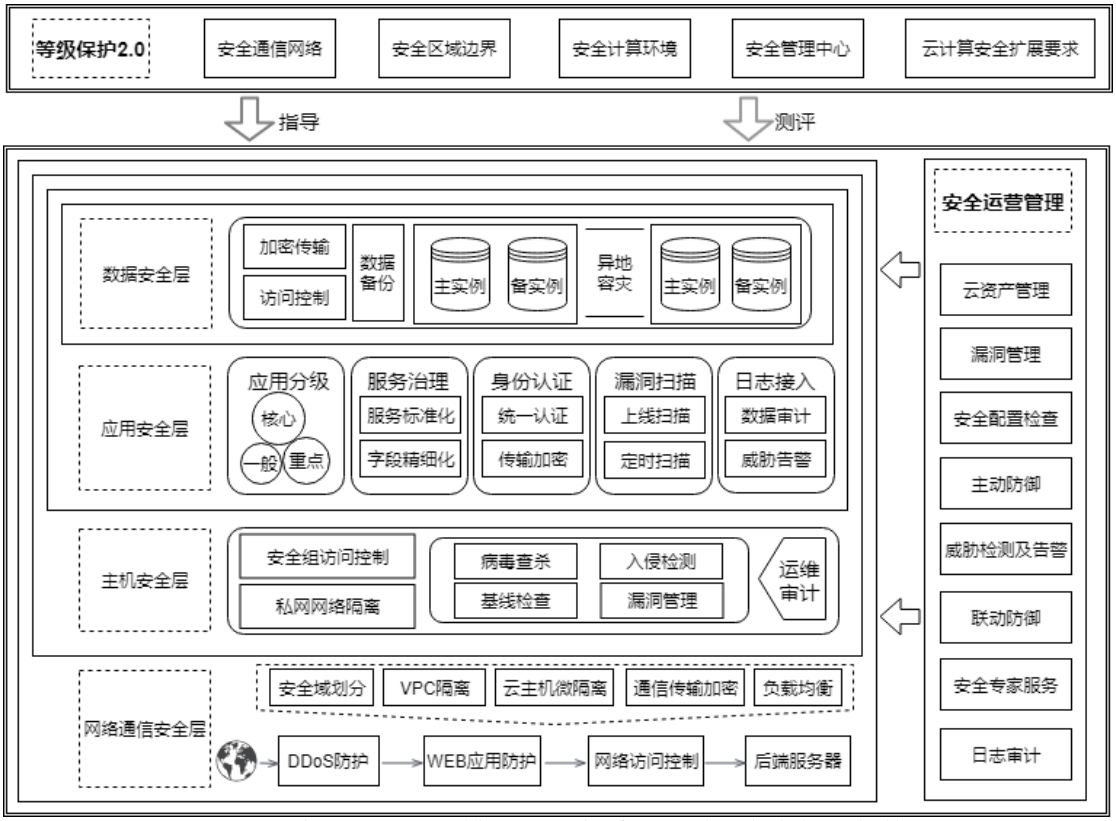


图 2 浙大一院混合云模式下信息系统外联的安全管理架构

下面内容将主要围绕我院基于混合云架构来实现医院信息系统外联的安全管控体系，重点从安全访问控制、安全数据存储、云安全等保合规三个方面进行详细的实践阐述。

三. 混合云模式下信息系统外联的安全访问控制

混合云的公有云区域作为互联网接入区域，等同于将传统网络安全架构下安全隔离区（Demilitarized Zone, DMZ）进行整体云化，依赖公有云提供的全栈安全防护体系进行集中防护，通过网络通信安全、云主机安全、应用与数据传输安全和安全运营管理四个层面的综合安全治理来实现信息系统外联的安全访问控制及安全风险规避[9,10]。

（一） 信息系统外联的云原生网络通信安全

为保证医院信息系统外联的安全可控，实施同质化、细粒度的安全访问控制，首先对外联信息系统的网络流量出入口进行统一化，实现以资源为中心的安全保护，能有效的实施并建立访问控制点，收敛资产暴露面。其次通过对公有云的 VPC 规划，同时执行细粒度的边界安全防护及策略优化。图 3 展示医院信息系统外联的网络通信安全层面的整体架构设计。

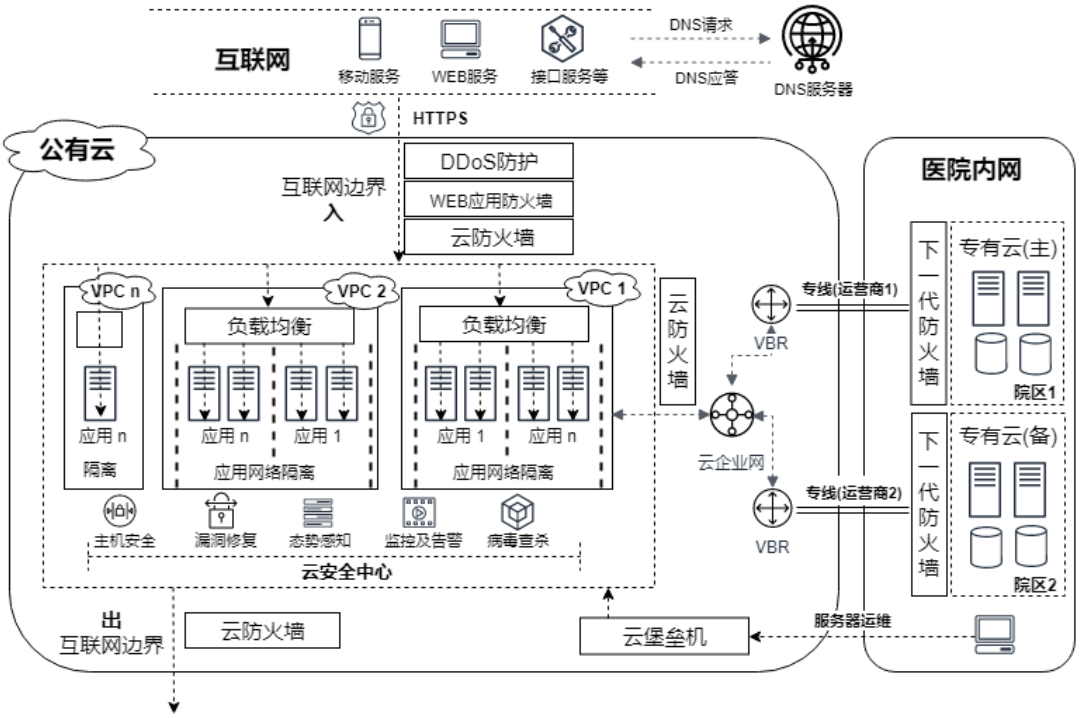


图 3 医院信息系统外联的云原生网络通信安全架构

（1）VPC 规划：VPC 基于 OverLay[11]技术在物理网络基础上构造虚拟网络以作为云上用于构建逻辑隔离分区的基本域，VPC 内可自定义网络环境（子网、路由表等）以创建并运行云资源，划分的不同 VPC 间使用 Vxlan 协议[12]对网络进行隔离，保证二层逻辑隔离，无法进行网络通信。我院以是否与院内核心信息系统存在数据交换和功能分区两个指标作为 VPC 划分的依据，如表 1 所示。基于 VPC 的分级划分，可通过物理专线接入云企业网实现云上 VPC 按需与医院自有机房的专有云平台互联，构建医疗混合云。

与院内信息系统数据交换	功能分区	VPC 划分	典型的应用
是	生产业务	VPC 1	互联网医院平台、患者移动服务平台等
否	生产业务	VPC 2	医院官方门户网站、图书馆网站等
否	运维服务	VPC 3	云资源个性化运维管理相关应用

表 1 浙大一院公有云 VPC 规划

（2）互联网边界安全防护：互联网边界安全防护主要针对 Web 资产在南北向流量的入口和出口进行有效管控，防范来自互联网的恶意入侵风险。流量入口部署原生 DDoS 防护、原生 Web 应用防火墙（Web Application Firewall, WAF）、原生云防火墙、原生负载均衡服务等安全产品进行融合联动，**相比于传统安全架构，云原生安全产品采用内嵌方式而非外挂部署模式，可充分利用云原生技术实现弹性扩容，以及获取云平台全网安全大数据进行整合分析，并且在处置风险时各安全产品间有效联动以实现安全运营的自动化闭环**。流量出口主要经由 NAT 服务后再通过原生云防火墙进行统一访问控制。

（3）混合云边界安全防护：混合云边界安全防护主要针对公有云 VPC 与医院内网专有云平台间数据交互的访问控制。在访问控制层面，首先依据 VPC 划分，保证 VPC 级别的网络隔离，其次公有云侧的互联 VPC 内通过云主机的安全组规则和云防火墙对入流量和出流量进行控制，同时医院内网侧采用下一代防火墙进行访问控制，实现双重异构防火墙策略进行安全防护。

（4）VPC 内应用边界安全防护：VPC 内应用边界安全防护主要针对单 VPC 内承载各应用运行的云主机间的东西向流量进行访问控制。**相比于传统网络架构，由于云计算的虚拟化、容器技术等实现了资源的细粒度隔离，物理设备不再是资源承载调度的最小单元，物理安全边界消失，因此**

应用之间的安全隔离变得十分关键。VPC 内云主机间的东西向流量访问控制主要通过安全组规则配置对云主机的入流量和出流量进行控制[13]。

（二）信息系统外联的云原生主机安全

云原生主机作为承载应用运行的容器，实现计算、存储等资源的即开即用以及弹性伸缩能力。医院外联信息系统主要由第三方服务商提供，需要将云主机提供给第三方服务商用于应用的部署及更新操作，云主机可以定义为一个不可信的环境。以下从三个层面进行策略的部署及实施。

（1）云主机网络隔离与访问控制：云主机的网络隔离设置保证云主机仅配置 VPC 内的私网地址，可避免在公网环境下暴露云主机实例从而减少外部攻击的风险。云主机的访问控制策略设置方面必须遵循最小授权原则，基于安全组策略可对传输协议、地址、端口等实施细粒度控制，实现对云主机的入流量和出流量进行严格的限制。

（2）云主机运维安全认证及管理：云主机的三方运维基于云原生堡垒机进行安全认证及管理，与安全组策略结合实现对我院云主机资产的统一管理、集中运维。

（3）云主机内部实时安全防护：云主机内部通过云安全中心统一管理，采集云主机内各类数据并结合云平台全网威胁情报数据，实现多维度、高效的关联分析，同时结合漏洞库及规则库的自动及时更新，为云主机提供操作系统漏洞自动修复、病毒自动查杀、智能化主动防御、威胁告警与处置等功能。

（三）信息系统外联的应用与数据传输安全

医院外联应用主要来自于第三方服务商提供，医院对于应用程序代码的安全性也无法及时掌握。鉴于应用自身安全的不可信，一方面通过基于网络通信安全防护，在互联网流量统一入口处部署云原生 WAF 来有效识别业务流量的恶意特征，另一方面我院通过应用分类分级管理、统一服务安全治理、统一身份认证及加密传输管理、统一漏洞扫描策略、统一应用日志管理等五个方面进一步强化应用安全。

（1）应用分类分级管理：基于与核心系统数据交换、互联网入、互联网出、服务人群、数据分类 5 项指标进行应用分级，将应用划分为一般应用、重点应用和核心应用，如表 2 所示。基于应用分类分级，一方面可实现按需安全防护，如与核心系统数据交换需求决定部署的 VPC 类型，对于医院官方门户网站等其他门户网站将开通网页防篡改策略等。另一方面在进行应用的安全评估过程中，通过对应用的交互设计方案进行重构实现对应用降级处理，如原本需要开放互联网入流量的变更为仅需开放互联网出流量等。

应用级别	分类指标					典型应用
	与核心系统数据交换	互联网入	互联网出	服务人群	数据分级	
一般应用	无	开放	开放	无限制	一般	医院官方门户网站、医院图书馆网站等其他信息展示类网站
重点应用	有	不开放	开放	系统	重点	医院接入移动支付服务、医疗数据上报服务等
核心应用	有	开放	开放	员工	重点	办公自动化系统、移动钉钉办公、CA 电子签名服务等
	有	开放	开放	患者	核心	互联网医院平台、患者移动服务平台等

表 2 应用分类分级

（2）统一服务安全治理：针对外联信息系统与院内核心系统间存在的数据交换，统一要求接入医院的服务安全治理平台，这也是从应用层面对混合云边界实施访问控制，实现数据服务标准化定义、数据字段精细化管理和数据交互安全管控。

（3）统一身份认证及加密传输管理：用户身份认证是应用进行访问控制的核心环节，而弱密码、Cookie 劫持、SQL 注入等应用漏洞严重影响身份认证的有效性。一方面尽量避免采用独立建设的用户认证模式，基于“零信任”理念[14]将应用接入微信、支付宝、钉钉等第三方可信应用认证网关平台，另一方面对于必须采用独立建设的用户认证模式的应用，对身份认证策略进行统一规范及

渗透测试，包括采用密码安全策略、验证码校验方式、应用错误返回自定义页面等要求。数据传输方面，云原生 WAF 可直接关联 SSL 证书，使用 HTTPS 协议对用户与应用间的交互访问进行全链路数据加密。

（4）统一漏洞扫描策略：云原生漏洞扫描针对应用可能存在的弱密码、Web 注入漏洞、文件漏洞等威胁进行全面检测。可按应用级别设定周期性的漏洞扫描任务。

（5）统一应用日志管理：外联信息系统运行过程中的业务日志统一接入云原生日志服务，实现业务日志的数据采集、数据存储、数据加工、查询与分析、可视化、告警、消费与投递等。结合日志审计服务对接入日志服务的外联信息系统业务日志进行威胁情报检测。

（四）信息系统外联的云原生安全运营管理

信息系统外联的一体化安全闭环管理机制，包括信息系统外联的需求评估及安全设计、安全策略配置及论证、安全运营管理及日志审计等，如表 3 所示，形成医院信息系统外联贯穿上线前、上线中、上线后的一体化安全闭环管理机制。同时依托云安全中心建设完善的安全运营体系[15]，基于态势感知[16]的体系化运营管理，通过安全服务联动管理，安全运营可视化，实现统一云资产管理、统一漏洞修复及基线检查、统一安全策略检查、统一主动防御配置、统一威胁检测及告警、日志审计等。

时间阶段	核心流程	流程工作描述
系统上线前	系统需求评估	对外联信息系统的需求分析及外联的必要性评估
	系统安全设计	依据应用分类分级标准进行划分，同时对网络安全架构、端口开放、数据接口服务、用户身份认证安全等进行重点分析并提出对应的优化策略
系统上线中	数据服务接入	应用统一接入服务治理平台进行数据服务的联调
	日志服务接入	应用产生的操作日志统一接入日志服务平台，可用于系统维护查询和业务日志审计
	身份认证论证	对用户身份证认证策略进行论证
	云资源分配	按需分配云数据库、云服务器、负载均衡服务、NAT 网关服务、域名服务、堡垒机服务等资源
	安全策略配置	安全控制策略配置，包括 WAF 域名绑定、云防火墙配置、云主机安全组配置、数据库访问控制、SSL 证书配置等
	系统漏洞扫描	系统对外开放访问前进行弱口令检测、身份认证漏洞、Web 注入漏洞、文件漏洞、前端漏洞、错误配置等进行全面渗透扫描及整改
系统上线后	定期漏洞扫描	按应用级别设定周期性的漏洞扫描任务
	安全运营管理	基于云安全中心统一管控所有的云资产并监控云上业务整体安全，实现态势感知的体系化安全运营管理
	日志审计	对流量日志、事件日志和操作日志等进行行为分析、安全分析、变更行为追踪和行为合规性审计

表 3 信息系统外联的安全评估及论证核心流程

四. 混合云模式下信息系统外联的安全数据存储及访问

基于医疗数据安全及隐私保护考虑，外联信息系统的数据将统一存储至构建于医院内部网络进行本地化部署的专有云平台，一方面利用云计算架构赋予数据存储高可用性、强扩展性及高安全性等能力，另一方面依托专有云的私有化部署模式实现数据存储本地化和云平台管控权限自主化，可避免因失去数据控制权所导致的数据风险[17,18]。混合云模式下信息系统外联的安全数据存储架构如图 4 所示。

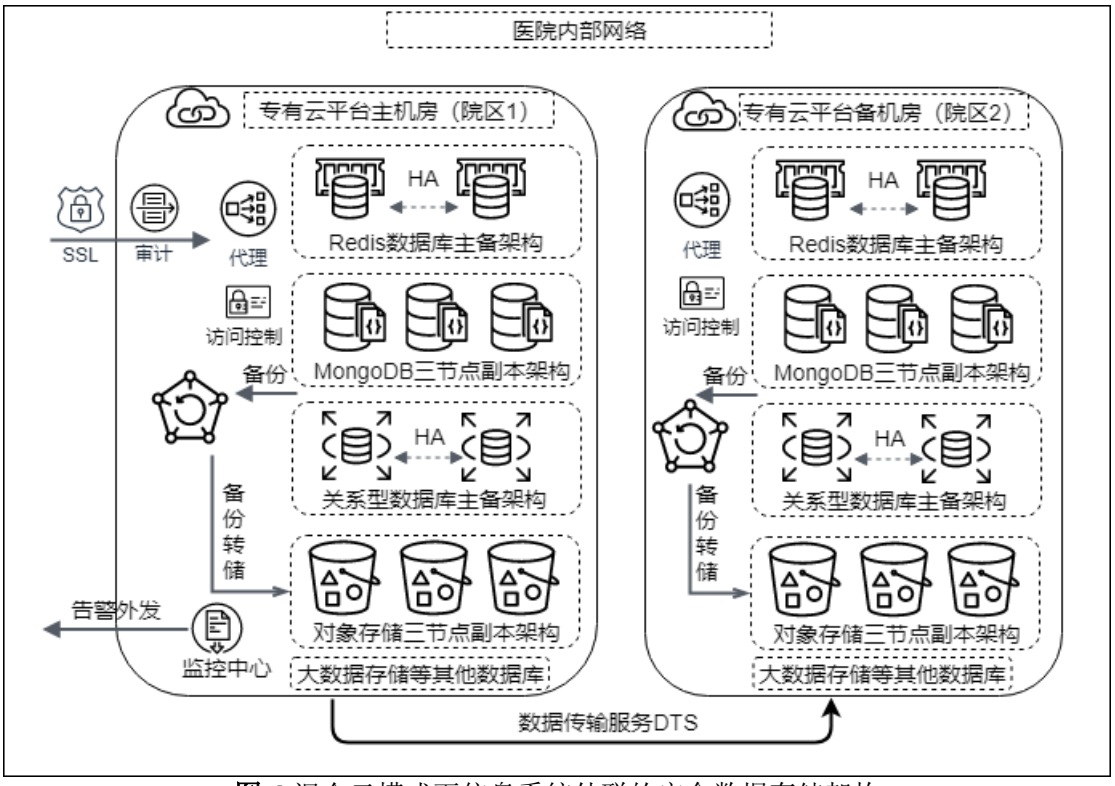


图 4 混合云模式下信息系统外联的安全数据存储架构

（一）数据统一存储及管理

基于云原生的数据统一存储管理，有利于采取统一的安全数据存储及访问管理策略，包括访问控制、传输加密、资源监控、备份与恢复、数据容灾、数据库审计、数据库可视化运维等。同时我院建设的专有云平台中包含了各种类别的云原生数据库以满足医院各类外联信息系统的数据存储需求，包括关系型数据库、NoSQL 数据库、分析型数据库、对象存储 OSS、大数据存储等。

（二） 数据存储高可用和高扩展性

基于专有云平台的云原生架构体系，云数据库采用主备高可用模式，同时实现宕机自动检测和故障自动转移。同时基于监控及告警，可实时监控数据库实例的相关性能，包括 CPU、内存、磁盘、连接等。分布式云存储的高扩展性主要包含水平可扩展性和垂直可扩展性两个维度，水平可扩展性是通过存储内部扩展管理的服务器节点数量实现，垂直可扩展性可在管理服务器节点不变的情况下，扩展单节点的性能[19]。

（三）数据备份及容灾

数据库备份主要通过设置自动备份策略，按天全量物理备份结合实时数据库日志备份，备份数据转储到对象存储 OSS 进行多副本存储。异地容灾主要通过创建异地灾备实例，基于我院在两院区分别建设的专有云，以院区 1 为主机房，院区 2 为备机房，通过数据传输服务（DTS）实现主实例和异地灾备实例之间的实时同步，实现 RPO≈0,RTO≈5 分钟。当前我院正在进行基于两院区云平台的容灾升级，最终将实现基于云原生架构下的业务同城双活。

五. 混合云模式下信息系统外联的云安全等保合规

等级保护 2.0 在等级保护 1.0 的基础上，为适应云计算等新技术的发展，专门针对云计算等新技术的个性安全保护需求提出安全扩展要求[20]，正好符合我院当前混合云模式下信息系统外联的安全合规性检查需求[21]。通过对标等级保护 2.0 的第三级安全通用要求和扩展要求，进一步完善我院混合云安全架构，并将混合云平台和其上部署的外联信息系统归并统一进行定级和测评，使其整体满足等级保护 2.0 的第三级要求[22]。

六. 讨论

构建医疗混合云平台用以解决医院信息系统外联的安全风险，相比于医院传统网络架构模式存在明显优势。我院目前信息系统外联数量将近 50 余个，均以混合云架构模式运行，互联网入口区域流量的日均检测威胁数量 3 万+，运行至今暂未发生安全事故。与医院传统网络架构下的安全管理模式相比，混合云平台基于云原生网络通信安全、云原生主机安全、应用与数据传输安全、云原

生数据存储安全以及云原生安全运营实现全栈安全管控，特别是云原生安全产品服务化、基于云原生态势感知的主动防御、云原生安全产品动态感知、云原生数据安全存储机制、经济效益等方面。

云原生安全产品服务化，基于软件定义安全架构，通过 NFV 和 SDN 技术相互融合，以编程的方式进行统一、按需、灵活的安全服务编排和管理[23]，能有效解决传统网络安全架构下的专用硬件安全设备面临部署成本高、升级困难、灵活性差等问题[24]，按需定制网络安全服务将实现更加灵活、动态的安全防护机制。

基于云原生态势感知的主动防御，传统网络安全架构的防御技术无法提供智能高效的网络安全分析与预测，只是通过建立防火墙、访问控制等多层次防御体系进行被动防御，基本遵循对已知安全风险的被动监测及响应，而云原生态势感知基于云平台全网安全大数据的充分共享、智能分析，可以实现对系统漏洞和威胁的理解，并据此对系统的安全状况进行更加全面的评估和预测[25]，从而对潜在的威胁进行及时预警，实现转被动防御为主动防御[26]；

云原生安全产品动态感知，云原生安全产品通过与云平台深度融合，实现对云资产更有力的控制，同时基于网络安全服务链[27]等技术，各原生安全产品间有效协同，一方面自动识别云资产，另一方面对风险资源进行主动处置，例如在发现安全威胁时，不仅仅生成告警信息，还能够自动联动相关云资产或其它原生安全产品，迅速对安全威胁采取处置措施和防护手段，实现从检测、告警到处置的安全运营自动化闭环。解决了传统网络安全架构下安全设备相互孤立、外挂式部署、安全数据分散、被动式处置的安全管理模式；

云原生数据安全存储机制，基于云原生分布式数据库进行统一存储管理，一方面采取统一的数据存储及访问控制策略，解决传统网络安全架构下数据存储分散、容灾备份机制不完善等问题[28]，另一方面对比将数据存储于公有云方案，可有效解决由于数据控制权及云平台管控权缺失等原因，导致医疗数据将面临泄露的安全风险及合规性挑战[29]；

经济效益层面分析，与传统分散式安全产品堆叠相比，基于混合云的一体化安全管理模式，同时依赖于云上安全产品的虚拟化技术及弹性伸缩能力，通过与我院核心业务内网的传统网络安全建设及管理方式对照，基于混合云架构的信息系统外联解决方案能有效降低安全建设及运维成本，同时提高了安全管理效率，在其他行业也有类似经验[30]。

PU-219

医疗行业云桌面系统技术架构与应用场景的探讨

雷保仓 何斌
漯河市中心医院

1 引言

2021 年，国务院颁布《关于推动公立医院高质量发展的意见》指出，要强化信息化支撑作用，“推动云计算、大数据、物联网、区块链、第五代移动通信（5G）等新一代信息技术与医疗服务深度融合推进电子病历、智慧服务、智慧管理“三位一体”的智慧医院建设和医院信息标准化建设。”随着云计算和虚拟化技术在医疗行业中的广泛应用，使医院的工作流程发生了巨大改变，为医务人员提供了灵活多样的工作方式；尤其是在智慧医院建设的进程中，面对信息系统数量增多、终端设备庞杂、应用场景日趋复杂、维护难度不断增大、故障修复时间急需缩短等问题，如何利用信息化手段改善患者就医体验，提高医护人员工作效率，为不同的终端用户提供个性化配置成为医院面临的重大挑战，相对传统 PC 云桌面技术作为一种新型的解决方案，在提升桌面故障响应效率和桌面安全性，降本增效，保障不同角色使用体验感方面具有明显优势。

2 云桌面技术概述

云桌面办公系统是一种基于云计算技术的虚拟化桌面环境，它允许用户访问云服务器上的虚拟桌面，使用各种软件和应用程序进行办公和工作。本质上是利用虚拟化技术将终端数据统一存储和分配对应资源，用户以接入设备即可通过网络访问到云桌面；实现终端满足个性化需求的同时对终端数据的安全管控和终端集中管理等功能。

目前，主流的云桌面架构主要有以下几种：VDI（Virtual Desktop Infrastructure 虚拟桌面基础架构）、IDV（Intelligent Desktop Virtualization 智能桌面虚拟化）、VOI（Virtual Office Infrastructure 虚拟操作系统基础架构）。

2.1VDI 架构

VDI 是采用集中计算、集中存储的虚拟桌面基础架构。其架构是所有计算资源都基于服务器，所有运算都在服务器上，前端只需要瘦终端通过网络连接服务器上虚拟桌面在显示器上显示，每一个用户是一个独立的操作系统，在逻辑上完全隔开。

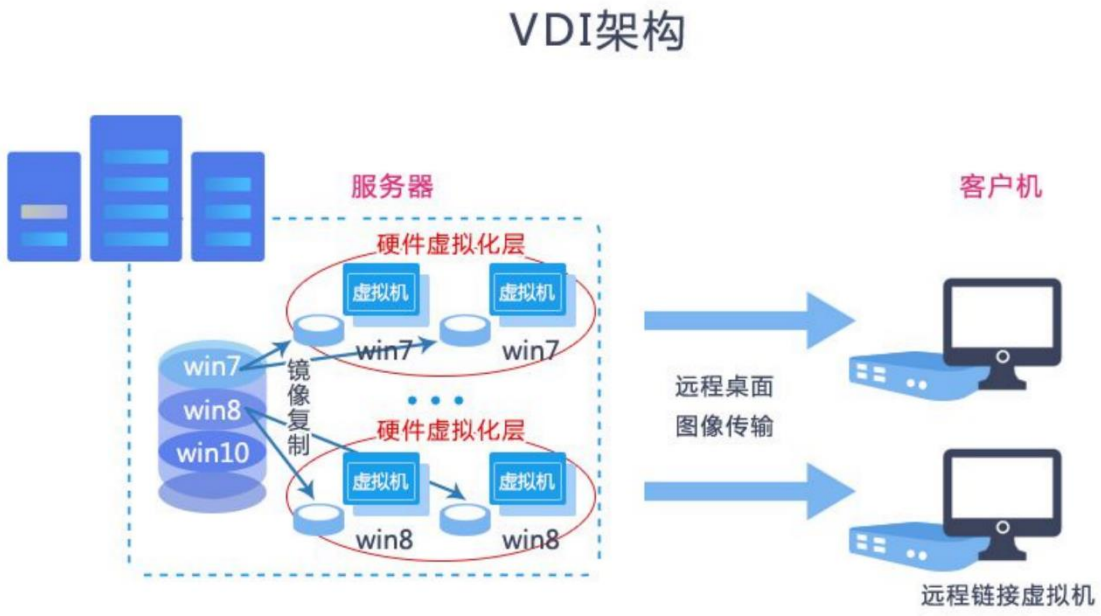


图 1 VDI 架构

VDI 让大规模的虚拟桌面得以高度集中化的管理，在该架构下可按需给每个终端分配 CPU、硬盘、内存，同一个账户使用者，即使在不同的终端登录，可以获取同样的桌面状态。另外，当终端用户系统出现故障时，只需在重启系统时，执行还原功能，即可对镜像进行快速恢复、解决问题。虽然 VDI 在资源按需分配、移动设备访问、集中管理控制、服务器架构设计、数据安全性等方面都具有很大优势。但是，VDI 要求所有桌面虚拟机集中在服务端运行，终端运行效果受到服务器性能和网络传输等诸多因素的限制，在性能、成本、兼容性等方面存在许多不足。

2.2IDV 架构

IDV 采用集中存储、分布运算的构架。IDV 的数据存储集中在后端，镜像存储在本地，配置并下发到终端机器硬盘上，终端通过虚拟机运行桌面，因此离线是可用的。IDV 不再对网络过度依赖，无需大量的图像传输，支持系统离线运行，同样可以统一管理终端桌面系统。

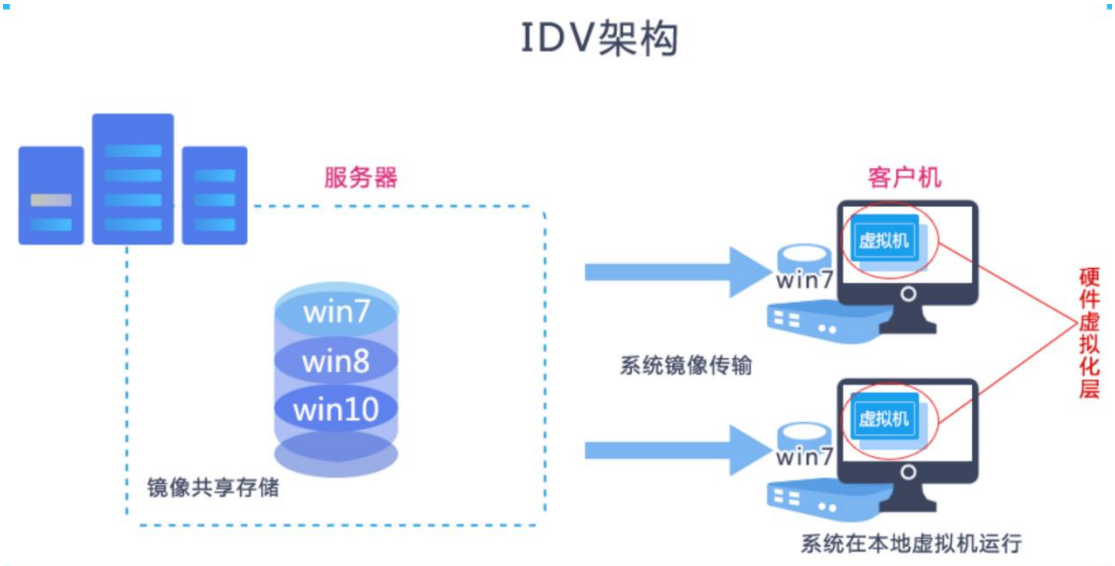


图 2 IDV 架构

IDV 架构的技术优势主要体现在采用分布式虚拟桌面技术，桌面系统理论上可以无限扩张，采取分布式计算并集中和简化管理、部署功能。得益于流行的“边缘计算”思路，虚拟桌面充分利用终端的硬件资源，每个终端都是虚拟桌面节点，桌面系统理论上可以无限扩张。其次，用户体验好。IDV 桌面体验与物理终端设备体验相当，接近本地体验效果。IDV 架构也存在一些缺陷，IDV 要求本地硬件必须统一，必须支持 VT 等带外特性，必须具有运行软件所需要的本地硬件资源；性能和兼容性还是没有办法和传统的 PC 机器相比，不支持按需分配、不支持多终端接入。

2.3VOI 架构

VOI 在传统云桌面架构基础上进行了改进，也采用“集中存储、分布运算”的构架，与 IDV 不同之处在于抛弃了硬件虚拟化层，而采用类似无盘工作站的方式启动客户端系统，让桌面完全运行在本地物理机只上，支持离线运行，支持桌面系统集中安全管理，但是性能有了很大的改善，桌面性能完全保持传统 PC 的体验，不存在兼容性问题，是新一代云桌面的代表。

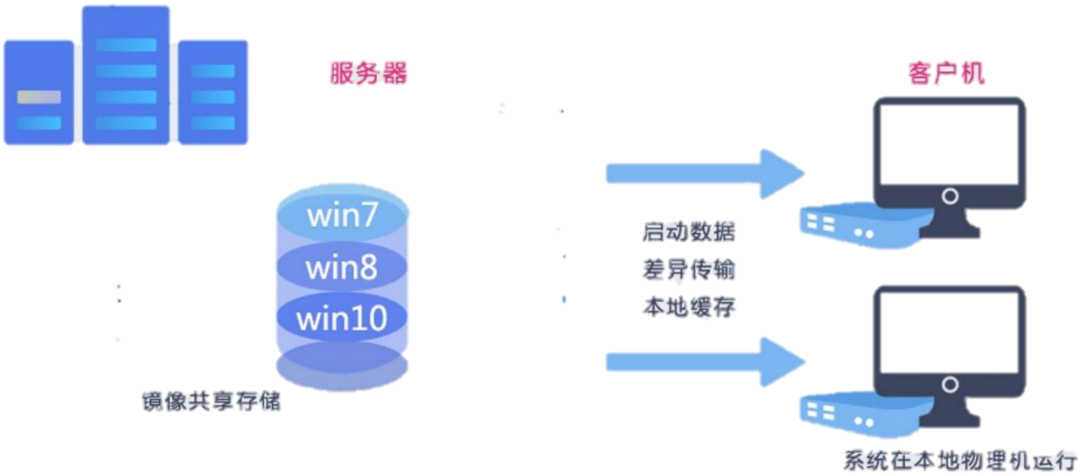


图 3 VOI 架构

VOI 架构也有明显的局限性。首先，因为它用前端计算，所以不支持移动设备。第二，VOI 架构。因为数据产生于前端，所以在广域网中，想保证完全的数据不落地是不容易的，需要单独、特殊的处理方式。第三，VOI 架构在广域网的环境下无法实现桌面漫游，必须搭配其他方式。

2.4应用技术优势比较

VDI 技术的优势非常明显，首先是移动性，其不受地域限制，无论在哪里桌面可以跟人走，还支持多种终端; 其次，数据安全可控，用户端只是桌面图像显示，而所有数据都会保存在云端。

IDV 支持离线,采用终端虚拟化技术，数据存放终端，不受网络影响，断网亦可访问; 主要成本来自于桌面终端成本，服务器端只是控制和管理，无需承载桌面，成本相对 VDI 架构大幅降低。

VOI 优势在于 100%使用本地的 CPU 和内存，外设不经过虚拟化层，与普通 PC 使用一致，不存在兼容性问题。另外，由于不经过虚拟化层，所以也不存在软件兼容的问题。

表 1 三种云桌面架构技术优势对比

云桌面技术	VDI	IDV	VOI
方案特点	集中管理、集中运行；	集中管理、本地运行；	集中管理、本地运行；
硬件资源要求	对服务器的性能要求较高，整体投资成本高；	对服务器硬件要求不高，对对终端硬件质量和稳定性依赖较大；	对服务器硬件要求不高，镜像需要依赖硬件配置；
网络资源要求	带宽占用较大,断网不可用；	对带宽和稳定性的要求相对较低,支持系统离线运行；	对带宽和稳定性的要求相对较低,支持系统离线运行；

外设支持	使用服务器资源，对外设支持受限；	终端设备支持 VT-D 技术,可以提供更好的外设支持；	充分利用 PC 计算性能，支持各类外设；
安全和易用性	数据统一存储于服务器端,数据安全性较高;支持多终端和移动端接入	数据可云端集中存储,也可本地存储,需要额外安全措施;不支持移动端	数据可云端集中存储,也可本地存储,需要额外安全措施;不支持移动端;3D 图像支持出色; 新机型不兼容 win7 等老系统

2.5云桌面融合架构模式

在医院的场景中，随着信息化的不断深化，对桌面支持的需求大量增加，传统的云桌面采用单一架构模式，难以适应医疗信息化“应用为王”的需求，用户对于计算机的需求远小于应用的需求，移动办公、远程办公在流行的同时，也推动了虚拟应用的长足发展。多架构（VDI/VOI/IDV）融合桌面云根据用户的需求，在不同业务场景采用定制化的表现方式，将服务器资源最大化利用，提供更高级别的数据安全保障，进行更细粒度的访问控制和安全策略设置，VDI 模式将计算、存储和网络资源集中管理，通过虚拟化技术为每个用户创建一个独立的桌面环境，实现资源的最大化利用。VOI 模式则将桌面环境与业务流程相结合，通过统一的管理平台为各个业务部门提供定制化的桌面服务。IDV 模式则完全将桌面环境交由用户自行管理和控制，满足他们对操作系统和软件的个性化需求。在数据安全保障方面，多架构融合桌面云采用了多种安全技术和措施，如数据加密、访问控制、防火墙等，确保医疗数据的安全传输和存储。同时，多架构融合桌面云还提供了严格的访问控制策略，确保只有授权的用户才能访问敏感数据和应用系统。此外，多架构融合桌面云还可以实现数据的备份和容灾，确保在发生意外情况时能够快速恢复数据和服务。

3. 融合架构配置场景

医院终端应用主要分为 6 个业务场景：门诊医生工作站、住院医生工作站、住院护士工作站、门诊收费窗口、医技科室工作站、行政办公内网。

3.1 门诊医生工作站

外设较为简单。诊室安装热敏打印机、扫码设备各一台，对终端稳定性能要求较高，对流动型办公、可靠性要求较高。由此，建议部署 IDV 终端设备。

3.2 住院医生工作站

住院医生工作站外设较为简单，外设终端共享，但对设备性能要求较高，运行业务系统较多，对流动型办公、可靠性要求较高。因此，建议部署 IDV 终端设备。

3.3 住院护士工作站

外设较多，存在多种打印机和扫描枪、读卡器等设备，且外设兼容性要求较高，业务单一，对流动型办公、可靠性要求较高。为此，建议部署 IDV 终端设备。

3.4 门诊收费窗口

门诊收费窗口外设最为复杂。拥有多种打印机、刷脸支付、读卡器等各类外设终端，对终端性能要求较高，对流动型办公、可靠性要求较高。所以，建议部署 VOI 或 IDV 终端设备。

3.5 医技科室工作站

外设种类复杂，存在超声、心电、核磁等众多专业性较高的设备。由于需要高性能、高分辨率功能，终端的前端计算压力比较大，外设的种类多，接口类型复杂，建议部署 VOI 终端设备。

3.6 行政办公内网

固定的工作内容，规模较小，外设简单，对网络占用不高，对业务数据获取的及时性、联系性和保密性要求较高，且非生产数据工作场景，大部分系统问题可通过后台工程师重启解决，减轻医院信息中心工作量，适合 VDI 终端。

表 2 医院终端场景部署分类

终端分类	占比	业务外设	适用终端类型

门诊医生工作站	27%	外设较为简单:配热敏打印机、读卡器	对桌面稳定性要求高：IDV
住院医生工作站	40%	激光打印（共享或网络）、读卡器	对桌面稳定性要求高：IDV
住院护士工作站	23%	激光打印机（共享或网络）、条码打印机、热敏打印机、平推打印机、读卡器等	外设复杂,稳定性要求高：IDV
门诊收费窗口（分诊）	3%	激光打印机、热敏打印机、平推打印机、读卡器、刷脸支付等	外设复杂,稳定性要求高，使用频率高：VOI/IDV
医技科室工作站	5%	图像采集卡、各种专用设备串口连接线等	要求高性能、高分辨率：VOI
行政办公内网	2%	外设简单：激光打印机	场景单一：VDI

3.7 融合架构云桌面在医疗行业的优势

融合架构云桌面在医疗行业的应用优势主要体现在以下几个方面：

3.7.1 降低运营成本

融合架构云桌面可以实现快速部署和统一管理，减少了医院在硬件设备、软件安装和维护方面的投入，降低了运维成本；同时，融合架构解决了医院电脑操作系统与新硬件平台不兼容问题，大多数医院的在用 HIS 版本，只支持 XP 及 WIN7 系统，一定程度上解决了医院投入高昂成本升级 HIS、电子病历等关键业务系统的问题；长期使用成本低，按照单台 PC 功耗 200W 算，而云终端的功耗平均在 50W 左右，中等规模三甲医院内网终端普遍超过 1000 台，从节能角度计算，长期使用成本远低于 PC 模式。

3.7.2 提高运维效率

在传统 PC 模式下，信息科在运维方面普遍感觉人力紧缺，一名工程师日常巡检极限在 50 台 PC 左右，故障响应时间长造成临床对信息部门满意度不高；在融合架构下，一名工程师可轻松负责超过 500 台终端的运维，相比于传统 PC，云桌面实现了集中管理、统一部署、统一升级，降低了 IT 运维的复杂度，管理成本较传统 PC 大幅降低，信息部门的工作效率和满意度大幅提升。

3.7.3 保证业务的连续性、稳定性

融合架构具有高扩展性和支持多节点等优势，使得云桌面不再受限于传统架构造成的性能与资源瓶颈，所有资源均可按需部署、灵活调度和动态扩展；分布式的部署方案，每个节点都有备份，即使某个节点出现故障，也不会对整个系统造成太大的影响，这就保证了业务的连续性和稳定性。融合架构采用的虚拟机技术可以在不同的硬件平台上运行，这样就可以避免因为硬件故障而导致的系统停机或数据丢失等问题。

4 结语

目前，大部分用户在桌面虚拟化架构选择上面临两难抉择，他们的理想目标是：用一套系统满足不同层次的桌面使用需要，既省钱，体验效果又好，既能在授权情况下支持移动办公，又能拥有最强的数据安全性和最佳的管理维护。这种情况下，仅采用 VDI、IDV、VOI 中的一种技术是不够的，只有 VDI、IDV、VOI、虚拟应用、PC 纳管真正实现融合，相互取长补短，才能实现这个目标。也就是说，多种技术融合的云桌面才是满足用户真实需求的云桌面，也是云桌面未来发展趋势。

在医疗行业信息化的大趋势下，云桌面技术作为一种新型的解决方案，具有广阔的应用前景。虽然还面临一些挑战，但通过不断的技术改进和优化，相信这些问题都将得到有效的解决。未来，随着 5G、AI 等新技术的应用，云桌面技术在医疗行业的应用将更加广泛，为医疗行业的发展提供更加强有力的支持。

PU-220

基于深度学习的血清类型分类研究

张志国[1] 吴红[1] 叶庆[2]
1.华中科技大学同济医学院医药卫生管理学院
2.华中科技大学同济医学院附属同济医院

目的 基于深度学习方法建立新的深度学习模型，开发一种能够快速准确识别并检测血清类别的方法，实现自动从血清样本数据中提取有价值的特征，对血清区域的识别以及对血清类型的准确分类。
方法 本研究选择 yolov7 模型用来进行血清检测结果的分类工作。yolov7 是一种目标检测算法，用来实现对图像中对象的定位和识别，yolov7 采用的方法是将输入的图像切割成一些特征图，然后通过多层由神经网络学习而成的卷积网络，将特征图特征化为检测框位置和类别概率，形成检测结果。
结果 通过观察 P-curve 图与 R-curve 图，可以发现识别血清样本是否发生溶血、黄疸、脂血的分类模型在不同置信度下都能保持较高的精确率，且在较高召回率条件下也保持了较高的精确率（大于 0.9）。此外，识别血清分类模型在 IoU 阈值为 0.5 时得到的 mAP 值均超过 0.98；在 IoU 阈值为 0.5-0.95 时得到的 mAP 值也都高于 0.90，显示出较高的性能。
结论 基于深度学习建立的血清分类模型验证了深度学习在医学领域的应用前景，优化了血清检测流程，提高了血清检测速度及准确率，具有应用于临床实践的潜力。

PU-221

人工智能在血液学诊断中的应用与展望

张立娟 方向东* 渠鸿竹 张昭军
中国科学院北京基因组研究所（国家生物信息中心）

目的 总结人工智能在血液学诊断中的研究进展，梳理当前人工智能在血液学诊断中的应用范围和面临的挑战，希望对血液诊断智能化的发展提供助力。
方法 综述国内外人工智能在血液系统应用中的相关工作。人工智能在血液学中的应用是广泛且有前景的，从细胞形态诊断智慧化、红细胞质量检测智慧化、血小板检测智慧化、对血液恶性肿瘤和地中海贫血贫辅助诊断等角度进行深入调研，并探讨当前人工智能所面临的问题及未来。
结果：人工智能技术对血液学诊断的进步体现在方方面面，降低大量繁琐、主观的人为操作，简单、客观、高效的提高了医疗效率，未来深度融合人工智能技术与血液学的发展，将推动更多的应用，极大地提高生产力。
结论 本文回顾了人工智能在血液学诊断中的应用，应用广泛且随着技术的发展准确性越来越高，基于人工智能的方法具有降低医疗成本的潜力，并且边际成本较低，高度可扩展，工作流程近乎自动化。然而不仅在血液领域，在整个医学领域，人类与人工智能（AI）的整合才刚刚开始，这里面有美好的前景、也面临一些挑战，我们要做的努力还很多，但人工智能在医学领域的蓬勃发展势不可挡。

PU-222

医院远程介入手术协作平台的设计与实现

林志刚 陈明
福建医科大学附属第一医院

为推进实施健康中国战略，提升医疗卫生现代化管理水平，通过优化资源配置，创新服务模式，提高服务效率，降低服务成本来满足人民群众日益增长的医疗卫生健康需求，国家中央人民政府陆续发布相关文件[1-4]，以促进互联网+医疗健康的发展。文件指出医疗联合体要积极运用互联网技术，加快实现医疗资源上下贯通、信息互通共享、业务高效协同，便捷开展预约诊疗、双向转诊、远程医疗等服务，推进“基层检查、上级诊断”，推动构建有序的分级诊疗格局。因此远程医疗作为一种异地、可交互为特征，综合运用各类计算机技术与医疗技术，涵盖指导、诊断、治疗等行为的医疗活动，因其突破空间限制、缓解地区医疗资源不均，带来良好社会效益和经济效益的作用受到了社会的关注[5-6]。

随着现代医学领域的科学技术不断发展，外科手术技术也日新月异。其中穿刺活检手术是诊断和治疗癌症的微创手术之一，也是医学诊断和治疗的主要方式，穿刺活检是在 CT 导引下，对全身各部位病灶经皮穿刺取得病理标本而最终取得病理诊断的介入技术；在穿刺手术中，常用的器械是穿刺针，医生凭借经验和水平，根据 CT 影像信息判定病变大小、形态、位置、坏死空洞区，以及其与临近血管、神经等的良好解剖关系，从而确定进针部位、角度和深度[7-9]。在基层医疗机构遇到需要高水平专家参与解决的疑难病例，一般只能通过会诊手术，由于距离遥远、天气阻隔、时间冲突等原因，常常导致会诊无法进行，延误患者的治疗。针对该问题，目前国内医疗机构还未在穿刺活检领域利用远程交互技术实施介入手术。因此医院如何运用相关诊断及治疗装置设备，过物联网技术远程实施介入手术诊疗是当下医联体远程协作服务发展需要研究的重要课题。

本文以某省属三级甲等医院为例，利用先进的计算机技术、全高清视频技术、网络通信技术、设计以患者为中心，以诊疗数据为基础，以音视频为纽带的手术协同体系，实现一种可行的远程介入手术平台。该平台基于可信网络，提供高效稳定的视频通讯、文件传输、穿刺路径同步、MPR 显示、数据统计、生成报告等模块化机制，不仅可以远程规划穿刺路径，还可以通过导航设备直接指示穿刺方向，手术端医生在专家的监控指导下仅需完成最基本的的进针动作，便可实现跨院区的介入穿刺手术，为国内医疗机构在该领域的远程手术实践开创先河，提供了一种可借鉴的模板。

1 相关技术

1.1 通信显示技术

1.1.1SocketIO 数据交换技术

基于 C#的 socketIO 通信交换技术，具有高性能、无堵塞的特点，同时支持文字、图像文件等内容的收发，可扩展性高。平台使用.NET 框架提供的流技术向远程主机发送流数据，或者从远程主机接收流数据，其套接字组成的房间概念与手术间概念高度契，与各客户端进行通信，使用方便。

1.1.2 WPF 图形显示技术

WPF 作为微软新一代图像系统，可为用户界面提供统一的描述和操作方法，利用 WAF 轻量级框架可以创建结构良好的 WPF 应用程序[10]，它提供全新的多媒体交互用户图形界面，特别对矢量图和三维强大的支持，使其在手术协作图像的展示可靠稳定，并有效减少提供最佳用户体验和通信逻辑所需的代码行数。

1.1.3 WebRTc 视频通讯技术

WebRTC 作为一项实时通讯技术，通过 NAT 穿越，可实现视频通讯的客户端数据点到点传输，避免了服务器中间跳转环节，大大降低了延迟[11-12]。其 NAT 穿越原理为（见图 2-1）：因为 NAT 会拒绝任何不请自来的报文段，所以 A1 客户端要与 B1 客户端建立端到端的通信，需要经过服务器的帮助，使 A1 得到 B1 在对方在 NATB 上映射的 IP 地址和端口，并且 B1 在 NATB 设备上建立了面向 A1 的映射，这样才能在 NAT 上“打通”一条畅通的通信通道。

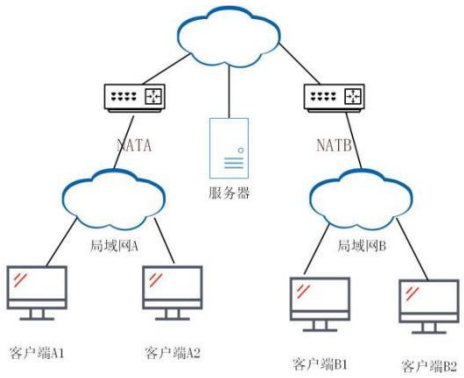


图 1-1 NAT 穿越原理

1.2 快速三维重建技术

三维重建技术可通过 CT 增强扫描仪，进行三维重建，通过精确测量穿刺点、穿刺深度、穿刺角度和穿刺水平面，在三维重建图上标注送针方向和送针距离[13]。三维中，固定平面为横断面（通过 CT 扫描获得），方便判断参照位置和图像对比；可动平面为矢状面和冠状面，医生以目标点（病灶位置）为中心，将可动面用于观看目标点附近的血管骨骼脏器等信息，从而避免术中对大的血管，重要的脏器，骨骼产生伤害。

1.3 影像实时交互技术

DICOM 是医学图像和相关信息的国际标准，用于医学数字成像和通信[14]。视频通讯收取 DSA 设备发生的 DICOM 序列文件，采用并显示 DICOM 图像及参数信息。平台可将位置信息可转为 3D 模型，并进行图像实时交互显示和标注[15]，便于手术方案(如穿刺路径等信息)的实时同步共享。

2 可靠通信网络的搭建

该省属三甲医院因业务需要，在城市建立三个分院区，医院的发展方式决定了其通信的网络的方式，下面主要针对内网专线和 5G 两种场景提出相对应的通信网络搭建方案。

1.若通过内网专线进行通信，三个院区同属一个城市，通过直连光纤进行互联（见图 2-1），因为地理位置相隔较远，不能共用一个局域网，因此通过组建各自的局域网进行各院区的通信，如图所示在城域网的节点上配置一台物理服务器，采用微服务架构，虚拟出 3 台服务器，分别提供音视频服务、DICOM 文件传输服务、数据库服务等，对图 3 中各个 NAT 设备的 Wan/Lan 口 IP 地址设置如下：

NAT:Wan 口 IP:61.54.12.1 Lan 口 IP:168.6.11.1

NATA:Wan 口 IP: 168.6.11.2 Lan 口 IP:10.10.1.1

NATB:Wan 口 IP: 168.6.11.3 Lan 口 IP: 172.168.0.1

NATC:Wan 口 IP168.6.11.4 Lan 口 IP:192.8.18.1

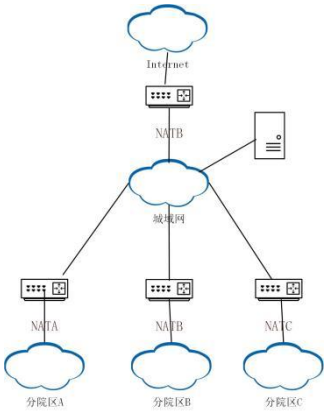


图 2-1 医院内网专线场景网络拓扑结构

2.若医院因业务需要，需要通过互联网 5G 进行通信，则相对应的组网方案（见图 2-2）如下：在总部设立服务器，分配一个公网 IP:125.65.165.187。各院区有自己的局域网并且接入 Internet。用局域网 C 的一台主机设置路由器 C,设置 Wan 口 IP：58.23.96.242，Lan 口设置为 211.80.181.1，主机 A 的 IP 地址设置为 211.80.181.2，默认网关是 211.80.181.1。

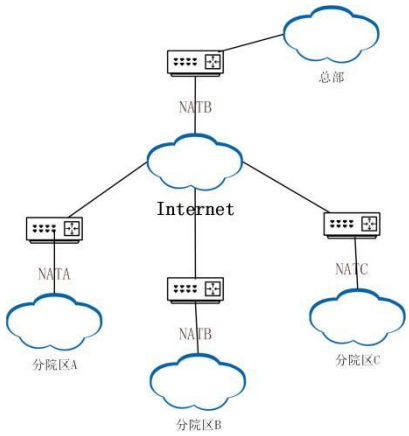


图 2-2 医院 5G 网络通信网

在两个场景中，整个网络由万兆核心交换机双机冗余，核心虚拟化互联采用 40G 以太网技术，楼宇汇聚采用万兆双链路，接入采用千兆双链路，提高了网络的容错率和可靠性。

3 需求分析与技术选型

3.1 平台需求

平台需支持高水平专家远程指导术端医师，从而跨越时空间距，避免患者与医生来回奔波，降低术端医师的手术难度，提高其手术效率；可支持多手术同时进行，高水平专家在此过程可进行规划与过程监督；可支持影像信息的 实时交互，专家与手术医生可实时讨论方案，及时处理术中突发状况；可提高穿刺精度，专家指导角度与实际角度误差不超过 1°，穿刺深度误差控制在 1mm 之内。

3.2 技术选型

3.2.1 交互界面

采用 BS 的优点是通讯与交互相关的开源组件与解决方案较多，测试和使用方便，只要有浏览器即可；其缺点是 GPU 客户端计算负载较重，性能较差，无太多高效的 3D 计算与显示组件。开发采用 CS 优点是客户端可以深度调用系统资源，降低服务器负载，如 GPU；3D 及 DICOM 组件资源丰富，性能较高；其缺点是开发周期长，成本高，测试不方便，且需要安装客户端，易用性较差。综上所述，3D 图像计算与显示为核心功能、基础功能、服务器 GPU 性价比远没有台式机高，医疗器械与软件对于开发周期低敏感，故选 CS 架构。

3.2.2 开发语言

采用 C#的优点是开发效率高，维护便捷，其缺点是性能相对较低，机器视觉相关算法库较少。采用 C++的优点是性能较高，算法相关的开源库较多，缺点是开发效率低，可维护性较差。结合两者特点，使用 c#中的 WPF 作为 UI 绘制界面，使用 c++做算法。

4 平台方案设计

4.1 流程设计

患者通过 PACS 系统预约 CT 检查后进行预登记，其登记完进行 CT 扫描，平台邀请远程医生接进入，医生进入平台便可讨论规划患者穿刺路径，方案确认后通过平台同步到显示屏，手术端医生根据引导进行穿刺，然后再进行 CT 扫描，判断手术是否到达目标位置，达到目标位置进行药物注射或者放射治疗，最后完成报告后结束手术流程（见图 4-1）。

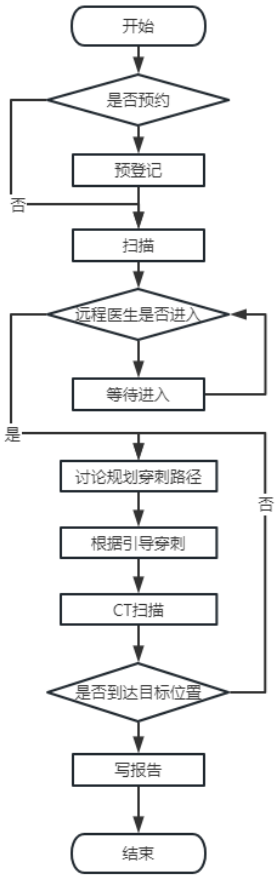


图 4-1 平台穿刺业务流程

4.2 架构设计

该平台包括基础框架层、组件层、模块基础层、及模块层。基础框架层包含日志和所有上层模块都会使用的基础功能；组件层包含系统核心组件，如 SocketIO 数据交换、WebRTc 视频通讯、DICOM 文件传输等； 模块基础层用于提供各个模块共用的组件，如显示、领域（业务逻辑）和接口；模块层包含如穿刺模块、报告模块、设置模块等；系统通过 IOC 方式调用，在组装层进行汇总组装。这 5 层独立的模块有机组成平台的架构（见图 4-2），其中顶层可以调用底层，同层模块间通过接口调用，不仅提高系统稳定性，而且使软件开发解耦，便于维护和测试。

4-2 平台架构图

4.3 功能设计

平台的协同手术室中专家可同时查看患者术中穿刺图像，同时接收手术直播中的术野等信号，通过音视频进行实时沟通，在手术规划、术中导航、手术支持等场景中有效指导术者，如调整穿刺针、评估疗效、协助并发症的处理，从而实现穿刺技术同质化，其核心功能设计如下：

4.3.1 视频通讯

该功能模块利用高速网络进行数字、图像、语音的综合传输，并且实现双向实时的语音和高清晰图像的交流，其中手术端配备手术端术野摄像头，可观看患者术中实时状态，客户端显示医生信息，方便在 DICOM 和路径图像之外提供额外需要展示或者指导的内容。

4.3.2 DICOM 文件传输

介入手术需要进行穿刺操作，穿刺过程中需要实时查看目前穿刺针的角度和深度，方便进行穿刺微调。该功能模块支持 DICOM 图像数据的获取与保存，如获取患者姓名、研究实例、CT 图像数据，并可在 DICOM 图像上实时标签，从而方便远程端和手术端医生之间进行深度的沟通与交流，避免理解偏差，奠定穿刺路径导航的基础。

4.3.3 穿刺路径同步

在穿刺手术过程中，可能会出现以下两个技术问题：1 是指导端和手术端间传输的 3D 平面间角度的异步；2 是沟通语音在时间上的延迟，导致其传达的穿刺角度和实际角度不符。而该功能提供穿刺路径规划的坐标位置信息以及术中针道与规划针道的对比，支持多条路径的设计、勾勒、保存和同步，可对穿刺路径进行设计与讨论，从而解决异步的问题，方便医生针对各穿刺路径的优劣进行实时、深度、精确的探讨确保避开患者血管，纵隔等关键部位。

4.4.4 多平面重建及同步显示

在传统的做法中，系统仅传输 CT 扫出来的平面图像，即横轴面。一张张 2D 图像的传输，难以直观的判断穿刺用针的所在位置，也不好掌控用针的精度，这就对医生的经验和水平提出了很高的要求，而该功能模块可将所有的 CT 平面图像转为 3D，通过多轴位旋转后可观看到患者任意平面角度的图像,也可直观显示穿刺的针道，实现脏器分割展示，使研究、检查、量化、分析患者体内信息变得更加直观便捷。

5 平台实现过程

5.1 软硬件配置

服务器配置为 OS:Ubuntu 、CPU:4310*2 、内存: 64G、 硬盘：4T、 Database：MySQL；穿刺主机配置为显卡 3050Ti 、CPU：i7、 内存：32G、 硬盘：2T、 OS:Win10、 Database:SQL Server；远程会诊端配置为显卡 CPU：i5 、内存：32G 、硬盘：1T 、Database:SQL Server。

5.2 使用.NET 具体实现

5.2.1 穿刺模块

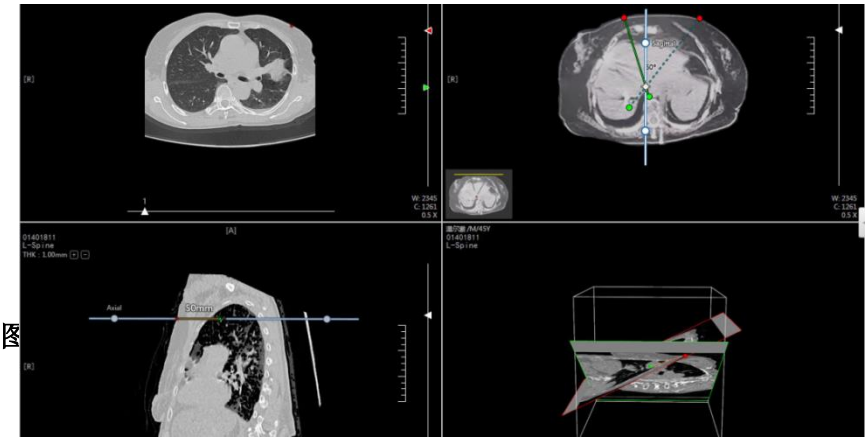
穿刺模块通过标记病灶位置和进针点形成穿刺路径，并可实现内脏的分割展示（见图 5-1），其关键方法有：

Stretch 拖拽目标点所在位置，包括更新 3D 位置信息；

Scale 更改切入点目标点，判断是否处于多平面重建模式；

Rotate 旋转切入点或者目标点平面；

Move 通过向量更改切入点目标点；



图

通过基于 HL7 协议的影像信息传输提供术中影像的获取、保存以及针道与规划针道的对比，其关键方法与基本代码如下：

Start 创建 DICOM 服务器实例，指定监听端口，启动 DICOM 服务器：

```
_server = new DicomServer<DicomCEchoProvider>(12345);
_server.Start();
```

OnCEchoRequest 设置响应的数据集，发送响应：

```
var response = new DicomCEchoResponse(request, DicomStatus.Success) ;
response.Dataset = request.Dataset;;
_server.SendResponse(response);
```

OnCStoreRequest 提取 DICOM 图像数据，将图像数据保存到数据库：

```
DicomDataset dataset = request.Dataset;
DicomPixelData pixelData = DicomPixelData.Create(dataset) ;
_context.DicomImages.Add(dicomImage);
_context.SaveChanges();
```

6 应用效果

该省属三级甲等医院开发的远程介入手术协作平台，于 2023 年 8 月开始上线，在网络的实际测试方面，院内本地 DICOM 传输，均值 3 秒/千张，院间 10 秒/千张；音视频延迟，均值 223 毫秒，最大 300ms，最小 120 毫秒；在临床应用方面，医院依托该平台成功实现两台远程 CT 引导下的跨院区腹膜后肿物活及肺肿物检穿刺术（见图 6-1），在此过程，导航机器视觉技术帮助医师自动识别、配准，增强现实技术立体地指示穿刺方向。在平台的支持下，医师约 4 分钟完成手术部署，3 秒钟完成校准，在介入手术环节的穿刺精度达到 2 毫米，完成复杂穿刺路径操作仅耗时 10 分钟，相比传统盲穿，针道调整次数平均减少 35%，手术时间缩短 50%，整个手术过程更加便捷、流畅、高效、可靠和安全。

图 6-1 主任医师跨院区协同团队开展介入手术

7 结论

该省属多院区三级甲等医院建成的远程介入手术协作平台，帮助医院建立有效的远程医疗工作机制，实现穿刺流程的优化，贯穿术前讨论、术前方案规划与定制、手术过程导航协同、术后复盘支持、术后归档等多个环节，手术超低延迟直播与示教、提升提高医生在手术的效率与质量；在平台的支持下，手术的便捷部署、穿刺能力及效率的提升让优质资源更好地通过“互联网+医疗”的方式下沉，为贯彻国家远程医疗、分级诊疗政策、推动公立医院高质量发展提供支撑作用。其作用主要表现在以下两方面：一打破地理局限性，构建全新的协作场景，促进学科之间的交流协作，提

高诊疗效率、准确率、患者治愈率及生存率，提高整体医疗服务质量；二、通过信息系统记录的手术全程画面等视频资料不仅可用于研究、教学和病历存档，也可以在有手术争议性时作为科学的判断依据。此外，现阶段该平台的展望：一、后续争取在省内实现广域网会诊手术；二、 结合 5G 实现国内远程手术，不仅仅限于介入手术。

PU-223

基于 HIS 系统的基础医嘱数据整理与思考

罗扬 刘征
重庆医科大学附属第一医院

引 言

随着我国医疗改革的不断深化、医疗技术的不断发展以及医院规模的不断扩展，医院业务部门对信息系统的需求不断增加[1]。为进一步满足医院服务质量、精细化管理需要，启动了信息优化工程，对医院信息系统进行全面升级和改造。数据作为新型生产要素，是数字化、网络化、智能化的基础[2]，已融入医疗服务各个环节。HIS 系统的切换是医院信息化建设最为重要和关键的环节[3]，其中基础医嘱数据的整理工作犹如地基建设，是系统运行的支撑和前提，具有重要的意义。

1 医嘱数据的定义

医嘱是指医师在医疗活动中下达的医学指令[4]。医嘱数据是围绕病人饮食、用药、检查、化验等方面的指令所产生的数据，包含诊疗医嘱、护理医嘱、西药和中成药医嘱、中草药医嘱、嘱托医嘱等内容。基础医嘱数据分为有费用和无费用的医嘱，其中最为核心的费用医嘱是基于当地医疗服务价格项目，由单个或者多个医疗服务价格项目封装而成，主要是为了方便临床使用。

2 基础医嘱数据整理思路

1.1 数据收集

基础医嘱数据整理是一项量大面广、耗时长的系统工程，数据的收集是第一步工作。通过现有系统与新系统数据结构和字段比对，发现存在较大差异，但原始数据仍保留着一定的使用价值。第一步，由信息中心导出现有系统原始数据进行分析，梳理涉及科室和医嘱类型，制定培训实施计划，明确培训对象、内容及目的（见表 1），分批开展培训工作；第二步，成立医嘱数据收集小组，制作时间表，明确工作进度安排；第三步，将数据收集模板表、原始数据发给科室，指定专人负责数据收集和答疑，推进数据收集、版本更新和管理工作。

表 1 基础医嘱数据培训计划及内容

培 训 计划	具体要求
培 训 对象	1、临床科室、医技科室：医生、物价员/网管员 2、职能部门：医务处、护理部、物价办、医保办、质控办、设备处、信息中心
培 训 目的	让科室了解新系统医嘱数据的填报要求，能够高效有序地收集基础医嘱数据
培 训 内容	1、根据数据收集表格模板，重点介绍新 HIS 系统医嘱数据表的字段、含义及作用 2、分析现有系统与新系统医嘱数据的区别 3、填报医嘱数据的注意事项及基本规则

1.2 数据审核与优化

1.2.1 数据分类

梳理医嘱与医疗服务价格项目的对应关系，分为 1:0、1:1、1:N、N:1 四种关系，详见表 2。按照执行科室范围可将基础医嘱数据分为全院公共数据和科室私有数据。公共数据主要由医院相关主管部门建立和维护，绝大部分公共收费类医嘱为单条医嘱，少数是全院多个科室执行的成组医嘱，如血气分析、保留导尿等医嘱；科室私有数据由科室与医院主管部门共同建立和维护。建立和维护医嘱基础数据最基本的原则是遵循医疗规范，符合用语习惯、系统规则、物价和医保政策要求，所以录入系统前有必要对医嘱基础数据进行审核和迭代优化。

表 2 医嘱与收费对应关系

医嘱与医疗服务价格项目对应关系	情况描述
1:0	一条医嘱不对应任何收费项目，即嘱托医嘱，也称无费用医嘱
1:1	一条医嘱对应一个收费执行项目，简称单条医嘱
1:N	一条医嘱对应多条收费执行项目，简称成组医嘱
N:1	多条医嘱对应一个收费执行项目

1.2.2 问题剖析

通过对收集的医嘱数据进行分析发现，主要存在以下几点问题：一是格式有误，新系统医嘱名称具备唯一性，与医嘱代码是一对一的关系，大多数科室报送的数据存在一对多、多对一的情况，不符合系统规则；二是部分成组医嘱不符合医疗规范，涉及物价收费规范、人员资质要求、医保三单相符政策等问题；三是医嘱名称不够规范，科室未按照命名规范命名，不符合用语习惯。

1.2.3 数据审核

因涉及科室多、数据相互交织、错综复杂，需把规则和存在问题梳理清楚，有针对性、分步骤地对数据的内涵和格式进行审核。

（1）组建数据审核小组，由医务处、物价办、医保办、护理部、医技科室、质控办、设备处、信息中心业务骨干组成，主要针对医嘱是否符合医疗规范、物价和医保政策、名称规范进行审核；遵循先易后难的工作原则，先审核公共数据，再审核需要多次求证的科室私有数据；审核后分批次组织科室讨论与修改，反复迭代优化。

（2）根据系统规则统筹全院医嘱数据，主要针对医嘱名称重复的情况进行规范。一是明确数据优先级，公共医嘱优先于科室私有数据；二是梳理医嘱名称重复的情况及处理方法（见图 1），利用 excel 表格对数据进行处理，保证医嘱名称的唯一性；三是跟科室现场确认后，打印纸质版返回科室再次核对签字。

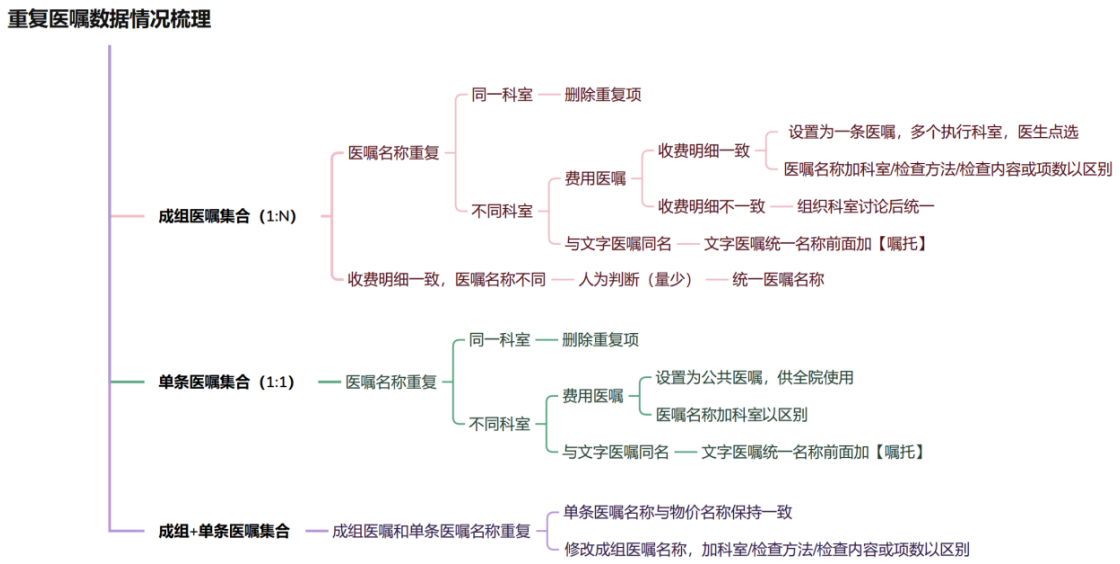


图 1 医嘱名称重复的情况及处理方法

1.3 数据导入

数据导入需谨慎，备份电子文档和科室确认签字的纸制文档，并注明整理人员、完成时间和版本状态，再由第三方工程师和信息中心人员批量录入系统，并配置医嘱执行科室。

1.4 数据核验

组织科室医生、物价员、网管员参与新系统医嘱数据测试，模拟真实环境开具医嘱，对数据质量进行核验，能有效减少系统上线时出现医嘱缺失、项目费用有误、执行科室错误等问题。

3 基础医嘱数据整理的思考

以 HIS 系统切换为契机，开展基础医嘱数据整理工作历时 4 个月，形成近两万条医嘱数据，是一项协调难度大、精细程度高的工作。通过思考总结了一些经验做法，以期为其他医院基础医嘱数据的准备和管理提供参考。

3.1 成立基础医嘱数据工作组

由医疗机构分管领导，抽调医务管理部门、价格管理部门、信息中心、临床科室和医药物资采供等部门业务骨干，组建基础医嘱数据工作小组，负责数据审核和迭代优化工作。

3.2 做好工作计划和进度管理

制作时间表，明确工作进度安排；梳理时间节点，安排专人来记录及追踪进度，可使用腾讯在线文档来共享实时数据；召开周会汇报工作进度，研究推进困难与解决办法。

3.3 数据留痕保障过程可追溯

数据整理是个迭代更新的过程，会产生很多版本的数据，需做好数据留痕保障过程可追溯。数据收集与发送由专人完成，备份电子文档，并注明整理人员、完成时间和版本状态；对于全院统筹的数据，根据处理情况做好记录，方便查询和追溯。

3.4 规范医嘱申报和审核流程

医嘱数据管理是一项长期、动态的工作，需规范申报和审核流程，明确责任部门、责任人及职责，保证医嘱的合规性。

PU-224

基于主题模型的青少年网络行为公众关注热点研究——以知乎平台“青少年网络行为”相关话题为例

刘灵涛 鲍俊 黄锦泉 刘玉文 陈兴智
蚌埠医学院卫生管理学院

0 引言

互联网的共享性与开放性给予了青少年多元化的世界。青少年通过互联网获得现实生活中无法获取的学习资源。如中国慕课，学习通，哔哩哔哩等为青少年提供了良好的学习平台。然而互联网在给予青少年多样化世界的同时，也给青少年带来了许多负面影响。《2021 年全国未成年人互联网使用情况研究报告》指出，近四成的青少年在上网过程中接触过不良信息，其中消极负面内容达 16.3%，血腥暴力内容达 15.8%[1]。随着互联网的普及，青少年容易割裂现实社会和网络社会，发生网络沉迷、网络犯罪、网络暴力等行为，青少年网络偏差行为接踵而至[2]。对于青少年而言，如果没有正确的引导，会对其思维方式、生活方式、道德观念等产生广泛的影响。

随着信息技术的发展，数据库计算能力与存储能力的提高，大规模网络文本数据的获取问题得到解决，为深入研究青少年网络行为提供了条件。基于互联网数据对青少年网络行为的研究逐渐成为研究热点。目前已有学者运用互联网大数据追踪分析青少年的网络行为。徐碧波等基于文献计量学和知识图谱研究方法，分析研究青少年的网络行为与网络心理问题，得出我国青少年互联网心理与行为研究多集中于网络成瘾、网络欺负、网络亲社会行为、网络自我表露四个领域[3]。张秀梅等基于文献计量学和知识图谱的研究方法，探究 2002—2018 年大学生网络行为的研究热点和发展趋势，得出大学生网络行为研究热点集中在高校行为失范研究、网络行为规范计策与教育创新问题、网络心理健康、自媒体背景下大学生网络行为特征及对策四个方面[4]。而李强等基于大数据挖掘的方法为大学生的上网行为进行分析，继而构建大学生校园网贷行为、网络沉迷以及风险用户的预警和干预方案[5]。

狄利克雷分配（LDA，Latent Dirichlet Allocation）主题模型是一种包含单词、文档和主题三层结构的三层贝叶斯概率分布模型。通常用作文档生成主题模型，识别语料库中潜在或隐藏的主题，被广泛应用于自然语言处理的任务中[6]。Abd-alrazaq 等人对 Twitter 上与 COVID-14 相关的推文进行了信息监测研究，以探讨与该疾病相关的话题关注[7]。赖显静等通过主题模型法实现了对慕课平台学习者评论的主题挖掘[8]。杜利等实现了对京东购物平台关于使用苹果手机用户的评论进行 LDA 主题建模以及情感分析[9]。通过对文本数据的主题建模可以获去公众对某一事件的主题关注，并为相关政府部门回应公众关注提供决策支持。在前期文献调查中发现，暂未有学者采用 LDA 主题模型法对青少年网络行为进行主题识别，因此，本研究拟从青少年网络行为入手，基于 LDA 主题模型法，挖掘青少年的网络行为的主题特征，为引导青少年开展积极的网络行为以及对网络舆情的分析与治理提供支撑和依据。

1 资料与方法

1.1 数据来源及采集

知乎是一个集知识分享与高质量问答社区为一体的综合性原创内容平台。截止 2022 年 12 月 31 日，知乎用户注册数量超 4 亿，累计 3.15 亿个问答，话题数量达 57.1 万[10]。知乎平台大量的实时、真实、多样的文本数据，为实现青少年网络行为公众关注特征的分析研究提供了重要依据。因此，本研究以知乎平台为数据源，以“青少年网络行为”、“青少年网络”、“青少年行为”为关键词进行搜索，通过八爪鱼采集器采集知乎平台有关青少年网络行为的文本数据，结果共获得 1537 条长文本数据。爬取的数据以 Excel 文件格式导出，采集时间截至于 2023 年 3 月 31 日。

1.2 研究方法

1.2.1 数据预处理

文本数据集包含大量的噪声和不相关的文本信息，如网页链接、字母、数字、特殊字符等，因此需要对文本数据进行预处理，以提高数据的质量。本文数据预处理包括数据去重、数据清洗、文本分词和去停用词四个部分。本研究借助 jieba 中文分词工具实现文本分词，使用哈工大停用词表实现去停用词。经数据预处理后共得到 1031 条有效数据。

1.2.2 LDA 主题建模

LDA 主题模型基本思想是将文档表示为潜在主题的随机混合物，其中每个主题都由单词的分布来表征[11]。本文使用 Python 中的自然语言处理库 Gensim 的 LdaModeal 模块实现青少年网络行为特征的主题建模，同时得到每个主题词以及各主题词在文档中的权重，将权重前 10 的关键词作为每个主题下的主题词。LDA 的概率图形模型如图 1 所示，概率图形的符号解释如表 1 所示。

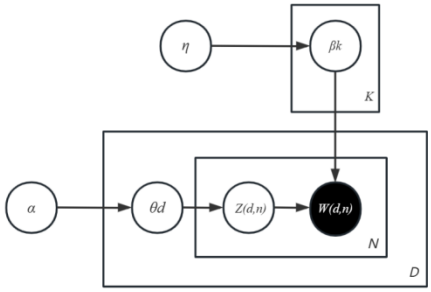


图 1 LDA 的概率图形模型

LDA 主题模型文档生成过程具体如下:

（1）对于每个主题 k={1,…,K}，从参数为 η 的 Dirichlet 分布中随机抽取一组主题词出现的概率 β_k;

（2）对于每个文档 d，从参数为 α 的 Dirichlet 分布中随机抽取一组混合主题概率 θd；

（3）对于每个文档 d 和文档 d 中包含的每个单词 W，使用混合主题概率 θd 随机抽取一个主题 $Z_{d,n}$，再根据主题 $Z_{d,n}$下的主题词概率分布 $P(W_{j,t}|\beta, Z_{j,t})$，得到 $W_{d,n}$。

整个过程，基于给定的 Dirichlet 参数 α 和 η，对应的观测变量和隐变量的联合分布计算见式(1)[12]：

$$P(W,Z,\theta,\beta|\alpha,\eta) = \prod_{j=1}^{M} P(\theta_j|\alpha) \prod_{j=1}^{K} P(\beta_j|\eta) \prod_{j=1}^{N} P(Z_{j,t}|\theta_j) P(W_{j,t}|\beta, Z_{j,t}) \quad (1)$$

表 1 LDA 概率图形的符号解释

符号	描述
α	表示主题词分布的狄利克雷先验
η	表示文档主题分布的狄利克雷先验
θ_d	表示第 d 个文档的主题分布

$Z_{(d,n)}$	表示第 d 个文档中第 n 个单词所属的主题
$W_{(d,n)}$	表示第 d 个文档中的第 n 个单词
β_k	表示第 k 个主题的单词分布
K	表示主题数量
D	文档的长度
N	文档中的单词数

1.2.3 困惑度

由于 LDA 主题建模是一个概率模型，其主题分类的有效性与主题数量的选择密切相关。如果选择的主题数量过少，每个选题下的含义就会过于宽泛；如果选择的主题数量过大，会导致数据的过度聚类，产生无用的主题或相似度过高的主题。为了解决这个问题，Blei 等人（2003）提出了一个困惑度模型来确定最佳主题的个数。他们通过计算不同主题数量的主题模型的困惑度来比较输出结果，困惑度越小，该主题数量的模型越好[13]。困惑度曲线如图 2 所示。困惑度的计算计算见式（2）：

$$perplexity(D_{test}) = exp\left\{\frac{-\sum_{d=1}^{M} log\, p(W_d)}{\sum_{d=1}^{M} N_d}\right\} \quad （2）$$

其中 M 是测试语料库的大小，Nd 是第 d 个文本的大小(即单词的数量)，p(wd)表示单词 wd 在文档中生成。

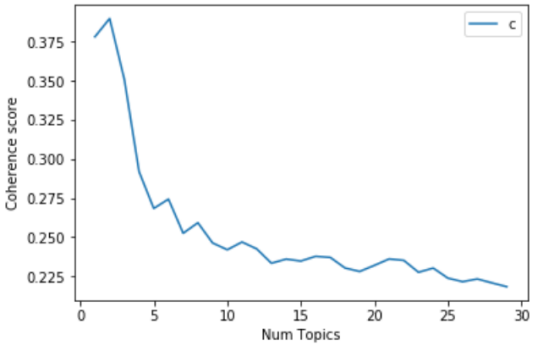


图 2 困惑度曲线图

2 研究结果

2.1 词频分析

本研究借助 Jieba 中文分词工具，对预处理后的青少年网络行为的文本数据进行词频统计，选取排名靠前的 25 位高频关键词。应用 TextBlob 中的 WordCloud 子包对高频关键词制作词云图，由图 3 可知，“心理问题”、“偏差行为”、“游戏成瘾”、“青少年犯罪”、“校园暴力”、“欺凌”、“抑郁症”、“网络暴力”等青少年网络行为的主要特征。

图 3 青少年网络行为高频关键词词云图

2.2 问答数据的主题特征

应用 LDA 主题模型对 1031 条青少年网络行为文本数据进行主题建模，由困惑度曲线图 2 可

知，第一个下降拐点对应的 k 值为 5，将主题数量设置为 5 并取得了最优的训练模型。进一步得到主题的前 10 位关键词及其所占主题词权重，最后由两名研究人员对 5 个主题进行人工标注，将 Topic1—Topic5 分别命名青少年欺凌行为、青少年沉迷网络、青少年犯罪行为、青少年心理健康问题、青少年自伤行为，结果如表 2 所示。

表 2 青少年网络行为主题以及主题词权重表

序号	主题词及权重	人工标注主题
Topic1	校园欺凌 0.005/网络欺凌 0.005/寻求刺激 0.004/打架斗殴 0.004/报复行为 0.004/霸凌 0.004/叛逆期 0.004/攻击行为 0.004/逆反心理 0.004/自暴自弃 0.003	青少年欺凌行为
Topic 2	孩子 0.007/网络 0.006/成瘾 0.004/青少年网络沉迷 0.004/网络文化 0.004/网络虚拟性 0.004/丧失自我 0.004/游戏世界 0.004/厌学 0.004/沉迷网络 0.004	青少年网络成瘾
Topic 3	青少年违法犯罪 0.004/低龄化 0.004/社会偏差 0.004/犯罪心理 0.004/极端行为 0.004/自控力差 0.004/反社会行为 0.003/心理扭曲 0.003/学校教育 0.003，/网络赌博 0.003	青少年犯罪行为
Topic 4	心理问题 0.005/青少年心理健康 0.004/负面情绪 0.004/精神疾病 0.003/心理疾病 0.003/行为障碍 0.003/病理性 0.003/抑郁 0.003/强迫症 0.003/应激行为 0.003	青少年心理健康问题
Topic 5	青少年自残 0.007/自伤行为 0.005/自杀风险 0.005/情感障碍 0.004/情绪 0.004/崩溃 0.004/精神分裂 0.004/自杀倾向 0.004/情绪调节 0.004/互联网 0.001	青少年自伤行为

根据文档—主题概率矩阵，将问答文本进行主题分类，得到青少年网络行为公众关注的文本文档占比情况，文本示例及其文档占比如表 3 所示。

表 3 青少年网络行为主题发帖数量及占比

序号	主题	知乎文本示例	发帖量	发数	发帖占比
Topic1	青少年欺凌行为	深圳一名刚上初一的女生小梦，发现朋友圈疯传指认自己是渣女的虚假图文，同学们对她议论纷纷，避而远之，小梦因此重度抑郁	08	2	20.17%
Topic2	青少年网络成瘾	13 岁少年沉迷某款网络游戏，疑似模仿游戏中翻墙跳楼动作而身亡	10	2	20.37%
Topic3	青少年犯罪行为	某上网少年迷上了网上赌博，没有钱就开始问家人要，家人不给就开始偷家里的钱，慢慢地胆子越来越大，最后抢劫杀人走上违法犯罪的道路	93	1	18.72%
Topic4	青少年心理健康问题	我的一个学生上一个暑假死了，跳河自杀的，死的时候 13 岁	15	2	20.85%
Topic5	青少年自伤行为	当刀划破我的手，再一次体验到那种疼痛时，我才感觉到自己活着	05	2	19.89%

主题 1 概括为青少年欺凌行为，发帖数量为 208，发帖占比 20.17%。其特征包括校园欺凌、

网络欺凌、霸凌等。主题 2 概括为青少年网络成瘾，发帖数量为 210，发帖占比 20.37%。其特征包括青少年网络沉迷、丧失自我、沉迷网络等。主题 3 概括为青少年犯罪行为，发帖数量 193，发帖占比 18.72%。其特征包括青少年犯罪行为、犯罪心理、极端行为和网络赌博等。主题 4 概括为青少年心理健康问题，发帖数量为 215，发帖占比 20.85%。其特征包括行为障碍、抑郁、强迫症和应激行为等。主题 5 概括为青少年自伤行为，发帖数量为 205，发帖占比 19.89%。其特征包括自伤行为、自杀风险、精神分裂和自杀倾向等。

2.3 主题可视化

应用 pyLDAvis 库对实现青少年网络行为主题进行可视化分析，pyLDAvis 库通过多维缩放模型视图呈现主题分布情况及主题词权重。LDA 主题模型可视化分析结果如图 4 所示，左侧面板显示每个主题之间的距离，可视化分析的结果可以看出 5 类主题并无交叉或重叠，均有明显的区别度。而右侧面板采用水平条形图的形式列出了每个主题下词频较高的前 30 个主题词。

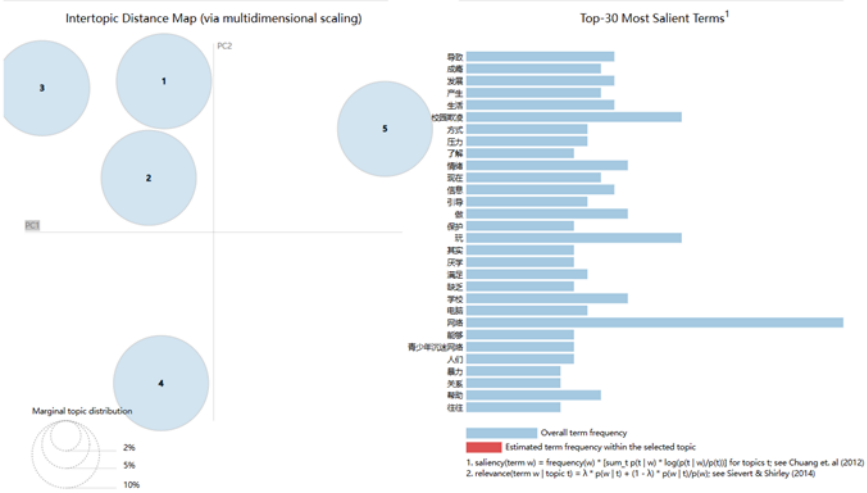


图 4 青少年网络行为 LDA 主题模型可视化结果

3 讨论

本研究从知乎平台青少年网络行为的文本数据入手，基于 LDA 主题模型对青少年网络行为进行主题建模，探索发现公众关注的热点话题。最终得到青少年网络行为集中在青少年欺凌行为、青少年沉迷网络、青少年犯罪行为、青少年心理健康问题、青少年自伤行为五个主题。

3.1 加强青少年网络行为法治教育，完善网络行为法律法规

由主题 1 青少年欺凌行为和主题 3 青少年犯罪行为可知，网络欺凌现如今已经成为影响青少年身心健康的重大问题之一。陈佳怡等在调查中发现，青少年网络欺凌行为的发生率高达 18.6%[14]。处于叛逆期的青少年心理发育尚不成熟，进行欺凌行为大多是为了追求刺激和恶意报复。并且由于网络的虚拟性，与传统的欺凌形式相比，网络欺凌更容易实现[15]。网络欺凌者可以通过匿名，隐身等方式，借助文字在网络宣泄现实社会中的各种不满与负面情绪。网络欺凌行为同样会对受害者产生负面影响，如自尊心的降低、抑郁、焦虑、自卑等[16]。在网络社交快速发展的世界里，尽管法律法规仍在完善中，但网络欺凌可以越界发展为网络犯罪，而互联网是青少年犯罪的主要根据地[17]。青少年网络犯罪主要包括以下几个方面：建立不良网站，传播非法信息；盗取他人信息，进行网络诈骗，传播网络病毒以及网络赌博等[18]。青少年网络行为和犯罪行为的发生反映出一系列的问题：青少年违法犯罪行为逐渐低龄化，学校教育、家庭教育的疏漏，青少年法律法规普及不足等。因此，要发挥家庭、学校、社会“三位一体”的作用，加强对青少年的网络欺凌教育、网络越轨行为认知教育、网络思想教育、网络道德规范教育等[19]，减少网络犯罪行为的发生。在法律层面上，应继续完善网络欺凌相关的法律法规，普及相关法律法规[20]，使网络施暴者意识到网络欺凌行为所需要承担的法律责任。

3.2 强化对青少年的人文关怀，净化网络文化环境

主题 2 为青少年网络成瘾。近些年，青少年沉迷网络的新闻屡见不鲜，最具代表性的就是青少年沉迷网络游戏 [21]。青少年普遍过早接触网络或电子产品，由于青少年的自控能力不足，很难抵挡互联网带来的各种诱惑[22]。过度的沉溺网络，会给青少年的身心健康带来严重的负面影响，如厌学、厌世、叛逆等[23]。为预防青少年沉迷网络：一是要落实家庭、学校、政府的主体责任，引导青少年科学合理地使用互联网；二是加大平台管理力度，净化网络文化环境，防治青少年过度沉迷网络。2021 年新修订的《未成年人保护法》在新增的“网络保护”专章中指出，玩游戏、看直播、刷短视频等娱乐休闲的比例较 2020 年有明显下降趋势[24]，加强平台监管力度是预防青少年沉迷网络的有效途径之一。

3.3 关注青少年心理健康问题，采取有效干预措施

主题 4 青少年心理健康问题和主题 5 青少年自伤行为可知，发帖总计 420 条，占发帖总量的 40.74%，这表明青少年心理健康已成为当今社会普遍关注的热点问题。青少年正处于身心发展的重要阶段，但存在心理健康问题的青少年人数不断上升，中国科学院发表的《2021~2022 年国民心理健康发展报告》指出，有 14.8%的青少年群体仍存在着不同程度的抑郁风险[25]。据统计，全球每七个青少年中就有一个青少年患有精神健康障碍，自杀则是世界上第四大导致青少年死亡的原因[26]。由此可见，青少年心理健康已成为一个亟待解决的重大问题，青少年心理健康工作应引起全社会的高度重视，共同营造有利于青少年身心健康发展的良好环境。及时发现青少年心理健康问题，采取有效干预措施。

4 结论

本研究借助 LDA 主题模型对青少年网络行为公众关注的特征进行文本挖掘与分析，识别出青少年欺凌行为、青少年沉迷网络、青少年犯罪行为、青少年心理健康问题、青少年自伤行为 5 类主题。在建立健全青少年积极网络行为方面，从这五个公众关注特征入手，改进青少年网络行为的工作。本研究还存在一些不足之处，由于采集文本数量较少及 LDA 模型自身的缺陷，在未来的研究中应进一步扩大数据来源并优化 LDA 模型，以期提升模型的准确性。

PU-225

ICU 患者医院感染现况调查研究

张亚锋[1] 刘新元[2] 尚小平[2] 杨林朋[2] 吕倩[2] 刘新奎[2]
1.南阳市中心医院 统计病案管理科
2.郑州大学第一附属医院 病案管理科

背景与目的 医院感染是影响患者安全最为突出的公共卫生问题之一，重症监护病房(Intensive Care Unit, ICU)患者的医院感染发病率、致死率和经济损失均明显高于其他科室，因此对 ICU 患者医院感染的预防与控制具有十分重要的现实意义。本研究以郑大一附院使用院感实时监测系统满一年的医院综合 ICU 科患者为研究对象，调查综合 ICU 患者医院感染发病情况、基础疾病患病情况、侵入性设备使用情况、抗菌药物使用情况、医院感染发生的季节分布和感染部位分布及医院感染风险因素等，为有重点地防控医院感染和构建医院感染风险预测模型提供依据。

方法 基于医院感染实时监控系统，收集 2021 年 1 月 1 日～2022 年 12 月 31 日郑大一附院使用该系统满 1 年的所有 ICU 患者的相关病历信息，应用 R studio 软件编程，对 ICU 患者人口学特征和医院感染患者进行统计描述，并应用单变量广义线性模型进行医院感染风险因素的分析。

结果 ①本研究符合纳入排除标准的共 516 例 ICU 患者，其中医院感染率为 12.02%（62/516），高于文献报道发达地区 ICU 患者医院感染率(1.62%~7.23%)。其中男性 314 例（60.9%），抗生素使用天数≤7 天的患者最多 371 例（71.9%），住院天数≥30 天的患者占比最少 236 例（13.24%），32.0%的患者有手术史。患者医院感染发病率最高的组为≥60 岁（25.56%）、男性（25.10%）、持续发烧天数≥6 天（50.70%）、住院天数≥37 天（58.90%）、抗生素使用天数≥13 天（49.02%）、床位数≥2000 张（32.82%）、需要病原菌培养（30.49%），其他还包括手术史、感染史、多器官功能衰竭、癌症史和糖尿病史，以上各项目组间比较均有统计学差异（P<0.05）。②医院感染患者更常发生在冬季（43.90%），感染部位构成比从高到低依次为呼吸系统 60.00%、泌尿系统 11.50%和血液系统 9.80%；最常见的病原菌为鲍曼不动杆菌 43.15%；最常见的多重耐药菌是耐碳青霉烯类鲍曼不动杆菌 23.35%。抗菌药物使用天数≤6 天、7~12 天和≥13 天对应的医院感染发生率依次为 6.15%、16.42%和 44.36%，三组之间差异具有统计学意义（P<0.05）。单变量广义线性模型发现可能导致医院感染的风险因素包括：患者年龄、入院临床科室、手术史、本次需要病原菌培养、糖尿病史、癌症史、住院天数、持续发热天数、抗生素使用天数、医院床位数、感染史等。

结论 郑大一附院 ICU 医院感染发病率高于发达地区，学习经济发达地区好的医院感染技术和经验，重点关注抗菌药物使用天数长、呼吸系统感染和由鲍曼不动杆菌引起的医院感染。

PU-226

神经退行性疾病：生活习惯与智慧医疗

毕成 沈百荣*
四川大学

神经退行性疾病（Neurodegenerative diseases，NDDs）是一类影响人脑神经元的慢性疾病。神经元神经变性导致神经元结构或功能的逐步丧失甚至死亡，由于神经元一般不会进行自我复制，所以这类疾病是无法治愈的且会导致神经系统功能受损[1][2]。神经退行性疾病的原因可能是多种多样的，通常没有治愈的方法，但可以通过药物和康复治疗来缓解症状和延缓病情的发展。神经退行性疾病的病理生理机制各不相同，有些主要引起运动症状，例如：帕金森症（Parkinson's disease，PD），肌萎缩性侧索硬化症（Amyotrophic lateral sclerosis，ALS），多系统萎缩症(Multiple system atrophy，MSA)，进行性核上性麻痹 (Progressive supranuclear palsy，PSP)，Tourette 综合征（Tourette syndrome，TS）；有些主要影响非运动性认知功能，例如阿尔茨海默症（Alzheimer's Disease，AD），克雅氏病（Creutzfeldt-Jakob disease，CJD），额颞叶痴呆(Frontotemporal Dementia，FTD)，路易体痴呆（Lewy body dementia，LBD）。在全球范围内，包括阿尔兹海默症和帕金森症等在内的神经退行性疾病的发病率和死亡率都在增长[3]。根据世界卫

生组织（World Health Organization，WHO）的预测，神经退行性疾病将取代癌症，成为继心血管疾病之后的造成死亡第二多的疾病[4]。阿尔兹海默症作为最常见的神经退行性疾病，它通常始于65岁以上的人群[5][6]。《阿尔茨海默氏病全球报告》（2018年）报告称，全球每3秒就会有1名痴呆症患者出现[7]。而在我国，65岁以上老年人帕金森病的发病率为1.7%，患者总数已经超过250万人，增长速度极为迅速，每年新增大约10万左右的患者。

神经退行性的生活习惯有很多，虽然这些生活习惯可能会导致疾病的发生，但是并不一定就是发生疾病的原因[8]，个性化的因素以及不同生活习惯的互相作用都是疾病风险增加的原因。对于神经退行性疾病，通常与生活方式有关[7]，如饮食、教育、爱好、体力活动、身体素质、抽烟、心理因素等。康复治疗可以帮助改善神经退行性疾病患者的日常生活和运动能力，提高生活质量，而保持健康的生活方式更是预防神经退行性疾病的重要措施，第一，是保持一个健康的饮食。以植物为主的饮食有助于降低阿尔茨海默病和帕金森病的风险，多吃鱼、坚果和富含叶酸的食物也可以促进脑部健康。第二，是进行适度的运动。适度的锻炼不仅有助于维持心血管健康，还可以提高个体的认知能力和记忆力，有氧运动、力量训练、瑜伽等都是适宜的运动方式。第三，是保持社交联系和认知刺激。经常参加社交活动和接受认知刺激可以预防神经退行性疾病的发生。可以多参与社区活动、与朋友聚会、阅读书籍等。最后，要保持心情愉悦和积极乐观的态度。

当前对于NDDs生活习惯的研究中，许多都存在异质性，例如：关于喝酒对阿尔兹海默症的影响，目前的研究结果并不一致。Sandra B Sipetic等人[9]在贝尔格莱德进行了一项病例对照研究发现PD与咖啡和酒精消费呈正相关，对110例新的PD病例和220例医院对照进行了访谈。病例和对照按性别、年龄和居住地（城市/农村）匹配。为了分析数据，使用了条件单变量和多变量逻辑回归方法。随着饮酒量的增加，PD的风险显着增加。相比于其他关于饮酒影响NDDs风险的报道，酒精的摄入量的不同，会对NDDs的影响也不同，适量饮酒可以降低患上NDDs的风险，但超过适量的饮酒会使患病风险增加，同时，过度饮酒也可能对大脑和整个身体健康产生负面影响，应该避免。Susanna C Larsson等人[10]使用与可修改风险因素相关的遗传变异作为工具变量的孟德尔随机化研究，实验共纳入了17008名阿尔茨海默病患者和37154名对照者。实验结果咖啡摄入量较高（每天每杯：1.26，95% CI：1.05-1.51；P=0.01）与阿尔茨海默氏症的几率较高相关，这可能和实验人群的差异导致的。较高的受教育水平通常作为许NDDs的保护因素，而在R Frigerio等人[11]的研究中，他们使用病例对照研究设计调查帕金森病与教育和职业的关联。作者使用两个独立的数据来源收集了有关教育和职业的信息：对系统中完整医疗记录的审查和电话采访。实验结果显示受过9年或以上教育的受试者患PD的风险增加（OR = 2.0；95% CI = 1.1-3.6；p = 0.02），并且随着受教育程度的增加，风险呈增加趋势，受过高等教育的受试者和医生患帕金森病的风险增加，这种结果的产生可能与样本的职业暴露密切相关。

随着老龄化在我国总人口中占比超过10%，因老龄化而导致的劳动力成本升高、养老保险缺口急剧增大、健康医疗负担急剧增大、医疗设备及设施急剧匮乏等问题在我国愈发明显和突出，主动健康和全民健身是有效缓解老龄化带来的经济社会问题的有效方法。2015年，随着世界卫生组织《关于全球的老龄化与健康报告》的颁布，“主动健康”的科学理念首次以全新的面貌出现在公众印象中。以主动健康为策略的疾病负担减轻、疾病早期治疗具有重要的科学意义和社会经济效益，探索人工智能在医学主动健康中的应用对于老年人的健康具有重要影响。神经退行性疾病作为一类慢性且不可逆的疾病，早期的发现和治疗是增加幸存者数量的关键，而生活习惯的改变在主动健康中占据着重要的位置，对生活习惯数据进行建模和应用在NDDs早期筛查和疾病预防方面具有重要意义。当前，ChatGPT为代表的聊天机器人显示了在知识查询、知识问答等任务中的巨大潜力。但是医学尤其是临床医学，是一个极其复杂的应用领域，现有的大语言模型因为其准确性、缺乏可解释性等备受争议，为医学知识的错误使用带来巨大的安全隐患。将泛化的或非专业化的知识应用于医学聊天机器人领域，将不可避免的出现知识盲区等问题。而以专业知识库或知识图谱为知识来源的医学聊天机器人将极大的改善因知识储备不足或专业化知识匮乏而带来知识错误应用问题，融合知识图谱的聊天机器人能够提供个性化的医疗信息，促进主动健康管理，降低医疗成本，提高医疗效率，并推动医疗科研和创新的发展。因此针对人群体质及疾病类型多样化与生活习惯的异质化，对神经退行性疾病的各种生活习惯构建基于专业知识库或知识图谱的聊天机器人代表了新一代人工智能在神经退行性疾病健康服务中的解决方案，通过结合神经退行性疾病可改变的生活习惯知识图谱和聊天机器人技术，可以提供个性化的疾病预防和管理建议，通过智能化的医疗咨询和指导，帮助用户早期发现潜在风险，及时采取相应的预防和治疗措施，人们可更好地管理自己的健

康，减轻医疗负担，提高健康状况和生活质量，也会对神经退行性疾病的早期诊断、风险评估方面提供帮助。

PU-227

“健康中国”背景下健康医疗大数据资源的应用与治理

侯雅琪 1,2，龚万里 1,2，于 琦 1,2*
1.山西医科大学管理学院
2. 临床决策研究大数据山西省重点实验室

1 引言

"健康中国"是中国政府提出的一个重要发展战略，旨在提高全体中国人民的生活质量和医疗健康水平。健康医疗大数据在这一战略中起着关键作用，可以帮助医疗机构提高服务质量，提升医疗效率，也可以帮助政府更好地了解公众的健康状况，制定更有效的政策。

2 国内外研究现状

2.1 国外健康医疗大数据的发展现状

健康医疗大数据应用与治理的研究在国内外得到了广泛关注和深入探讨，健康医疗大数据的应用领域非常广泛，包括临床医疗、公共卫生、医疗管理、药物研发等。目前在数据应用方面主要以利用大数据分析技术改进临床决策、提高疾病预测和预防、优化医疗资源配置。比如，Oracle 研发出医疗供应商供应分析软件与企业医疗分析软件，Kronos 研发出医护劳动力管理信息系统，IBM 研发出医疗防欺诈与滥用管理系统[1]。

英国医疗行业有机结合大数据与医药卫生科研机构，以“李嘉诚卫生信息与发现中心”为例，主要有两个机构：一个是靶标研究所，另一个是大数据研究所，目的是借助大数据技术获取、存储以及分析多样化医疗信息，明确新药物研发方向并降低研发成本，另外促进新型医疗手段的研发[1]。

国外研究健康医疗大数据更加注重整合多源数据，包括社交媒体数据、传感器数据等。许多国家都在建立医疗数据共享平台，以整合来自不同医疗机构和保健服务提供者的数据。

这有助于提高医疗数据的可靠性和准确性，同时推动医疗研究和实践的发展。

健康医疗大数据的发展现状不断演进，许多国家正在积极推动大数据在医疗领域的应用，以促进医疗服务的改进和创新，提高健康管理水平，进一步改善公众的健康状况。

2.2 国内健康医疗大数据的发展现状

国内健康医疗大数据发展现状十分活跃，随着信息技术的不断发展和医疗领域的数字化转型，利用大数据、云计算、人工智能等技术为提高医疗服务质量保驾护航已经成为普遍趋势。2022 年印发的《促进大数据发展行动纲要》和 2016 年印发《关于促进和规范健康医疗大数据应用发展的指导意见》 进一步确定了健康医疗大数据作为国家重要基础性战略资源的地位[2]。

我国正推动建立全国性的医疗大数据平台以及全国范围内的电子健康档案和电子病历的建设，将健康档记录和医疗信息数字化，整合各级医疗机构的数据资源，实现数据的利用与共享。国内的健康医疗大数据发展正处于一个快速发展的阶段，主要集中在医院和医疗机构的电子病历数据、实验室数据等，近些年，科研机构和企业共同构建健康医疗大数据实验室以及各类共享平台，积极调动可调动的力量共同参与健康医疗大数据的应用、治理，为医疗领域的现代化提供了有力支持。我国的健康医疗大数据发展规划主要是建设 1 个国家数据中心、7 个区域中心,各地建设 X 个应用发展中心,从而形成“1+7+X”的发展格局[3]。在信息化发展过程中,国家拟定的 7 个区域中心缩减为 5 个[3]。如今根据“1+5+X”的发展规划,我国目前已建成东部江苏省、西部贵州省、南部福建省、北部山东省以及中部安徽省五大健康医疗数据中心,东西南北中铺开了大数据应用发展的网络[4]。

总体而言，国内外研究都在探索如何充分利用健康医疗大数据的潜力，提高医疗质量和效率。然而，仍然存在一些挑战，如数据隐私保护、数据质量控制、数据共享和合作等，需要进一步研究和解决。

3 健康医疗大数据的应用

在数字化时代，健康医疗大数据的产生与积累日益增多，大数据技术与医疗健康领域的深度结合为改善医疗服务质量、推动精准医疗和优化资源配置提供了巨大机遇。

3.1 健康医疗数据支持临床决策

医疗数据在临床决策中扮演着重要的角色，可以辅助诊断和治疗、预测疾病风险、监测疾病进展和效果评估，以及支持临床研究和实证医学的发展。通过分析和应用医疗数据，医生可以更准确地进行疾病诊断、制定治疗计划和预测疾病风险。这有助于提高医疗的准确性和效率，为患者提供更好的医疗质量。

3.2 健康医疗数据推动精准医疗发展

医疗数据的应用可以实现个性化医疗。通过充分利用医疗数据包括病历、影像、实验室结果等，根据患者的病情和个体特征如基因组信息、病史、生活习惯等数据，医生能够更加准确地判断疾病、制定治疗计划，并为患者提供个性化的医疗服务。通过分析大规模的医疗数据，可以挖掘出不同人群之间的健康差异，为个体化的医疗决策提供依据。

3.3 健康医疗数据促进药物研发和临床试验

利用大数据分析技术，可以加速药物研发过程和临床试验。通过分析大量的临床数据和基因组数据，可以发现新的药物靶点、预测药物的副作用和疗效，从而提高药物研发的效率和成功率。

3.4 健康医疗数据预警疾病传播

通过分析大量的医疗数据和社交媒体数据，可以实时监测疾病的流行趋势，高风险人群和相关因素，提前预警和采取措施。通过分析社交媒体上的关键词和地理位置信息，可以追踪疾病的传播路径和速度，及时采取防控措施。这对公共卫生决策和政策制定非常重要，有助于采取针对性的预防措施、健康教育和干预措施。

3.5 健康医疗数据建立队列研究品牌体系

队列研究是医疗领域的关键基础性研究，是国际公认的慢性病研究首选设计。 队列研究通过纳入足够大的人群，系统收集流行病学数据及生物样本，进行长期（为期数年至数十年）的跟踪随访，阐明慢病发病机制，探索高效关键防治技术，提供重要医疗决策支撑，是具有战略意义的国家资源。队列研究旨在长期跟踪一组人群并收集相关数据来探索某种暴露因素与特定结果之间的关系。医疗数据在队列研究中扮演着重要的角色。医疗数据可以提供大量的信息和观察数据，为队列研究提供了基础。

医疗健康大数据的应用可以帮助提高医疗效率、个性化治疗、加速药物研发、提前预警和控制疾病的传播、以及改善健康管理和预防。大规模的医疗数据同样也可以为科学研究提供宝贵的资源和参考依据。医疗数据的应用可以推动学术进展，促进新的发现和医学知识的累积，从而推动医学的发展和进步。

4 数据治理

大数据的“4V 特征”，即海量（Volume）、高速（Velocity）、多样（Variety）、真实（Veracity），揭示其能在健康医疗领域发挥重要作用[2]。健康医疗数据主要来源有：临床医疗数据如电子病历、医学影像数据、生化实验报告、病人自我评估报告；生物组学数据如生物医学知识库、基因工程数据、各类组学数据；物联网数据如可穿戴设备数据、环境传感器数据、医疗设备传感器数据等；互联网数据如社交媒体平台数据、移动手机 APP 数据、虚拟健康社区数据。

海量的健康医疗数据的意味着对数据的治理及应用有着更高的标准、更严格的要求。数据量增长迅速、数据多源异构、非结构化的自由文本、先有数据后建立管理模式、大数据管理与传统的数据管理区别较大，正是这些特点使得在大数据时代医疗数据的应用面临新的挑战。 如何确保数据的质量、保护数据信息的安全以及更好地进行数据共享成为了待解决的问题。

4.1 数据质量

医疗大数据的质量直接影响其应用效果，为确保医疗数据的准确性和完整性，应建立标准化的数据采集和整理流程，实际上，并非所有医院都能够正确认识到数据资产的价值，这就导致了忽视数据质量、对数据资源利用不充分等问题。实际上，部分医院存在对数据资产价值认识不足、利用不充分、忽视数据质量等问题，导致系统中存在大量“脏数据”[5]。提高医务人员的意识并且建立标准化的数据采集和整理流程是首要之举。提高数据质量控制和清洗能力，用人工智能和机器学习等技术来处理大规模数据，避免错误和噪声的干扰也能大大提升数据的质量。

4.2 数据安全

在数字经济时代，健康医疗大数据不仅是实现全民健康的重要技术手段，其本身也是具有重要国家安全价值的重要信息资源，具有多重战略属性。（“健康中国”背景下的健康医疗大数据发展 ）要想加强数据的存储和传输安全，采用先进的加密技术和安全协议，保护数据不被非法访问、篡改或泄露具有重要意义。同时建立健全的安全管理体系，包括定期的安全审核和漏洞扫描，及时更新补丁和安全策略；制定严格的隐私政策和法规，确保医疗数据的收集、使用和传输过程中的合法性

和透明度；加强监管和执法力度，对违反数据保护规定的行为进行处罚，提高违法成本，维护广大公众和个人的数据安全和隐私权益；加强公众对于医疗大数据的理解和认知，提高对于数据隐私和安全的重视程度。

4.3 数据共享

各大医疗机构和健康管理平台之间进行数据共享与整合，利于医疗资源的优化配置和互联网医疗服务的提供。建立全面的健康医疗大数据基础设施：建立统一的健康医疗大数据平台，整合来自各个医疗机构、诊所、实验室、保险公司和政府卫生部门的数据，促进跨机构和跨地域的数据共享，加强合作和协作机制，制定数据共享的框架和规范，明确数据的使用目的和访问权限，保护数据提供方的权益，确保数据的互操作性和安全性。同时，建立健全的数据交换平台和标准化的数据格式，方便数据的集成和共享。采用统一的数据标准和格式，以确保数据的一致性和可比性。这有助于不同数据源之间的交互和分析。

综合来说，更好地利用健康医疗大数据需要跨部门合作、技术创新和制度改革。这将有助于提高医疗体系的效率，促进疾病预防和治疗，最终提高国民健康水平。

5 小结

健康医疗大数据的应用为提升医疗服务质量和资源配置效率提供了强有力的支持。然而，同时也面临着隐私保护、数据安全、数据质量和数据共享等治理问题。只有通过制定相关政策和法规，加强技术和管理手段的应用，推进数据共享和开放，才能实现健康医疗大数据的可持续发展。

PU-228

基于医院综合信息集成平台的 DRGs 管理系统的分析

郭佳
黄河三门峡医院

一、背景

疾病诊断相关分组(Diagnosis Related Groups ,DRGs)20 世纪 60 年代起源于美国，80 年代后开始在欧洲和澳洲及亚洲等地区开始传播，目前，DRGs 以及相似病种管理已成为世界各国卫生服务发展研究的重要课题。DRGs 的本质是数据统计分析工具，它是根据就诊人员的年龄、性别、住院天数、临床诊断、诊疗过程、手术、疾病严重程度、住院费用、合并症与并发症及转归等因素，将相似病例分到同一组，然后决定应该给医院多少补偿[1]。DRGs 肩负着重要使命：创建方便医疗服务质量或医疗费用管理的病例分组体系，或者一种医疗产出的测量模型。它通过将特征相同或相似的病人归为一组，并通过科学的统计分析方法分析特征，构建医生之间、医院间相互同质且可比的基础单元，解决了精确科学衡量医院医疗服务补偿及质量的标准问题，为医疗费用补偿或医疗服务评价提供了衡量标准。

DRGs 数据采集的内容来自病案首页，然而病案首页信息繁多，许多数据需要人为录入或者二次计算，极大的增加了临床和病案人员的工作量[2]。但目前医院信息化发展的主要以临床科室的需求为导向，科室提出需求，信息部门配合上线相新的系统，新系统通过接口与院内在用的其它信息系统联通，通过这种方法，虽然较为简单方便，可以较为满足单个科室的业务需要，但是多个科室提出需求，就会多次重复上述步骤，院内信息系统间反复进行相互对接，这种做法不仅从成本上是对医院的浪费，而且由于各厂商系统的数据标准不同，有些信息还不能实现共享，形成了一个个信息孤岛，不能实现信息的互联互通，不利于病案首页数据的提取；

基于信息集成平台的 DRGs 管理能够结合信息集成平台与 DRGs 数据采集中的各个子系统，各个系统不仅要进行对接，并且要完善自身的标准化建设，实现平台数据和 DRGs 数据的互联互通，是医院信息化建设发展的目标。

二、医院信息集成平台的实现

医院完整的信息系统通常有近百个子系统组成，涉及许多专业领域。如此专业且复杂的系统需要非常精细化的开发分工，整合不同厂商有特色的专业系统是医院信息系统的发展趋势。为了保证医院信息化能够取得成功，必须保证各个系统的有效集成和数据互联互通[3]。但是这些系统通常是由于科室需求独立且分散建设的，这些信息系统来源于不同厂家，有着各自的技术，没有统一标准的信息交换方式，这些系统的集成整合已经成为医院信息化发展必须解决且难以解决的主要问题[4]。

集成信息平台为临床数据中心的数据来源提供了技术基础和保障，通过信息标准、交换原则的

制定，对业务系统提供标准的信息交换服务，确保数据交换过程的安全性、可靠性，实现数据在系统平台范围内自由、可靠、可信的交换。通过医院信息平台建设，一方面可以规避“点对点”式的信息共享与交换，并使得医院可以基于信息平台整体上进行业务流程优化与管理，另一方面利于医院信息系统建设的持续性发展，以适应未来的需求变化，避免信息化建设的大范围的推倒重来,系统集成平台的构建主要面向两个核心问题：这就是医疗数据交换标准和互操作性[5]。

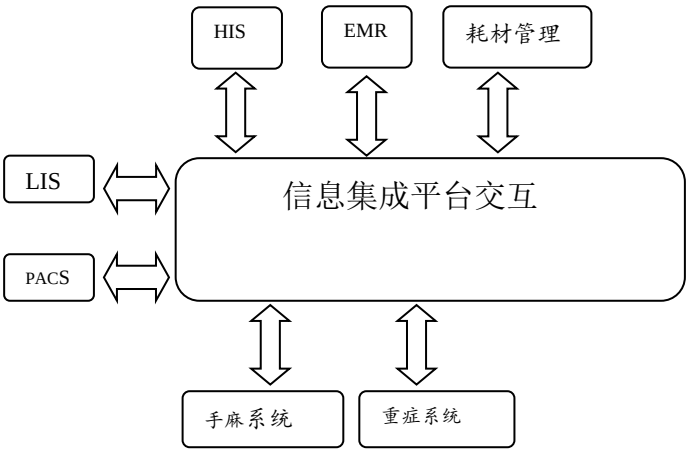
医疗数据交换标准根据平台制订标准的相应的表结构，统一的字典库，完善的知识库、统一的编码方式等。分散在各个系统的数据集中到一个数据中心的数据库里，必须保证递过程中的数据完整性、及时性、正确性、关联性等，制定合理的医疗数据交换标准是信息平台的基础，是各个接入的厂家必须遵守的设计标准。

医疗信息互操作行主要有以下三个方面：

1、基本互操作性。基本互操作指信息系统 A 的数据可以被信息系统 B 所接收，但不能被有效识别并利用。这是最基本的互操作性，可以依靠目前的互联网或大多数局域网采用的通用传输标准技术实现，如 HTTP、TCP/IP 等。语法互操作性是指通过定义数据交换的结构或格式，确定数据信息互相交换的语法规则，这样可以是信息系统间进行信息交换时可以有效识别相关数据，为了实现这一目标。需要以 HL7 标准系列、DICOM、NCPDP 等作为建设标准[6]。

2、语义互操作性既能使数据交互的结构互操作性，又能使有对交换数据统一编码化，这样接收信息的系统就可以分析理解和利用数据的所代表的含义，这是最高一级的互操作性。各地区可以依据各地的实际情况制定语义标准，术语标准，主要有 ICD、LOINC CPT 等。

3、功能互操作性是指信息系统之间在进行交换数据的同时进行相互作业的流程协调行。目前主要采用 IHE 标准框架，包括 XDS 系列、ATNA、PIX、等[6]。



图（1） 信息集成平台建设效果

通过建立集成信息平台，集成各类应用系统以及日常运营的业务，通过该平台整合医院内部业务应用系统，形成一个互联互通的医院业务协作网络。信息集成平台效果图如图（1）

三、基于信息集成平台的 DRGs 管理系统

（1）DRGs 数据管理遇到的问题

在 DRGs 应用过程中，为了达到综合准确分析数据要求，需要收集患者年龄、诊断、治疗方案、治疗效果等方面的信息，但是，由于医院信息系统、国家相关政策、人员等原因，在实际操作过程中有一些难点和问题。

1、信息收集困难：目前各个子信息系统建设时间、技术标准及建设方案均不相同，没有统一的、通用的数据标准，为病案科和临床的病案数据提取与分析带来诸多困难如：信息提取不准确，人工工作量繁重，大量的人为二次计算等；

2、编码错误：DRGs 以医疗服务的诊断编码为基础来进行分组如果医疗人员对编码规则不熟悉或出现错误，可能导致分组结果不准确，影响医疗服务的正确计费和报销。

3、分组不准确：DRGs 分组是依据医疗服务的诊断、手术、病情等诸多因素进行分类，有时病例之间存在相似性，但可能被分到不同的 DRG 组中，导致资源的不合理分配，还有，部分罕见

病例无法被准确分组等。

（2）院内综合信息平台的建设标准

信息集成平台建设要达到基于统一的数据标准实现，集成厂商必须遵守相关行业标准进行行业建设标准，国外的标准有 ICD9，ICD10，SNOMED（系统医学术语命名法），DICOM（医学数字化影像通讯标准），HL7（医院电子信息交换标准，UMLS(一体化医学语言系统)，LOINC（观测指标标识符逻辑命名与编码系统），DRGS（疾病相关诊断分组）等。

国内标准如《电子病历系统应用水平分级评价标准（试行）》《电子病历基本架构与数据标准》卫生部下发的 112 种临床路径，《医院信息互联互通标准化成熟度测评方案》《医院信息基本数据及研究报告》，《医院信息系统基本功能规范》，《中国医院信息系统标准代码研究报告》等。

（3）基于信息平台的 DRGs 管理系统建设

DRGs 的数据采集应用和实施涉及到多个科室与人员: 医务部门（临床医师）需要准确把握主要诊断及主要手术、操作以及其他诊断及手术等；病案科（编码员）需要对主要诊断及主要手术、操作的进行准确判定，并且对于医师书写的诊断及手术、操作的要有正确理解，且编码员需要有丰富的编码知识、经验，并且编码员需要了解相关的临床知识，通读患者病历等、医保科需要根据医保政策需求制定 DRGs 结算与实施规则，对特殊病例（例如未入组病例、费用极高病例、费用极低病例）进行界定并给出相应的结算规则，及时向上级部门沟通反应情况、分析原因、提出建议、合理协商，更切实提高医院医保管理水平与实效，减少可能发生的政策性亏损；以及信息化平台的自动提取等，需要各个科室和相关工作人员的协调配合。

医院各部门的相关工作人员应熟悉 DRGS 数据采集内容和运作机制，有利于各相关部门各司其责，通力协作更好地推进医院 DRGS 应用取得更好的成效[7]；此外，院内的各个信息系统，如 HIS 系统、临床路径系统、手术麻醉信息系统、重症信息系统、耗材物资信息、检验检查信息等相关应用系统都应满足综合信息平台的接入要求和数据互联互通的要求，及时进行系统标准化改造，尽早达到“书同文、车同轨”建设效果，减少医生及病案人员的人为录入量，提升系统自动提取效率；

（4）基于信息平台的 DRGs 管理系统的使用效果

基于信息平台的 DRGs 管理系统建设完毕后，可以实现患者信息管理、DRGs 分类与评价、医疗资源管理、医疗费用管理、医疗质量评估、绩效报表等功能[8]：可对医院各科室的工作量、病种构成、治疗费用等进行分析，为医院管理人员提供了客观的数据支持，便于制定科学的管理决策；通过 DRGs 分类方式对医疗服务进行评价，可以更加客观地衡量医疗服务的质量水平，进而发现问题并采取相应措施进行改进以提高医疗服务质量；能改善医院 DRGs 付费亏损情况，为医保、医务、病案、财务等各职能部门的 DRGs 管理提供了良好的数据支撑；且具有分组器动态优化、分组匹配率高、支持多个分组器、同时满足医院各部门要求；信息系统互联互通与智能分析，临床医师和病案人员只需选择或者少量输入即可完成病案首页信息录入工作，最大限度的减少了重复录入与计算的时间和精力，提升工作效率。

四、结语

随着我国医疗制度的深化改革及医疗保险制度的实施，在现阶段的医疗行业中，采用 DRGs 付费方式对公立医院的成本进行管理是医院在新时期和新形势下保持优势的关键，实施病种医疗质量的全面管理和病种医疗服务的分析及评价，在病种及其费用控制的基础上对各个不同的医疗机构进行考核也是必行之道[9]。如何保障 DRGs 的顺利实施，不仅需要临床科室、病案科、医保科等人员熟悉 DRGS 数据采集内容和运作机制，做到全程把控，更要有坚实和完善的信息系统支持，基于信息平台的 DRGs 管理系统的使用则能把两者有效的有机结合，发挥最大的效果，提升医院工作人员的整体水平，进而提高医院的管理效能。

PU-229

医疗保险对农村中老年人健康状况和生活满意度的影响--基于 CHARLS 队列研究

路亚柯[1,2]，赵艳婷[1,2]，岳馨培[1]，杨林朋[1,2]，尚小平[1]，赵敏[1]，刘新奎[1,2]
1 郑州大学第一附属医院
2 河南省医院管理研究院

老龄化是经济社会发展的重要趋势，根据第七次人口普查数据可知，2021 年中国 65 岁及以上的人口占总人口的比例已经上升至 14.2%，这标志着我国已经正式迈入“老龄化社会”[1]。老龄化是国家治理现代化不可回避的现实问题，《“健康中国 2023”规划纲要》指出要把健康摆在优先发展的战略地位[2]。与此同时，为积极践行“十四五”规划战略部署，加快完善老年健康支撑体系，“成功老龄化”理应成为我国老龄治理的策略[3]。成功老龄化是指个体在老化过程中保持身心健康，维系自身与外界之间的良性互动关系[4]。

农村中老年人是一个特殊的群体，由于农村人口流动及外出务工子女增多，农村中老年人的生活满意度和生活品质受到了相应的冲击[5, 6]。此外，中老年人的身体机能产生生理性衰弱，以及做我保障功能的逐步弱化，更容易受到外部干扰因素的影响，导致发生各类中老年慢性疾病的风险增加[7]。因此，对于医疗保险的需求也相应的增加，医疗资源的合理配置对中老年人的身体健康状况与生活满意度有着重要影响[8, 9]。

在老龄化日益严峻的中国，关注中老年人的身体健康状况，提升中老年人的生活满意度，应是我国未来国计民生的重点关注的问题。本文以中国 2018 年健康与养老跟踪调查数据为基础，探究医疗保险对中国中老年人健康状况和生活满意度的影响。

1 研究方法

1.1 调查对象

本研究数据来源于中国健康与养老追踪调查（China Health and Retirement Longitudinal Study, CHARLS），该队列从 2011 年启动，每两年进行一次随访，旨在收集中国 45 岁及以上中老年人家庭和个人的高质量微观数据。问卷内容主要包括：个人基本信息，家庭结构和经济支持，健康状况，体格测量，医疗服务利用和医疗保险，工作、退休和养老金、收入、消费、资产，以及社区基本情况等。

本研究使用的数据是 CHARLS 2018，涉及变量包含基本信息、生活习惯、医保类型、健康状况和生活满意度。剔除年龄小于 45 岁或大于 70 岁，健康状况和生活满意度变量信息缺失的人群，最终纳入分析的人群样本量为 13266 人(图 1)。

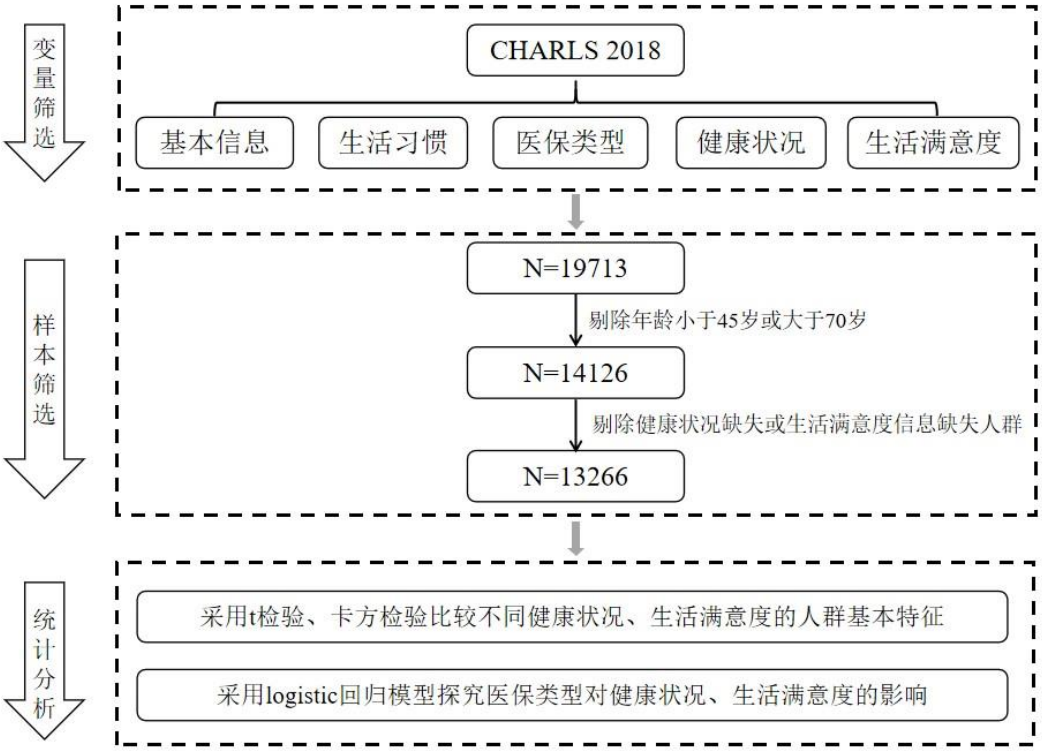


图1 本研究的技术路线图

1.2 研究方法

1.2.1 一般情况 包括患者的年龄、性别、教育水平、婚姻状态、吸烟状态、饮酒状态以及医疗保险。

1.2.2 医保类型 医疗保险类型通过问题“您本人目前是否参加以下医疗保险”，分为没有参加任何医保，城镇职工医疗保险，城乡居民医疗保险，城镇居民医疗保险，新型农村合作医疗保险，其他医保类型（公费医疗，医疗救助，单位购买的商业医疗保险，个人购买的商业医疗保险，城镇无业居民大病医疗保险以及长期护理保险），至少参加两种及以上的医保类型。

1.2.3 健康状况 健康状况通过是否具有慢性病以及自评健康状态两个条目来进行评估。通过问题“是否有医生告诉过您有以下这些慢性病？”定义慢性病的状态（无慢性病 *VS.* 有慢性病），无慢性病定义为不具有以下任何一种慢性病，有慢性病定义为至少有以下一种慢性病。通过问题“您认为您的健康状况怎样？”定义自评健康状态（健康 *VS.* 不健康），回答“很好、好、一般”认为是健康，回答“不好、很不好”认为是不健康。

1.2.4 生活满意度 生活满意度的定义通过问题“总体来看，您对自己的生活是否感到满意？是极其满意，非常满意，比较满意，不太满意还是一点也不满意？”分为满意（极其满意、非常满意、比较满意）和不满意（不太满意、一点也不满意）。

1.3 统计分析

采用 SPSS 25.0 软件对数据进行描述统计，计量资料以均数和标准差（$\bar{x}\pm s$）描述，计数资料以频数和百分率或构成比表示。采用 t 检验、卡方检验进行两组间的差异比较，采用 logistic 回归模型探究医保类型对健康状况和生活满意度的影响。检验水准 α=0.05。

2. 结果

2.1 调查对象基本情况

该研究人群平均年龄为（57.42±6.73）岁，男性 6221 人（46.89%），女性 7045 人（53.11%）。5644（42.54%）人患至少一种慢性病，3138（23.65%）人认为自我身体健康状态较差，1586（11.96%）人生活满意度较低。

与无慢性病的人群相比，患慢性病人群的年龄较大（57.90 *VS.* 57.07），吸烟（4.82 *VS.* 3.12），不饮酒（65.57 *VS.* 62.17）占比较高，且差异有统计学意义（P<0.05）。

与自评健康状态健康的人群相比，不健康人群的年龄相对较大（58.86 *VS.* 56.97），女性

（59.11 *VS.* 51.24），初中及以下（92.26 *VS.* 83.49），非结婚状态（23.65 *VS.* 15.63），不吸烟（97.08 *VS.* 95.85），不饮酒（73.93 *VS.* 60.42）和无医保（2.84 *VS.* 2.44）的人群占比相对较大，且差异有统计学意义（P<0.05）。

表 1 不同健康状况人群的基本特征比较

缺失变量：吸烟状态（n=4904），饮酒状态（n=3）

因素	慢性病（n，%）			自评健康状态（n，%）		
	无（n=7622）	有（n=5644）	P	健康（n=10128）	不健康（n=3138）	P
年龄	57.07±6.67	57.90±6.78	<0.001	56.97±6.67	58.86±6.72	<0.001
性别						
男	3585（47.03）	2636（46.70）	0.142	4938（48.76）	1283（40.89）	<0.001
女	4037（52.97）	3008（53.30）		5190（51.24）	1855（59.11）	
教育水平						
初中及以下	6501（85.29）	4850（85.93）	0.162	8456（83.49）	2895（92.26）	<0.001
中等	1056（13.85）	761（13.48）		1578（15.58）	239（7.62）	
大学及以上	65（0.85）	33（0.58）		94（0.93）	4（0.13）	
婚姻状态						
已婚且居住在一起	6380（83.71）	4684（82.99）	0.274	8545（84.37）	2519（80.27）	<0.001
其他	1242（16.29）	960（17.01）		1583（15.63）	619（23.65）	
吸烟状态						
不吸烟	4621（96.88）	3419（95.18）	<0.001	6081（95.85）	1959（97.08）	0.013
吸烟	149（3.12）	173（4.82）		263（4.15）	59（2.92）	
饮酒状态						
不饮酒	4737（62.17）	3701（65.57）	<0.001	6118（60.42）	2320（73.93）	<0.001
饮酒	2882（37.83）	1943（34.43）		4007（39.58）	818（26.06）	
医疗保险						
无医保	212（2.78）	124（2.20）	0.127	247（2.44）	89（2.84）	<0.001
城镇职工医疗保险	860（11.28）	679（12.03）		1335（13.18）	204（6.50）	
城乡居民医疗保险	915（12.00）	629（11.14）		1184（11.69）	360（11.47）	
城镇居民医疗保险	285（3.74）	231（4.09）		407（4.02）	109（3.47）	
新型农村合作医疗保险	4798（62.95）	3557（63.02）		6133（60.55）	2222（70.81）	
其他医保	100（1.31）	85（1.51）		155（1.53）	30（0.96）	
医保种类≥2	452（5.93）	339（6.01）		667（6.59）	124（3.95）	

与生活满意度高的人群相比，生活满意度低的人群相对年轻（56.81 *VS.* 57.50），女性（61.10 *VS.* 52.02），初中及以下（90.73 *VS.* 84.86），非结婚状态（24.78 *VS.* 15.49），吸烟（4.07 *VS.* 3.82），不饮酒（69.46 *VS.* 62.83）以及无医疗保险（4.73 *VS.* 2.23）的人群占比较大，且差异有统计学意义（P<0.05）。

表 2 不同生活满意度人群的基本特征比较

因素	生活满意度（n，%）		P
	满意（n=11680）	不满意（n=1586）	
年龄	57.50±6.74	56.81±6.62	<0.001
性别			
男	5604（47.98）	617（38.90）	<0.001
女	6076（52.02）	969（61.10）	
教育水平			
初中及以下	9912（84.86）	1439（90.73）	<0.001
中等	1675（14.34）	142（8.95）	
大学及以上	93（0.80）	5（0.32）	
婚姻状态			
已婚且居住在一起	9871（84.51）	1193（75.22）	<0.001
其他	1809（15.49）	393（24.78）	
吸烟状态			
不吸烟	7027（96.18）	1013（95.93）	0.689
吸烟	279（3.82）	43（4.07）	
饮酒状态			
不饮酒	7337（62.83）	1101（69.46）	<0.001
饮酒	4341（37.17）	484（30.54）	
医疗保险			
无医保	261（2.23）	75（4.73）	<0.001
城镇职工医疗保险	1458（12.48）	81（5.11）	
城乡居民医疗保险	1394（11.93）	150（9.46）	
城镇居民医疗保险	457（3.91）	59（3.72）	
新型农村合作医疗保险	7224（61.85）	1131（71.31）	
其他医保	170（1.46）	15（0.95）	
医保种类≥2	716（6.13）	75（4.73）	

缺失变量：吸烟状态（n=4904），饮酒状态（n=3）

2.2 医保类型对健康状况和生活满意度的影响

单因素回归结果显示，与无医保的人群相比，参加城镇职工医疗保险（OR=1.35, 95%CI: 1.06-1.72），城镇居民保险（OR=1.39, 95%CI: 1.05-1.84），新型农村合作医疗保险（OR=1.27, 95%CI: 1.01-2.09），其他医疗保险（OR=1.45, 95%CI: 1.01-2.09）的慢性病检出风险较高。调整其他协变量后结果显示，城镇职工医疗保险（OR=1.55, 95%CI: 1.14-2.13），城镇居民医疗保险（OR=1.60, 95%CI: 1.13-2.28），新型农村合作医疗保险（OR=1.44, 95%CI: 1.08-1.92）以及医保种类≥2（OR=1.51, 95%CI: 1.08-2.11）的慢性病检出风险增加。

与无医保的人群相比，参加城镇职工医疗保险（OR=0.42, 95%CI: 0.32-0.56），其他医疗保险（OR=0.54, 95%CI: 0.34-0.85）以及医保种类≥2（OR=0.52, 95%CI: 0.38-0.70）的健康自评状态相对较好。调整其他协变量后发现相似的结果，城镇职工医疗保险（OR=0.53, 95%CI:0.37-0.77），其他医保（OR=0.53, 95%CI:0.29-0.98）以及医保种类≥2（OR=0.64, 95%CI:0.43-0.95）的健康自评状态相对较好。

单因素回归结果显示，与无医保的人群相比，参加医保的人群的生活满意度相对较高。调整其他协变量后发现相似的结果，城镇职工医疗保险（OR=0.26, 95%CI:0.17-0.40），城乡居民医疗保险（OR=0.40, 95%CI:0.27-0.59），城镇居民医疗保险（OR=0.50, 95%CI:0.31-0.79），新型农村合作医疗保险（OR=0.62, 95%CI:0.44-0.86），其他医保（OR=0.35, 95%CI:0.16-0.74）以及医保种类≥2（OR=0.44, 95%CI:0.29-0.68）的人群生活满意度均相对较高。

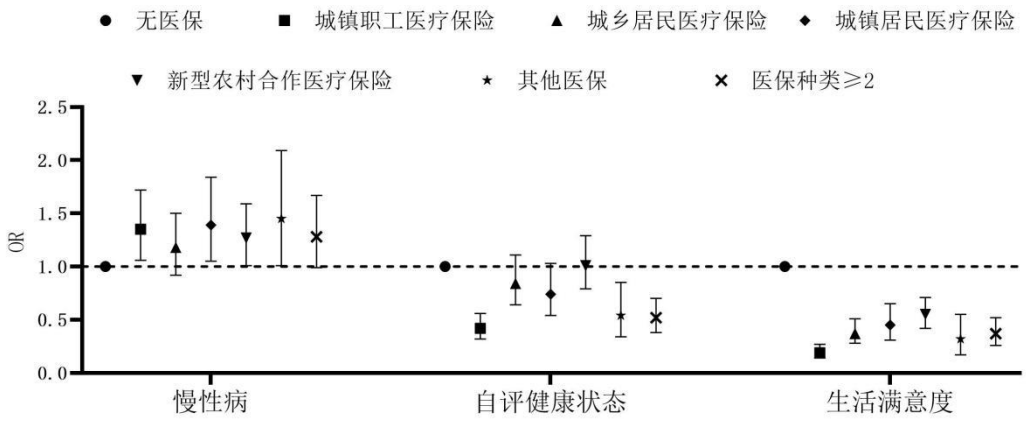


图 2 医保类型对健康状况、生活满意度的单因素 logistic 回归分析

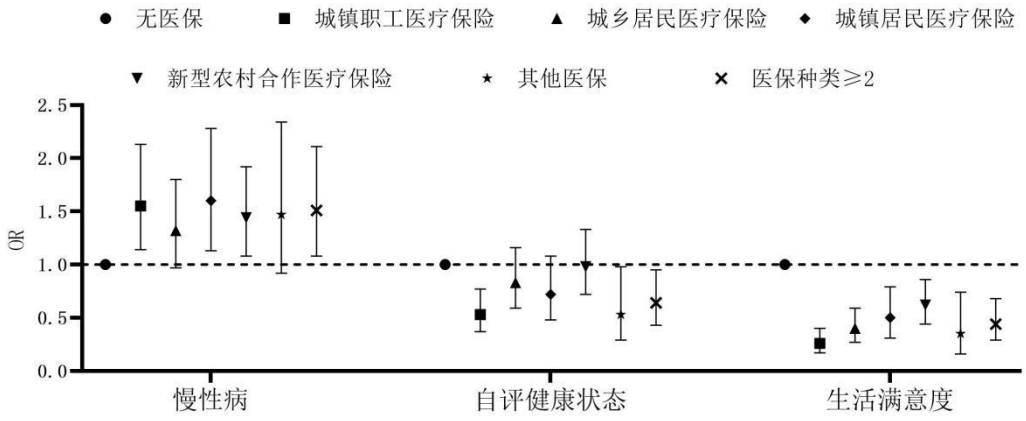


图 3 医保类型对健康状况、生活满意度的多因素 logistic 回归分析

3 讨论

在老龄化日益严峻的中国，关注中老年人的身体健康状况，提升老年人的生活满意度，应是我国未来国计民生的重点关注的问题。针对这一现状，建立基础医保为本的多层次医疗保障体系，对提升中老年人的生活质量有着重要指导作用。本研究以 2018 年中国健康与养老跟踪调查数据为基础，发现医疗保险有助于中国中老年人慢性病的早期发现和预防，提高中老年人的自评健康状态和生活满意度。与无医疗保险的中老年人相比，基础医保中，城镇医疗保险的改善程度较好；参加其他保险或者参加多种医疗保险的人群对改善中老年人的健康状况和生活满意度同样具有重要的意义。

3.1 不同医疗保险类型有助于中老年人慢性病的早期发现和预防。本研究结果发现，城镇医保，新型农村合作医疗保险以及其他种类的医保提高中老年人慢性病的早期发现的概率，而参加多种医保类型的人群慢性病的早期发现能力行对较弱。可能原因有[9, 10]，基础医保参保人员比例相对较大，机制相对而言更为成熟；而商业保险缴纳费用参差不齐，市场机制仍有待提高，致使参保人员比例相对较小。该研究结果也提示我们，为了更好的为慢性病患者提供健康管理，应建立健全医疗保障体系，有效发挥社会医疗保险兜底作用，政府积极引导商业健康保险以及其他补充保险参与，建立多层次医疗保险制度。

3.2 不同医疗保险类型改善中老年人健康状况。与不参加医疗保险的人群相比，参加城镇职工医疗保险，其他医保和多种医保的中老年人的健康自评状态相对较好。而城乡居民医疗保险、城镇居民医疗保险以及新型农村合作医疗保险没有明显的改善作用。究其原因[11, 12]，一方面是职工医保和居民医保的报销政策存在较大的差异，参加职工医保的报销比例相对高于居民医保。另一方面，参加职工医保的人群有足够的经济实力支付医疗费用，也会更注重日常体检等项目检查，能更早的对疾病进行预防。

3.3 不同医疗保险类型提高中老年人的生活满意度。与不参加医疗保险的人群相比，参加医疗保险的人群对生活的满意度显著升高，在调整其他协变量之后，医保对中老年人生活满意度的提升

作用依然存在。该研究结果在其他研究中也得到了证实[13]，基于 CHARLS 2015 的数据研究发现医疗保险能促进老年人社区居家养老生活满意度的提高[14]。一项基于 CHARLS 2013 的数据研究发现选择城镇职工医疗保险、城镇居民医疗保险或是新型农村合作医疗都会对老年人的生活满意度产生显著影响，且回归系数显著为正，该研究认为参加这三种医疗保险的老年人对生活的满意度更高[15]。

基于以上研究，本研究认为政府应完善社会医疗保险制度，提高医疗保险政策的覆盖率，满足老年人就医的医疗服务需求，是保障老年人生活满意状况的重要基础。

PU-230

大数据时代医院统计如何为医院管理服务

赵鹏飞
郑州大学第一附属医院

医院的统计数据主要是指对每位患者的病历数据进行相关整理以及数据统计，以此为医院对病人的相关治疗与后续诊断提供参考标准，由于当前医院医疗数据的大量产生，致使医院工作人员难以针对大量数据进行有关统计分析，因此，医院工作人员应依据大数据形式对相关医疗数据进行统计与管理，通过对大数据管理模式的学习与掌握，来对相关医疗数据进行研究与分析。医院对于医疗数据的统计方式也间接对医院综合发展产生了一定影响与作用。由于传统医疗数据统计方式无法满足现代社会对数据统计分析的大量需求，因此，医院工作人员必须充分意识到大数据统计方式的重要性以及对医院未来发展的相关影响。

一．大数据时代对医院统计的重要作用及意义

面对新时期大数据时代的全面到来，使医院面临着前所未有的重大机遇与挑战，医院统计部门利用先进大数据进行医疗数据的相关信息统计工作，使医院数据统计各个流程（信息数据录入，汇总，统计等环节）的速度及效率得到显著提升，节省了大量投入资源（人力，物力，财力等资源），使医院数据统计管理人员因计算失误所带来的数据偏差发生概率明显降低，从一定程度上消除了数据信息管理统计人员因主观因素致使医疗统计数据整体质量下降的不良现象，降低医院统计部门对各类医疗数据进行大量统计的工作难度，促使医院数据统计管理工作更加标准化规范化，致使医院数据统计管理工作整体质量得到显著提升。伴随现代社会医疗事业的不断发展与进步，医院统计部门通过利用大数据可使医院的综合发展能力及竞争力水平得到快速提升，使医院能够不断满足现代社会对医疗领域的大量需求，顺应时代发展。

二．医院大数据统计现状

由于当前医院各部门的重点仍在于对患者治疗及药品研发上，同时受到传统医疗观念的相关影响，忽略了对大数据统计领域方面的研究，致使医院整体统计能力及统计水平无法显著提升，各类医疗数据无法充分发挥内在利用价值及意义，同时，负责医疗数据统计的相关工作人员在统计领域的整理规划水平及专业操作技能水平有所欠缺，在对相关医疗数据统计形式上过于片面单一，致使医疗数据的统计工作效率难以进行大幅度提升，使医院综合发展能力及综合统计水平受到了严重影响及阻碍。

三．医院数据统计的存在问题

（一）轻视医院统计部门

传统的医院数据统计方式主要是医生通过对病人的相关诊断，在相互交流讨论中来实现对医疗数据的统计。但在当前社会，医生需通过大数据统计方式来对相关数据进行分类整理与统计，对各类统计数据展开详细研究与分析，来实现对医疗数据信息的快速有效统计，由于医院数据信息部分管理人员缺乏对医疗数据的统计力度和有效手段，导致医生难以针对大量医疗数据进行合理有效的分析及深入研究，除此之外，由于医院在数据统计管理领域没有投入足够的资金及相关医疗统计设备，致使医院数据信息统计管理人员缺少对数据统计的丰富经验及专业知识，医院数据信息管理人员无法对大量医疗数据进行科学合理的收集与整理，导致医院整体数据统计水平降低。

（二）数据利用价值无法充分发挥

很多医院信息统计人员对医疗数据统计过于固定单一，缺乏对医疗统计数据的系统性、统一性的整理与规划，甚至将各类医疗数据进行相互混杂，致使重要信息数据无法充分发挥自身利用价

值，使所有统计数据无法正确地展现自身利用特性及特点，导致医院统计部门难以从各种医疗数据中查找到准确的相关信息，对医院患者的相关疾病治疗产生了不良影响。

（三）数据统计方式相对固定单一

很多医院仍然按照传统的统计学方式对医院各类医疗数据进行相关统计，数据统计方式相对固定单一，缺乏一定灵活性和相关新颖性，致使大量宝贵时间流失，对医疗数据信息统计的整体效率提升产生严重影响，导致医院对医疗数据信息的整体统计能力水平降低，甚至部分乡镇医院依然采用人工记账的信息数据统计方式，致使医院统计部门在对各类医疗数据信息进行统计计算时，医疗数据统计出现偏差和错误现象时有发生，严重影响了各类医疗统计数据的正确率与精确度，同时由于人工记账方式统计速度较慢；难度较高，严重拖延医疗数据统计时间，致使医院统计人员整体工作质量降低，导致医院各类医疗信息数据的统计出错率与偏差率大幅度上升。

（四）无法实现大数据统计实际应用

由于部分医院统计基础硬件设备较差，导致医院统计人员在对各类医疗数据信息进行整理统计时，无法运用大数据手段及方式对各类医疗信息数据进行准确统计计算，致使难以对各类医疗信息数据进行信息化及科学化统计，导致统计数据文件的归档和整理工作难以顺利进行，使医生在对各类医疗信息数据研究分析时浪费大量工作时间，导致整体医疗统计分析效率降低，对医院的综合发展水平产生一定阻碍。同时，由于无法实现对大数据统计方法的实际应用，无法对各类医疗信息数据进行科学合理的有效运用，为后续各类医疗信息统计数据的审核及上报工作带来一定压力与负担，导致相关医疗信息数据无法充分发挥自身的利用价值，致使医疗数据整体统计质量水平有所下降。

（六）缺乏医疗数据统计评价反馈体系

由于目前很多医院统计部门缺乏相应的医疗信息数据统计评价反馈体系，导致统计人员无法及时对各类医疗统计数据作出明确评价及统计结论，无法将各类医疗数据信息统计结果上报至医院相关审核及考察部门，难以获得医院审核部门的相关有效反馈，及时对各类错误医疗统计数据作出相应改正，导致医院信息统计人员对医疗数据信息统计的质量水平有所降低。由于医疗数据统计评价反馈相关制度体系的缺失，使医疗数据信息统计管理人员难以严格要求自己，导致各类医疗数据统计错误现象时有发生，使各类医疗数据的真实性与可靠性有所降低，对医院整体医疗数据综合能力管理统计水平产生不良影响，对未来医院综合发展能力水平的快速提升产生一定阻碍。

四．大数据时代医院统计为医院管理服务的方法及策略

（一）提升统计人员素质

由于大数据时代的来临，医院信息统计人员应对自身统计专业水平及综合统计素养进行全面提升及强化，应不断改善统计形式及统计操作技能，来适应大数据时代的发展，相关医疗信息数据统计人员应根据社会发展局势，深刻意识到自身在统计领域发展的不足与欠缺，通过对网络科技及社会统计管理专业技能的不断学习与大量应用，强化自身在数据统计领域的专业理论知识及统计操作技能，同时要注重对相关数据的整合及归档，便于医生的后续研究分析工作得以有效进行，此外，医院信息统计人员要及时做好对统计数据的上报与回馈工作，使各类医疗统计数据能够充分发挥自身的利用价值。同时，医院数据统计部门需进一步加强对统计工作人员的管理，不断在数据统计工作的各个流程及环节进行深入完善和升级，努力提升医院数据统计人员的统计专业技能，以确保各类医疗统计数据的真实性及可靠性，全面提升统计数据信息的总体质量。

（二）改进完善统计方式

为确保所有医疗统计数据能够充分发挥自身的重要意义及作用，应针对医院统计制度体系进行相应改进及相关完善，医院统计工作人员在对医疗数据进行有关统计与整理时，需借助先进信息数据统计方式，以此来做到对数据高速有效的统计。统计部门内部要相互配合、相互协作，共同为数据统计工作作出自身贡献，对各类医疗数据的存在问题要做出及时反馈，使统计数据更好地展现出自身价值，从而全面提升统计数据的正确率与精确度。同时，应建立健全医疗数据统计制度体系，使各类医疗数据能够以系统化、统一化的形式呈现，致使医院研究管理人员能够通过系统化数据对其进行正确分析与精确判断，促使医疗数据统计方式能够进一步标准化与规范化。

（三）加强对大数据的应用

医院统计人员在对医疗数据进行整理统计的同时，应借助大数据技术进行信息化、科学化数据统计，通过对大数据信息统计技术，来对统计数据进行科学合理的整理和归档，进而为医生在病历分析上节省大量时间，提升对医疗数据分析的效率，通过对大数据统计手段的应用，建立健全医疗数据统计制度体系，促使医院统计工作人员能够发挥最佳水平，对医疗数据进行合理有效的运用，

进一步对医院物资设备的运营及病历数据的分析作出更为严格规范的管理，利用大数据统计方式，及时对错误数据进行有效纠正，从一定程度上来降低统计数据的偏差，促使医疗统计数据的准确率及精确度得以大幅度提升。

（四）提升医疗数据统计质量

医院信息数据统计部门应不断改善对传统医疗数据的统计收集方法及数据整合方式，不断提升医疗数据的统计质量，致使医院对医疗数据的统计方式能够顺应现代社会的发展与需求。其次，医院统计部门要充分利用移动互联网等现代信息技术手段来全面提升医院对医疗数据的统计效率，在使各类医疗统计数据的整体统计质量及效率进一步提升的同时，促使各类医疗数据能够充分发挥自身应用价值。医院统计部门可通过建立各类医疗数据检测小组，定期对统计部门的上报医疗数据信息进行严格确认检查，制定医疗统计数据的相关台账并及时对各类医疗数据的做好记录，确保各类医疗统计数据的真实性和可靠性，全面提升医院对各类医疗数据的统计质量。

（五）建立医疗数据统计评价反馈体系

医院信息数据统计部门在作出对医院大量医疗数据进行相关统计工作的同时，应建立对各类医疗统计数据的评价反馈体系，及时对各类医疗统计数据作出相关评价，并按时将各类医疗统计数据的相关评价结果上报至相关审核部门，通过医院审核部门的相关反馈，及时对各类错误医疗统计数据作出相应改正，促使各类医疗数据的统计质量能够显著提升。医院统计部门在制定相关评价反馈制度的同时，应对医院信息数据统计评价及相关反馈内容进行定期考核及严格审查，作为医院数据统计管理人员的年度考核的重要参考标准，以此来确保医院数据统计相关管理人员能够严格要求自己，进一步提升医院各类医疗统计数据的正确率，从而全面加强各类医疗统计数据的真实可靠性。

结语：面临大数据时代的到来，医院各部门应通过大数据统计手段来对数据统计质量和效率进行更为全面的提升与完善，促使大数据统计工作能够更好地为医院及现代社会发展进步服务，在利用大数据统计方式及手段的同时，不断提升医院统计部门人员的专业素养及综合应用能力水平，从而全面提升医院的综合数据管理能力，为未来医院的不断发展与进步作出应有贡献。

PU-231

浅谈医院信息网络安全建设与管理

张红霞、周麒、薛五洋
河南省儿童医院郑州儿童医院

一、引言

目前，医院信息化建设呈现出多科室、多平台、多系统协同的特性。各科室自有系统通过医院内网相互联系，实现资源整合及数据共享，极大缓解医疗资源不足的问题，提高患者就医体验。这种情况下，网络安全问题被急剧放大，传统的安全防御体系开始显得捉襟见肘。一方面，信息化建设的任一环节，如系统、网络、数据等都可能成为潜在的安全隐患。另一方面，工作人员对于网络安全的认识程度不深、专业技术不足，难以应对日益复杂的网络系统和越来越多变的网络攻击手段。如何解决管理制度不完善、人员专业性不足、技术手段薄弱等问题带来的危害，提升信息系统在“互联网+”时代的安全防护能力，成为医院信息化管理人员面临的主要挑战。

二、医院网络安全面临的问题

（一） 网络安全技术应用问题

近年来，云计算、物联网、大数据等信息化技术在三甲医院的应用日益广泛。信息化技术的应用促进了医院内网信息系统和卫生系统、互联网、金融机构类等服务专网或平台内外网融合。通过多种渠道如：网页、微信小程序、自助机和手机 App 等，为患者提供咨询问诊、挂号缴费、报告自助打印、信息推送、在线沟通等功能，极大提高了医院的服务质量，改善患者就医效率和体验。

高度依赖于内外网是信息技术应用的主要特点，然而，当网络出现波动或故障时，会给医院业务造成影响，甚至造成巨大损失。尽管各医院越来越重视网络安全，也制定了多项安全监管制度，

部署多种安全设备，仍无法完全消除安全隐患，大型医院信息系统遭受勒索病毒攻击的情况屡见不鲜，因此加强网络安全管理是医院信息化建设工作的重中之重。

（二）医疗数据安全问题

信息数据是医院重要的核心内容，已成为医学研究重要的分析基础。患者个人敏感信息，包括姓名、身份证、医疗数据等已成为国外敌对势力重点攻击对象。目前，大型三甲医院因中勒索病毒造成全面网络瘫痪、数据泄露甚至数据丢失等重大安全事件，大多是由于数据安全管理措施不到位导致：由于移动存储类介质随意接入内网引发病毒传播，或服务器终端开放端口没有按照“最小使用”原则开放，又或者是医疗数据的没有经过脱敏及加密等。

（三）网络安全管理人员不足

随着医院网络规模扩大，安全运营、安全分析、应急响应、攻防渗透等领域专业人员数量和能力不足等问题日益凸显。医院无法保留高端网络安全专业人才，而自主培养又面临成本过高、周期过长的问题。网络运营人员往往身兼数职，无法投入时间和精力提升专业技能，人才缺失导致医院在网络安全防御上主动性不足。

三、医院信息网络安全建设思考

（一）规范网络安全管理制度

医院网络安全要坚持“管理、技术双管齐下，七分管理、三分技术”的建设思路。参照等保2.0 标准、医院发展现状和行业最佳实践，制订医院信息安全管理总体方案，优化院内安全策略、管理制度和工作流程。落实日常管理方法及不定期检查方法，提高信息网络安全防护能力，切实做到“以此为据，有章可依”。

（二）规划网络安全建设方案

在进行网络建设方案的规划时，首先要对信息资产进行评估，识别医院面临的安全风险。其次，依据安全管理需求进行分析，形成符合医院的安全运营体系框架，在框架下对组织体系、制度体系、流程体系等进行详细设计。最后，明确安全建设方向和步骤，形成医院网络安全中长期建设实施蓝图。

（三）完善网络安全防护措施

服务器是医院信息系统的核心，在医院信息系统安全运行中起着主导作用。随着医院规模日益扩大，单核心的服务器架构无法再安全性和稳定性上满足医院发展需求。建设双机热备的网络架构是提高网络安全、确保服务器稳定、可靠、高效运行有效措施。双机热备包含两个部分：由备用服务器和共享磁盘阵列组成的备份系统和用于实现负载分担和故障切换的 VRRP 网络协议。当主服务器出现故障时，网络可以无缝切换到备用服务器继续运行，保障医院的各项业务正常运行。

此外，优化防火墙配置、合理设置访问规则也是网络安全管理的关键。首先，通过不断调整和升级拦截规则，及时发现和隔离可疑高风险。使用管理软件检查外网的使用权限，保证医院内外部网络的独立运行和业务协同。

其次，使用访问软件加强访问限制，避免非授权人员私自接入内网环境。最后，定期更新医院所有终端和服务器的病毒数据库，定期进行全面扫描，提高服务器终端的防病毒能力。

（四）加强管理人员安全培训

维持医院信息系统的稳定运行，专业的技术人才是重中之重。医院要从两个方面做好安全培训工作：一是树立安全意识，定期开展全院全员网络信息安全培训，明确安全责任，提高安全防范意识，形成安全管理意识，安全服从意识，从管理到运营，做到制度有落实，保障安全规划的执行。二是提升安全管理人员的专业技能，培养合格的管理团队，能够合理规划医院网络架构，快速定位并处理各种复杂网络安全问题。

四、结语

解决医院信息化建设面临的安全问题，需要从多个角度分析面临的问题。随着信息化技术的飞速发展和广泛应用，需要才能以动态的方式迎接各方面的挑战。除了技术上加强防范，建设健全的网络安全管理体系、明确各项规章制度、培养优秀管理人员，才是提高医院信息网络安全、保障业务正常运行的重中之重。

PU-232

基于可解释性机器学习的慢性心力衰竭合并肺部感染患者的生存分析

申采玉[2]，王帅[2]，周锐盈[1]，汪雨贺[3]，陈兴智[4]，杨枢[1]
[1]蚌埠医学院卫生管理学院
[2]蚌埠医学院公共卫生学院
[3]蚌埠医学院附属蚌埠第三人民医院急诊内科
[4]蚌埠医学院基础医学院

1 引言

慢性充血性心力衰竭（chronic heart failure, CHF）是一种复杂的临床综合征，通常由心脏结构和/或功能异常所引起[1]。据《中国心血管健康与疾病报告》报告显示[2]，中国 35-74 岁的人群中约有 400 万人患 CHF，患病率达 0.9%，其中病死率为 2.8%。重症监护病房（Intensive Care Unit, ICU）心力衰竭患者的肺炎发生率很高，死亡风险比其他患者高 4 倍[3]。但早期临床诊断不明显，难以诊断。再加上重 ICU 病房工作压力大、临床信号复杂繁多可能导致治疗延误、护理欠佳、甚至出现临床决策失误[4]。在以前，虽有一些常用的传统方法和指标被用来预测 CHF 合并肺部感染的诊断和预后，但效果并不明显。

随着人工智能的发展，机器学习技术已经广泛运用于临床，高性能的模型可以有效预测患者的预后提供临床决策[5]。本研究通过构建和对比 RF、LR、XGboost、LightGBM 四种机器模型，选择最优模型对 CHF 合并肺部感染患者进行生存结局的预测。用 SHAP 加性解释对最优模型进行可解释性分析。

2 对象与方法

2.1 数据来源

本研究是一项回顾性队列研究，数据来源于 MIMIC-IV（v.1.0）数据库[6]中提取的队列数据。该数据库是一个大型的、单中心的、公开可用的去身份化的患者数据库，包含了全面的患者信息，例如人口统计学、入院记录、实验室检查、药物、程序、体液平衡、出院小结、床边进行的生命体征测量、护理人员的笔记、放射学报告和生存数据等资料。数据库中的所有患者的隐私信息均去身份化处理。本研究通过考试获得了 MIMIC-IV 数据库的访问权限，且获得了贝斯以色列女执事医疗中心、麻省理工学院机构审查委员会的批准（证书编号：39168475）。

2.2 研究对象

本研究选择了 MIMIC-IV 库中 ICU 内诊断为 CHF 合并肺部感染的所有成年患者，通过人工审查国际诊断代码[7]（ICD-9）和（ICD-10）确定诊断 CHF 和肺部感染的患者。患者的纳入标准为：（1）患有 CHF 合并肺部感染的患者；（2）首次入住 ICU 的患者；（3）在 ICU 停留时间≥24 小时。排除标准包括：（1）年龄<18 岁；（2）严重肝病患者；（3）恶性癌症患者。

根据国际诊断代码 ICD-9 和 ICD-10，按病因学[8] 肺炎分类标准提取细菌性、病毒性、真菌性、非典型性肺炎患者，将 CHF 合并肺部感染患者分为细菌性肺炎和非细菌性肺炎（包括病毒性、真菌性、非典型性肺炎）两个亚组。本研究以患者的第一次住院时间为开始统计的起点，记录患者是否死亡，以患者在数据库记录的时间段内死亡或未死亡为统计终点。根据患者出院时的状态划分生存组和死亡组，研究的结局指标是 CHF 合并肺部感染患者的 30 天院内死亡率。图 1 是筛选患者的具体流程图。

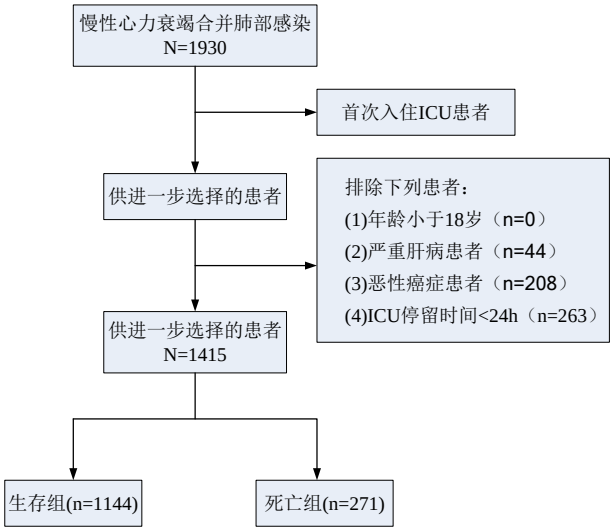


图 1 患者筛选流程图

2.3 数据收集

本研究从 MIMIC-IV 数据库中提取的变量有：人口统计学特征（性别、年龄、体重）、基本生命体征、入院后患者的病因诊断、并发症、各种实验室指标、治疗、简化急性生理学评分 II（SAPSII）、序贯器官衰竭评分表（SOFA）等资料。提取患者基本生命体征的平均值；实验指标是患者在 ICU 24 小时内的平均值；肌酸激酶同工酶 MB（CK-MB）、心肌肌钙蛋白 T（TnT）、B 型利钠肽（NT-proBNP）、氧分压（Po2）、二氧化碳分压（Pco2）；乳酸最小值（lactate_min）、乳酸最大值（Lactate_max）、碳酸氢盐最小值（Bicarbonate_min）、碳酸氢盐最大值（Bicarbonate_max）；24 小时累积尿量。所有数据脚本均可从 GitHub（https://github.com/MIT-LCP/mimic-iv）获取。

2.4 研究方法

2.4.1 数据处理与统计分析

缺失值小于 5%的用均数填补，大于 5%的用 KNN 分类算法进行多重补插，缺失值大于 30%的患者删除。连续性变量以 X±S 或 M（Q1,Q3）表示，分类变量用频率（%）表示，采用 χ2 检验或 Fisher 精确检验比较组间差异。绘制细菌性肺炎和非细菌性肺炎患者 Kaplan-Meier 生存曲线,使用对数秩检验进行检验。以 P<0.05 为有统计学意义。

2.4.2 四种模型

LightGBM（Light Gradient Boosting Machine）是一个迭代提升树系统,是梯度提升决策树（Gradient Boosting Decision Tree，GBDT）经过改进后的一种变体[9]。基于 Histogram 的决策树算法使用最大深度限制的叶子生长策略，缩短算法的时间，提高了整体的执行效率。

XGboost（eXtreme Gradient Boosting）是陈天奇等人开发的一个开源的机器学习项目。具有复杂度低、运行速度快、准确等优点[10]。该算法依据损失函数在梯度下降方向上组合多个 CART 树，以最小化损失函数，能够自动利用 CPU 的多线程进行分布式学习和多核计算，在保障预测准确度的前提下可以提高计算效率，适合处理大规模数据。

RF（Random Forests）是一种基于决策树的集成学习算法，采用自举重采样技术，从 N 个原始训练样本集中反复随机选取 B 个样本作为训练集，剩余样本作为测试集[11]。优势在于对样本量无要求，对异常值和噪声的容忍度高，不受变量间共线性问题的影响，尤其适用于具有交互作用的多维、复杂数据分析。

LR（Logistic Regression）也称为 logistic 回归,是一种线性回归模型。逻辑回归的应用依赖于用于应用一元和多元线性回归的相同假设。与多元线性回归类似，逻辑回归可以灵活地包含多个预测变量[12]。机器学习和 LR 在某些疾病预测方面表现相似，且可能比一些预测性能更高的机器学习

的偏倚性小。

2.4.3 SHAP 模型解释

Shapley 加性解释[13]（Shapley Additive explanation, Shapley）是一种解释机器学习算法结果的可视化方法，用于解释“黑盒”模型是如何衡量特征影响因变量的。机器学习的开发增加了模型的复杂性，加大了模型的理解难度。在医疗领域中，许多决策关乎病人的生死，复杂且缺乏可解释性的预测模型会削弱人们信任度。

为了应对该问题，许多学者对此展开了大量的研究。如生物信息学中[14]，Kang M 等人基于可解释机器学习为解释复杂和异构的生物数据提供了新的依据；在影像医学中[15]，Ma M 等人基于可解释机器学习模型预测乳腺癌分子的亚型，使放射科医生区分乳腺癌分子亚型的准确性得到了显著提高；在药学领域中[16]Simon ST 等人基于可解释机器学习预测药物是否诱发 qt 综合征（diLQTS）的研究。因此本研究基于 SHAP 加性解释，对模型的全局变量和单样本变量进行分析，解释这些变量怎样影响模型中的死亡率。

3 结果分析

3.1 基本信息

研究共纳入 1415 名 CHF 合并肺部感染的患者，271 名患者死亡，1144 名患者存活。据统计分析患者的平均年龄为 69.8 岁，死亡者的年龄均大于幸存者，女性占 46.3%，男性占 53.7%。在死亡组中，年龄、体重、心率、平均收缩压、呼吸率、红细胞（RBC）、白细胞（WBC）、血小板（Platelet）、 血尿素氮（BUN）、B 型利钠肽（NT-proBNP）、乳酸、碳酸氢盐、24 小时累积尿量存在显著差异。关于合并症，死亡组患有脑血管疾病、慢性肾病等并发症的占比大于生存组。两组患者的基线特征、生命体征和实验参数的比较详见表 1。

表 1 人口统计学基线特征表

M(Q1,Q3)

指标	总人数（n=1415）	生存组（n=1144）	死亡组（n=271）	t/Z/χ2 值
人口				
年龄/岁	69.8（61.0，80.0）	69.2（61.0，80.0）	72.4（65.0，82.0）	<0.001
性别（%）				
女	655（46.3）	538（47.0）	117（43.2）	
男	760（53.7）	606（53.0）	154（56.8）	
体重（kg）	84.2（65.8，97.0）	85.5（66.5，98.5）	78.9（63.5，90.0）	<0.001
基本生命体征				
心率/（次·min-1）	85.2（73.7，95.0）	84.4（72.8，94.3）	88.8（77.2，101.2）	<0.001
平均收缩压（mmHg）	116.6（104.5，126.2）	117.3（105.0，127.4）	113.5（103.3，122.4）	<0.001
平均舒张压（mmHg）	61.2（53.8，67.1）	61.4（53.9，67.3）	60.4（53.5，66.6）	0.157
呼吸/（次·min-1）	20.6（17.7，23.1）	20.4（17.6，22.7）	21.4（18.6，24.1）	<0.001
体温（/°C）	36.8（36.6，37.1）	36.9（36.6，37.1）	36.7（36.4，37.1）	0.003
氧饱和度（/%）	96.6（95.3，98.3）	96.6（95.3，98.3）	96.6（95.2，98.6）	0.319
并发症/例（%）				
脑血管疾病	221（15.6）	157（13.7）	64（23.6）	<0.001

				1
轻度肝病	136（9.6）	97（8.4）	39（14.4）	0.003
截瘫	66（4.6）	46（4.0）	20（7.3）	0.018
CKD	578（40.8）	452（39.5）	126（46.5）	0.035
实验室指标				
血糖（mg/dL）	211.0（115.4，179.7）	151.5（115.1，177.9）	462.3（116.7，183.3）	0.038
红细胞（m/μl）	3.3（2.8，3.8）	3.4（2.8，3.8）	3.2(2.7，3.7)	0.024
白细胞（10^3/μL）	11.4（7.5，13.5）	10.8（7.2，12.9）	14.3（9.1，17.3）	<0.001
血小板（10^3/μL）	226.6（150.0，288.0）	231.2（152.0，292.0）	209.4（130.0，273.0）	0.004
血红蛋白（g/dL）	9.8（8.3，11.0）	9.9（8.4，11.1）	9.6（8.1，10.8）	0.057
血清钾（mmol/L）	4.2（3.7，4.5）	4.2（3.7，4.5）	4.2（3.8，4.6）	0.162
血清钠（mmol/L）	139.4（136.0，142.0）	139.3（136.0，142.0）	139.7（136.0，144.0）	0.240
血清钙（mmol/L）	8.5（8.1，9.0）	8.6（8.1，9.0）	8.4（7.9，8.8）	0.021
尿素氮（mg/dL）	36.9（18.0，48.0）	34.4（18.0，43.5）	47.4（24.0，66.0）	<0.001
肌酐（mg/dL）	1.8（0.8，2.0）	1.7（0.8，1.8）	2.0（0.9，2.4）	0.037
肌酸激酶（ng/ml）	11.6（2.6，6.2）	10.9（2.5，6.0）	15.4（2.8，9.2）	<0.001
肌钙蛋白-T（ng/ml）	0.7（0.08，0.53）	0.64（0.08，0.52）	0.88（0.07，0.62）	0.037
脑钠肽（pg/mL）	13207（6192，18491）	12852（5555，17906）	14703（8280，20015）	<0.001
氧分压（mmHg）	108（71.0，123.0）	110（71.2，126.0）	100（69.0，116.0）	0.038
二氧化碳分压（mmHg）	43.6（38.0，47.0）	43.6（38.7，47.0）	43.4（36.0，48.0）	0.777
乳酸最小值（mmol/L）	1.6（1.0，1.8）	1.5（1.0，1.8）	1.9（1.1，2.1）	<0.001
乳酸最大值（mmol/L）	2.8（1.5，3.3）	2.6（1.4，3.0）	3.4（1.6，4.1）	<0.001

续表 1
M(Q1,Q3)

指标	总人数（n=1415）	生存组（n=1144）	死亡组（n=271）	t/Z/χ2值
碳酸氢盐最小值（mEq/L）	25.9（19.0，26.0）	23.0（19.0，26.0）	21.3（18.3，25.0）	<0.001
碳酸氢盐最大值（mEq/L）	25.7（22.0，29.0）	26.1(22.0，29.0)	25.0(21.0，28.0)	0.003
尿量（ml）	1759（905，2295）	1848（994，2395）	1382（584，1885）	<0.001

重症评分				
SOFA	6.9（4.0，9.0）	6.3（3.0，9.0）	9.2（6.0，12）	<0.001
SAPSII	42.0（32.0，50.0）	40.2（31.0，48.0）	49.4（40.0，58.0）	<0.001
SIRS	2.6（2.0，3.0）	2.5（2.0，3.0）	2.9（2.0，4.0）	<0.001

3.3 最终纳入模型的特征

本研究基于 LightGBM 筛选出 20 个重要特征构建模型。LightGBM、XGBoost、RF 模型都可获取特征的重要性且排序结果差异不大。我们以 XGBoost 为例对特征进行排序后确定最终纳入模型的特征，包括 SAPSII 评分表、SOFA 评分表、白细胞、血尿素氮、心血管疾病、乳酸最小值、血小板、尿量、体重、年龄、体温、心率、呼吸、血清钙、血清钠、血糖、二氧化碳分压、氧分压、脑钠肽、舒张压。如图 2 所示。

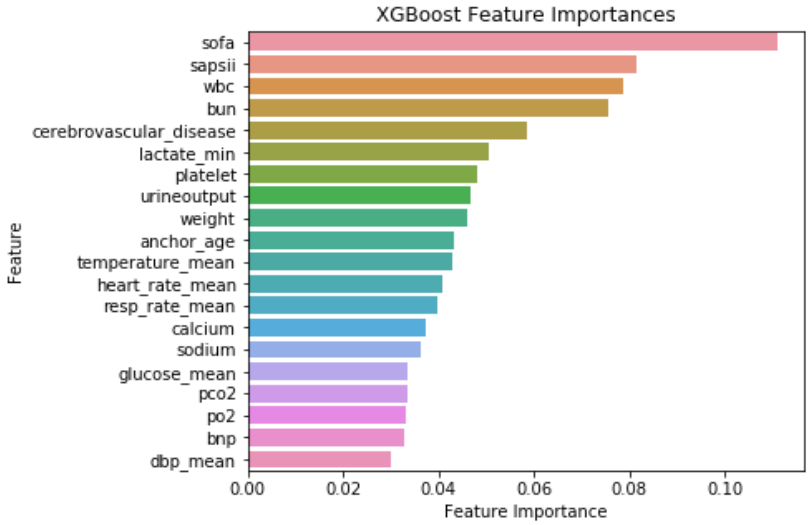


图 2 XGBoost 重要特征排序图

3.4 模型构建和比较

3.4.1 两组肺部感染患者的 30 天生存分析

本研究将患者分为细菌性肺炎（841 人）和非细菌性肺炎（574 人）两个亚组，绘制 Kaplan-Meier 生存曲线比较两组患者的 30 天生存率。并使用对数秩检验（Log-Rank Test）比较进行检验，如图 3 所示，结果显示 P>0.05，因此尚不能认为两组患者的生存曲线有显著性差异，即细菌性肺炎和非细菌性肺炎患者的生存差异不大。

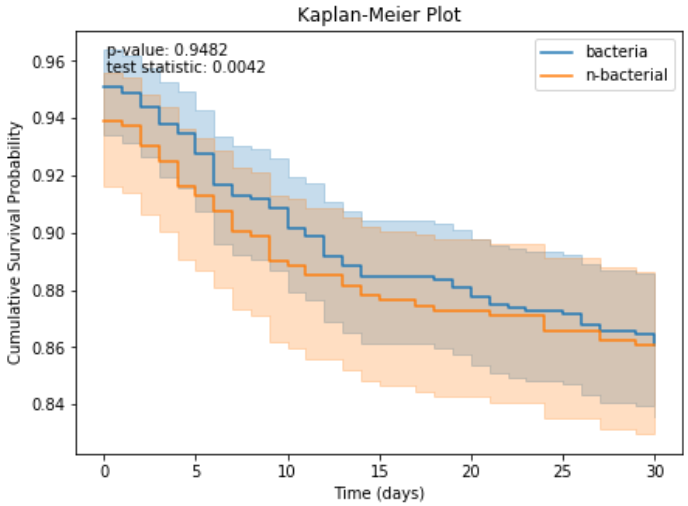


图 3 两组患者的生存曲线

3.4.2 两个亚组的模型构建和比较

基于 LightGBM 模型筛选相关性最大的 20 个特征，将数据集分为细菌性肺炎（n=841）、非细菌性肺炎(n=574)两组数据。两亚组数据都随机分为两组数据，取 80%作为训练集， 20%作为测试集，对特征进行归一化处理，构建四个不同的分类模型：XGBoost、LightGBM、逻辑回归（LR）、随机森林（RF）模型，对每个模型进行训练，使用训练集和测试集评估模型。评估指标包括准确性、精确度、召回率、F1 值、ROC 、AUC。ROC、 AUC 表示分类模型的区分能力，详见图 4。

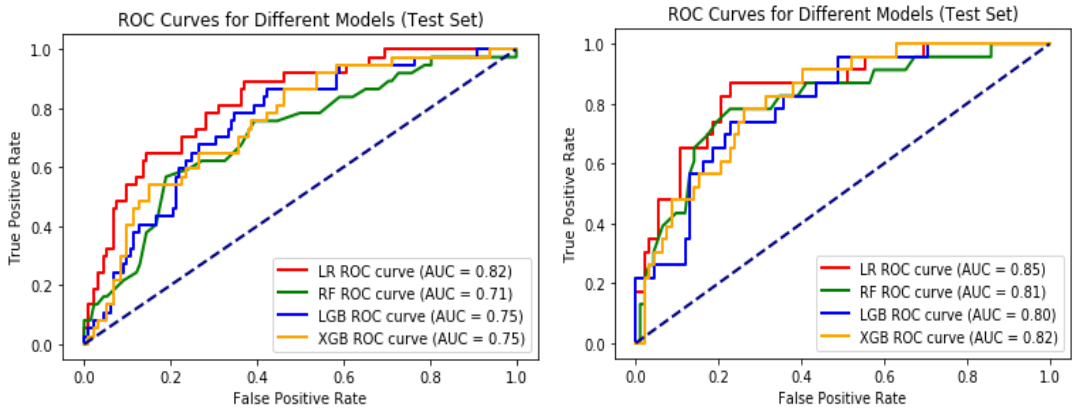


图 4 两组患者的 ROC 曲线图

表 2、表 3 结果显示，两个亚组构建的机器学习模型和经典的逻辑回归模型的预测性能相差不大。XGBoost 模型相对 RF、LightGBM 两个模型性能更好。由于本研究数据集的生存组和死亡组的患者比例不平衡，数据类别不平衡时召回率（Recall Score）比准确性、精确度等指标更有参考性。召回率得分越高，机器学习模型在识别正面和负面样本能力越好。而 XGBoost 两个亚组的召回率为 0.541、0.500，F1 值为 0.478、0.524 均高于其他模型，综合所有指标 XGBoost 的预测性能最好。

表 2 细菌性肺炎模型性能

模　型名称	准确性	精确度	召回率	F1-Score
LR	0.805	0.591	0.351	0.44

RF	0.722	0.407	0.595	0.484
LightGBM	0.728	0.390	0.432	0.410
XGBoost	0.763	0.465	0.541	0.500

表 3 非细菌性肺炎模型性能

模型名称	准确性	精确度	召回率	F1-Score
LR	0.844	0.727	0.348	0.471
RF	0.826	0.636	0.304	0.412
LightGBM	0.800	0.500	0.261	0.343
XGBoost	0.826	0.579	0.478	0.524

3.5 基于 SHAP 的模型解释性分析

3.5.1 全局样本特征解释分析

图 5 通过平均绝对 SHAP 值确定每个变量对 XGBoost 模型预测结果的全局重要度。SOFA 评分的均值在所有预测时段中具有最强的预测价值，SOFA 评分越高，患者的死亡风险也越高。其次白细胞、血尿素氮、呼吸、年龄、B 型利钠肽（NT-proBNP）、心率、乳酸、脑血管疾病等特征的均值也具有重要的预测价值。

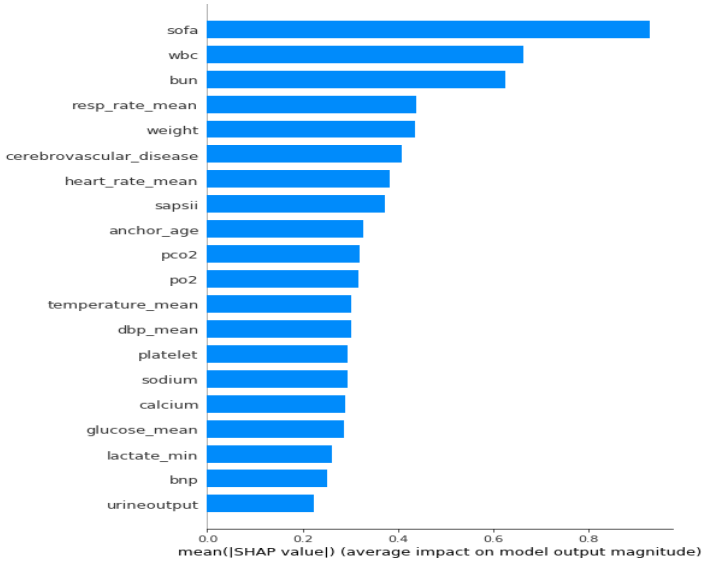


图 5 重要性排序图

图 6 显示了预测变量与目标结果的正负关系，用不同颜色的点绘制所有患者对结果的归因，红色点表示高风险值，蓝色点表示低风险值。可以看到，白细胞、尿素氮的增加对预测结果有正向影响，使预测结果倾向于死亡；尿量、体重的增加对预测结果有负向影响，使预测结果倾向于生存。

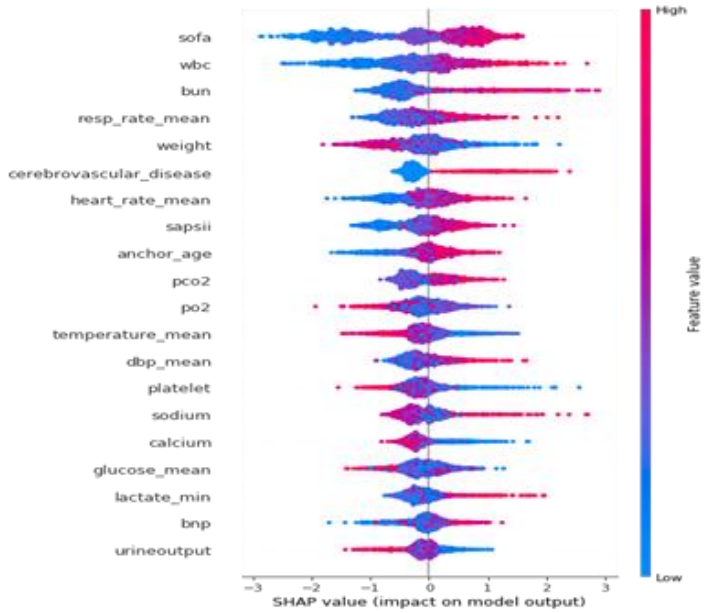


图 6 SHAP 摘要图

3.5.2 单样本预测特征解释分析

为了进一步探索 XGboost 模型在个体患者和临床应用方面的特征贡献，本研究绘制了 SHAP 的单样本特征图。图 7 Force Plot 图是 CHF 合并细菌性肺炎存活患者的 SHAP 特征贡献图，蓝色是患者的生存影响因素，该患者存活的原因是 SOFA 评分低、体重较大、年龄较小、尿素氮水平较低，无其他临床并发症。

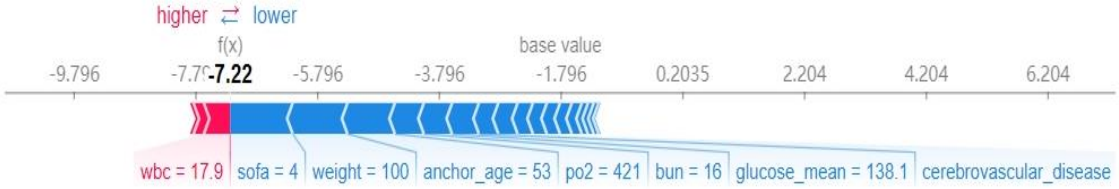


图 7 CHF 合并肺部感染存活患者示例图

4 讨论

4.1 疾病机制

研究表明，约有 59.6%的 CHF 患者合并肺部感染[17]。慢性心力衰竭和肺部感染相互影响。慢性心力衰竭患者心脏收缩力下降，肺静脉回流受阻，导致肺循环淤血和肺水肿，肺泡液积聚影响细菌清除和局部防御机制从而引发肺部感染。肺部感染会导致体内炎性因子释放和呼吸功能障碍导致心肌缺血缺氧损伤，心脏的负荷增大，进一步使患者的心力衰竭症状加重[18]。

在 ICU 中慢性心力衰竭合并肺部感染的患者临床表现及实验室检查往往不典型，易延误诊断影响患者的预后。本研究构建的 CHF 合并肺部感染患者的预测模型，有助于医务人员预测患者 30 天的院内死亡率，降低患者的死亡风险。据模型显示，年龄、体重、心率、平均收缩压、呼吸率、红细胞计数（RBC）、白细胞计数（WBC）、血小板（Platelet）、 血尿素氮（BUN）、B 型利钠肽（NT-proBNP）、乳酸、碳酸氢盐、24 小时的累积尿量等因数与患者的死亡率有关。

4.1.1 死亡率增加的可能机制

心力衰竭的患病率随着年龄的增加而翻倍，成年人群的发病率占 1%左右，60 岁以上的人群发发病率则超过 5%，且多属于发病急性期[19]。结果显示，年龄是 CHF 合并肺部感染患者的重要危

险因数（P<0.05），本研究总体患者的平均年龄为 69.2 岁，死亡组患者的平均年龄为 72.4 岁，远高于生存组患者的平均年龄，这与 Li Shen[20]等人的研究一致。

在模型中 WBC 的重要性排第二，CHF 合并肺部感的生存组患者的白细胞明显低于死亡组患者，且全身炎症反应综合征（systemic inflammatory response syndrome, SIRS）评分死亡组评分高。这与心力衰竭合并肺部感染后心肌细胞损伤，患者心肌抑制和血管通透性增加导致体内炎症反应迅速激活有关[21]。血尿素氮和尿量是临床上常用的检测肾功能的指标。尿素升高和尿量减少提示患者可能有肾功能损伤[22],少尿和肾功能损伤会导致液体潴留增加心脏的容量负荷，使 CHF 合并肺部感染患者的症状加重。

在模型中死亡患者的 BNP 水平明显升高，B 型利钠肽（NT-proBNP）是影响患者生存的重要因素。NT-proBNP 是心室合成分泌的一种具有抑制肾素-血管紧张素-醛固酮系统活性、利尿、排钠等作用的神经激素。稳定性好，半衰期长，反应心力衰竭患者病情严重的程度 [23]。研究结果显示，患者体内的乳酸水平与 CHF 合并肺部感染患者的死亡率成正相关（P<0.001）。动脉血乳酸可以评估组织和器官的微循环功能，是组织氧气可用性、新陈代谢和身体细胞灌注的重要指标，动脉血乳酸在患者预后方面起着重要作用[24]。

4.1.2 生存率增加的可能机制

与传统研究不同的是，在模型中体重是 CHF 合并肺部感染患者的保护因素，生存组的体重均较死亡组的体重略高。这与国外研究结果一致[25]，在慢性心力衰竭的情况下，与正常体重的患者相比超重和轻至中度的肥胖患者生存率有显著的提高，这一现象被称为“肥胖悖论”。本研究中，钙离子水平是 CHF 合并肺部感染患者预后的关键因素。据研究[26]，低钙血症和高钙血症都与患者的死亡风险增加有关。游离血清钙离子是人体中重要的电解质，对心肌细胞起到激发、收缩、耦合的作用，钙离子水平的稳定有助于患者的预后。

4.2 研究局限性

本研究有以下局限性：首先，本研究是一项回顾性研究，选择性偏倚几乎无法避免。但本研究严格设定纳排标准使纳入的案例信息更加真实准确。第二，本研究的数据是来自单中心的数据库，存在研究数据量小、不平衡等问题。第三，尽管我们的模型通过内部验证集进行了广泛的测试，但研究中的结果是否可以应用于其他人群尚不能确定。因此，未来我们还需要在各种临床环境中进行进一步研究以确认我们的研究结果。

5 结论

目前模型包含的预测变量有 SOFA 评分、白细胞、血尿素氮、呼吸、年龄、B 型利钠肽、心率、乳酸、脑血管疾病、氧分压、二氧化碳分压等 20 个。在本研究中，我们开发和验证了三种机器学习模型和经典的逻辑回归模型对比。XGBoost 两个亚组模型的准确性为 0.763、0.826。精确度为 0.465、0.579 比 RF 和 LightGBM 模型更高。召回率分别为 0.541、0.500，F1 值分别为 0.478、0.524 较 LightGBM、LR、RF 三个模型更好。综合来看在不平衡的小数据样本中 XGBoost 模型的性能较其他模型的性能更好。而基于 SHAP 方法对 XGBoost 模型的预测特征进行解释，确保了模型的临床可解释性，有利于临床医生更好的理解模型决策的过程，提高患者的存活机会。

PU-233

肾透明细胞癌微环境预后相关基因的组学分析

李卓阳，于琦*，贺培凤*
山西医科大学管理学院

目的 探讨肾透明细胞癌（clear cell renal cell carinoma，ccRCC）微环境对患者预后的影响，分析预后相关基因的预后价值。

方法 结合TCGA（The cancer genome atlas）数据库中ccRCC转录组数据和相应临床信息，采用ESTIMATE算法进行免疫评分和基质评分，采用Kaplan-Meier法分析ccRCC微环境与患者预后的相关性；采用R语言程序包limma筛选差异表达基因（DEGs），通过GO、KEGG富集分析确定DEGs潜在功能，通过COX回归分析和蛋白质互作（PPI）网络分析筛选与预后相关的核心DEGs；最后采用差异分析、生存分析和临床相关性分析以评价核心DEGs的潜在预后价值。

结果 从TCGA数据库下载611例ccRCC患者的转录组数据和相应临床信息，Kaplan-Meier生存曲线显示免疫评分低组患者的生存时间比高分组长，基质评分组无明显差异。差异表达基因分析得到43个上调基因和42个下调基因，富集分析显示DEGs主要与免疫细胞增殖和分化、免疫反应、细胞因子活化和质膜外侧等相关。COX回归分析和PPI网络分析显示IGLL5为预后相关核心DEG，进一步分析显示IGLL5在肿瘤样品中的表达水平高于正常样品，同一患者的肿瘤和正常样品中观察到类似结果；CD2低表达与BC患者预后呈正相关，其表达水平随着分期的进展而升高。

结论 ccRCC微环境免疫细胞含量与患者预后呈负相关，IGLL5表达水平可能有助于ccRCC患者预后的判断。有必要开展进一步的基础研究加以验证。

PU-234

基于在线健康社区的老年健康信息需求研究——以肺癌、高血压为例

张楚[1] 郑颖[1] 黄锦泉[2] 谢晖[1] 刘玉文[2]
1 蚌埠医学院 护理学院
2 蚌埠医学院 卫生管理学院

0 引言

目前，人口老龄化是世界各国共同面临的全球性议题[1]，在中国，据第七次人口人口普查报告显示[2]，60 岁以上老年人口约 2.64 亿，占总人口的 18.7%，我国将进入老龄化急速发展阶段。人口老龄化给我国的老年健康事业带来巨大挑战，据中国健康与养老追踪调查（CHARLS）数据显示[3]，我国 60 岁及以上老年人的慢性病患病率为 86.2%，共病患病率为 65.1%，在我国仅有 43%的老年人身体健康状况良好[4]，疾病的困扰极大程度上降低了生活质量，老年人迫切需要疾病治疗、预防保健和营养健康等方面的健康信息，以解决自身健康问题。所以，当前医疗卫生人员应考虑如何更好地为老年人提供精准的健康信息服务，满足老年人健康信息需求。

在漫长的疾病自我管理过程中，老年人通过阅读书籍、报纸、看电视、听广播和关注社区健康宣传教育等方式获取健康信息[5]。近年来，随着互联网技术在我国迅速普及与应用，老年群体加速融入网络社会，越来越多的老年人借助互联网获取健康信息[6]，在线健康社区平台成为广受欢迎的专业健康网站，用户可以在平台上进行健康知识分享、专家咨询和经验交流[7]。据统计，近年来在线健康社区的问诊量激增[8]，OHCs 为用户获取健康信息提供了一个安全、隐私、专业及实时的健

康交流环境，另外，OHCs 中储存了海量的老年问诊信息，为研究人员捕捉老年人复杂的健康信息需求提供了充足的数据资源。

但是，成全等[9]在 OHCs 进行评价时发现：国内 OHCs 存在信息参差不齐、知识表达不规范等问题，对老年人进行信息选择、健康决策造成干扰；而且，目前关于老年人健康需求领域的研究多采用问卷调查、质性访谈的研究方法，而基于文本数据结合自然语言处理（NLP）等技术深度挖掘老年健康信息需求的研究较少。因此，本文研究目标主要有以下两个方面：第一，借助 OHCs 梳理老年健康信息需求分布，协助老年患者及医疗卫生人员进行健康决策；第二，基于 OHC 深入了解老年健康信息需求，为 OHC 优化老年健康信息服务提供依据。

1 对象与方法

1.1 数据来源

本研究根据各大健康医疗网站在“站长之家”（www.chinaz.com）上的综合排名，选择影响力广泛、信任度较高的“好大夫”在线健康平台作为数据来源。使用“八爪鱼”采集器爬取自 2022 年 11 月-2023 年 4 月年龄≥60 岁老年人的问诊信息，爬取内容包括就诊人性别、年龄、疾病描述、希望获得的帮助、就诊科室和医生的职称等信息。共爬取数据 47792 条。

1.2 数据处理

对爬取的问诊数据进行去重、人工填补、统一内容格式等数据清洗工作，并删除与老年问诊数据不相关的内容，最终得到 44354 条老年健康问诊数据。由于数据量大且混杂，本研究首先由两名具有医学背景的研究者根据“国际疾病分类第十版（ICD-10）”标准抽取 1%的问诊数据进行人工标注，并对归并结果进行一致性检验，经计算，本研究 Kappa=0.936,归类结果具有较强的一致性。Kappa 值计算结果见表 1。

表 1 一致性检验结果

项目		值	渐近标准误差[a]	近似 T[b]	渐进显著性
协议测量	Kappa	0.936	0.013	54.315	0
有效个案数		445			

a. 未假定原假设；b. 在假定原假设的情况下使用渐近标准误差。

利用 Python 的 Jieba 包对老年问诊数据进行分词，在分词的过程中使用“搜狗医学词汇”和“爱问文库医学词汇大全”作为分词词典供机器学习，使用哈工大停用词表、百度停用词表和四川大学机器智能实验室停用词表去除问诊数据中无实际意义的词语，将老年问诊数据集转化成“文本-词汇”矩阵；运用 TF-IDF 算法对“文本-词汇”矩阵中的特征词进行权重调整，最终建立新的“文档-词汇”语料库，以此作为后续研究的实验数据。

1.3 K-Means 聚类分析

K-Means 算法是一种无监督机器学习的聚类算法[10]，简单高效，在主题聚类分析中应用广泛。K-Means 聚类算法基于欧式距离将给定的数据集划分为 K 个簇，簇内的各个点紧密相连，而簇间的距离尽量的大，由此进一步归纳为不同的主题类别，以探究文本主题分布及主题演化趋势。在对文本数据集进行规范数据结构后，基于肘部法则确定聚类数（k），轮廓系数法评估聚类效果，借助 Python 的机器学习工具 Scikit-learn 执行 K-Means 聚类训练。

1.4 统计学方法

采用 Excel2019、SPSS25.0 进行数据统计分析，计量资料采用均数±标准差($\bar{x} \pm s$)表示，计数资料采用频数、构成比（%）和百分比（%）表示，卡方检验比较不同性别和年龄组间的健康信

息需求的主题差异，以 P<0.05 为差异有统计学意义。

2 结果

2.1 问诊用户基本特征

本研究收集的数据中，老年女性问诊数量 22532 条，老年男性问诊数量 21831 条，两者问诊量之比为 1.03：1。从年龄上看，年轻老年人（60-74 岁）的问诊量较大，占总问诊量的 78.2%，随着年龄的增长，问诊量呈现下降趋势，见图 1。其中，36851 名用户描述了患病时长，患病时间<1 年的用户占 89.7%，患病 1-5 年的用户为 6.5%，患病 6-10 年的用户为 1.7%，患病时间≥11 年的用户为 2.1%。

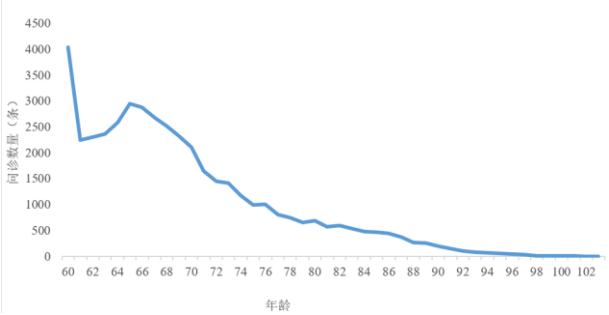


图 1 在线健康社区老年人年龄-问诊数量趋势

2.2 被咨询者所在科室、职称特征

统计 40585 名用户所咨询科室的分布，老年问诊科室主要集中在内科、外科、肿瘤科、五官科、皮肤性病科，占比分别为 41.4%、28.5%、7.5%、4.2%、4.1%，其中呼吸内科和骨科是内外科中在咨询最多的科室。统计已知的 40130 名被咨询者的职业情况，被咨询者以医生为主，占比 99.8%，其他包括护士、康复师、药师、技师、心理治疗师、营养师共占比 0.2%，其中主任医师占比 43.2%，副主任医师占 32.2%，主治医师和医师共占比 24.6%。

2.3 疾病问诊分类特征

根据“国际疾病分类第十版（ICD-10）”标准及新冠肺炎问诊情况，将疾病问诊分为以下 17 大类。见表 2。

表 2 疾病问诊分类情况

疾病分类	男	女	占比（%）	平均年龄
肿瘤	6466	4967	25.78%	68±6.36
新冠肺炎	3045	3341	14.40%	72.78±9.34
循环系统疾病	3316	3069	14.40%	69.49±7.46
呼吸系统疾病	1587	1501	6.96%	68.66±7.39
消化系统疾病	1331	1512	6.41%	68.83±7.69
泌尿生殖系统疾病	1073	1029	4.74%	68.53±7.31
肌肉骨骼系统和结缔组织疾病	789	1263	4.63%	68.07±6.69

症状、体征和临床实验室异常所见	902	1150	4.63%	69.42±7.9
内分泌、营养和代谢疾病	537	987	3.44%	67.16±6.35
神经系统疾病	671	818	3.36%	70.25±7.62
皮肤和皮下组织疾病	653	772	3.21%	71.54±9.41
损伤、中毒和外因的某些其他后果	360	571	2.10%	70.92±9.0
眼和附器疾病	340	468	1.82%	68.71±7.3
精神和行为障碍	191	436	1.41%	68.41±6.95
某些传染病和寄生虫病	296	288	1.32%	69.09±7.39
耳和乳突疾病	133	208	0.77%	67.52±6.82
血液及造血器官疾病	141	143	0.64%	69.07±7.36

2.4 确定聚类数量

在肿瘤和循环系统疾病中，关于肺癌、高血压的问诊量最多、而且这两类疾病在慢性病种类中具有一定的代表性，因此进一步探讨老年人对这两类疾病的信息需求情况。基于规范后的 2538 条肺癌问诊数据，计算老年问诊信息的聚类数量，结果显示，当 K=20 时，聚类曲线出现明显的拐点，同理，基于规范后的 921 条高血压问诊数据，计算老年问诊信息的聚类数量，结果显示，当 K=10 时，聚类曲线出现明显的拐点，在该拐点处所形成的主题类别能够概括文本特征。

2.5 主题分析结果

利用 K-Means 方法对 2538 条肺癌问诊数据进行聚类分析，合并各别具有相似性的聚类结果，最终确定 7 大类老年人对肺癌的健康信息需求，见表 3。采用同样的方法对 921 条高血压问诊数据聚类分析，最终确定 5 大类老年人对高血压的健康信息需求，结果见表 4。

表 3 在线健康社区老年人对肺癌疾病信息需求的主题特征

主题	子类目	关键词	数目	占比（%）
新冠时期肺癌疾病管理与风险把控		医院,方案,控制,新冠,发烧	445	17.53%
肺癌的检查与诊断		咳嗽,方案,检查,症状,医院,住院,建议,做手术,基因,检测,突变,化疗,靶向,多发,双肺,结节,检查, 病变	569	22.42%

肺癌的病理学特征		转移,淋巴,腺癌,基因,方案 右肺,淋巴结,结节,多发,纵隔,诊断,细胞,检查,穿刺,病理,复查,手术,玻璃,报告,病理,解读,术后,浸润性	604	23.80%
肺癌的治疗	药物治疗及用药副作用	建议,用药,提供,不适,分析 放疗,化疗,周期,细胞,方案,手术,免疫治疗,靶向,检测,基因,术后,服用,呕吐,替尼,头晕,心脏	665	26.20%
	手术	手术,检查,切除,医院,左肺		
肺癌术后管理		复查,术后,手术,报告,医院	114	4.49%
肺癌疼痛管理		疼痛,控制,感觉,难忍,手术	97	3.82%
健康生活方式及预防		睡眠,饮食,体检,咳嗽,大小便	44	1.73%

表 4 在线健康社区老年人对高血压疾病信息需求的主题特征

主题	子类目	关键词	数目	占比（%）
高血压的诊断		检查,血压高,舒服,心脏,体检,升高,就诊,诊疗,治疗,舒张压	347	37.68%
高血压的治疗	用药及药物调整	稳定,一粒,拜新同,增加,剂量,缬沙坦,高压,降到,氢氯,噻嗪	255	27.69%
	药物治疗效果及副作用	调整,药物,偏高,毫克,心率 心跳,咳嗽,控制,治疗,头晕		
高血压的症状		高压,低压,头晕,治疗,症状, 血压高,头疼,恶心	131	14.22%
高血压的预防管理		治疗,测血压,睡眠,饮食,控制	110	11.94%
高血压的监测与控制		控制,不住,血压高,治疗,升高	78	8.47%

2.6 健康信息需求差异化分析结果

不同性别肺癌患者的健康信息需求分布如图 2 所示。不同年龄段肺癌患者的健康信息需求分布

如图 3 所示。

不同性别高血压患者的健康信息需求分布如图 4 所示。不同年龄段高血压患者的健康信息需求分布如图 5 所示。

老年人肺癌的问诊数据中，不同性别与健康信息需求主题的卡方检验结果（χ^2=18.0，P=0.006）、不同年龄与健康信息需求主题的卡方检验结果（χ^2=254.54，P=0.002），均表明不同性别、年龄的肺癌患者对健康信息的需求存在显著差异；高血压问诊数据中，不同性别与健康信息需求主题的卡方检验结果（χ^2=3.53，P=0.473）、不同年龄与健康信息需求主题的卡方检验结果（χ^2=140.14，P=0.297），这表明不同性别、年龄的高血压患者对健康信息的需求没有显著差异。

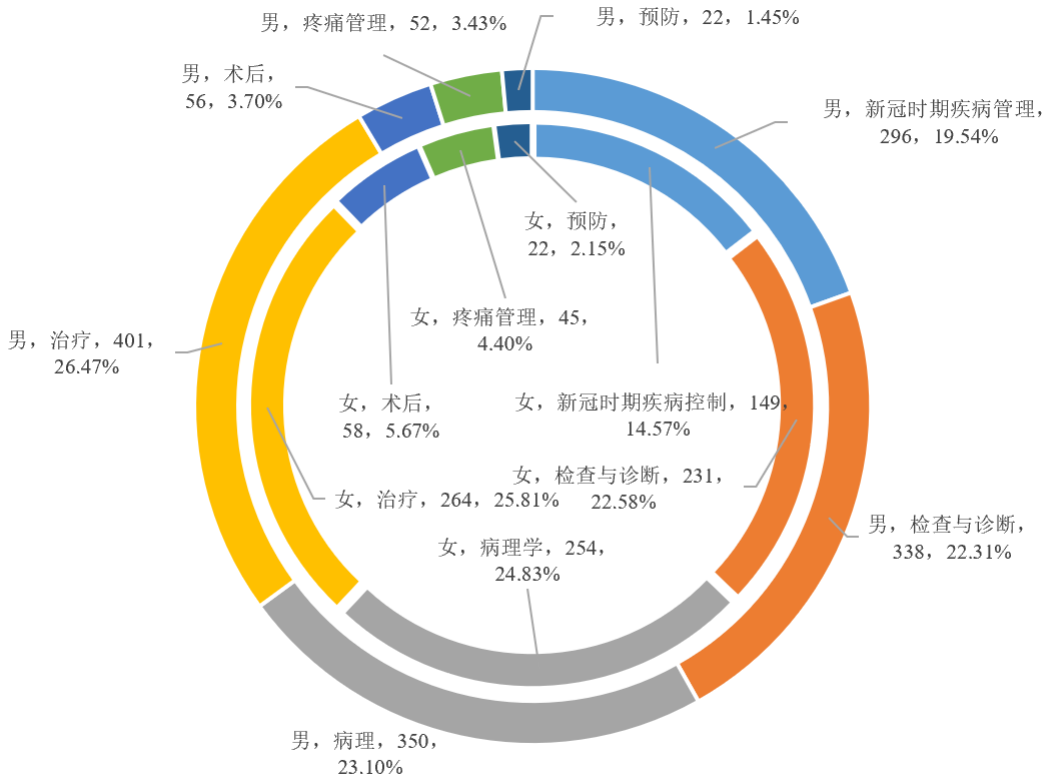


图 2 不同性别肺癌疾病信息需求分布

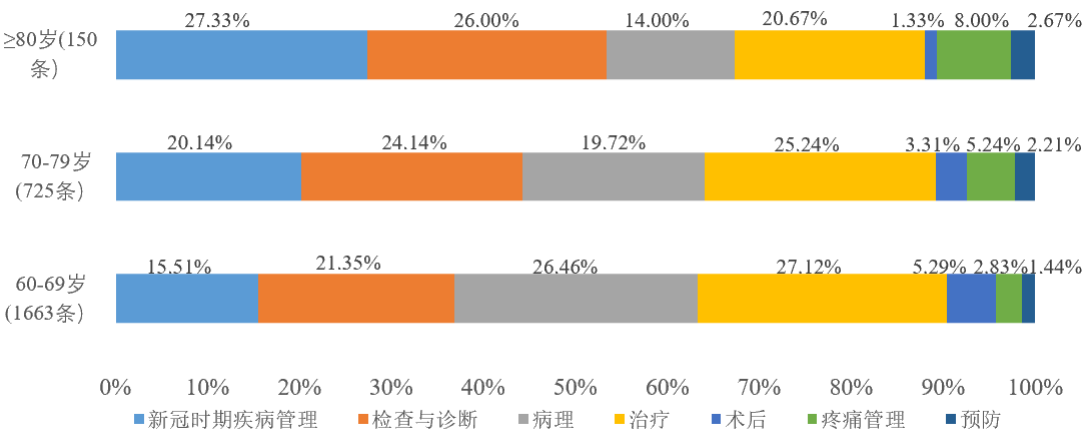


图 3 不同年龄肺癌疾病信息需求分布

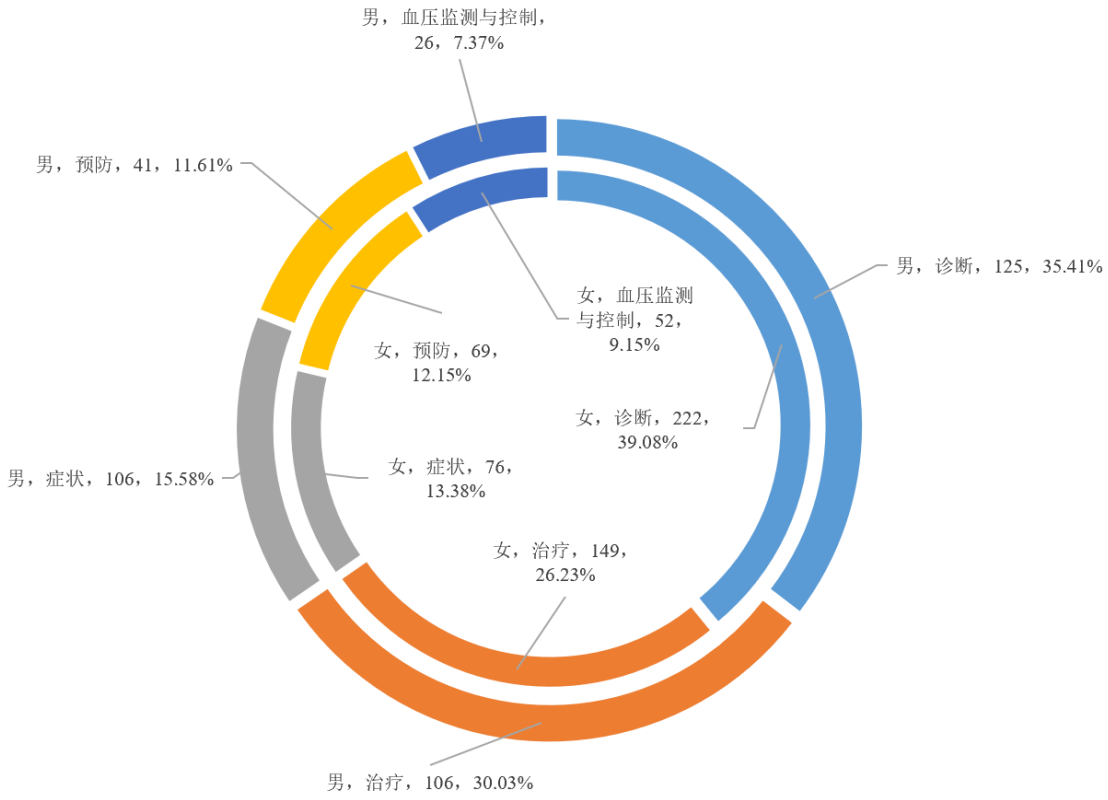


图 4 不同性别高血压疾病信息需求分布

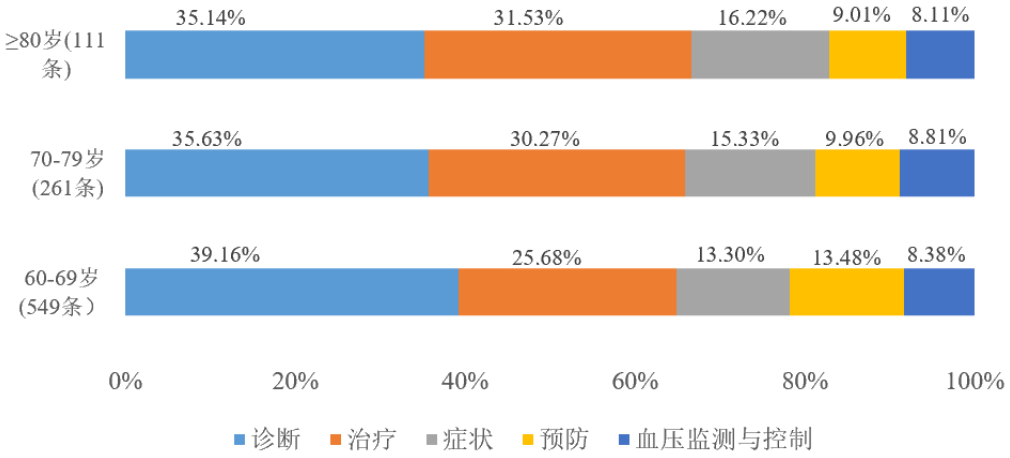


图 5 不同年龄高血压疾病信息需求分布

3 讨论

在线健康社区平台的用户数据能够反映当前老年人所患疾病的流行病学特征。根据疾病问诊分类结果来看，老年人对肿瘤、循环系统疾病、呼吸系统疾病及消化系统疾病的问诊量较高，健康咨询较为频繁，说明老年人存在该方面的疾病问题，这与既往探究医疗机构病例疾病构成情况相似[11,12]。通过患者的问诊情况来看，女性问诊量高于男性，由于受到生理、心理及社会因素的影响女性的健康意识往往高于男性，女性更倾向于关注自身的健康变化[13]。从年龄角度分析，老年人的平均年龄为（69.36±7.73）岁，年轻老年人的问诊量高于其他年龄段老年人的问诊量，一方面可能是因为数字鸿沟的问题[14]，年轻老年人接触互联网的机会，接受新鲜事物的能力要高于高龄老年人，在数字技术方面代沟较小，另一方面，年轻老年人刚刚步入老年阶段，面临更多老年相关的生理和健康挑战，而且年轻老年人健康意识较高[15]，所以更关注自身健康问题。此外，老年人对健康信息

的需求主要集中在内科、外科、肿瘤科、五官科和皮肤性病科，由于增龄和老龄化的迅速发展的缘故，出现视力障碍、听力下降的老年人骤增[16,17]，导致老年人对五官的健康信息需求激增。同样，老年人的皮肤屏障保护作用减弱，更容易出现湿疹、红斑等疾病，而且在问诊文本中发现部分老年人存在性健康方面的信息需求，相关研究表明性健康问题在老年人群中很常见，但是他们很少咨询医生[18]，因此在传统思想的影响下，由于羞涩老年人会选择通过在线咨询性功能障碍方面问题，他们希望从互联网上得到有用的信息。通过分析老年人的问诊记录发现，在线健康社区平台上存储的老年问诊数据可以反映老年群体所关注的健康问题，在一定程度上能够反映当前主流疾病的流行病学特征，而且从在线健康平台上通过数据挖掘的手段获取的数据内容客观、真实，对社区工作者、医务人员具有一定的参考和使用价值。

在线健康社区为老年人的健康教育提供信息支持。在本研究中，肺癌、高血压疾病分别是肿瘤和循环系统疾病中问诊次数最多的疾病，基于这两个特定疾病我们从中探讨老年人对疾病健康信息需求的共性和特性。对于肺癌患者，健康信息需求主要包括治疗、病理学特征、检查与诊断、新冠时期疾病的管理与风险把控、术后、疼痛管理及预防 7 个主题类目，其中治疗、病理学特征及检查与诊断的问题是他们迫切性需要了解的，早期肺癌难以诊断，导致病情延误，影响生存质量，快速的诊断和精准的治疗对与改善肺癌患者的预后，延长生命周期至关重要[19]，病理学特征有助于肺癌的确诊和分期，有助于监测治疗反应，能够及时评估治疗的有效性并调整治疗方案；而肺癌的预防信息需求在整体的信息需求中占比较小，这与之前国内学者的研究结果一致[20]。健康信息需求受到特定环境的影响，新冠肺炎流行时期给肺癌患者带来了许多额外的挑战和冲击，回溯问诊文本，新冠肺炎的流行给肺癌患者的就诊、随访、治疗计划调整、疫苗接种和心理健康带来直接或间接的影响，肺癌患者对诊断延迟、治疗延误、免疫力下降的问题更加关注，促使更多的患者利用互联网寻求健康信息支持[21]。术后管理有利于提高和恢复患者的生活质量，术后的肺癌患者仍可能发生复发或远处转移，患者往往需要在术后进行临床检查、实验室检查以随访和监测病情，在线健康社区平台能够为患者供术后病情监测及检查报告结果解读的信息服务，以求降低肺癌的复发率。癌痛是肺癌患者的常见并发症之一，可能发生在患病的各个阶段，肺癌患者的需求在癌痛时期的信息需求集中在如何控制疼痛。而高血压患者的健康信息需求与肺癌患者相比，高血压的症状识别以及高血压患者病程中血压的监测与控制是该疾病管理的两个关键特点，高血压被称为“沉默的杀手”，早期高血压患者出现头晕、头疼的症状，这些症状也极易与其他疾病混淆不易识别，此时需要提高警惕，而且高血压长期存在，暂时没有根治的办法，所以血压的监测与控制显得对整体健康非常重要。由以上分析可以发现，老年人对慢性病的信息需求存在相似之处，对于疾病的治疗、诊断、预防和症状是患者希望对某一疾病的基本了解，但是针对不同的疾病或者在不同的社会背景环境下，对疾病的信息需求又有专一性，所以，医务工作者在进行健康教育的宣传工作时，一是了解疾病特征，二是了解受众目标及其最关心的信息，以患者的信息需求为导向，多渠道提供持续的健康教育和支持。此外，在遇到紧急情况（疫情或大流行）时，需加强医院和社区的合作力度，二者的紧密联系才能及时为患者提供准确的信息，以平衡医疗资源应对突发状况。

基于在线健康社区老年人健康信息需求的差异性，制定个性化的健康教育和宣传策略。在明确老年人存在的健康信息需求后，通过划分年龄和性别进一步探讨老年人存在的信息需求差异，有助于开展针对性的宣传教育。对于肺癌患者，整体来看男性问诊量多于女性问诊量，比例为 1.48（1515）：1（1023）。据报告显示[22]，2020 年全球新发肺癌中男女的比例为 1.75（140 万）：1（80 万），男性肺癌的患病率高于女性，但是近年来，男性和女性在肺癌发病率的趋势上有所转变，男性的肺癌发病率稳步下降，而女性肺癌肺癌发病率正在上升[23]。本研究结果显示，男性对肺癌在新冠时期疾病管理与风险把控和治疗方面的信息需求略高于女性，而女性对其他方面的健康信息需求均高于男性，这表明男性更关注疾病的风险因素和治疗方面，而女性更倾向于肺癌的早期诊断、筛查和预防措施，更注重健康的维护。另外，60-69 岁和 70-79 岁年龄段老年人较为关注肺癌的治疗，80 岁以上老年人则更为关注新冠时期疾病管理与风险把控和肺癌的检查与诊断。且不同性别、不同年龄的肺癌患者在健康信息需求方面存在统计学差异，因此健康教育和宣传工作需要凸显出个性化、多样化、精准化。对于高血压患者，男性女性问诊量方面与肺癌问诊量相反，高血压的女性问诊量高于男性问诊量，比例为 1.61（568）：1（353）。与男性相比，女性更关注疾病的诊断，国外的一项研究表明，在高血压患者中仅 48%的高血压患者能明确自己的诊断，而且女性比男性更能意识到自身的患病诊断情况[24]，与女性相比，男性比女性更关注高血压的治疗，由此可见，女性更倾向于高血压的早期诊断，男性更倾向于寻找解决问题的办法，如：治疗。另外，从年龄上看，60-69 岁、70-79 岁、80 岁以上老年人对诊断的信息需求都较为关注，但是呈现略微下降

的趋势，这与唐晓波等[25]对在线健康平台上高血压患者健康信息需求主题演化分析结果一致，各个年龄阶段的患者对高血压的治疗需求一致较高。本研究结果中，不同性别、不同年龄的高血压老年人对健康信息的需求无差异，这表明老年人普遍全面关注高血压，但是在对不同性别、不同年龄的高血压患者进行宣传教育时，仍需突出信息需求的侧重点。综上，深入了解不同性别、年龄老年人的健康信息需求差异有利于改善健康信息传播和健康管理的方式，有助于优化医疗资源分配和医疗服务供给，从而提高老年人的生活质量，促进社会的整体健康。

4 结论

本研究以肺癌、高血压为例采用主题分析和统计分析相结合的研究方法，探究了老年患者的健康信息需求分布及差异性。基于肺癌和高血压这两个代表性疾病我们发现，不同疾病类型其信息需求不同，其中疾病的诊断、治疗、预防是两类疾病健康信息需求的共性。与高血压患者的信息需求相比，新冠时期肺癌的疾病管理与风险把控、病理学特征、术后管理及癌痛管理是肺癌患者信息需求的特性；与肺癌患者的信息需求相比，高血压的症状和监测与控制是高血压患者信息需求的特性。男性患者在患病的过程中更倾向于寻找解决问题的办法，如：治疗；女性患者在患病过程中更倾向于健康的维护，如：诊断、筛查和预防。研究结果有助于相关部门了解老年人健康信息需求的现况，为社区工作者、医务人员开展个性化、多样化、精准化的健康教育工作提供借鉴依据，为优化在线健康信息平台的服务提供思路。本研究存在一定的局限性，首先，数据来源仅采集了“好大夫在线”的健康数据，在今后的研究中应扩大数据来源平台，以确保研究结果更加全面；其次，本研究中仅针对肺癌、高血压进行了聚类分析，今后的研究中应从多个疾病角度出发，探究老年人健康

PU-235

县域三甲医院的数据治理实践

田梅 侯彦东
山西省汾阳医院信息中心

1 项目背景及研究意义

智慧医院的数据治理，参照医疗行业数据治理的基础标准，是将医院业务系统产生的数据，利用先进的数据处理和加工技术，通过多种采集方式，汇聚到统一的平台上进行数据清洗和加工，最终形成有价值、可被利用的数据资产[5]。同时在数据安全体系保障下，提供数据权限控制、数据导出脱敏和访问加密、以及数据泄露追溯等核心功能，保障数据安全[1]。

在实际应用中各个医院不管对智慧医院还是数据治理都有着不同的探索，有的建大数据平台，有的建科研大数据挖掘，还有的建 BI 数据统计。我院关于数据治理也搭建了一系列的框架体系，下面就我院基于医院数据治理与服务的实际业务需求，初步建立的数据治理体系，进行简单的介绍，其主要目标有以下几点：

一.数据服务：对各个业务系统的数据进行统一管控，保证数据质量，提供数据共享机制，实现数据上报。

二.数据展示：采用各种可视化展示方式，提供直观、美观、互动的 UI 界面。

三.权限控制：通过统一授权方式实现对数据访问权限的有效控制。

四.决策支持：对业务数据采集建设数据仓库，对数据进行挖掘，为业务及管理决策提供支持。

五.数据安全：提供数据的加密、脱敏、认证、泄密追踪等数据安全服务。

2 项目整体架构

主数据层通过统一的数据标准，结合制度约束、系统控制等手段，提高数据的规范性、开放性和共享性；数据采集层建立管理异构数据库的连接信息，调用中间件按照设定的规则进行数据集成工作，保证数据的共享；建立一套标准的数据转化、清洗、整合规则，并内置数据质量、一致性、增量核查的功能，保证数据的规整性；安全服务应用层采用消息摘要、数字签名防止数据篡改和进行身份认证；数据应用层采用学习神经网络算法等先进的 AI 技术针对不同角色提供可用、易用的数据展示。

整体架构图

式

3.1 术语主数据平台

通过术语主平台进行数据标准的定制和发布、元数据管理、主数据管理、数据模型管理，并可以实现数据资产展示、数据资产检索、数据资产血缘分析、数据资产追溯等，提供数据的规范性、开放性和共享管理，提高数据管理水平[1]。为后期数据治理标准化提供依据。

3.2 数据集成平台

采用 ESB 模式建立 SOA 松耦合架构，由数据总线、接口适配器、HL7 标准引擎、消息中间件、主数据管理、主索引、管理监控平台、运营和医疗管理数据中心八大模块实现医疗业务数据的采集、共享及互联互通，解决了医院信息集成度低，缺乏统一的信息标准及有效的整合机制等问题。同时采用数据加密、控制访问权限、日志记录等方法保证了数据的可靠性、安全性。

3.3 订阅式发布系统

基于 SOA 的消息发布订阅模块收集异构、不同平台的医疗信息系统的消息，为用户提供可用、易用的消息订阅界面，并对订阅内容进行统一管理；通过订阅匹配模块行进消息和订阅的匹配，快速查询订阅者；消息传输模块提供可靠、稳定的消息传输服务，将消息推送到用户终端。

3.4 CDR 临床数据中心

是一个整合多个来源的临床数据仓库，以医患为中心的统一视图数据库[5]。通过 CDR 可以提供灵活的结构化、标准化检索服务并与业务系统实现无缝集成，满足多角度、多方式的查询、浏览、调用需求。

3.5 CDSS 临床辅助决策

基于深度学习神经网络算法等前沿技术自我完善知识库、规则库以及决策引擎模型，利用医院大数据、知识图谱、以互联网主流的信息推荐技术，实现精准、高效的智能综合分析与判断，为医生从海量对象中迅速有效的筛选出符合自己偏好的对象，为诊疗过程的各个环节提供实时提醒和精准推送，提升医生诊疗服务能力、规范诊疗行为、减少误诊漏诊[2]。

3.6 BI 商业智能

通过对数据仓库中用来存储经过处理加工过的面向决策支持主题的数据集合，应用 BI 内置的统计分析算法和人工智能算法实现数据的挖掘及全景展示[3]。解决了传统的报表分析口径和方法不够灵活，无法进行多维度和深层次的分析的不足。为管理提供一站式决策支持。

3.7 CA 电子签名认证

主要由基础设施层、安全服务层、安全应用层三个部分组成。基础设施层实现证书的签发、管理、证书源目录、密钥生产与存储、司法取证服务等；安全服务层主要提供应用系统运行所需要的 PKI 安全中间件和证书应用服务器，向上为应用系统提供开发接口，向下提供统一的密码算法及设备驱动；安全应用层主要对医院内部的业务系统进行 CA 认证服务集成接入[4]。不管是响应国家政策还是医疗机构解决自身的安全问题，都有非常好的应用效果。

4 结论

数据治理的核心，也即医院数据资源建设的重点，是持续提升数据质量。有质量的数据才具有价值。数据的价值在于分析利用，所有的数据资源管理、技术平台，最终都是为了数据的分析利用，数据利用的形式可以是建立基于数据的常规化运行的应用系统，也可以是“即席式”的数据分析，即针对特定的临床和管理问题，挖掘数据规律、建立分析模型。对于数据利用，要制定吸引医务人员参与的措施，建立便于数据开发利用的技术平台和服务流程，提高医务人员的数据利用能力，构建一种鼓励数据资源开发利用的环境氛围。

5 总结

从智慧医院的概念来讲，我院的信息化建设的路还很长很长，数据治理方面也显得比较凌乱还有许多需要改进的地方，比如说数据安全管理还比较薄弱，数据处理的流程监控也不够完善，但作为一个县域医院我们的信息化建设一直都在不断学习中前进，在不断探索中完善，且必须宏观符合国家政策，微观符合实际使用需求，是一项功在当代利在千秋的系统工程,也是我们医院信息人肩负的长期使命和责任。

PU-236

在线问诊平台用户情感影响因素及知识图谱构建研究

杨 爽 王 琰 王雅晴 刘曙霞 安新颖
中国医学科学院医学信息研究所

1 引言

随着互联网时代的到来，手机等移动设备的急速发展，促进了传统医疗的数字化、移动化，令互联网+医疗的模式越来越普遍，尤其是在新冠疫情爆发后，用户对线上医疗的需求急剧增加。在线医疗平台的应用在有效缓解线下医院压力的同时，又能减少交叉感染的风险，拥有很好的发展前景。

本文通过爬取在线医疗平台在线问诊服务的用户评论，采用主题分析、情感细粒度分析深入挖掘用户满意度，发现用户在线上问诊过程中关注的重点，结合当下现状提出优化在线医疗平台在线问诊服务的对策建议，以进一步提升全社会的健康水平。

2 数据及方法

2.1 数据来源及预处理

本文基于好大夫在线医疗平台在线问诊模块，利用 python 爬取精神心理科医生的线上服务评论内容，共采集到 8523 条数据。由于数据来源于社交媒体，评论内容存在格式、拼写不规范的问题，因此数据清洗的质量很大程度上决定了分析结果的可靠性。数据清洗的主要步骤包括删除含有错别字的评论、删除空值，最终纳入评论 8413 条。

2.2 基于 BERTOPIC 主题挖掘

BERTopic 是一种基于 BERT 嵌入表示和聚类算法的主题建模技术，首先使用基于 sentence-transformer 的词向量模型进行词嵌入，在此基础上用 UMAP（Uniform Manifold Approximaiton and Projection for Dimension Reductiong）和 HDBSCAN（Hierarchical Density - Based Spatial Clustering of Applications with Noise）算法对词向量降维和聚类，最后使用 c-TF-IDF 算法对主题中的词汇进行重要性测量。

由于 BERT 在多项自然语言处理任务中展现出了优异效果，BERTopic 利用 BERT 生成的嵌入向量代表文本句子，使得模型在识别上下文语义方面更加灵活和准确。相较于传统主题建模技术（如 LDA），BERTopic 可以通过聚类算法自动确定最佳主题数目，减少了人为干预，增加了模型的可靠性。同时，BERTopic 对于随机初始化敏感性较低，能够产生稳定的结果并能够很好地抵抗噪声和异常值的影响。除此之外，BERTopic 在进行短文本主题识别具有较好的效果，因此本文使用 BERTopic 主题模型进行主题识别，构建方面词典。

2.3 情感分析

本文基于 snowNLP 库训练在线问诊用户评论情感分析模型，将随机爬取到的好大夫在线问诊

平台 100 位医生在线问诊评论数据按照训练集和测试集为 7：3 的比例训练模型。采用准确率指标判定模型训练质量，准确率是指模型正确分类的样本数占总样本数的比例。本研究将用户情感划分为积极、消极两种极性，即情感分数>0.5 判定为积极，情感分数=0.5 判定为中立，情感分数<0.5 判定为消极。

2.4 情感三元组识别

本文基于句子的依存关系抽取情感三元组，即（方面实体，情感词，情感极性）。情感三元组抽取过程可分为 2 步：第一步，抽取（情感词，情感极性）二元组。若在遍历语料的过程中出现自定义情感词典中的词汇，则触发情感词提取程序，将对应情感词和情感极性保存为二元组。第二步，抽取方面实体，组成（方面实体，情感词，情感极性）三元组。触发情感词后，借助 Spacy 库遍历以该情感词为节点的依赖语法子树上的所有节点，筛选其中词性为“NOUN”的词与构建的方面词典进行匹配，确定方面实体，组成情感三元组。

2.5 知识图谱构建与查询

知识图谱能够直观描述“实体、属性、关系”，目前在金融、电商、医药等领域得到了广泛应用。而 Neo4j 因其轻量级、高性能的优势受到了众多关注，因此本文使用 Neo4j 数据库进行知识图谱构建。首先构造节点表、关系表，通过 csv 格式导出，节点表和关系表如表 1、表 2 所示。最后通过 Cypher LOAD CSV 命令将所有 CSV 文件导入 Neo4j 数据库中。

表 1 节点表

字段名	类型	说明
id	int(10)	主键
node__name	varchar(45)	节点名
node__category	varchar(45)	节点类型

表 2 关系表

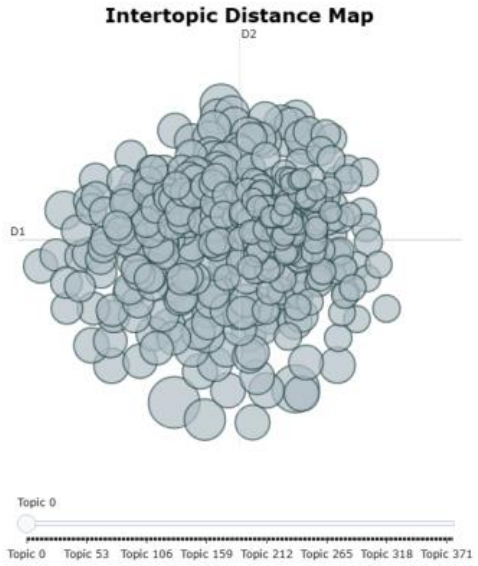
字段名	类型	说明
id	int(10)	主键
id__start	varchar(45)	头节点 id
node1__name	int(10)	头节点名
id__end	varchar(45)	尾节点 id
node2__name	int(10)	尾节点名
relationships	varchar(45)	两节点间关系

3 结果

3.1 基于 BERTopic 主题挖掘结果

为了避免混淆词对提取结果的影响，因此在进行主题提取前对纳入评论数据进行了二次清洗，去除人名、表情符号等影响提取结果的停用词，对字符数小于等于 2 的词汇予以删除。数据清洗后，利用 BERTopic 进行主题建模，本文将“nr_topics”设置为“auto”，使模型能够自动生成主题，共

得到 372 个主题。随后，对主题关系进行可视化，发现有较多主题仍存在重叠现象。因此需要人工调整主题数量，对重叠主题进行合并，消除离群文档，最终提取得到 7 个主题。再次对提取后的模型进行主题可视化，发现各个主题在主题距离图上有明显的距离，说明聚类效果较好。主题数量调整前后的距离图如图 1 所示。经过主题聚类，共得到 7 个主题，分别为体验感、疗效、态度、医德、医术、回复速度、服务时长。



（1）主题调整前

（2）主题调整后

图 1 BERTopic 主题距离图

3.2 情感倾向识别

利用预训练文本对模型进行训练，模型准确率达到 98.19%，训练效果较好。采用训练好的模型进行情感分析，发现有 8055 条评论为积极（95.74%），358 条评论为消极（4.26%），无持中立态度的评论。可见，患者对于该问诊平台的在线问诊服务的评价多为积极态度，说明该平台的在线问诊服务质量普遍较高。但是也存在医生在诊疗过程中安抚了患者的心态，患者出于对医生的感谢，给予好评。部分评论数据情感计算结果如表 3 所示。

表 3 评论情感分类

序号	评论	情感分数	分类
1	医生很专业，问诊很详细，态度和蔼亲切，分析问题细致，对这次问诊非常满意	0.999944482	积极
2	医术精湛，调药精准，医德高尚	0.995716982	积极
3	石医生很专业，几句话就解答疑惑了，非常有经验，很专业。非常好，态度也很好，非常有耐心。	1	积极
4	是一位医德高尚的医生，值得信赖！	0.989169462	积极
5	状态慢慢好起来了，感恩！！！	0.997310825	积极
6	没有在规定时间内开药	0.067695634	消极

7	无任何病情沟通及药物建议	0.00992128	消极
8	完全没回答问题	0.19297951	消极
9	非常不好的感受	0.02588183	消极
10	感觉比较糊弄	0.02478403	消极

3.3 满意度影响因素分析

在对评论语句进行情感分析的基础上，从方面级别对评论内容进行了情感三元组识别，如表 4 所示，以期能够从整体角度更加直观地观察到患者对在线问诊服务各方面的情感极性。分别统计了积极、消极情感词修饰的不同方面实体出现频次，从而探析影响患者在线问诊情感倾向的因素。

对于积极情感倾向的方面实体，可以划分为体验感、态度、医术、医德、回复速度、服务时长、疗效 7 类。如“感受”“感觉”“体验”等与患者体验感有关的评价词在积极情感中出现次数较多，出现次数占比达到 85%以上，说明评论者大多对于就诊体验表达了积极的情感倾向。而与医生诊疗态度等相关的评价词出现次数排在第二梯队，说明医生对待患者是否有耐心、是否仔细也是影响患者情感倾向的重要因素。其次，“专业”“调药”“诊断”等与医生医术有关的评价词出现频次仅次于第二梯队，表明医生的专业水平能够在一定程度上影响患者的情感。此外，诸如医生的医德、沟通能力、回复速度、服务时长和服务疗效等内容也是积极情感倾向情感词评价的主要方面。

对于消极情感倾向的方面实体，可以划分为体验感、疗效、态度、医德、回复速度、沟通能力、服务时长、费用。患者的体验感是影响其消极情绪的主要因素，其次是医生诊疗产生的效果，以及在诊疗过程中医生对待患者的态度。诸如“感受”“心情”“体验”等方面出现频次较高，说明体验感能够直接影响患者的情绪；而诸如“症状”“问题”“帮助”等和疗效有关的方面词出现仅次于体验感，说明诊疗效果能够明显影响患者的心情；诸如“回复”“回答”“敷衍”“简单”等代表医生态度的关键词出现频次较高，说明医生对患者的态度一定程度上导致了患者的消极情绪。而医生医德、回复速度、沟通能力、服务时长、费用等方面也是容易导致评论者消极情绪的重要原因。

表 4 情感三元组（部分）

序号	评论内容	方面	观点词	情感极性
1	很耐心回答我所有问题。	回答	耐心	积极
2	医术精湛，调药精准，医德高尚	医术	精湛	积极
		调药	精准	积极
		医德	高尚	积极
3	感谢阳主任耐心专业的回复	回复	耐心	积极
		回复	专业	积极
4	非常感谢钟主任，替病患着想，解答清晰明了，消除了我的疑惑，很专业	解答	清晰	积极
5	利用休息时间回答很及时，专业的指导，	回	及时	积极

	感恩困境中给予的帮助	答导	指专业	积极
	非常不好的感受	感受	不好	消极
6	回复过于简单敷衍，这次只回复让复诊，再问没回复	回复	敷衍	消极
7	就是我的症状还是没有解决	症状	没有解决	消极
8	48 小时回复那么几条，最后还剩下 14 小时提问，直接拖到最后 48 小时告知问诊时间到了。期间每条不超过 10 个字。专家费真太好赚了。在平台最差的一次问诊体验。	体验	差	消极
9		专家费	好赚	消极
10	回答太过粗糙	回答	粗糙	消极

3.4 在线健康平台知识图谱呈现

将数据导入 Neo4j 数据库中，得到共计 332 个节点数据和 2988 个关系数据。借助 Neo4j 浏览器端可以查看知识图谱的效果，如图 2 所示。图中不同实体类别采用不同颜色加以区分，其中紫色节点表示医生，橙色节点表示影响患者情感的方面，关系包括积极情绪和消极情绪两种。例如若医生在某一方面的患者评论情感包含消极情绪，则两个不同节点之间为消极关系。从知识图谱可知，不同医生在各方面的患者情感倾向，有助于患者选择合适的医生就诊。同时 Neo4j 提供了查询语言，通过编写查询语句能够了解每个医生在不同方面的服务情况，将医生的情况推送给患者。

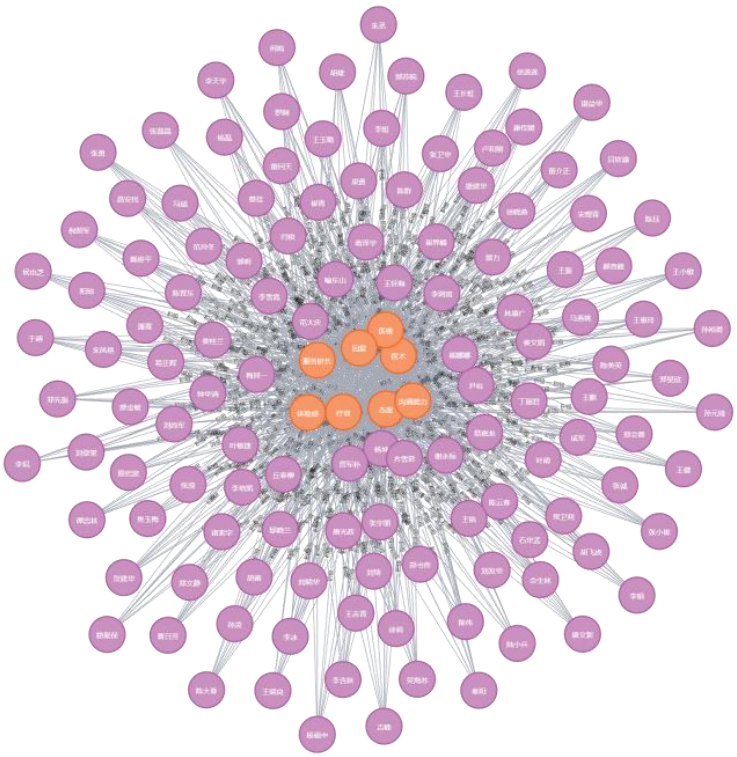


图 2 在线问诊情感知识图谱

4 结论

在线诊疗作为一种新型的医疗服务形式，为患者提供了便利的医疗咨询和在线诊疗，越来越多的用户开始借助在线医疗平台开展线上问诊服务。随着在线问诊平台的快速普及，有必要了解患者对于在线服务的情感倾向和影响因素，以进一步提高在线问诊服务质量和用户体验。因此，本研究基于方面级情感分析，对在线诊疗平台的用户情感倾向及其影响因素进行了分析，同时构建知识图谱有助于直观观察到医生在不同方面的服务质量。

本研究基于医疗在线问诊平台的用户评论数据构建了一个情感分析模型，发现大多数患者评论态度为积极，说明当前在线问诊平台的服务质量较好，但是也存在患者处于对医生的尊重和感谢给予好评的情况。同时，运用 BERTopic 模型对评论内容进行主题提取，共抽取得到 7 个主题。随后，对评论数据进行情感三元组提取，发现体验感、态度、医术、医德、回复速度、服务时长、疗效、费用、沟通能力等 9 个方面是影响患者情感倾向的重要原因。而体验感、态度、医德、回复速度、服务时长、疗效是影响患者积极、消极情绪的共同因素。因此，未来医生可以通过提升服务态度、提高回复速度、注重服务效果等方面改善患者满意度，同时平台可以通过增加服务时长，完善医生考核制度，提高其医德医风的途径改善在线问诊用户满意度。

本研究的创新点包括采用方面级情感分析方法总结了影响医生在线服务质量的影响因素，同时构建了在线问诊情感知识图谱，能够直接了解到不同医生在各方面的患者情感倾向，有助于患者选择合适的医生就诊。本研究尚存在一定的局限性。在研究数据的选取上，数据量较少且不具有代表性。因此，未来可以扩展数据来源，采集多个在线问诊平台的评论数据。在方面级情感分析上，未考虑评论中的隐含信息。仅从句子中明显的依存关系进行方面提取，忽略了句子中表达情感但是未出现的方面词。因此，未来可以在研究方法上做进一步的完善。

PU-237

大数据技术在中医药传承与创新中的应用

黄尤江[1] 郑利荣[1] 贺 莲[2]
1 广东省医学学术交流中心（省医学情报研究所）
2 广东省卫生健康宣传教育中心

中医药(TCM)，是包括汉族和少数民族医药在内的我国各民族医药的统称，反映了中华民族对生命、健康和疾病的认识，具有悠久历史传统和独特理论及技术方法的医药学体系。“中医药学是中国古代科学的瑰宝，也是打开中华文明宝库的钥匙。”习近平总书记高度重视中医药工作，要求把中医药工作摆在更加突出的位置。随着人工智能、大数据采集和分析等技术的发展与应用，对中医药的发展带来了机遇。加强中医药健康医疗大数据开发与应用，对于加强中医药传承、突显中医诊疗规律、推动中医药科技创新、优化健康产业布局、助推中医药健康产业发展和健康中国建设等具有重要作用和社会价值[错误!未找到引用源。]。

1 中医药学术与思想传承

1.1 研究方法

基于文献计量学分析中医传承当前研究概况，通过构建检索式“(中医 or 中药) and 传承”在CNKI中进行主题检索，从主题角度进行可视化分析，进而提出建议和想法。

1.2 结果分析

根据检索结果显示，共检索出 8789 篇相关文献，主要分布在 2004 年以后，主要主题分布如图 1。从图 1 可以看出，目前中医传承研究主要集中在用药规律研究（组方规律）、名老中医经验（学术思想）传承，对于传承创新相关文献虽然占有一定比例，但主要从管理和教育角度进行了相关研究，对于技术研发和开发性研究相对较少。目前研究中医药传承的主要方法为中医传承辅助平台（系统）和数据挖掘技术。

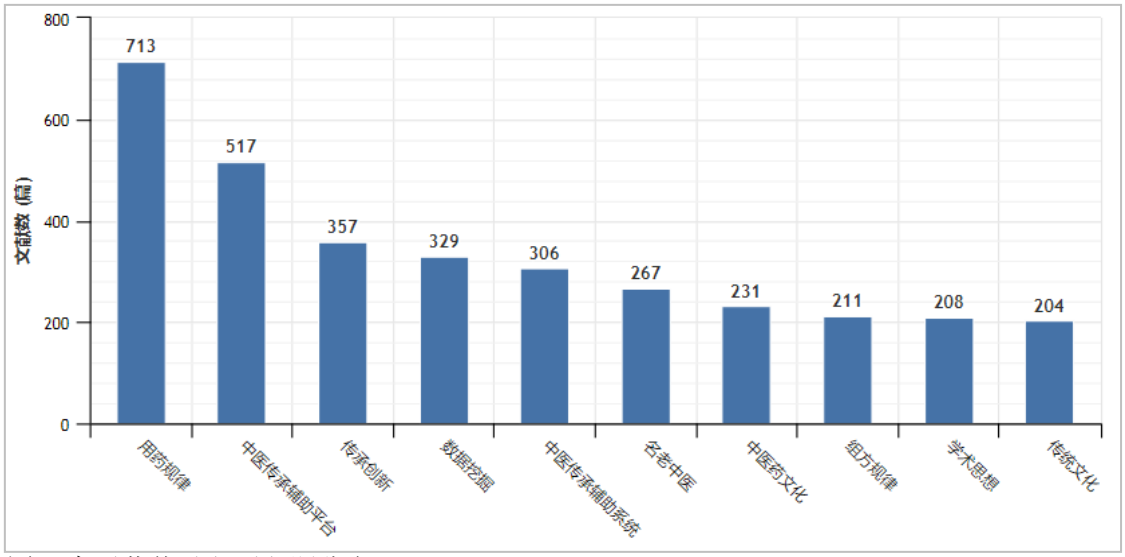


图 1 中医药传承主要主题分布

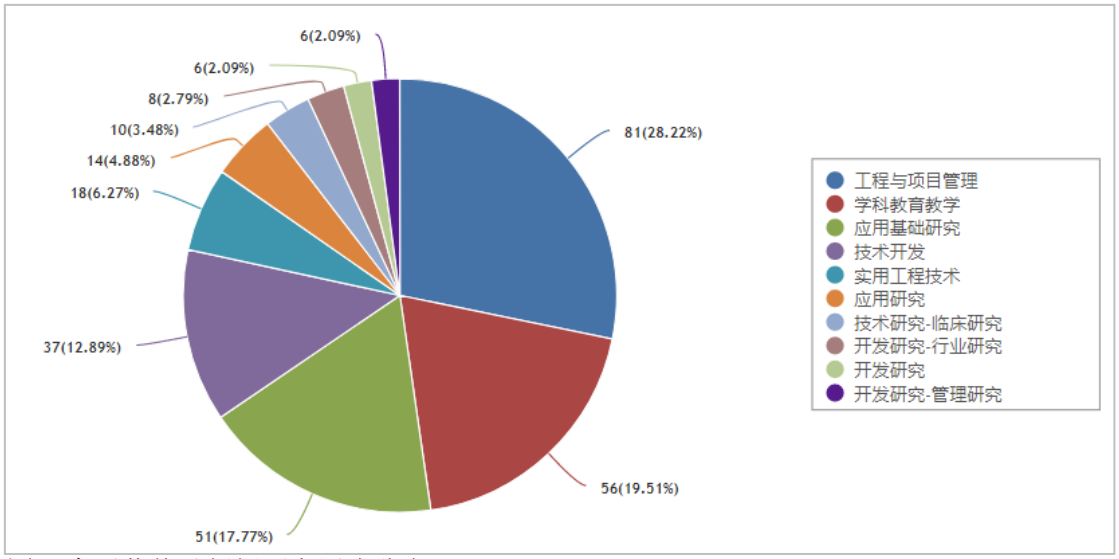


图 2 中医药传承创新研究层次分布

1.3 讨论与建议

目前对于中医药传承创新研究主要依赖于中医传承辅助平台（系统），通过数据挖掘技术探究用药规律或组方规律，另一方面则是通过对名老中医学术思想加以总结。中医传承辅助平台(TCMISS)目前主要集中于名老中医经验总结、用药规律研究、医师个体经验研究和组方配伍分析四个方面[错误!未找到引用源。]，在中医推拿及针灸处方选穴规律方面的研究相对较少。研究内容多局限于思想传承与总结，有待创新发展。对于名老中医经验和学术思想的总结，主要以论文产出为研究成果，缺乏系统的知识仓库和数据仓库建设。因此有必要构建中医药传承数据仓或知识库用于专题研究，以适应大数据时代发展。

2 中药网络药理学及小分子药物研究

网络药理学近年来已成为中医药领域的热门研究工具，其系统性和整体性特点为复杂中药体系的研究提供了新思路，被广泛用于筛选中药有效成分、阐释作用机制和揭示方剂配伍原理等[错误!未找到引用源。]，有力促进中药现代化进展，如图 3 所示。网络药理学聚焦所需要研究的蛋白、基因和通路，为中药活性成分和小分子药物研究提供了有力手段。

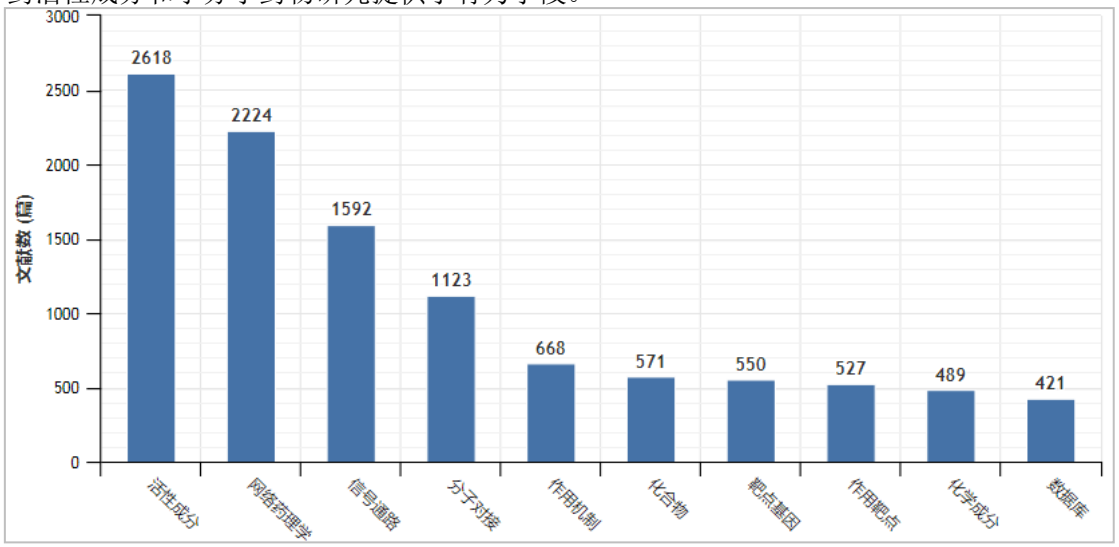


图 3 网络药理学研究热点分布

新药研发流程繁琐，需要经历寻找靶点、设计药物、生物实验、临床试验和审批上市等多个流程，新药提取或合成衍生物合成是研发新药的关键点。对于中药而言，传承治疗思想形成的经验方或验方已经不少，而临床医生根据中医体质、八纲辨证及中药基本药理等中医理论还可以按照君臣佐使配伍原则拟定新的组方。对于活性成分或小分子的研究缩小了研发范围，提升了新药研发进度，并且可以根据疾病发病机制和药物作用机制双向出发验证药物疗效（图 4），药物研发可以基于基因、蛋白、代谢组学、影像组学、信号通路等作用机制从验方或组方追踪到单药，甚者具体到

药效活性成分和小分子。相较抗疟疾药青蒿素、抗肿瘤药三氧化二砷（AS2O3，砒霜）及心血管药物丹参酮等中药的研发与临床应用，新冠肺炎中药方剂的临床试验与应用，其研发周期大幅度缩短，成本-效益与国外抗逆转录病毒药物及 RNA 干扰药物相比更具优势。

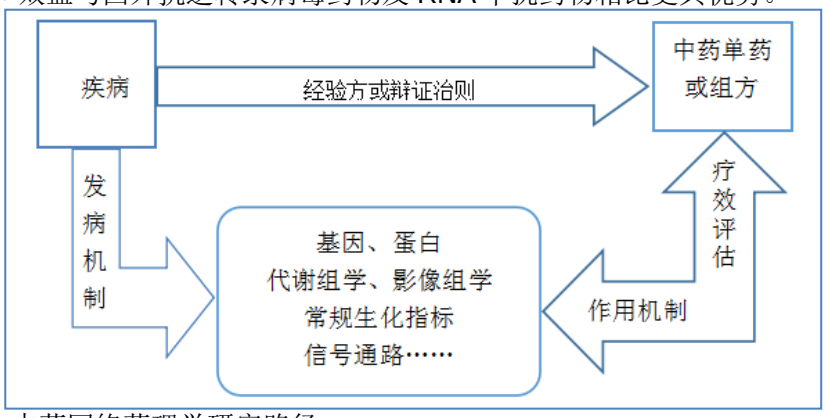


图 4 中药网络药理学研究路径

3 “治未病”预测模型

“治未病”思想是我国传统医学的瑰宝，《素问·四气调神大论》中的“圣人不治已病治未病，不治已乱治未乱”和《灵枢·逆顺》中的“上工治未病”等观点，揭示了治未病理论的重要性[错误!未找到引用源。]。经过两千多年的发展，中医“治未病”的技术方法得到了丰富和发展。“治未病”思想对于预防和治疗慢性病，延缓疾病进展，改善患者生活质量，减轻患者经济负担等均有明显作用[错误!未找到引用源。]。

大数据是信息化、网络化、智能化的核心。中医健康管理从数据采集、数据传递到数据处理，都与大数据密不可分[错误!未找到引用源。]。然而近几年大数据技术在医学领域的应用主要集中于信息科学研究，即大数据技术本身的研究，以及疾病诊断学和数学概念三个方面，其中构建预测模型用于疾病风险评估、诊断预测和预后预测是研究的热点，常用大数据技术包括 Nomogram 列线图预测模型、深度学习、LASSO 回归及随机森林算法[错误!未找到引用源。]等，但对于中医治未病领域的涉及并不多。中医治未病与体质和辩证论治思想息息相关，北京中医药大学中医学院吴涛涛[错误!未找到引用源。]研究认为以中医体质为预测变量构建相关疾病的临床预测模型是可行的。黑龙江中医药大学王超杰等[错误!未找到引用源。]研究发现高年龄组、低学历、中医体质中的痰湿质是不稳定性心绞痛发病的重要影响因素，阳虚体质、痰湿体质老年人是发生不稳定性心绞痛的高危人群。由此足以说明，在慢性病预防管理方面，除了常规因素外，还应从“治未病”思想出发，考虑中医体质等多种因素。构建大数据预测模型探索不同中医体质人群疾病风险是必要的。

4 中医养生信息个体化推送

大数据用于信息推广最成功的模式在商业领域，众所周知我国几大电商平台及搜索引擎，只要有在电商平台或搜索引擎搜索商品、疾病、兴趣，接下来就会推送类似或相关的商品或服务。但在医疗行业内基本属于空白，即使是某些图书馆，也不能做到个体化精准信息推广服务。当前，医院信息推广主要以微信小程序为主，以我省某中医院为例，其推广的信息包括医疗技术上的成就，获得的荣誉，根据时令季节变化推送养生之法，很明显，推送消息属于普遍性质类而缺乏针对性。

个性化信息推送是根据用户预先定制的需求，把符合要求的内容发送到用户的邮箱或 RSS（聚合内容，Really Simple Syndication）阅读器，及时为用户传递最新数据，省去用户登陆数据库的繁琐[错误!未找到引用源。]。对于养生，邓铁涛老先生推崇“上工治未病”的思想，认为养生重于治病[错误!未找到引用源。]。实现中医养生信息个体化推广可以从以下几个方面入手：①基础条件：拥有大量中医养生数据，即数据库或知识仓库。②系统中有患者中医辩证信息或健康个体的中医体质信息。③传统人工智能信息分类推送或人工标签模式推送。④对于即时咨询者构建结构化数据，借鉴 ChatGPT 构建语言训练模型，研发智能机器人程序或智能小助手。

5 结语

大数据时代，医院信息系统(HIS)数据、生物信息数据、医学研究成果库、临床医学研究数据、医疗卫生服务平台数据、公共卫生普查数据、居民健康管理数据等各类医疗大数据备受，通过分析医疗大数据可以为临床决策与医学科技发展提供重要信息，进而促进人类健康事业的发展。在中医药研究中采用医疗大数据分析与挖掘不仅可以发现中医药使用的特征与规律，中药主要药效成

分和小分子，也有利于中医药的现代化与创新，加速中医药国际化进程，使中医药造福于全人类[错误! 未找到引用源。]。但值得一提的是，大数据有多大才算大数据，如按照 TB 级别到 PB 级别这个标准来衡量，很多数据库的确是满足的。但随着新网络技术概念的提出和知识产权保护意识的增强，医疗数据共享情况较十年前并没有明显改善。

本文基于目前查新工作中常见中医药项目主题和媒体效应，从中医药传承、药物研发、“治未病”预测模型和中医养生信息个体化推送四个方面论述了大数据技术在中医药传承与创新中的应用，涉及的具体领域有限。但无论信息技术如何革新，应用领域如何拓宽，数据库和知识库是数据分析和应用的基础，除了已有的数据库和知识库，还应建立中医药领域专有的数据库，比如中药活性成分（小分子）数据库。中医药的传承与创新任重而道远，大数据和人工智能的应用势必为其提供技术支撑。

PU-238

基于选举多方安全计算实现的临床研究联邦学习框架

陈廷寅 [1,2] 冯嵩 [1,2]
[1] 中南大学湘雅医院（中南大学医院管理研究所）
[2] 国家老年疾病临床医学研究中心（湘雅医院）

各类数据以惊人的速度增长，也促进了数据采集、汇聚、分析和应用相关技术的快速发展[1]。随着全球许多国家对敏感信息和个人数据的收集实施了更严苛的隐私保护法规，传统的收集数据和处理数据的方法不再适用。如果只是收集非敏感或非个人数据，许多重要的价值将被错过。即使已获同意收集个人敏感数据，未经授权共享数据也会引起数据泄露的威胁[2]。收集的数据同样可能被不当处理，从而侵犯个人和公司的基本隐私权[3,4]。近年来，医疗领域产生了很多结合敏感和个人信息以实现其价值最大化的先进技术的发明，已成为该领域最大的科研增长动力。让这种增长变得更加重要的一个方面是，有许多场景都需要结合多个数据源获取的敏感信息，才能获得非常有用的价值。然而，在实践中，跨机构或多合作者的数据共享有各种限制，包括数据所有权、法律遵从性和技术约束等[5]。由谷歌人工智能研究团队推出的联邦学习[6]，支持多台设备或机构的去中心化协同机器学习。训练数据保存在每个客户机设备上，只有本地训练的模型参数被传输到中央聚合服务器，中央聚合服务器将这些模型聚合为一个全局联合模型。随后，联邦模型被发送回客户机，用于迭代地训练和改进模型，具体如图1所示。

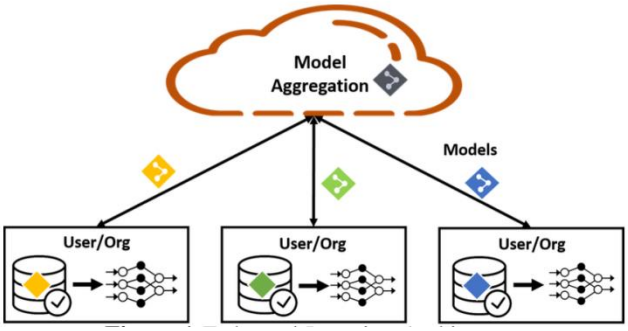


图 1 联邦学习基本模式

虽然联邦学习技术似乎可以确保数据保留在参与方本地，但已有研究表明，原始隐私数据仍可以被从联邦学习技术中提取出来[7]。他们公布的方法表明，机器学习模型参数可以被实施逆向工程，从而提取设备或机构本地提供的数据集[8,9]。一般认为，为了保证所有信息的保密性，防止间接数据泄露，需要对模型参数进行安全加密[10]，以防止联邦学习模型训练受到任何可能的逆向工程和第三方攻击。因此，常利用隐私保护技术与联邦学习协同作用以保护模型聚合数据机密性，其中的隐私保护技术，如：安全多方计算(MPC)[11]、同态加密(HE)[12...16]和差异隐私(DP)[17]等是比较常见的候选技术。很显然，谷歌已经认识到这一缺陷，随后提出了一种基于MPC的联邦学习实用安全聚合方法[18]; 此外，还有微众银行提出的

采用同态加密保护梯度的纵向联邦学习系统SecureBoost[19];值得一提的是OpenMind[20,21]在联邦学习框架中结合了MPC和DP功能。

MPC是一种加密技术，允许多个参与方在他们的敏感数据集上计算一个联合函数。只有联合函数的输出是公开的，而不透露参与者的私有输入。由于MPC协议、软件实现和底层计算基础设施的快速发展，安全的MPC性能得到了显著提高。目前，在典型应用中，MPC的实现速度比全同态加密(FHE)快数千倍。另一种确保数据安全性和机密性的方法是差分隐私(DP)。DP通过在计算过程中加入随机噪声来保护数据隐私，但由于噪声会影响联邦学习模型的准确性，因此不适用于所有联邦学习场景。

基于上述原因，本文选择基于MPC来实现临床科研联邦学习的模型聚合，提出一个基于选举的多方安全计算实现联邦学习框架。基于选举的多方安全计算引入了一种层次结构，在这种结构中，需要交换更少的秘密共享来进行隐私保护模型聚合。当参与方很多，而委员会成员的数量只是联邦学习参与方数量的一小部分时，这种技巧就变得更加有优势。点对点和选举的MPC实现的联邦学习框架都是在秘密分享协议的基础上实现的。联邦学习框架与实现多中心医学研究的科研平台进一步集成，该平台可以部署在不同机构间或公有云上。在我们的实验中使用了一个联邦学习建模的用例，以证明所提出的联邦学习框架在模型准确性和执行效率方面的有效性。

1 相关工作

联邦学习作为是隐私计算体系的重要组成技术，在医疗健康领域的实践已能支持较大规模应用的实施。在具体实践中，医疗机构、基因测序机构、科研机构等作为数据提供方和潜在的数据使用方，通过联邦学习实现各机构间数据互联互通，进行跨机构的精准防疫、药物开发、辅助诊断、基因分析、临床医学研究等应用。联邦学习技术的应用能够有效帮助多个医疗参与机构在满足用户隐私和数据共享方式符合政府相关法规要求的前提下，使得各医疗数据参与方不用向主中心共享原始数据，通过算法交互实现多方机构数据融合和机器学习联合建模，从技术上打破数据孤岛现象，实现医疗大数据科研协作。因此，联邦学习模式在医疗领域非常契合跨机构医疗大数据联合分析的需求，一方面保证了医疗机构对个人患者数据的隐私信息安全保护，另一方面也提高了医疗数据模型应用分析的准确性能。

联邦学习是一个非常宽泛的主题。因此，联邦学习可进一步分为横向联邦学习和纵向联邦学习。在横向联邦学习中，所有参与方(如：区域医院)都有完整的特征集和标签可用来训练其本地模型。从不同参与者计算的本地梯度被发送到服务器/聚合器，以训练一个全局模型。在纵向联邦学习中，所有参与方(如：同一城市内的银行、保险和金融公司)收集具有不同特征空间的数据，只有一方可以访问标签。由于缺乏信息，部分参与方(没有标签)无法自己训练模型。因此，为了安全地对齐样本空间，增加了额外的复杂流程，并且训练过程需要交换部分模型更新。本文仅对横向联邦学习进行研究。

目前，只有少数几个敢于挑战集中式机器学习范式的生产级联邦学习方案是可用的。谷歌已于2019 年 3 月推出名叫 TensorFlow Federated (TFF)[22]的联邦学习框架。CrypTen[23]是一个基于 MPC 的保护隐私机器学习框架，是由 Facebook 在 PyTorch 上实现的。FATE (Federated AI Technology Enabler)[24]是微众银行发起的另一个开源项目，旨在为构建联邦 AI 生态系统提供安全计算框架。

与现有的一般联邦学习框架不同，我们的联邦学习框架更为注重隐私保护和计算性能。通过将分散在不同医疗机构间的数据集成到具有可视化的多中心医学科研平台。虽然 MPC 已广泛应用于保护隐私的联邦学习框架，但在提高 MPC 支持的联邦学习框架的性能和可扩展性方面的工作还很有限。OpenMined 为模型聚合提供了一个非常基本的 MPC 开销评估。谷歌提供了联邦学习实用 MPC 协议的复杂度分析和通信开销评估，本文提出了一种基于选举的 MPC 操作，以实现高效的模型聚合。

2 基于 MPC 的联邦学习框架

在本节中，我们将说明传统的点对点 MPC 模式和基于选举的 MPC 模式支持的联邦学习框架的体系结构和实现细节。此外，还对交换的消息的数量和大小进行理论分析。

2.1 多方安全计算

多方安全计算是一类允许对敏感数据进行机密计算的加密技术。目前 MPC 技术占主导地位的两种技术是混淆电路和秘密分享[25]。混淆电路是一种基于布尔电路的密码协议。它的提出是为了解决很流行的百万富翁问题。两个百万富翁想知道谁更富有，但不透露他们的实际财富。近年来，基于秘密分享的 MPC 协议在生产系统中得到了广泛的应用[26]。一个典型的秘密分享 MPC 协议如图 2 所示，每个敏感数据被分解成秘密分享，这些秘密分享结合起来产生原始数据。它们通过 MPC 协议交换秘密份额，相互交互计算机密函数。只要大多数参与方诚实地遵循协议，将不会有任何参与方学到任何超出最终输出的内容。由于每个计算节点都在一个秘密分享上工作，只有在一定数量的计算方被破坏的情况下，攻击者才能从秘密分享中重建私有数据。因此，MPC 可以被视为一种提供可信方功能的密码学方法，可信方将接受私人输入，计算一个函数，并将结果返回给利益相关者，而不需要参与者之间的相互信任。联邦学习框架中采用了 Additive[27]和 Shamir[28]秘密分享 MPC 协议进行隐私保护模型聚合。

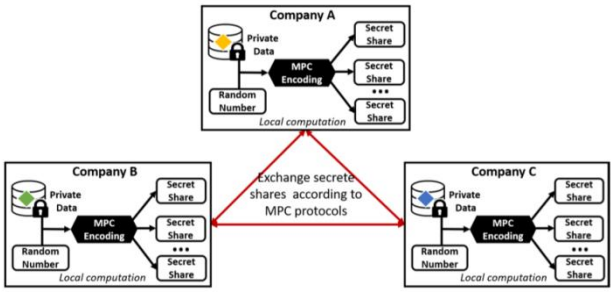


图 2 多方安全计算协议概览

为简单起见，本研究仅关注逻辑回归机器学习模型，模型聚合计算各方训练的局部逻辑回归模型之间的平均权值。由于联邦学习参与方的数量是已知的，因此只需要添加 MPC 操作。加法操作(如：计算 n 个参与方的私有数据 V(i)的 sum) 采用加性秘密分享的 MPC 协议，每一方使用使用随机数的值生成 n - 1 个属于自己的秘密分享。最后一块秘密份额是 $V(i,n) = (V(i) - \sum_{j=1}^{n-1} V(i,\ j)\ mod\ Q$。毫无意义的秘密份额与其他各方交换，然后每一方计算一份来自各方的私有结果的秘密份额的总和。再将各方计算出的总和进行交换，积累得到相加结果 S。采用加性 MPC 协议，各方对无意义的秘密股份进行保密计算。引入的随机性可以通过两轮消息交换和相加来消除，并且可以直接证明 $S = \left(\sum_{i=1}^{n} V(i)\right)\ mod\ Q$。机密计算只公开了最终输出 S。如果各方都是诚实但好奇而不勾结的，则不可能推断出各方的隐私值 V。相反，Shamir 秘密共享协议使用多项式插值，在有限域下是安全的。为了生成秘密共享 v，它使用了一个随机的 d 次多项式 $q(x) = a_0 + a_1 \times x + a_2 \times x^2 + a_d \times x^d$ 其中 $a_0 = r,\ \forall i \in [1,t],\ a_i$ 是随机数。给定任意 d+1 份额，多项式 q(x)可以用拉格朗日插值法重构。因此，秘密 v 可以用 v = q(0)简单地恢复。我们选择 d = n - 1，即所有参与者为每个私有数据生成 n 个秘密分享。秘密分享与其他方交换，以在不损害隐私的情况下对其私人数据进行加法操作。Shamir 秘密共享协议的安全级别和计算成本远高于 Additive 协议。但是，交换的消息的数量和大小保持不变。

2. 2 基于点对点多方安全计算的联邦学习

我们考虑联邦学习的标准场景，其中参与者在他们的私有数据集上同时训练他们的本地逻辑回归模型。目标是找到一个张量 T 的参数/权重的逻辑模型，以最小化经验损失函数。为了在不牺牲精度的前提下增强隐私性，采用 MPC 技术安全地聚合了张量 T。一旦满足收敛准则，训练结束。MPC 引入了计算和通信两方面的开销。因此，为使隐私保护联邦学习切实可行，提高性能是非常必要的。用于解决大规模、复杂的问题，导致机器学习模型越来越复杂。对模型的单个参数/权重进行连续的 MPC 操作将会带来巨大的开销。并行 MPC 操作的效率在之前的文献中得到了验证，因为批量操作可以显著降低计算和数据交换的成本[29]。图 3 说明了通过计算张量平均值进行模型聚合的过程。单个局部模型的张量是隐私敏感的。通过添加随机化，它们被分成多个张量的秘密分享。交换秘密份额后，每个参与者持有每个张量的一份。参与者对秘密分享进行局部聚合，然后进行全局聚合，以抵消随机化，从而得到所有参与者张量(即局部模型)的平均值。通过这种方式，参与者获得了全局训练时的联邦模型，同时保持他们的数据和局部模型是私有的。模型聚合中交换的消息数量 *Msg_Num* 可以按照如下计算：

$$Msg_Num = \left(n * (n-1)\right) * 2 * e = 2 * n\text{^}2 * e - 2 * n * e \quad (1)$$

模型聚合交换的消息大小计算如下：

$$Msg\_Size = Msg\_Num * s = 2 * n^2 * e * s - 2 * n * e * s \quad (2)$$

由于上述两个方程的计算复杂度为 O(n^2)，传统的点对点 MPC 实现的联邦学习框架的可扩展性较低[28]。另一方面，在实践中，当越来越多的公司加入联邦学习生态系统时，所取得的效益也越来越显著。

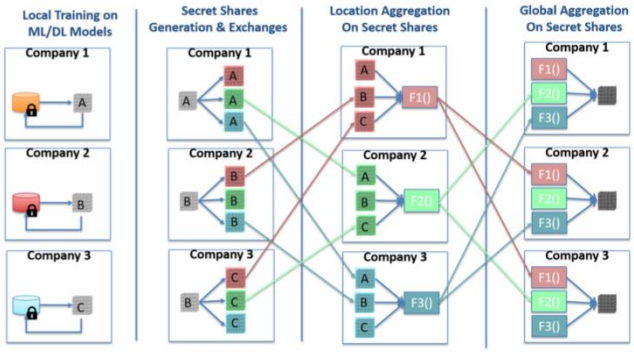


图 3 多方安全计算实现的模型聚合流程

2. 3 基于选举多方安全计算的联邦学习

为了提高可扩展性，我们提出了一个基于选举 MPC 模式的联邦学习框架，该框架避免了在所有联邦学习参与方之间交换大量的巨大的模型参数/权重的张量(以秘密分享的形式)。如图 4 所示，阶段一使用点对点 MPC 来选举联邦学习参与方的一个子集作为模型聚合委员会成员。阶段二使用委员会成员提供的 MPC 服务来聚合所有联邦学习参与方的本地模型。阶段一委员会选举，从 n 个联邦学习参与方中选出 m 个参与方，建立一个小范围的聚合委员会。各方同时执行该函数后，将得到委员会成员名单。各方生成一批随机数作为自己的初始投票。通过 MPC 协议将所有参与方的投票加在一起，在不透露所有参与方的初始投票的情况下，得到聚合的投票。第一阶段主要是边际成本，因为它只对每个联邦学习训练执行一次，交换的消息是一小批委员会投票(以秘密分享的形式)。在第一阶段交换的 Msg_Num 消息数为:

$$Msg\_Num = \left(n * (n-1)\right) * 2 = 2 * n^2 - 2 * n \quad (3)$$

用于交换 Msg_Size 的消息大小为：

$$Msg\_Size = Msg\_Num * b = 2 * n^2 * b - 2 * n * b \quad (4)$$

联邦学习各方对模型进行 t 轮迭代进行本地训练，得到本地模型 t (i, j, t)。他们生成 m 个本地模型的秘密共享，并上传给相应的委员会成员。随后，委员会成员一起进行机密模型聚合，然后将这个 epoch 的聚合模型 G(j)广播给联邦学习的所有各方。联邦学习训练和模型聚合在预先知道的全局 epoch 数或直到模型收敛。通过将巨大张量的秘密份额从 n 减少到 m 来控制第二阶段的代价，其中 m 远小于 n。第二阶段交换的 Msg_Num 消息的数量可以通过添加以下项来计算:

1) 各方都向委员会上传秘密份额;
2) 委员做模型聚合;
3) 委员会成员向所有各方传播聚合模式。

$$Msg\_Num = (n * m + (m-1) + n) * e = (n * m + n + m - 1) * e \quad (5)$$

在第二阶段中交换的消息的大小，表示为 Msg_Size 公式如下：

$$Msg\_Size = Msg\_Num * s = (n * m + n + m - 1) * e * s \quad (6)$$

通过累积第一阶段和第二阶段的 Msg_Num 和 Msg_Size，提出的基于选举的 MPC 实现联邦学习框架的通信成本如下方程所示：

$$Msg\_Num = 2 * n^2 + n * (m * e + e - 2) + m * e - e \quad (7)$$

$$Msg\_Size = 2 * n^2 * b + n * (m * e * s + e * s - 2 * b) + m * e * s - e * s \quad (8)$$

通过与传统的点对点 MPC 模式相比，我们提出的基于选举的 MPC 实现联邦学习框架可以显著减少消息交换的数量和大小。通过避免生成 n 个巨大张量的秘密分享(以秘密分享的形式)并与各方交换，可扩展性也得到了显著提高。下一章将给出了具体用例和参数配置的实验结果

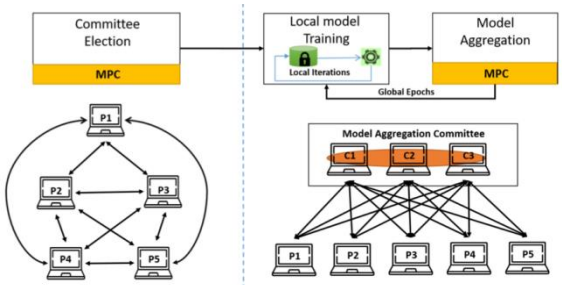


图 4 基于选举的多方安全计算实现联邦学习

2. 4 基于联邦学习的多中心临床科研框架

在进行多中心临床科研时，保护患者的隐私是至关重要的。然而，迄今为止，大多数的医疗机构都是采用集中存储（即所有参与研究的数据集中到一起存储在某个指定机构内部或指定的第三方）的数据库中的数据进行研究。这就导致了在进行数据分析和建模时完全可以访问到来自各自医院的数据。通过联邦学习的模式，不在需要这样集中的数据库，可以很好的解决因为数据集中而引起的安全性和隐私泄露等问题。通过联邦学习的模式可以保证机构内的临床实验数据无需离开机构本地，更广泛的可以接入患者的不同终端数据（如：手机、可穿戴设备或植入设备等）。此外使用联邦学习进行多中心临床研究的另一个好处是在数据不出域情况下，可以更好的促进联盟的形成，增大模型训练的数据规模，进而提高模型的质量。

结合当前国内医疗机构间网络状况现状，为了更好的促进联邦学习在多中心临床科研场景的应用，实现更好的联邦优化策略、更高效的聚合模式、更低的网络通信成本均已成为重要的研究主题。结合本文中提出的基于选举 MPC 的联邦学习聚合方法实现了面向多中心临床科研的联邦学习框架。框架训练的主要流程如图 5 所示。

框架执行的主要流程：

- 将医疗机构内部的数据进行预处理，处理后的数据作为联邦学习的输入数据；
- 机构内部用各自的本地数据建模；
- 周期性的将本地模型加密上传到虚拟的聚合节点（选举后的 MPC 节点）；
- 由虚拟节点完成全局模型的聚合并将聚合后模型下发到所有机构；
- 机构内部更新本地模型。

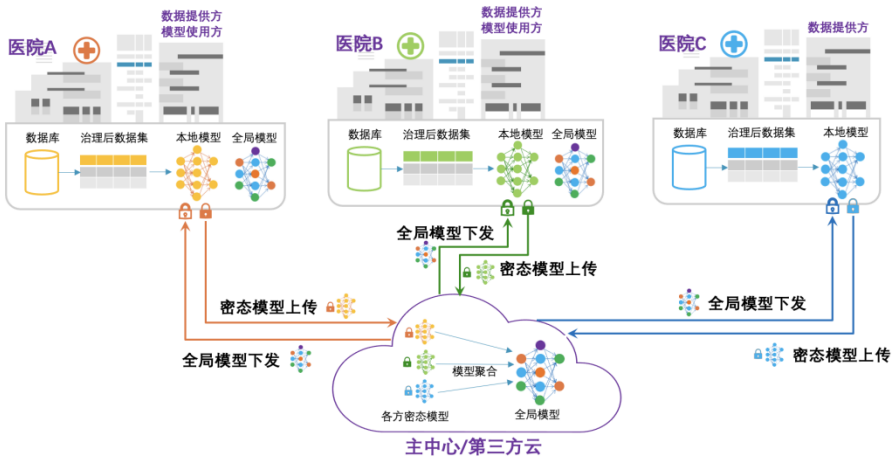


图 5 面向多中心临床科研的联邦学习框架训练流程

3 实验及结果

3. 1 实验设计

在实验中，我们使用 Additive 和 Shamir 秘密共享 MPC 协议评估了联合学习平台的性能。以下图表中分析的结果主要以 Additive 秘密分享结果进行对比。通过实验比较了传统的点对点 MPC 的联邦学习和本文提出的基于选举 MPC 的联邦学习在交换的消息规模和执行时间两方面的差异。实验在以下环境中进行：

同局域网环境：每个参与方运行在独立的 Docker 容器中，容器以 Centos7.5 OS 为基础镜像，分配 8GB 的 RAM，1000 Micro Core CPU。所有的 Docker 容器均匀部署在 8 台物理主机上。

3. 2 模型预测与准确性

本实验的主要目的是通过比较不同训练模型的准确性来验证联邦学习的有效性。

- 本地模型（Local）：仅基于本地数据集，用 3×15=45 次迭代训练单个模型。
- 集中式模型（Centralized）：上传数据集到集中式服务器，服务器根据合并的数据集通过 3×15=45 次迭代训练模型。
- 联合模型（Federated）：经过 3 次迭代训练本地模型后，加入点对点 MPC 支持的模型聚合，共 15 个全局的迭代。

为了衡量预测的准确性，我们采用三种流行的指标。

- 召回率：正确预测的正向事件的百分比。
- 准确率：预测的正向事件中，真实情况为正向的百分比。
- 平衡性：模型的总体性能考虑了正向和负向，不用担心数据集的不平衡。

我们选择任意三个机构的数据集作为训练数据，其余的一个作为测试数据，采用循环的方式。Additive 秘密分享和 Shamir 秘密共享 MPC 分别对逻辑回归模型取得了相同的实验结果。如表 1，联邦学习的预测精度与集中学习相当，验证了联邦学习对医学科研的有效性，且不泄露数据隐私。联邦学习模型优于局部模型，但对不同医疗机构的数据分布可能有所不同。

表 1 不同模式训练方法预测精度对比

训练方法	召回率均值	准确率均值	平衡性均值
Local	0.922	0.91	0.89

Centralized	0.939	0.97	0.93	0.941
Federated	0.933	0.94	0.93	0.935

3. 3 通信成本分析

我们根据所有联邦学习参与方和模型聚合委员会成员之间交换的消息的数量和大小来评估通信成本。在传统的点对点 MPC 框架中，只有当所有的节点都在一起工作时，才能重构共享模型。相比之下，在选举 MPC 框架中，共享模型可以由所有选定的委员会成员之间的共谋重构。基于>= 3 参与者之间没有共谋的假设，我们选择采用 3 方委员会。第三部分为理论分析部分。暂未进行形式化证明，使用逻辑回归模型;局部训练需要 3 次迭代(即 t = 3)，全局模型聚合需要 15 个 epoch(即 e = 15)。对于选举的 MPC 实现联邦学习框架，选择三个参与方作为模型聚合委员会(即 m = 3)。根据经验，一轮就足以决定委员会成员。

我们在同一个局域网内进行了一系列实验，模拟不同数量的参与方，控制参与方的数量从 4 个增加到 64 个。统计了消息的总数量和消息总大小。结果如图 6 和图 7。传统的基于点对点 MPC 实现的联邦学习框架因为消息的数量和大小随着参与方的增多急剧增加导致其可伸缩性较差。正如理论分析所预期的那样，本文中提出的基于选举的 MPC 方法大大减少了消息的数量和大小，更好的提高了系统的可伸缩性。选举阶段交换的消息数量与 MPC 真正执行阶段交换的消息数量相比是微不足道的，因为在选举阶段的每个消息都只是委员会投票过程中所使用的一个小规模数组(以秘密份额的形式体现)。

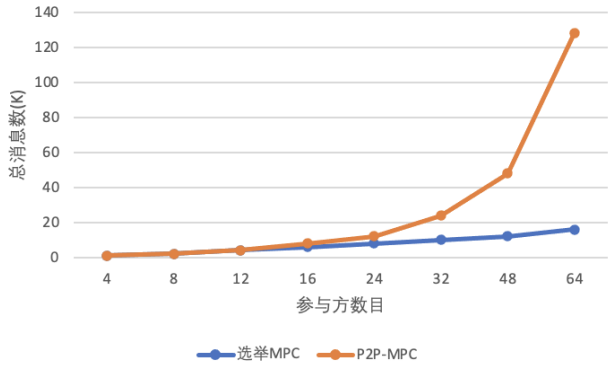


图 6 不同方式 MPC 消息总数随参与方变化对比图

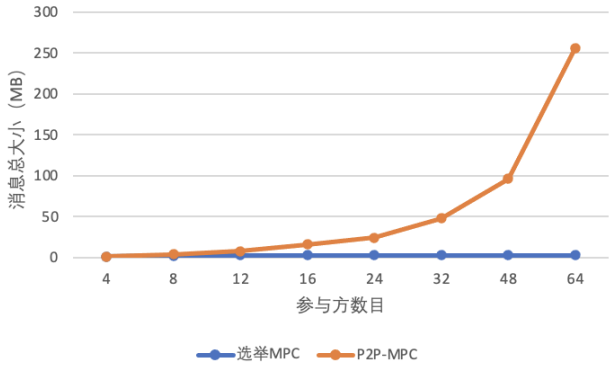


图 7 不同方式 MPC 消息总大小随参与方变化对比图

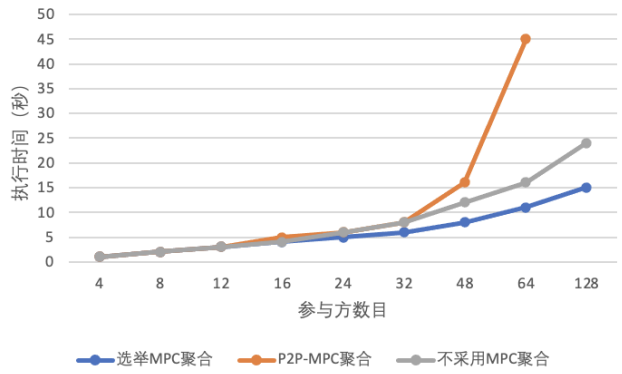


图 8 不同方式聚合执行时间随参与方变化对比图

执行时间方面同局域网环境下的实验结果如图 8，传统的基于点对点 MPC 的联邦学习框架随着参与方数量的增加也出现了扩展瓶颈。通过采用选举 MPC 实现联邦的聚合可以大幅的提升系统的扩展性。当节点数目增加到 128 个时，执行时间缩短了 25 倍。如前面分析的结论一致，基于选举的 MPC 联邦学习框架执行时间更可控。

4 结论

本文中，我们提出了一个基于选举的多方安全计算实现的联邦学习框架。它支持多个组织共同学习一个机器学习模型，同时保证内部的私有数据不被泄露。框架采用 Additive 和 Shamir 两种秘密分享 MPC 协议，以保护隐私的方式对局部模型进行聚合。通过联邦建模预测的实验结果，反应出联邦学习可以训练出比彼此隔离的数据集上训练的模型更好的模型，并达到与传统集中式机器学习相当的预测精度，从而验证了联邦学习的有效性。通信成本和聚合效率方面，本文与传统的点对点 MPC 聚合模式进行了对比，结果是基于选举的 MPC 实现的联邦学习框架大大降低了通信成本，提高了系统的可扩展性。根据参与者数量的不同，的逻辑回归模型，本文提出的方法可以在联邦学习的执行性能上提高 2 ~ 25 倍的速度。

在未来，我们将扩展我们的工作到更多应用场景，包括分诊诊疗、慢病防控、疾病早筛、医保控费等方向，并在实践中发挥联邦学习的优势。除了横向联邦学习外，我们还将研究纵向联邦学习和迁移学习，以实现跨多个领域的全面协作。此外，我们还将继续尝试提高系统的性能、可伸缩性以及隐私保护和网络安全保障，以应对在不可信环境中对联邦学习效率产生负面影响等问题。

PU-239

基于大数据技术的医院资源规划系统应用与思考

冯标斌 刘阳
中南大学湘雅医学院附属海口医院（海口市人民医院）

0 引言

在新医改政策的推动下，国内大中型医院的综合管理水平需要不断提高，医院精细化管理需要

一个有效的一体化管理信息系统支撑。显然 HRP 的建设能给医院带来管理上的变革，但目前市场上提供 HRP 系统主要来源于传统的 ERP 系统，虽然 ERP 在企业中的应用已经相对成熟，但是医院内部资源管理相比企业要复杂的多，企业管理的成功案例并不能有效解决医院管理存在的问题。因此，很多医院在实施 HRP 的过程中，直接照搬企业经验并不能达到预期的效果，反而导致“人财物”不能在同一个体系下很好的运作。

1HRP 系统在医院应用中的必要性

1.1HRP 系统概述

HRP（Hospital Resource Planning）系统全称医院资源规划系统,是基于企业 ERP（Enterprise Resource Planning 企业资源计划）的现代化技术和管理理念，结合医院业务特点和管理特色，整合医院已有信息资源，实现“人财物”、“医教研”、“护药技”管理科学化、规范化、精细化、可持续发展和战略转型的支撑环境。能够为医院领导提供医院经营数据，随时随地掌握医院运行情况，有助于及时做出决策的有利工具。

在发达国家已经有超过 50%的大中型医院已经部署 HRP 系统[1]，国内大多数医院的 HRP 建设刚起步，很多医院没有认识到 HRP 建设的重要性，认为医院相应的系统都有，何必要重复建设。在国家推进的医疗体制改革中，医疗机构内部的经济运营改革是重要的组成部分，精简、集约、高效、减少内耗成本，显然 HRP 的建设能给医院带来管理上的变革。

放眼全国，HRP 系统的开发商主要来源传统的 ERP 厂商，但都存在专业化程度与医疗行业不匹配的瓶颈，因此如何寻找合适的 HRP 产品就成为医院在选型时需要重点考虑的问题。

1.2HRP 系统基本功能特点

HRP 系统主要包含财务核算、成本核算、供应链管理、资产管理、专项管理、人力资源管理、绩效管理、全面预算等几大业务领域。

1.2.1 以预算为主线

当前 HRP 系统支持以医院的战略目标为基础，编制医院的年度工作计划和预算，并将年度预算分解到各个职能部门及业务科室，确立各项工作指标，建立基于预算的过程控制与管理体系，实现对收入、支出、工作量等相关指标的追踪，实现事前、事中、事后全过程管控。

1.2.2 全成本核算

结合最新全国成本核算办法及医院运营管理的要求，对医院成本核算进行规划及设计，以医院信息管理系统（HIS）为数据基础，创立标准成本核算体系，建立一套适合医院业务管理特点的成本核算办法。满足精细化管理要求，优化资源配置，为医院内控提供成本数据的支撑。全面支持按病种付费为主，按人头、按床日、总额预付等多种付费方式相结合的复合型付费方式。支持多方协同监管与动态监管，为医疗资源配置、医疗服务价格调整降低医院运行成本等，建立科学合理的补偿机制数据。

1.2.3 精益绩效管理

构建统一、灵活扩展的人力资源管控平台，以满足医院战略发展目标的人力资源规划为框架的规范统一、支持决策、全员参与、可灵活扩展和有效集成的人力资源管控平台，整合人力信息，实现全员人力资源集中管理。支持按照工作量、工作质量、平衡记分卡等多种方式的融合型绩效体系，将各个科室的核算方案批量计算提升工作效率，根据指标多维度考核，建立不同的指标评价

人。

1.2.4 供应链闭环管理

HRP 系统可以将物资的采购、验收入库、领用、结算、盘点、付款、维修保养、报废等各个环节纳入体系化管理，实现全物资、全过程的闭环管理。优化临床科室物资申请流程，可根据各医院实际情况定义适合的审批流，在每个节点设置工作流跟踪，在线实时反馈申请业务进度[2]。对申请采购环节做到层层监管，责任落实到个人，避免科室出现重复申请推诿责任的情况。

2HRP 在医院中的应用效果

长期以来，许多医院仅部署 HIS 系统和其它一些独立的业务系统，医院的信息化建设主要还是围绕以 HIS 收费为核心的医疗管理系统和以电子病历为核心的临床信息系统方向发展，但是人力资源系统、财务管理系统、医疗物资管理系统等各个系统的独立，导致人事绩效分析、设备效益分析、财务收支结余分析等无从下手。这直接导致信息脱节、流程脱节、管理脱节等现象发生，信息孤岛现象严重，医院运营产生的数据流还需要深度整合，无法为领导决策提供更有说服力的数据支撑。因此，如何以 HRP 系统落地医院发展战略、管理目标才是根本。以下将从“人”、“财”，“物”三点，并结合项目实施经验来说明。

2.1 人力资源管理

医院人力资源作为医院的核心资源，是医院管理中的重要组成部分，将医院高效的人力资源管理与现代信息系统结合在一起，可提高医院整体的工作效率和质量。从战略角度重新审视现代医院人力资源管理模式，结合现代信息系统的医院人力资源管理可以推动医院人力资源管理的职能从传统的人事管理到战略的人力资源管理的转变，提升医院战略决策的管理能力，增强医院核心竞争力[3]。

早期医院采购的都是独立的人力资源管理系统，系统功能相对比较简单，主要用来记录职工档案，个人信息，能够对基本的组织架构进行管理。从功能的角度是可以满足人事部门的使用，但对于现代化更全面的管理模式，对岗位管理、薪酬管理、保险福利管理、绩效考核管理等模块，并没有很好的跟财务对接。例如每个月的薪资报表还是要通过传统的 EXCEL 表格方式去提交，重要的是不能多维度分析人员薪资数据，形成统计报表。利用 HRP 的平台优势，把人作为一种投资，将人力资源转换为人力资本，把“人”这一医院最有价值的资源开发利用好，是推进医院可持续发展的基础。

在人力资源模块实施的过程中，HRP 的主要工作是梳理组织机构、理清层级关系、规范科室命名、统一编码规则。大多数医院都会遇到这种情况，在不同的系统中对同一个科室的命名不一样，对于实现全院信息互联互通会增加很多麻烦，所以 HRP 的人力资源模块需对科室名称和层级关系进行统一的梳理，并在政策上明确，所有系统涉及到的部门命名应以 HRP 系统为准，才能提高整体的管理效率。

2.2 财务管理

医院在实施 HRP 系统时应该由财务管理开始，再逐步扩展到成本管理、资产管理、物资管理和预算管理，医院所有的资源花费始于预算，终于结算，可以说财务管理是整个 HRP 的核心模块。HRP 能够将医院内部与财务管理相关的管理体制管理机制、运行方式、预算管理、人力资源等方面进行有效的整合，从而实现资源的共享和分类管理，这也在一定程度上能够对医院的财务支出进行有效的规划和控制[4]。

以财务作为顶层，建立集“预算、核算、监督、分析”于一体的闭环财务管理模式，将财务从单纯的记录、统计只能扩展到医院经济活动的业务过程中去，实现从核算型财务到战略性财务管理职能的转变。在预算管理的基础上，强化资金管理和成本管控；优化财务成本核算体系，细化财务

成本核算单元，建立内部服务计价体系，精准归集科室成本，以精细化财务核算辅助医院精细化管理举措；加强财务监督和分析，实现财务与业务的一体化协同，实时跟踪各项业务的开展情况。

HRP 系统在财务管理中的应用主要是扁平化的管理方式，应建立在 HIS 系统的基础上实现对医院各项业务的管理，打通各业务流程和财务流程的数据传递路径，自动生成相关凭证，保障账账相符。实现财务对业务的监督与控制职能，提高财务管理水平的同时，也能加强对 HIS 系统及其他业务系统的监管力度。最终的目的就是将业务数据转变为财务数据，再将财务数据转变成经济效益分析，为领导层提供直观有效的数据用于决策。

2.3 物的管理

物资的管理是现代化管理科学的重要内容，医院物资管理是医院为完成医疗、教学、科研等工作，对所需各种物资进行计划、采购、保管、供应、维修保养等各项组织管理工作。医院物资管理主要研究对象是物资在医院内的流转过程和管理方法，包括医院物资分类、物资定额管理、物资供应计划编制、物资的采购运输、物资仓库的管理和组织领导等。

医院对于医疗物资管控一直是个风险点，业务量越大物资的消耗也随之增多，相应也会出现各种错漏现象，物资管理部门的管控难度也随之增加。耗材的使用经常会出现出入库不平衡的情况，有些耗材更是难以追溯。由于一些耗材会经常进行调价，导致医院都会存在“一种耗材两种价格”的情况频繁出现，卫生监管部门每年都会对医院高值耗材的使用情况进行针对性的检查，使得医院在耗材管控方面存在不小的压力。由于医院 HIS 系统建设通常都早于 HRP，开始都是物价部门来对耗材收费项目进行命名定价，难免会存在耗材命名不标准，规格型号不统一的情况，这也会导致入库和实销对不上的情况，也会给 HRP 的上线增加大量工作。

诸如上述问题，医院在运用 HRP 对物资进行管控的时候，需要做大量的准备工作。要结合医疗业务预算合理性确定物资采购预算、科室消耗预算，强化二级库的管理，加强收费材料和高值耗材的“进销存”监督，做好非收费材料消耗定额，以切实降低材料成本消耗，并与预算、财务、成本、HIS 医嘱实现一体化联动。重点做好以下几个环节的工作：（1）采购环节：首先要加强资质管理，确保医疗安全；其次是要引入库存规划，规范采购策略，减少物资积压和资金占用；更重要得事要强化科室采购计划管理，杜绝无计划采购现象和跨期采购行为，同时也建立合理的应急采购制度。（2）保管环节：实现物资分类管理，完善物资属性，重点强化高值耗材管理机制，可对高值耗材赋予唯一标识码，做好物资出入库（含寄存入库）和移库管理工作，主要的目的就是理清家底。（3）使用环节：通过二级库管理、收费材料审核、高值耗材跟踪机制，清楚记录材料领用的时间，使用时间，使用人，用给谁，是否计费。强化设备巡检、维护、计量检定等运行过程的管理，实时了解设备运行动态，为医疗服务顺利开展提供保障，同时为设备效益分析提供设备运行成本数据。（4）核算环节：加强对库存物资的存货核算，定期对物资进行盘点，优化供应商往来结算流程，监督科室材料成本核算等，算清物资流动的明白账[5]。

3 医院对 HRP 系统应用的思考

3.1 统一顶层设计、分布应用实施

医院 HRP 系统建设的需求更多是着眼于当下业务需求和管理堵点，随着基础管理的完善，大量业务数据的产出，医院管理的精细度将再上一个台阶，对医院发展战略层面的考量将是重中之重。HRP 系统建设最大的难点在于管理弹性和柔性大，这就要求医院在构建系统时，必须依托医院发展战略及组织运行机制进行统一顶层设计，才能使系统建设更好的适应管理需求变更和业务流程变化的需求。

3.2 以先进的平台技术构建柔性 HRP

HRP 系统建设范围较广、应用深度高、管理可变性大、延伸性较强，建议医院在有条件的情况下引入先进的平台技术，以构建 HRP 系统来满足医院柔性管理的需求。系统建设不仅仅是满足当下的业务需求，更要基于医院信息化发展战略统筹规划，构建适合医院未来管理可持续升级的

信息化总体架构。

3.3 以一体化系统建设落地精细化管理

大多数医院存在信息系统都比较分散，信息共享不规范的情况，日新月异的需求导致接口频繁变动，接口开发成本逐年增高，数据的综合利用率偏低。因此，在 HRP 系统建设阶段应该从一体化角度统筹考虑：

（1）HRP 内部一体化：精细化管理需要一体化信息系统的支撑，通过建立一体化、闭环管理的 HRP 系统，实现财务业务一体化、临床运营一体化、集团化管理一体化，彻底消灭“信息孤岛”现象；

（2）HRP 外部一体化：HRP 建设会涉及较多的业务系统接口，如人事接口、财务接口、物资接口、设备接口等。传统点对点的接口方式已不利于系统设计、规划、维护、运营。建议通过数据集成平台统一进行数据交换，以便于异构系统集成、建立医院运营数据中心、主数据管理以及综合决策分析。

4 结语

通过 HRP 系统的构建支撑医院内涵建设，建立“以预算为主线，人财物精细管理为支撑，绩效管理为保障”的管理循环，有效帮助医院发展战略、管理目标的实现，并进一步将管理能力打造成医院核心竞争力。HRP 系统的确是帮助医院走向现代管理精细化的有效工具，但也需要大量基础数据作为支撑，而这些数据都是需要一点一滴建立起来，才能更好挖掘数据应有的价值，从而更加精准的实现“搞清人、核准财、管住物”的目标。

PU-240

新时代数据治理技术促进智慧医疗发展

黄虹

复旦大学附属华山医院

一、引言

进入大数据 AI 时代，数据质量的重要性不言而喻。而数据治理，则是大数据质量的重要保障和前提。按照国际数据管理协会（DAMA）的定义，数据治理是对数据资产的管理活动形式权利和控制的活动集合（规划、监控和执行）。

对于医疗领域来说，医疗大数据除了呈现有大数据的 Volume（大量）、Velocity（高速）、Variety（多样）、Value（价值）、Veracity（真实）的 5V 特点以外，还具有鲜明的冗余性、异构性及隐私性等特点。随着医院信息化发展，越来越多的医院意识到了数据治理的重要性，但是由于各个医院的发展水平参差不齐，只有部分信息化程度较高的三甲医院开展了数据治理工作，大部分医院仍然处于起步阶段，因此，在医疗领域发展数据治理方法和技术具有很强的必要性和紧迫性。

二、医院为什么要开展数据治理

2.1. 医院的数据现状

目前医院的数据现状导致医疗大数据的利用存在一些现实的困难，主要原因有以下几点：

①数据多样性：临床数据包含各种不同格式的多样化数据信息，如医嘱、用药、费用等结构化

数据：如医务人员问诊过程中产生的主诉病史记录、诊治过程形成的病程文字描述；如采用仪器检查产生的数字、图片、影像、视频、声、光、电信号数据等。

②数据异构性：临床数据分散在不同的系统中，包括有 HIS、LIS、PACS、EMR 等不同的信息系统，从临床数据类型来说，有不同的临床表型数据和生物样本数据，这些数据对接整合不足导致存在数据孤岛，数据不完整等问题。

③非结构化数据：临床诊疗过程中往往存在较多的不确定因素，无法在诊疗之前明确所有需要收集的数据格式，同时，医务人员习惯采用文本方式来录入临床数据，因此沉淀了大量的非结构化大文本数据。

④数据敏感性：医学伦理、法律法规遵从、患者的隐私保护、社会属性等，也是医务人员在处理临床数据时，不可避免的问题，造成临床数据的管理和治理成本较高。

2.2. 数据治理的意义

医院开展数据治理工作的意义主要包括以下几点：

①形成可用的数据资产：数据治理能够对各个业务系统医疗数据进行统一采集存储和处理，可以有效解决医疗信息孤岛问题，形成可用的临床数据资产平台，提供医教研管各个业务位面进行数据统一调度，满足不同方向的数据应用需求。

②提升研究效率：经过数据治理形成高质量临床数据能够减少临床研究者在患者纳排、特征数据提取等工作中消耗的大量精力，提高临床研究的效率，将医生从高强度工作中解放出来。

③助力临床工作：医生、护士、药师、检验检查技师等不同的医务人员角色需要快速、高效的对患者就诊数据进行查阅。数据治理能够对数据更好地进行整合和展示，从而帮助医务人员进行更加准确的诊疗工作。

③支撑医院运营管理：数据治理工作能够合理评估、规范以及管理医院的信息资产和运营情况，在有限的信息技术投资和管理成本降低的同时，能够挖掘和发挥数据资产的价值，为医院运营管理提供数据化的决策支持工具。

三、传统的数据治理方式和不足

3.1. 早期的数据治理：初步实现数据集成与连接

经典的数据治理过程初步奠定了数据应用的技术基础。经过数据汇聚，把分散在 HIS、LIS、PACS、EMR 等不同医疗系统中的数据整合到一个数据池中，重复信息被筛选剔除，相同类型的数据被规范统一，从而方便研究中对数据进行比较和分类。

3.2. 传统数据治理方法的不足

现代循证医学研究理论对于大数据应用管理有两个必要要求：1）可根据复杂需求灵活查找数据；2）导出的数据可直接用于统计分析。尽管过去不断有新技术、新标准和新组织推动数据质量的提升，然而我们仍然感觉到，完全支持可查找、可分析的数据资源仍然很少。具体来看，问题体现在以下几点：

- 找不到想要的患者

随着医学研究领域的不断深耕，研究目的不断细化，对入组患者的定位要求逐渐提高。例如，外科手术后出现术后并发症的 II 期肿瘤患者，在数据中往往很难定位。

- 导不出想要的指标

全时序临床数据描绘了患者全生命周期临床事件的轨迹，在此基础上进行纵向分析，进而以CRF 表单的形式导出，成为临床研究中的重要部分。然而，时序数据完整度不足等问题不仅阻碍了数据使用，往往还要进行大量的人工数据补录核查。

- 导出的数据不能统计分析

统计软件对数据质量要求较高，常见的问题主要有：数据填充缺失、数据分布异常、数据逻辑冲突、数据维数过多、数据标准化程度低。这些问题不仅会降低研究的可靠性,还会给分析结果带来偏倚。

传统的数据治理体系只对数据进行了“物理加工”，数据没有产生质变，若要成功催化医疗数据的“化学反应”，使得数据发生质变，传统的数据治理技术需要进一步进化。

四、新一代数据治理方法

4.1. 数据治理中的“化学反应”

实际工作中，医生需要的是患者全生命周期数据，然而从临床信息系统的数据库中，我们会发现存在大量的多源异构数据。这些情况造成数据并没有在临床场景中发挥更大的作用，反而在系统中产生信息冲突和错误。同时，由于医疗行业对信息错误的容忍度很低，错误的信息会导致不良的医疗结局。因此需要对临床数据进行深度治理，就如光合作用一样，让多源异构数据通过数据治理产生化学反应，如下图所示：

$$CO_2+H_2O \xrightarrow[\text{叶绿体}]{\text{光能}} (CH_2O)+O_2$$

1.原料：临床问题，临床证据　　2.产物：临床数据，临床成果
3.条件：医学知识，药物/器械　　4.场所：门诊，病房

通过数据治理后得到的真实世界数据，覆盖了患者全生命周期的多源异构信息，能够更全面且精准地体现症状和追踪疾病趋势，具有更好的时效性，能有效协助医生了解患者病情发展和疾病预后转归情况。

4.2. 阳光数据治理工厂——“化学反应”的制造工厂

正如树叶是光合作用的发生的绿色工厂，数据治理过程也需要一个阳光数据治理工厂来支持化学反应的进行。在阳光数据治理工厂中，集成了必要的技术和组织，保障了真实世界数据治理过程中的三个基本任务能够高效、安全、高质量的完成：

（1）将原始数据拆解为基本组成部分（即“分解”过程）；
（2）按数据模型对数据进行重新组装（即“重组”过程）；
（3）基于医学知识对数据进行标准化（即“归一”过程）。

如果将数据治理过程比作一种化学反应，用大数据和人工智能技术助力真实世界“数据”变“证据”的顺利转化，如下图所示：

数据治理工厂首先要支持建设真实世界数据标准，从底层建立规范，坚实基础。然后运用多重技术手段，支持院内外多中心、随访及移动端多源数据的融合与标准化处理，从而保证数据质量。

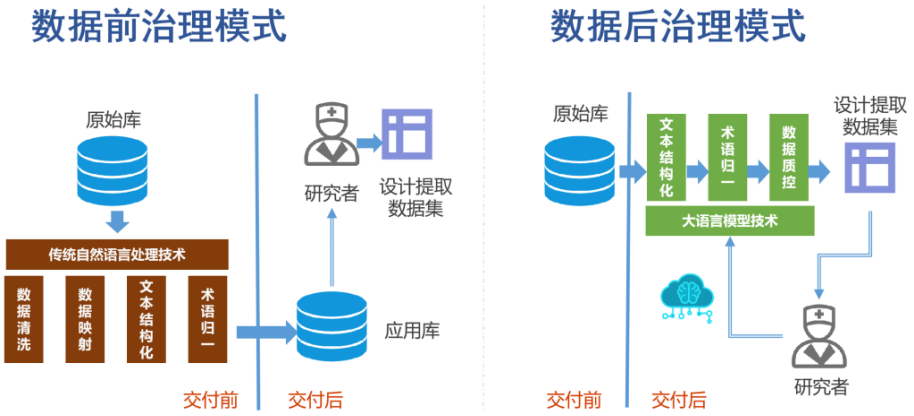
完善的技术支持和健全的数据治理组织结构，是全面开展数据治理工作的基础。数据治理 2.0 的能否落地实现，主要依赖以下三点：

- 通过打破原有信息系统数据结构，实现患者维度的数据重组，是数据治理的“化学反应”发生的前提条件。
- 构建治理真实世界数据的阳光数据治理工厂，为“化学反应”提供适当的孵化环境。
- 通过大数据技术，打通数据产生与应用的屏壁，实现数据在临床和科研两个场景之间的良性循环。

五、基于大语言模型的数据治理技术

自从 ChatGPT 横空出世以来，大语言模型成为整个人工智能行业的高频词。从大语言模型的特征分析，其先进的算法、规模庞大的参数、精准强大的语言理解能力、生产能力和逻辑推理能力，已经有足够的潜力给医疗领域带去深远影响。

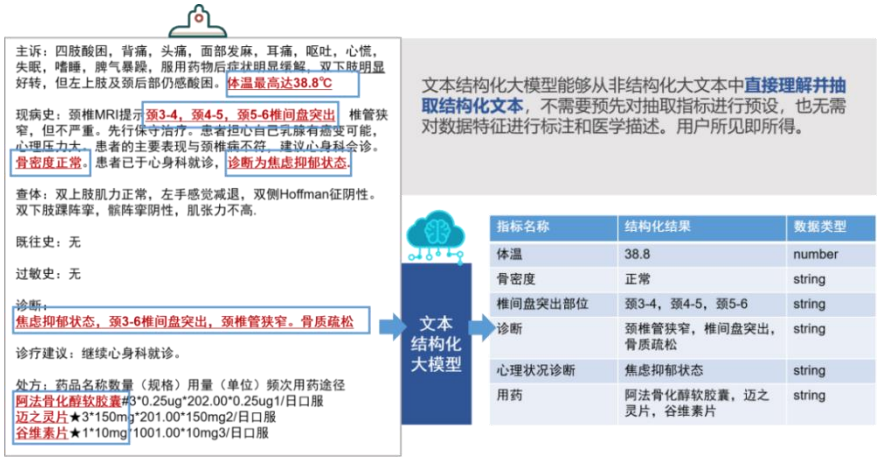
当前常见的医疗数据治理方式是一种前治理模式，即在数据与终端用户交互前对数据统一完成治理，而大模型技术的优势在于，由于其优异的可交互特性，结合其对大文本优秀的理解能力，能够实现一种后治理模式，能够根据终端用户的具体应用需要进行治理。



5.1. 大模型在电子病历文本结构化的应用

电子病历数据后治理的第一步便是对大文本进行信息提取。与直接提取词语的实体识别技术不同，大模型能够直接理解文本，然后输出这段文本中的对应信息，进而将非结构化的输入最终转化为结构化输出。

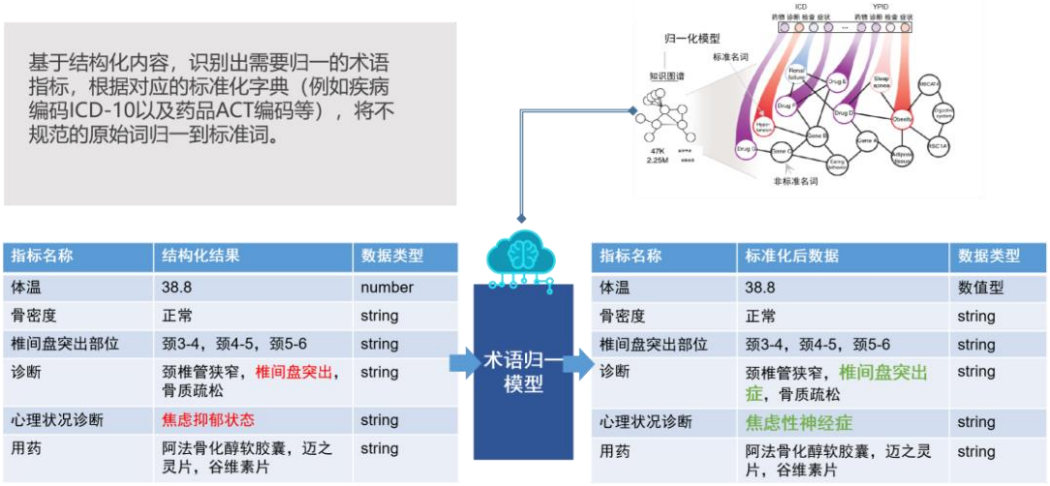
大语言模型本质上是一种生成式模型，在生成文本时，即便是目前最优秀的 GPT-4 也会生成非预期的输出，这为实际应用带来了潜在风险。为了缓解这个问题，我们采用多种干预策略，除了在模型侧使用对齐调整策略，使用多种提示词优化入参，同时基于业务逻辑对输出结果进行校验。在这个基础研发的文本结构化大模型，无需对数据指标进行预设或标注，即可智能化输出结构化文本信息。



5.2. 大模型在医学术语标准化的应用

大语言模型在计算资源有限的情况下并不适合直接用于术语标准化，但是利用大语言模型对文本信息的总结和摘取能力，可以让其在诊断相关的输出与医生的表达方式基本一致，再输入给医学术语标准化算法模型，最终得到术语归一化结果。

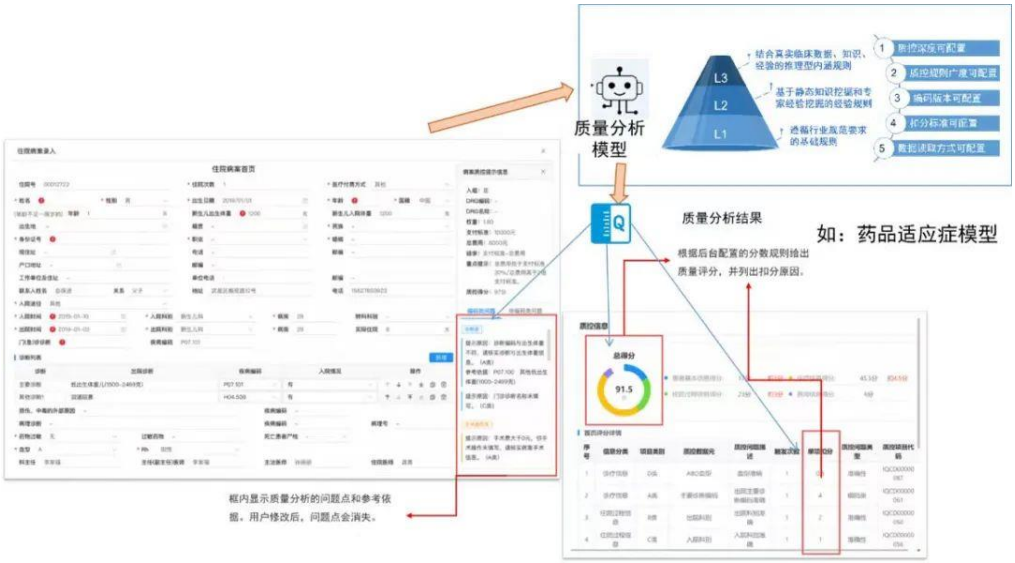
通常，应对术语标准化采取的技术方案框架是“多路召回+精排”。然而在实际业务落地场景中存在很多“一对多”的情况。因此我们将技术方案框架优化为“实体识别+向量编码+快速检索”。同时引入模型量化加速技术，以适配不同的业务场景需要。通过引入大模型，可以基于原始数据直接生成标准化数据，使输出的临床指标更规范。



5.3. 大模型在数据质控的应用

数据质量评估的目的是服务于上层临床应用，对数据过程中每个可能产生的错误进行发现和溯源，其本质是对病历文本所包含医学知识的理解和评审。

利用大语言模型能够从两个方面提升数据质控质量：在构建质控规则引擎方面，利用大语言模型解析药品说明书、临床路径等，构建专家质控知识库，同时在临床文本数据中挖掘诊断、手术、药品、治疗项目等关联度，构建后验质控规则。在质控应用方面，大语言模型语言理解能力提取的关键审核信息可以直接与质控规则引擎对接，根据后台配置规则给出量化评分，列出扣分原因，同时显示审核结果问题点和参考依据，可以进一步稳定数据治理的工程质量。



六、小结

近年来，随着医疗水平发展、硬件设备的迭代升级、算法模型的优化改进，AI 在医疗场景中技术积累越发成熟，应用场景日益丰富。基于深度学习算法和核心技术壁垒，同时结合由海量临床研究数据等积累的医学知识体系和科技工具进行定制开发，在电子病历结构化治理、临床决策支持等方面能够实现新一代数据治理技术体系的应用构建和落地。

经过 30 多年医疗信息化建设，医疗行业的关注焦点已经从信息系统建设延伸到了数据资源管理和利用。而要实现数据资源的有效利用，数据治理是关键环节。只有在数据治理领域奋发创新，让数据应用有效支撑医院发展，才能进一步提高医院管理水平，实现数据对医疗业务发展的价值赋能，持续推动我国医院的高质量发展。

PU-241

基于泛在物联网感知的大数据双碳管理及决策系统建设方案研究

段涛、邹小红
川北医学院附属医院

引言

2021 年 11 月，国管局、国家发展改革委、财政部和生态环境部印发了《深入开展公共机构绿色低碳引领行动促进碳达峰实施方案》，明确提出了有条件的地区 2025 年前实现公共机构碳达峰、全国公共机构碳排放总量 2030 年前尽早达峰。医院建筑作为较大的公共建筑之一，在日常运营中用能情况复杂，传统的水电表记录会因为没有对能源消耗的分项计量数据来诊断及系统的节能改造，导致能源的浪费。而节能减排是医院后勤管理的重头任务，为了反应国家的“双碳”政策号召，针对医院现有存在的用能管理方面的缺点：（1）计量体系不健全，即医院各区科室的能源设备分布广泛，但信息化管理水平低，不清楚能耗浪费位置，无法精确统计能源使用，导致无法有效节能；（2）环境质量检测要求高，即“智慧医院”和“绿色医院建筑”的评价标准规定了楼宇和室内环境质量的要求，如室内温湿度、光照、有机污染物浓度等。需要使用温湿度传感器和物联网检测仪实时监控，并为管理决策提供数据支持；（3）能耗设备智能度不高，无法做到精准控制，即医院的照明设备众多，尤其在办公区和地下车库。中央空调大且耗电。由于不同区域需求差异，空调耗能高。因此，采用智能控制策略对能耗进行分时管理显得尤为关键，可有效帮助医院节能减排；（4）消防安全及电梯安全需防范预警，即医院的核心功能区如空调机房和电梯机房通常依赖人工巡检，缺少预警。用电设备容易短路，消防智能化不足可能导致火灾。电梯使用频繁，需要持续监控以减少故障。

针对以上问题进行研究分析，设计了一套为医院建筑提供基于能耗、管理和安全的解决方案，该方案利用物联网技术创建设备管理平台，实时监测能耗、环境质量等，并智能化控制照明和空调，同时强化消防和电梯的安全管理。精准的实现物联网感知技术在大数据双碳管理和决策系统建设的应用，实现对各个单一物体的“高效、节能、安全、环保”的“管理、控制、运营”一体化[1]。

泛在物联网感知技术

物联网[2]是一种通过网络将物理设备连接起来的系统。这些设备可以收集并交换数据，使得物体能够互相“通讯”并做出响应。它可以通过各种有线和无线网络实时、准确地传输联网的目标信息，对检测对象进行智能识别和精确定位。

物联网的整体架构如图 1 所示，可以从多个层次进行划分，但一个常见的架构主要可以分

为：感知层、网络层和应用层。

感知层是物联网架构中的第一层，主要负责对物理世界的信息进行感知、收集并初步处理。它是物联网的“眼睛”和“耳朵”，直接与物理环境互动。感知层包括二维码标签和识读器、RFID 技术、Zig-Bee 技术、摄像头、GPS、传感器、M2M 终端和传感器网关等，主要功能是识别物体、采集各种信息，与人体结构中皮肤和五官的作用相似[3]。

网络层负责确保从感知层采集的数据可靠、高效地传输到处理层或其他设备。它为物联网设备之间以及设备与中心系统或云之间的通信提供了基础。

应用层是架构中的顶层，直接面向最终用户。它将底层（如感知层和网络层）提供的数据和服务转化为实际的应用和解决方案，为用户、企业或其他实体提供价值。

随着网络技术的飞速进步，泛在物联网这一概念应运而生。与传统的物联网相比，其显著特征在于其广泛且无处不在的设备互连能力。5G 和 IPv6 技术为这种广泛的物联网带来了关键技术支持。这些先进的技术大大增强了物联网的数据传输速度和覆盖范围，真正实现了时间和空间的无缝连接。

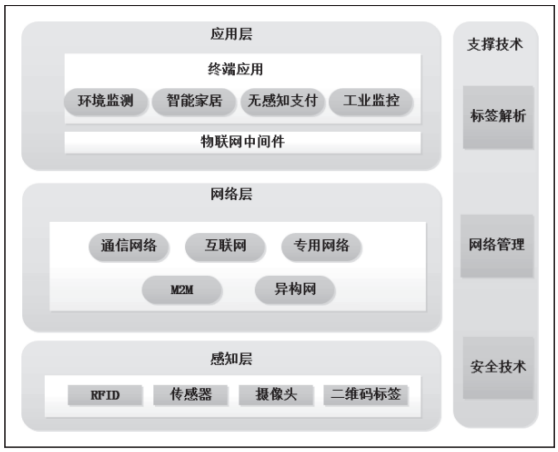


图 1 物联网架构图

大数据双碳管理及决策系统建设方案

医院建筑由于其特殊性，其能源消耗和碳排放情况与一般建筑有所不同。医院需要 24 小时不断电、有大量的医疗设备和设施需要稳定的电力供应，因此，实施大数据双碳管理及建设决策系统对于医院建筑来说尤为重要。以下是一个专门针对医院建筑的双碳管理及决策系统建设方案。

0.1 **建设理念与思路**

基于医院能源管理方面的现状，借助新一代信息技术，充分发挥多源数据融合应用能力和智能模型决策能力，为医院能源管理提供科学、高效的信息化平台，实现对医院能耗信息的数据采集与分析显示，实现分建筑、区域、科室以及重点设备等的能耗计量；对区域环境质量相关指标进行实时监测；对照明灯具与空调进行智能化策略管控；对院区消防安全与电梯安全智能化防范。其整体框架如图 2 所示。

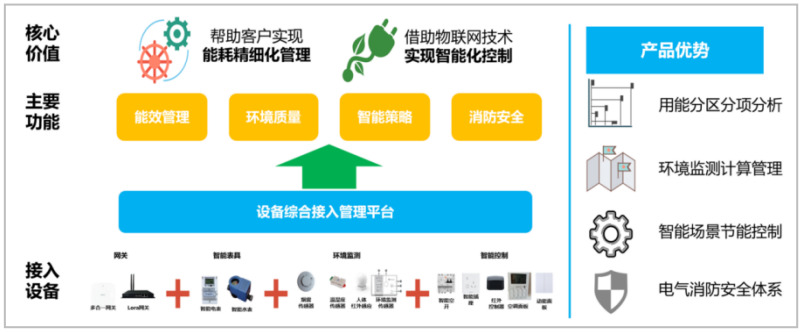


图 2 双碳管理与智慧决策系统框架图

以基于 LoRa、4G 等无线通信技术和有线方式，通过智能网关、控制设备，在医院的关键能源使用点（如中央空调系统、电力系统、照明系统等）安装传感器和监测设备，如光照度传感器、智能开关等，将采集的数据汇聚至平台，实现对建筑设备的精确控制、精细管理。其网络架构如图 3 所示。

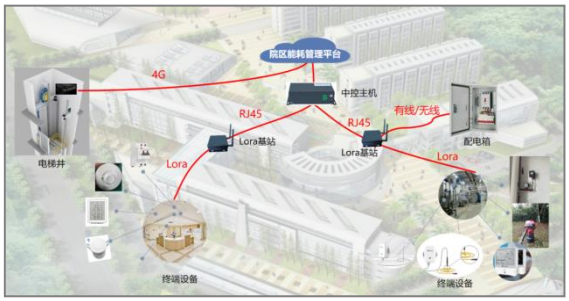


图 3 院区能源管理平台网络架构图

0.2 搭建设备综合管理平台

院区数据一张图呈现，通过能耗情况“一屏全览、一路追踪、一体智治”，提供了数字化、透明化的管理手段和决策依据，实现设备在线感知管理、节能减排有术可用、用能安全全程保障。为此设计的院区能耗驾驶舱可视化界面如图 4 所示，使得医院管理人员可以实时监控和管理能耗，以帮助医院更加高效、经济和环保地运营。

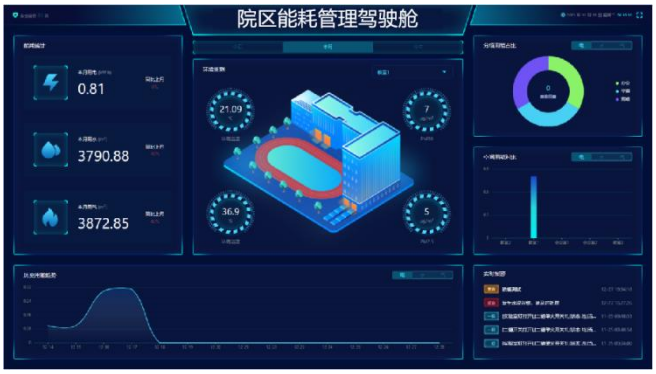


图 4 院区能耗管理驾驶舱示意图

开发能耗监控模块，实现节能增效、高效生产

院区物联系统实现对水\电等多维能耗数据采集，建立能源数据库，提供能耗分项统计、用能质量分析、能源审计、能效公示等功能。如图 5 所示，智慧能源高效管控云平台在线监测医疗建筑的能耗动态过程，收集生产过程的能耗数据，提供实时及历史数据分析、对比功能，以发现能源消耗过程和结构中存在的问题，通过优化运行方式和用能结构以及建立医院能耗评估、管理体系，提高医院现有供能设备的效率，实现节能增效、高效生产。

图 5 智慧能源高效管控云平台示意图

实现环境监控，打造安全健康环境

在温湿度 \ 二氧化碳浓度 \TVOC\ 光照度等环境感知数据、重点区域门窗状态数据采集基础上，联动告警策略，与新风、空调、照明设备联动，帮助院区打造安全、视觉健康、空气健康环境，如图 6 所示。（应用场景如天台门窗开启预警、空气质量监测、照明亮度监测）

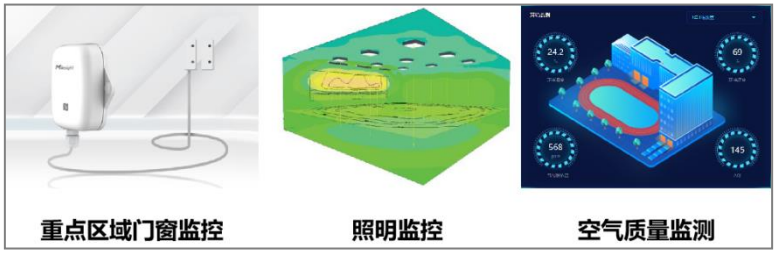


图 6 环境监控示意图

环境质量检测参考数据主要根据不同地区的标准、法规和建议进行制定。表 1、表 2 是一些常见的环境质量指标及其一般的参考值。

表 1 CO_2浓度参考

二氧化碳	说明
400ppm	同户外环境空气水平。
400-1000ppm	通风良好的室内环境。
1000-2000ppm	空气质量差，需要通风。
≧2000ppm	空气不流通，可能导致头疼、困倦、注意力不集中、心率加快甚至轻度恶心状况。
5000ppm	不要在此环境下超过 8 小时，否则可能出现人体机能混乱等其它严重症状。
≧40000ppm	暴露在空气中可能会导致严重的缺氧，永久性脑损伤、昏迷，甚至死亡。

表 2 TVOC 浓度参考

AQI 指数	说明
0-50	空气质量优。
51-100	通风良好的室内环境。
101-150	空气质量良好。轻度污染，需要通风。
151-200	中度污染，可能对健康人群心脏、呼吸系统有影响，需要加大通风量。
201-250	重度污染，健康人群普遍出现症状，需要优化通风系统。
251-350	严重污染，健康人群有明显强烈症状，应进一步优化通风系统并减少出席。
>350	极度污染，出现明显健康问题，应极大限度优化通风并避免出席。

注：TVOC[4]（“Total Volatile Organic Compounds）总挥发性有机化合物，室温下饱和蒸气压超过了 133.32Pa 的有机物，在常温下可以蒸发的形式存在于空气中，它的毒性、刺激性、致癌性和特殊的气味性，会影响皮肤和黏膜，对人体产生急性损害。AQI[5]:(Air Quality Index,空气质量指数)描述了空气清洁或者污染的程度，以及对健康的影响。

开发智能策略功能，构建院区智能环境

基于实时在线物联网采集传感控制硬件，支持高自由度自定义智能策略功能，实现对于照明、

空调、新风系统或指定设备运行控制、智能节能减排等应用，如图 7 所示。（应用场景如设备远程巡检、智能会议室、智能温控）

图 7 智能策略功能示意图

碳放管理，提供精准、高效的碳排计算工具

参考国际通用和我国重点行业制定碳核算模型的权威方法，结合本地实际情况，建立垂直行业的碳排放核算模型和核算方法。提供精准、高效的碳排计算工具，为未来碳资产变现奠定了基础，碳放管理流程如图 8 所示。（应用场景如辅助 CCER/CDCER 交易、碳排分析、减碳规划）

图 8 碳放管理流程图

用能安全，提供预警信息

在过压、欠压、缺相、过温保护基础上，提供用电监测、自动预警与自定义高级告警等功能，为医院提供全方位的智慧电气火灾预警防护体系。主要功能包括：当用电异常时，系统可自动判断并发出预警信息，可实现故障快速预警、快速响应；针对用电安全可以直接快速设定基础告警（过压告警、欠压告警、过流告警、短路告警、漏电告警、温度告警等）；设定自行根据需要进行高级告警（市电电压不稳、自定义过载告警）。智慧电气火灾预警防护体系如图 9 所示。（应用场景如分区夜间设备用电统计、线路老化预警、过载预警）

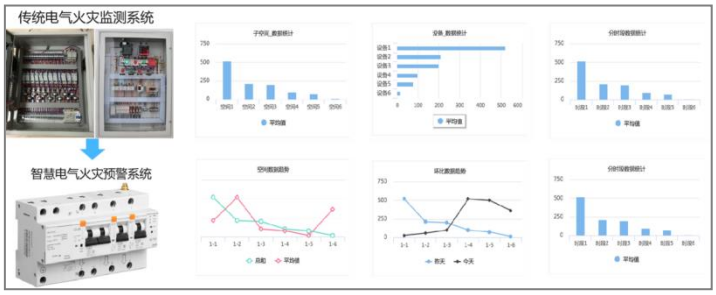


图 9 智慧电气火灾预警防护体系

能效分析

通过设备能效管理和统计分析实现对医疗设备的精细化管理[6]，利用设备开机启动频率，结合使用时间长短，可为设备折旧、维保核算、能耗提供分析报表，为医院管理人员提供参考依据，为重大决策提供数据支撑，真正实现设备的精细化管理。使用频率高、时间长的设备：设备的维护次数增加以延长设备的使用寿命；提高资源利用效率，为医院设备采购、配置、维保提供合理优化。能效分析实现过程如图 10 所示。（应用场景如设备/空间使用率监控、设备远程控制）

图 10 能效分析过程示意图

其他应用：

对电梯、配电和其它设备运行/告警状态监控、数据融合融通，提升智能、安全。

（一）智慧消防

通过消防水源设备、智能探测器设备、电气火灾等设备实现温感报警、液压监测、液位监测、消防栓管控、电梯安全等功能，为医院提供全方面的智慧消防体系，如图 11 所示。（应用场景如燃气泄露预警、消防设施远程巡检）

图 11 智慧消防示意图

（二）智能梯控

依托电梯物联网，围绕落实电梯监管部门、使用单位、维保单位主体责任，实现电梯设施数据有效采集、多系统间数据互通，如图 12 所示。（应用场景如电梯信息化维保、电梯智能监管、电瓶车进梯监管）

图 12 智能梯控示意图

1 预期效益

医院节能降耗是一个长期而细致的工作，降低医疗成本，对打造节约型医院具有十分重要的现实意义。基于本建设方案，为医院带来以下效益。

(一）智能节能减排

通过搭建院区能耗管理平台，对医院的开关、设备等加装物联网采集传感器，采集照明设备、电梯、中央空调、医疗终端设备等耗能数据，上传至院区能耗管理平台进行集中管控。

(二）环境的实时监控

通过加装温湿度、二氧化碳浓度 TVOC 光照度等环境感知传感器，对医院环境进行监管，帮

助院区打造安全、视觉健康、空气健康的环境。

(三）智慧用能安全

通过部署智能电气火灾预警系统，在过压、欠压、缺相、过温保护基础上，提供用电监测、自动预警与自定义高级告警等功能，为医院提供全方位的智慧电气火灾预警防护体系。

(四）碳排放智能管理

提供园区碳排放核算模型和核算方法、工具，实现追踪碳足迹，辅助规划减碳路径，为未来碳资产变现奠定了基础。

(五）智慧消防体系

依托电梯物联网，为医院打造智能梯控系统，实现院区电梯智能化管理，同时通过消防水源设备、智能探测器设备、电气火灾等设备实现温感报警、液压监测、液位监测、消防栓管控、电梯安全等功能，为医院提供全方面的智慧消防体系。

(六）智能优化策略

基于能源、管理、安全的建设可以优化医院能源结构、降低能耗、提升能效，进而提升医院用能管理水平，可以改善大型综合性医院在现有能源供应紧张、价格大幅上涨的大环境下的能源费用支出持续上升状况，同时可以加强用能安全、改善医疗环境，在能源费用降低的同时促进医院利润增长。

2 小结

基于医院的能源管理现状，本文设计了一个基于泛在物联网感知技术的大数据双碳管理及决策系统。这是一个融合了多种先进信息技术的方案，为医院打造了一个科学且高效的能源管理信息化平台。它不仅支持能耗和环境的实时监控，还融合了智能策略、碳排放管理以及用能安全等多项功能，旨在优化医院的碳排放平衡，为医院创建一个智能、健康、环保的环境，进一步推动绿色医疗和可持续性发展。

PU-242

基于 ETL 的心肺血管疾病专病数据库建设与实践

窦一峰 单淑香* 庞东旭 蒙文涛
天津市宝坻区人民医院

目的 探索心肺血管疾病患者诊疗数据的采集方法，构建心肺血管疾病专病数据库，为提高临床医师和科研人员提供准确全面的数据资料。

方法 以天津市宝坻区人民医院心肺血管疾病患者的临床诊疗数据为基础，采用 ETL 的方式实现多源异构数据的抽取与转换、数据治理、数据质控以及数据分级分类和数据安全。

结果 完成了专病数据库的建设并投入使用，包括 11 个一类域，21 个二类域和 866 个数据单元，共计 19998 例患者的数据资料，并取得了较好的应用效果。

结论 心肺血管疾病专病数据库能够明显提升医务工作者的科研效率，有效助力挖掘真实世界医疗数据的价值。

PU-243

大数据时代背景下医院信息管理创新发展的工作讨论

潘传红① 雷保仓① 何斌① 关鹏①

河南省漯河市中心医院 漯河市医专一附院 信息中心

1 引言

随着科学技术的不断发展，新技术的出现使医院信息呈现出增长速度快、形式多样化、数据量大等特征[1]。生物医学大数据蕴含了极其丰富的信息和知识，是关乎国家持续发展、国人生存健康的重要战略资源。大数据概念的提出让人们意识到真实海量的医院数据相关分析比少量抽样数据分析更有价值[2]。信息管理是信息分析利用的前提，医院信息新特征的出现使当前医院信息管理无法满足海量数据分析利用需求。近年来，人们对大数据背景下的医院信息管理提出了新的要求，即创新理念、转变思维、注重应用[3]。医院信息管理如何适应新时代的要求,推动医院发展是我们需要思考的问题。

2.医院信息管理发展历程及其特征

2.1 信息时代下的医院信息管理

20 世纪是信息时代，其目标是“信息化”，利用信息化提高业务工作效率，计算机技术是其技术支撑。医院信息化建设经历了“单板机与 PC 机”技术应用阶段，局域网技术应用阶段，互联网技术应用阶段三个阶段[4]。近年来，在政策支持及社会需求推动下，我国医院信息化由系统建设期逐步进入以整合为主阶段[5]。以满足医院管理和医疗职能部门的信息管理需求为主要目标的信息化建设，使医院基本实现了各职能部门的信息化，管理和医疗工作得到了长足的发展，但医院信息标准化滞后，“信息管理”仅作为各部门完成各自职能管理的一种手段，其信息标准和信息流程规范仅以本部门传统职能要求来制定，管理系统仅成为实现各部门职能管理工作流程信息化的工具。未考虑医院整体发展战略需求，形成各自为政，信息共享程度不高、信息资源利用价值低下，信息化投资效益和信息资源率不高的局面[6]。

2.2 大数据时代下的医院信息管理

21 世纪，全球进入了数字化互联网时代，物联网、云计算、人工智能、传感技术、移动通讯网络等技术已逐步渗入医院，给医院信息管理带来新的挑战。大健康概念的提出[7]、“互联网+医院”[8]、“智慧医院”[9]的出现，使得健康服务者与被服务者能够使用多种移动终端设备、从不同位置、通过多种手段来接入互联网，并在这一过程中不断创造新的健康信息，越来越丰富的在线健康应用和服务，不断激励医院健康服务者与被服务者创造和分享信息。医院数据将不仅仅局限于围绕疾病而产生的医疗数据更多的将是围绕健康服务而产生的大健康数据，数据规模大，增长速度快，数据种类多。这些特征昭示着医院大数据的到来。医院数据来源于医护人员向人们提供健康服务及以该服务活动为中心的延伸活动，其反映的内容与医院真实世界中的发生息息相关。在大数据时代，人们对医院真实数据进行大数据分析比小样本数据分析更能揭示问题的本质[10]。这一方法使人们不受限于传统的思维模式和特定领域里隐含的固有偏见，可以更真实更深刻的洞悉医院需求，提高医院管理效率、提升业务水平。21 世纪是大数据时代，其目标是“数据化”，大数据技术是其技术支撑。

3.医院信息管理的创新发展思路

医院信息新特征对医院信息管理提出了新的要求,本文从三个方面审视当前医院信息管理，并采用新的观点对当前医院信息管理所遇到的相关问题进行解答,提出医院信息管理创新发展思路。

3.1 从基于“信息”到基于“数据”的概念创新

在当前医院信息管理中，信息的价值是相对的，例如对于工资管理系统而言，患者的治愈率并不是该系统要管理的信息，相反对临床信息系统而言，患者的治愈率是一条重要信息，当需要依据患者治愈率来考核医生绩效从而确定工资时，就有两种选择，一是重复采集每一个医生所治疗的患者的治愈率数据；二是治愈率数据提取转换到工资管理系统上，但该方法将面临着数据格式不能转换或数据转换格式后丢失信息等棘手问题。

分析原因，我们不难发现，当前信息管理的“信息”在采集时由于人们表达信息的方式与管理信息的需求不同而被动的进行了加工，呈现出不同的数据形态。在“信息”这一概念的指导下，各系统间数据共享困难导致医院面临着数据重复采集、利用率低、维护费用高等困局。为了进一步解释该问题，本文认为数据是信息的供体，信息来源于数据。“数据”是管理和服务对象行为活动的表征和描述，是以数字、符号和文字来标记，具有属性特征，且能被机器识别和处理的概念或资料，不应随个人表达信息方式与管理信息需求改变而改变。

3.2 从“信息管理”到“数据管理”与“数据资源管理”的模式创新

随着科学技术的发展，大数据时代的到来，医院信息管理的思路与要求也发生了变化。过去，“信息管理”仅作为医院各部门完成各自职能管理的一种手段。医院的各部门根据工作需要，提出各自信息管理需求，并以部门职能和工作流程设计了信息项（管理指标）和管理信息模型，通过计算机编程，完成信息管理系统设计，并应用到些部门的中，实现管理信息化。现在，在大数据的视角下审视医院信息管理不难发现，信息管理是受制于科学技术，而不得已开展的一种活动，是为了解决局部、具体的管理自动化问题以提高管理工作效率为主要目标，是应用计算机技术来实现的一种管理活动。

“数据管理”是在云技术和大数据理念支撑下，提出的一种新型管理模式。它的主要特征是在管理对象行为活动描述域和数据量的极大扩展，它不仅可高质量和高水平的满足和覆盖信息管理所有需求和目标，且可通过数据的知识化分析和数据挖掘技术来发现新知识和新规律。

“资源管理”是管理学范畴的概念，也是经济管理学主要研究的内容[11]，它关注的是如何实现资源的优化配置和合理利用，是以追求资源利用的最大效益为目标的一种管理活动，资源配置、资源利用、资源效益是资源管理的主要内容，“配置”、“利用”和“效益”是资源管理的三要素。“数据资源管理”较“数据管理”，在管理理念上将数据管理提升到一个服务于医院全局的资源管理层级，其管理目标是数据资源的优化配置、共享利用、效益最大化。数据资源管理要求对医院数据进行统筹规划和标准化处理，采取集约化管理方式，实现数据资源的分类整理、知识化分析、主题化汇聚，达到管理目标。相对于“数据管理”，“数据资源管理”在管理目标、管理技术理论与方法等诸多方面均有重大提升和质的跨越。

3.3 从“信息系统集成平台”到“数据资源管理与服务平台”的技术创新

“信息系统集成平台”是将孤岛式、分散的信息管理系统主要通过“接口”和“视图”等技术实现信息在医院内信息共享与交换[12]，是在原系统基础上的整合和技术改造。这种方式保持了各个信息系统原有的信息管理目标、模式和信息流，以及系统的相对稳定性。但从医院信息管理战略定位和管理理论、技术方法层面而言，这种系统集成仍然仅是一种“技术改良”，虽然这种集成也可实现所谓的院内“信息共享”但仍然是“信息”级别，而不是“数据”级别的共享与交换。

“数据资源管理与服务平台”是基于医院管理理论、卫生经济学理论、计算机技术方法和医院总体发展战略，而设计的一个数据资源管理平台，可实现全院数据层级的管理，它的“总体”宽度和“资源”深度属性是医院信息管理技术体系构建理论的深化和本质上的跨越。一方面它将“信息管理”提升到“数据管理”模式，大大扩展了体系的管理域和数据量，同时将“各部门”信息管理需求来规划设计系统扩展到“医院总体”管理需求来统筹规划体系结构和数据的标准，保证了医院数据共享的顺畅度和高水平，是真正地高质量实现“数据共享”的解决办法和模式。另一方面，它将管理学的卫生经济学理论和方法作为体系构建的依据，将资源的“优化配置”、“合理利用”和“最大效益”等管理三要素作为体系的核心目标，是在“信息管理”目标基础上的重要和关键性延

拓。有利于数据的多元化采集、知识化处理、主题化汇聚、集约式管理。有利于医院形成大数据能力和发挥信息化建设效益，实现医院管理现代化、智能化。

4.创新发展的特色优势分析

当前医院信息管理是从信息管理的角度对医院进行管理，在这种管理理念下，医院积累了大量的数据。随着移动网络终端在医院的运用医院数据将会快速增长，数据产生方式和种类繁多。在这样的背景下，传统的“信息”——“信息管理”——“信息系统集成平台”将不能满足医院的发展需求。此时需要结合大数据时代构建新的管理概念、管理模式与理念、管理技术，即“数据”——“数据资源管理”——“数据资源管理与服务平台”（以下简称“新管理平台”）。

4.1 强调信息化建设投资效益

新管理平台的核心是数据资源管理，其在数据管理的基础上引入了卫生经济学理论与方法，把数据当作一种资源，强调数据资源的合理利用、优化配置与效益最大化。真正做到用数据管理医院、服务临床与科研，较过去信息管理更强调信息化建设的投资回报。

4.2 强调数据资源规划与标准化

医院数据来源于医院管理和医疗行为活动，其真实地反映了整个医院行为活动。为了将医院数据转化为数据资源，需要对医院的整体数据进行规划，同时为了避免数据理解表示不一，还需要相应的数据标准作为支撑。

4.3 强调信息技术与大数据技术相结合

新管理平台要实现“用数据说话、用数据管理、用数据决策”的管理目标，需要将信息技术与大数据技术相结合，在信息技术的支撑下利用大数据技术实现大数据分析与挖掘，并在知识工程相关理论与方法的指导下，从数据中获取知识，进而用知识指导行动。

5.思考

信息管理发展策略创新。在新管理平台的指导下，医院应在现有系统稳定运行的情况下，将数据资源建设工程作为医院信息化建设的首要任务，即以数据资源为核心驱动，构建数据资源管理与利用一体化平台，建设以医院数据资源为核心的智慧医院数据生态圈，实现医院数据资源的最大化开放与共享利用。

医院信息管理人才需求。信息管理部门作为医院信息化建设的责任部门，是新管理平台下的信息管理主体，肩负着保证数据质量、科学管理数据、合理利用和优化配置数据资源以及数据知识化处理的任务，然而当前从事医院信息管理的人员的知识背景主要为计算机[13]，缺乏知识工程、数据工程和信息标准化等专业技术人员，信息化技术团队人员结构性问题比较突出，无法适应资源管理和大数据技术的需求。因此未来的医院信息化建设者应该由具有医学、医院管理学、标准化科学、卫生经济学、计算机学、知识工程学等多学科知识背景的人组成。

综上所述，以大数据的视角看待医院信息管理，使人们从最基本的概念理解认识医院信息管理，为解决当前医院信息孤岛和信息化医院转向为智慧化医院指明方向。然而在新思路指导下的医院数据资源管理需要在实践中不断完善。

PU-244

多模态医疗健康数据分类与应用

葛令乔、渠鸿竹、方向东

1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室

2. 中国科学院大学 北京 100049

3. 基因组与精准医学检测技术北京市重点实验室 北京 100101

1. 引言

近年来，我国的医疗水平伴随着社会的进步与发展，得到了极大的提升。同时，随着科技的进步与发展，各级医院的病历记录系统与诊断检查结果保存系统都得到了更为规范化、系统化的提升。随着各类检测与影像技术的广泛应用，医疗领域也面临着海量的数据冲击[1]。Intel 公司的统计报告预测：2020 年医疗数据存储量增值 35ZB（1ZB=230TB），是其 2009 年的 44 倍。在多模态医疗数据成为医学界公认的研究资源的背景下，将其应用于科研问题探索发现与临床诊断治疗指导的案例已不胜枚举。因此，对多模态医疗健康数据分类与应用的总结可以帮助临床医生更好地对数据进行应用挖掘，从而产生更加有效的治疗策略与手段。

2. 数据类型

本文将多模态医疗健康数据，按照其数据类型，分为了基础数据、成像数据、化验结果、病历数据、遗传学数据 5 个方面，并分别对各分类所包含数据的具体内容以及数据的应用手段进行总结归纳。

2.1 基础数据

多模态医疗健康数据中的基础数据主要涉及病人的基础信息，包括年龄、性别、身高、体重、腰围、臀围、视力、血压等。同时，基础数据也包括病人的流行病学特征，如居住史、影视、体力活动、疾病史、家族史、睡眠精神问题等内容。

基础数据的获取方法众多，首先，基础数据可以通过医生的问诊记录中提取；其次，也可以通过各大体检部门或单位的信息记录系统中提取；同时，也可以在进行某些医疗相关的前瞻性队列研究的计划中，通过调查问卷的形式从参与者处获取。

基础数据在医学研究中具有不可替代的重要性。身高体重等基础物理信息的应用有评估个体健康状况的重要指标、疾病的诊断和监测、药物剂量的计算、人群研究和流行病学研究和个体化医疗。病史、居住史、家族史等背景调查可以提供关于患者疾病诊断和分类的重要线索、评估患者的预后，同时也可以用于疾病流行病学研究，通过分析大量患者的病史信息，可以了解疾病的流行趋势、高危人群、风险因素等，为疾病的预防和控制提供科学依据。

2.2 成像数据

医学影像是指为了医疗或医学研究，对人体或人体某部分，以非侵入方式取得内部组织影像的技术与处理过程。随着技术的进步与发展，逐渐形成了以 X 射线（包括 CT）、磁共振成像、超声和核医学为代表的多种医学影像方法。目前常见的医学图像有 X 射线图像、核磁共振图像、超声波图像、内窥镜图像、眼底摄影图像和核医学图像。

生物医学影像目前已经成为了疾病诊断和治疗中不可或缺的一部分，特别是在癌症的诊断中，将医学成像与活检结果相结合，使其成为确诊的重要依据[2]。但随着科学技术的发展和成像技术的进步，医学影像的数量大幅提升，越来越多的医学影像需要得到专业医生的解读。然而，医学图像的解读是一个非常主观的工作，医生有可能会因为经验不足或者疲劳而产生解读错误，使一些疾病漏诊，导致假阴性的出现[3]。由于多模态数据的统筹应用，将病人病历中的诊断数据与成像数据相结合，使用人工智能的手段，通过计算机辅助诊断，已经成为了多模态医学数据利用中卓有成效的一个案例。如何充分利用人工智能深度学习方法处理这些超大规模的医学图像大数据，为临床医学中各种重大疾病的筛查、诊断、治疗计划、治疗图像引导、疗效评估和随访提供科学方法和先进技术，是当前医学图像分析领域急需解决的重大科学问题和前沿医学影像关键技术[4]。

2.3 化验结果

医学化验结果是指通过对患者的生物样本（如血液、尿液、组织等）进行实验室检测和分析后得出的数值或定性结果，这些结果可以反映患者的生理状态、疾病风险、疾病诊断和治疗效果等信息。

医学化验结果可以根据不同的指标和检测项目进行分类，常见的分类包括以下几个方面：（1）血液学指标；（2）生化指标；（3）免疫学指标；（4）凝血学指标；（5）微生物学指标。

医学化验结果的应用主要有以下几个方面：（1）诊断和筛查：医学化验结果可以作为辅助诊断的依据，帮助医生确定疾病的类型、程度和预后，例如，血液学指标可以用于诊断贫血、感染等疾病；生化指标可以用于评估肝功能、肾功能等；（2）监测疾病进展和治疗效果：医学化验结果可以用于治疗效果的检测和评估，例如肿瘤标志物的变化可以用于评估肿瘤的治疗效果，糖化血红蛋白的变化可以用于监测糖尿病的控制情况；（3）预测疾病风险和预后：医学化验结果可以用于评估个体的疾病风险和预后，例如，血脂指标可以用于评估心血管疾病的风险；（4）个体化治疗：医学化验结果可以用于制定个体化的治疗方案。例如，基因检测可以帮助选择合适的药物和剂量，提高治疗的效果和安全性。

2.4 医学病历数据

医学病历数据是指医生在诊疗过程中记录的患者相关信息的集合，包括患者的个人信息，各种检查结果、诊断方案、治疗方案、用药记录等。

医学病历数据是医生进行诊断、治疗和研究的重要依据。医学病历数据可以用于科学研究和医学教育，医学研究人员可以通过分析大量的病历数据，探索疾病的发病机制、风险因素、治疗效果等，为疾病的预防和治疗提供科学依据。医学教育中，医学病历数据可以作为案例进行讨论和学习，帮助培养医学生的临床思维和判断能力。电子病历蕴含疾病描述、诊断、检查、治疗等大量临床数据，在医学专家和信息学家的共同参与下，利用人工智能技术挖掘电子病历数据的研究急剧增加。电子病历包含大量临床信息，有诊断、化验结果等结构化数据和临床表现等非结构化数据。基于分类编码系统，结构化数据可以被快速检索，然而，对于非结构化数据，其在自然语言处理下仍较难准确检索，且相关诊断可能只在临床记录中提及，需要不断完善算法[5]。

2.5 遗传学数据

遗传学数据是指从生物医学实验室、临床领域和公共卫生领域获得的基因组学数据、蛋白质组学数据、代谢组学数据等。遗传学数据通常都具有大规模、多样化特点，其分类可以根据数据的类型和来源进行分类：（1）基因组学数据：包括基因序列、基因表达、DNA 甲基化等数据，用于研究基因与疾病之间的关系、基因功能等；（2）蛋白质组学数据：包括蛋白质表达、蛋白质互作数

据等，用于研究蛋白质结构、功能和相互作用等；（3）代谢组学数据：包括代谢产物的测量数据，用于研究代谢途径、代谢物与疾病之间的关系等。

遗传学数据在医疗领域方面的应用众多，通过整合和分析大规模的遗传学数据，可以揭示疾病的发病机制、病理生理过程等，为疾病的预防和治疗提供科学依据。同时，也可以通过分析个体的遗传学数据，制定个性化的治疗方案，提高治疗效果和安全性。通过分析大规模的生物医学数据，可以用于药物靶点的发现、药物筛选和药效评估，加速新药的研发过程。

3 多模态数据综合应用

如何综合管理与运用这些规模极大且信息有效价值极高的数据就成为了一个急需解决的问题，也是促进多模态医学数据的进一步利用的关键问题。

目前，针对多模态医学数据统筹管理与运用最为有效的方式就是队列数据库的建立。队列研究是通过追踪是否暴露于某种可以因素或者按照暴露程度分为两组或多组的人群，并比较个组的结局以及结局的发生率，来判定该种可疑因素与结局的关联以及关联程度。大型队列研究对上万人群的健康状况以及疾病发展进程进行持续数年的追踪，成为生命组学样本和表型数据的重要来源[6]。同时，征得病人许可后从医院系统中获取的病人的基础信息以及病历与检验结果等，也可与大型队列研究中的目标队列相互补充，相互完善。在大数据和精准医学时代，大型队列研究是慢性病病因研究的首选方法之一，队列遗传资源已成为我国重要的战略资源。中国目前已建成了若干大型队列，例如中国慢性病前瞻性研究项目[7]、泰州人群队列[8]、上海女子队列[9]等。

推动多模态医学数据与队列遗传资源的标准化建设与利用可有效促进生物医药领域研究与技术开发的原始创新[10]。关于多模态数据与队列数据的处理、系统化分析目前也开发出了包含生存分析、全基因组关联分析、孟德尔随机化、荟萃分析、医学影像挖掘、临床文本挖掘等众多分析手段与方法。

4 总结与展望

虽然经过了多年的建设，我国已形成一定规模的多模态医学数据资源，这些数据资源也为人民群众的健康水平的提升、精准医学相关研究以及疾病的防控做出了非常大的贡献，但中国多模态医学数据统筹管理系统的建设仍然面临着一系列的挑战。多模态医疗健康数据还需要朝着不同医疗系统资源规范化、格式统一化、规模扩大化、组学内容丰富化的方向发展。不同医疗系统资源规范化有助于数据资源的统筹管理，极大提升后期数据的利用效率；格式统一化有助于各个医疗系统之间的数据资源互通共用，使数据流通效率大大提升；规模扩大化有利于提升数据分析结果的有效性，使得研究结果在更大的人群规模中得到验证；组学内容丰富化有助于生理过程中基础机理的探索，揭示发病机理，从而促进治疗手段的提升。多模态医疗数据资源的统筹建设是中国医疗系统建设中不可或缺的一部分，是中国未来生命健康科学领域向上攀登的最为坚实的台阶，需要得到各级医疗系统的重视。

PU-245

基于人工智能的多组学分析技术助力精准医学

郑光敏 [1,2,3] 渠鸿竹 [1,2,3] 张昭军 [1,2,3] 方向东 [1,2,3*]

1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室

2. 中国科学院大学

3. 基因组与精准医学检测技术北京市重点实验室

1 引言

过去二十年来，各种组学技术正在以前所未有的规模和分辨率表征不同的分子层特征，为疾病的发生提供了整体视图，也为系统描述分子机制、找寻关键生物标志和开发生物医学应用算法等方面提供了数据资源。但剧增的数据分析给研究人员带来了重大挑战，人工智能已证明能够分析多模式多组学数据流，多组学技术与人工智能算法的同步发展，推动了精准医疗的发展。

2 多组学技术

随着高通量技术的蓬勃发展，各种组学技术已经被开发出来，主要包括基因组学、表观基因组学、转录组学、蛋白质组学和代谢组学等。各种组学技术通过描述从 DNA 到代谢物的几乎所有生物分子，能够对复杂生物系统进行系统、定量的表征。

基因组蕴藏着遗传信息，它调控基因表达并决定细胞的结构和功能[1]。基因组分析在可鉴定疾病相关基因变异[2, 3]，监测疾病治疗反应和描述耐药机制[4]。

表观遗传变化修饰影响基因的表达和功能[5]。在人类疾病的发生和发展过程中起着根本性的作用[6, 7]。部分表观遗传生物标志物和治疗策略已经纳入医保[8]。

转录组动态变化反映了机体的生理重塑[9]，空间转录组学表征了转录组概况，同时又保留了有关组织空间背景的信息，为癌症生物学提供了重要见解[10]。转录组技术在生物标志物发现、疾病诊断、分类以及治疗方案优化方面具有广泛的用途[11, 12]。

蛋白质作为最终行使功能的分子，参与 DNA 复制、遗传信息传导、催化代谢反应到驱动细胞运动各个方面[13, 14]。蛋白质组学分析可为临床确定新的治疗靶点和可操作的生物标志物[15, 16]。

代谢组是参与代谢的代谢物或小分子化学物质的集合[17]，可以直接反映生化反应的功能。最近的研究表明，肿瘤细胞可以重新编程细胞内代谢，以支持增殖的需求，同时它对于免疫细胞也至关重要[18-20]，与细胞分化、功能和命运相关。代谢失调导致的特征性代谢表型，可用于早期疾病诊断、监测或疾病治疗靶点[21, 22]。

单细胞组学技术的出现使得以前所未有的高分辨率探究基因组改变、DNA 甲基化、染色质可及性、基因扰动、转录组动力学、蛋白质丰度和代谢重编程。此外，各种高阶单细胞多模式组学技术允许在同一细胞中同时分析不同组学层面的信息[23]，从而提供对细胞行为和身份的全面了解。因此，单细胞组学技术有潜力改善精准医学的多个领域，包括早期检测、诊断和风险分层、非侵入性监测和药物靶点发现[24, 25]。

3 基于人工智能的多组学分析技术

人工智能（Artificial Intelligence，AI）方法在不同的组学分析领域取得了显著的成功，在基因组学数据分析中，包括 3D 基因组结构预测、表观基因组修饰和染色质可及性建模、基因组注释、sgRNA 设计和基因组编辑结果预测等。几篇最近的综述总结了基因组学分析的机器学习方法[26, 27]，但是基于深度学习模型的基因组学分析技术也在迅速发展。在转录组学领域，人工智能方法已被广泛用于研究基因表达、选择性剪接和转录因子结合，促进疾病进展分子机制的发现。转录组学技术与机器学习相结合可以作为早期疾病检测[28]、分类以及预后和复发预测的辅助诊断工具[29]。

人工智能方法在蛋白质组学和代谢组学分析中也得到了很好的应用。人工智能适用于蛋白质组学分析，特别是在识别未知肽方面[30]，具有很高的准确性。此外，人工智能方法还被用于解析蛋白质与核酸和蛋白质与蛋白质单键相互作用[31]，这对各种生物过程具有重要意义。同样，代谢组学研究也高度依赖人工智能方法，它常用于建模过程中代谢建模参数的优化和代谢网络重建[26, 32]。

目前所有的组学分析技术都离不开利用人工智能的方法来加强分析。人工智能方法也承担了单细胞组学分析中的大部分任务，包括数据去噪和插补、批量效应校正、细胞聚类、轨迹推断等。在多组学范围内，一系列人工智能方法，已被广泛应用于多组学数据的集成分析[33]，通过人工智能算法整合多组学数据和非组学数据，例如临床和电子健康记录数据以及临床实验室测试，可以提供链接组学信息和临床表型并全面分析疾病状态的机会。因此，基于人工智能的多模态多组学数据综合分析对于精准医疗的临床诊断和决策具有广阔的前景[34]。

4 基于人工智能的多组学分析在精准医疗中的应用

4.1 基于人工智能的多组学分析在疾病早期检测和筛查中的应用

早期疾病检测受到极大关注，因为它有助于及时治疗疾病来降低死亡率。近年来，液体活检（例如血液检测）已被证明是一种微创且有利的早期疾病检测和指导治疗方法[35]。然而，血液中循环生物标志物含量相对较低，这限制了它们的临床应用。基于人工智能的多组学分析可以显著提高液体活检的灵敏度和准确性（表 1）。Cohen 研究组在血液测试中，通过将循环蛋白表达和循环肿瘤 DNA 突变与 AI 相结合，开发了 CancerSEEK 来检测早期癌症[36]，CancerSEEK 以 70% 的准确度识别出 8 种癌症类型和五种之前未能早期检测的癌症，其敏感性为 69-98%。Wang 研究组通过使用血清外泌体的 microRNA 谱和肝细胞癌标志物甲胎蛋白（AFP）的蛋白质表达[37]，开发了一种区分肝硬化和肝细胞癌的诊断模型 MedCalc，AUC 为 0.93。在另一项研究中，Chabon 研究组开发了一种基于 cfDNA 的机器学习方法 Lung-CLiP[38]，该方法使通过将肿瘤特征与 cfDNA 突变/片段结合起来，使其达到了与肿瘤相关 ctDNA 分析相似的灵敏度。随着人工智能算法、液体活检和多组学技术的发展，人工智能整合多种数据类型推动早了早期疾病检测和筛查[39, 40]。

表 1 基于 AI 的多组学分析在疾病早期检测和筛查中的应用

应用	组学技术	人工智能算法	代表成果	研究来源
疾病早期检测和筛查	cfDNA、蛋白质组数据	CancerSEEK	识别多种癌症类型	[36]
	miRNA、蛋白质组数据	MedCalc	区分肝硬化和肝细胞癌。	[37]
	游离 DNA	Lung-CLiP	检测肺癌的早期阶段	[38]
	mRNA 表达	SVMs, Random Forests	识别癌症和正常组织的准确率高达 97.89%	[39]
	mRNA 表达、基因组学、表观基因组学	AMARETTO	通过分析 CNV、DNA 甲基化和基因表达数据来识别泛癌驱动基因	[40]

4.2 基于人工智能的多组学分析在疾病诊断、分类和分级中的应用

随着高通量技术的发展，癌症的分类已从传统的“形态学”发展到“分子分型”的新时代。从多组学技术中提取的分子特征已用于诊断和分类疾病亚型（表 2）。基于深度神经网络（Deep Neural Networks，DNN）的算法可以有效地处理大量数据并整合多个数据以进行疾病分类和诊断。基于 DNN 算法，Sun 研究组开发了基因组深度学习（GDL）模型检测 TCGA 数据库中癌症相关的点突变[41]，通过分析癌症特异性点突变，构建了分类器模型可以精确区分 12 种癌症类型和健康组织，AUC 为 0.94。Capper 研究组开发了一种根据 DNA 甲基化图谱对肿瘤类型进行分类的算法[42]，并将诊断的准确性与病理学家的组织病理学诊断进行比较，基于机器学习（Machine Learning,ML）的分类器与病理学家的分类一致，尽管部分病例 （15.5%） 被分配到肿瘤子类别，但进一步的分子分析表明，分类器准确预测了 93% 的不匹配病例。一项类似的研究还验证了使用 ML 把 DNA 甲基化和基于拷贝数的分类用于癌症诊断[43]。此外，DNN 算法也被应用于识别癌症的起源，一个基于 DNN 开发的模型，利用 TCGA 中 18 种癌症类型 DNA 甲基化数据来预测癌症起源和类型[44]。总体而言，这些研究证明了将人工智能与多组学数据相结合在疾病诊断、分类和分级中的应用前景[45]。

表 2 基于人工智能的多组学分析在疾病诊断、分类和分级中的应用

应用	组学技术	人工智能算法	代表成果	研究来源
疾病诊断和分类	WES	DNN	通过检测 WES 数据中的基因组点突变，基因组深度学习 (GDL) 模型可以识别 12 种健康组织和癌症组织	[41]
	DNA 甲基化数据	随机森林算法	基于 DNA 甲基化的分类器用于诊断中枢神经系统肿瘤	[42]
	mRNA 表达、甲基化	随机森林和不同的分类器	识别乳头状肾细胞癌 (PRCC) 驱动基因并区分 PRCC 的早期和晚期	[45]
	DNA 甲基化数据	深度神经网络	识别癌症类型和起源	[44]

4.3 基于人工智能的多组学分析在预测疾病预后和治疗反应中的应用

过去几年，人工智能算法已将基因组学和转录组学等多组学数据整合到风险预测模型中，用于预测疾病复发风险的评估和指导治疗反应[46-48]（表 3,4），例如 MesoNet、Decipher[49]和分子预后评分 (mPS) 系统[50]，可以为医生提供更准确、更适合的治疗策略。Erho 研究组证明针对 22 个基因表达谱进行训练的 Decipher 评分分类器可以精确预测前列腺癌患者前列腺切除术后是否发生转移转移[49]。同样， Shimizu 研究组建立的基于 AI 方法的通用分子预后模型， mPS 系统，可以精确预测乳腺癌患者的生存情况[50]。

表 3 基于人工智能的多组学分析在预测疾病预后中的应用

应用	组学技术	人工智能算法	代表成果	研究来源
疾病预后	mRNA 表达	Decipher score	预测前列腺癌的早期转移和患者的生存	[49]
	mRNA 表达	分子预后评分 (mPS)	基于 23 个预后相关基因 mPS 可精准预测乳腺癌生存期	[50]
	甲基化、mRNA 和 miRNA 表达	深度学习	用于预测肝细胞癌生存的 DL 模型，一致性指数为 0.68	[46]
	甲基化、CNV、mRNA 和 miRNA 表达	最小冗余和最大相关性 (MRMR)	使用多组学数据预测卵巢癌的生存率	[47]
	mRNA 表达、CNV、SNP	小组正则化深度学习癌症预测 (GDP)	使用多组学数据训练 GDP，以预测泛癌症的生存情况，例如多形性胶质母细胞瘤、肾透明细胞癌和膀胱尿路上皮癌	[51]

AI 算法预测药物治疗的反应更具临床相关性，有助于临床医生提出更好的治疗建议。Johannet 研究组使用黑色素瘤患者的组织学数据和临床数据开发了一个多变量分类器[52]，可以预测免疫检查点抑制剂的治疗结果，AUC 为 0.80。还有研究组使用人工智能衍生的基因特征（AIGS）来预测膀胱癌的临床结果[53]，其结果表明，AIGS 准确地区分了对免疫疗法或其他治疗策略敏感的患者。此外，基于多组学数据分析，AIGS 还对一些基因突变和拷贝数变异进行了表征，为膀胱癌患者提供了潜在的治疗药物。这些研究表明，整合多组学数据的人工智能算法在评估恶性肿瘤患者免疫治疗的反应方面显示出了前景。

表 4 基于人工智能的多组学分析在预测疾病治疗反应中的应用

应用	组学技术	人工智能算法	代表成果	研究来源
治疗疾病反应预测	转录组、基因组、表观基因组和蛋白质组数据	贝叶斯多任务 MKL	经过训练的模型可根据多组学数据预测药物敏感性。	[48]
	蛋白质组学数据	BEMKL	该模型根据蛋白质组数据进行训练，以预测癌细胞对化疗药物和分子靶向抗癌化合物的药物敏感性。	[54]
	组织学数据和临床数据	CNN	预测晚期黑色素瘤的免疫检查点阻断效率。	[52]

近年来，单细胞测序正在成为指导疾病精准治疗的有力工具。越来越多的研究表明，单细胞测序是剖析疾病内异质性、克隆进化、肿瘤微环境、肿瘤转移的潜在机制以及对疾病相关疗法的反应的有效方法[55]。随着单细胞组学时代的出现，人工智能将进一步推动基于单细胞组学的精准医学化。

5 结语及展望

人工智能方法已成功解决多组学数据的复杂性和异质性，并广泛应用于临床。尽管基于人工智能的多组学分析已经取得了很大进展，但其在精准医疗中的应用仍面临许多挑战。如高质量、完整化配对样本的数据稀缺、多模态数据的高度异质、人工智能模型的低解释性。目前研究人员仍在努力解决这些问题并促进多组学分析的临床转化。随着在临床实践中越来越多地产生多组学数据，我们肯定人工智能技术将提供有意义的信息来指导临床决策，最终推动精准医学管理的创新。

PU-246

欧洲 ELIXIR 生物医学数据工具服务平台体系实践与启示

刘盛宇 胡拯涌 段一凡 钱庆 吴思竹

中国医学科学院医学信息研究所

近年来，生物医学研究中高通量技术的广泛应用导致了基因组、核酸序列和蛋白结构等生物医学数据的急剧增长。面对这些高维度且结构复杂的数据，传统基于人工筛选和初级智能算法的分析手段遭遇重大挑战[错误!未找到引用源。]。因此，研究人员迫切需要高效处理各类数据并从中提取生物医学意义的专业数据分析工具。值得关注的是，国际上如欧洲生命科学基础设施中心（European Life-Science Infrastructure for Biological Information，ELIXIR）和美国国家生物技术信息中心（National Center for Biotechnology Information，NCBI）等数据中心已经提供了此类工具服务。

尤其是 ELIXIR，其在工具注册、管理以及容器化运行环境的建设上已经形成了成熟体系，为我国解决生物医学数据处理中的问题如工具短缺、管理体系不完善、标准化不足等提供了宝贵参考。鉴于此，本文深入探讨 ELIXIR 工具平台的实践经验，并为我国相应平台的建设提供建议。

1. ELIXIR 生物医学数据工具服务平台体系建设概述

2013 年 12 月，ELIXIR 正式成立。这是一个欧洲跨政府的生物医学资源整合机构，由欧洲分子生物学实验室（European Molecular Biology Laboratory，EMBL）、21 个成员国和两个候选国（塞浦路斯和奥地利）共同组成[错误!未找到引用源。]。为了响应生物医学数据的分析、整合科研需求，ELIXIR 构建了完整的生物医学数据工具服务平台体系，建立了工具标准化注册、运行环境管理、工作流集成和工具验证等系列工具子平台，形成工具收、管、用体系。整个工具平台的研究和服务是在 Horizon 2020 和 Innovative Medicines Initiative（IMI）项目资金支持下开展建设[错误!未找到引用源。]。

ELIXIR 生物医学数据工具服务平台体系由 bio.tools、WorkflowHub、BioContainers、UseGalaxy 和 OpenEBench 五个子平台组成。表 1 列出体系中的五个子平台的基本信息。其中，bio.tools 主要实现工具的标准化注册和检索发现，已注册 7247 个工具。WorkflowHub 致力于生物医学数据工作流的注册与共享，整合了 300 余个工作流，这些工作流定义了顺序化的数据处理步骤，支持完成特定分析任务。BioContainers 为工具提供标准化运行环境，覆盖了 8982 个工具。UseGalaxy 是工具和工作流的运行平台，涵盖 7851 个工具和工作流，bio.tools、WorkflowHub 与 Galaxy 之间已实现工具和工作流的双向整合和执行功能，bio.tools 和 WorkflowHub 的部分工具和工作流能在 Galaxy 中执行，而 Galaxy 的部分产出也可以共享存储到 bio.tools 和 WorkflowHub 中。最后，OpenEBench 负责多维度的工具性能评估，记录了 23748 个工具。

技术研发方面，尽管五个子平台共同采纳 HTML、Shell 和 Python 开发语言，但也结合了各自特有的开发语言以实现特定功能。bio.tools 采用 XSLT 语言来支持 XML 格式的工具文件注册；BioContainers 使用 Perl 和 Makefile 语言来完成工具运行环境容器自动化构建；Jinja 和 TeX 语言用于 UseGalaxy 中工作流各代码模块生成和格式化；而 OpenEBench 利用了 Facebook 研发的 Hack 语言，来满足多线程编程需求。平台服务方面，各子平台均有其开源协议选择，表 2 列出生物医学数据工具服务平台中各子平台的服务详情。bio.tools 和 OpenEBench 选择 CC BY 4.0 协议，注重工具共享；WorkflowHub 遵从 BSD 3-Clause 协议，主张结构化代码复用；BioContainers 选择 APACHE LICENSE, VERSION 2.0 协议，强调完善版权和免责声明；UseGalaxy 则使用 MIT License 协议，最大限度的保证研究人员二次开发的自由度。API 设计方面，bio.tools、UseGalaxy 和 OpenEBench 各自开发专属 API，以满足特定功能需求，从而实现更加精细化的功能支持。相比之下，WorkflowHub 和 BioContainers 侧重接口通用性，因此选择如 JSON API、Restful API 和 GA4GH 等公共的标准 API，用于促进平台间的融通和整合。

表 1 各平台属性比较

属性	bio.tools	WorkflowHub	BioContainers	UseGalaxy（欧洲服务器）	OpenEBench
诞生年代	2014	2020	2015	2019	2017
工具	7247	>300	8982	7851	23748

量					
开发语言 (特有)	XSLT	无	Dockerfil、Perl 和 Makefile	Jinja 和 TeX	Hack
许可证	CC BY 4.0	BSD 3-Clause 许可证领衔的 83 个许可证	APACHE LICENSE, VERSION 2.0	The MIT License	CC BY 4.0
API 接口	bio.tools Web API	1. JSON API 2. GA4GH 的 TRS API	1. GA4GH Tool Discovery API 2. Restful API	Galaxy API	1. OpenEBench 工具监控 API 2. OpenEBenchScientificAPI

2. 子平台架构和主要功能特点

2.1. 生物医学数据工具注册平台—bio.tools

bio.tools 是一个生物医学数据工具的注册和检索平台，提供规范化工具注册，存储了丰富的工具元数据描述，包括工具的唯一标识、基本信息、特征信息、功能信息、使用信息和相关学术资料等。平台支持研究人员通过 biotoolsID、工具名称等主要元数据项进行工具检索和发现[错误!未找到引用源。-错误!未找到引用源。]。

bio.tools 为研究人员提供工具的注册和汇交通道。当上传自研工具时，需详细描述其基本信息、输入输出格式、数据类型和 biotoolsID 等元数据信息。biotoolsSchema 数据模型采用了 18 个结构化的生物医学相关词表，确保元数据信息的标准描述。此数据模型基于生物医学数据分析和管理（EMBRACE Data And Methods，EDAM）本体框架，实现工具属性的细粒度描述。EDAM 本体由 Jon Ison 等人设计，为工具分类提供了层次化结构，包含首选术语、同义词、相关术语及其定义和其它信息[错误!未找到引用源。-错误!未找到引用源。]。具体细分为五类：“主题”描述应用领域（如“蛋白质组学”）；“操作”定义功能（如“肽识别”）；“数据”明确输入输出类型；“biotoolsID”标识工具（如“biotools:signal”）；“格式”描述支持的数据格式（如“Thermo RAW”）。通过元数据标准的使用保障了 bio.tools 中工具描述信息的标准化集成和高效管理。图 1 绘制了 biotools 的服务模式：其中，1A 是研究人员获取 bio.tools 工具服务流程，而 1B 描述基于 biotoolsSchema 数据模型的工具注册流程。

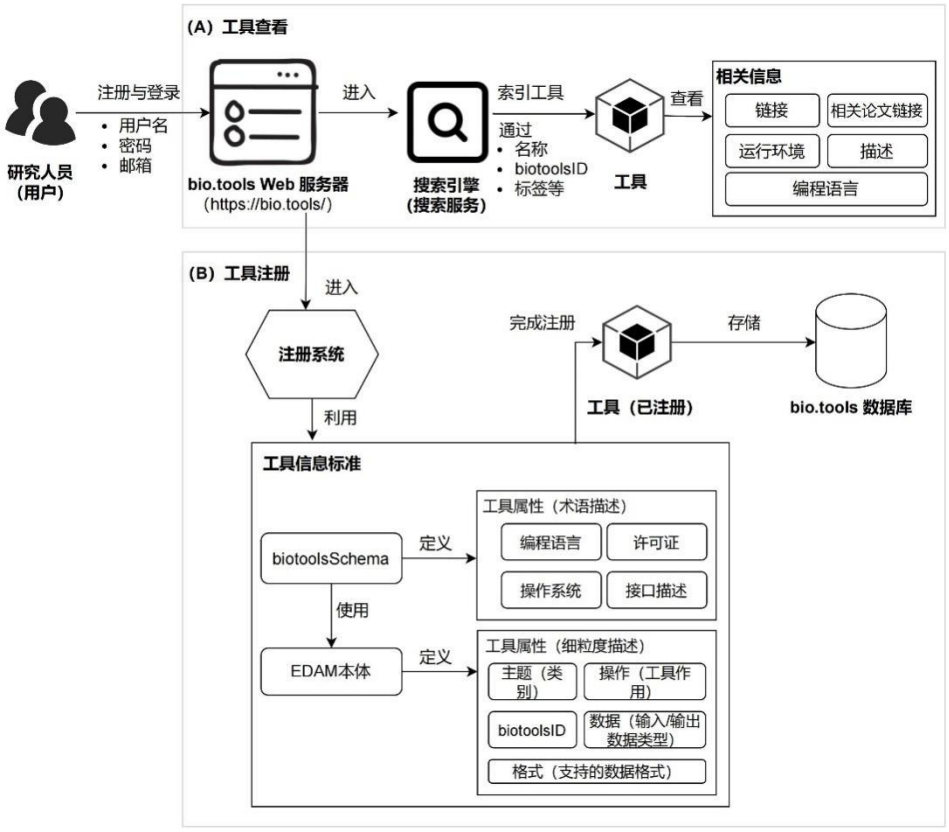


图 1：bio.tools 平台服务模式图

2.2. 生物医学工作流资源集成平台—WorkflowHub

WorkflowHub 是针对生物医学数据工作流的注册和发现平台，不仅提供丰富的工作流资源，还支持按照创建时间、名称、工作流类型（如 Common Workflow Language、Galaxy、KNIME、Nextflow 和 Snakemake 等）和发布团队等关键词进行多维度检索。研究人员可以通过平台检索所需工作流并了解工作流相关源文件、资金来源和详细描述文件（类似于 README.md 文件）等详细信息。WorkflowHub 支持研究人员注册集成工作流，其提供了一整套完善的注册机制，包括填写工作流元数据信息、上传文件（支持本地和远程链接的方式）等并提供多次审核质控。每个工作流都会被赋予一个 DOI 标识用于唯一标识和定位该工作流，有助于工作流的传播和引用。除了由研究人员的主动上传，WorkflowHub 的管理团队也积极地从 myExperiment（共享科学工作流和实验设计的社交网站）、nf-core（生物医学工作流的开源数据库）、UseGalaxy 和 GitHub 等权威渠道筛选并整合工作流，确保其在数据库中的工作流都经过了严格的验证，并可在 UseGalaxy 平台上复现[错误!未找到引用源。,错误!未找到引用源。]。

为便于研究人员并行注册多个工作流，WorkflowHub 平台搭建了核心工作流标准化框架，并严格遵循 RO-Crate、Common Workflow Language（CWL）和 Bioschemas 三个数据标准模型，从而确保工作流信息的结构化、规范化，并在 Web 服务器中提升其检索效率。其中，RO-Crate 负责将各工作流的文件、目录结构、数据类型和其它相关描述整合进 JSON-LD 的元数据文件；CWL 明确规定了工作流的输入输出数据格式，并规划任务执行顺序；而 Bioschemas 则标记了工作流 Web 端的 28 个关键属性，以增强工作流在线可查找性[错误!未找到引用源。]。通过以上标准化模型技术，WorkflowHub 为研究人员提供了一个高效、规范的工作流注册平台。图 2 展示了 WorkflowHub 的主要功能和服务模式，2A 是基于 Bioschemas 标准的工作流检索服务，2B 揭示了依赖 RO-Crate 和 CWL 的多工作流并行注册机制。

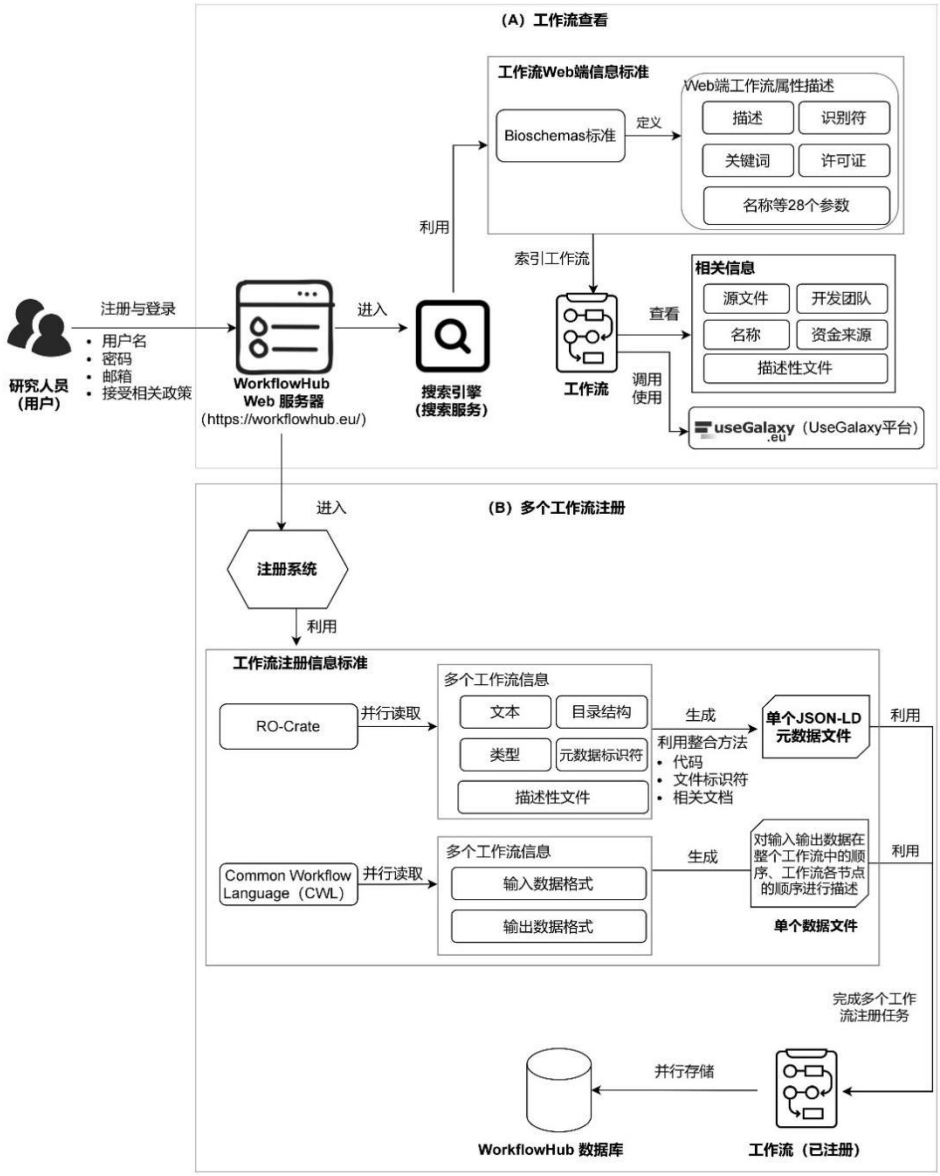


图 2：WorkflowHub 平台服务模式图

2.3. 生物医学数据工具可执行环境容器化平台—BioContainers

BioContainers 使用容器化技术为生物医学数据工具提供了标准化运行环境，确保工具在不同系统中的稳定性。该平台包括三个关键组件—Docker Containers、Conda Containers 和 Specifications。Docker Containers 组件通过 Dockerfile 脚本自动配置工具的运行环境。Conda Containers 组件采用两步构建方法：首先使用 Conda 脚本生成 Conda 包，而后基于此构建 Docker Containers 组件，以保证容器的统一性和可追溯性。Specifications 组件明确了容器的相关规则，如许可证[错误!未找到引用源。]。

BioContainers 为研究人员提供 BioContainers Registry 作为 Web 端服务器。这个服务器索引了各工具及其容器镜像，使得研究人员能够基于工具名称、类别或开发团队等信息进行精确筛选和定位[错误!未找到引用源。]。每个工具都配有详细的描述文件（类似于 README.md 文件），其中包括工具名称、版本、功能描述和操作指南。有了这些详细的指南，研究人员可以轻松地在本地执行工具，避免了复杂的环境配置步骤。

考虑到研究人员可能有自定义需求，BioContainers 也允许研究人员构建特定工具的运行环境容器，允许其选择适当的容器化技术并编写相应的配置脚本。BioContainers 积极鼓励社区成员参与，并由社区贡献者进行审核和验证，通过验证后，自定义容器会被上传到 BioContainers 的核心数据库，并在 BioContainers Registry 中检索和使用。此外，当工具的容器需要更新时，BioContainers 支持研究人员通过“容器请求界面”提交相关需求。社区贡献者会对这些需求进行评估和审查。审查通过后，相关的更新会被纳入核心数据库，供后续检索和使用[错误!未找到引用源。]。图 3 展示了 BioContainers 的完整组件，从研究人员的初步请求到容器的部署阶段。

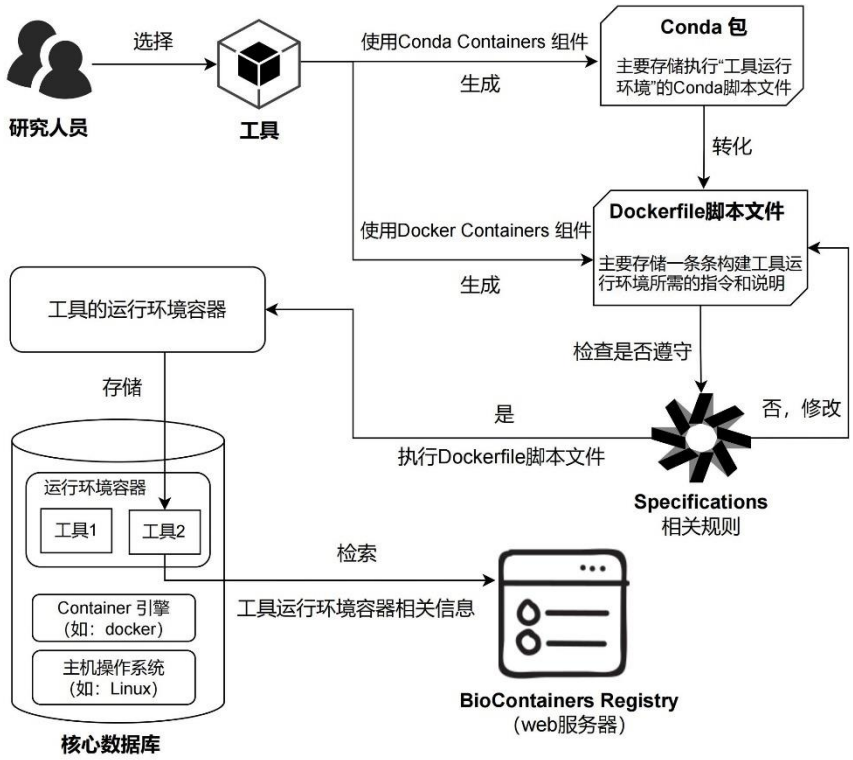


图 3：BioContainers 运行环境容器架构图

2.4. 生物医学数据工具/工作流分析平台—UseGalaxy

UseGalaxy 是开源的工具和工作流（本章节统称为工具）集成分析平台，为非编程背景的研究人员提供了一系列精准且可复现的数据分析工具。它支持团队协作和数据结果的共享。研究人员可以通过使用 UseGalaxy 集成的各类工具，完成多种生物医学分析任务，如 DNA/RNA 序列比对、变异检测、基因组注释、基因表达量分析、蛋白质交互网络分析和代谢途径模拟等。平台还允许研究人员创建、保存和根据项目需求调整自定义分析工作流。

UseGalaxy 平台由五大核心组件构建，包括公共 UseGalaxy 服务器、GalaxyToolShed 存储库、UseGalaxy 云服务模块、培训服务模块和 UseGalaxy 社区。公共 UseGalaxy 服务器采用并行计算策略和高效的工作流调度算法，满足高通量数据处理需求。GalaxyToolShed 存储库作为持续更新的工具和算法集合，囊括了材料科学、机器学习和系统生物学等领域的分析工具。为确保工具和工作流的高质量，该库设立了严格的审核标准，涵盖代码效率、API 标准化、功能完整性和稳定性，保证和公共 UseGalaxy 服务器的无缝对接。UseGalaxy 的云服务模块提供弹性计算框架，分别为学术用户和商业用户的不同使用需求，提供了学术云和商业云两类计算资源。学术云主要为研究和教育提供服务，而商业云为企业大数据分析提供增强能力。UseGalaxy 通过培训服务模块为研

究人员提供专门的教学内容，确保他们能有效利用平台。UseGalaxy 通过社区为研究人员、开发人员和教育工作者等更广泛的用户提供交流平台，并举办通过论坛、交流会和年会分享和宣传推广最新研究和技术成果[错误!未找到引用源。,错误!未找到引用源。]。图 4 描述了 UseGalaxy 的技术架构和应用流程。

UseGalaxy 遵循用户需求导向的技术更新策略，及时将最新的生物医学数据分析技术结合到平台服务中。为提高数据处理的灵活性，平台支持数据类型的批处理转换，同时集成了多项关于数据流开发的建议，从调整输入到处理不连续数据[错误!未找到引用源。]。平台引进 Galaxy-ML 支持全套的监督机器学习工具，包括数据标准化、特征选择、模型定义、超参数优化和交叉验证[错误!未找到引用源。]。结合最新的生成式技术，UseGalaxy 引入了 OpenAI GPT 辅助工具为研究人员遇到的编程问题提供实时的问答解决方案。

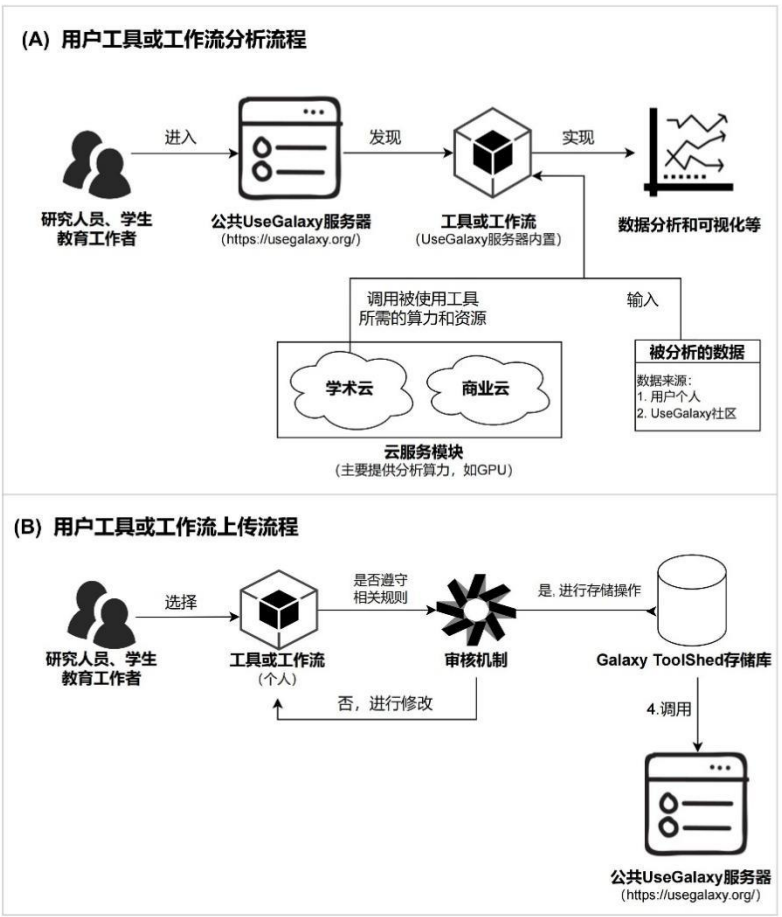


图 4: UseGalaxy 技术架构和应用流程图

2.5. 生物医学数据工具性能评估平台—OpenEBench

OpenEBench 是 ELIXIR 工具服务平台体系中的重要一环，由巴塞罗那超级计算中心（Barcelona Supercomputing Center，BSC）和 ELIXIR 及其合作伙伴联合推出，是一个专门针对生物医学数据工具和工作流的标准化评估平台。在生物医学领域，基准评估为工具性能提供了客观标准。因此，OpenEBench 主要为研究人员提供了一种机制，使用存放于 ENA（European Nucleotide Archive）和 EGA（European Genome-phenome Archive）数据库的基准测试数据集进行工具的测试和评估[错误!未找到引用源。]。

OpenEBench 的工具评估流程可概括为“预测结果生成”和“基准测试”两大阶段。在“预测结果生成”阶段，先将选定的基准测试数据集以及工具导入至数据分析平台（如 UseGalaxy 或

Nextflow）。此后，通过工具在数据分析平台中对基准测试数据集进行的处理，进行预测性比较，生成“预测结果数据”支持后续评估。

“基准测试”阶段包括三个关键步骤：（1）验证阶段（2）指标计算阶段（3）工具评估阶段。首先，验证阶段的目标是检验“预测结果数据”的完整性和准确性，并生成“验证结果数据”。其次，指标计算阶段会根据既定基准深入对比“预测结果数据”，进而识别并产出如匹配精度等关键性能指标，并形成“评估结果数据”。最后，工具评估阶段将前两步得到的数据将汇集到工具评估专题数据库中，生成可视化分析图表，呈现最终结果。此整合过程在虚拟研究环境中进行，以确保所有数据都经过统一且严格的评估标准。

OpenEBench 的工具评估环节涉及初、中、高三个参与层级的社区成员，其中每个层级拥有不同的权限。初级参与者主要负责基准测试结果的存储，确保实验的可重复性和数据来源的明确性。中级参与者除了拥有存储功能外，还能运用基准测试工作流评估其他成员的工具性能，这包括使用参考数据集来计算评估指标。高级参与者则拥有平台的全部权限，能够直接在 OpenEBench 中启动整个工具评估流程。每个层级的功能都是基于其前一层级进行扩展，例如，高级参与者生成的数据将按照中级层级的评估指标进行处理，并依照初级层级的数据结构进行保存[错误!未找到引用源。-错误!未找到引用源。]。具体的工具评估流程详见图 5。

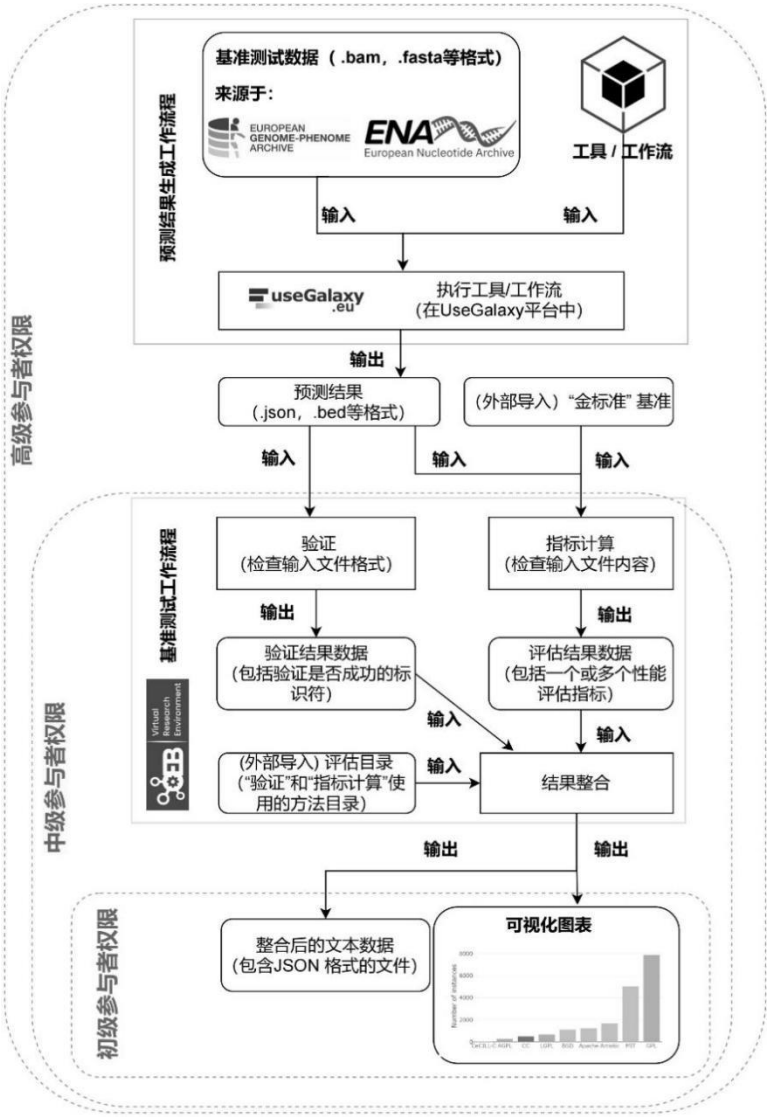


图 5：OpenEBench 工具评估流程图

3. 子平台之间和外部平台的交互式模式

在 ELIXIR 的生物医学数据工具服务平台架构中，精心设计的策略确保了其自身及其与其它平台的高效交互，打造了一个高度协同的环境。图 6 展示了数据工具服务平台的五个子平台之间以及和计算平台、培训平台的交互。

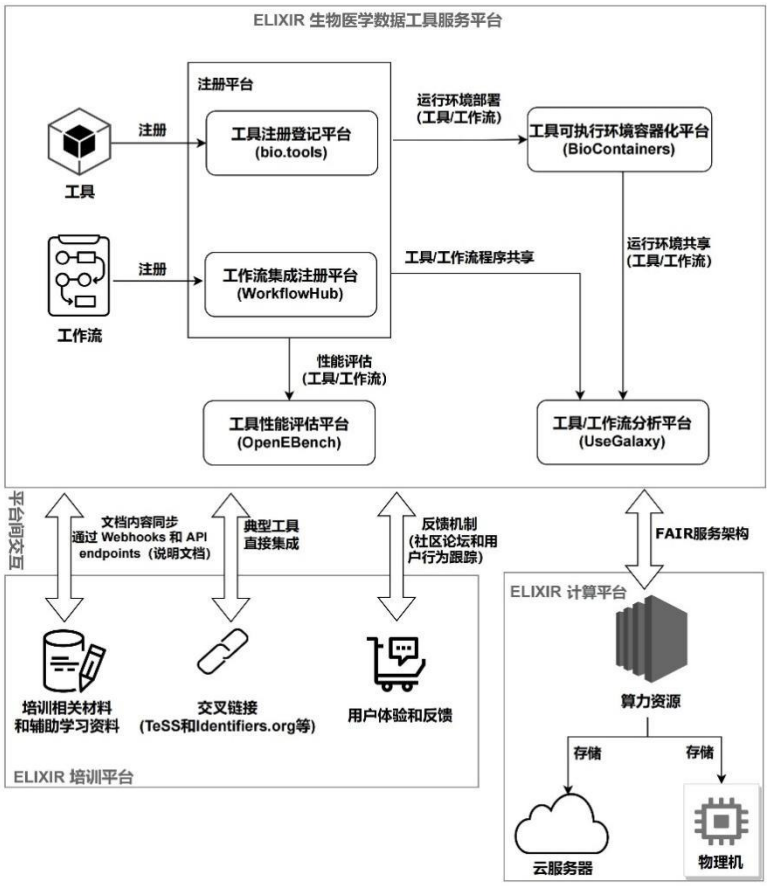


图 6: ELIXIR 数据工具服务平台体系内部交互及外部交互图

在资源方面，各子平台遵循统一的协同共享策略。特定的工具和工作流首先被注册到 bio.tools 或 WorkflowHub 数据库。接着，BioContainers 为这些工具和工作流提供稳定、适用于多平台的容器化运行环境。完成配置后，这些工具和工作流在 UseGalaxy 平台上直接调用并执行。执行结束后，其性能会在 OpenEBench 平台受到评估。各子平台共同构建了一个高效的资源共享网络，确保工具和工作流在生物医学数据分析任务中具备完整性和互操作性。

除五个子平台之间形成有序的生物医学数据工具服务网络外，ELIXIR 还建立了高性能计算平台与培训平台。高性能计算平台支持大规模基因组分析和蛋白质结构模拟等复杂任务。培训平台则致力于指导如何高效地应用分析工具与数据处理技术。工具服务平台与计算平台基于 FAIR 服务架构交互，推动标准化数据接口的建立，并确保数据符合 FAIR 模型特征（可发现性、可访问性、互操作性和可重用性）。高速数据通道技术进一步提高了与计算平台的数据传输效率。培训平台为工具服务平台提供教学材料和服务，并通过 Webhooks 和 API endpoints 定时更新工具与工作流信

息。反馈机制则有助于优化工具服务平台的资源与服务。培训平台还集成了工具服务平台的典型应用，使学者在实践中能直接操作这些典型工具和工作流[错误!未找到引用源。]。

4. 对我国生物医学领域数据工具服务平台建设的启示

本文对 ELIXIR 生物医学数据工具服务平台和五个子平台进行了描述和分析，该体系有效地满足了生物医学数据工具的注册、描述、管理、集成、测试和使用等方面的需求。这对我国生物医学领域数据利用和数据工具平台建设具有重要启示，主要体现在以下方面。

4.1. 注重工具服务平台生态体系建设

ELIXIR 生物医学工具服务平台和其五个子平台—bio.tools、WorkflowHub、BioContainers、UseGalaxy 和 OpenEBench，综合打造了一个全方位、互补的生物医学数据工具服务生态体系。这些子平台实现了工具和工作流的统一注册、索引和资源集成，为研究人员提供了迅速、便捷工具获取和发现功能；汇集多种高效分析工具，为研究人员打造了综合、便捷的数据工具服务平台，通过先进的容器化技术提供按需配置的统一工具的部署和运行流程，显著减少了各工具间的兼容性问题；实现了对工具功能、性能等的系统评估和验证，确保了工具可用性、有效性和可靠性。这些子平台的建设系统性地形成了工具创建注册、汇集整合、安装运行、检测应用的全链条服务体系，并且规范了工具管理和服务流程，对于我国生物医学领域面临的工具碎片化、开发冗余和缺乏评估机制等问题具有重要的借鉴作用。

4.2. 建立健全工具管理标准规范

ELIXIR 五个子工具平台建设均注重遵循统一标准，不仅基于 BioSchema 结合领域提出了具有针对性的核心数据模型、工具格式和利用了领域本体进行表达规范控制，确保了数据一致性和工具间的互操作性。这种规范化策略不仅简化了数据管理流程，同时也增强了工具的适应性和跨平台、跨实验室的数据共享能力。目前，我国生物医学数据工具和平台建设存在缺乏遵循标准的情况，因此，应深入研究并借鉴 ELIXIR 的经验，加强我国的生物医学数据和工具标准规范的研制和应用，包括引进国际标准和建立健全我国特色的生物医学工具管理标准规范，确保工具的高效和规范化应用，从而提升生物医学科研的效率和促进科研成果产出。

4.3. 统筹工具运行环境管理

由于存储资源和算力资源是有限的，将存储和计算资源进行整合，建立运行环境的统筹管理策略能够更有助于提升资源的有效利用。BioContainers 通过采纳如 Dockerfile 的标准化描述文件，结合 Docker 和 Singularity 等先进的容器技术，确保了工具在多样化的计算环境中都保持高度的可移植性和一致性。这种技术不仅简化了工具的部署和配置过程，还大大减少了因环境差异导致的运行错误。鉴于我国生物医学数据的迅速增长及其对高密度数据科研的迫切需求，实践 BioContainers 的容器化方法具有重要意义。它不仅能有效处理大量生物医学数据，而且有助于构建统一、高效的工具运行环境架构。

4.4. 强化工具基准测试和验证

生物医学领域中有大量数据处理和分析工具，但工具的质量和性能参差不齐会影响数据处理或分析的结果的准确性和可信性。OpenEBench 的基准测试框架为生物医学工具提供了一套完备的评估体系，确保工具性能的科学、准确分析。该框架不仅涵盖了多种评估指标，而且整合了丰富的基准测试数据集，确保评估的深度和广度。这种综合性和结构化的评估方法不仅提高了工具的信任度，还为研发人员提供了性能优化的方向。面对生物医学数据工具的复杂性和多样性，考虑到我国

研究领域的特色和需求，我们可以在 OpenEBench 的评估策略基础上进一步完善、定制评估体系，确保其更符合国内实际，推动我国生物医学研究的持续健康发展。

4.5. 重视技术跟进和融合

ELIXIR 工具平台能够实现对工具和工作流的版本管理，并对技术进行持续追踪和整合采取了一套系统化的策略，它不仅持续关注前沿技术动态，包括引进 OpenAI 的 ChatGPT 和 EOSC Hub 的混合云科学服务（结合本地私有云、国家社区云、公共/商业云和欧洲创新研究云）等最新技术，将新技术集成和融合到已有功能中，而且对引进的技术进行严格的科学评估和实验验证。通过集成其它开源工具平台的优质工具和工作流，扩展可用、高质量的工具资源。不断的更新和资源扩展促使该平台的工具能够持续满足科研人员的使用需求。在我国的生物医学数据工具和平台建设中也应关注最新技术发展，不断改进和优化工具和系统平台的性能，融入新的智能化技术和资源，从而帮助科研人员更好地开展创新性研究。

4.6. 用户社区驱动创新发展

ELIXIR 工具平台突出的特点就是采用用户社区驱动的发展策略，建设中强调了用户需求和工具开发的同步性。其建立不同子平台的用户社区，并深入挖掘社区的意见和需求，确保工具与实际科研场景紧密结合。他们非常重视和社区用户的参与，让用户参与信息审核、工具检测、使用测评，并积极获取用户的反馈，不断进行工具产品迭代和优化，从而极大地提高了工具的可用性和易用性。此外，通过定期组织学术研讨、会议和多种形式的培训，积极为研究人员、开发人员等各类用户搭建技术与知识互动的桥梁。我国的生物医学数据工具和平台建设时，也应高度重视用户社区的参与性，建立健全的交流平台和意见建议反馈通道，以此推动满足中国生物医学领域科研用户使用需求的工具产品。

5. 结语

本文全面地探讨了欧洲 ELIXIR 生物医学数据工具服务平台体系，并对其在工具注册标准化、运行环境容器化、分析架构以和性能评估等领域的成果进行了系统性论述。通过参照其经验，为我国生物医学数据工具服务平台的建设和发展提供了启示和建议。然而，考虑到 ELIXIR 工具服务平台体系的丰富内容以及笔者的研究能力和学科背景，本文在内容完整性和深入分析方面可能存在短板。这将是笔者未来研究中需持续努力和完善的方向。

PU-247

基于研究者视角的生物医学转录组数据库发展现状和建议

李晨龙[1,4]，王琪[1,4]，刘格良[3,4]，张俣翔[2]，于琦[3,4]，贺培凤[3,4*]

1.山西医科大学基础医学院

2.山西医科大学第二临床医学院

3.山西医科大学管理学院

4.临床决策研究大数据山西省重点实验室

1 引言

生物医学数据是一类与医学和生物学相关的数据资源，这些数据涵盖了人类健康、疾病的发病机制、治疗方法以及生物多样性等方面的信息。生物医学数据的收集、存储和分析对于推动医学和生命科学的研究和创新至关重要。这些数据可以来自基因组学、转录组学、蛋白质组学和其他领域，为科研人员提供了深入研究人类健康和疾病的机会。生物医学数据的不断积累和广泛分享对于推动医学科学的进展和改善人类生活质量具有重要意义[1]。

转录组学数据库在生物医学数据领域担当着至关重要的角色[2]。这些数据库承载了大量的基因表达数据，为科学家和研究人员提供了宝贵的资源，以深入研究基因的表达方式和模式。这些数据库不仅推动了基因功能解析、疾病研究、药物开发和个性化医学等领域的发展，还促进了全球生物医学领域的国际合作和知识共享[3]。作为生物医学数据的重要组成部分，转录组学数据库的贡献不可低估，它们在推动医学科学进步和创新方面发挥着关键作用。

国际上有许多重要的转录组学数据库，它们致力于收集、存储和提供基因的表达信息以及与基因表达相关的数据。其中最富盛名的有 TCGA（The Cancer Genome Atlas）和 GEO（Gene Expression Omnibus）。这些数据库汇集了来自不同实验和研究的大量基因表达数据，覆盖多种生物学领域，包括基因功能注释、调控网络、组织特异性表达、疾病相关基因等。这些数据库的开放性和互操作性促进了全球科学合作，有助于更广泛地分享和利用生物医学数据[4; 5]。TCGA 和 GEO 数据库已经在全球范围内确立了它们作为公开可获得的生物医学研究数据的领先地位。这两个数据库的权威性和广泛性使它们成为研究人员的首选数据资源。

GSA（Genome Sequence Archive）和 CNSA（CNGB Sequence Archive）是国内研发成熟的转录组学数据库，相比 TCGA 数据库和 GEO 数据库，它们的数据存储更加规范，结构更加合理[6; 7]。然而，由于它们相对较晚诞生，尚未得到广泛的应用和认知。

目前，无论是国内还是国外的转录组学研究，主要集中在美洲和欧洲人群的数据资源，这些数据主要来自于国际上的 TCGA 和 GEO 数据库。由于数据库的局限，对中国人群的转录组学研究受到了限制，同时也导致了中国人口的转录组学数据未得到充分利用。因此，发展国内的转录组学数据库变得至关重要，这将有助于推动针对中国人群的转录组学研究，避免浪费宝贵的中国人口数据资源，并且有望为中国的医学研究和进步提供更坚实的支持。

本研究以研究者的观点为基础，对国际和国内的转录组学数据库特征进行深入分析，探讨这些数据库的优点和不足之处，旨在为国内的转录组学数据库的未来发展提供建议和指导。

2 国际转录组数据库特点分析

2.1 TCGA 数据库

TCGA 数据库是一个重要的癌症基因组学数据库。它是一个国际性的合作项目，旨在收集、存储和分享多种类型癌症的分子特征数据以及与癌症患者相关的临床信息。TCGA 的目标是通过深入研究癌症的基因组学特征，帮助科学家更好地了解癌症的发病机制，以及发展更有效的癌症治疗方法[8]。TCGA 数据库具有以下特点：

1、癌症数据覆盖全：数据库几乎涵盖了人体所有主要器官的癌症样本数据，这意味着研究者可以在这一资源中找到来自不同部位的癌症病例的数据，为跨器官的比较和研究提供了丰富的信息基础。

2、数据质量高：TCGA 数据库的卓越之处在于其数据源严格来自官方合作项目，经过精细筛选和高质量控制，确保了数据的卓越质量。这一特性为研究者提供了可靠的数据基础，可用于各种生物信息学和生物医学研究，确保了研究的科学性和准确性。

3、数据规范：TCGA 数据库以其统一规范的数据结构、丰富而完整的数据内容以及详尽的临床信息而脱颖而出。这一特点极大地方便了研究者的数据获取和分析工作。研究者可以轻松地访问多种数据类型，从而为他们的科学研究提供了充分的素材和信息资源。

4、数据开放获取：TCGA 数据库坚持免费开放原则，任何研究者都可以自由地获取所需的数据，无需支付任何费用或面临限制。这一开放的数据政策促进了科学研究的广泛参与和合作。

然而，尽管 TCGA 数据库在癌症研究领域表现卓越，提供了丰富的官方合作测序数据，但需要强调的是，该数据库的应用范围相对有限，因为它主要关注癌症领域。这意味着除了癌症之外的其他疾病，如心血管疾病、自身免疫性疾病等，未在 TCGA 中找到广泛的测序信息支持。此外，尽管 TCGA 的癌症样本数量庞大，但每种癌症类型的样本量是固定的，这可能不足以满足某些研究的需要，特别是在需要更大样本规模以应对癌症异质性和多样性时。因此，对于广泛研究多种疾病或需要更多样本数据的项目，TCGA 数据库可能存在一定的限制。

2.2 GEO 数据库

GEO 数据库是由美国国立生物技术信息中心创建和维护的一项重要生物信息学资源。它是一个全球性的生物信息数据库，致力于收集、整理和共享高通量分子生物学实验数据，尤其是基因表达数据[5]。GEO 数据库具有以下特点：

1、数据多样：数据库涵盖了多种癌症类型以及各种其他疾病的不同组织样本的测序数据。更令人印象深刻的是，每一位研究者都有权将自己的测序数据上传至 GEO 数据库，从而使得数据库中的数据种类非常丰富。这意味着对于相同类型的疾病，您可以找到在不同实验条件下获得的多个数据集，每个数据集都代表了一种特定的实验设置或病例群体。这种丰富的异质性使得 GEO 数据库成为了一个独一无二的资源，为研究者提供了大量的选择和比较的机会，有助于更全面、深入地探索各种疾病的分子特征和机制。

2、用户友好的界面： GEO 提供了直观且易于操作的在线界面，旨在让用户能够轻松地访问和利用其中的数据资源。此外，GEO 数据库还内嵌了 GEO2R 在线分析工具，这个工具使研究者能够方便地查看研究样本的临床信息，同时还可以在线对数据进行各种分析。更为重要的是，

GEO 还提供了相应的 R 语言分析代码，使用户能够深入探索和定制分析过程，这对于需要特定数据处理和统计方法的研究项目非常有帮助。

3、数据开放共享：和 TCGA 数据库一样的是，GEO 的数据是开放共享的，任何研究者都可以自由访问和下载数据。这一特点促进了科学研究的合作和数据重复使用，有助于加速科学进展。

然而，尽管 GEO 数据库拥有多样的数据类型和用户友好的界面，但数据的规范性和质量面临着巨大的挑战。由于 GEO 允许每位研究者自行上传其研究数据，这导致了数据库中存在一定程度的数据不一致性和质量差异。有些数据可能由于上传者提供的格式不规范或未提供足够的样本信息而难以被其他研究者有效利用。更为严重的情况是，某些数据可能存在上传错误或与其描述不符的问题，这可能会给研究社区带来混淆和误导。因此，在使用 GEO 数据时，研究者需要谨慎验证和评估数据的质量和准确性，以确保其科研工作的可靠性。

3 国内转录组数据库发展现状

GSA 是中国国家基因组科学数据中心的一个数据库，其主要任务是存储和管理生物学研究中产生的原始 DNA 序列数据。作为一个全球性的资源，GSA 旨在为科研社区提供一个集中的平台，让研究人员能够方便地存储、共享和获取公开可用的 DNA 序列信息。此外，为了保证数据传输的质量和稳定性，GSA 提供了自主开发的软件工具：Edge turbo，这一工具由国家生物信息中心自主研发。Edge turbo 支持 Linux 版本，为数据传输提供了强大的保障。该软件的使用可确保数据传输的高效性和安全性，从而满足了研究人员对数据完整性和可用性的高要求[6; 9]。

CNSA 是中国国家基因库旗下的数据库，用于存档组学数据，包括原始测序数据和其进一步分析的结果。它以项目、样本、实验、运行、组装和变异等六种数据对象的方式组织这些数据。CNSA 采用了全球生命科学领域常用的数据标准，致力于构建一个全面且经过策划的数据存储库。研究者可以通过 CNSA 进行数据的存储、管理、检索和引用，以促进组学数据的有效利用。值得一提的是，CNSA 创建了一个关联模型，将生物样本、样本信息和分析数据关联在一起。这意味着研究者可以从生物样本出发，追踪数据的完整生命周期，从生物样本到样本信息再到分析数据[7; 10]。

GSA 和 CNSA 都具有以下特点：

1、数据存储规范：在 GSA 中存储的转录组表达数据非常规范，格式也非常整齐。

2、数据安全隐私保证：它提供了受控访问机制，研究人员必须提交访问受保护数据的申请，并经过审核批准后才能获取受控数据。这种安全措施有助于确保数据的机密性和隐私性，防止未经授权的访问和使用。

尽管 GSA 和 CNSA 数据库自成立以来，在数据结构方面表现出了高度规范和合规性，优于其他数据库，但目前必须承认其数据存储容量相对较小，无论是在数据类型的多样性还是样本数量方面，它们需要未来的不断扩展和发展，这是一个长期的过程。此外，GSA 数据库中存储的转录组数据是原始的测序数据，未经过下游数据的处理和加工。这意味着研究人员需要投入相当大的时间和精力来学习和掌握相关的基础知识和技术，以便有效地处理这些数据。因此，对于研究者而言，使用 GSA 数据库可能需要具备较高的技术门槛和数据处理能力。

4 国内转录组数据库发展建议

尽管国外转录组数据库在发展和应用上取得了显著成就，但它们仍然存在一些不足之处。最重要的问题是，这些数据库缺乏针对中国人群的转录组学数据。这个问题对于中国的研究者来说尤为突出，因为缺乏中国人群的相关数据限制了他们进行针对中国人群的转录组学分析和研究。这也可能导致一些医学研究在适用于中国人群方面的滞后和不足。因此，发展拥有大规模中国人群转录组数据的国内数据库对于促进中国的医学研究和发展至关重要。这样的数据库将为中国医学科研提供宝贵的资源，有助于加速医学领域的进步，更好地满足中国人群的特性。虑到国际和国内转录组学数据库的演进趋势以及科研界的需求，本研究提出以下几项专业建议，以促进国内转录组学数据库的发展：

4.1 提供高质量数据

为了确保数据库能够有支撑性的典型数据，我们必须确保其中包括了常见癌症和其他备受关注的疾病的转录组测序数据，这些数据应该是通过官方合作采用了最先进和成熟的技术进行测序的。尽管在数据样本数量上不需要追求过多，但必须确保每一个样本都具有高度的代表性，以满足科研需求。

4.2 自由上传且数据规范

用户自行上传的数据可作为典型数据的有益补充，提升数据库的数据丰富性和多元性。然而，为了充分发挥这一潜力，允许用户自由上传的同时，必须采取严格的审核措施，以确保数据的规范性和可用性。因此，必须确保每份数据的存储格式方便研究者使用，数据描述信息也必须充分准确，以便为研究者提供切实可用的数据资源。这一举措将有助于提高数据库的整体质量和可信度，使其更好地满足研究者的需求。

4.3 用户友好的界面

一个清晰简洁的操作界面是至关重要的，它可以显著降低研究者上手的难度，从而提高数据的利用效率。因此，数据库必须具备用户友好的界面，以确保用户能够轻松地进行数据检索和分析。

4.4 提供内置工具

数据库应该提供一些便捷的内置工具，以辅助研究者更好地解读和分析数据，从而进一步提高数据的利用效率。这些工具可以帮助用户在不需要离开数据库界面的情况下进行基本的数据处理和分析，为他们节省时间和精力。这种用户体验的改进将促使更多的研究者积极利用数据库中的数据资源，从而推动科学研究的发展。

4.5 继续坚持数据开放获取且提供受控访问

数据的开放获取是实现数据最大化利用的前提条件。只有当更多的研究者能够自由地访问和使用这些数据时，才能够充分挖掘其中潜在的、尚未被发现的信息和知识。同时，应该坚持提供受控访问功能，以满足研究者对上传的特殊数据的需求。

5 结论

本研究通过对国际和国内的四个代表性转录组学数据库进行详细分析，考察其特征和优缺点，为国内的转录组学数据库发展提出了五项建议，分别是：提供高品质数据、支持自由上传并遵循数据规范、设计用户友好的界面、整合内置工具、继续坚持实行数据的开放获取并同时提供受控访问机制。

PU-248

健康医疗大数据的应用与治理

温丹 党雅娜

开封市人民医院病案统计科

摘要：健康医疗大数据在治理领域的应用正在成为一种趋势。本文摘要将介绍健康医疗大数据在治理方面的重要性和应用领域，探讨其对社会治理和公共健康的影响，并讨论可能面临的挑战和隐私问题。

第一部分将介绍健康医疗大数据在治理中的重要性。随着社会的发展和人口的老龄化，健康医疗问题已经成为重要的社会治理议题。而大数据的应用可以提供全面、准确、实时的健康信息，有助于政府和相关部门制定更有效的政策和干预措施。同时，大数据还可以用于监测和预测疫情、评估社区健康状况以及改善医疗资源配置，为社会治理提供重要的决策依据。

第二部分将探讨健康医疗大数据在治理中的应用领域。健康医疗大数据在治理中的应用非常广泛，包括公共卫生管理、医疗资源调配、医疗质量监管等方面。通过分析大量的健康医疗数据，政府和相关部门可以更好地了解社会的健康状况和需求，提供个性化的医疗服务，并制定更具针对性的公共卫生政策。此外，大数据还可以帮助政府和医疗机构合理分配医疗资源，提高医疗服务的效率和质量，从而提升社会治理的水平。

第三部分将讨论健康医疗大数据应用于治理可能面临的挑战和隐私问题。由于医疗数据的敏感性和隐私性，保护个人隐私和数据安全是健康医疗大数据应用于治理时的重要问题。政府和相关部门需要制定相应的法律法规和隐私保护政策，确保数据的安全性和合法性。同时，由于医疗数据的多样性和复杂性，数据整合和标准化也是一个挑战，需要采取相应的措施来解决。

综上所述，健康医疗大数据在治理中的应用具有重要的意义和潜力，可以为社会治理和公共健康的提升做出贡献。然而，其应用仍面临着一些挑战和隐私问题，需要政府、医疗机构和社会各方共同努力来解决。未来，随着技术的进一步发展和政策的规范，健康医疗大数据将为社会治理带来更多的机遇和挑战。

PU-249

大数据在驱动智能精准多学科协作中的实践

杨谨菲 左秀然 郭涵怡*

华中科技大学同济医学院附属武汉中心医院

0 引言

2021 年 6 月 4 日国务院办公厅发布《国务院办公厅关于推动公立医院高质量发展的意见》，文件中指出：一是引领公立医院高质量发展新趋势，推进医疗服模式创新，推广多学科诊疗模式，强化信息化支撑作用，推动云计算、大数据等新一代信息技术与医疗服务深度融合，大力发展远程医疗和互联网诊疗。二是要建设公立医院高质量发展新文化，强化患者需求导向，为人民群众提供安全、适宜、优质、高效的医疗卫生服务[1]。

1 现状分析

目前患者就医形式以专科就诊为主，而很多病种涉及多专科，患者奔波于多个专科，得到的治疗方案不统一，体验感差，互动性弱，治疗方案也主要依靠专科医生经验，且多次就诊记录依靠患者主动携带病历资料，就诊完成后医生无法回顾性地干预需要复诊的患者，也无法多维度统计和分析。

2 项目目标

通过数据分析，约 80%的专病患者散在其他专科或体检，多学科联合诊疗可将散在其他科室（包括门诊/住院/体检）的患者进行集中管理；对入组患者进行长期的诊疗管理，随时查看患者历次就诊记录，且可通过微信公众号或者 AI 语音电话与患者互动，患者在移动端填写问卷，所有互动数据汇集至数据中心，根据不同检查检验数据及患者互动数据，通过大数据分析，可对患者进行分组管理。改变以往患者走出医院即石沉大海模式，所有患者病情按照不同级别管理，根据既定诊疗路径实现双向提醒（医生端/患者端）。

从一个病种探索一种全病程管理模式，逐步推广至多个病种。医疗服务转被动为主动，信息化提供全面支撑，多学科联合诊疗给患者提供一站式医疗服务，省去多科室东奔西走的烦恼；避免多学科诊疗方式不统一给患者带来困扰，针对病种，建立规范分组路径，以患者为中心，全程诊疗数据展示及分析；提高患者参与度，让患者全程参与诊疗计划，知晓全病程,诊后提供主动随访/复诊/上门服务，全病程跟踪诊疗。

3 医疗服务模式创新

3.1 建立多学科联合诊疗团队

从呼吸病肺结节病种入手，如图 1，首先组建多学科联合诊疗的团队，以分管领导作为首席专家，牵头科室作为组长并配备工作秘书，与肺结节病种相关的几个专科纳入团队，启动时至少包含有：呼吸科、胸外科、肿瘤科、影像介入科、病理科、核医学科，团队里专家以副高及以上作为纳入条件。每个角色各有分工，首席专家负责运营决策，组长负责人员排班管理及诊疗决策，工作秘书负责联系各专科及患者沟通事宜。在门诊设置肺结节 MDT 专家门诊，该门诊至少三位以上专家共同坐诊，在门诊设立 MDT 诊室，该诊室以会议室的形式，家庭会客的模式，设有大屏供专家为患者及家属讲解病情，同时设有团队照片墙，营造温馨气氛，让患者体验家庭式的医患沟通。[2-3]

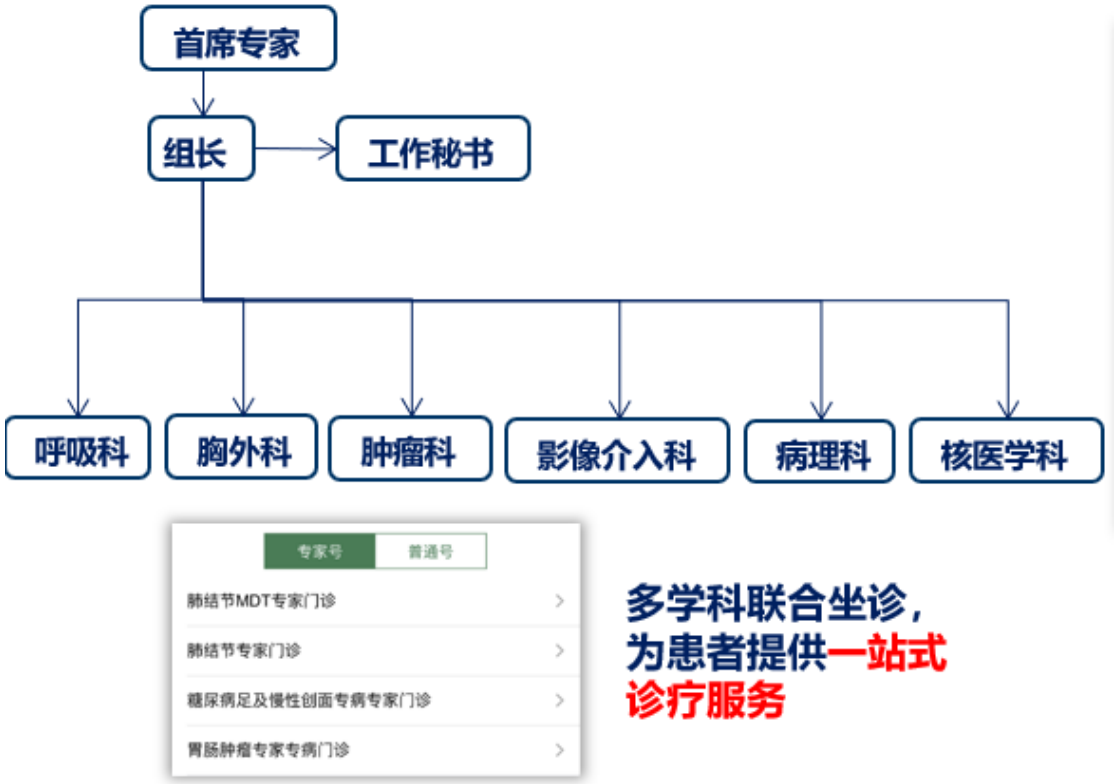


图 1

3.2 规范诊疗路径及就诊流程

基于以上专家团队，对肺结节病种进行归纳总结，形成院内专家共识，通过影像结果对肺结节大小进行分类和大小分级，对患者进行分组，如图 2，不同组别制定不同的诊疗路径和规范，屏蔽不同医师的经验性判断，形成规范化的多学科统一诊疗路径。

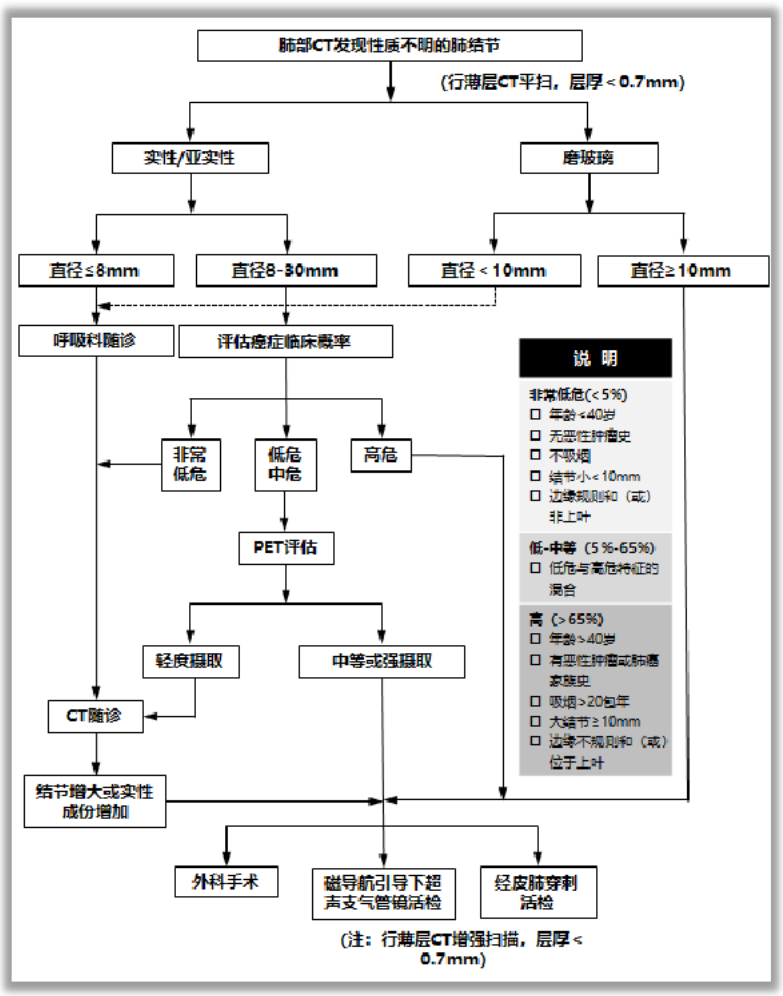


图 2

如图 3，对诊疗过程制定规范的就诊流程，患者来源主要有门诊/住院/体检，门诊和体检的阳性患者直接进入肺结节全病程管理，住院患者经系统消息提醒至工作秘书，工作秘书联系相应科室，将患者纳入肺结节全病程管理团队。入组后，对患者进行检查结果分析和风险评估，按照图 2 的规范进入路径，当患者本次住院结束后，依然在管理团队的管理下，进行长期的跟踪随访管理，定期复查，及时干预。[4-6]

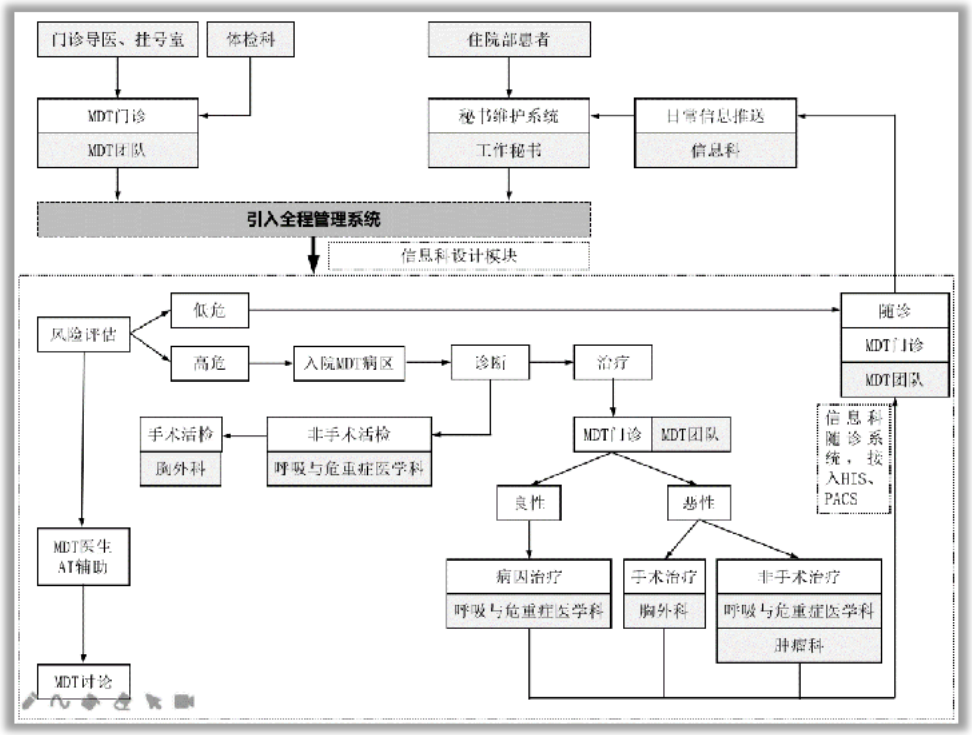


图 3

4 数据服务能力提升

从数据汇集到数据治理，再到数据利用，让数据多跑路，让患者少跑腿，及时发现潜在风险患者的同时，专家团队及时干预，并实现互联网医疗线上线下一体化，通过物联网健康管理、智能随访全病程管理等手段达到患者服务精细化。通过运营决策驾驶舱、DRG 精细化管理等手段达到运营决策科学化。[7-10]

4.1 运行原理

如图 4，首先，把院内各大系统的数据进行汇集，包括医嘱、病历、影像、病理、超声等数据全部汇集至数据中心；二是，对这些数据进行处理，通过创建患者主索引，把每一份数据纳入患者档案的一部分，形成以患者为中心的个人电子健康档案；三是，通过不同检查结果的值域判断和组合值域判断，建立“患者筛选器”，该筛选器可根据组合值域判断筛选患者，比如肺结节大于多少，检验值高于多少则进入一号队列；四是，把专家团队形成的诊疗路径按照标准化的指标利用患者筛选器形成不同队列，比如实性结节大于 8mm 队列。根据不同队列的诊疗路径，可对进入专病组后的患者，进行长期的跟踪管理。

图 4

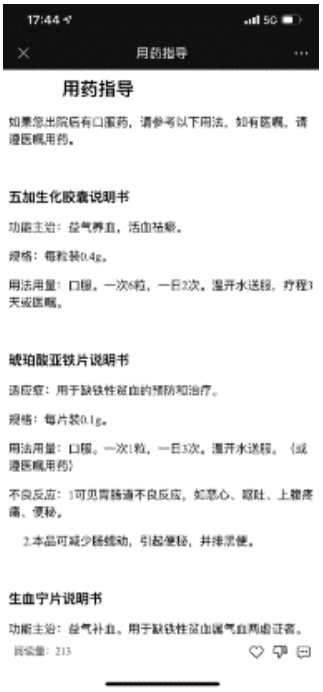
4.2 电子病历系统智能化

预制模板、套餐，辅助高效系统操作，把医生的时间还给患者，系统内置电子病历结构化模板 7416 个、医嘱套餐 13.5 万个、检验检查申请模板 4.1 万个、诊断模板 8506 个，在医生高效工作的同时，提高数据质量，结构化标准化的数据更利于大数据分析和精准定位。引入人工智能影像辅助诊断，改变人工诊断 40 分钟，实现 AI 辅助诊断 5 分钟，将医疗质控前置，做到智能规则预警，避免漏诊。

提高诊疗服务的连续性及可及性，肺结节 CT 报告已然结构化，当影像诊断为肺结节时，检查报告上将打印二维码，扫码可直接链接到肺结节专家团队介绍并提供挂号链接，让医疗服务更加可及。

4.3 “互联网+”与医疗健康服务深度融合

通过智能化手段结合“互联网+”应用，实现 80%就诊流程移动化，平均排队次数由 10 次减少至 4 次、平均候诊时长由 120 分钟减少至 35 分钟；宣传、检查报告链接指引、患者参与 MDT 病例讨论，专家团队实时查看全程诊疗记录，并预设下一次随访/复诊时间，诊后提供互联网医院全程服务，包括 16 项互联网在线医疗服务，包括上门护理/康复、送药到家等。从医生体验上可视化全病程跟踪查看，制定计划，预制提醒；从患者感受上智能语音外呼、个性化图文宣教、音视频宣教、随访等手段，让患者参与度更高。



5 MDT 成果成效

血糖控制影响着外科术后恢复及愈后情况， 围手术期血糖控制显得至关重要，为了降低手术风险，我院于 2021 年 3 月上线全院血糖管理平台。约 80%的糖尿病患者散在其他专科或体检科，根据糖尿病管理团队的专家共识，形成规范的诊疗路径，对标准化的数据通过“患者筛选器”筛查出具有糖尿病高风险的患者，并发送提醒给糖尿病专家管理团队，专家团队甚至比外科管床医生更早获悉患者的血糖风险，通过动态血糖监测、药物控制、健康宣教等方式及时对他科患者进行干预，保证围手术期的血糖控制。同时，将该类患者纳入我院糖友会进行长期管理，从外科出院后转化至糖尿病专家管理团队，对血糖进行长期的控制，效果甚佳。[11-14]

随着肺结节及血糖管理的应用成功，我院逐步上线了共计 11 个病种的多学科联合诊疗，成果如下：

2021 年 01 月，肺结节线上线下一体化专病全病程管理，共筛查出 31784 人次，并将 1400 余名患者纳入管理；

2021 年 03 月，全院血糖管理平台共筛查出近 14 个月数据共 143876 人次，5863 位患者纳入管理；

2021 年 03 月，VTE 高危患者人群专病，共筛查出近 14 个月数据共计 19278 人次；

2021 年 04 月，颈动脉狭窄 MDT，共筛查出近 10 年数据，30 余万人次；

2021 年 05 月，房颤专病管理，共筛查出近 12 个月数据共 2423 人次；

2021 年 06 月，慢性淋巴细胞白血病，共筛查出近 1 年数据共 12 人次；血液科检验危急值管理，11 月份至今共计 1782 人次；

2021 年 08 月，丙肝和乙肝，共筛查出近 10 个月数据共 8319 人次；

2021 年 09 月，胃肠道肿瘤 MDT，共筛查出近 9 个月数据共 11243 人次；

2021 年 09 月，无痛医院建设，共筛查出近 9 个月数据 9980 人次；

2021 年 09 月，体检专科推荐共计 19598 人次；

2017 年 01 月，营养多学科管理，共筛查出 6 年的数据共 79949 人次纳入管理。

6 项目总结

大数据驱动智能精准多学科协作信息化实践实现 “三个升级”，一是让就医体验升级，从诊前-诊中-诊后提高患者参与度，一站式医疗服务提升患者归属感，规范化诊疗路径减少患者就医时间周期，提高患者对多专科诊疗方案的信任度，对疾病的早期干预减少了患者总体就医费用。二是让临床服务升级，多学科联合诊疗提升医院整体诊疗效率、让诊疗服务更加标准化和精准化。三是让医院管理升级，科学决策及智能管控，把 DRG 支付等规则纳入大数据，进一步为管理层提供决策支持。

PU-250

基于条件对抗生成网络的数字病理图像标准化方法

柴振新[1] 田雨[1] 邓博洋[1] 叶前呈[1] 周天舒[2] 章琦[3] 梁廷波[3] 李劲松[1,2]

1. 浙江大学

2. 之江实验室

3. 浙江大学医学院附属第一医院

目的 不同医疗结构在制备病理切片时，使用不同的制备程序、染色剂和扫描仪等，导致不同 医疗机构制备的数字病理图像存在数据异质性。本研究旨在采用基于条件对抗生成网络的数字病理图像标准化方法，以削弱数据异质性，使得基于单一医疗机构数据集训练的深度学习模型可以在其他医 疗结构的数据集上取得与单一医疗机构数据集上相近的结果。

方法 本方法将原始数字病理图像块的 灰度图作为限定条件，保留结构信息而去除染色信息，输入到对抗生成网络的生成器中，生成标准化 的数字病理图像块。同时，通过训练判别器来判断输入的图像对是否为生成的数字病理图像对。在训 练过程中，生成器和判别器相互竞争，并进行反复迭代优化。在针对特定的下游任务时，可以使用任 务模型对训练完成的模型进行微调，以更好地适应下游任务。

结果 通过在收集的肝癌数字病理图像 数据集上验证所提出的方法，证明了方法的可行性。通过使用 UQI、PSNR、SSIM 等指标对生成图像 质量进行评估，结果显示在浙江大学医学院附属第一医院的验证集中，生成的图像相比原始数字病理 图像具有较高的 UQI（0.9944±0.0030）、PSNR（34.5752±1.3270dB）和 SSIM（0.9865±0.0167）。

结论 基于条件对抗生成网络的数字病理图像标准化方法生成的标准化数字病理图像质量接近原始数字 病理图像。因此，该方法可以有效完成数字病理图像的标准化任务。

PU-251

“健商”培养对高校学生体育锻炼的影响研究——以江西中医药大学为例

邵哲

江西中医药大学

青少年的体质健康是全民健康的基础，与国家前途、民族希望、中国梦等紧密联系，增强青少年体质、促进青少年健康成长是党和政府高度关心的大事。习近平总书记十分关心青少年的成长和发展，多次在重要场合强调“没有青少年健康成长，国家就没有远大发展”[1]。自党的十八大以来，党和政府对提高青少年体质健康作出一系列重大决策，出台了一系列相关文件，如 2016 年国务院办公厅印发的《关于强化学校体育促进学生身心健康全面发展的意见》、2020 年中共中央办公厅、国务院办公厅印发的《关于全面加强和改进新时代学校体育工作的意见》等，为学校体育教育改革掀起了一波新的高潮。

长期以来，学校体育作为促进青少年健康成长的重要环节，在青少年健康促进方面发挥了重要作用，直接影响着青少年的体育锻炼行为、运动兴趣及终身体育意识的培养[2]。然而，受传统体育教育体制的影响，教师在体育教师中重视体育知识、体育技能的传授，忽视与学生健康行为、健康习惯相关的健康教育，影响了学校体育在促进青少年体质健康方面作用的发挥，造成了青少年体质健康指标呈整体下滑趋势[3]。健商作为一种全新的健康理念，与青少年体质健康之间存在密切联系[4-5]。然而，当前学校体育教育中对健商教育的价值认识尚不够清晰，影响了健商教育的开展。本课题以江西中医药大学为例，通过对在校学校健商状况及体育锻炼状况的调查，探讨健商培养对高校学生体育锻炼的影响，为高校更好的开展健商教育提供一定参考，旨在帮助高校学生提高健康的自我管理能力，培养高校学生健康的健商理念和终身体育观念，促进健康的体育锻炼方式的养成，提升其健康水平。同时，也为高校体育教育改革提供参考。

1 研究对象与方法

1.1 研究对象

以“健商”培养对高校学生体育锻炼的影响为研究对象，以江西中医药大学学生为调查对象，对高校学生“健商”和体育锻炼之间的关系进行研究。

1.2 研究方法

1.2.1 文献资料法

以中国知网（CNKI）、读秀学术搜索和 Web of Science 为主要检索平台，以健商（Health quotient）、健商（Health quotient）并含体育锻炼（physical training exercise）为关键词组合，检索国内外核心期刊上的相关研究成果，了解该领域的研究现状、主要观点，为研究提供相关的理论基础与参考。

1.2.2 问卷调查法

（1）调查问卷的编制

健商是健康商数的简称，这一概念最早由加拿大籍华人医学专家谢华真教授在其出版的《健康商数》一书中提出，是个体所拥有的健康意识、健康知识和健康能力的综合体现[6]。自健商概念提出以后，引起了各国学界的广泛关注，围绕健商的研究成果日益丰富，并涌现出一系列相关的健商调查量表。围绕研究主题与目的，本研究以谢华真教授的健商自测量表为基础，增加受试者基本情况与体育锻炼情况的调查内容，形成用于本研究的健商与体育锻炼调查问卷。

（2）效度与信度检验

根据研究需要，本课题采用验证性因子分析来进行效度检验。在进行验证性因子分析之前，进行 KMO 和 Bartlett 检验，发现 KMO 值为 0.954，KMO 值大于 0.8，研究数据非常适合提取信息（从侧面反应出效度很好）。因子分析的结果显示，共提炼出 23 个公因子，旋转后累积方差解释率为 75.520%，意味着研究项的信息量可以有效的提取出来；结合因子载荷系数可见，所有项目的因子载荷均在 0.6 以上，说明本课题修改后的问卷具有良好的结构效度。

量表的内部一致性信度最常用的评价指标克朗巴赫 α 系数（Cronbach α 系数），通常 Cronbach α 系数高于 0.6 既满足信度要求，本量表总体的 Cronbach α 系数为 0.971，一致性信度很好，量表信度的初步考核得到较为满意的结果。

（3）问卷发放与回收

受新冠肺炎疫情的影响，本研究采用在线问卷的方式，通过在线问卷平台（问卷星：https://www.wjx.cn/）完成调查问卷的发放与回收。

表 1 本次调查样本的人口统计学特征一览（N=467）

	变量	频数	样本比例
年级	一年级	99	21.20%
	二年级	89	19.06%
	三年级	93	19.91%
	四年级	60	12.85%
	硕士在读	102	21.84%
	博士在读	24	5.14%
性别	男	251	53.75%
	女	216	46.25%
专业	工学	59	12.63%
	理学	84	17.99%
	艺术学	51	10.92%
	医学	78	16.7%
	管理学	46	9.85%
	经济学	74	15.85%
	文学	38	8.14%
	应用心理学	27	5.78%
	其它	10	2.14%
户籍	城镇	210	44.97%
	农村	257	55.03%

在正式调查之前，采用 G*Power 3 分析方法对调查样本量进行估算(Faul, Erdfelder, Buchner, & Lang, 2009) [7]，根据中等的效应量（f^2=0.20, Cohen, 1998），0.01 水平的α水平及课题研究的具体情况来计算确定样本量，在确保统计力度分析(Statistical Power Analysis)达到 0.95 的条件下，确定样本量不低于 460 人。从表 2 可以看出，课题组采用分层随机抽样的方式对江西中医药大学 467 名学生进行了调查，满足了样本量要求，受访群体的男女比例比较均衡，学生所在年级及专业分布比较广泛，说明本课题的调查样本具有一定的代表性。

1.2.3 数理统计法

采用统计学软件 SPSS 22.0 对问卷调查调查获得的数据进行统计学处理，包括简单频数分析、交叉分析和相关分析等，为研究提供数据支撑。

1.2.4 逻辑分析法

以数理统计结果为基础，运用分析、综合和归纳等方法，对数据结果行思维加工，由表及里理清和把握健商培养与高校学生体育锻炼的关系。

2 结果与分析

2.1 高校学生“健商”与体育锻炼基本情况分析

2.1.1 高校学生“健商”的基本状况

健商作为近年来新生的健康概念，反映了一个人的健康智慧及其对健康的态度。本课题中所采用的健商量表共包括自我保健、健康知识、生活方式、精神状态和生活技能 5 个维度，其中每个维度 4 个因子，各因子 5 个条目，共计 100 个条目，且每个条目 0-6 分，将 5 个条目总分除以 3 为该维度的健商指数，4 个维度总分除以 4 为该因子的健商指数。健商指数评价满分为 10 分，依据得分高低分为“理想”状态（9～10 分）、“较为理想”状态（6～8.99 分）、“值得警惕”状态（3～5.99 分）以及“糟糕状态”（0～2.99 分）四个等级[8]。

表 2 江西中医药大学学生“健商”状况总体评价结果（N=467）

		样本量	最小值	最大值	均值	标准差
健	自我保	467	3.46	10.0	8.06	1.44
识	健康知	467	2.57	10.0	8.01	1.61
式	生活方	467	2.91	10.0	8.46	1.36
态	精神状	467	4.26	10.0	8.36	1.29
能	生活技	467	3.92	10.0	8.39	1.56
	总评	467	3.42	10.0	8.26	0.98

表 2 为江西中医药大学 467 名大学生健商各维度和整体得分情况，由表可知江西中医药大学学生健商的五个维度得分均在 8～9 分之间，其中生活方式维度的得分最高，为 8.46 分，其次是生活技能维度与精神状态维度，得分分别为 8.39 分与 9.36 分，与之相比自我保健维度与健康知识维度的得分相对较低，分别为 9.06 分与 8.01 分。江西中医药大学学生的整体健商指数为 8.26 分，处于“较为理想”的状态。吕健力（2018）等的研究表明，河北省高校大学生的健商处于“较为理想”的状态，自我保健、生活方式、精神状态处于“较为理想”的状态，而健康知识、生活技能处于“值得警惕”的状态[9]。本课题的调查显示高校学生的整体健商水平及五个维度的得分均处于“较为理想”的状态，分析其原因可能与近些年来高校体育教学改革不断深入及对学生健商教育重视程度的提高有关。近些年来，党和政府为提高青少年体质健康出台了很多政策、措施，指导各地制定确保学生体质健康的文件和指标，学校体育教学改革进入深水区，在体育教学过程中更加重视对学生的健康教育，促进了学生健商指数的提升。

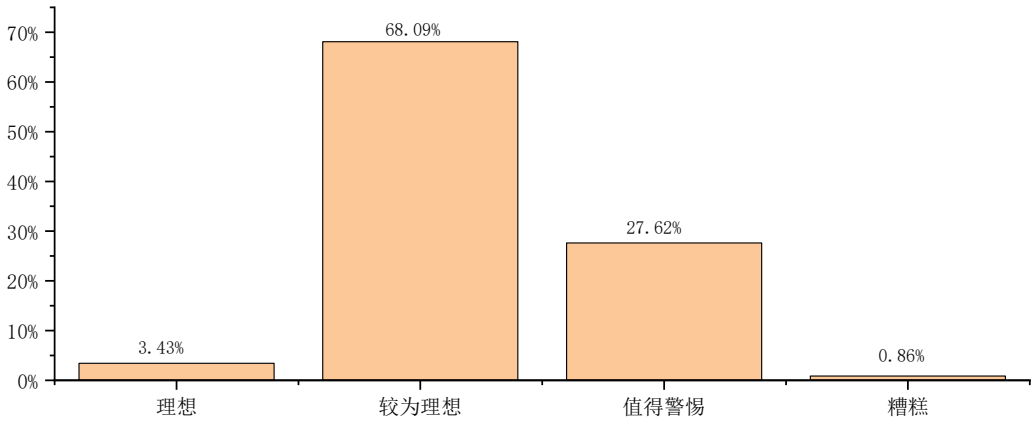


图 1 江西中医药大学学生“健商”水平人数总体分布（N=467）

统计结果显示，江西中医药大学学生健商的五个维度的人数比例整体上呈现为“两头小中间大”的趋势，其中处于“较为理想”水平的占 68.09%，处于“值得警惕”水平的占 27.62%，处于“理想”水平的占 3.43%，处于“糟糕”水平的占 0.86%。可见，在受试样本中大部分学生的健商处于中等水平，但有接近三分之一的学生健商水平处于“警惕”水平，有提高自身健康的需求。

2.1.2 高校学生体育锻炼的基本状况

研究和实践均充分表明，系统的体育锻炼是提高青少年体质健康水平的基本保障，而体育锻炼不足是导致大学生体质健康水平下滑的重要原因。根据研究需要，课题组对江西中医药大学学生的体育锻炼情况进行了调查。

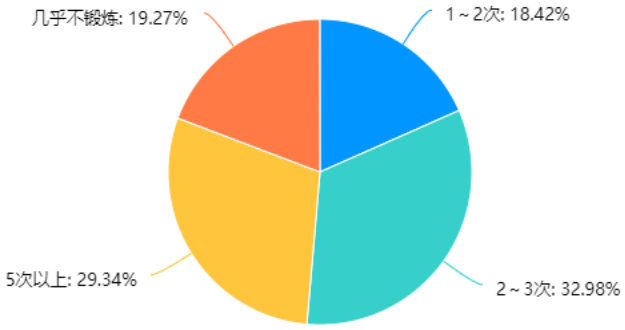


图 2 江西中医药大学学生每周参加体育锻炼的次数一览（N=467）

在调查的江西中医药大学 467 名大学生中，每周坚持参加体育锻炼的占比超过 80%，其中每周参加 2～3 次体育锻炼的比例最高，为 32.98%，每周参加 5 次以上体育锻炼的占比 29.34%，每周参加 1～2 次体育锻炼的占比 18.42%。

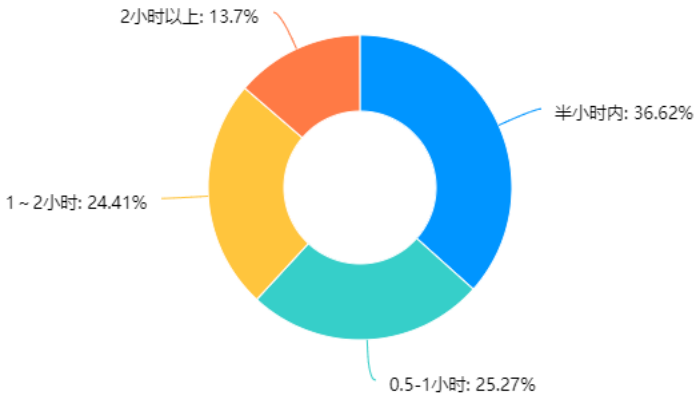


图 3 江西中医药大学学生每次参加体育锻炼的时间一览（N=467）

从江西中医药大学大学生参加体育锻炼的情况看，每次参加体育锻炼在半小时以内的比例最高，为 36.62%，每次参加体育锻炼在 0.5～1.0 小时、1.0～2.0 小时的次之，占比分别为 25.27%、24.41%，每次参加体育锻炼超过 2 小时的占比最低，为 13.7%，不同大学生每次参加体育锻炼持续的时间差异较大。

根据国际上判定体育人口的标准：（1）每周参加体育锻炼频度在 3 次（含 3 次）以上；（2）每次参加体育锻炼的时间在 30 分钟以上；（3）每次体育锻炼的强度在中等强度以上，根据该标准（1）、（2）并结合调查结果可知，在受试群体中达到体育人口标准的占比只有 26.98%，说明虽然参与调查的大部分学生都有一定的体育锻炼习惯，但与体育人口的标准还有一定差异，要想改善自身的体质健康水平，需要增加每周的体育锻炼次数或延长每次的体育锻炼时间。

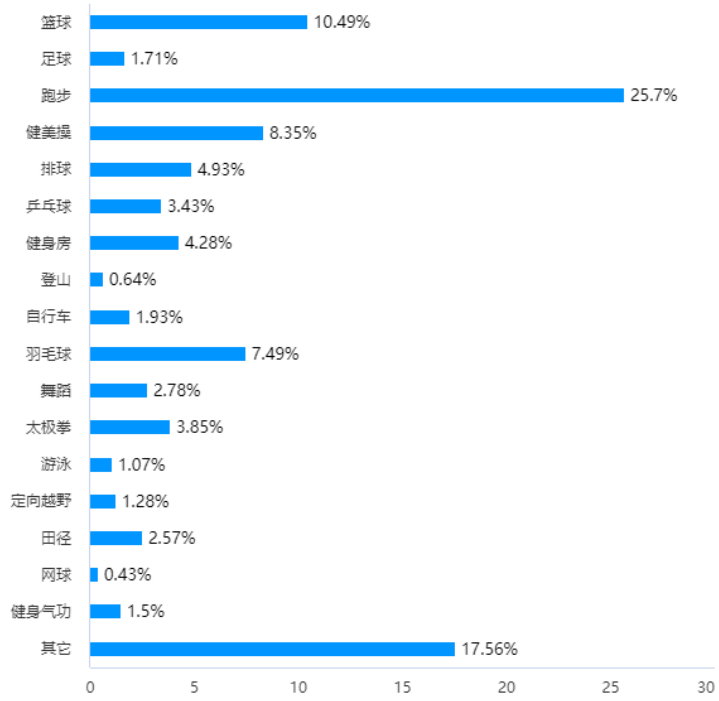


图 4 江西中医药大学学生参加的体育运动项目一览（N=467）

调查显示，高校学生参加的体育运动项目较为多样，集中在篮球、足球、跑步、健美操、排球、乒乓球、健身房、登山、自行车、羽毛球、舞蹈、游泳、太极拳、定向越野、田径、网球、健身气功等项目方面，其中参加跑步的学生比例最高，为 25.7%，其原因是因为跑步对场地设施要求不高，锻炼时间较灵活；其次是篮球、健美操、羽毛球等，选择比例分别为 10.49%、8.35%和 7.49%，与之相比参加游泳、健身气功、定向越野、登山及网球等的学生较少。

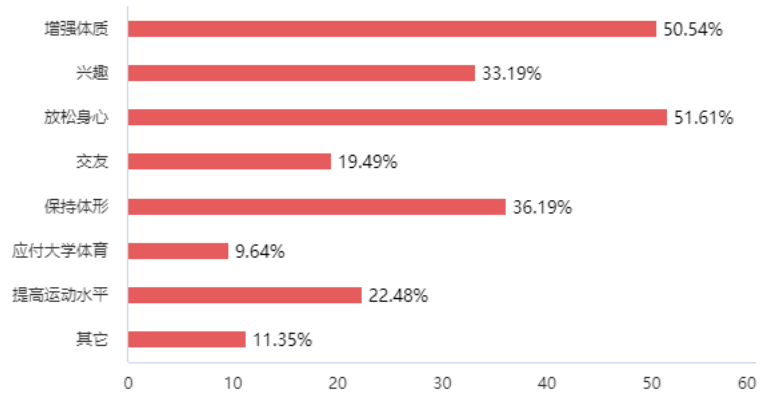


图 5 江西中医药大学学生参加体育锻炼的主要目的一览（N=467）

心理学研究表明，良好的参与动机是大学生积极主动参与体育锻炼的内在动力。如图 5 所示，高校大学生参加体育锻炼的动机较为多元，主要集中在放松身心、增强体质、保持体形、兴趣爱好、提高运动水平及交友等方面，其中出于放松身心、增强体质目的而增加体育锻炼的比例最高，说明高校学生参加体育锻炼的动机较为端正，但部分学生参与体育锻炼存在应付心理。

2.2 高校学生“健商”的异质性分析

2.2.1 不同年级学生健商状况比较

表 3 不同年级大学生健商指数的对比（N=467）

	一年级	二年级	三年级	四年级	硕士	博士
自我保健	8.01±1.04#	8.35±0.98△*	7.95±0.93#	7.99±1.07#	7.98±1.12#	7.97±0.95#
健康知识	8.00±1.10#	8.47±1.09△*	8.01±1.03#	8.02±1.01#	7.99±1.36#	8.01±1.36#
生活方式	8.40±0.99	8.59±0.98	8.51±0.92	8.48±0.93	8.45±1.18	8.44±1.16
精神状态	8.35±0.92	8.41±0.92	8.32±1.13	8.18±0.83	8.32±1.25	8.36±0.99
生活技能	8.35±0.93	8.47±0.88	8.26±0.88	8.48±1.06	8.48±1.36	8.52±1.31
总评	8.22±1.36	8.46±0.97	8.24±1.02	8.23±0.94	8.24±1.27	8.26±1.08

注:#、 △、*分别表示 P<0.05,#表示与二年级比,△表示与三年级比,*表示与四年级比

如表 3 所示，对不同年级学生健商调查结果的多因素方差分析发现，二年级大学生与一年级、三年级、四年级大学生以及硕士、博士在自我保健、健康知识两个维度上得分存在显著性差异（P < 0.05），二年级大学生在这两个维度上的得分显著高于其它年级大学生，其它维度上的得分无显著性差异（P > 0.05）。分析其原因，二年级大学生已经适应了大学的环境与生活，学校的健康教育课程也已经完成，该阶段的大学生已经掌握了基本的健康知识，且随着专业知识的积累和成长经历的丰富，对健康知识的认识逐渐加强。相对于一年级学生而言，这一阶段的大学生已经能很好的适应大学生活；相对于三年级、四年级学生而言，这一阶段的学生没有毕业的压力，课余时间能够通过多种形式、途径参与业余文化生活，拓展自己的知识领域，较为关注自身的健康的保健。然而，在日常教学中观察到，学校较为关注一年级、二年级学生的体育教学，体育课堂内外的教学改革也多集中在一年级、二年级，忽视了高年级学生对课余体育活动的需求，因而相对于一年级、二年级学生，高年级的健商指数有所下降。

2.2.2 不同性别学生健商状况比较

不同性别学生健商状况比较如表 4 所示，可见男生、女生在健商总分上的差异不具有统计学意义（P=0.632 > 0.05）。从不同维度看，男生与女生在生活方式维度上的得分存在显著性差异（P < 0.05），男生在生活方式上的得分明显低于女生，而在其余 4 个维度上的得分均不存在显著性差异（P > 0.05）。

表 4 不同性别大学生健商指数的对比（N=467）

	男生	女生	t	p
自我保健	8.07±1.02	8.03±0.94	0.278	0.781
健康知识	8.08±1.46	7.98±1.24	0.665	0.507
生活方式	8.22±1.34	8.58±0.97	-2.520	0.013
精神状态	8.34±1.33	8.45±1.24	-1.450	0.148
生活技能	8.21±1.37	8.32±1.09	-0.812	0.418
总评	8.14±1.22	8.24±0.91	-0.913	0.362

健商量表中的生活方式维度共 4 个因子，分别是“抽烟、喝酒、吸毒等不良生活习惯”、“饮食、营养与饮食习惯”、“锻炼与健身”、“日常生活”，进一步的统计学检验显示，男生与女生除“日常生活”因子的差异不显著之外，其余 3 个因子的差异均有统计学意义（P < 0.05）。其中，男生在“抽烟、喝酒、吸毒等不良生活习惯”因子上的平均得分只有 5.83 分，处于“值得警惕”状态，结合日常观察发现很大一部分男生有抽烟、喝酒等不良生活习惯；在“饮食、营养与饮食习惯”方面，男生与女生均处于“理想”状态，但女生在该因子上得分显著高于男生，说明女生的饮食、营养与饮食习惯比男生好；在“锻炼与健身”因子上男生处于“理想”状态，女生处于“较为理想”状态与“值得警惕”状态的临界点，说明女生普遍参与体育锻炼的人数较男生存在明显差异。

2.2.3 不同专业学生健商状况比较

表 5 不同专业大学生健商指数的对比（N=467）

	自我保健	健康知识	生活方式	精神状态	生活技能	总评
文学	7.18±0.97△*※	7.17±0.35*※	7.34±0.77△*※	8.15±0.22△#*	7.21±0.15#△*	7.38±0.13△#☆
理工学	7.24±0.49△*※	7.44±0.45*※	7.57±0.55△*※	8.71±0.29△*※	8.31±0.19◇○※	7.85±0.94*
医学	8.51±0.61#*◇○	8.68±0.48#☆◇○	8.45±0.55#☆◇○	8.58±0.51*#※	8.45±0.30◇○※	8.53±0.47#△◇○☆
艺术学	8.42±1.02#◇○	7.64±0.48*※	7.45±0.28☆#※	8.55±0.32#◇○☆	8.27±0.19◇○※	7.92±0.26*
管理学	7.33±0.87△*※	7.52±0.46*※	7.24±0.72△*※	8.22±0.29△#*	7.25±0.44#△*	7.64±0.42*
经济学	7.42±0.94△*※	7.29±0.35*※	7.18±0.27△*※	8.26±0.32△#*	7.27±0.15#△*	7.63±0.64*
心理学	8.37±1.21△*※	8.11±0.45#☆◇○	8.13±0.48#*◇○	8.46±0.57#◇○☆	7.36±0.54#△*	7.88±0.86*

注：*、#、△、◇、○、☆、※分别表示 P＜0.05 ，其中#表示与理工学比，△表示与艺术学比，*表示与医学比，◇表示与管理学比，○表示与经济学比，☆表示与文学比，※表示与心理学比

如表 5 所示，不同专业大学生健商状况差异较大，从健商总分看医学专业学生的健商总分最高，与其它专业学生健商总分均有显著差异，其原因是因为医学专业学生的专业特性决定了他们在大学期间能够学习到更多与健康有关的专业知识，掌握了一些应对、治疗、预防与干预疾病、提升健康水平的技能方法，这些与健康相关的知识与技能随着他们专业学习的深入而潜移默化的渗透到他们的日常生活中，因而其健商水平明显高于其它专业学生。相对而言，其它专业大学生在大学期间接受的健康教育不足，同时在日常学习、课余锻炼中没有意识自觉主动的去学习次方面的知识，故健商水平明显比医学专业学生低。

2.2.4 不同地域学生健商状况比较

表 6 不同地域大学生健商指数的对比（N=467）

	城镇	乡村	t	p
自我保健	8.16±1.04	8.01±0.92	1.009	0.314
健康知识	8.05±1.35	7.89±1.32	0.983	0.325
生活方式	8.58±1.18	8.43±1.12	1.190	0.235
精神状态	8.51±1.26	8.31±1.31	1.119	0.264
生活技能	8.74±1.25	8.27±1.15	2.244	0.026
总评	8.32±1.00	7.21±1.00	1.565	0.118

如表 6 所示，城镇学生的健商总分及各维度得分均高于农村学生，统计学检验显示城镇与农村学生在生活技能上的得分有显著差异（P < 0.05），城镇学生的生活技能得分明显高于农村学生，其它维度得分均无统计学意义。生活技能维度包括生活技能、学习与环境、社会关怀与支持、人际关系技能共 4 个因子，分因子分析发现，城镇学校与农村学生在生活技能因子上的分无显著差异，在学习与环境、社会关怀与支持、人际关系技能 3 个分因子上得分存在显著差异，说明与农村学生相比，城镇学生对自己专业的满意度更高，对环境的适应能力更强，来自社会的支持以及关注感知度更好，在人际交往方面的能力更强。

2.3 “健商”对高校学生体育锻炼的影响分析

随着学界对健商关注和重视程度的日益提升，健商与学生体育锻炼之间关系引起了越来越多学者的关注[10]。课题通过对江西中医药大学学生的调查，通过交叉分析和相关分析来分析健商对高校学生体育锻炼的影响。

表 7 健商平均得分*体育锻炼参与度交叉分析表（N=467）

等级		体育锻炼参与度				
		1～2 次	2～3 次	5 次以上	几乎不锻炼	合计
糟糕	数量	0	0	0	4	4
	百分比	0	0	0	100%	100%
值得警惕	数量	41	33	24	31	129
	百分比	31.78%	25.58%	18.60%	24.03%	100%
较为理想	数量	44	113	106	55	318
	百分比	13.84%	35.53%	33.33%	17.30%	100%
理想	数量	1	8	7	0	16
	百分比	6.25%	50.00%	43.75%	0	100%

如表 7 所示，为江西中医药大学健商平均得分与体育锻炼较差分析表，由该表可以看出处于不同健商状态的大学生其体育锻炼状况明显不同，由此可以得出结论，健商培养对大学生体育锻炼有直接影响，培养健商、提高大学生的健商水平是提升其体育锻炼积极的有效手段。其中，健商水平处于“糟糕”状态的大学生每周几乎不参加体育锻炼；健商水平处于“值得警惕”状态的大学生，每周参加体育锻炼的次数以 1～2 次最多，占比为 31.78%，每周参加 2～3 次体育锻炼的次之，占比为 25.58%，每周参加 5 次以上体育锻炼的最少，占比 18.60%，还有 24.03%的大学生几乎不锻炼；健商水平处于“较为理想”状态的大学生，每周参加 2～3 次体育锻炼的最多，占比为 35.53%，每周参加 5 次以上体育锻炼的次之，占比为 33.33%，每周参加 1～2 次最多体育锻炼的最少，占比 13.84%，其余 17.30%的大学生每周几乎不锻炼；健商水平处于“理想”状态的大学生每周均参加体育锻炼，每周参加 2～3 次体育锻炼的最多，占比为 50%，每周参加 5 次以上体育锻炼的次之，占比为 43.75%，每周参加 5 次以上体育锻炼的大学生最少，占比为 43.75%。比较可见，健商水平处于“理想”状态的大学生均养成了一定的体育锻炼习惯，每周均会坚持参加体育锻炼，而健商水平处于“糟糕”状态的大学生尚没有养成良好的体育锻炼习惯；健商水平处于“较为理想”状态的大学生，每周坚持参加体育锻炼的比例明显高于健商水平处于“值得警惕”状态的大学生，每周参加 2～3 次、5 次以上体育锻炼的比例更高。与之相比，健商水平处于“值得警惕”的大学生，虽然有 75.97%的大学生每周都参加一定次数的体育锻炼，但每周只参加 1～2 次体育锻炼的比例偏高。由此说

明，随着健商平均得分的提高，高校学生体育锻炼的整体水平也随之提高，结果也进一步表明，健商培养与高校学生体育锻炼之间存在积极的影响。

表 8 健商平均得分*体育锻炼参与度交叉分析显著性检验表（N=467）

	相关值	P 值(双尾检验)
皮尔逊卡方检验	125.397[a]	.000
有效变量数	467	

如表 8 所示可见，高校学生健商平均得分与体育锻炼参与交叉检验的相关值为 125.397，P=0.000，说明两者之间具有相关关系。

3 结论与建议

3.1 结论

3.1.1 本研究纳入调查的高校学生健商整体水平处于“较为理想”状态的中等水平，五个维度的人数比例呈现为“两头小中间大”的趋势。

3.1.2 本研究纳入调查的高校学生每周坚持参加体育锻炼的学生占比超过 80%，每次参加体育锻炼在半小时以内的比例最高，参加的体育运动项目较为多样，以跑步为最多，参加体育锻炼的动机较为端正。

3.1.3 高校学生的健商存在异质性，不同年级、不同性别、不同专业及不同地域大学生的健商存在一定差异。其中，二年级大学生的健商水平高于其它年级学生，女生在生活方式分维度上的得分高于男生，医学专业大学生的健商水平高于其它专业大学生，城镇学生的生活技能得分明显高于农村学生。

3.1.4 本研究纳入调查的高校学生健商平均得分与体育锻炼参与度存在相关关系，随着健商平均得分的提高高校学生体育锻炼的整体水平也随之提高，结果表明健商培养与高校学生体育锻炼之间存在积极的影响。

3.2 建议

3.2.1 建议江西中医药大学在今后的体育教学改革中将健商教育的理念融入其中，以健商教育理念为基本思想来指导体育教学改革，优化教学内容，改革教学方法，建立学生健商健身计划，定期对学生的健商进行测试与分析，及时发现健商教育中存在的问题，提高健商教育水平，为体育教学改革助力。

3.2.2 江西中医药大学应开设一些体育保健、健康教育等公共课或选修课, 增加大学生的健康知识，增强他们的健康意识。在学校实施健商教育与健康促进教育时，需要注意性别、年级、专业及学生所在区域的差异，针对学生个性心理发展特点，区别对待与引导教育，这样不仅能起到事半功倍的效果，同时也减少学校健康教育的盲目性。

3.2.3 江西中医药大学在体育理论课中应该加强生活方式、生活技能等健康教育内容，使大学生对于健商有更全面深入的认识，充分认识到体育运动的重要性，提高学校体育与健康教育的教学质量。

PU-252

基于知识图谱的全科未分化疾病小样本辅助推理方法研究

郑奕骁[1] 田雨[1] 刘强华[1] 周天舒[2] 尚勇[2] 刘颖[3] 任菁菁[3] 李劲松[1]

1. 浙江大学

2. 之江实验室

3. 浙江大学医学院附属第一医院

目的 全科医学是一门综合性医学专业学科，目的在于向社会提供全面的医疗保健服务。未分化疾病即通过病人描述的临床症状及相关的理化检查仍无法确诊的疾病。早期未分化的疾病可能会恶化

为严重疾病。全科未分化疾病的未能及时诊断，会导致医疗资源的浪费，和医疗成本的增加。现实中的全科医生队伍人数不足并且综合质量有待提高，且现实中全科数据集的大部分疾病具有小样本的特点。因此研究一种小样本条件下有效的全科疾病辅助推理方法辅助全科医生进行临床诊断，提高全科医学效率。

方法 本研究通过手动构建和自动构建的方式构建知识图谱，通过知识图谱计算症状的邻接矩阵，设计引入自注意力机制的图卷积神经网络学习节点特征向量，训练使用了对比学习中三元组训练的方法使向量空间上同类样本接近异类样本远离，最后计算疾病的原型表示。预测时根据输入样本与各个疾病原型表示的欧氏距离给出辅助推理结果。

结果 在来自浙江大学医学院附属第一医院消化内科以及急诊科室门诊记录的数据集上，该方法达到了 0.9196 的准确率。

结论 本文的研究结合了知识图谱和对比学习、原型网络的方法和思想，为全科未分化疾病的辅助诊断提供了一种全面的解决方案。在基层全科医疗条件不完备的情况下能有效节省医疗资源，高效给出可靠的疾病推理结果，改善全科医疗服务的质量和效率。

PU-253

体数据的自由投影与平行切割体素绘制实现

李心舟，李晨，王彤，季达峰
南通大学医学院

医学影像体数据是对检测结构一种完全化的信息呈现，在图像配准上，可提供更多的特征信息[1]。在图像分割上，可实现体据内的空间生长分割，减少信息的丢失[2]。体素作为体数据最基本的单元，在算法加持下，可实现超体素的目标结构分割[3]，体数据中的特征场往往是基于目标结构的组织特征，在切割时可体现出不同的组织结构[4]。目前对体数据的切割主要以兴趣框(Region of Interesting, ROI)调节为主，ROI 在调节时呈方块状切割体数据显示部分并有助于图像特征的提取[5]，当目标结构呈现不规则突出时，无法满足显示要求。与主流的光线渲染相比，体素绘制虽然消耗太多的内存，但其对体素信息的还原度要高于光线渲染，可基于二值掩模零化法实现切割。本研究首先利用自由二维切割函数形成投影掩模，对三维体数据进行零化，再利用体素绘制对零化后的体数据进行可视化，以探讨自由二维切割函数与体素绘制个性化显示体数据的方法。

1 资料与方法

1.1 CT 数据资料

腰椎 CT 数据 1 例，图像分辨率 512pix512pi，空间分辨率 0.2773mmx0.2773mmx1mm，扫描范围 1~5 腰椎，轴位扫描。

1.2 软件

图像预处理：3DSlicer5.2.3（www.slicer.org, open source, Havard Medical Shool Univertisy, USA）；

平台软件：MATLABR2016a（Mathworks Comp.USA）；

集成软件：微创骨科三维定位系统 V2.0（Dis3Ds,南通百影数字科技有限公司,中国）。

1.3 方法

1.3.1 数据预处理 将数据导入 3DSlicer 中，分离出其中所需的连续序列，通过 convert 模块导出为一个 DICOM（Digital Imaging Communication of Medcine）序列。

1.3.2 分辨率转换 将上述 DICOM 序列导入 MATLAB 中，利用文件头信息实现图像分辨率向空间分辨率转化（公式 1）。

$$R_{space}=R_{image}*R_{perpixel} \quad （公式 1）$$

其中 R_{space} 为物理分辨率，R_{image} 为图像分辨率，$R_{perpixel}$ 为单像素代表的空间分辨率。

1.3.3 非零像素提取 利用 MATLAB 中 find 函数实现每张图像的非零像素提取。

提取后将体数据内的非零灰度像素坐标转化为三维度数据，同时保存灰度值作为绘制颜色，图 1 展示了体素信息的转化过程。

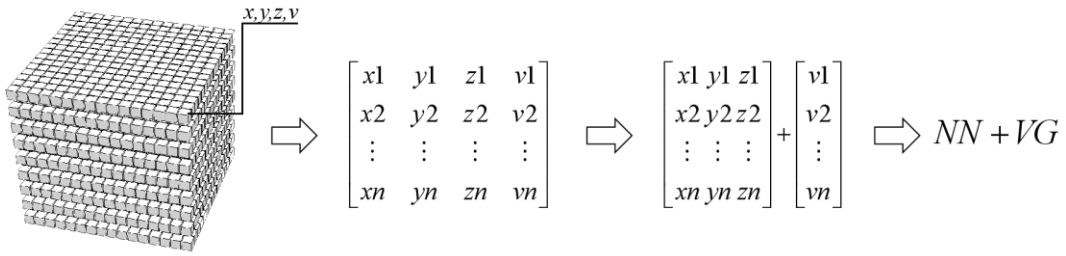


图 1 体素信息转化过程

1.3.4 体素绘制 利用 MATLAB 中 scatter3 函数实现体数据绘制。图 2 展示了体数据体素绘制可视化效果。

图 2 体数据体素绘制，左：前面，右：后面(colormap: pink)

1.3.5 不规则切割 利用体数据重组函数将轴位数据重组出矢状切面后，调整好切面位置，利用自由二维切割函数沿椎间孔进行切割，将获得的范围进行二值化，根据重组后矢状面厚度，对二值化掩模进行扩展，形成体掩模数据。通过体掩模与原数据点积获得切割后的体数据，设置好体素偏色后形成伪彩体素绘制模型。

图 3 展示了不规则切割的流程。

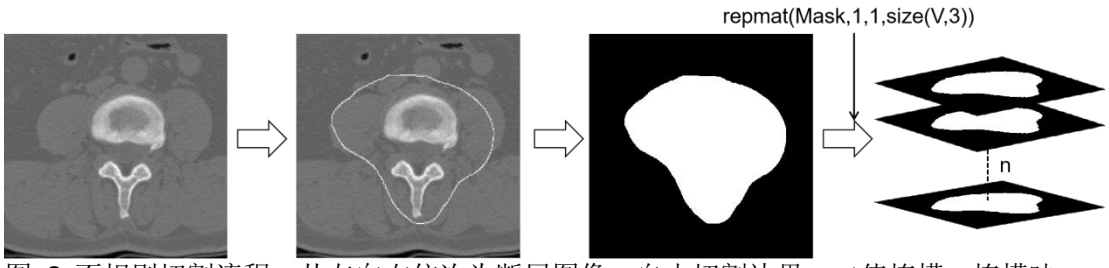


图 3 不规则切割流程，从左向右依次为断层图像，自由切割边界，二值掩模，掩模映射扩展体积化，n=size(V,3)

1.3.6 切割后体数据绘制 将 1.3.3 与 1.3.4 组合成绘制模块，然后利用该模块实现切割后体数据绘制。

1.3.7 多阈值体绘制 通过设置高低阈值，利用绘制模块对体数据中不同组织进行绘制。

1.3.8 基于轴方向的平行切割 平行切割原理是将 x，y，z 轴方向进行限制性零化，通过三方向掩模点积形成体积掩模，将体积掩模与原体积数据再次点积，最后通过绘制模块实现体素绘制。图 4 展示了平行切割原理。

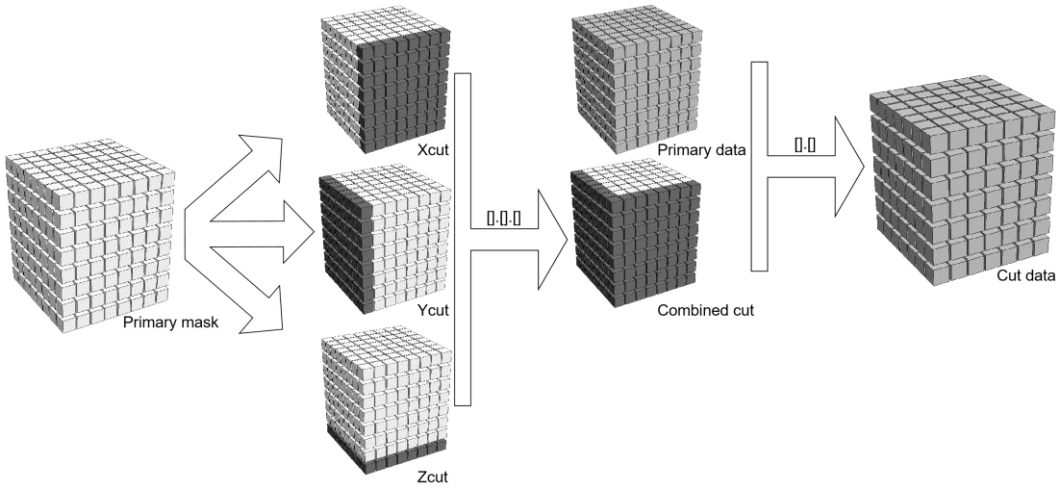
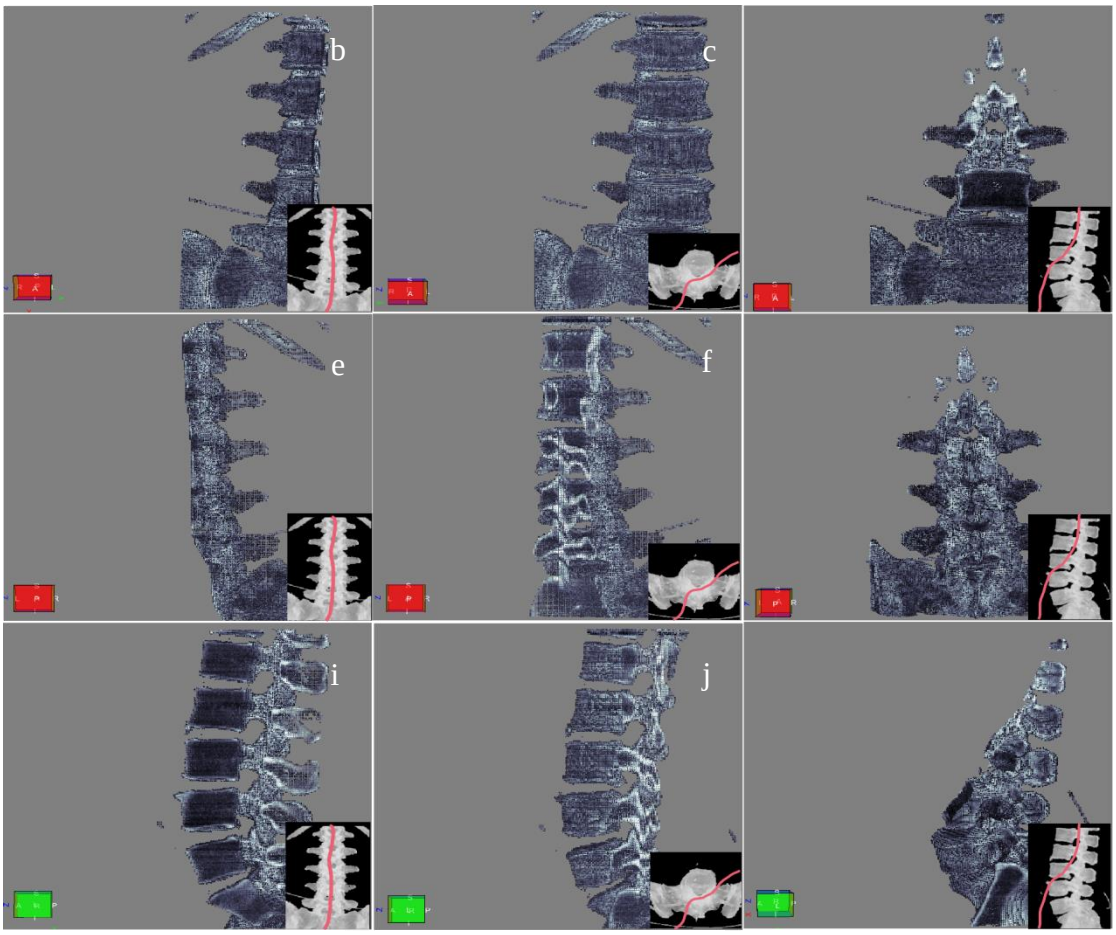


图 4 平行切割掩模生成与点积过程

2 结果

通过在冠状断层、水平断层以及矢状断层上进行自由切割，实现多种切割效果。图 6 展示了基于冠状，水平与矢状位的不规则切割效果。



k l m

图 6 不规则切割效果，a,d,h 为冠状位切割效果；b,e,i 为水平位切割效果；c,f,j 为矢状位切割效果；k-m 为冠状位，水平位和矢状位切割模型体素值分布图(colormap: bone)

3 讨论

3.1 体数据自由切割的必要性

体数据包括影像扫描时患者的全部特定信息，根据不同的临床需求针对性地提取其中的信息[6]。对兴趣区的设置关系到特征提取与可视化效果，因而提出了体数据切割的概念，目前所用的切割方法主要是包围框裁剪[7]，这种方法与本研究中的平行切割中正向效果（图 4）相似，自动探测体数据边缘点，找出 X，Y，Z 的最大边界，根据边界形成包围框，再给予包围框上下左右前后分布好控制图形柄，拖动相关的图形柄达到对体数据切割的目的。但这种切割方法存在只能从外向内平行进行，最后形成与图 4 一样的抽样体数据小方块，并不能实现在水平位、冠状位与矢状位自由切割的目的，也阻碍了图像内部结构的探索。为了弥补这一缺点，在某些超声软件中，提供了基于自动边界的不规则二维图像兴趣区提取[8]。这些兴趣区也有助于图像特征的提取[9]。

MATLAB 提供了一些实现自由切割的基础函数，比较常用的有 impoly，imfreehand，前者是进行不规则角点曲线切割，在封闭切割曲线后，可以对形成的角点进行移动来改变切割曲线，但

impoly 只能形成折线，当目标结构复杂时，很难精确掌控切割边界。imfreehand 则可以精确地掌控切割边界，当缺点在于一旦形成切割边界，无法进行修改。本研究基于 imfreehand 在二维断层上形成掩模，再扩展到体数据中，由图 6 可以看出，达到了任意切割的效果。

3.2 点绘制映射与纹理映射

光线投影(Raycasting)是将体数据中的灰度信息，根据光线传播方向进行投影，将灰度信息进行叠加并以此为阻光率，对像素显示量进行调节，从而达到体数据投影为二维图像的效果。这种渲染是基于点绘制映射[10]，优点是对体数据信息高度还原，且可起到小核滤波去噪的作用，但缺点是光线叠加算法会丢失部位细节。纹理映射则通过对多边形三维物体进行 UV 拆解后，将二维图像映射至 UV 上，从而实现三维模型的细节体现，优点是基于多边形的表面渲染算法可弥补模型本身的细节不足，映射速度较快。本研究中所采用的体数据为 213 个断层，在显示时利用 camzoom 对镜头距离进行有效调节，由图 6 可见，软组织与骨骼均得到较好的体现。

3.3 不足与展望

MATLAB 平台是一款比较全面的矩阵处理平台，相较于可视化工具包(Visualization Toolkit, VTK)，它的优势在于图像处理算法的多样化与个性化，但基于 C++的 VTK 在跨平台兼容上显然明显优于 MATLAB[11]。此外，MATLAB 平台还存在程序运行的具体不足，首先在 MATLABR2016a 这个版本调用 VTK 需要安装插件[12]；其次，对象(图形句柄)需要用代码进行回收，对内存释放产生的较大的影响；再次，相机处理呆板，旋转对象分为旋转轴与旋转对象两种，常用的旋转轴函数 rotate3d 存在轴限制，虽然旋转对象可达到取消轴限制的效果，但随着对象被删除，相机会报错而不能工作，非常不利于三维可视化交互；最后，体素绘制也存在体素点间隙固定，无法实现细节观察。在不规则切割上，目前仅限于冠状位、水平位与矢状位的投影切割，无法实现基于相机视角切割，这也是今后研究的重点。

4 结论

自由投影与平行切割的体素绘制可实现 CT 体数据中投影切割，更好地选择性展示 CT 数据内部结构。

PU-254

硝苯地平的多中心临床疗效评价及患者概况研究

王婷婷 赵文龙*
重庆医科大学

目的 该模型旨在帮助高血压患者合理用药，提高治疗的有效率、避免并发症的发生，并使患者生活质量得到提升。

方法 本研究使用 XGBoost、LightGBM、GBDT、AdaBoost、CatBoost、RF、Logistic7 种机器学习方法，并使用准确度、灵敏度、特异性、阳性预测值、阴性预测值、曲线下面积对模型进行综合性能评估。

结果 通过对模型进行重要性排序和特征选择后，纳入 HCT、AST、MPV、RBC、K、TBIL、TP、TT、C19 个变量。再使用 LightGBM、XGBoost、GBDT 构建减少变量后的机器学习模型，LightGBM 模型性能最优（AUC=0.880，ACC=0.788）。

结论 当患者 HCT、RBC、K 值较高，而 TT、MPV、AST、TBIL、TP 较低时，使用硝苯地平可获得更佳的治疗效果。

PU-255

公立医院绩效考核背景下我国远程医疗的发展契机与思考

王硕果① 刘新奎②

①郑州大学公共卫生学院

②郑州大学第一附属医院

引言

近年来，随着我国经济综合实力的高速发展，居民的生活水平也大幅提升，大众越来越重视自身及家人的身体健康，这一现状对各级医疗机构提供高质量医疗服务提出了更深层次的要求。而目前经过局部的医疗体制改革，我国卫生医疗整体服务水平，特别是高端卫生医疗服务水平在快速提升，但是高端医疗资源仍然集中在大城市，医疗资源供需不平衡的问题依然突出。在此前抗争新冠疫情的过程中，远程医疗充分发挥灵活快捷、无时空限制的特点，北京、河南、广东等地通过远程会诊方式抢救重症患者，为有效降低病死率及医护人员感染率起到了不可忽视的积极作用。因此，各级医疗机构应该充分利用好远程医疗，推动远程医疗服务落地，以此来提升医疗服务水平和医疗质量。

1 远程医疗的定义

“远程医疗”于 20 世纪 70 年代由美国学者 Thomas 提出，可简单理解为远距离治疗”，其目的在于通过远程通讯技术和计算机信息技术，实现医务人员跨地区、远距离为病人提供临床医疗服务。大量实践证明，远程医疗能够增加患者就医机会、消除时空障碍、实现优质医疗资源共享、提升基层医疗机构诊疗水平，在慢病管理及老年人疾病预防保健方面有着突出作用。世界卫生组织定义远程医疗为：远程医疗是在传统医疗服务的基础上，通过现代信息通讯技术来实现邀请方与受邀方远距离、大空间内的医疗卫生资源优化配置，达到疾病会诊与治疗、健康管理、医务人员再教育等目的。我国国家卫生计生委定义远程医疗服务为邀请方与受邀方医疗机构之间借助通讯、网络等技术，由受邀方为邀请方的患者提供更权威的医疗帮辅。我国国家卫健委将远程医疗分为“远程会诊”与“远程诊断”两种模式，其中，远程会诊是“受邀方提供诊断治疗意见，邀请方明确诊断方案”，远程诊断是“由邀请方实施医学影像、病理、心电、超声等辅助检查，由受邀的上级医疗机构进行诊断”。

2 公立医院绩效考核背景下远程医疗发展契机

2.1 有力的政策指引

近 5 年，国家层面先后发布互联网医疗、远程医疗相关政策文件超 50 余个，特别在公立医院绩效考核启动后，更有相应指导文件相继发布。2020 年，工业和信息化部办公厅、国家卫生健康委办公厅联合发布《关于进一步加强远程医疗网络能力建设的通知》，特别提到建设医疗云计算和大数据应用服务体系，持续提升医疗信息化基础能力，实现信息资源共享，为远程会诊远程影像、远程心电等网络应用场景提供技术支撑，提升远程医疗系统互联互通能力，研究推进远程医疗系统音视频通讯互联互通机制。2021 年，国务院办公厅发布《关于推动公立医院高质量发展的意见》，特别指出推动云计算、大数据等新一代信息技术与医疗服务深度融合。推进电子病历、智慧

服务、慧管理“三位一体”的智慧医院建设和医院信息标准化建设，同时特别强调大力发展远程医疗和互联网诊疗。

2.2 多元化健康之需

云计算、大数据、人工智能等信息科技的不断成熟进步也催生了互联网与医疗的深度融合，“互联网+”健康、互联网诊疗、远程医疗以蓬勃之势进入人们的健康生活中，为老百姓的多元化健康需求注人新的活力和新的动能。远程医疗作为“互联网+”健康最早期、最成熟的应用之一，近 10 年来已为医务工作者、卫生管理者、患者及广大老百姓所熟悉。

2.3 全民健康之策

远程会诊、远程诊断、远程问诊、远程预约、远程随访、远程开方、远程教育、云端论坛、视频会议等众多便捷的就医模式融入到人们生活中，不仅大大方便了就诊的患者，也为推进全民健康贡献了一份举足轻重的力量。远程医疗与现代医学的深度融合应用，在加固重大疾病救治防控体系建设方面正发挥着越来越重用。

3 以公立医院绩效考核为契机的远程医疗应用实践

3.1 基于五级远程网络构建整合型远程医疗应用体系，提升远程医疗可及性、普惠性

公立医院绩效考核启动以来，多家医院进一步统筹优质医疗资源、优质教学资源，以医联体、专科联盟建设为纽带，持续加固五级远程医疗网络构架，推动构建区域医疗协同发展的上下联动远程医疗创新应用体系。该体系以“省-市-县-乡-村”五级远程网络为支撑，融合云计算、大数据等关键技术，向上联通国内 56 所知名医院，向下联通市、县级医院并辐射至县域医共体，逐步建立起了一个集“远程医疗、远程教育、区域协同、应总处突、健康维护”为一体的数字化整合型远程医疗应用体系该体系远程医疗功能日趋成熟完善，快速服务于医联体、联网医院以及区域内的医院、专家、医生和患者群体，充分发挥了远程医疗特有的优势和效能。

3.2 以诊疗同质化为立足点，构建“四位一体”远程诊断体系，推进区域检查检验结果互认

随着“互联网+”医疗的深人推进，数字化逐步成为解锁区域检查检验结果互认的“最佳钥匙”。近几年，多家医院综合既往远程医疗实践成果，不断拓宽远程医疗应用领域，以远程医疗关键技术体系建设为切人点，完善构建集“远程心电诊断、远程影像诊断、远程病理诊断、远程超声诊断”四位一体的远程诊断平台，并将该平台集成于远程医疗大数据平台中，以医联体和专科联盟为纽带，围绕医学影像、心电诊断、病理诊断、超声医学等方面，扎实推进远程心电远程病理、远程超声、远程影像等远程诊断级应用，逐步实现“基层检查、上级诊断”。超声诊断。“四位一体”远程诊断体系的建立，极大解决了基层医疗机构诊断资源贫乏、诊断能力严重不足的问题，为推进区域同质化诊疗和区域检查检验结果互认奠定良好的基础。

3.3 启动远程 MDT，开启多学科协作“线上互动”模式，提升疾病诊治的“质”和“效”

远程医疗的会诊类型既往多局限于普通会诊、点名会诊和急诊会诊，缺乏针对疑难杂症、多器官损害疾病的远程协作，而这类疾病的远程指导应该是基层医疗机构“急难愁盼”的会诊问题。远程医疗多学科联合会诊，即远程 MDT，为联网医疗机构在线开展多学科疑难危重症的会诊难题找到了解决路径。通过实施远程 MDT，一方面可以与知名三甲医院签署远程医疗合作协议，开展跨省区远程医疗合作，解决三级医院疑难危重病例的会诊需求，减少跨区域就诊患者；一方面辐射至区域内各级各类联网医疗机构。远程 MDT 为联网医疗机构提供了一个跨区域、跨学科在线交流的互动平台，通过线上多学科联动、多专家互动有效促动了不同区域优质医疗资源的线上协作和交流，有效提升了区域疾病诊治的“质”和“效”，为疑难危急重患者提供了切实有效的诊疗保障。

3.4 构建医联体协作体系，促进区域医疗服务能力、医疗服务质量“双提升”，助力分级诊疗

随着医改深入，医联体建设已成为深化医改和推进分级诊疗的有效举措和重要抓手。近年来，各地医院结合区域地理人文特点、群众健康需求以及大型公立医院改革发展方向，将分级诊疗作为医联体工作的重要抓手，在已趋成熟的远程医疗协作网络体系基础上,探索构建医联体协作体系。该体系主要运用一个远程医疗大数据集成应用平台，内容主要包括互联网远程会诊、远程教育培训等远程医疗业务协同。远程医疗作为医联体实效合作的有效手段，同时也作为医联体实效合作的重要平台，在规范医联体内各级医疗机构诊疗行为，协助建立推行临床路径，促进医联体内同质化管理、同质化诊疗发挥了很大的作用，很大程度上推进了基层医院的规范化、制度化建设，为努力实现“小病不出乡、大病不出县、疑难危重病不出省”的分级诊疗格局奠定了良好的基础。

4 公立医院绩效考核背景下远程医疗的发展思考

4.1 合力打通数据联通制约壁垒，提升远程医疗平台的联通性、兼容性和共享性

随着信息科技的不断发展进步，以及人民群众对健康需求的多元化，远程医疗在其发展应用过程中亟待解决的一些问题也日益凸显，特别是不同医疗机构间诊疗信息无法联通共享的问题、不同信息系统平台间不能兼容互动的问题、标准化远程医学质控体系建设的问题以及需求剧增的便捷高效的移动端远程会诊问题等，这些问题一直以来都是困扰远程医疗发展的短板。公立医院绩效考核背景下，这些远程医疗的短板已经沉积为严重制约远程医疗提质增效的瓶颈问题。目前，迫切需要从信息、技术、管理、政策等各个角度、各个层面联合发力，打通这些问题的制约壁垒，建立互通共享的数据平台，才能真正提升远程医疗平台的联通性、兼容性和共享性，更大更好更优地发挥远程医疗效能。

4.2 深化远程医疗应用，提升远程医疗的协作性、增效性和普惠性

远程医疗应用最多、基层需求最大、最趋常态的功能之一就是远程会诊，包括医疗类远程会诊和医技诊断类远程会诊，以远程心电、远程影像、远程病理为代表的医技诊断类远程应用星蓬勃之势。打通数据联通制约壁垒是一个方面，从医联体建设、区域医疗中心建设等方面深化与远程医疗的实效融合，发挥好、利用好、应用好远程医疗大平台的各项功能，完善构建符合区域医疗协同发展的远程医疗应用体系，不断提升远程医疗的协作性、增效性和普惠性，才能更好地应用数字化手段助力优质医疗资源扩容下沉，发挥好公立医院在整合型医疗卫生服务体系中的领头作用和骨干作用。

5 小结

我国远程医疗的发展尚处于发展阶段，受统筹地区范围的限制，各地发展形态各异。不同地区纳入远程医疗的服务项目不同，各项目的定义及服务范围也存在差异，加上不同地区医疗成本的不同，造成远程医疗服务定价参差不齐。因此国家应加快制定指导性政策文件，规范远程医疗的服务项目及价格标准，在充分了解各地发展现状的基础上合理界定远程医疗服务范围、考量医疗服务成本，进一步加强我国医疗资源的共享。患者是远程医疗的受益者，只有不断提高患者参与度，远程医疗才拥有更广阔的市场和发展空间，从而推动远程医疗健康可持续发展。

PU-256

区域医疗中心模式下远程云服务平台的构建与探索

王娜* 褚震宇 周潞荣 陈文豪 焦皓男
长治市人民医院

目的 本文在介绍远程问诊现状的基础上，以长治市人民医院为例，建立区域医疗中心模式下远程云服务平台，旨在实现优质医疗资源扩容下沉，打造一个可循环造血，高性能、高可靠、高扩展的云服务平台，为医护人员提供更便捷的诊疗工具，为患者提供更优质的诊疗服务。

方法 远程云服务平台基于统一规划、顶层设计思路，利用现有院内局域网、卫健委专线网络及 5G 互联网为系统提供基础信息整合通道，利用云计算、大数据等新一代信息技术的先进理念，规划部署一个横向到边、纵向到底覆盖医疗资源均衡的远程云服务平台，满足多级部门实时互动、常态化视频，以及加强统筹融合、共享公用的需求。

该平台以长治市人民医院院内会诊平台为中心，向外辐射药店、社区、乡镇卫生院等医疗机构，实现音频、视频、数据、内容、即时通讯、业务全覆盖，满足数据的统一管理、统一运维。

该平台最底层为网络层，支持院内局域网、卫健委专网、5G 互联网、语音网络；网络层之上时视频服务层，实现 API 开方接口；视频服务层之上是接入层，支持多点部署和移动个人终端部署；接入层之上是业务层，满足院内应用，多学科 MDT，区域远程会诊，远程开方，云上服务台等远程服务，实现 HIS、LIS、PACS 等数据共享。

结果 本院自建立区域远程云服务平台以来，不断完善区域医疗中心工作内容，根据各医疗机构的不同需求，开展多方面的技术支持，为医护人员提供了更便捷、更高效、更精准的工作流程，为患者提供了更优质、更智能、更贴心的诊疗服务。

随着远程云服务平台功能的不断完善，云 MC 平台自上线以来，共举办协同会议 28 次，会诊 400 余次，疑难病例讨论 284 次，远程查房 60 余次；云上服务台自上线以来，共接通次数达 2 万余次；云诊室自上线以来从最开始的 33 家药店增加至 168 家，处方量从 750 张增至 1584 张，累计共完成 10 万张，同时，不断缩短每一单的完成时长，从最开始 60 秒缩短至 20 秒以内；云探视自上线以来累计共完成 8000 余例，有效的改善了本市居民的就医体验。

结论 1、实现统一管理。该平台建立区域医疗中心群组，将所有设备都标签化加入群组，从院区、科室可以精确到具体个人终端，监控每台设备的在线使用情况，从而实现所有终端设备的的全流程管理；

2、成本低。该平台支持多种网络接入，省掉了现有商用微信视频、腾讯会议等产生的二次视讯服务费用，从而降低了运营成本，同时，支持多种业务场景，省去不同应用场景需要配置的软硬件成本；

3、开放度高。该平台支持医疗设备高清接口共享、PC 端桌面共享、一键无线双流共享三种数据共享方式；

4、可扩展性好。该平台支持二次开发，对接本院互联网医院，实现按需使用、长期运营增效，从而满足该平台的可持续性发展。

PU-257

我国远程医疗发展现状分析

秦超雨

郑州大学第一附属医院

1 引言

随着信息技术的快速发展和互联网的普及，远程医疗成为了当前医疗领域的热门话题。远程医疗利用先进的通信技术和网络平台，实现医疗资源的跨时空传输和共享，为患者提供诊断、治疗、咨询等医疗服务，极大地改善了医疗服务的效率和便捷性。李海洁[1]指出，在我国，远程医疗发展得到了广泛的重视和推动，政府和医疗机构纷纷投入资源和精力，推动远程医疗的应用和普及。陈润卿[2]等指出，随着技术的不断创新和应用的深入推广，远程医疗将在我国医疗体系中扮演越来越重要的角色。通过本文的研究和分析，希望能够为我国远程医疗的进一步发展提供科学的参考和有益的借鉴，推动我国医疗服务的现代化转型，提高人民群众的健康水平和生活质量。

2 远程医疗的内涵

远程医疗指的是通过远程通信技术，实现医疗服务的提供和医患交流的方式。通过远程医疗，患者可以在不必亲自前往医院的情况下，通过电话、视频和互联网等方式与医生进行沟通、咨询和诊疗。姜艺佼[3]等学者认为远程医疗需要大量信息技术的支持、投入成本耗时较大，且偏远地区存在通信匮乏等问题。

远程医疗减少了患者的交通和住院费用，同时也减轻了医院的负荷。除此之外，陈芳[4]认为随着人口老龄化和慢性病的增加，对医疗服务的需求也日益增长。远程医疗为老年人和慢性病患者提供了更便捷的医疗管理方式。他们可以在家中接受远程监护和远程诊疗，减少因疾病而频繁就医的困扰，提高生活质量。

3 我国远程医疗研究进展

中国的远程医疗研究在过去几年取得了显著的进展。刘叶[5]等学者认为在三甲医院发展远程医疗心电诊断中心的平台建设是互联网 + 和医学诊疗高度融合的结果。不仅可以体现“互联网 + 远程医疗”从理论向实际的转变，而且还能改善传统的医学诊疗模式，缩短基层医疗机构急性冠脉综合征患者的“院前急救时间”，降低基层医疗机构心血管疾病的病死率，并且在远程医疗心电诊断过程中还可以联系急诊科，为患者安排绿色通道，由此看来，远程医疗心电会诊具有良好的社会应用前景。

于立升[6]提出，目前“互联网+远程会诊”应用存在诸多问题。现有的研究成果对于“互联网+”医疗服务实际的运行环境与患者诊疗体验缺乏专业医学背景和临床使用的反馈数据，使用单一的疾病评测功能量表或是患者体验测量工具必然产生诊断偏差。而侧重于研究用户感受的服务内容，或是仅仅分析人机交互效果及流量转化数据的应用研究，必然存在放大医疗服务中的商品属性，偏离医疗救助服务本质的问题。研究远程会诊和远程手术技术的应用，包括远程专家会诊、远程手术指导等。这些研究可以帮助解决医疗资源集中在大城市的问题，提高偏远地区患者的就医便利性和医疗服务水平。

远程监护与慢病管理方面，需要研究远程监护技术在慢性病管理和健康监测中的应用，其中包括远程心电监护、远程糖尿病管理等。这些研究可以实现医疗资源的有效利用，提高患者的生活质量和健康管理效果。

远程医疗数据安全方面，需要研究远程医疗中的数据安全和隐私保护技术，包括医疗数据传输加密、身份验证等。这些研究对于保护患者个人隐私和医疗数据的安全具有重要意义。

社会经济影响评估方面，需要研究远程医疗对社会经济的影响，包括医疗费用、患者就医便利性、医疗资源利用效率等方面的评估。这些研究可以为政府决策和资源配置提供科学依据。

值得注意的是，以上进展领域并不是独立存在的，它们往往相互关联和相互促进。我国的远程医疗研究在技术创新、政策支持和应用推广等方面都取得了长足进步，并且在改善医疗服务和健康管理方面发挥了重要作用。未来，随着技术的不断进步和政策的进一步完善，中国远程医疗研究将继续推动医疗服务的创新和发展。

徐健[7]指出美国在 2020 年 3 月将远程医疗服务纳入国家卫生计划，互联医疗政策中心远程医疗服务收费编码等一同增加。据调查，美国远程医疗面临应用方面等问题，其对网络质量、设备清晰度、用户意愿与技术能力等均有较高要求，由于对远程医疗了解程度有限、需掌握更多新知识和新技能、不确定其有效性等因素，多数临床医生不愿应用远程医疗。

4 我国远程医疗发展面临的问题

我国远程医疗在发展过程中面临一些挑战和问题，其中包括技术标准与互操作性、法律法规和隐私保护、医保支付和商业模式、医疗资源分配不均、患者接受度和信任等问题。

技术标准与互操作性：远程医疗涉及多个技术领域，如网络通信、医学影像传输、远程监测等。因此，需要制定统一的技术标准和互操作性规范，以确保各个系统和设备之间能够无缝协同工作，实现信息的安全、准确和及时传输。

法律法规和隐私保护：崔文彬[8]学者提出，远程医疗涉及到医疗信息的传输和存储，需要符合相关的法律法规和隐私保护要求。加强对远程医疗的监管，明确数据隐私保护的责任和义务，保障患者个人隐私和医疗信息的安全。

医保支付和商业模式：远程医疗的发展还需要解决医保支付和商业模式的问题。目前远程医疗服务的费用结算、医保支付标准和政策尚未完善，需要建立统一的费用制度和支付机制，推动远程医疗的可持续发展。

医疗资源分配不均：远程医疗可以弥补医疗资源分布不均的问题，但在实际应用中，仍存在医疗资源的不平衡现象。一些偏远地区和农村地区的医疗机构和患者可能无法充分享受到远程医疗的便利和优势，需要进一步推动资源的优化配置和覆盖。

患者接受度和信任：张炎亮[9]等学者认为，远程医疗需要患者主动参与和接受，但一些患者对于远程医疗的接受度和信任度较低。需要加强对患者的宣传教育，增加对远程医疗的认知和理解，提高患者对远程医疗的信任和接受度。

同时，孙倩倩[10]等学者认为我国面临的问题有服务定价和补偿机制不完善、远程医疗系统平台功能不完善、法律法规不健全。目前对于远程医疗的规范多见于政策性文件，尚未上升为国家层面的法律法规，对医疗机 构、从业人员、医保基金监管、医疗责任认定、数据安全等方面的规定多等同于线下诊疗，缺乏应有的权威性和科学性。

面对这些问题，需要政府、医疗机构、技术提供商和学术界共同努力，加强合作与协调，推动远程医疗的规范化、标准化和可持续发展，以提高医疗服务的质量和效率，更好地满足人民群众的健康需求。

4 结论与展望

远程医疗作为一种创新的医疗服务方式，具有巨大的潜力和机遇。通过综合分析现有研究和实践经验，我们认为远程医疗已经成为解决医疗资源不平衡、提高医疗效率和服务质量的有效手段。然而，远程医疗仍然面临着信息安全隐私保护、技术可行性等方面的挑战。在未来，我们需要加大技术研发和创新力度，完善相关法律法规和伦理规范，为远程医疗的健康发展提供支持和保障。只

有克服挑战，并与多学科共同合作，才能实现远程医疗在改善人们生命质量和医疗服务中的长远价值。

刘川[11]认为远程医疗服务可及性包含四个方面，分别是医疗服务的距离可及性、医疗服务的可得性、医疗服务的经济可行性、医疗服务的可接受性。距离可及性是指通过某种途径可以随时得到远距离的医疗服务;可得性是指真正有迫切需求的人在相同政策下获得医疗服务; 经济可行性是指基于当前的医疗服务价格以及用户的支付能力等方面考虑能否获得医疗服务;可接受性是指用户自愿接受的程度。

基于对现有理论和文献的梳理和思考，本文最后对未来研究作些展望，具体如下：在技术创新方面，相信在未来，随着 5G 技术的普及和应用，远程医疗将更加高效和稳定。同时，人工智能、大数据和物联网等新技术的应用，将进一步提升远程医疗的精准性和个性化服务。在政策支持与规范方面，政府应继续加大对远程医疗的政策支持和引导，完善相关法律法规，明确远程医疗的管理要求和操作规范，保护患者利益和医疗安全。在跨界合作与整合方面，远程医疗需要医疗机构、科研机构、技术企业等多方合作，形成良好的生态系统。跨界整合医疗资源，推动远程医疗与医疗机构、药店、保险公司等的深度融合，提供全方位的医疗服务。在人才培养与教育方面，加强对远程医疗人才的培养和引进，提高医疗人员的技术水平和应用能力。同时，加大对患者和公众的宣传和教育力度，提高对远程医疗的认知和接受度。在数据安全和隐私保护方面，应当加强对远程医疗数据的安全保护，建立健全的数据隐私保护机制，确保患者个人隐私的安全和医疗数据的合法使用。

PU-258

基于数字化平台的智慧健康社区建设案例分析-以深圳市为例

赵靓

深圳市卫生健康发展研究和数据管理中心

目的 本研究旨在通过深圳市典型案例分析，深入探讨数字化平台在智慧健康社区建设中的关键作用，阐述数字化平台在智慧健康社区建设中的价值，为未来的健康社区建设提供创新的思路和解决方案，以应对不断演变的健康挑战。

方法 通过案例分析，探讨了以社康为中心的数字化平台建设以及以社区为中心的数字化平台建设的体制机制建设、主要做法和效果，从核心建设主体、资源整合能力、辐射带动能力和建设发展趋势等方面对两种建设路径进行对比分析，提炼两种路径的优势和不足，为智慧健康社区的建设发展提供合理有效的经验借鉴。

结果 以社区为核心的福田健康云管家，强调社区层面的资源整合和横向纵向协作，为社区工作提供了强大支持。以社康为核心的大冲智慧健康社区平台，侧重于标准化的服务流程和居民互动，提高了社康人员的工作效率。两种路径在核心建设主体、资源整合能力、辐射带动能力和建设发展趋势等方面存在明显不同。福田健康云管家在资源整合能力和辐射带动能力方面表现出色，而大冲智慧健康社区平台以社康为核心，专业服务供给能力强，侧重于标准化的服务流程和居民互动，提高了社康人员的工作效率。

结论 以社区为核心的建设和以社康为核心的建设各有其优势和不足。各城市和地区应根据自身情况，选择符合特定需求的路径，结合地区资源、发展趋势和服务需求推进数字化平台建设，助力推进智慧健康社区建设，推进医疗服务的远程化、智能化发展，提高居民的健康管理体验和医疗服务效率。

PU-259

基于逻辑组学及药物表观精准对接技术筛选新型铁死亡诱导药物治疗 c-MYC/Bcl-2 重排的高级别 B 细胞淋巴瘤及基础实验验证

李泓毅[1,2]，卢学春[1,2,3*]
[1]解放军医学院
[2]解放军总医院第二医学中心血液病科
[3]山西医科大学

目的 存在 c-MYC 和 BCL-2 重排的高级别 B 细胞淋巴瘤(High grade B cell lymphoma, HGBCL)属于高度侵袭性 B 细胞淋巴瘤，患者预后差，耐药性强。本研究拟探明 c-MYC 和 BCL-2 重排的 HGBCL 与预后不良相关的核心基因，为后续研究 c-MYC 及 BCL-2 重排的 DHL 提供基础。

方法 通过公共数据库检索 HGBCL 的基因表达谱数据。差异分析后筛选核心基因，并通过 FerrDb 铁死亡数据库进行铁死亡相关核心基因筛选。通过对市面所有药物及化合物的转录组数据整合后，将其与转录组数据分析整合，根据逻辑组学原理进行配对整合，最终筛选可能对 HGBCL 具有治疗作用的药物。将筛选的铁死亡相关药物处理 DOHH-2 及 DHL-10 细胞系（c-MYC 和 BCL-2 重排的 HGBCL）后进行转录组高通量测序并进行 KEGG 富集分析；使用 Western Blot 验证 DOHH-2 药物处理前后凋亡蛋白及铁死亡相关蛋白表达；使用 CCK-8 法验证铁死亡诱导剂杀伤效果及联合铁死亡抑制剂 Fer-1 是否可以挽救细胞死亡；使用 GSH 试剂盒及 MDA 试剂盒分别检测药物处理前后 DOHH-2 内谷胱甘肽及脂质表达变化情况进一步验证铁死亡通路激活。

结果 共得出 287 个核心基因，并通过自建数据库筛选获得药物 TMX。CCK-8 验证 TMX 对 DOHH-2 具有杀伤作用，IC50 平均值为 11.04umol/L。药物处理后 DOHH-2 细胞的 Bcl-2、Bax、PARP-1 等凋亡蛋白表达均无显著差异，转录组测序的通路结果显示来曲唑的作用机制可能为铁死亡通路激活，富集的基因为 VDAC3、FTL、FTH1 等。TMX 与 Fer-1 联合处理细胞发现，Fer-1 能够挽救来曲唑对淋巴瘤细胞的杀伤作用；此外药物处理后 DOHH-2 内谷胱甘肽及脂质表达变化验证来曲唑通过铁死亡通路对细胞起到杀伤作用。

结论 通过生信方法筛选出药物 TMX 可能对 c-MYC/Bcl-2 重排的 HGBCL 具有治疗作用，通过 CCK-8 及转录组测序结果发现，TMX 对其的杀伤效果可能是通过铁死亡途径发生作用。

PU-260

智能医学的理念创新和范式突破的研究

智鹏
解放军医学院

目的 本研究旨在探讨智能医学领域的理念创新和范式突破对医疗健康领域的影响，以及这些变革如何为医疗系统和患者带来价值。

方法 通过查找现有文献进行了文献综述，分析了相关数据和趋势。进行些案例研究，以了解智能医学在不同医疗环境中的应用情况。此外，进行一项问卷调查，以了解医疗专业人员和患者对智能医学的看法和期望。

结果 研究发现，智能医学的理念创新和范式突破正在迅速改变医疗健康领域。通过人工智能、大数据分析和远程医疗技术的应用，医疗决策变得更加精确和个性化，患者获得了更广泛的医疗服务可及性。案例研究显示，智能医学已经在疾病预测、药物研发、临床诊断等多个领域取得了显著的

成功。问卷调查结果显示，医疗专业人员普遍认为智能医学将提高医疗质量，并减轻医疗工作的负担，同时患者更加倾向于使用远程医疗服务，以便更方便地获得医疗建议和治疗。

结论 智能医学的理念创新和范式突破为医疗健康领域带来了巨大的机会和挑战。通过更准确的诊断、个性化的治疗方案和更便捷的医疗服务，智能医学有望改善患者的医疗体验，提高医疗质量，并为医疗专业人员提供更多的工具和支持。然而，数据隐私和安全性问题需要得到妥善解决，医疗专业人员也需要不断更新他们的技能，以适应这一快速发展的领域。综上所述，智能医学的发展将为医疗健康领域带来积极的变革，但需要全球社区的共同努力来实现其最大潜力。

PU-261

MTOR and type 1 diabetes risk: a Mendelian randomized study

Can Wang[1], Yu-Xin Chen[2], Lei-Lei Zheng[3], Pei-Feng He[4], Qi Yu[1, 4]*

[1]School of Management, Shanxi Medical University, Taiyuan, China.
[2]School of Basic Medical Sciences, Shanxi Medical University, Taiyuan, China.
[3]The first clinical Medical College, Shanxi medical university, Taiyuan, China.
[4]Medical Data Sciences, Shanxi Medical University, Taiyuan, China.

Introduction

Type 1 diabetes(T1D) is a complex and invasive autoimmune disease, which is first manifested as impaired glucose tolerance in clinic, and then gradually progresses to symptomatic hyperglycemia[1].In recent years, the global burden of diabetes has increased significantly, and the trend and extent of this trend vary greatly among different regions and countries[2] . According to a study, the largest relative increase in the prevalence of diabetes between 2021 and 2045 is expected to occur in middle-income countries (21.1%). By 2045, the global health expenditure related to diabetes is estimated to be 1054 billion US dollars[3]. It is noteworthy that the overall pathophysiological feature of T1D is the T-cell-mediated self-destruction of the insulin-secreting pancreatic islet β cells in the pancreas[4, 5]. In addition, environmental factors, viral infection and intestinal microbiome also affect the occurrence of T1D[6]. However, its exact etiology and pathological mechanism are still unclear. A large number of recent studies have shown that inappropriate mTOR signaling can lead to metabolic diseases such as diabetes[7, 8].

The mechanical target of rapamycin (mTOR) is a highly conserved serine / threonine kinase that belongs to the phosphatidylinositol-3 kinase-associated kinases (PI3KK) superfamily[8, 9]. MTOR can regulate nutrient, energy and insulin signaling by forming catalytic subunits of two different protein complexes, called mTOR complex 1 (mTORC1) and 2 (mTORC2) [10]. Among them, mTORC1 is responsible for regulating two downstream effectors, namely ribosomal protein-S6 kinase 1 (RP-S6K 1) and eukaryotic translation initiation factor 4E-binding protein (EIF4EBP)[8], mTORC2 mainly phosphorylates AGC kinases [11], including Akt (also known as protein kinase B, PKB), protein kinase C (PKC)

family members, etc. Dysregulation of mTOR is associated with many pathophysiological conditions, such as aging, diabetes mellitus, abnormal lipid metabolism, and cancer[12].

However, whether mTOR mediates T1D is unknown. As a result, We used Mendelian randomization analysis to explore the causal relationship between mTOR protein(including AKT, ATF6, ATG7, BECN1, EIF4A3, EIF4B, EIF4E2, EIF4EBP2, EIF4G3, HIF1A, ULK3, LST1, PRKCA, RP-S6KA6) and T1D.

Methods

Aggregate summary statistics of Genome-wide Association Studies (GWAS) were performed in this study.

Genetic datasets

Strong instrumental variables were obtained from GWAS on plasma RP-S6K for this MR analysis. The publically available INTERVAL study Proteomics- GWAS data, which involved 3301 healthy individuals of European descent whose mean age is 43.7 years, 51.1% are men, and the initial INTERVAL study collected 45,264 eighteen years or older healthy blood donors, provided GWAS exposure data of genetic prediction towards the mTOR-related gene. [29875488] A total of 3,622 plasma protein concentrations are quantified by the SomaLogic method, which applies a protein-capture Slow Off-rate Modified Aptamer (SOMAmer). https://www.phpc.cam.ac.uk/ceu/proteins/[8] We employed the genetic predictors of mTOR downstream targets AKT, ATF6, ATG7, BECN1, EIF4A3, EIF4B, EIF4E2, EIF4EBP2, EIF4G3, HIF1A, ULK3, LST1, PRKCA and RP-S6KA6 to determine the causal association between mTOR-downstream proteins and T1D.

To prevent bias caused by population confounding, we limited the genetic background of T1D patients to European ancestry. We used a publicly available GWAS dataset for T1D from the Integrative Epidemiology Unit (IEU) GWAS database(https://gwas.mrcieu.ac.uk/), including 462,933 Europeans (38,472 and 424,461 controls).

Genetic variants selection criteria

The MR study was based on the three main assumptions:(1)Assumption of the relevance: the chosen independent variables (IVs) (mTOR-related genotype) should be straight related to the modifiable exposure (mTOR proteins level). (2) Assumption of independence: the chosen IVs had no relationship with the confounding factors (age, gender, birth weight, etc.) both in exposure and outcome. MR analysis considered that all the confounder variables did not emerge and affect the genotype random assortment after birth (3) Assumption of the exclusion restriction: the chosen IVs impact the outcomes (T1D) only through their relationship with the exposure. We utilized the two-sample MR analysis to appraise the causal association between proteins level concerning mTOR(exposure) and T1D(the outcomes). To satisfy the assumptions which ask the chosen IVs closely related to the exposure robustly, we set the threshold ($p < 5 \times 10^{-6}$, and no intersection attained with $p < 5 \times 10^{-8}$) to ensure the significance at a genome-wide level. To remove the effect of linkage disequilibrium, we applied relevant, rigorous standards ($r^2<0.001$).

The PhenoScanner database is a tool curated large-scale GWAS studies designed to scan phenome, binding the SNPs with disease phenotype and participation path. It

operates based on the European samples from 1000 Genomes and Hapmap, providing the function of searching the associations with trait deputize and the input variable. By inputting the given SNPs related to mTOR circulating protein levels of interest, we explored their relationship with the disease phenotype. We picked out the SNPs associated with age, height, weight, education, obesity, and other confounding factors and excluded them from our analysis[8, 13].

Statistical analysis

Two-sample MR analysis was employed by TwoSampleMR (version 0.5.6) and MR-PRESSO packages (version 1.0.0) in R software (version 4.0.5). The F-statistic was utilized to estimate the strength of the genetic instruments (F-statistics > 10 discharged the week instrument bias). To obtain a relatively reliable causal evaluation, we primarily used the classic inverse variance weighted model (IVW), equipped with multiplicative random effects and regarded P<0.05 as the threshold of positive results for causal effects. We applied the weighted median estimator model (WME), weighted model-based method (WM), and MR-Egger regression model (MER) to assess causal effects. A robust result can be acquired when invalid IVs account for more than half weight by the WME, and the WM is applicable to achieve a robust general causal evaluation when valid IVs offer most of resembling individual evaluation. When no effect on the validity of IVs, the MER is conducive to attaining a relatively robust assessment. Generally, the power of MER and WM methods is moderately compromised compared to the IVW methods. (35703935 32249995) [14, 15]For sensitivity analyses, Cochran's Q statistic was utilized for heterogeneity between each chosen SNP(P<0.05 manifested significant heterogeneity). Usage of the PhenoScanner assisted in the pleiotropy assessment of the genetic instruments. To examine the effect of SNP outliers, the MR residual sum and outlier (MR-PRESSO) and leave-one-out analyses. We excluded IVs one by one through leave-one-out analyses. The intercept of the MER analysis, the MR pleiotropy residual sum, and the outlier test (MR-PRESSO) were conducive to the check of directional pleiotropy. Moreover, an accurate estimate can be obtained by MR-Egger when the assumption of independence is established. (32978410)[8]

Results

As has been stated, for mTOR genetic prediction of exposure of the upstream and downstream targets, we acquired SNPs that could forcefully and independently be predicted circulating plasma levels of mTOR-related proteins (AKT, ATF6, ATG7, BECN1, EIF4A3, EIF4B, EIF4E2, EIF4EBP2, EIF4G3, HIF1A, ULK3, LST1, PRKCA, RP-S6KA6). The number of correlated SNPs for each protein was 8, 13, 9, 11, 10, 15, 14, 13, 8, 12, 11, 15, 8, and 14, respectively.

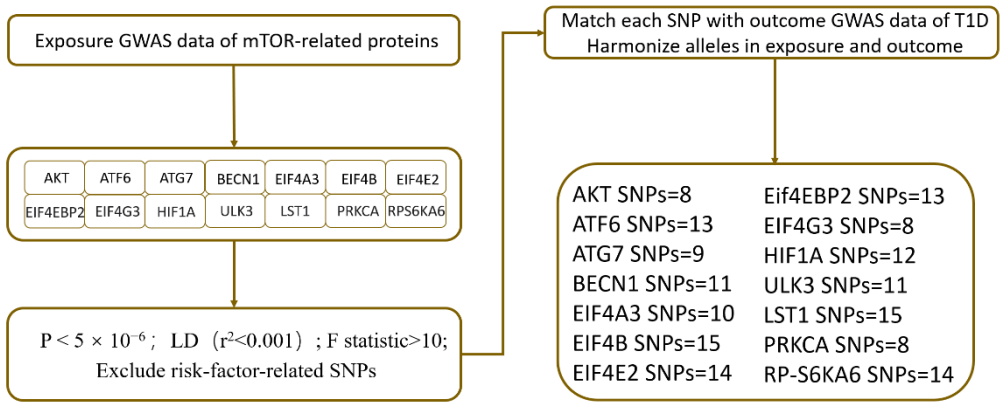


Fig.1: Train of thought framework.

The IVW analysis primarily indicated that genetically predicted RP-S6KA6 was causally associated with a higher risk of T1D [odds ratio (OR) 0.89, 95% confidence interval (CI) 0.82-0.97, P = 0.008], and similar risk estimates were obtained using the MR-Egger regression [odds ratio (OR) 0.91, 95% confidence interval (CI) 0.78-1.06, P = 0.24] and WM [odds ratio (OR) 0.90, 95% confidence interval (CI) 0.81-1.01, P = 0.07], although the association was not statistically significant. No outlier between the circulating plasma level of RP-S6KA6 and the risk of T1D was identified via the MR-PRESSO test. The leave-one-out analysis verified the robustness of the outcomes. The relevant records were shown in Fig.2.

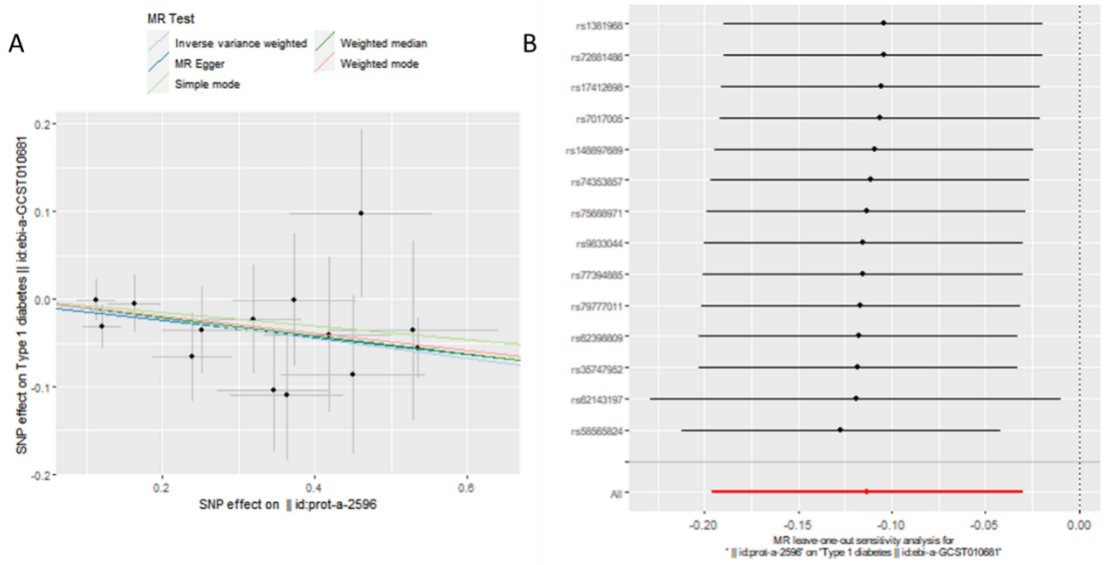


Fig.2: The results of MR analysis for RP-S6K. (A) MR-EGGER: the slope of each line indicating the evaluated MR effect in different models. (B) Leave-one-out of sensitivity tests. Calculate the MR results of the rest IVs after being discharged the IVs one by one.

As for other protein complexes, negative results were reported using MR methods, whose results with IVW displayed as that AKT[odds ratio (OR) 0.99, 95% confidence interval (CI) 0.86-, P = 0.86,ATF6 [OR 1.04, 95% CI 0.93-1.13, P = 0.60],ATG7 [OR 1.04, 95% CI 0.90-1.19, P = 0.61],BECN1 [OR 0.90, 95% confidence interval (CI) 0.80-1.10, P = 0.09],EIF4A3 [OR 0.99, 95% confidence interval (CI) 0.86-1.14, P = 0.91],EIF4B [OR 0.94, 95% confidence interval (CI) 0.84-1.06, P = 0.32], EIF4E2[OR 0.97, 95% confidence interval (CI) 0.89-1.06, P = 0.50], EIF4EBP2[OR 1.00, 95% confidence interval (CI) 0.89-1.12, P = 0.98], EIF4G3[odds ratio (OR) 1.04, 95% confidence interval (CI) 0.89-1.22, P = 0.60], HIF1A[odds ratio (OR) 1.05, 95% confidence interval (CI) 0.94-1.17, P = 0.91], ULK3[odds ratio (OR) 1.08, 95% confidence interval (CI) 0.96-1.22, P = 0.21], LST1[odds ratio (OR) 0.99, 95% confidence interval (CI) 0.89-1.10, P = 0.83], PRKCA[odds ratio (OR) 1.02, 95% confidence interval (CI) 0.76-1.36, P = 0.90]. The relevant records were shown in Fig.3.

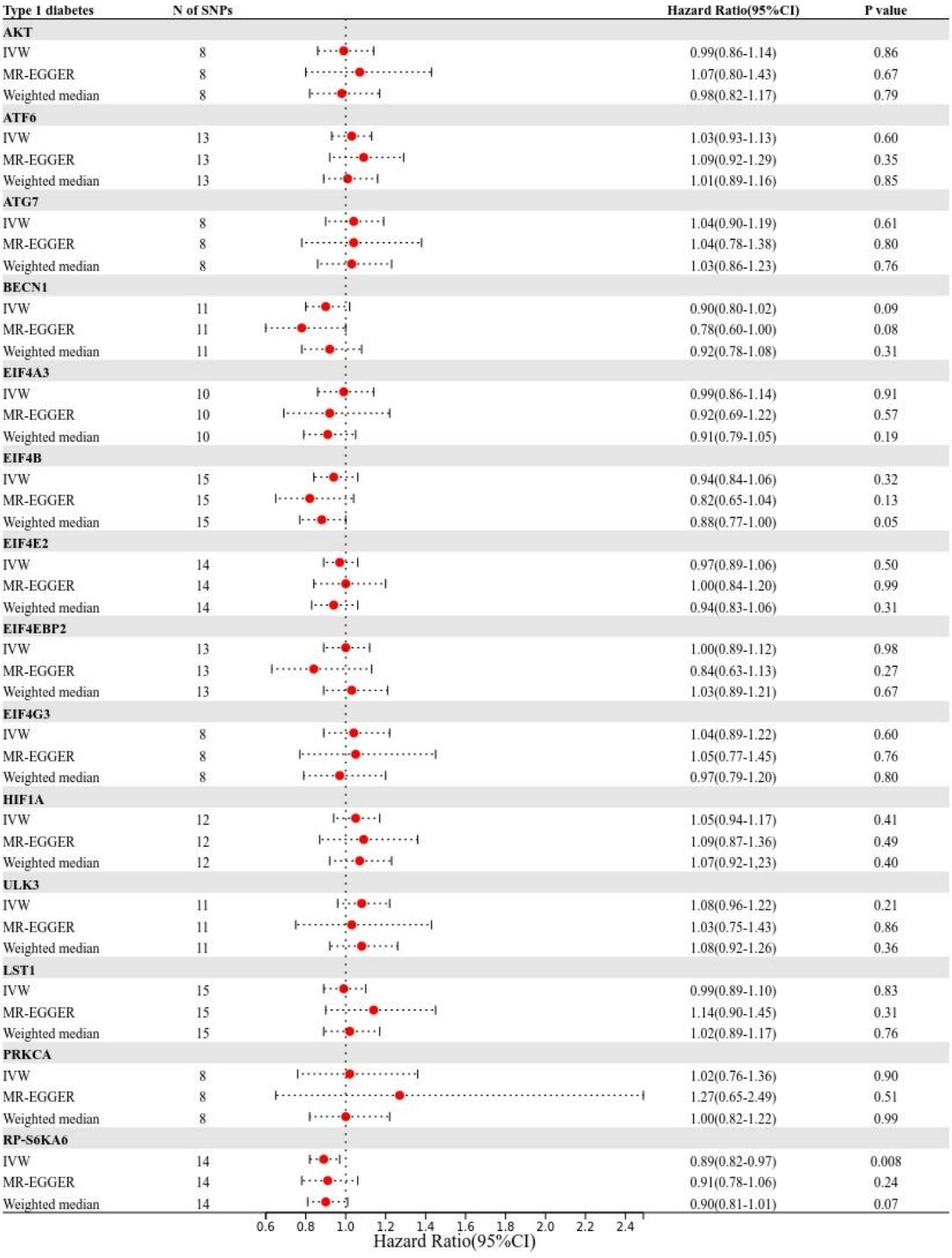


Fig.3: Forest plot of mTOR downstream targets (AKT, ATF6, ATG7, BECN1, EIF4A3, EIF4B, EIF4E2, EIF4EBP2, EIF4G3, HIF1A, ULK3, LST1, PRKCA) and causality in type 1 diabetes

Discussion

MR found robust evidence of a causal association between T1D and mTOR. MTOR may be targeted for intervention by repurposing existing therapeutics to reduce the risk of T1D.

T1D is an autoimmune disease characterized by insufficient or complete absence of insulin production, resulting in patients being dependent on exogenous insulin therapy. Despite significant research progress, the exact cause of T1D is still not fully understood. Plays a key role in cell growth and metabolism.

RPS6KA6 (Ribosomal Protein S6 Kinase A6) is a Protein Coding gene. It participates in the protein synthesis and proliferation process of cells by phosphorylating protein S6. However, in T1D, hyperactivation of RP-S6KA6 has been observed and is associated with insulin resistance and autoimmune processes that destroy pancreatic β cells.The expression level of RP-S6KA6 was significantly increased in T1D patients. This elevation is associated with abnormal activation of T cells during insulin resistance and autoimmunity. RP-S6KA6 hyperactivation may lead to abnormal T cell differentiation and function, which triggers an attack on pancreatic beta cells. In addition, RP-S6KA6 can further aggravate the development of T1D by regulating the insulin signaling pathway, affecting the secretion and action of insulin.

However, in T1D, hyperactivation of RP-S6KA6 has been observed and is associated with insulin resistance and autoimmune processes that destroy pancreatic β cells.

Although the association between RP-S6KA6 and T1D has been initially revealed, more in-depth studies are still needed to elucidate the molecular mechanism of this relationship. Further studies can explore the exact role of RP-S6KA6 in the pathogenesis of T1D and whether new therapeutic strategies can be developed by modulating the activity of RP-S6KA6 to improve insulin sensitivity and disease management in T1D patients.

PU-262

全球电子素养的现状及影响因素：一项 meta 分析

钟华 刘茜文 陈红*
四川大学

目的 本研究旨在总结全球电子素养的现状并探讨其影响因素。

方法 采用针对横断面研究准备的标准化 Joanna Briggs Institute（JBI）关键评价工具进行方法学有效性评估和数据提取，同时使用标准化 β 系数、95%置信区间（CI）和 I^2检验来展示 meta 的相关性度量。电子健康素养的预测因素根据"素养与健康"概念框架和技术接受模型（TAM）进行整合。

结果 地理信息图纳入 69 项研究，其中 17 项研究符合 meta 分析的标准。大多数国家/地区的参与者表现出足够的电子健康素养，特别是在发达国家。然而，老年人口较多或处于老龄化社会的国家/地区往往具有较低水平的电子健康素养水平。影响较高 eHealth 素养的因素包括：行动因素：互联网使用（β = 0.24, 95% CI 0.06－0.42, I2 = 87.1%）；决定因素包括：（i）年龄（β = -0.11, 95% CI -0.19－ -0.01, I2 = 80.3%），（ii）种族（β = -2.61, 95% CI -4.11－ -1.11, I2 = 80.2%），（iii）收入（β = 0.49, 95% CI 0.11－ 0.86, I2 = 82.3%），（iv）就业状况（β = -1.63, 95% CI -2.32－ -0.95, I2 = 99.7%），（v）教育（β = 0.18, 95% CI 0.16－0.19, I2 = 58.2%）；个人能力：感知有用性（β = 0.83, 95% CI 0.13-1.52, I2 = 68.2%）和

自我效能（β = 0.18, 95% CI 0.13-0.23, I2 = 0.0%）；健康状况因素：疾病（β = -0.15, 95% CI -0.8- -0.16, I2=84.2%）。

结论 大多数国家/地区的参与者表现出足够的电子健康素养。而电子健康素养影响因素是难以改变的社会健康决定因素和人口统计因素。因此，我们建议未来的 eHealth 技术可以通过根据用户特征对消费者进行分类，采用更有针对性和适应性的方式来开发。此外，可以实施旨在提高消费者自我效能并改善其对应用程序的感知有用性的干预措施，以促进其电子健康素养水平。

PU-263

创建高级别电子病历助力医疗质量管理实践与创新

侯彦东
山西省汾阳医院信息中心

2018 年 12 月 7 日，国家卫生健康委办公厅发布了《关于印发电子病历系统应用水平分级评价管理办法（试行）及评价标准（试行）的通知》；《通知》要求，到 2020 年，所有三级医院要达到分级评价 4 级以上，二级医院要达到分级评价 3 级以上。新标准对医院电子病历系统的功能提出了更为详细和具体的要求，更加注重了医疗质量管理中的电子病历所具有的功能。《国务院办公厅关于建立现代医院管理制度的指导意见（国办发[2017]67 号）》提出，医院要建立覆盖临床诊疗服务全过程的医疗质量管理与控制制度；加强重点科室、重点区域、重点环节、重点技术的质量安全管理；严格落实手术分级管理、抗菌药物分级管理、临床用血安全等医疗质量安全核心制度。《医疗质量管理办法》中也明确提出，要强化基于电子病历的医院信息平台建设，提高医院信息化工作的规范化水平，充分利用信息化手段开展医疗质量管理与控制。山西省汾阳医院在基于电子病历为核心的信息化建设过程中对标电子病历五级功能要求，充分利用互联网+技术，实现了高效、精准、实时的医疗质量管理和控制，在促进医院突破传统医疗质量管理模式，创新智能化、信息化的质量管理新模式，在提高医疗质量控制水平、改善医疗安全、提高医疗信息资源利用等方面有了新实践。

1. 传统医疗质量管理模式存在的问题

现代医疗质量管理实现的是全程质量管理，包括基础医疗质量管理、环节质量管理、终末医疗质量管理；但由于各地医院管理和技术水平不一，条件受限，无法实现全流程管理。存在管理的片面性，监管范围少，局限于病历、药品、感染，未能覆盖全部诊疗过程，监管范围覆盖不全，只能用抽查方式以点盖面。滞后性，偏向事后统计，通过事后统计发现问题，不能在事前、事中及时有效的警示临床，越级行为事后发现，缺少医师资质判断。被动性，质控管理与临床医生脱节，交互性差，不能实时掌握重点病人动态。单一式临床事中质控功能简单、表象、事中质控与业务流程相关性低，缺乏结构性、逻辑性、内涵方面的质控。数据利用缺失，数据收集零散，没有有效的统计学价值。以上问题严重制约着医院的医疗质量管理水平。

2. 应用电子病历系统控制医疗质量的思路

以优化医疗质量管理流程为目标，以落实 18 项医疗安全核心制度为抓手，以诊疗指南、临床路径为指导，通过电子病历系统应用的流程节点为基础，进一步规范诊疗行为，建立全新的院内质量控制体系，实现事前预防、事中监控、事后控制的全流程质量控制体系，以数据管理为依据，综合评价质量安全管理效果，为医疗质量和安全管理提供有效工具，提升医院整体质量安全管理水平。

3. 基于电子病历医疗质量控制的应用实践

3.1 依托电子病历加强质量管理，优化管理流程

3.1.1 加强病历管理制度落实：住院病历的质量管理是医院质量管理的重点工作，我院在住院病历的管理中实现了事前、事中、事后的全流程病历质量控制体系。事前管控：实现病历的不同内容的时限质控，超期病历医生站弹框提醒，时限超过 2 倍时间病历锁定，需医务科解锁；实现病历内容的逻辑质控，比如：对男性病人病历记录中，出现女性的生理、病理描述；体温、血压、心界等超出合理范围等，出现不符的情况医生站会弹出提示。对缺项、漏项进病历保存时会进行提示。事中管控：实现病历的环节质控，质控员通过系统进行质控，质控结果及时通过医生站、短信或微信通知主管医生，及时进行整改；出科病历实现出科病历的一级质控评分。事后管控：实现病历的终末质控，院级质控医师通过系统打分，结果及反馈内容同样通过医生站、短信或微信及时通知主管医生。此外对于新生儿科、产科、神经内科等科室设置了专科化结构化模板，专科化评分内容。对于超期病历、环节、终末质控评分可进行统计分析，为医务科提高病历质量提供管理依据。

3.1.2 加强临床科室运行质量管控

3.1.2.1 通 CIS 系统实现了临床科室人员分级管理，根据不同级别的医生，给予了不同的权限，例如医生的中西药处方权，精麻毒药品处方权，抗菌药物分级权限，用血权限，会诊权限等。

3.1.2.2 在会诊管理制度落实方面，在传统会诊流程注重医疗信息采集的基础上,通过引入会诊审批功能、会诊后受邀、邀请科室评价功能，加强了对医院科间会诊的管理，提高了医院科间会诊的效率和质量。同时我们实现了通过短信、微信等通知医生及时会诊、患者等候会诊。

3.1.2.3 在合理用药方面，一是抗菌药物分级管理制度落到实处，我们不仅有权限管控，还实现了特殊抗菌药物的在线申请，会诊审批等功能，通过信息技术实施抗菌药物临床应用管理，其中包括抗菌药物监测数据信息发布，实现院、科两级抗菌药物使用率、使用强度等指标实时统计、分析、评估和预警。二是：实现了信息系统处方前置审核、处方点评，不合理处方用药设置了风险级别，在医生下达医嘱时进行提醒，超五级以上的风险提醒由临床药师进行干预，加强了临床药师与临床医生的沟通；每月对科室处方可以进行自动汇总点评，对特定药品、人员可以单独点评；提升了临床药学工作的服务水平。

3.1.2.4 手术管理方面不仅实现了分级管理，同时实现了手术的审批，重大手术的上报，并且医务管理部门可以对手术量、手术级别、重大手术、非计划手术进行质控监测。

3.1.2.5 临床用血管理方面，在实行分级管理的基础上，实现用血审批根据同一患者一天（24 小时内）申请备血量：大于 800ml 小于 1600 科主任审批，大于 1600ml 医务科审批。同时开展了用血质量监测：按科室、血液种类等指标统计输血量使用情况；统计不同科室的不同种类输血的费用情况；输血文书的专项监管，输血申请单文书缺失率、知情同意书数缺失率、输血记录文书缺失率、输血反应记录缺失率；24 小时内输血量≥1600ml 的用血申请率、审批率等。

3.1.2.5 临床路径管理方面，实现了电子化的临床路径诊疗流程，诊疗表单，医务科具有对科室申请的路径审批管理功能，审批路径标准和表单，确保合适病种进入临床路径；对各科室临床路径进行监测，包括路径工作管理监测：入径率、完成率、变异率、退出率；路径质量管理监测：治愈率、好转率、死亡率、15 天再住院率、31 天再住院率、非计划、再次手术率、院内感染率、手术部位感染率；路径变异与退出监测：路径经济效益监测：临床路径变异原因分析、退出原因分析；平均住院日、术前平均住院日、平均费用情况对比、路径患者平均费用、疗程平均费用等。

3.1.2.5 治疗管理方面，充分利用信息化优化康复、血透等治疗流程，通过血透管理系统提供面向科室运营的管理服务、面向医护人员的医疗和护理服务。同时全面实现与透析机、体重秤、血压计等硬件设备互联互通，以及院内 EMR\HIS\LIS\RIS 等系统的数据共享。引进康复治疗系统，实现了从治疗预约、评估到治疗记录的全院数据共享。通过系统对治疗数量、质量等进行相应的分析，解决了治疗工作效率低、评估难的问题。

3.1.3 加强医技科室运行质量管控：医技科室质量管控一直是各个医院的困惑，我院充分利用电子病历 PACS 等系统的优势，实行了检查报告的结构化，可对检查报告和检查图像进行在线质控，质控结果形成相应的分析。检验科室可以通过 LIS 系统，开展了相应的室内质控，形成了失控

预警机制，上线了流水线检验系统，按照事先设定的审核规则实现了报告自动审核。病理科通过病理系统对术中冰冻诊断符合率，术中诊断及时率，染色切片、组织诊断情况进行分析，对免疫组化质量进行分析。医技科室业务信息化以后极大的提升了工作效率，也具备了相应的质量管理的手段。

3.2 智能辅助，提升安全避免差错

建立了完善的知识库系统，包括静态知识库：覆盖疾病知识（9000+）、检查知识（1100+）、药品知识（30000+）、手术知识（1100+）、护理知识（500+）、医学评估量表（150+）等；动态知识库：检验规则、检查规则、手术规则、报告解读、药品、输血、膳食等形成知识图谱反哺临床；病历内容智能质控：根据主诉、诊断等内容，对病历内容进行质量控制；检查检验合理性校验提醒：综合过敏信息、诊断、性别、年龄等内容自动提示适应症、注意事项等；用药、膳食类合理性校验：综合诊断、性别、年龄的内容自动提示药品、膳食等注意事项；手术合理性校验：综合药品、诊断、性别、年龄的内容自动提示手术、操作适应症、禁忌等。避免了临床医护人员产生医疗差错行为，降低医疗安全风险。

3.3 数据有效利用，提升管理效能

通过完善的，覆盖全业务流程的电子病历系统以及医疗闭环质量管理体系及智能辅助功能的建成，院内形成了完整的数据中心。建立了运维管理平台，通过对门急诊、住院重点部门、重点环节、重点病人的数据管理和分析，为管理者提供了更加方便、准确、及时的决策依据。目前我院建有：质量管理平台，着重医疗质量运营管理数据的管理和分析；门急诊管理运营平台，可以对预约情况、出诊时间、门诊流量、门诊手术、门诊病历、急诊人次等进行实时监控；住院管理运营平台，可以对住院病人数量、危急重病人、手术病人、费用情况等进行实时监控。为医院决策部门提供了数据支撑，使决策更为科学化，管理更为精细化。

4. 总结：高级别电子病历的评价标准目标明确，就是为医院提供高级别的临床决策支持功能，我院通过高级别电子病历的创建，进一步优化了医院的质量管理体系，创新了线上和线下一体化的管理模式，使医院医疗质量的目标管理更加程序化、可操作化、在管理上提高了一个层次。为医院迈入高质量发展的轨道上奠定了坚实的基础。

PU-264

基于价值医疗视角下医后付服务应用研究

汪滔 彭杰

连云港市东方医院

0、引言

2010 年 9 月，原卫生部委托北大人民医院探索优化服务流程，试运行“先诊疗后结算”服务模式；2013 年“先诊疗后结算”服务模式再次被写入卫生部年度工作要点，并在一些试点医院推行“先诊疗后结算”服务模式；2023 年 05 月 26 日，卫生健康委和中医药局联合发布《改善就医感受提升患者体验主题活动方案（2023-2025 年）》提出探索推行“医后付”服务模式，特别强调门诊就诊简化流程、创新模式，提升患者门诊体验。

本文基于价值医疗视角，选取连云港市市直医院（以连云港市东方医院为例）的“医后付”服务为研究对象，对其开展的“医后付”服务建设应用情况进行分析研究。理论上，可进一步丰富便民化医疗服务供给的价值研究；实践中，可减少缴费次数，提升患者体验，实现就诊流程，完善信用体系建设及加快“医后付”服务推广提供借鉴[1]。

1、基于价值医疗视角下医后付服务

价值医疗是美国哈佛大学商学院管理学迈克尔·波特教授在 2006 年出版的《Redefining Health Care Creating Value-Based Competition on Results》中提出的概念，是一种追求最高性价比的医疗，也就是以相同的或者更低的成本提升医疗效果，其核心是提升医疗质量和医疗效果和降低医疗费用[2]。而健康管理结果包括健康产出、治疗效果、与其他治疗方法的比较效果和服务可及性、服务对象的满意度。“医后付”服务本质上是一种信用就医行为，通过先享受“医后付”服务，再评估患者信用资质，这样在诚信就医和文明城市建设的基础上，增进医患互信，摆脱以往“医后付”先授信额度，才能享受的局限性。

从政府部门角度看，基于价值医疗视角下医后付服务是积极落实国家深化医药卫生体制改革工作要求，是着重推进医院信息化建设等便民举措，通过实施“医后付”切实解决长期以来患者就诊缴费“堵点”问题，极大地提升了患者的就医体验。从而加快便民、利民、惠民的高品质医疗卫生服务体系建设，凸显价值性。对于医院而言，门诊实行“医后付”服务将提高信息化建设，提高服务供给效率，减少人群医院逗留时间，同时门诊“医后付”就医简化就医流程，提高患者满意度，将推动“以患者为中心”全流程门诊服务价值的实现[3]。

连云港市卫健委直属医院在 2022 年 5 月份正式启动“先诊疗后付费”试点工作，并在 2022 年 10 月正式更名为“医后付”。“医后付”是在原有付费模式的基础上，充分尊重患者的选择权，在不设患者使用门槛前提下新增加的一种纯信用服务方式，在该服务下，病人无需诊前预存押金，只需在挂号时与医院签订“医后付”协议，即可享受诊疗结束后一次结算，诊疗费用选择在医院或者回家后将费用一次结算。患者在医院门诊、急诊看病可跳过缴费环节先就诊，仅需当日就诊结束离院前到自助机或收费窗口刷社保卡完成医保缴费或回家后通过移动支付缴清即可。若用户发生逾期支付，医院则会通过短信向欠费患者进行催收，并暂停其“医后付”服务，但并不影响其正常就医；对于超过 7 天欠费逾期的患者，还会有人工催收电话进行再次催收；若逾期催收超过一个月但仍未补缴的患者，医院会将就医信息上传至连云港市公共卫生信用平台，对拖欠费行为起到信用惩戒和警示作用[4]。

2、“医后付”服务系统建设

依照连云港市卫健委的文件精神，连云港市东方医院成立了多部门组成的工作专班，召开专题会议，梳理问题清单并出台解决方案。在充分尊重患者的就医习惯的前提下，尽量做到系统“无感”改造。设计思想是以自助机、医院小程序电子签约为主，人工为辅的多种签约方式并存，随着签约人数的增长、“医后付”服务全面铺开，最终以自助机、医院小程序为主要签约途径，从而实现无纸化、信息化数字签约。

2.1、信息系统流程改造

信息系统的改造是“医后付”工作的重点和难点。医院先后改造了多个系统和医院小程序，实现了指引单打印、门诊收费项目确认等功能，由“预付费”变为“后付费”模式，欠费提醒也更加人性化、智能化。患者可选择自助机、人工窗口、手机移动端等任一途径签约“医后付”，并由手机号码二次验证，确保签约患者实名就诊。

2.2、新增“医后付”标志

为方便医生区分“医后付”签约患者和使用“医后付”的患者，在医生工作站患者列表新增签约标志和使用“医后付”标志；只有当签约标志和使用“医后付”标志都标记的情况下，本次诊疗才可享受“医后付”服务；针对已经签约的患者，医生也可在征得患者同意的情况下，在本次诊疗过程中使用“医后付”，并通过指引单进行拿药、检查等。

2.3、收费确认改造

针对科室自做医技项目，可以通过收费确认模块完成项目确认收费，这样有效解决医生开了医嘱项目而没有检查的问题；对应的 LIS、PACS 等系统也进行了全面改造，药房、医技科室及所有治疗相关科室全部上线收费确认功能，确保“医后付”催款工作“不错催，也不漏催”。提升患者就医体验，提高医院满意度。

2.4、隔日缴费功能改造

在使用“医后付”服务过程中患者因某种原因未能及时当天缴费,隔日缴费功能能将患者的缴费信息保留，患者可以通过人工、自助机和医院小程序多种渠道进行缴费。

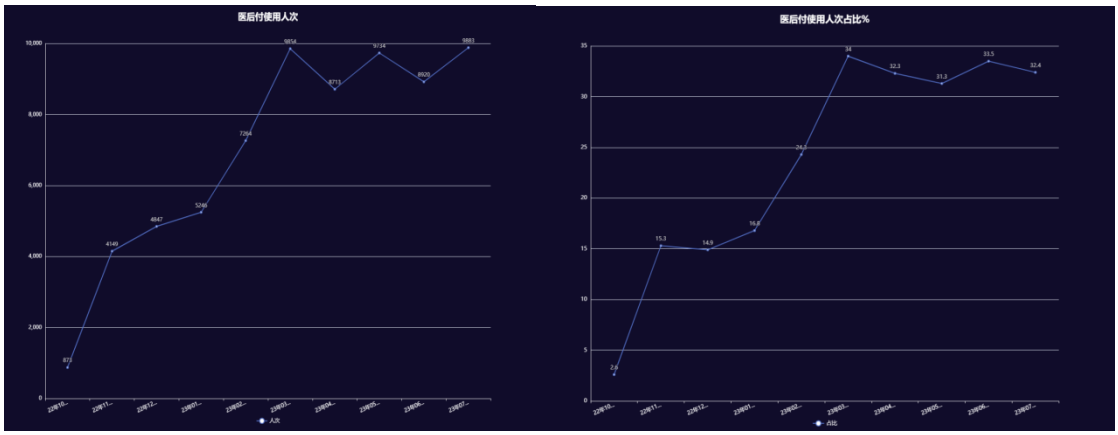
2.5、欠费查询系统建设。

欠费查询系统是“催款”工作的基础，医院通过对信息系统的升级改造，建立了一套完善的欠费查询系统，将患者的就诊记录、医嘱执行情况、缴欠费情况、检验检查结果等内容统一到一个操作界面，欠费管理人员可以方便地了解患者在医院的就诊全过程和欠费情况，为“催款”工作提供数据依据。结合市卫健委联合惩戒平台，提高欠费和催款的综合管理能力，实现“查、催、改、提”四个环节的全流程管理。

3、“医后付”应用效果

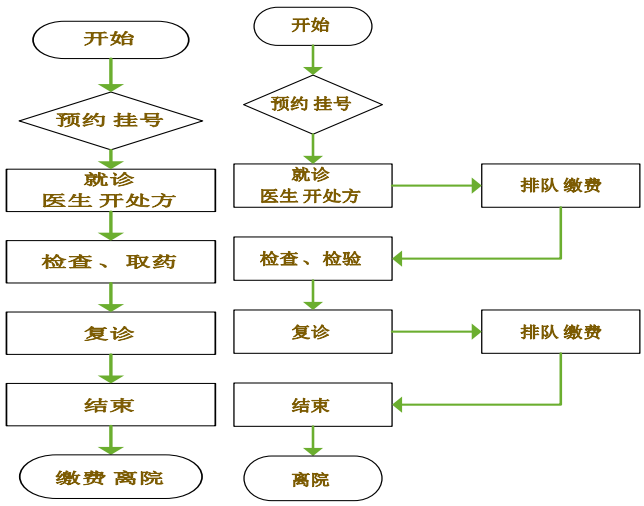
3.1、“医后付”运行情况

目前“医后付”服务模式已在连云港市东方医院门急诊全面应用，通过院内外多媒体宣传和工作人员宣教，来院患者对“医后付”服务的知晓率有了很大提高。截至 2023 年 7 月，连云港市东方医院“医后付”签约人数近 4 万，使用约 7 万人次，近 5 个月的使用率一直稳定在 35%左右，使用金额约 1800 万元，金额占比约 39%，欠费 7 天以上金额约为 1.2 万元，欠费金额和人次保持在可控范围之内。



医后付月使用人次和人次占比

3.2 优化了就医流程



医后付就诊流程传统的就医流程

研究表明,患者在就诊过程中,仅排队的时间就占总就诊时长的 70%左右，传统门诊就医管理方式已无法满足患者及时就医的迫切需求,尤其针对急危重症患者,传统繁琐的就医流程有延误救治时机的风险。相比较于传统就医流程，“医后付”的服务模式如图所示，可以大幅度减少患者就医时的缴费次数，减少排队次数，不用反复排队，缩短了就诊时间,让患者能够更加专注于医疗过程，增加了就医的便捷性。

3.3 减轻患者经济负担

在传统模式就医过程中，在无法预估诊疗费用的情况下，患者一般准备远高于本次诊疗的费用，而有时候由于种种原因，患者并没有享受部分的诊疗服务，造成了患者的经济上的浪费。“医后付”服务可以让患者根据实际的诊疗情况，在诊疗结束后，医院或者回家后一次性支付真正接受的医疗服务的费用即可，有效的减轻了患者的经济上负担。

3.4 提升医院满意度

医院成立了由主要领导、分管领导、门诊、财务、信息、审计、宣传等多部门组成的“医后付”项目领导小组和工作专班，定期召开“医后付”专题会议，及时梳理问题清单并讨论解决方案。通过对系统和服务流程的不断改进和完善，形成一套 PDCA 循环。医院在门诊患者中还开展了“医后付”问卷调查，收集患者真实体验和意见建议，持续改进“医后付”相关工作，切实提升服务内涵和患者就医体验。

4、结束语

价值医疗，其核心为以患者为中心，提高患者满意度是体现价值医疗最有效体现，连云港市市直医院实施“医后付”的门诊流程，依托了本地医保制度和市卫健委统一的信用平台，积极响应国家、省、市卫健委关于改善医疗服务的相关要求，遵循了方案论证、先行试点、全院同步、逐步实施的改造思路，较传统就医模式至少缩短了患者 60%缴费等待时间。同时延时支付缴费，更是解决了部分患者资金短缺的困境，带来一种从容就医新体验。窗口收费环节得到改善，就诊高峰期窗口排队现象得到很大程度的缓解，有效解决“缴费时间长”问题。根据使用“医后付”患者的回访，一致认为此模式提供了实实在在的便捷服务，对提高患者对医院的满意度起到了非常积极的作用。医院也结合催款过程中患者反馈的问题整改，实现该项工作的持续改进和群众满意度的不断提升。为探索适合我国国情的“医后付”服务提供参考，积极为“健康中国”建设贡献东方力量，努力在中国式现代化新实践中展现东方担当!

PU-265

基于先进医院信息系统的智慧医院建设

崔建国
濮阳市人民医院

1 引言

我国医院信息化经过近 30 多年的持续发展，大致经历了小型机的摸索、微机初级应用、全院规模管理信息系统、临床信息系统与区域医疗协同摸索、基于电子病历的医院信息平台、人口健康信息化等 6 个发展阶段[**错误!未找到引用源。**]。伴随着医院临床要求及管理需求的不断增加，以及信息化和大数据技术的不断发展，医院内部使用的各类系统可多达上百套，部分医院内部依然存在着“信息孤岛”[**错误!未找到引用源。**]、“数据鸿沟”，即由于系统间不互联不互通导致的数据及业务流程的割裂；另一方面，随着医院规模和体量的发展，医院的运营管理和医疗协同亟需先进的信息系统作为支撑，以实现跨平台、跨系统的信息交换和业务联动。

在深化医改的背景下，公立医院需突破原有空间限制，转向高质量发展格局的需求更为迫切。因此，如何利用一体化的信息系统、以信息化手段支撑临床业务协同、提高患者满意度及以数字化赋能医院运营管理决策，具有重要的研究价值和意义。

2 新一代医院信息系统的特点

医院信息系统(Hospital Information System, HIS)是智慧医院建设的信息化和数字化基石，对医院的医疗服务质量与运营管理水平的提高起到了无可替代的作用[**错误!未找到引用源。**]。为有效推动医院在智慧医疗、智慧服务、智慧管理等方法的发展，对医院信息系统也有了更高的要求，具体体现在：

2.1 数据标准化

高质量的数据是智慧医院建设的核心基础，虽然有各种技术手段可以对数据进行治理，但从源头上对数据进行结构化、标准化具有不可替代的重要意义，此外，各信息系统之间的资源共享互联互通使得数据标准统一化的工作显得尤为重要[**错误!未找到引用源。**]。因此，新一代信息系统需要有完善的底层信息模型的设计，使用规范的、符合国际国内标准的数据模型，一方面确保数据可利用、可分析，同时为高效的互联互通打好基础。

2.2 业务一体化

新一代的医院信息系统需要避免以往由于多套系统导致的业务割裂、无法互联互通、无法形成闭环的问题，着力实现 HIS+EMR 一体化、门诊住院一体化、院内院外一体化、线上线下一体化、固定移动一体化、多院区一体化，在确保医疗业务场景连续的同时，也避免了医护人员反复切换不同的业务系统，有效提升工作效率、优化资源配置、降低服务成本[**错误!未找到引用源。**]。以电子病历系统为例，它以患者为中心，建立面向专科专病的电子病历，将患者在医疗机构中的历次就诊时间、就诊原因、针对性的医疗服务活动以及所记录的相关信息有机地关联起来，并对所记录的海量信息进行科学分类和抽象描述，使之系统化、条理化，进而实现了结构化、标准化的数据采集、存储和管理，适合临床、教学、科研、患者服务、外联协同等不同应用场景，覆盖评估、预测、计划、干预、促进、评价全闭环过程的，实现了对病人医疗和健康数据的全面收集、及时管理，实现高层次的统计学数据分析和数据挖掘。

2.3 临床智能化

充分利用规则库、知识库、大数据、人工智能等技术手段，以智能化的方式赋能临床业务，在提升医护人员工作效率的同时，有效避免不规范、不准确、乃至临床事故的发生。例如，合理用药系统实时进行用药风险检测、对诊断与用药进行匹配检测、对风险用药进行拦截；临床决策支持系统（CDSS，Clinical Decision Support System）可以基于患者差异化的病情辅助医生做出“正确”的决策，排除非正常的诊疗风险，在提高临床效率的同时，也有助于对年轻医生的成长和知识积累，同时，通过在电子病历界面设置卡控分级管理等功能,能够更好地提高临床规范性,不断提升医疗质量管理规范落实的可操作性[**错误!未找到引用源。**]。

2.4 数据驱动

智慧医院所需要的不仅仅是将流程和数据通过信息化和数字化的手段进行实现和存储，更重要的是基于数据进行主动式的分析和辅助决策，以数据驱动医院运营管理效能和医疗行为质量的提

升。因此，一方面需要进行医院数据中心的建设，基于数据质量体系实现对临床数据、科研数据、运营数据的有效管理；另一方面要基于对数据的分类整理和分析，将隐形的信息显性化、碎片的信息结构化，形成对医院医疗和管理行为的洞见，以数据驱动医疗质量的提升、用路径化的方式进行以患者中心的全生命周期健康管理。研究表明，以数据驱动决策,赋予医院业务新应用,对提升健康医疗服务效率和质量等都具有重要意义[**错误!未找到引用源。**]。

2.5 国产替代

近年来的信息系统建设浪潮中,越来越强调基于全国产化生态环境下的信息系统自主可控性要求[**错误!未找到引用源。**]，医院开始根据业务场景也逐步对相应系统进行国产化替代，并努力达到“应替尽替，真替真用”的目标。在此大背景下，新建的医院信息系统应当满足客户端、服务端、数据库等能够全面支持国产化的要求，可以根据医院所使用系统的国产化进程可灵活选择和切换。

3 新一代医院信息系统赋能智慧医院建设

为全面推进健康中国建设，更好地满足人民日益增长的医疗卫生服务需求，国务院办公厅于2021 年 6 月印发《关于推动公立医院高质量发展的意见》，以信息技术作为强有力的发展支撑，推进电子病历、智慧服务、智慧管理“三位一体”的智慧医院建设和医院信息标准化建设有序进行[**错误!未找到引用源。**]。智慧医院的建设需要以信息系统为抓手，从医疗、服务、管理三个维度进行突破，以评促建，不断推动医院高质量发展。

3.1 智慧医疗

以电子病历为核心的临床信息化涉及临床业务的全流程，通过电子化流程改造，优化和重塑临床业务，应重视与完善流程的闭环管理，并需在基于知识库的智能提醒中扩大应用范围和场景，用以减少人为差错。同时，对数据质量的治理与利用是电子病历建设的最终使用落脚点[**错误!未找到引用源。**]。计虹等[**错误!未找到引用源。**]建议，电子病历建设过程中需要重点关注以下关键点：1）深度集成化便捷医生工作台；2）大数据支持的智能临床决策辅助；3）覆盖全链条临床闭环追溯管理；4）基于可信数字的全流程自动归档体系；5）加强数据治理驱动数据利用。通过完善电子病历系统的建设，确保源头数据的准确性、真实性，进而促进对数据的利用和基于患者全生命周期的数据提供精准医疗、数字医疗服务；与此同时，基于 AI、自动化方式对医疗流程进行智能决策支持、智能质控、自动化存档等，可以有效提升医疗行为的准确性、提高效率、降低成本。

3.2 智慧服务

随着“互联网+”、物联网设备、一体机等智能硬件的不断深入应用，患者服务场景逐渐由院内延伸至院外、服务路径由线下延伸至线上，在满足患者医疗需求的同时，提供更加人性化、差异化、智慧化的服务。智慧服务已成为重要发展趋势,有利于医疗服务质量的显著提高，有助于患者就医体验的显著提高[**错误!未找到引用源。**]。在此过程中需要医院信息系统具备足够的开放性和对固定、移动等不同数据源的承载和处理能力，并基于对数据的统一利用、服务的贯通，形成一体化、多样性的服务模式。例如，在常规在线建档、在线预约、在线取号、在线挂号、在线查询等基础服务之上，提供多样化的支付方式（小程序支付、诊间支付、“先看病后付费”信用支付、线下窗口支付等）、便捷的预约方式（在线检查预约、在线体检预约）、互联网医院服务（在线咨询、

在线问诊、一键续方、在线药房、处方流转等）、以及其他丰富的个性化服务（线上订餐、送餐服务，给特殊病患的定制营养膳食套餐、停车信息推送等）。

3.3 智慧管理

有了数据的持续支撑和驱动，可以在人、财、物、科研、教学等多方面多维度通过数据分析赋能医院管理，将原本不可感知、不够清晰、没有规则的信息基于数据进行量化、具象化、可视化。并进一步基于数据提供智能辅助决策分析[**错误!未找到引用源。**]，可以基于涵盖关键绩效指标（Key Performance Indicator，KPI）进行多主题分析与展示，便于管理者从门诊、住院、医技、药品、耗材等全方面准确掌握医院的运营状况，并可以分科室进行汇总展示。同时，基于信息系统的数据互通，实现财务业务一体化联动，支持全面预算管理、发票付款与费用核销、合同线上流转管理等。通过医院数据资产的有效积累，也为基于数据驱动的科研、教学提供了扎实的基础和丰富的真实世界数据与素材，可以缩短医生的培养周期，提升科研产出的速度和质量。

4 展望

随着信息技术、人工智能、互联网及物联网等技术的飞速发展，将进一步夯实医院高质量发展的数字新基建，一体化信息系统与数据中心的建设将逐渐使更加广泛的医院受益于技术的发展、同时也为医生的临床业务、医疗协同、科学研究、教学管理等方方面面带来实质性的便利，通过数据驱动的方式提升医生的工作效率，尤其是将医护人员从繁琐的、重复性的日常工作中解放出来，专注于专业性更高的医疗及护理业务。未来，人工智能、大语言模型、数字孪生等技术都会更深入融合进智慧医院建设中。例如，随着大语言模型的成熟和深化应用，有望进一步改变医护人员的工作模式和方式，以大模型作为医护人员的得力助手，例如对于问诊过程中病历内容的智能理解、总结、填写，医生只需要进行核对并确认即可，将显著提升病例录入的效率与质量。

5 结语

本文从信息系统建设与效果的角度，分析了“三位一体”的智慧医院建设过程中对新一代信息系统的要求，并从智慧医疗、智慧服务、智慧管理的角度阐述了如何以新一代信息系统为抓手进行智慧医院的建设、怎样以新技术新方法新模式赋能医疗效率与质量的提升，并对未来的发展进行了展望。当前，人民群众的医疗需求日益增加，国家和有关部门对智慧医院建设的要求也逐步提升，医院需要充分了解和利用信息技术和数字技术推动医院高质量发展。

PU-266

影像检查智慧服务流程优化实践

汪伟，郭晓宙，王海星，张靓
山西医科大学第一医院信息管理处

近年来，为改善人民群众就医体验，医疗机构充分发挥信息技术的支撑作用，创新建立智慧医疗服务模式。随着诊疗技术的发展，医技影像检查对于疾病的诊治发挥着重要作用，由于临床诊疗对影像检查结果的依赖性以及患者数量的上升，导致医技检查工作量不断攀升。影像检查的全流程涉及各临床科室、各医技检查科室、财务、医辅等多部门交叉配合。因此，患者完成检查环节多，办理手续复杂。优化检查流程，节约患者时间，让患者少跑腿，是智慧服务流程优化的主要目标。对影像检查进行全流程智慧服务改进，设计流程更科学、服务更高效的检查流程，是提升医疗服务

质量的重要任务。

1 影像检查服务现状

大型综合性医院的医技影像检查项目多、资源分布广，影像检查资源管理不统一，患者检查过程中往返各科室次数多，预约、缴费、取报告流程繁多，给患者诊疗带来不便。当前检查服务流程主要存在以下问题：第一，医院影像检查的预约流程多为“多站式”流程，各部门分隔明显，患者往返次数多，等候时间长的矛盾突出，整体预约流程耗时且耗费医技科室人力资源。第二，影像检查未实现全流程信息化覆盖，部分环节存在“断点”，有待提高手续办理的自助性和便捷性。第三，影像检查资源需统一管理，提高资源利用率，满足患者多样化的检查需求。

随着医院智慧服务流程的不断完善，影像检查全流程服务已趋于“线上+线下”一体化协同模式，通过信息技术建立临床、医技、财务等相关环节的信息互联互通，以创新改进影像检查智慧服务流程，促进影像检查高质量发展。本文所述流程优化主要在 CT、超声、核磁、普放四项影像检查中开展。

2 影像检查智慧服务流程优化措施

为了优化影像检查流程，提高影像检查服务的智能性、便捷性，减少患者往返次数，缩短患者影像检查等待时间，我院借助信息化手段，从申请单开立、检查预约、诊间缴费、报告获取、复诊预约、全流程信息引导等方面进行智慧服务优化。主要实施 4 项优化措施，包括：开发智能预约平台进行检查自动预约、升级自助服务提高诊间服务便捷性、丰富互联网医疗功能延伸服务范围、提供引导式消息推送服务全流程覆盖。

2.1 设计自动预约流程，开发智能预约平台

我院利用信息化手段设计运行智能化检查自动预约平台（简称“平台”），实行集中预约、分时段预约和碎片预约结合的预约诊疗模式，以此整合全院影像检查资源，将人工预约向智能化预约转变，提高预约效率。平台统一管理全部预约资源，供线上线下患者同时使用，提高影像检查资源的使用率。

2.1.1 预约流程设计

平台面向门诊患者、住院患者、互联网医院患者提供影像检查自动预约服务。系统通过查询检查申请单信息，匹配可预约资源，实现预约、改约、退约，返回预约导引单的功能。

门诊患者由医生开立检查单后，患者可获知检查时间后再缴费，患者于 30 分钟内缴费，则锁定此检查时间并打印导引单，包含预约检查名称、检查时间、检查地点及注意事项。若患者未缴费，则系统将自动释放号源，待缴费后重新进行检查预约。住院患者由住院医生开立检查医嘱，护士执行医嘱时进行自动预约，患者于预约当日到指定检查科室进行登记、检查。

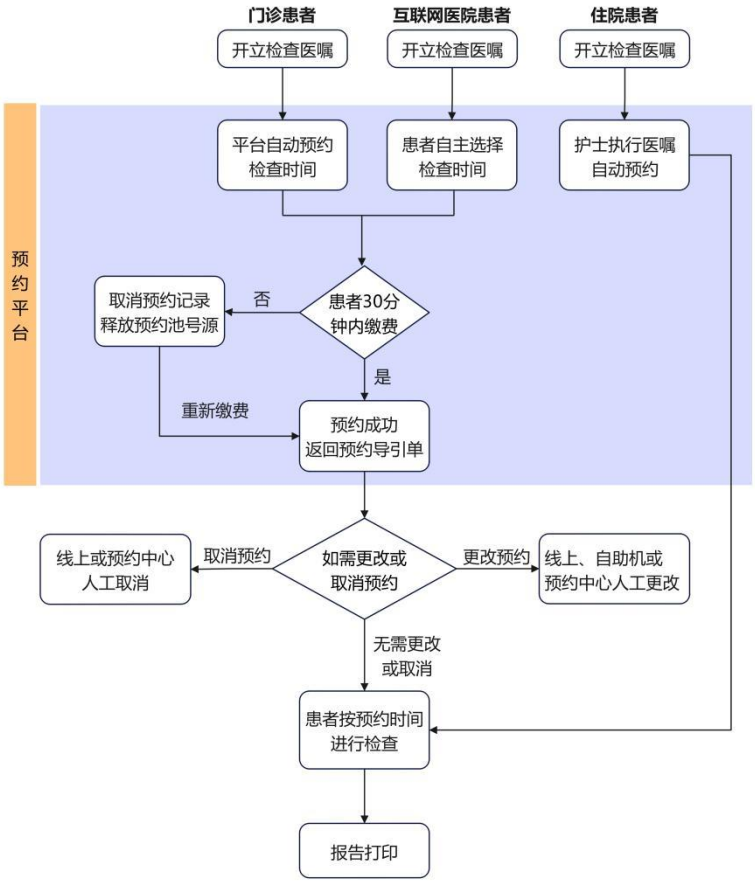


图 1 智能化检查自动预约流程

2.1.2 预约资源统一管理

平台整合全院可预约检查号源，根据患者的检查类别、设备的排班,自动完成患者的各项检查预约。经过前期科室调研，按照往期各项检查工作量，预设各检查项目的自动预约号源数。以检查设备为单位，分时段、分项目设置可预约号源数。同时，为了应对紧急需求，系统为每台设备设置加号数量，以应对危急重症患者应急需求。

2.1.3 预约规则设计

根据影像检查实际应用场景，平台设计多种预约算法和预约规则，包含多重条件，如检查项目、检查部位、检查用药、单项检查时间、特殊要求等，以此设置各项目之间的合并规则、互斥规则、分时段规则。除此之外，根据医院实际情况，可相应设置优先原则、地理位置就近原则等。

（1）互斥、合并规则

互斥规则是为了安排同一患者多项检查的检查时间和检查顺序，避免检查时间冲突，形成合理的预约时间和顺序。如将同一患者的 B 超检查和 CT 检查互斥预约到不同时间，需要使用药物的检查项目与其他检查间隔规定时间，或按照检查顺序先安排不使用药物的检查。合并规则是将同一检查类型的不同部位进行合并，减少患者往返次数和等待时间。如胸部 CT 和头颅 CT 合并，患者只排一次队。当医生向同一患者开立多条检查医嘱时，平台可有效触发检查互斥规则及相同检查的合并规则，进行检查排程。

（2）分时段预约规则

分时段规则是根据检查项目的特殊要求分时间预约。如将超声检查需要空腹的项目预约在早上固定时间段。在分时段的前提下，如有设备空间，进行碎片预约，提高设备使用效率，缩短患者等

待时间。

（3）个性化预约规则

一是位置就近预约规则。平台根据院内科室位置分布，智能化就近匹配预约资源，使患者往返路程最短。二是特殊患者的优先预约规则。优先规则是为了给急诊患者、绿色通道患者、小儿患者提供优先预约。平台区分预约渠道，为急诊患者无资源预约、绿色通道病人无主索引、无资源批量预约登记，小儿患者优先排期。

2.2 升级自助服务，优化诊间流程

为探索“一站式”自助服务模式，简化患者就医流程，减少患者等待时间，医院通过升级自助服务，开通影像检查的诊间自助医保支付、自助报告打印服务，提高影像检查的诊间服务便捷性、高效性。

2.2.1 开通多渠道医保支付方式，节约缴费等待时间

以往的就医过程中，医保支付多依赖人工窗口刷个人医保卡结算。医院通过全面升级自助服务，对接医保平台，改造计费系统，使用消息推送技术实现了了手机端医保支付服务功能，此外还开通了自助机端医保支付服务、窗口医保“刷脸”支付服务。减少患者在影像检查过程中的跑腿缴费、排队缴费，提高就诊便捷性，提升患者就医体验。

2.2.2 检查报告线上查询、自助打印

患者做完检查后无须原地等待报告，也无须往返多个医技科室拿取报告，可通过医院微信公众号绑定就诊人信息，在线查询就诊人检查报告。此外，就诊人还可自主选择方便的时间在自助机端一次性打印影像检查报告。此措施一方面减少医技科室人员拥挤现象，另一方面方便患者灵活安排时间，无须长时间等待。

2.3 拓展互联网医疗功能，线上咨询、复诊预约移动端完成

为了方便异地患者、复诊患者就医，自主选择来院时间，接受影像检查的患者还可以选择互联网医院预约检查、改约检查、预约复诊、线上咨询。互联网患者由互联网医院开立检查单，患者自主选择检查时间，进行在线改约、退约。患者还可以根据需要，选择互联网医院图文报告咨询，减少往返医院次数。

2.4 智能引导式消息推送服务全流程覆盖

为了让患者在就诊过程中更加清晰地了解各诊疗环节的待办事项，给予患者关于就诊时间、地点、注意事项等信息的详细提醒，避免患者不清楚就诊流程而多次往返于各窗口造成时间浪费、造成体验感下降，医院针对影像检查全流程设计引导式消息推送服务。

全流程智能引导式消息推送提供检查过程中的自动预约提醒、待支付提醒、检查报告提醒、满意度调查的信息推送，使患者从预约挂号开始即可建立与本次检查的信息关联，实时向患者手机端推送下一个环节的相关信息，减少患者的无序流程，为患者提供更高质量、更高效率、更加体贴的医疗服务。

3 影像检查智慧服务流程优化的应用成效

经过影像检查全流程信息化支撑改造，医院以信息通促进服务通，让患者在影像检查就医全过程中无感办理各项手续，显著减少了患者往返次数、等待时间，提升患者就医体验。

第一，智能化检查自动预约平台实现了影像检查资源统一管理，可预约医技资源全院共享，患者数量和空闲资源实时动态展示，增加了临床医生、医技科室工作的协同性，解决了医院低谷期无患者,造成医院资源浪费问题,提高了医技资源利用率，使得患者周转加快，有效缩短了患者等待时间。各影像检查科室自动预约比例逐步提高，高于 70%，检查等待时间均有不同程序的缩短。

第二，流程优化后减少了患者排队次数和排队时长，方便患者就医，改变了患者往返于多个医技科室、缴费窗口、诊室的现状，至少减少患者排队次数 3 次，解决了患者在就诊高峰期排队时间长、候诊时间长的问题，合理分流患者，有助于维护良好的就诊秩序，有效节约了患者就医的时间成本。

第三，自助流程代替传统人工流程，逐步释放相关科室人工压力，提高整体工作效率。各医技科室预约工作人员转变工作职责，从单纯预约操作转为预约管理，进行可视化预约监管、协调各检查科室预约资源等。后期，通过整合预约窗口，设置全院统一的预约服务中心，将有效节约预约人力资源，推进人力资源优化配置。

4 总结与讨论

影像检查智慧服务流程优化体现了以患者为中心的服务思想，为患者实现“一站式”自助服务，在改善患者就医体验、节约医院人力成本、进一步提升医疗服务质量等方面取得了较好的应用效果。

下一步，医院仍将聚集患者看病就医的难点堵点，关注特殊群体就医需求，对老年人、残疾人等特殊群体，进行适老、助残服务改进，进一步提升就医全流程服务质量。

PU-267

信息化在医院后勤管理的应用分析

周麒 张红霞

河南省儿童医院郑州儿童医院

引言

后勤服务作为医院管理的重要组成部分，包括医院物资、总务、设备、财务、人力资源、基本建设等诸多方面，是临床、教学和科研等工作的基础保障。后勤服务具有不间断性、技术性、服务性等特点，其工作效率和管理水平直接影响医院的正常业务运转和健康发展[1-2]。随着医院的不断发展，粗放的管理模式不能满足实际需求，向智能化、精细化管理转型已成为必然趋势[3]。

大数据具有高频实时、深度定制化、跨组织数据整合、多主体协同等新特性[4-5]，现有管理方法、技术、工具以及人员水平难以应对当前的复杂局面和挑战。探索新的医院后勤管理理念和方法，以更好的发挥大数据动能和价值，对提升医院管理效果、降低运营成本、满足临床服务需求具有重要意义。

近十年来随着医院规模日益扩大,该院通过在医疗最前端的信息化工程,大大提升了医院效率和医院水平。但是作为医疗运营主要保障的医疗后勤保障业务一直处在医疗信息化建设的边缘,作为医疗经营的一个"短板",制约了医疗后勤保障业务的进展和水平的提高。并且由于医疗信息化的持续深入,各后勤单位科室所的信息化网络系统间出现信息孤岛，使整个后勤保障无法进行高效、统一的管理。如何运用先进的、科学的、智能化管理工具和管理手段来帮助医院后勤保障部门解决实际问题是目前的研究热点。

基于此，本文提出一种数据驱动医院后勤管理模型“1+1+3+N”，即一个后勤运营中心、一个智慧后勤管理平台、三个综合管理系统、N 种影响因素协同，包括一系列业务系统软件为载体的信息化协同等。首先通过定性、定量的分析方法，梳理医院后勤管理过程中的影响因素，剖析其特征和作用机制。继而构建数据驱动医院后勤管理模型，并制定相应保障措施以完备智慧后勤管理体系。本研究探索数据驱动时代医院后勤管理向智慧后勤转型的思路，能够为促进医院精细化管理提供对策及建议。

2 医院后勤管理信息化现状分析

目前，医院后勤管理逐渐展露出多机构、多院区、多部门协同的特性，由此产生的数据资源呈现多源、异构与海量增长等趋势。如何解决面向医院管理过程中的数据驱动、数据价值转换问题，是医院应对大数据环境开展智慧型管理服务、提升管理质量和效率必须解决的关键问题，也是学者和管理人员持续关注和研究的热点。

目前医院后前管理相关研究主要集中在以下两个方面：（1）数据驱动医院管理领域的理论模式研究。主要研究大数据对于医院管理领域的重要性或推动作用。（2）数据驱动在医院管理领域的应用研究，主要针对少量或具体医疗健康问题提出基于大数据技术的解决办法[6-7]。相关研究表明，大数据在医院后勤管理领域的知识发现、科技创新、管理决策支持等方面发挥了积极作用。合理应用大数据及新的处理技术和分析手段，可以为管理人员提供优质高效的数据分析和决策支持。

然而，与其他行业相比，医院后勤管理领域在利用先进数据分析工具和方法方面仍有所欠缺，数据价值在医院后勤管理的真正落地还存在诸多困境与挑战，主要表现在：

首先，现有研究缺乏对数据驱动医院后勤管理规律和过程的深入探讨。目前研究多关注面向与临床的医疗数据挖掘和应用，少有对数据驱动医院后勤管理的研究，更缺少从原理层面进行系统研究，没有形成可借鉴、可复制的系统性的模式和理论模型。

其次，现有研究多探讨数据驱动医院后勤管理方法和技术的应用，更关注具体细节，少有从宏观和中观角度进行全景式探讨，没有针对和后勤服务相关联的各层面、各要素进行分析。

再次，数据驱动医院后勤管理是一个多学科交叉问题，现有研究多是单一维度的，缺乏多学科深度交叉融合的研究，导致研究存在局限性。

综上所述，数据驱动医院后勤管理还没有形成成熟的理论模型，对复杂多变的应用场景、多层次多源头的数据资源、多元化的技术手段等影响因素的研究和挖掘深度稍显不足。亟需提出一种高效的智能化后勤管理系统，以提升医院后勤管理的能力和水平。

3 平台设计

为统筹管理医院后勤业务中涉及的原相对独立的安防、消防、病房呼叫等子系统，上层统一管理平台的设计至关重要。提出基于大数据、云计算等技术的“1+1+3+N” 一体化后勤信息化建设方案，满足各级医院后勤智慧管理的综合需求，打造安全、高效、节能的医院后勤。医院也可根据信息化建设发展需要和实际情况，按照“整体规划，急用先行，分步实施”的原则，进行模块化选择、定制，在满足医院当前所急需解决问题的前提下实施后勤信息化建设工程，逐步完善，不断更新，最终达到医院后勤信息化整体建设目标。

3.1 智慧后勤运维综合管理平台

智慧后勤运维综合管理平台是将本次建设的各业务支撑系统采集到的数据、各后勤部门所提供的服务、能耗管理采集的能耗数据集中到平台中，进行数据的分析及展示，同时对于前端科室汇集到的后勤需求，通过平台进行统一的调度指挥，建设后勤保障统一对外联系“窗口”、对内统一指挥“大脑”，最终实现一个中心解决所有后勤相关问题的目标。

3.2 后勤设备安全智能监测系统

后勤安全智能监测系统，结合物联网、大数据、人工智能等新一代信息技术，对医院后勤设备设施的运行数据与环境因素进行不间断监测与统计分析，实时发现设备设施存在的故障隐患，实现故障的可视化定位及精准预警，并根据故障级别及人员层级分别提供 APP 通知、短信、电话

等多方式的分级告警。基于积累的海量数据库资源，利用大数据技术对故障规律进行有效预测，提前进行精准高效的预测性维护，以精准化、可视化、智能化的管理方式，实现医院安全从源头到末端的全流程闭环式智慧化管理。同时将数据上传至后勤综合管理平台，为后勤整体运行管理提供设备安全数据支撑。

3.3 后勤综合服务管理系统

后勤综合服务管理系统是以一站式服务调度中心为核心，将后勤维修、保洁、运送、被服管理等业务纳入到调度中心范畴，临床有诉求只需拨打调度中心一个电话便可以解决问题，为临床工作做减法，为临床服务做加法。临床诉求以数据流的形式在信息系统内流转，利用手机实现跟单处理，满足工作需要，实现无纸化管理。系统建设以配合智慧后勤综合管理平台对后勤服务进行管理为主要目的，已达到提高后勤整体工作效率，提升后勤服务水平的目标。

3.4 医院能源管理系统

医院能源管理系统依托平台，能够实时、全面、准确地采集水、电、气等各种能耗数据，动态分析能耗状况，发现用能规律和特点，快速定位用能负荷高峰，可视化显示能耗 数据信息，科学管理降低能耗使用成本。对能耗数据进行分析处理，实现能源消耗可视化、 提示、预警和故障监测，对高耗能设备实时能效进行监控诊断出现异常及时告警，以此实现能耗的科学安全管理，可以实现 5%-15% 的节能降效收益。

4 应用分析与讨论

4.1 平台对业务的基础支撑

制定后勤业务统一平台、统一数据标准、统一系统架构，系统实现以用户中心、设备中心、空间中心、消息中心、协作中心和数据服务中心来支撑后勤各项业务应用，打通各子模块子系统之间的壁垒，实现数据信息的流通。

4.2 平台对管理决策的支撑与管理

根据后勤保障平台和各个业务子系统采集的数据和信息，按照医院的运维管理要求， 进行抽象、分析和汇总，以不同的维度提供给院方，对其决策进行支撑。具体包括：报告报表：全功能分析报告子系统，结合云端数据，生成医院领导层需要的各种报告和分析报表。

5 总结

数据驱动时代，传统医院后勤管理模式已经无法满足临床一线、患者需求，本文梳理分析我国现有医院后勤服务现状、存在的问题和发展趋势,并与国内外现代化医院进行比较，研究在大数据、云计算等信息化技术飞速发展的背景下如何构建医院后勤服务信息化系统。本文提出一种数据驱动医院后勤管理模型“1+1+3+N”，为完善医院后勤服务体系提出对策和建议。具有重要的理论价值和实践价值。

PU-268

智慧医院建设

李若斌
开封市人民医院病案统计科

1.智慧医院的起源和发展

2019 年 3 月，国家卫健委首次提出智慧医院的建设问题，主要包含三个层面：智慧医疗、智慧服务、智慧管理。智慧医院的产生与医疗信息化的发展密不可分。2000 年以前，我国医疗信息化建设已经开始，但多为财务核算系统以收费为核心功能。2000-2010 年，大数据技术的逐渐完善，医疗大数据也随之发展，医疗信息化迎来爆发期。2010 年起，国家开始对电子病历实施的探索标志着医院由信息化建设走向智慧医院的建设。2010 年开始，创新技术层出不穷，大部分医院尝试将云计算、5G、物联网、AR、VR 等技术与医院内的业务场景融合，以达到提升医院运营效率、为患者提供便利就医服务、为医生减轻工作负担的目的。与技术发展配套而来的是相关的政策与标准规定，电子病历规范和智慧服务标准的发布都为智慧医院的建设提供了方向。由此可见，推动医院从信息化医院走向智慧医院的两大因素分别为创新技术和宏观政策[1]。

目前，我国已基本形成智慧医疗、智慧服务和智慧管理“三位一体”的智慧医院建设体系。国内医院则在智慧化建设过程中以此为基础，逐渐向面向医务人员的智慧医疗、面向患者的智慧服务及面向医院管理的智慧管理的体系靠拢，并进行了不少探索尝试。下面分开简单介绍智慧医院三个方面的内容。

1.1 智慧医疗

我国智慧医疗建设起航于 2011 年，原卫生部印发的早期的《电子病历系统功能规范（试行）》中明确提出了电子病历系统建设要求以及涵盖的范围。智慧医院建设以电子病历建设为核心，重视临床诊疗信息化以及医疗数据的互联互通。医院信息化也不再局限于财务收费等基础业务，开始逐渐向临床诊疗业务转移，形成了以门（急）诊、病房等临床信息系统以及检查检验、病理、影像、心电、超声等医技科室信息系统为基础的电子病历系统。不同医疗场景的业务流程通过信息技术进行标准化，原来的手工等工作方式被信息化系统及流程所取代，一方面使医护人员的工作效率得到了很大提高，另一方面，数据存储形式发生改变，有效地促进了医院数据资产的转化与积累。在医疗业务信息化基本完成的前提下，随着系统持续运行，医院数据资产不断地沉淀，数据体量达到了须要从“量变”到“质变”的阶段后，数据共享、数据治理、数据利用成为智慧医院建设的新目标。该阶段智慧医疗的核心在于集成平台、临床数据中心、BI 数据平台建设。基于统一的数据管理，医疗业务系统的数据质量得到了很大提升，为数据的临床应用打下了坚实的基础。在业务标准化和数据标准化的双线驱动下，智慧医疗从传统应用演变为智能化应用，如临床路径管理、DRG 医保控费、临床决策支持、科研支持等，为医护人员提供各类智能提示或提醒，注重医疗工作质量管理与控制。信息化的作用从早期的替代逐渐转变为对医疗工作的主动干预、自动优化、闭环管理。

1.2 智慧服务

在智慧医院建设中，面向患者的智慧服务的重点在于患者体验，本质上是实现就医流程和部分医疗服务的在线化。近年来伴随着移动互联网的发展，智能手机在居民工作、生活中扮演了越来越

重要的角色，居民利用手机进行全流程就医也成了医院信息化建设的重点。该阶段中，智慧医院建设围绕“以患者为中心”，充分利用移动设备不受时空限制的特点，开展全面线上就医服务建设，实现了患者直接利用手机就能挂号、缴费、预约等目标，有效改善了长期以来由于医疗资源相对短缺而困扰居民就诊的“三长一短”等问题。医疗服务在线化的典型代表是远程医疗和互联网医疗的发展，医疗服务不再遥不可及，如慢性病复诊、续方、随诊等可在线上完成，无须经历在医院的挂号、候诊、开方的烦琐流程。一方面线上医疗服务的就医时间更加灵活、就医流程更加简化，提高了患者的就医体验；另一方面互联网医疗服务将医疗资源从院内延伸到院外、从城市延伸到基层，有力改善了由医疗资源配置不均衡带来的“专家难求”等问题。随着《智慧服务分级评估标准体系》的推出，智慧服务的建设重点转移到内外互联、实现院内院外电子病历共享等方面，以实现医院、基层医疗机构、居家等场景的医疗健康数据和服务全打通为目标，进一步实现真正覆盖患者诊前、诊中、诊后的全生命周期管理。

1.3 智慧管理

早期的医院管理通常采用经验管理的方式，以医疗事务为主，以学术为导向。随着医院的不断发展，收支规模逐渐扩大，经济运行压力也持续增加。国家提出高质量发展目标后，医院亟须向现代化管理转型，通过企业化经营、管理创新、控制成本、提高效率等方式实现有限资源的最佳利用，实现社会效益与经济效益的最大化。

在医院基础信息化建设实现阶段性目标之后，利用新一代信息技术服务医院管理、支撑决策，达到智慧管理的目标成为可能。借助物联网、大数据、人工智能等新技术，在实现互联互通的基础上，构建后勤管理、医疗供应链管理、手术室管理、医疗设备管理、安防管理、指挥调度等精细化管理应用，升级医院管理工具，是面向医院内控的新一代管理模式。国家发布智慧管理分级评价标准，进一步推动了我国医院现代管理模式的转变。基于智慧管理的管理理念，医院管理由传统的经验型转变为数据驱动型，运营管理由粗放型转变为全程可追溯的精细化，后勤管理由体力型转变为信息驱动的专业化，多院区管理由离散型转变为资源共享及同质化，医院整体管理水平将步入新台阶。

2.智慧医院建设概括

智慧医院的建设可以分为两条主线，分别是医疗信息化和医疗建筑智能化。医院信息化系统和智能化系统都是为了提高医院的医疗服务水平和服务体验而设计的，但它们采用的技术手段和侧重点不同。主要区别在于它们的目的和所采用的技术。概况来说，信息化系统主要关注医院内部流程和业务的优化，而智能化系统则强调为患者提供更自动化和智能化的服务。

2.1 医院信息化系统主要利用电子病历、医疗信息管理系统和医疗数据分析等工具，旨在管理和优化医院的各项流程和业务。它的目的是提高医院运营效率，优化诊前、诊中、诊后医疗服务环节的体验。

2.2 医院智能化则采用传感器技术、数据处理技术、人工智能等技术手段，强调为患者提供更自动化和智能化的服务。智能医院能够实现患者、医务人员、医疗机构、医疗设备之间的互联互通，以提高医院运营水平和服务质量。其中，智能预约、智能导诊、智能检查和智能诊断等应用可以减少患者等待时间，提高就诊效率。

2.3 医院智能化系统的设计要遵从顶层设计的方法论，在整个顶层框架设计时遵从自上而下的原则，在实现步骤设计时采用自下而上的方式，按照顶层设计的原则、方法和关键点，结合传统的医院智能化系统的设计方法，形成一套完善的新型国际化智慧医院设计思路。遵循“结构合理、系统稳定、适当冗余、适度超前”的设计理念，设计体现渐进性（留有扩展余地、可分步实施）。

“智慧医院”是一个由技术和结果推导出的概念，它不是单一的技术堆积，也不是某一个功能的代名词。智慧医院是具备信息化、互联网化、智能化特征的医院。信息化是指，医院建立了不同维度的数据系统以及各维度数据的集成系统；互联网化是指，医院推出了移动应用，为医务人员及患者提供诊前、诊中、诊后环节数据的输入与输出；智能化是指，医院运用了大数据、云计算、物联网技术、自动化设备、机器人、智能工作流与运营管理系统。智慧医院能够实现患者、医务人员、医疗机构、医疗设备之间互联互通，提高医院运营水平，优化诊前、诊中、诊后医疗服务环节的体验。

医院智能化子系统之所以众多，是为了满足医院多样化的需求，提高医院运营效率，为患者提供更优质的医疗服务。具体来说，医院智能化子系统具有以下几方面优势：

2.3.1 提高医护人员的工作效率。智能化子系统可以实现医疗设备、患者、医护人员之间的互联互通，提高医护人员的工作效率，减少医疗差错。

2.3.2 提供更好的就医体验。智能化子系统可以为患者提供便捷的就医流程和个性化的服务，如自助挂号、缴费、查询检查结果等。

2.3.3 实现医疗资源的合理分配。智能化子系统可以实现医疗资源的实时监测和管理，根据患者的病情和需求合理分配医疗资源，提高医疗资源的利用效率。

2.3.4 提高医院的运营效率。智能化子系统可以实现医院各类数据的整合和分析，为医院管理决策提供数据支持，提高医院的运营效率。

2.3.5 保障医院的信息安全。智能化子系统应具备高水平的信息安全防护能力，保障医院的信息安全。

3.智慧医院未来发展趋势

随着我国医疗市场的逐渐放开，新进入医疗行业的跨行业参与者带来了新的发展思路。智慧医院在未来将持续在以电子病历为核心的智慧医疗、以标准化及高效为目标的智慧管理和以患者为中心的智慧服务三大方面继续发展，并呈现出如下趋势[2]。

3.1 线上线下一体化

2020 年 12 月，国家卫生健康委联合国家医疗保障局、国家中医药管理局发布《关于深入推进“互联网 + 医疗健康”“五个一”服务行动的通知》，要求加快推进线上线下一体化的医疗服务模式，推动便民惠民医疗服务向纵深发展。如何运用互联网数字技术优化智慧医疗服务流程，让病情轻微的患者少跑腿，让情况严重的患者第一时间得到救治，成为智慧医院发展的重点方向。

3.1.1 基于电子病历的互联互通

电子病历互联互通是智慧医院建设的基础所在，其普及将打通医院内部信息系统，使患者信息在院内乃至不同院区间实时流动并共享，使远程医疗平台、影像辅助诊断云平台和区域信息集成平台、区域医疗中心、区域影像中心、区域检测中心等重点建设领域得以落地。在未来，电子病历有望进一步演化为个人电子健康档案，进一步延伸应用范围。

3.1.2 数据驱动医疗

智慧医院的建设将涵盖丰富的数据，从而为大数据、人工智能及健康设备等新兴技术在医疗领域的创新应用提供有力支撑，助力提升医院诊疗、管理及服务水平。目前超九成的受访专家认同智慧医院大数据驱动将使医疗从普适化模式逐渐向个性化模式转变，进而形成包括预防、诊断、治疗和康复的全产业链条，精准满足居民的就医和健康管理的需求，并对医院的运营、管理和技术体系的设置产生深远影响。

3.2 就医流程再造

看病流程复杂、排队时间长、沟通时间短等仍是目前患者就医的主要痛点，传统实体医院就医流程分散、碎片化是导致这些痛点的主要原因。越来越多的医院开始践行“以患者为中心”的服务理念，在保证医疗质量及安全的前提下对流程和服务进行重塑，比如引入“一体化诊区”和“中心制医院”等先进理念，并结合智慧服务大幅改善患者体验。另外，未来医院的医疗服务须要进一步由院内向院外扩展延伸。增加挂号咨询、诊后答疑和用药指导、慢性病居家监测、一键就医或急救等服务是患者对智慧医院建设的希冀，智慧医院的建设将打破医疗边界，将医疗服务下沉到社区，再进一步延伸至家庭和个人，通过智能技术将医院与个人、院内与院外的健康数据结合，从而提供全生命周期的健康服务。

3.3 新技术赋能

通过新一代互联网技术、人工智能、大数据、物联网和可穿戴设备等新技术的应用，能为患者带来无障碍体验，为医生提供诊断和治疗辅助，为管理者提供决策和管理支持，真正打造以信息化为基础、数字化为手段、智能化为目标的智慧医院，全面提升医院的服务能力，如自动记录并分析患者过敏史、自动提醒用药剂量、将检验检查危急值及时告知临床医生、云端智能诊断、看病全流程智能化、智能语音识别、突发公共卫生事件的感知和预警、智能发现临床诊疗规范与标准的改进点等。

综上所述，智慧医院为患者带来了新的希望，提升了患者体验，改善了就医感受。智慧医院的建设既能打通目前患者就医的难点、堵点，又能不断帮助临床提高医疗技术与质量，同时智能辅助决策系统帮助医院管理层统计分析医院各种数据指标。智慧引领科技，创新改变未来！

PU-269

基于大数据的科室耗材可视化闭环管理实践

[1]周继光 [2]李咏梅 [2]李玲 [2]吴晓菁 [2]陈晓燕 [3]方秀云 [3]张丽芬
[1]福建医科大学附属漳州市医院信息科 [2]财务科 [3]眼科

随着医疗技术的发展，在治疗中需要的耗材品种规格多样，同时随着科室业务发展，业务量也在逐渐增加，耗材使用量、领用量也在逐年增加，因此如何有效、高效做好耗材管理成为当前临床科室管理的焦点。目前，医院耗材相关的信息系统主要有 HIS 系统，用于收集开单与收费管理；HRP 财务管理系统，用于耗材的请领管理，两个系统分属不同的公司，在耗材编码，科室编码与名称，使用规格单位等方面都存在不一致，信息孤岛现象严重，医院缺少深入一线的可视化运营信息化平台，耗材管理人员无法通过系统统计分析耗材请领及使用情况。同时科室领用耗材随意性较大，存在科室超量领取、库存积压，过时浪费等现象，在各种盘点过程中出现盘点数据不平衡的问题。为改变这一现象，采用基于大数据的科室耗材闭环管理对提高工作效率，减少差错，实现耗材有效管理有着重要意义。

一、 科室耗材闭环管理的定义[错误!未找到引用源。,错误!未找到引用源。,错误!未找到引用源。,错误!未找到引用源。]错误!未找到引用源。：

科室耗材闭环管理是以采购处仓库为耗材一级库，各临床科室耗材仓库为二级库，耗材从二级库库开始，通过实时抓取系统间数据，建立【原有数量+领用数量=现有数量+使用数量】数据模型，以可视化手段实时展示盘点、领用、使用和库存间数量平衡关系，以闭环为管理单元，实现耗材的完全管理。

二、构建科室耗材闭环管理的可视化模型[错误!未找到引用源。]。

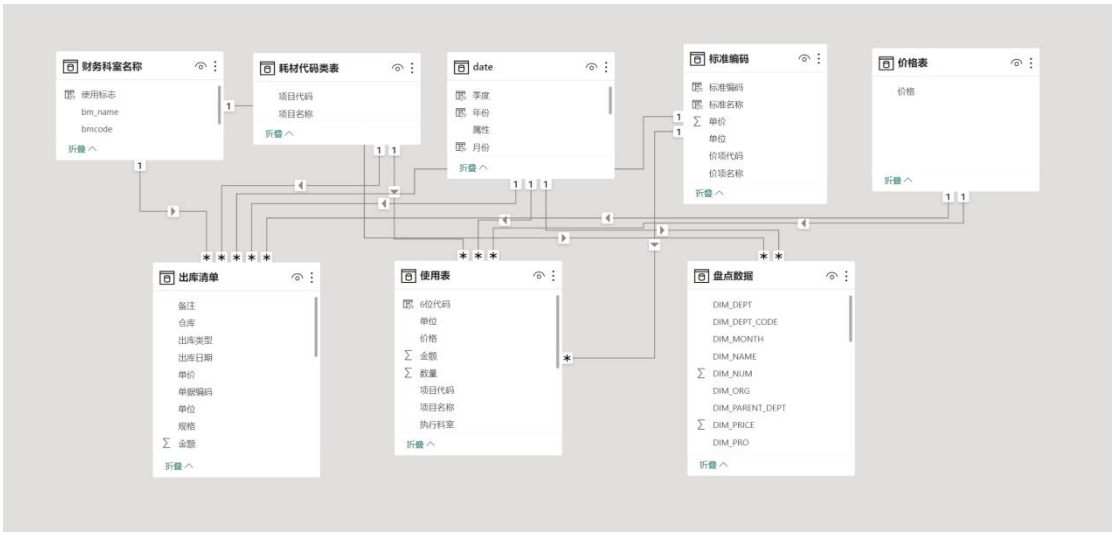
2.1 可视化模型利用的开发工具是微软 power BI，该工具源于微软公司，具有图形化操作界面和灵活多变的 DAX 公式语言相结合的特点；支持各种形式的数据对接，嵌入了 R 与 python 等当下流行的可视化建模工具，数据建模能力强，现已广泛应用在各行业的可视化管理工作中。

2.2 实现闭环管理需要获得的数据源：1.二级库出库数据，HRP 财务管理系统中以采购处为耗材一级库，请领耗材的科室为二级库，此次模型建立以二级库出库数据为基础，从物资系统中提取；2.使用库数据，该数据需要从 HIS 收费系统中提取，以上数据均以执行时间为准。3.出库耗材编码与收费编码间对应关系表。按照医保要求，耗材必须与收费项目内涵一致，否则属于除外内容，需要单独申请一个编码，因此会出现了同一个耗材在整个系统中存在多个编码的现象，要实现出库、使用、收费的工作流中数据间的正确归集，此时必须建立起出库耗材编码与收费编码间一对多对应关系表；4.不定期盘点数据表，建立不定期盘点制度，收集不定期盘点数据，是实现闭环的关键。任何盘点数据既是下一个闭环的开始，也是上一个闭环的结束，通过多次盘点即可形成多个数据闭环，通过核对以闭环为基础的理论库存数与实际盘点数差异情况既可以分析出盘点工作的正确性，也可以实现耗材全过程有效管理目的。

2.3 建立模型的数学原理：任何一个闭环的数学依据均为：原有数量+领用数量=现有数量+使用数量。通过查阅各种文献，仓库管理工作中所有工作都是围绕着这四个因素的平衡关系而开展，最终评判管理是否到位，耗材是否存在损耗等都需要通过这四个因素来评判。因此此次模型建设的原理依据就是该关系式。

2.4 表间关系建立。数据源导入后，利用自带的 PowerPivot 进行关系建模，数据建模过程类似于 SQL 语句中的关联[错误!未找到引用源。]，根据逻辑关系依次将各个表的字段进行关联，本次建模采用了星型结构作为建模的基本结构，将事实表与维度表之间建立一对多的关系，维度表位于“一”端，所有业务表位“多”端，“一”端表中的主键必须唯一。本次模型的建设以财务科室名称，耗材代码类表，日期表为维度表，以出库清单表，使用清单表和盘点数据为事实表，通过建模，使用多个表变成了一个整体，便于跨表数据分析。见图 1。

图 1 构建耗材闭环管理的数据基本模型



2.5 利用 DAX 语言编写度量值。数据分析三要素[错误!未找到引用源。]：指标、维度、分析方法。指标有：使用量，领用量，出入库数量差，使用金额，加权价格等，维度指分析的角度，如按照科室名称、时间维度的年度、季度、月份、星期等，分析方法时间序列，对比等方法。所需要的度量值主要包括第一次盘点时间，第一次盘点时实际库存，第一次盘点后领用量，第一次盘点后使用量，第二次盘点后的理论库存，第二次盘点时的实际库存，此六个度量值构成了第一个闭环，第二次盘点时间，第二次盘点后的使用量，第二次盘点后的领用量，第三次盘点时的理论库存和第三次盘点时的时间库存，结合前一闭环的实际盘点数据构成了第二个闭环。

2.6 利用 PowerBI 可视化视觉工具制作可视化模型 汇集以上维度表，度量值，利用 Power BI 自带的可视化视觉工具即可制作成一个界面内容丰富、查询内容多样、维度多、响应速度快的 BI 模型，见图 2，3，4。

图 2：通过 BI 工具建立的可视化闭环管理模型

图 3：物资领用可视化图

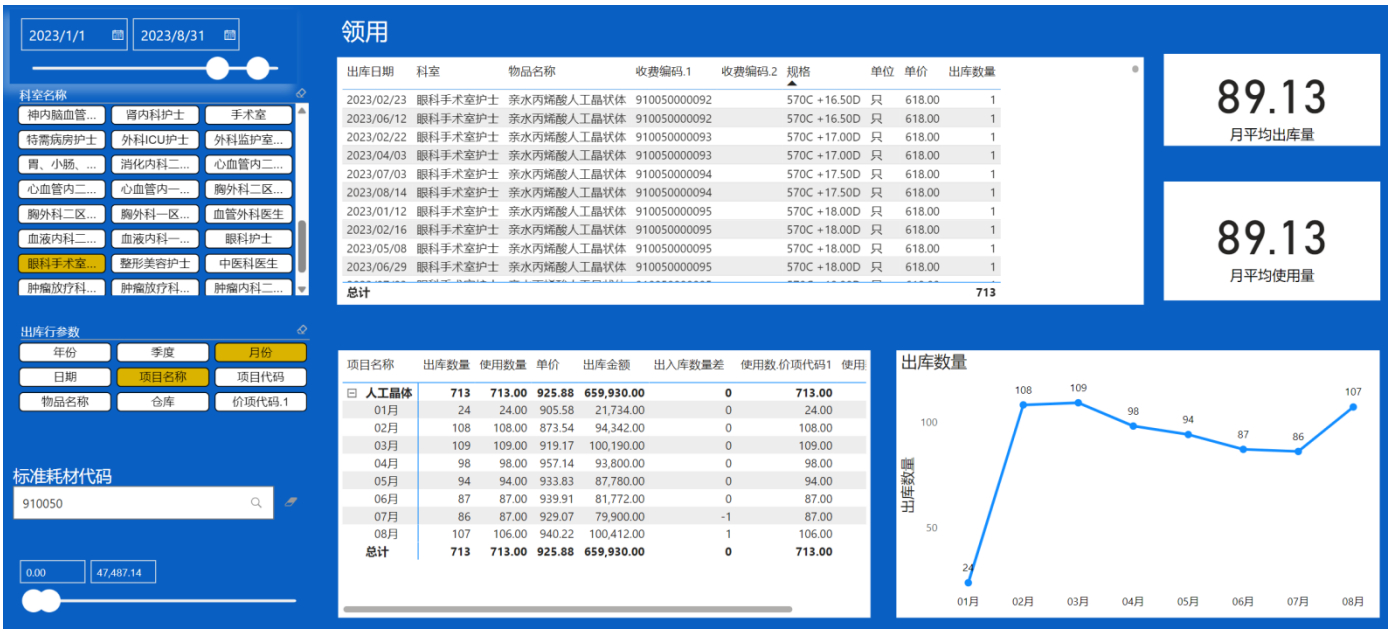
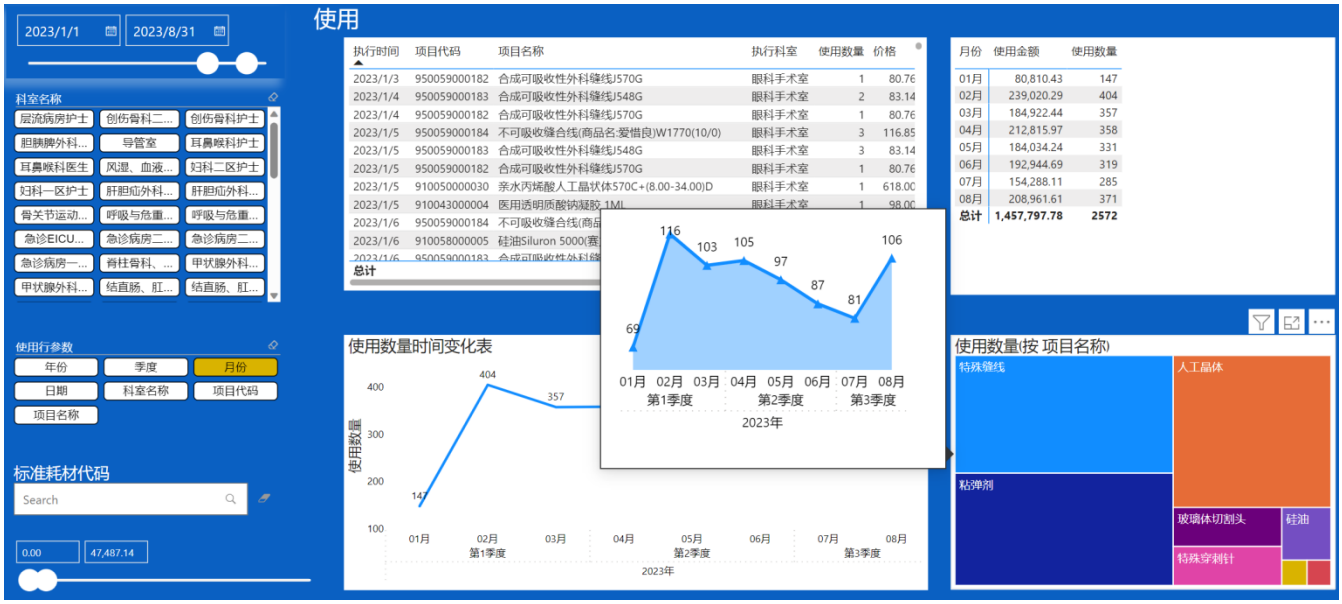


图 4：耗材使用可视化图



三、科室耗材可视化闭环管理的实践效果。

1、提高了工作效率。

每种耗材经过每两次盘点后即可根据原有盘点数、领用数、使用数和现有盘点数形成完整的数据闭环，即第一次盘点数（起始数据）+领用数=使用数+理论库存数，再根据理论库存数和第二次盘点数是否相等即可判断这段时间内耗材管理是否到位，也可通过该表直观批量判断出哪些耗材存在管理问题，极大地方便了工作。经过多次盘点后，即可形成多个数据闭环，通过观察多个闭环内数据的平衡关系可以判断出盘点数据质量，耗材管理质量，一改以往耗材管理中耗时耗力，数据来源多，核查繁琐的弊端，同时通过 Power BI 的刷新功能，及时更新数据，与系统同步，解决了医

院缺少深入临床一线的可视化运营信息化平台的弊端[错误!未找到引用源。]。如图 5 眼科耗材可视化闭环图，玻璃体切割头 3 月 31 日份盘点时 17 个，4 月份领用了 16 个，4 月份使用了 20 个，按照前面数量关系理论库存必须是 13 个，实际盘点是 13 个，数据关系平衡。但是 5 月份盘点后出现了理论盘点与实际盘点数的差异，相差 1 个，通过这个不平衡关系，顺藤摸瓜就可以发现并纠正存在的问题，同样情况出现在特殊穿刺针，可吸收缝线和透明质酸钠等耗材的管理中。往往这种关系的失衡，大多由于收费漏收多收而造成的，通过该关系式上半年纠正了 4 例收费漏收情况，挽回了科室损失。

图 5：眼科耗材可视化闭环图

标准名称	第一次盘点时间	第一次盘点时实际库存	第一次盘点后领用量	第一次盘点后使用量	第二次盘点时间	第二次盘点时理论库存	第二次盘点时实际库存	第二次盘点后领用量	第二次盘点后使用量	第三次盘点时间	第三次盘点时理论库存	第三次盘点时实际库存	第一次盘点后至目前的使用量	第一次盘点后至目前的领用量	第一次盘点后至目前的理论库存
⊟ 玻璃体切割头															
一次性23G双刃高速玻切头&G-42302	2023/3/31	17.00	16.00	20.00	2023/4/30	13.00	13.00	20.00	14.00	2023/5/31	19.00	18.00	82.00	97.00	32.00
⊟ 特殊穿刺针															
一次性使用眼科穿刺系统&G-46311	2023/3/31	17.00	16.00	20.00	2023/4/30	13.00	13.00	20.00	14.00	2023/5/31	19.00	18.00	82.00	82.00	17.00
⊟ 特殊缝线															
不可吸收缝合线&W1770(10/0) 12根/盒	2023/3/31	60.00	12.00	10.00	2023/4/30	62.00	63.00	12.00	13.00	2023/5/31	62.00	61.00	62.00	60.00	58.00
合成可吸收性外科缝线&J548G	2023/3/31	120.00	96.00	49.00	2023/4/30	167.00	170.00	48.00	48.00	2023/5/31	170.00	167.00	215.00	276.00	181.00
合成可吸收性外科缝线&J570G	2023/3/31	141.00	96.00	50.00	2023/4/30	187.00	189.00	48.00	30.00	2023/5/31	207.00	206.00	201.00	252.00	192.00
可吸收性外科缝线&6/0	2023/3/31	1.00	12.00		2023/4/30	13.00	12.00			2023/5/31	12.00	12.00	3.00	12.00	10.00
⊟ 粘弹剂															
医用透明质酸钠凝胶&1ML	2023/3/31	1.00	180.00	100.00	2023/4/30	81.00	83.00	120.00	91.00	2023/5/31	112.00	106.00	489.00	590.00	102.00

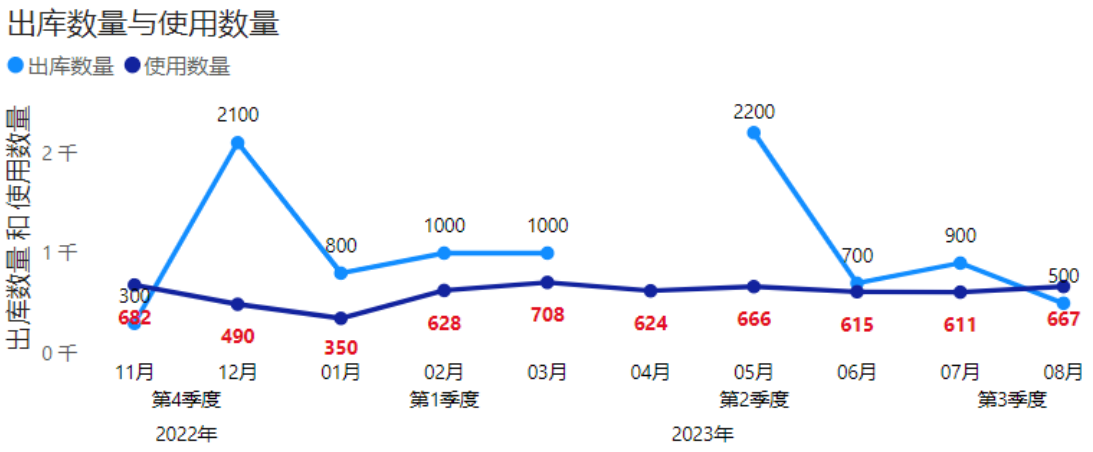
第一个闭环

第二个闭环

2. 有利于纠正“以领代支“的管理弊端，便于日常合理安排采购和物资领用。

由于采用大数据 BI 可视化方法，能清晰核查出申领和使用的过程，可以按时间维度，科室维度，出库类别等多维度观察分析使用量和领用量，使用量和领用量在同一个界面，异常使用和领用情况可以直观展示，由于采用了时间序列分析方法和对比分析方法，任何一种耗材的使用与领用一般都需要符合一般时间变化规律，通过图示方法可以快速查找某些产品申领过大，会导致丢失和过期浪费，如图 6 某科一次性使用静脉留置针的使用量是维持在 600-700 之间，其中 2022 年 12 月超使用量 3 倍达到 2100 只，2023 年 5 月一次性领用 2200 只，大量库存往往导致丢失与过期现象。通过简单分析每一种常用耗材的日常领用关系就可以做出正确的判断，通过汇总全院数据即可以安排全院每一种耗材的采购规律，避免一级库需求暴增与积压。

图 6，一次性静脉留置针出库情况图



3. 分散盘点，工作时间灵活。科室耗材管理原则要求是“账物相符”，定期对库存物品进行清点盘查盘点，是医用耗材库房管理中的一项不可缺少的工作。目前采用盘点方法一般也是固定时间大盘点，由于品种多，规格多，时间紧，因此在集中盘点过程中难免会出现盘点出错，导致数量关系不能平衡且无法纠正。采用该模型作为闭环管理工具，盘点机动灵活，可以利用任何有利时间挑选单一品种进行盘点，并不需要一定是月末或者月初某一具体时间点（如图 7），给予盘点科室极

大的工作灵活性，大大减少了出错机会（如图 5），保证了数据质量。

图 7，任意盘点时间下的可视化闭环图

标准名称	第一次盘点时间	第一次盘点时实际库存	第一次盘点后领用量	第一次盘点后使用量	第二次盘点时间	第二次盘点时理论库存	第二次盘点时实际库存	第二次盘点后领用量	第二次盘点后使用量	第三次盘点时间	第三次盘点时理论库存	第三次盘点时实际库存	第一次盘点后至目前的使用量	第一次盘点后至目前的领用量	第一次盘点后至目前的理论库存
⊟ 分流管															
脑脊液分流管及附件&41101	2023/8/28	4.00											6.00	6.00	4.00
脑脊液分流管及附件&42866	2023/8/28	4.00											5.00	5.00	4.00
脑脊液分流管及附件&43103	2023/8/28	4.00											5.00	5.00	4.00
⊟ 扩张器															
切口皮肤组织牵开扩张器(薄膜型)120/130-120/250、220/230-220/250	2023/8/21	0.00													0.00
切口皮肤组织牵开扩张器(硅胶型)&AF-G型（50/40-25/25)	2023/8/21	0.00											74.00	100.00	26.00
⊟ 脑脊液分流器															
脑脊液分流器及其组件&RE-1021	2023/8/28	3.00											1.00	1.00	3.00

四、小结

随着 2019 年国家卫健委 29 号文[错误!未找到引用源。]《卫生健康委、中医药局关于印发医疗机构医用耗材管理办法的通知》的颁布，明确要求医疗机构应逐步建立医用耗材信息化管理制度和系统。医疗机构耗材管理信息系统应当与医疗机构其他相关信信息系统整合，做到信息互联互通，实现每一种医用耗材的闭环可溯源。本研究及为本次研究而设计的模型，契合了卫健委关于耗材管理办法文件要求，建立了科室耗材监管模式可视化管理方法。将 HRP 物资请领耗材库与 HIS 系统耗材使用库进行了有效整合，同时利用高效灵活的盘点数据，形成了有效的数据闭环，实现了科室耗材高效、可溯源、直观的管理，简化了工作，提高了效率，同时利用大数据寻找规律，实现零库存，降低了医院运营成本，对医院管理有利有益，值得推广。

PU-270

浅析以党建引领智慧医院建设实践与思考

刘琦[1] 陈小蕾[1] 周伟[1] 肖勇[2]

海南省中医院

引言

随着信息技术的发展，“互联网+”深入各行各业，而新时代中国特色社会主义医院信息化建设也需要党建赋能，用党建引领发展。这是我党在大数据时代应对时代变迁的自我转型和自我革新。近年来，国内以党建设引领各行各业信息化建设取得了显著成效。海南省“三医联动项目”是省委部署的重点民生工程，通过建立健全以新技术、新标准、大数据、中台等信息技术为抓手的监督管理机制，助力医疗服务、公共卫生、疾病预防、计划生育、医保等，2020 年 12 月启动，海南省中医院为隶属于八家省属医疗机构建设对象之一。我院坚持将“党建+信息化”新思想新模式贯彻落实在三医联动项目中，用党建点亮信息化建设之路，更好地助力智慧医院发展。

党建引领信息化建设意义

提高思想站位，统一认识，压紧压实责任

2021 年 5 月 20 日下午，我院召开三医联动省中医院信息化建设启动部署会。医院领导班子、各党支部书记、各职能部门负责人出席，会议由医院副院长主持。党委书记在会议讲话：广大干部

职工要有高度的政治责任感，把思想和认识统一到中央、省委和省卫健委各项决策部署上来，三医联动省中医院信息化建设项目是我院信息化能力提升的一次重要机遇，全体员工要以加快医院高质量发展助推海南自贸港建设、建设健康海南为己任，切实增强能力提升建设的紧迫感和使命感，要项目实施过程中“提高思想站位、统一认识、全力配合”，为医院各项重点工作任务落地提供坚强保障。

坚持以“患者”为中心发展理念

医院深入贯彻学习全国卫生健康工作及全国中医药局长工作、全省卫生健康工作及省委工作报告中关于卫生健康和中医药的内容等会议精神。医院发展始终坚持人民至上、生命至上的价值理念，在信息化建设过程中牢牢把握党建中“以患者”为中心指导思想，就诊过程中突出便捷、实用、智慧，进一步实现“让患者少跑路，数据多跑路”的建设理念，努力双升“质”与“量”医疗服务，不断提升医院的核心竞争力。

全局性谋划、战略性布局、整体性推进

2022 年是“十四五”的开局之年，也是我院江东新区的启用试运行之年。我院坚定不移贯彻新发展理念，坚持稳中求进工作总基调，系统、准确总结了信息化建设发展历程，紧扣医院今后五年发展工作的总体要求，适用新时代医院长远发展，立足医院实际，对标《“健康中国 2030”规划纲要》、《电子病历系统应用水平分级评价标准》、《医院信息互联互通标准化成熟度测评方案》等标准，按照“全局性谋划、战略性布局、整体性推进”的原则，以更高的站位制定医院信息化建设“十四五”规划，以推动高质量发展为主题，以建设数字医院及切实为人民健康服务为总目标。

党建引领信息化建设举措

组建党委领导下的工作专班

为贯彻落实本次信息化建设项目，医院建设一支讲政治、敢担当、能作为的高素质专业化的队伍，来“凝力量、保落实、提能力、聚发展”。成立以院长为组长的领导小组负责负责总协调、做出管理决策，同时保持与省医保局、省卫健委等上级主管单位沟通协调；成立由三医联动项目监理组成的监理小组负责监督管理；成立由各科临床科室主任组成的院内执行小组，负责执行领导小组决策；成立信息部全体成员、各科室信息专员、各信息公司派驻技术团队组成技术保障组。强调相关工作小组要扛起主体责任，制定应急方案，严格落实各项工作措施，切实保障院内工作平稳有序运行。

强化扁平化管理体系，充分发挥基层网格作用

医院党委坚持以问题及需求为导向、列出清单，迅速组织需求调研小组，包括项目承建方、项目监理与院内信息科、临床科室，深入到医院重点业务部门，开展对现有系统的业务流程、系统应用沟通、运营管理等相关工作现场调研，充分发挥基层网络作用。历经 1 个多月对医院现有系统进行 3 轮全面摸底，对我信息化建设形成初步调研报告，主要包括系统运行情况、业务流程、存在的问题、拟采用的手段等。医院信息化建于 2009 年，最建设的有基础 HIS 系统、医技系统、护理系统等，2016 年进行升级优化及扩充（体检系统、病案首页等）。以 HIS、EMR 为主线的业务系统版本老旧，底层架构兼容性较差，系统之间个别内容数据不一致，缺乏患者主索引、部分系统难以实

现协同，临床业务与运营管理建设达不到医院的人财物精细化管理要。同时对于患者而言，需要多次出入医院，平均看病消耗时间比较长，智慧化服务程度不高，且存在自身研发能力较弱的通病。

盯紧突出问题、统筹力量提能力

2022 年是海南省开展能力提升建设活动年。从国家中医药管理局办公室印发的《关于全国三级公立中医医院绩效考核国家监测分析情况的通报》，近两年，我院三级公立医院绩效考核整体排名有所提升，但电子病历应用功能水平分级一直未达到“2020 年三级公立中医医院电子病历功能应用水平 4 级”的目标要求。以电子病历水平 4 级为主线，分解“查堵点、破难题、促发展”的清单，强化信息系统建设，加强医院服务能力与质量，促进医院信息化向数字化、智慧化转型。统筹力量集中以下几点提能力。

一是按照电子病历评级 4 级标准，完善各系统模块功能，统一数据标准，提高数据质量、完善业务闭环，为电子病历 5 级和互联互通四级甲等评级目标打下良好基础；二是建立以院内集成平台为核心的数据及接口交互体系，保障系统在标准下有序业务协同，加强追溯及闭环监管能力；三是新建临床数据中心，对接上级主管部门数据采集等系统。

坚持系统治理观念，制定技术实施路线

医院党委在实施部署会上强调：“完整、准确、全面贯彻新发展理念，必须坚持系统观念”，要从整体看待发展，实施方案设计过程中主要采用系统科学法等方法，加快进度，严控潜在风险点，将信息系统与临床业务协同发力、形成合力，整改突出问题，制定全面、完整的技术实施路线，不断规范及优化门诊挂号、收费、就诊、取药、随访及住院相关等全环节流程，争创人民满意的医院。

技术路线包括需求调研及需求收集、技术方法及建设内容对标、基础数据梳理、重点业务流程确定、软硬件环境、数据备份、信息安全、紧急预案等。技术路线如下所示。

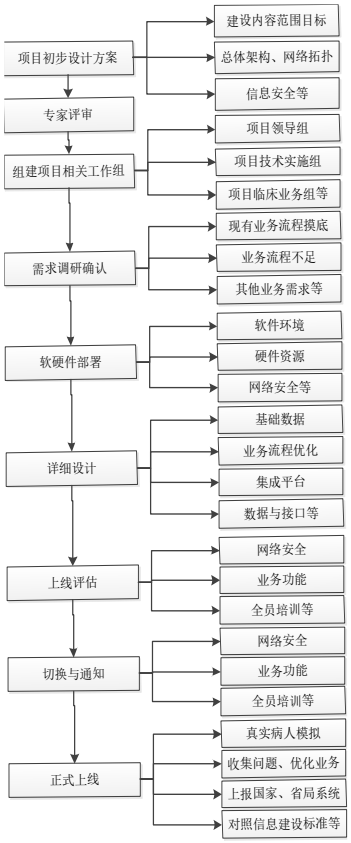


图 12 技术路线图

党建引领信息化建设成效

截至目前，我院在本次建设项目中建设了医院信息集成平台、临床数据中心、360 视图及 360 助手，新建 8 个业务系统，完善及优化 30 多个业务系统，其中涉及到 HIS、电子病历、医技系统、移动医疗、移动护理、临床辅助决策、运营管理、综合管理等系统。从全国三级公立中医医院绩效考核数据来看，近两年门诊患者满意度、住院患者满意度呈现逐年上升的趋势，离不开医院信息化对患者预约挂号、线上交费、医保报销、电子病历文书、医疗诊疗服务、患者全息视图等各项业务支撑，有效缓解“三长一短”现象，目前我院信息化建设已满足三级公立中医医院电子病历功能应用水平 4 级条件。

面向患者的智慧服务建设

我院认真贯彻落实党中央、国务院“以人民为中心”的重要决策部署，用智慧服务着力解决医疗就诊过程中“三长一短”和“看病难、看病贵”的问题，努力提升患者就医体验、满意度。本次建设中医院本着“信息共享、服务智能”的方向及为患者“优质、便捷、智慧、人文”服务的目标，优化与完善智慧服务信息系统。

诊前服务：基于医院微信公众号提供当日挂号、预约挂号等，目前诊疗预约中应用电子系统预约的人次数占总预约人次数比例约为 80%；同时为老年群体等提供“互联网+护理服务”上门服务；

诊中服务：目前微信公众号、支付宝等渠道应用信息技术开展信息患者看病期间产生的相关就诊费用、药品使用情况、医技报告、电子发票等推送服务；

诊后服务：近两年门诊患者满意度平均分为 84%、住院患者满意度平均分为 87%，药品调剂与配送我院中药智慧药房大约有 18%的具有药品调剂与配送服务系统并有配送，标识与导航通过处方单、检查检验申请单提供标识及就诊地址导航；提供门（急）诊、住院及健康体检患者智能导诊单，实现“一单一号”完成院内所有医嘱及检查检验项目，方便患者快速就诊室看病；

全程服务：目前老年病、心血管科室等已经开展了健康宣教，近 3 个月已有 507 人次数患者推送宣教内容；为充分发挥中医预防、养生、康复的优势基于海南省智慧医院 APP 开展全省人民的中医体质辨识服务。

远程医疗：目前远程区域影像已有 10 家中医医院开展 DR 与 CT 项目的远程诊断，近两个月人次数最多的定安 CT 为 800 人次数，DR 为 1704 人次数；屯昌 CT 为 496 人次数、文昌 106 人次数；远程会诊系统已有 1 家省级中医院及 17 家市县中医医院；海南省中医药数据中心：目前中医馆健康信息平台已初步建立省、市（县）、乡镇三级覆盖基层中医医疗服务保障网，已有 261 家乡镇卫生院、二级大部分中医医院接入系统，开展中医诊疗服务。中医馆 2.0 建设升级 70 家乡镇卫生院等系统。

面向医务人员的智慧医疗建设

医院贯彻落实习近平新时代中国特色社会主义思想——“以人民为中心”是新时代坚持和发展中国特色社会主义根本立场。医院主体对象主要为患者及医务人员。其中智慧医疗主要为医务人员服务，通过智慧信息系统辅助临床医生及医技工作者的业务工作，同时记录患者就诊过程中的临床全生命周期及医生临床诊疗行为。“精进、协同、便捷、智慧”

本次建设过程中最突出的智慧医疗信息服务为：一方面突破院内集成平台的“从无到有”新阶段，集成平台由信息行业标准、基础设施、集成平台、数据中心、临床辅助决策、安全防护、医疗服务、医疗管理、运营管理、医疗协同、便民服务、医院门户等几个部分构成。通过患者、科室、员工主索引接口及在平台上注册标准服务等内容，便于各业务系统中患者信息及医疗行为数据的高度共享及互联互通，同时便于对接上级管理部分的系统及平台服务；另一方面新建基于集成平台的 360 患者全息视图（以下简称“360 视图”）及 360 助手，两者“一内一外”，高效配合临床。为院内临床医生在诊疗过程中提供 360 视图做到“一屏观患者”，360 视图主要包括该患者在本院门（急）诊、住院、体检等历史就诊记录、患者基本信息、医技报告、手术情况等便捷辅助决策。360 助手的作用主要为院内临床医生提供经过省卫生健康委三医平台认证的其他医疗机构的既往诊疗的患者信息、检查检验、医保等内容，方便医生了解患者在其他医疗机构中的诊疗过程、体征各项指标，助力全省“检查检验互认一件事”实施推广及应用。

面向管理人员的智慧管理建设

截至 2020 年底，全国医院共 35394 个，二级及以上公立医院有 51.2%开展预约诊疗，91.6%开展临床路径管理，63.2%开展远程医疗服务，86.7%参与同级检查结果互认，93.7%开展优质护理服务等。为贯彻落实以习近平同志为核心的党中央推行全面深化改革工作，我院在 2022 年的智慧管理举措有以下。

临床管理：为让助力海南人拥有一份完整、可共享的电子健康档案。在本次建设中，党委提出：临床管理要与系统融合应用，注重患者就医全过程及医护人员行为管理。系统按照医护人员岗位职责、职称和业务，进行分组角色权限管理，细化到处方权、手术权限、抗菌药物及系统功能模块等；患者就医过程中的各类病历文书启用我院新修订各类病历文书质控标准，系统被赋予了质控标准的内涵，其中环节质控、院级质控、终末质控覆盖了病历文书书写全过程，以系统自动与手工质控相结合的方式，双管齐下。

综合运营管理：指人、财、物、协同办公等的管理。人力方面则通过完善人事管理系统对人员档案、岗位、薪酬等细化管理；财力方面：新增绩效核算管理管理系统，对各级各类科室成本、诊次成本、病种成本及科室收入收益等进行全过程核算，追溯成本来源及成本组成，优化各项资源配置；物力方面：新建固定资产管理系统，对各类资产进行分类管理；协同办公管理：已日常化实现无纸化公文流转、行政审批流程、院内信息发布与公告、会议信息等。

数据管理：数据是资产，管理通过临床数据中心，临床数据中心是医院实现数据治理、成为数据资产的最佳手段。通过信息技术汇聚临床数据，依据数据标准处理数据、数据挖掘与分析主题化、标签化管理数据。自建成临床数据中心，以来，对接《海南省全民健康信息平台数据对接工作方案》由从独立存储各业务系统变为由临床数据中心统一采集、聚合、处理再上报。截至 8 月底，我院上传平台数据表单由 26 变为 29 张，三级医院排名从 16 名变为第 7 名，数据质量综合得分增加了 27.84 分、增长率为 51.45%，数据质量治理数据一致性、规范性、完整性、及时性增长率分别为 120%、20.43%、28.86%、52.97%。由临床数据中心集中处理、统一对接省平台的效果是显著的。目前，临床数据中心共采集 65091 份门诊病历，174591 份住院病历，处方数量 185752 张，挂号人次数为 28743 人次等。后期将历史数据持续采集到临床数据中心，不断探索及强化其服务能力。

信息安全管理：没有信息安全就没有医院安全。在本次建设中，我院在信息安全上为两方面。一是管理：不断完善网络安全管理制度体系，建立定期安全检查机制，系统日常运行、系统漏洞、数据备份、安全技术等措施，及时发现安全问题及隐患；二是技术：以《网络安全等级保护基本要求》为基准，开展备案为“三级”的重点业务系统安全等级测评工作；开展网络安全及策略管理、数据安全管理、数据库系统、用户访问等安全策略检测、安全加固及风险分析等巡检，并制定的应对措施。

结语

新时代中国特色社会主义医院信息化发展实施“党建+”新模式。新时代公立医院信息化建设要接轨国家、卫生健康系统、医院等“十四五”规划，“雁行千里靠头雁”，充分发挥党建引领信息建设高质量高标准医院。从加强顶层设计出发，坚持以人民服务为本、守正创思考新等原则，将党建工作的思想内涵融合到业务系统中，强化党建赋能新时代信息化建设工作。

PU-271

业务驱动医疗服务项目收费实行研究

徐挺玉[①②],黄忠秋 [①②],殷悦楚楚[①②]，郭建军[①②]*

江苏省人民医院

1 目的

如何有效地控制医疗费用不合理增长，提高医保基金使用效率，成为社会医疗服务可持续发展亟需解决的问题。开展医疗服务项目收费与合理控费研究，从多因素角度分析医疗业务与收费之间的关联关系，提出业务驱动的合理控制医疗费用的关键内容和环节，实现先执行后收费，形成业务收费统一管理平台，实现关键策略，成为当前和今后一个时期重要而迫切的课题。

业务驱动医疗服务项目收费研究目的

1）通过梳理规则形成业务与收费项目绑定的知识库，帮助临床避免触犯医保规则限制导致扣分，减少既往需要反复核查是否超出限制的时间，增加临床满意度，让临床工作人员能够将更多的时间用于救治病人。

2）形成智能化反馈机制，通过对历史数据的分析、比对，评估制定指标的合理性，从而针对性的改进本系统，进一步优化业务驱动医疗收费体系的构建。

3）业务驱动医疗收费体系构建，提高医院的管理效率，扩大了医院的影响力。

4）为院内绩效改革提供支撑，通过信息化手段针对工作量合理分配提出有实际依据的建议。

2 方法

有针对性地对本院历史医疗业务、医疗收费项目进行调研，通过了解实际业务场景，分析数据，提出需亟待解决的实际问题。密切关注国内外研究的相关进展，通过资料检索、大数据分析、专家咨询和理论演绎与归纳，形成医疗业务与医疗服务项目收费关联方案。基于协同理论与价值链理论方法分析本项目中的管理决策与服务优化领域的科学问题。在技术上，采用逐步推进的路线，通过建模、优化、仿真、系统建设的方式，运用定性和定量相结合的研究方法，建立和优化业务费用统一信息平台。

1）组建项目研发小组、制定与业务流程关联的收费标准

与计财处联合组建业务与收费体系研发小组，就诊疗项目信息、疾病手术信息、药品信息、对照目录、实际开展的具体医疗业务等，并针对基于医疗业务流程的收费管理，进行标准化设计，统一规范管理。

2）构建医疗业务与收费之间的知识库

（1）根据不同的医疗业务场景，梳理基于手术科室、内镜科室、放疗科室、导管室等执行科室的专科业务收费知识库，构建标准化业务编码、收费编码映射关系，形成智能清单。

（2）各业务执行科室按要求梳理的知识库，须提交至医院医保、物价部门进行合规性审核，审核通过后统一纳入院级知识库系统进行管理。

（3）联合财务及物价部门对最新物价内涵设立出更符合规范的医嘱收费项目，在这个过程中不断补充更新知识库及规则库，完善对知识库的智能管理。

3）优化业务执行科室操作流程并实现智能提醒

（1）对院内HIS等医疗管理系统计费模块进行改造，实现先执行后收费，系统同时记录每一个操作步骤的操作人、操作时间。

（2）在医护人员通过信息系统进行收费时，系统自动调用对应的业务知识库，对照智能收费清单，对医疗业务及收费行为进行合规校验，并对用户智能提醒。通过业务流程的优化，最终实现医疗业务的合理收费，降低医保审核带来的扣费风险，同时提高了业务数据质量，支撑绩效改革对于工作量的合理分配。

本项目总体研究理论框架基础上，以层层递进的方式展开。流程图如图1所示：

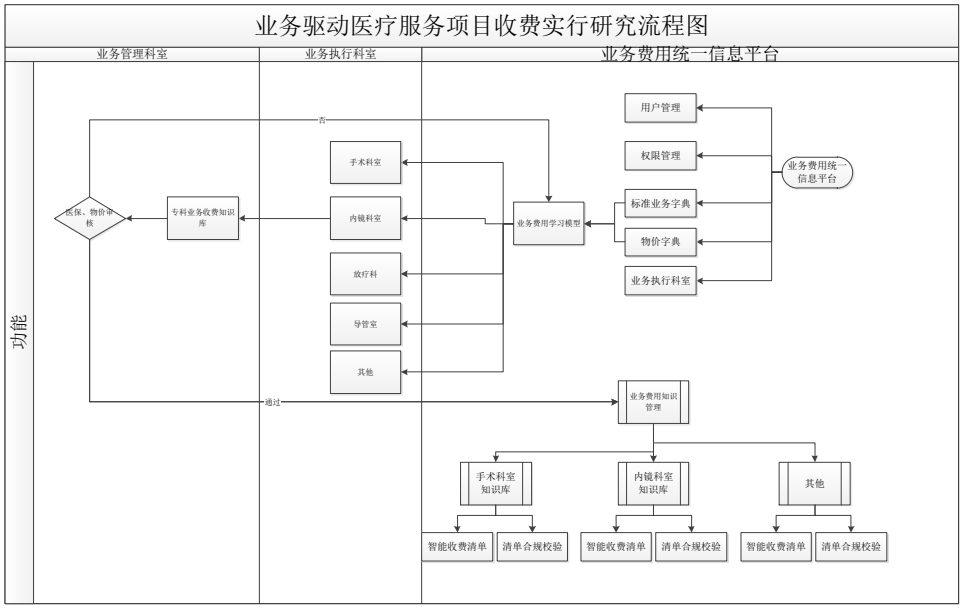


图 1 业务驱动医疗服务项目收费实行研究流程图

3 结果：

业务费用统一信息平台的构建，如图2：

（1）基础字典管理

①标准业务字典，根据不同的业务操作类型，形成标准化的操作业务字典库，包括手术操作、内镜操作、放疗操作、导管室操作、其他检查或治疗类操作等业务字典。

②物价字典，与院内物价信息管理系统中的物价项目实时对接

③业务执行科室字典，根据所开展的医疗业务的最小执行单元，设置业务执行科室字典，并与院内HIS系统机构字典、财务核算单元字典对照。

（2）业务费用知识库管理

开发知识库管理系统，可对手术、内镜等不同操作类型的医疗业务进行收费项目关联，同时针对不同的业务设置对应的合规性校验，系统界面支持对知识库和规则的灵活配置及修改，主要功能包括模板定义、元素定义、事件定义、表单多集关联等。

（3）权限管理

系统对财务、医保、物价等管理部门就临床科室人员进行用户授权，可以访问对应的模块。临床科室可指定对应的知识库维护人员在系统中维护对应规则内容。

（4）基于信息平台的业务数据统计分析

开发数据分析系统，可从业务执行科室、时间等不同维度对某个具体的医疗活动，手术、操作等业务开展与费用控制情况进行全流程数据报表汇总及分析，并将数据分析结果及时上报、反馈，使医院管理者“心里有数，管理有据”。同时，根据多维度数据分析结果，指出存在的问题，明确改进方向，指定精准有效的管理措施。

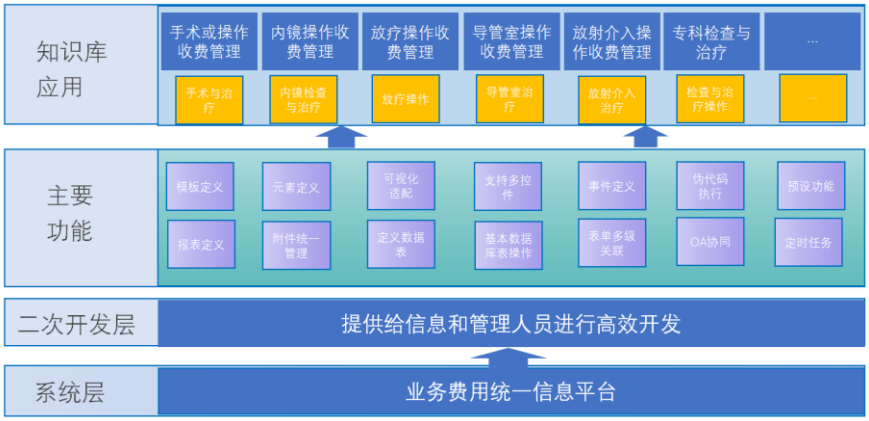


图2业务费用统一信息平台架构图

4 结论

通过业务驱动医疗服务项目收费实行研究，建设“业务收费统一管理平台”，实现业务驱动医疗服务项目收费，具有以下意义：

（1）减少超标准多收费、超范围乱收费、套用项目收费、重复收费、不按项目内涵收费等现象，实现医疗行为全过程监管；

（2）促进医疗机构管理措施的科学合理化；

（3）促进医保制度的优化完善；

（4）促进绩效考核管理与医疗服务价格管理

本文的研究可促进绩效改革对于工作量的合理分配，确保工作量可细化至个人，促进绩效考核的公平性、合理性、精细化；本课题建设优化医疗服务价格管理权限配置、完善医疗服务价格管理的支撑体系，推进公立医院补偿机制、绩效改革、医疗控费、医保支付等相关改革，增强改革的系统性、整体性、协同性，形成综合效应。

PU-272

基于患者视角的智慧医院建设研究综述

李欣悦 王安莉

湘雅三医院 中南大学生命科学学院 医学信息研究湖南普通高等学校重点实验室

1. 智慧医院概述

1.1 智慧医院的内涵

智慧医院的内涵涵盖面向不同受众对象时所定义的不同方面。针对医院的三大类对象，即患者、医护人员和管理者，智慧医院的定义也不同[1]。面向医务人员的“智慧医疗”主要指以电子病历为核心的信息化建设，通过医院局域网实现病历和影像、检验等其他系统之间的互联互通。面向患者的“智慧服务”则主要借助智能化设备让患者感受更加方便和快捷的就医体验，如医院挂号缴费一体机和自助报告打印机的应用，患者可通过手机预约挂号、预约诊疗和出院结算等。而面向医院的“智慧管理”，则指精细化的成本核算是医院精细化管理的重要依据。管理者可借助信息化管理系统，使用手机或办公室电脑查看全院运转状态，包括办公自动化(office automation, OA)系统及医院内部后勤管理等。这就是医院智慧管理领域的定义[2]。

1.2 智慧医院建设的关键技术

智慧医院建设的关键技术主要是指将信息化技术应用于医疗服务中的各种技术手段和工具。主要包括以下医疗信息化系统、物联网技术、人工智能技术、大数据分析技术、云计算和边缘计算、虚拟现实和增强现实技术等，这些技术的应用可以提高医疗服务的质量和效率，改善患者就诊体验，推动医疗行业的数字化和智能化发展。

1.3 智慧医院的发展现状

2019 年，国家卫生健康委发布了《医院智慧服务分级评估标准体系》，这意味着未来推进智慧医院建设的基本框架已经初步建立[3]。我国医疗行业在发展过程中，在医疗技术和基础设施建设方面取得了重要进展，但智慧医院功能仍然存在不完善的情况，信息技术并没有被充分整合到医疗过程中[4]。随着医疗卫生改革的持续深化，医疗机构的智慧建设成为推动医院发展的重要驱动力[5]。经过多年的信息化建设，医院在为患者提供服务方面取得了显著进展，实现了方便患者预约诊疗、方便智能导医分诊和方便候诊提醒；实现了快速查询检验检查结果、快捷诊间结算和便利移动支付；患者就诊往返医院次数减少、门诊候诊时间缩短、院内重复排队次数减少以及平均住院日减少[6]。

2.智慧医院建设优化医疗服务流程

2.1 患者信息管理与隐私保护

智慧医院建设对患者信息管理与隐私保护具有重要影响。一是电子病历与患者信息管理方面，智慧医院建设利用电子病历系统实现患者信息的数字化和集中化管理。电子病历记录了患者的病史、诊断、治疗方案等信息，对医生提供了全面的医疗数据。通过电子病历系统，医生可以快速查阅和更新患者的健康信息，提高医疗服务的准确性和效率。二是数据的共享与互通，智慧医院建设将不同部门和医疗机构的数据进行整合和共享。通过数据共享，医生可以从其他医疗机构获得更全面、准确的患者信息，提高诊断和治疗的水平。数据的安全和隐私保护这两个方面是患者信息管理的重要方面。应制定明确的数据共享协议和规范，明确数据使用的目的和范围，并履行必要的数据保护义务，以保证患者的隐私权不受侵犯。最后是患者的参与权与控制权，智慧医院建设为患者提供了更多的参与到医疗决策和管理过程的机会，同时患者对自己的数据享有控制权和决策权，医院应确保患者知情同意、数据访问、删除和更正等权利，同时提供透明的隐私政策，保护患者的隐私和个人信息。

2.2 医疗过程的优化和标准化

智慧医院建设对医疗过程的优化和标准化产生的积极影响主要是体现在流程优化、标准化管理、实时监测与跟踪等方面。流程优化：智慧医院建设利用信息化技术对医疗流程进行优化，提高医疗服务的效率和质量。通过引入电子化的预约、挂号、排队系统，患者可以在家中或移动设备上预约医生，减少等待时间和人群聚集。智能医疗设备和系统可以实现自动化的数据采集和传输，减少医务人员的重复操作和人为错误，提高工作效率。标准化管理：智慧医院建设将医疗业务过程进行标准化管理，制定和推行各项规范和流程，提高医疗质量的一致性和可控性。通过制定标准化的诊疗方案、手术流程和护理操作规范，可以减少人为差异和医疗错误的发生，提高医疗过程的安全性和可靠性。实时监测与跟踪：智慧医院建设引入物联网与传感器技术，可以实时监测和跟踪患者的生理参数、病情变化以及医疗设备的状态。通过及时采集和传输数据，医务人员可以远程监控患者的健康状况，提供及时的医疗干预，减少潜在的风险和意外情况的发生。

3. 智慧医院建设提升患者就医体验

3.1 便捷的医疗服务体验

智慧医院建设通过引入信息化技术和创新的服务模式，提供了更加便捷的医疗服务体验，如网上预约与挂号、电子病历和资料管理、远程诊疗和远程监护、移动医疗应用程序和数据驱动的个性化医疗等措施，提供更加便捷的医疗服务体验。这些创新的技术和模式使患者可以更方便地获得医疗服务，减少等待时间和就医步骤，提高医疗过程的便利性和效率。

3.2 个性化的医疗服务

智慧医院建设带来的明显变化之一就是数据驱动的医疗决策。智慧医院利用大数据分析和人工智能技术，对病历、检查结果、实时监测数据等数据进行详细分析。通过医疗诊断数据分析，医生可以对患者病情做出更严谨的分析和判断、疾病可能的发展趋势以及最优的治疗方案。医生可以根据患者的个体特征和需求，提供更加个性化的诊断和治疗方案。制定个性化诊疗计划包括治疗方案、用药剂量、手术方案等，更好地满足患者的需求和条件，提高治疗效果和患者满意度。同时智能设备和传感器应用、远程医疗和远程监护、移动应用和健康管理平台等服务也能够提供更加个性化的医疗服务。

3.3 增强的信息透明度与参与度

智慧医院运用数字化和信息化技术，带来了以下好处。首先，患者能够方便地访问和获取个人病历、检查结果以及诊断报告等医疗信息。这样的举措提高了信息的透明度，使患者更全面地了解自身病情，有助于他们更准确地理解医生的建议和治疗方案。此外，智慧医院还公开了医疗服务质量指标、医生资质信息以及医院排队时间等等，为患者提供了更多决策所需的清晰信息。

另一方面，智慧医院的建设促进了患者的主动参与和决策权。患者可以通过智能设备、移动应用等各种渠道随时与医院进行沟通和咨询，提出问题和关注点。他们还能参与制定个性化的诊疗计划，与医生共同讨论和决定治疗方案。这种积极参与使患者更加主动地管理个人健康，并与医生建立更紧密的合作关系。整体而言，医患互动质量得到提高，医患之间的沟通更加顺畅和高效。这样的改善有助于建立更为信任和良好的医患关系，从而提高整体的医疗质量。

3.4 智能化的健康管理

智慧医院建设在智能化健康管理方面具有重要影响。通过实时监测和远程护理、个性化健康管理方案、疾病风险评估和预测、智能健康咨询和辅导以及健康数据的共享和交流，智慧医院建设为个人提供了更加方便和有效的健康管理方式。基于人工智能技术，系统可以提供个性化的建议和指导，解答患者的疑问，为患者提供及时的健康支持。智慧医院建设使得医疗数据的共享和交流更加便捷和安全。患者可以授权医护人员、家庭成员甚至第三方应用使用其健康数据，实现跨机构的医疗信息互通。这有助于医护人员更好地了解患者的健康情况，提供更精准的医疗服务。

4.结语

本论文从患者视角出发，探讨了智慧医院建设对医疗服务流程和患者就医体验的影响，包括智能化的健康管理、更加便捷的医疗服务、就诊流程标准化等方面。虽然当前智慧医院建设取得了一些可喜的成果，但是未来的智慧医院建设仍面临一些挑战，如患者就诊数据的隐私保护和智慧医院建设评价方法等方面。在未来，应继续加强研究和实践，提升智慧医院建设的科学性和实用性，为智慧医院的高质量发展提供可行性建议。

PU-273

新冠肺炎疫情防控常态化下如何提升智慧医疗服务

李欣悦 王安莉

湘雅三医院 中南大学生命科学学院 医学信息研究湖南普通高等学校重点实验室

1.我国新冠疫情发展现状

新冠肺炎疫情爆发以来，我国国务院联防联控机制迅速对新型冠状病毒感染的肺炎疫情防控工作进行全面部署。我国卫健委迅速成立新型冠状病毒感染的肺炎应对处置工作领导小组，指导地方做好疫情应对以及防控工作，尽快开展新冠疫苗的研发项目。全国各个相关省份以及有关地区进一步落实自身责任，采取切实有效措施，落实新冠疫情防控中的各项工作。随着疫情的发展和有关政策的落实，疫情防控指挥政策也在因“况”而变，我国的疫情防控由最初的局部防疫阶段，到全民

抗疫阶段，最后到现如今的疫情防控常态化，统筹兼顾，保证国民健康和人身安全的基础上，有秩序地恢复各个企业的日常工作，加快恢复国民经济的正常运行。

现如今，新冠肺炎疫苗的研发取得了突破性进展，我国国民大多已经完成新冠肺炎疫苗的接种，大大降低了人民感染新冠肺炎的概率，但是随着新型冠状肺炎病毒的变异，疫苗并不能完全保证人身安全，尤其是最近上海疫情的全面爆发，单日新增确诊人数突破新高，给医院的接诊带来了极大的困难，普通患者就诊面临着诸多不便，纵观国内外疫情的发展趋势，新冠肺炎疫情正在朝着大流行的方向发展，疫情防控形势仍旧不容乐观。

2.我国智慧医疗服务现状

随着我国互联网技术的发展，基于电脑终端和移动设备终端的患者预约就诊、健康咨询、医药电商等服务行业逐步兴起。早在 2014 年国家卫健委就第一次提出了医疗机构可以通过信息技术手段对患者人群开展远程服务，也就是远程医疗服务。最近几年，我国对于智慧医疗服务的发展给予了高度的重视，并且出台了有关政策，如《国务院办公厅关于促进互联网+医疗健康发展的意见》、《关于推进新冠肺炎疫情防控期间开展“互联网+”医保服务的指导意见》等，这都是在社会治理信息化的过程中国家重视智慧医疗建设的有关依据。智慧医疗服务是在信息技术的持续高速发展和智慧城市建设的不断推动下，人民群众的健康管理与医疗行业信息化、医疗行业智能化相互融合的高级阶段。智慧医疗是在信息化技术的支持下，由传统医疗走向创新的结果。

目前，我国的智慧医疗服务主要可以用于疾病诊断、病理学诊断、远程医疗、居民健康管理、替代传统的门诊挂号以及优化就诊流程等方面，目前，我国还处于智慧医疗发展的初级阶段，智慧医疗相关立法工作以及配套的法律法规尚未完善，也在一定程度上制约了智慧医疗的发展。因此，在加快我国智慧医疗建设的同时，智慧医疗服务的法律法规体系构建也迫在眉睫。此外，虽然我国医疗大数据的体量庞大，但是各级医疗机构的大数据以及互联网等不同来源的医疗大数据之间尚未实现统一整合及信息数据的互通。医疗大数据的整合、挖掘和优化，有效数据的提取，并通过相关软件或医疗信息平台应用到我国日常的医疗工作之中是一个巨大挑战，还需要进行更多的研究和探索。不仅如此，现阶段并不是所有患者都能完全接受智慧医疗这一新生事物，无论是对传统医疗的固有思想还是对于接纳新生事物的抵触，都导致了这部分人群认为智慧医疗可信度不高，难以接纳。因此，对于智慧医疗的规范化宣传以及试点运行尤为重要，通过社会主流媒体官方进行宣传以及使用人群之间相互宣传，可以进一步提高居民对于智慧医疗的认识，进一步实现群众对于智慧医疗的接受度，进一步促进智慧医疗服务的发展和完善。

3.新冠肺炎疫情防控常态化下智慧医疗服务的适应性调整

3.1 保持与国家疫情防控政策一致

自疫情发生以来，我国相继出台了一系列有关财政、国民安全等的有关政策和通知，包括《国务院办公厅关于进一步精简审批优化服务精准稳妥推进企业复工复产的通知》、《关于改善一线医务人员工作条件切实关心医务人员身心健康的若干措施》和《财政部关于对中小微企业贷款实施临时性延期还本付息的通知》等在内的一系列政策，智慧医疗服务的调整应当与这些政策保持高度一致，以政策为基础和出发点，以人民的身体健康和核心利益为原则，优化智慧医疗服务，提升智慧医疗服务患者服务质量，保障人民的身心健康。

3.2 简化智慧医疗服务的使用方法

因新冠疫情的发生，医院降低了其他患者的住院率，采取非必要不留院的安排，那么对于一些处于隔离和封控中的人员，远程医疗已成为必要之选，这其中不乏老年群体，但他们中的很对人对于智慧医疗服务并不了解，甚至不能熟练使用智能手机，那么智慧医疗服务于他们而言并不能充分发挥作用，因此，智慧医疗服务的使用要更加趋于简单化和便捷化，以便于各个患者的注册和使用，提高智慧医疗的使用效率。

对于智慧医疗服务而言，最重要的一个特性就是其便利性，相对于传统医疗服务模式而言，智慧医疗服务不仅大大缩短了患者在医院滞留的时间，患者可以通过在线预约、在线挂号的方式，减

少不必要的排队时间，还有效地降低了医院因人员密集带来的踩踏危险。通过推行预约挂号和自助挂号，分散某一时间段挂号就诊的人次，从而减少患者排队挂号的时间。

推进电子就诊卡的使用，实现电子就诊卡一卡通，包括在线预约挂号、在线缴费、远程医疗等各个服务于电子就诊卡一身，将就医流程简单化，优化信息发布系统、自助查询系统和患者随访平台，降低患者终端的操作性和技术性，真正实现惠民利民的智慧医疗服务。

3.3 加强智慧医疗服务的基层建设

智慧医疗服务目前在大城市之中应用较为广泛，但是对于一些偏远地区和一些经济不发达地区尚未开通智慧医疗服务，而且在疫情期间医院的资源配置倾斜，一些地区的患者去医院就诊十分困难，甚至有些患者得不到有效的治疗，所以应该在与防疫政策保持一致的基础上，加快这些地区的基层设施构建，比如利用 5G 通讯技术，而且 2019 年，我国工业和信息化部和国家卫生健康委员会专门制定了数字化医疗方面的有关文件，进一步加强了我国智慧医疗的建设，在政策的扶持下更应该加快这些基层建设，早日实现智慧医疗服务全国范围全覆盖。

3.4 完善智慧医疗反馈和评价机制

智慧医疗服务在我国目前仍处于发展阶段，各个方面的设施和建设并不完善，我们在借鉴发达国家的已有智慧医疗服务的经验的同时，应该结合我国的基本国情做出调整，在这一不断摸索过程中，群众的声音至关重要，智慧医疗服务最终服务的也是群众，故智慧医疗服务系统的构建和完善需要倾听群众的声音，接受有建设性意义的评价和反馈，改进智慧医疗服务中存在的问题和不合理之处，所以，智慧医疗服务的反馈和评价和机制的设置是十分必要的，不仅要设置反馈和评价机制，还要保证群众反馈意见的真实性和便利性，以便于智慧医疗服务的更好发展。

3.5 加快智慧医疗法律体系的建设

智慧医疗依托的是网络数据，既然是网络数据就会出现患者人群信息安全等相关问题。无论是在线咨询服务，还是远程医疗服务，患者人群的就诊信息、疾病诊断，甚至是患者的基本信息都有被第三方恶意窃取的可能，一旦发生患者隐私泄露等问题，其实很难能够追踪到根源，明确哪一主体的责任。当患者浏览网站时，不慎点击网站外部链接广告，患者信息就有可能遭到泄露。所以说，智慧医疗的发展建设离不开国家加快健康信息保护的法律法规体系建设。不仅是信息安全，使用智慧医疗的患者群体在接受医疗服务的过程当中，一旦出现了侵权损害，或者相关医疗事故时，就会面临服务主体不明确，医疗执业机构与相关网络平台责任划分不清晰等困境。因此智慧医疗相关立法工作应与时俱进，不断更新修订。

4.结语

新冠肺炎疫情防控常态化给智慧医疗服务带来了诸多困难和挑战，但是对于智慧医疗服务的优化升级来讲未尝不是一个机遇，我国智慧医疗仍处于发展阶段，需要在借鉴欧美一些发达国家智慧医疗发展经验的同时，需要结合我国的基本国情做出适应性改变和调整，需要归纳总结出适合我国国情的智慧医疗发展道路。在发展智慧医疗的同时，应根据使用者对于智慧医疗的评价，进一步反馈于智慧医疗的发展，实现智慧医疗的健康、可持续发展。

PU-274

公立医院连续性医疗服务体系的信息化路径研究

左秀然 彭明毅 杨国良
武汉市中心医院

1. 引言

近年来，国家高度重视以人民健康为中心的医疗服务体系构建，提出三级公立医院在紧密型城市医疗集团中要发挥龙头作用，完善网格化服务链条，统筹预防、治疗、康复、健康促进等一体化、连续性医疗服务[1]。明确了要科学合理网格化布局紧密型城市医疗集团，推动医疗服务供给侧结构性改革[2]。在深化医药卫生体制改革 2023 年下半年重点工作任务中，明确要促进优质医疗资源扩容和区域均衡布局。在推动临床专科能力建设方面，国家提出坚持“以患者为中心、以疾病诊疗为链条”，探索打破医学学科和诊疗科目壁垒。综上可见，公立医院转变医疗服务模式构建连续性医疗服务体系迫在眉睫。

文献 3 基于卫生服务连续性概念框架构建了卫生服务连续性评价指标体系，其中信息连续性是其中三个研究的维度之一。但支撑连续性医疗服务的信息化支撑仍存在着较为普遍的供需不平衡问题。需求侧，居民希望全程管理、服务可及、个性定制，医院人员需要有粘性的患者群体、高质量的数据支持、高效的信息系统支撑，医院管理者需要加大资源整合，将条块化管理模式向横向贯通相结合的模式转变，政府部门则需要区域医防融合管理，实现区域医疗资源集约化利用及公共卫生事件应急联动。以上需求需要信息化强有力的支撑，但供给侧，以应用为中心的条块式信息化建设方式无法从根本上满足以人为中心的管理需要。

2. 连续性医疗服务链条分析

对公立医院开展连续性医疗服务体系的业务链条做分析，从而指导开展信息化支撑体系设计。通过分析以人为中心的健康医疗全过程，总结开展连续性医疗服务需满足“四全”，即：全周期，从出生到死亡；全人群，从普通人群到重点人群；全链条，覆盖诊前、诊中、诊后、全程服务；全闭环，包含服务闭环、医疗闭环、管理闭环。以连续性医疗服务业务指导原则分解涉及的服务链条，包含：一是以患者为中心、以疾病为链条的临床诊疗协同，解决疾病诊疗患者以个人为中心系统性诊疗的需要，二是打开围墙线上线下一体化的延续服务，解决患者院外管理线上线下一体化管理的需要，三是区域医防融合的一体化管理，解决疾病预防、健康管理以及公共卫生应急管理协同的需要。通过三个链条，以居民为中心的角度，单体医疗机构服务需要探索向区域网格化整合型、连续性进行转变。

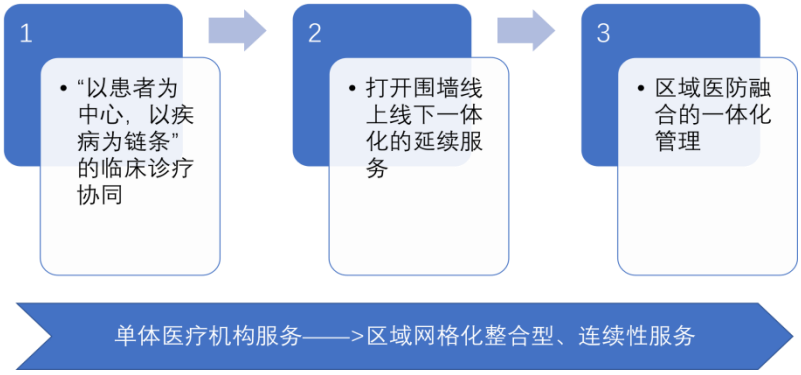


图 1 连续性医疗服务业务链条

3. 信息化支撑体系

近年来，国家推动“三位一体”智慧医院顶层设计，取得成效显著[4]。但公立医院更高水平的智慧医院发展，需要强化横向贯通区域医疗资源协同，从区域层面为居民提供连续一体化服务的能力。基于以上所分析开展连续性医疗服务体系的业务需求，提出信息化支撑体系的“六个协同+一套支撑”，其中“六个协同”为临床多学科协同、医教研协同、多院区协同、线上线下协同、区域医联体协同、城市公共卫生资源协同，“一套支撑”则为云网边端基础设施支撑体系。

图 2 “六个协同+一套支撑”

3.1 六个协同

（1）临床多学科协同：解决一人多病、一病多科诊疗的信息化支撑。

依托医疗大数据中心，建立疾病风险分级定量评估指标中心，以数据驱动预警，面向不同的诊疗学科推动多学科 MDT 联动协同，由平台管理科室制定干预措施，患者所在科室执行措施，并开展效果评价及干预措施的质控管理。进一步，患者出院后，与院外所在社区的基层机构对其联合开展协同共管，延续院内多学科 MDT+院外医联体基层机构 MDT。

（2）医教研协同：实现临床科研的单中心向多中心、区域专病大数据应用转变的信息化支撑。

医疗技术的提升，也有赖于科学研究创新成果的产出，传统做法下基于单体医疗机构的医疗大数据往往无法满足科学研究对大数据的需求，因此将单体医院的单中心，提升到多中心，进一步依托区域卫生信息平台开展区域专病大数据中心研究，可较好发挥大数据价值，提升临床科研能力。

（3）多院区协同：实现医疗机构多院区集约化资源配置、高效化资源协同统一调度指挥的信息化支撑。

公立医院发展“一院多区”是公立医院高质量发展的重要实施路径。2022 年以来，国家卫健委对公立医院多院区发展进行了总体部署，引导高水平的公立医院在控制单体规模的基础上，适度发展多院区，促进优质医疗资源扩容和区域均衡布局，同时提升重大疫情救治能力[5]。支撑一院多区、集团化发展，是信息化支撑体系中必须考虑的范围。在集约化资源管理方面，多院区各类号源资源池统一管理，包含专家专科资源、检查设备资源、床位资源等，管理决策应用覆盖决策层管理层、执行层系统覆盖电脑端、手机端、大屏端院区维度、专科维度、专病维度等。多院区协同，同时需建立统一的运营指挥系统，实现多院区智能运营和管理， 一键转换、一网监测、一体指挥。

（4）线上线下协同：解决线上线下医疗健康服务一体化协同、无缝衔接的信息化支撑。

建立互联网医院系统，并与线下业务无缝一体化衔接，满足患者诊前、诊后环节对医疗健康延续服务可及性的需求。可开展的服务有音视频问诊、图文咨询、复诊在线延方、名医直播、药师在线审方、用药指导、药品物流配送、自助线上检验检查开单、院外延续上门护理等。其中面向个人的应用，体现移动化部署特点，以武汉市中心医院为例，80%就诊流程实现移动化就医等待次数减少 4-6 次、等待时间减少 1/3-1/2。

（5）区域医联体协同：解决区域医联体内医疗资源同质化服务、资源有序下沉的信息化

支撑。

建设远程医疗协同平台支撑分级诊疗，包含远程会诊、远程培训、远程联合查房、远程联合门诊等远程协同服务。面向基层社区卫生服务中心实现紧密型业务联动，转诊号源基层开放、上转住院床位预约、患者自助入院登记直接病区登记入住、联合病房患者上转绿色通道等，线上线下无缝衔接提升转诊患者跨院就诊体验。

（6）城市公共卫生资源协同：满足所在区域城市对医疗机构开展医防融合、公共卫生事件应急指挥协同的信息化支撑。

依托医院承担的区域慢病管理中心牵头机构的定位，建立医院-社区-家庭-个人的四级智慧化血糖管理新模式[6]，医院端将多学科专病协同管理系统，对接至区域卫生平台的基层卫生系统，社区管理的居民的慢病公卫数据，可纳入风险预警指标，实现统一预警及干预管理。在与城市公共卫生事件应急的指挥联动机制方面，公立医院一是提高数据上传区域卫生信息平台的数据质量，满足区域对疫情防控多点触发预警监测的需求，同时院内应急指挥系统与区域平台对接实现应急会商联动，医用物资纳入区域平台实行统一应急物资调度。

3.2 一套支撑

在基础设施的支撑方面，业务功能条块式的信息系统建设需要转变提升为数据驱动业务闭环，并强化数据治理和数据资产管理，以提供高质量的数据开展健康风险评估预警的精准性。同时，在系统的建设理念上，建立以人为本的应用体系，实现全生命周期一体化医疗健康管理，建立任务驱动、智能辅助的医生工作门户，

医管融合、业财融合的联动应用，数据支撑的全面质量综合应用。为了支撑个人全生命周期管理，需要实现人通、资源通、信息通，例如医院、基层社区医生、第三方管理团队共组签约服务团队，预约资源打通共享、电子健康档案融合共享。在新基建方面，技术赋能、融合创新，例如物联网可穿戴实现动态连续体征数据采集分析、健康预测、人工智能 AI 辅助诊疗、智能语音识别技术、智能机器人、医疗健康大数据科学决策模型、 数字孪生技术等的全面应用。

4. 结语

在新的形势下，公立医院的服务主体发生了变化，将从单一主体到全人群，从疾病诊疗到全生命周期，最终还是将以患者为中心，进一步实现向前预防式管理以及向后的诊后延续服务。如何适应这种新的定位以及保持可持续发展，体现以人为中心的连续性健康医疗服务，数字化改革驱动观念和流程变革，医防协同一体化是发展思路，也是公立医院后续创新医疗服务模式的新课题，下一步，将继续以此主题继续研究深化连续性医疗服务体系中各环节的实现路径及方法。

PU-275

医院智能化建设及其在智慧医疗发展中的意义

郑健 薛志刚
哈尔滨医科大学附属第一医院

1 引言

随着我国社会经济的迅速发展，信息技术在我国各个领域都得到了广泛的应用，起到了举足轻重的作用,医院智能化建设也得到了人们的广泛关注。医院的智能化建设建立在信息化基础之上。所以，如何利用信息化技术来提高医院的智能化水平，就显得尤为重要。

2 智能化技术

医院信息化的发展加大了经费、人力资源的投入，逐步改善了医院的基础设施。同时，在日益增长的医疗信息数据中，传统的信息处理方式所存在的弊端也日益凸显出来。应积极推动“智慧医疗”的建设，科学应用现代科技，充分利用计算机仿真的功能。当前，医院“一卡通”、电子医疗档案等已经在全国范围内推广开来，数字医院的建设也取得了一定的成果。国家制定了一系列政策，增加了财政支持力度，引入了多种先进的科技，推动了传统医疗行业的现代化和智能化。在“云物大智移”中，特别是在智能化方面取得了长足进步，智能医疗是一种基于计算机技术的医疗技术，它可以通过计算机来模拟人类大脑处理信息的过程，代替人工信息的采集、处理、分析、存储等工作[1]。智能技术是一门集图像处理、计算智能、数据挖掘、数据融合于一体的人工智能技术。在信息处理领域，主要有两种类型：一种是以传统的计算机为核心，另一种是以神经计算网络信息处理为核心。

3 医院智能化建设在智慧医疗发展中的作用

3.1 智能系统的紧密集成

传统的智能楼宇系统是医院发展不可或缺的一部分，但部分医院由于缺少相应的业务管理流程，缺少高效的智能化处理方式，造成了资源利用的不平衡。所以，我们所要面对的问题就是，如何让智能楼宇系统变得更有效率。紧密集成智能系统，实现资产、服务及空间管理，进而提升医院日常运作的智能化和高效化。

3.2 实现对设备的智能化控制

智能医院是通过网络来智能化地控制设备，通过网络来完成智能的安全监控，智能的环境控制，移动的医疗，以及人员和设备的定位。与此同时，通过智能化系统对有关信息进行处理，可以让医院的环境变得更加舒适、节能、高效，从而更好地降低运营成本，提升医患关系的满意度，也对减少医疗事故的发生起到了重要的作用。

3.3 推动医院数字化与智慧化建设

在智能化医院的建设中，要以病人为本，使病人不用再进行现场排队。通过网上挂号，设备自动打印检查单，当天就能完成检查、住院费用的结算服务，为病人提供了极大的便利。这样可以使病人在就诊时体验到更好的治疗服务，缩短了接诊时间。同时，可减少病人的复检和反复收费。另外，该系统还可应用于病人的病情监测、病人的满意度调查、病人的后续治疗和病人的出院评价等。

4 医院智能化建设

4.1 概述

智慧医疗中心的规划设计与建设，涉及数字影像、设备集成、电子医学，临床应用、科学研究、远程医疗等多个领域。楼宇建筑的智能化是实现其建设目标的关键保证[2]。构建一套基于智能的新型信息管理体系，推动医院内部管理与诊疗的顺利开展，使之与现行的诊疗流程相适应，创造出属于自己的医院服务品牌。当前，医院楼宇智能化建设存在着明显的自主性。在推动智慧医疗的

发展中，要改变这一现状，必须将安防、暖通、电力、消防等多种系统有机地结合起来，实现内部的有效协调。应坚持信息一体化、过程数字化、建筑智能化、管理规范化的“四化”原则[3]。智慧医疗的构建是一个长期而系统的过程，并非一朝一夕之功。要从医院的实际发展出发，对每个步骤进行科学地规划，逐步推进，使数字化建设与医院的业务流程相结合，从而减少医院的管理费用，达到资源的最优分配，保证医疗质量。

4.2 综合布线系统

综合布线系统是信息传递的主要通道，也是医院信息传递与共享的主要渠道。医学图像数据和诊断信息可以在综合布线系统中传送。为了使该系统具有多样化的功能，使其能够支持不同的计算机网络协议，以达到信息互联互通的目的。要加强技术的研究和开发，加强网络解决方案的灵活性和容错能力，建立语音系统、医学图像传输、医学数据存储等基本功能模块。综合布线系统是医院数字化信息化建设的主要保证，必须提高医院 HIS,CIS,LIS,PACS 等设备的可靠性，并为其提供一个良好的通信环境。从而使医院各个部门信息的采集、存储和处理工作更为顺畅，以满足广大用户的需要。

4.3 建筑设备管理系统

硬件设施对软件系统的稳定性起着决定性的作用。要对各种设备进行科学地控制，保证其在较长的时间内安全、稳定地运转。为降低设备维修成本，提高设备运行效率，需建立一套能够自动检测、监控、故障定位、上传的设备管理系统。按照医院的具体要求，研制出与之相适应的建筑设备系统，例如：暖通系统：可同时控制空调机组和新风机组；电气系统：主要负责照明系统的管理，直流电源的控制；电力设备的监测：交通设备主要负责对扶梯、电梯进行监视。

4.4 门禁管理系统

门禁系统是当前医院普遍采用的一种管理系统，它的应用范围很广，如电梯间、楼梯间、病室入口等。通常采用监控系统等方式来进行管理。主要的管理区有：财务室、电脑室、高值耗材库房、药品库房、贵重仪器库房等。

4.5 数字监控系统

医院的属性是公共服务类型，人流量较大。除了大量的出院、入院病人之外，医院里的工作人员也有很多，很容易引起一些安全隐患问题。利用数字监控系统对患者进行实时监测与记录，能够为病人营造一个良好的医疗环境，为医务人员营造一个安全的工作环境，保障全体人员的生命与财产安全。

4.6 智能化医护对讲系统

在每一间病房的床头都可安装一套智能医疗对讲系统，通过呼叫机与护士站进行通讯，提高通信效率，防止护理人员产生手忙脚乱的现象。而且一旦有什么突发状况，也能及时地通知护理人员，到了夜晚，也能用对讲机进行自动循环监听，保证每一位患者的生命安全。

5 对医院智能化建设的几点建议

5.1 增强智能化意识，加强领导重视

在我国，要加强医院员工的智能化意识，加强医院的管理，提高医院的智能化水平。在实施这一过程中，医院的管理者或领导者应充分认识到医院智能化建设的重要意义及发展方向，并据此制定出更为完整的发展方向与策略。

5.2 制定智能化建设地总体规划

为推动医院智能化建设的顺利进行，有关部门必须对医院智能化项目进行全面规划。这就需要有关的管理者在智能化建设之前，对整个项目展开一种科学、合理地评估、分析和总结，进而对医院智能化的建设过程中可能出现的问题和影响因素进行全面的分析，从而推动医院智能化工程建设的顺利进行。

5.3 加强人才的培养

要使医院的智能化工程得以成功的开展，就必须加强各种类型的人才培养。在这个过程中，有关的管理者要对所有的工作人员进行信息技术、医学知识和职业技能的训练，来推动医院的智能工程建设，保证所有的工作人员都能够肩负起自己的责任，并严格遵守有关的标准和要求。

6 发展方向

就其发展趋势而言，与我国相比，国外在智慧医疗方面的研究起步较早，在深度与广度上也比较成熟。近年来，伴随着社会对智慧医疗的需求日益增长，其研究结果较为丰富。就其研究内容而言，国外学者对于智慧医疗的研究，多是围绕着技术创新，信息安全，以及构建智慧医疗的应用场景。目前，我国对智慧医疗的研究多是从新技术应用、标准规范、服务模式和影响因素四个方面进行的。智慧医疗是“健康中国”战略实施的有力支持与保证，也是当前国家卫生事业发展的一个主要方向。智慧医疗是提升患者就医便利性、改善健康管理水平、满足不断增长的健康服务需求的重要手段与策略。推动云计算、大数据、物联网、区块链和 5G 等新型信息技术与健康服务的深度融合，是未来智慧医疗的发展趋势[4]。

7 结语

综上所述，我国的科学技术已经有了很大的进步，医院的智能化、科学化和信息化正在不断地发展。大力发展智慧医院，增加人力物力的投入，加强信息化建设，建立“互联网+”的智慧医疗服务体系。为进一步提高医院的内部管理，提高医院的工作效率与品质，医院的智能化将会得到更多的应用。要牢固把握智能化建设对智慧医疗的重要作用，即紧密结合智能系统，智能控制设备，数字智能化建设。医院应加大资金和人才的投入，为患者着想，充分利用人工智能技术，促进医疗行业的现代化。

PU-276

基于内控设计的非医疗服务项目收费平台设计与应用

陈云 刘剑锋* 梁诗弋
重庆市妇幼保健院（重庆医科大学附属妇女儿童医院、重庆市妇产科医院、重庆市遗传与生殖研究所）

1、非医疗服务项目收费平台建设背景研究

1.1、非医疗服务项目增多

随着人们生活水平的提高，医疗机构的高质量发展，患者对非医疗服务的需求增加，医院为能够满足患者对于生活品质的需求，提高患者在医院的整体就医体验和优质服务，通过增加自身服务或者与第三方机构合作为患者提供丰富的非医疗服务，如陪护、陪检、基因外检、就餐、停车、便民、租借押金等。随着服务项目的增多，对多收费场景、收费实效有了更高更灵活的要求，传统的手工单据和人工收费已经无法满足医院的内控管理要求。

1.2、医疗机构收费内控设计的重要性

多而杂的非医疗服务项目收费与管理有别于传统的医疗收费、退费、发票、资金账户归口、信息存储的管理模式，这促使医院管理者必须根据非医收费的独特性重新构建内控方案，制定合规性风险控制策略，从管理、执行、监督等角度出发，使支付功能、业务流程符合医院内控管理要求，规范资金业务活动，保证医院资产和资金安全，提升服务效率，提高客户满意度。

1.3、非医疗服务收费给医院内控管理带来挑战

随着公立医院改革的推进，精细化管理将成为医院高质量发展的必然趋势，医院在提供越来越多的便民服务，有的医疗机构引进第三方服务，让病人一站式完成就医体验，在给患者带来方便的同时，也给医院带来了巨大的管理挑战。

1.3.1、合法可靠性、

医院必须遵守国家相关法律法规和政策文件，确保其收费行为符合法律规定。未经患者同意不得强制消费或乱收费。

1.3.2、流程标准透明性

医院应向患者详细介绍非医疗服务项目的收费流程、价格及项目内容，在清楚明白的情况下由患者自主选择是否消费。

1.3.3、监督管理

医院应建立相应收费科室的信息联络员，确保各部门严格执行收费标准、流程和信息公开原则。周期性地对非医疗服务的收费核算数据进行审计，确保其合法性和合规性。

1.3.4、资金监管风险

医院需建立独立的非医收费资金账户，在收费、退费控制、资金对账、第三方机构对账分账健全管理制度，严格管控非医收费套现等漏洞。

2、基于内控管理的非医疗收费平台设计要点

非医疗服务费收费场景多、支付渠道的多样性，给医院财务工作带来管理模式的变化与挑战，必须通过建设完善的非医疗服务收费平台（以下简称非医平台），集中管理、控制结算风险，解决非医疗收费管理工作中的时效、人力和内控管理方面的问题。

2.1 非医平台内控需求分析

（1）非医平台要满足现阶段的主流支付方式：现金、银联、第三方支付（微信、支付宝），并实渠道流程现动态分配。（2）支持线上、线下快捷缴费，减少病人缴费排队时长。（3）病人信息与 HIS 系统进行数据联动，保证信息一致性和数据隐私安全。（4）建立独立的微信、支付宝收

款商户号。（5）对账的智能性，对账周期、对账异常、数据安全设计。（6）票据管理无纸化、电子化，核销全流程管理

2.2 非医平台主要功能设计

非医平台主要面向全院有收费需求的科室，针对不同支付场景设计多种缴费渠道，同时考虑每种收费方式下的财务风险，根据资金管理的内控要求，利用信息化手段整合转账、银行卡、微信、支付宝、现金、支票等多种支付渠道，并给财务管理者提供方便、快捷、高效、准确的数据结算汇总、对账服务，提升财务管理质量和运行效率。

2.2.1 权限定义：由财务管理者开、停相应临床科室的收费权限，并赋予可以缴费的非医疗服务收费项目，指定科室的收费渠道，1）线下缴费，科室打印二维码收款码，引导病人线下便捷缴费，2）线上开单缴费，诊间能进行移动支付，如需现金支付，需要去收费室进行缴费。通过诊间移动支付解决病人少跑路缴费问题。

2.2.2 服务项目管理：自定义服务收费属性，项目名称、规格、单价、服务商、收费模式、项目财务分类、使用票据类型（自制收据、增值发票），退费流程等等，

2.2.3 日报管理：平台完成财务管理与收费员的日常财务稽核。

2.2.4 退费控制：平台支持退费流程动态配置，满足多中支付下的灵活退费及财务风险管理。

2.2.5 报表查询：平台可根据财务管理及绩效管理需求生成多种固定格式的收费报表，统计支付方式汇总及明细。按需授权，使临床科室和财务管理者看的数据有区分度。

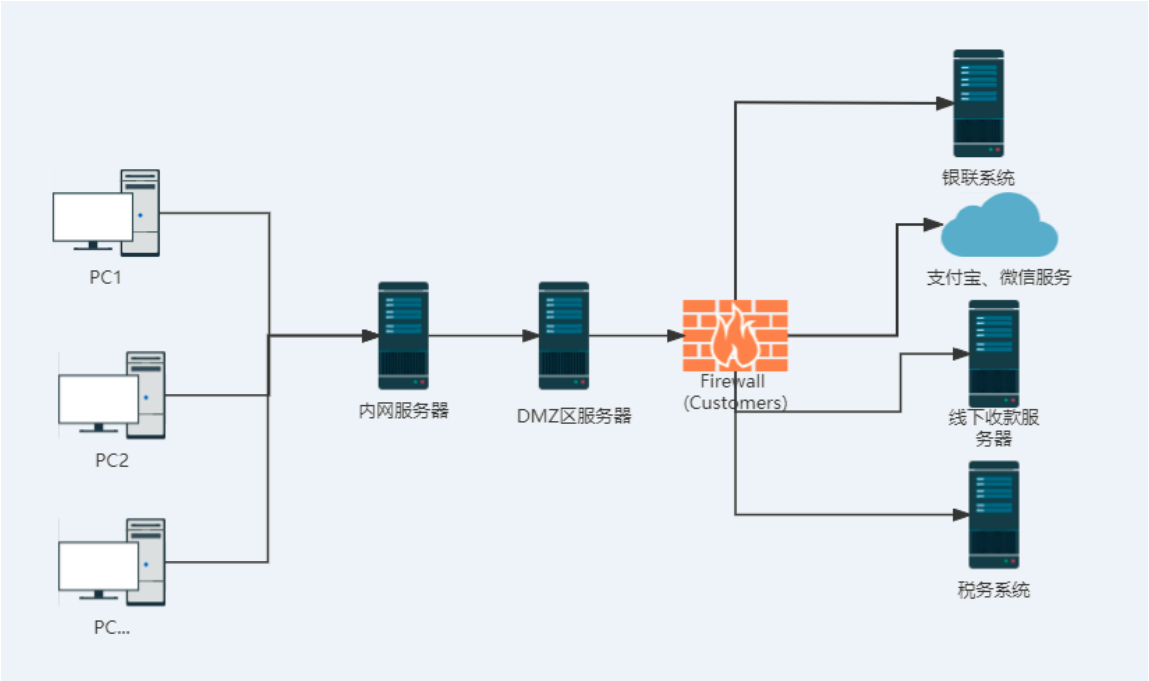
2.2.6 单边账处理：由于网络原因，可能会造成患者移动支付时，出现移动端端已扣款，非医平台未结算成功，形成“单边账”。系统会在次日自动下载账单比对核实,将单边数据展示出来，经财务人员核查后确实是“单边账”的，则通过原路返回的方式退回，

2.2.7 三方对账分账

非医平台会自动下载账单对账，与移动支付平台，银联平台进行对账，对账通过后，第三方服务收费统一管理，根据项目定义完成收费分账。按月度结算。

2.2.7 电子票据管理：系统引入公众号平台移动端入口，对接税务系统，实现院内发票电子化，无纸化流程管理。

2.3 非医平台系统架构



3、非医平台配套措施

1、财务制度

完善的制度是信息化平台建设的基础。移动支付的兴起与快速发展，影响了患者支付思维和习惯，也影响了医院资金管理模式。因此，资金管理制度建设必须跟进配套完善，以制度为基础进行平台功能设计和岗位使用管理，从而实现完善的内控体系。制度建设应该包含，非医疗服务收费管理办法、第三方合作方管理制度、各方对账管理制度、退费管理制度、收费全面稽核制度和票据管理制度，资金 T+1 回款制度等。定期对科室加强非医疗收费培训。

2、宣传公示制度

非医疗服务收费一定要跟患者做好宣传、做到收费透明性、自愿性、公正性。接受社会监督，给患者提供全流程知晓服务。

3、信息管理制度

信息安全是首位任务，医院管理要制订和完善信息化管理制度及具体操作规范，合理设置岗位操作权限，严控技术人员后台修改数据，确保数据的一致性，不断推进风险防控机制的建立和完善。

4、非医平台建设成效

非医平台采用迭代开发模式，于 2020 年 6 月底完成第一版本发布至今，共收纳 100 余项收费项目，使用科室 20 多个，第三方服务机构 12 家，服务人次 70 万人次，收费金额占医疗收费的 3%。

5、结束语

随着医疗高质量发展，病人的就医需求逐步升级，对非医疗服务的需求越来越多，通过对非医平台的建设，切合医院的精细化、数字化、信息化管理，基于财务内控机制，展开科学合理、针对性强的项目设计及应用实践，探索项目建设资金控制、应用场景灵活扩展、异构网络信息安全交换共享等工作机制，提高财务管理规范性，提升患者服务质量，丰富了医院医疗数字化建设内涵，通过调研，国内没有类似系统，通过本系统设计，对于各级区域妇幼服务平台建设具有指导和借鉴作用。

PU-277

一院多区可追溯智慧系统助力医疗物资精细化管理的应用实践

张司露 李慧慧 郭金鸽

郑州人民医院

目的 为构建智慧医疗、智慧服务、智慧管理“三位一体”的智慧医院，探索一院多区全程可追溯智慧医疗物资的一体化管理模式。

方法 以患者为中心，保证医疗质量安全为前提，构建多系统联动的信息架构，建立从医疗物资采购到医患使用的全生命周期可追溯体系。

结果 引入三方合作模式统一归口管理，理顺管理机制，提高工作效能。通过系统互联互通，使业务形成数据闭环，促进管理模式变革，提升医疗行为的规范性，减低财务风险，从而体现医疗质量管理的价值。

结论 一院多区全程可追溯智慧医疗的最大优势在于以患者为中心，保证医疗质量安全为前提，智慧医疗、智慧服务、智慧管理联动，数据互联互通、监管前移，建立预警机制提高供应链风险控制能力，专业的人干专业的事，最大化的保障临床价值和物有所值，助力智慧医院高质量发展，为医院精细化管理及清廉医院建设提供经验借鉴。

PU-278

门诊慢特病跨省直接结算实践与思考

张继民[1] 王宇彤[1] 刘志国[3] 王冰[4] 孙学勤[1] 朱卫国[1,2] 韩丁[5]
[1]中国医学科学院北京协和医院医疗保险管理处
[2]中国医学科学院北京协和医院全科医学科（普通内科）
[3]中国医学科学院北京协和医院财务处
中国医学科学院北京协和医院信息中心
中国医学科学院北京协和医院

1 引言

《社会保险经办条例》于今年 9 月 1 日正式公开发布[1]，第二十三条提出应当做好异地就医医疗费用结算工作。跨省就医需求持续增长，对异地直结工作提出更高要求，截止今年 8 月底，全国住院、普通门诊和高血压、糖尿病等 5 种门诊慢特病的跨省直接结算达到了 7216.71 万人次，减少群众垫付 972.1 亿元，与去年同期相比，分别增长了 245.10%和 91.77%[2]。其中，门诊慢特病相关治疗费用跨省直接结算作为减轻群众经济负担、疾病负担的有力举措进展显著，目前，每个统筹地区至少有 1 家定点医疗机构能够实现跨省直结。全国门诊慢特病跨省联网定点医疗机构达到 2.46 万家，累计惠及 73.02 万人次[3]。从文献分析的角度，通过系统检索，既往关于门诊慢特病跨省直结的文献研究较少，贵州省[4]、长三角[5]和攀枝花市[6]医疗保障局就门诊慢特病跨省直结情况、管理路径、进展与难点解决途径进行了报告和分析。新时代同样赋予了公立医院医保治理新的理念，应该放弃“等、靠、要”的传统思维，主动作为，建章立制，实现自我管理。

在相关政策和条例不断出台、异地就医需求不断增长的新时代下，某大型三甲医院-X 医院已于今年 8 月开通门诊慢特病跨省直接结算功能，本研究围绕流程设计中的重难点及实践进行总结分析，围绕问题提出展望和建议。

2 方案制定

2.1 成立专项小组

以“实现信息多跑路，群众少跑腿”为基本原则，以便利患者、减轻临床医生和职能部门工作负担为目标，由医保处牵头，会同信息中心、门诊部、财务处成立慢特病跨省直接结算项目专项工作小组，医保处、门诊部和财务处结合各部门主要业务及患者就诊涉及的挂号、换号、就诊、缴费等节，制定改造方案，信息中心根据方案进行系统建设。在试运行及正式运行过程中，各部门遇到问题立即沟通协商，持续优化和改进患者就诊流程。

2.2 流程设计

X 院目前可实现一种特病（恶性肿瘤）和两种慢病（糖尿病和高血压）的跨省直接结算，患者在参保地完成慢特病资格认定之后，需在就诊前完成医院内部慢特病身份识别。经前期调研，目前常见的流程模式有两种（见图 1），两种模式各有优劣，模式一采用收费员勾选的方式对患者身份进行识别并确认本次就诊病种，但多数患者不止有一种慢特病病种且收费员并无医学相关专业背景，导致就诊病种判断错误的情况时有发生，可造成患者多次往返诊室和收费窗口，减缓就诊速度，增加收费窗口压力。模式二设立独立窗口，由医保部门工作人员对本次就诊病种进行判断，这样做虽然保证了身份识别的准确性，但是增加了人力成本，且由于其并非患者就诊过程中的必要环节，很多患者忽略此环节直接进行后续治疗，导致结算时不能享受慢特病待遇进行报销。因此，亟需寻找一种准确且高效的方式，达到避免患者造成损失，同时减轻医院工作负荷的双重目的。X 院在模式一和模式二的基础之上制定了“收费员识别患者身份，医生判断就诊病种”的新模式，在换取医保号时收费员告知患者本次就诊是否可以享受慢特病待遇，患者也可以在挂号凭证上看到“异地慢特病”标识，在开具医嘱时，系统会首先提示医生患者有慢特病身份，进而由医生端对患者本次

就诊的病种进行判断，这种“收费员识别-患者确认-医生判断”的机制使得因误判等主观行为而导致患者延误诊疗、延误报销的比例大幅降低。

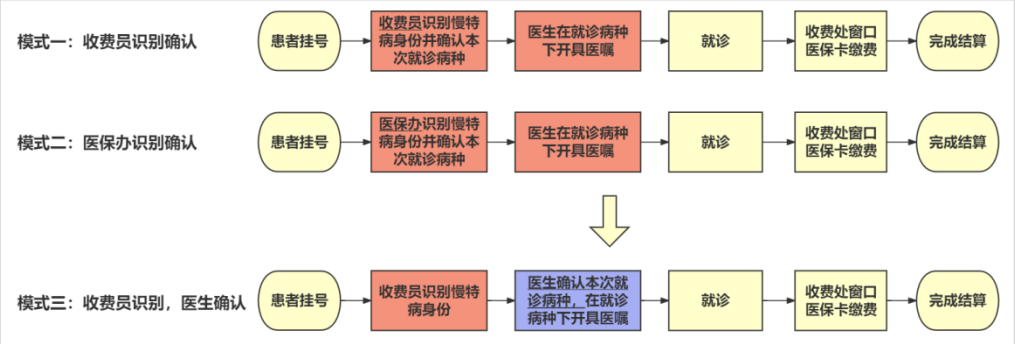


图 13 异地门诊慢特病结算流程模式

2.3 慢特病药品目录维护

遵循“就医地目录、参保地政策”的原则，各医院均按照北京药品目录对慢特病身份患者药品进行报销。部分医院采用人工核实药品的方式，由医保办、药剂科等相关主责部门对本次就诊药品是否为特病药品进行筛选，人工成本增加，且不能完全保证结果的准确性。X 院通过增强信息建设的方式对于医保药品目录进行审核，前期由药剂科对目录进行整理，由信息中心将目录补充进药品决策支持系统中，医生在开具医嘱时决策系统会进行相应提示，避免特病药品开在医保直结页签下或者非特病药品开在特病页签下而造成患者报销的损失。

2.4 宣传提示

患者政策宣贯是工作落实推进中的重要一环，X 院于第一时间更新微信公众号“门诊异地医保慢特病直接结算问与答”，阅读量达 3.5 万人次，同时同步更新了“异地普通门诊、慢特病门诊就医跨省直接结算须知”宣传海报，让患者多渠道、全方位了解相关政策，提高患者对慢特病跨省直结政策的知晓率，减轻收费窗口、医保咨询窗口和患者服务中心的压力。

2.5 系统试运行

医生对于系统改造的不熟悉可能造成误判患者就诊病种、在普通门诊页签下开具医嘱等多种情况发生，因此 X 院在正式上线该项功能前，在全院层面对于系统改造的几个特点进行了公开宣传，同时以 X 院放疗科作为试点进行运行，于正式上线前一周邀请四位放疗患者来院进行测试，一方面测试系统的稳定性和联通性，另一方面关注医生在操作过程中是否存在困难和堵点，进而保证在全院范围铺开后顺利进行。

3 阶段性进展

开通异地门诊慢特病结算服务功能以来，总体运行平稳，无因流程设计不当而造成患者投诉情况的产生。图 2 展示了从 8 月 16 日跨省直结功能正式运行以来至 9 月 15 日运行满一个月的门诊结算情况，开通前半月慢特病跨省直接交易量显著升高，至 9 月逐渐趋于稳定。9 月 11 日至 15 日日均门诊量约 251 人次。直接结算患者医保垫付资金平均为 79.65%，不仅减轻患者的疾病负担，更减轻患者的经济负担。

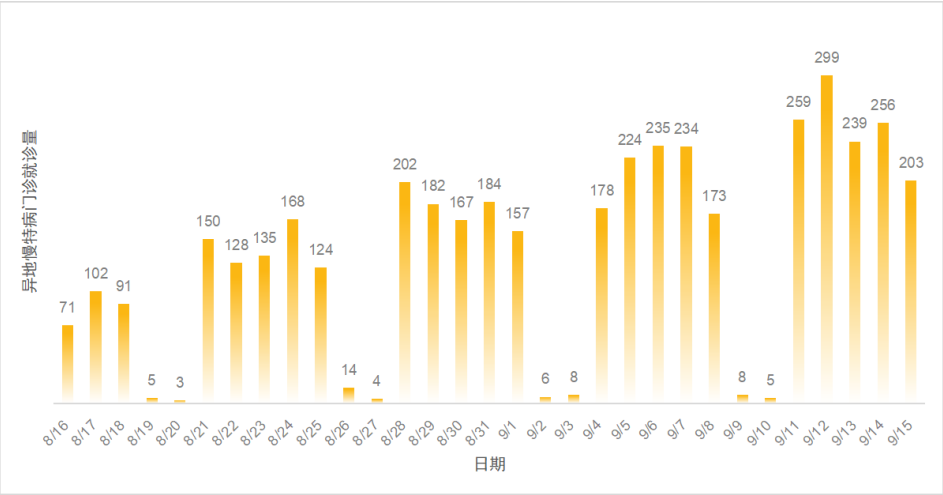


图 14 异地门诊慢特病结算人次分布情况

4 主要问题及解决方法

在开通该项服务一个月过程中，X 医院十分重视患者反映的问题，各部门及时汇总反馈并建立问题台账，不断完善结算流程。当前主要问题及现有解决方法如下。

4.1 获取患者门诊慢特病信息失败

在开通该项服务一个月过程中，通过进行舆情监测发现，患者门诊慢特病信息获取失败是造成门诊投诉的主要原因。X 院医保部门会第一时间会同信息中心分析失败原因，若因患者中途变更备案信息，参保地未及时更新，导致医院获取失败，则告知患者联系当地医保局进行确认；若因网络不稳定、传输速度慢等多种原因获取失败，且多次尝试后仍未获取，则向患者解释后，由患者决定是否按照普通医保直接结算或是全额结算后或当地手工报销。

4.2 患者门诊慢特病跨省直结政策咨询量大

运行一月以来，医保咨询窗口关于慢特病跨省直结政策咨询的比例明显增加。目前，X 院将就诊流程及注意事项整理并印发成宣传页，在窗口分发给患者，减少排队等候时间。若患者需解决的问题过于复杂，不能当天解决。会让患者自行填写“特殊事项申请表”，在填写一周内由医保工作人员负责联系，进行解答。若涉及其他手续需到院完成，则提前联系患者，在统一时间集中办理。

4.3 门诊慢特病退费流程复杂

很多参保地原则上不支持门诊慢特病隔笔退费，进而造成退费手续十分复杂、周期较长。X 院在微信推送中以温馨提示的形式让患者缴费时务必慎重，避免后续退费造成的等待时间过长、发票等报销材料丢失造成的不便。同时，为减少因检查检验项目预约等待时间过长而造成退费的情况，X 院门诊部积极配合，组织各临床医技科室为患者缩短预约时间。如果患者确需退费，X 院会向患者解释退费平均周期及可能造成的不便，由医保专业人员负责对接进行办理。

5 展望及建议

5.1 建议医保经办机构尽快实现慢特病隔笔退费

隔笔退费情况存在较为普遍，同时也是患者的正常需求，建议还不支持隔笔退费的统筹地区尽快更新政策要求，为更多患者提供便利。同时，由于各统筹地区对于是否支持隔笔退费的要求并不统一，医疗机构在解答患者问题，对不同地区患者难以采取统一管理措施，建议经办机构及时告知医院当地最新政策要求，医院通过短信提示等多种方式告知患者，进而达到“信息多跑路，患者少跑路”的目的。此外，建议医保相关部门完善慢特病跨省退费相关流程，减少不必要的环节和手续，节省医疗机构和参保人的等待时间。建议接诊医疗机构优化流程，对于一些预约周期长、退费频率高的项目实行先检查后缴费，减少退费交易。

5.2 建议国家医保平台、参保地和医疗机构加强信息建设

加强异地就医，特别是异地慢特病就医患者信息建设，需要国家医保平台、参保地和医疗机构

协同推进，建议国家医保平台尽快制定患者信息的对接方案，包括数据接口的规范、数据传输的方式等。统一与参保地和医疗机构的数据标准，确保数据对接的准确性和一致性。参保地和医疗机构应配合建立数据接口，在对接过程中需要注意患者信息的安全性，包括患者信息的加密和权限管理，以保护数据的隐私和安全性。参保地还需定期对患者备案数据进行维护和更新，确保患者的基本信息完整，慢特病身份准确。

5.3 建议尽快实现异地门诊慢特病线上诊疗及医保移动支付

门诊慢特病跨省直接结算、线上诊疗和医保移动支付都是医保领域的重要创新，将最新的科技和医保、医疗相结合，是政府智慧城市建设规划、国家医保深化便民服务的重要举措。当前常见病和慢性病患者线上复诊已日趋成熟，医保移动支付范围不断扩大，在国家政策的统一要求下，建议各医疗机构积极开展医院自由移动端线上诊疗及医保移动支付工作，不仅是服务异地门诊慢特病，更是全体人民群众，提高支付效率，提升服务质量。

6 结论

总的来说，异地门诊慢特病直接结算是医保领域的又一项便民惠民举措，能够减轻患者的就医负担、减少垫资压力、提高报销效率、促进医疗资源公平分配，是“医保公平”理念的体现。X 院采用的“收费员识别患者身份，医生判断就诊病种”新模式，是门诊慢特病跨省直结方式的一次创新性探索，主动解民难、排民忧、顺民意，在医保领域解决好服务群众的“最后一公里”问题，在方便参保人就医的同时，有效促进医院医保信息化建设，推动智慧医院的建设和发展，具有重要的社会意义和价值。

PU-279

基于 RFID 技术构建固定资产全生命周期管理体系实现全流程智能化管理

金迟
郑州市中心医院

一、案例背景

当前新医改进程不断深化以及党中央提出“过紧日子”要求，需要医院方面积极优化资源配置，提升资产利用率并降低医院运营成本，而优化固定资产管理方式是落实此要求的重要手段。

另外医院的固定资产数量多、分布广、差异大，使得资产的信息管理工作量较大，工作效率较低、管理成本较高，资产数据质量较差。郑州市中心医院针对以上问题通过完善固定资产管理制度、防范资产活动风险、监测资产运行情况等手段。构建了基于 RFID 技术构建郑州市中心医院统一的资产全生命周期管理体系，通过对医院的固定资产进行智能数据采集，优化固定资产配置，提高固定资产使用率，利用现有资源发挥最大效益，有助于医院提高自身医疗服务能力，从而有效提高公立医院的竞争力，最终为医院可持续发展提供强有力支撑。

表 1 医院资产管理现状 SWOT 分析矩阵模型

优势（Strength）	**劣势（Weakness）**
1. 拥有经验丰富的专业人员和资产管理团队，具备医疗设备和设施管理方面的专业知识。 2. 拥有较大的规模，配备了各种先进的医	1. 管理过程的复杂性：医疗设备和设施的种类繁多，科室和部门之间可能存在信息孤岛，这可能导致资产管理过程复杂、信息不准确以及数据丢失的问题。

疗设备 3. 配备了 RFID 技术和资产全生命周期管理平台包含扫码盘点、扫码维修等，有助于提高资产管理的效率和准确性	2. 维护成本高：医疗设备和设施的维护成本相对较高，包括设备维修、设备更新、设备保养等，这可能对资产管理预算造成压力
机会（Opportunity） 1. 自动化技术的应用：可以考虑利用物联网、RFID 等自动化技术来实现资产管理的智能化和自动化，提高管理效率和准确性。 2. 数据分析与优化：通过大数据分析和人工智能技术，可以深入了解资产使用情况、维护需求和优化策略，以提高资产利用率和运营效率。	**威胁（Threat）** 1. 技术变化：技术的快速发展可能导致当前使用的资产管理系统或设备过时，需要不断跟进和升级。 2. 资产丢失和损坏风险：由于医院环境的复杂性，资产丢失和损坏的风险较高。这可能导致资产损失和额外的维修成本。 3. 管理和监管要求：医疗设备和设施的管理和监管要求不断提高，包括法规、合规性和安全性等方面，这对资产管理提出了更高的要求。

二、主要做法

1. 建立资产“身份证”，摸清医院资产“家底”。

资产入库后，一个资产创建一张资产卡片，用于资产实物管理；资产卡片生成时候，可根据资产信息，通过 RFID 打印机打印写入资产卡片信息，同时将打印完成的 RFID 标签，贴于资产表面，用于后期可视化管理及 RFID 快速扫描盘点使用。

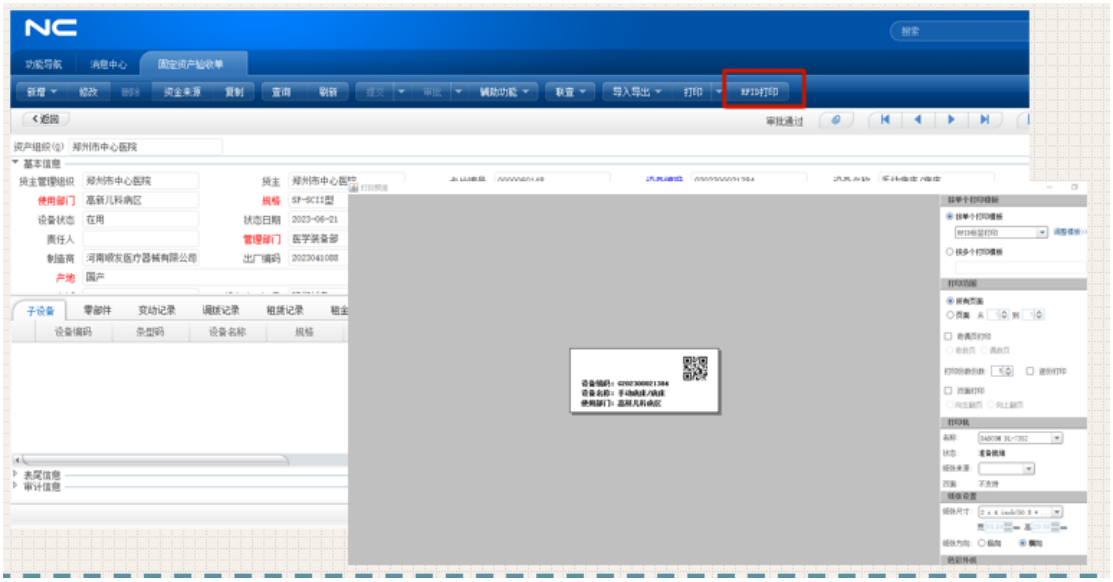


图 1 RFID 资产标签打印

2. 打破时间、空间限制，资产盘点一键生成。

使用 RFID 射频技术，在射程范围内自动感知并识别资产，无须近距离查找条码标签并扫码读数，可提高资产清查的效率及准确性。基于安卓或 IOS 操作系统的 APP 或者开发小程序功能，业务功能包括:固定资产清查、维修处理、PM 实施等。移动终端工作台支持扫描资产二维标签准确定位和读取设备资产的基本信息，要求支持通过 PDA 或智能手机方式实现相关业务的处理

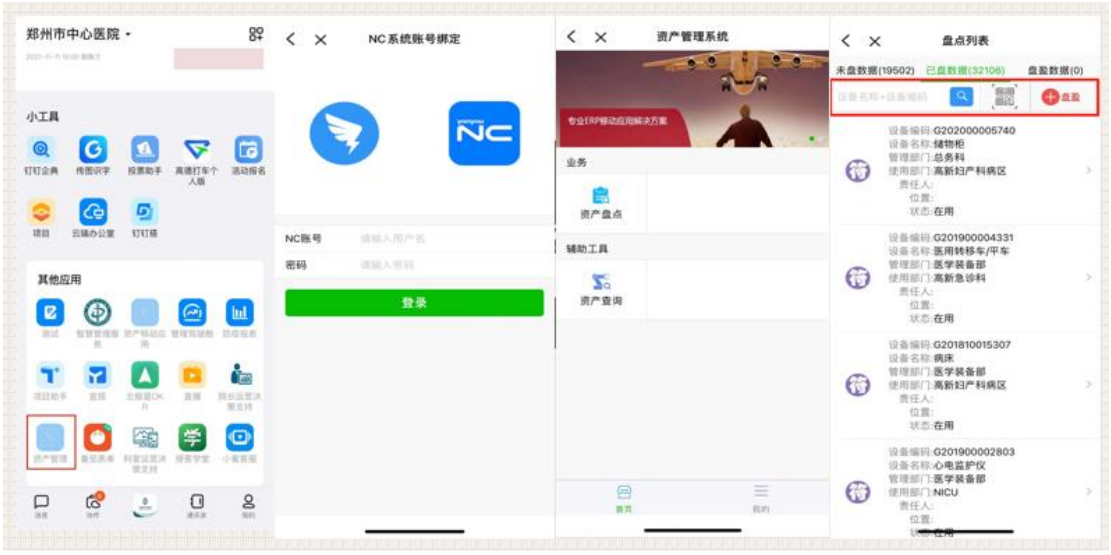


图 3 移动端资产管理小应用

3. 全流程线上管理，一体化平台管控。

构建资产的全生命管理流程，以设备资产为核心，实现设备购置、设备使用、设备变动、设备盘点、设备维修维护等全流程管理流程；配备掌上设备移动管理平台，集成到钉钉微应用中，科室管理员可以使用资产管理系统的账号登录钉钉小应用端扫描资产标签上的二维码实现设备移动报修、移动盘点、资产查询等移动管理功能；也可以随时在手机端查询科室内资产数据、实时掌握科室资产动态。

图 4 移动端资产管理系统

4. 重点设备智能监控，细化资产运营维度。

通过与其他业务系统对接，抓取设备运行收入数据、成本数据，建立大型设备运行绩效指标，方便查看大型设备的成本和收入等数据；制作设备运营智能分析可视化图表，进而指导医院设备精细化运营。

图 5 大型设备经济效益月度分析报表

图 6 设备运营分析可视化图表

5、梳理资产管理流程，建立健全资产管理制度。

我院对固定资产实行三级管理。一级管理部门：资产管理委员会，围绕医院质量与安全管理委员会年度工作计划制定资产工作计划,统筹规划全院的资产管理工作，并监督执行；对资产构建、处置、清查等重大经济事项进行综合性管理和监督。指导开展全院资产的清产核资（每年组织一次全院的清查核资工作）、统计报表、总结、考核、评估、评优、培训等工作。二级管理部门：财务科、医学装备部、总务科、信息科、基建科、车管科。归口管理科室负责固定资产的采购计划编制、采购执行、验收组织，做好固定资产的初始分类，建立固定资产档案、设备台账；监督配合各科室做好设备的使用和维护，定期保养、检修和维护固定资产，建立设备技术档案和维修记录，确保设备完好提高利用率；固定资产转移、报废等实物动态管理及审核，对各部门的固定资产管理进

行指导和监督。三级管理部门：固定资产的使用科室，负责固定资产的管理工作，指定固定资产管理员专人管理，承担固定资产的日常管理工作并对所管资产的安全完整负责；按照相关要求做好各项培训，严格执行技术操作规程和维护保养制度，确保设备的完好、清洁和安全使用；定期自查固定资产，资产发生盘盈或盘亏要及时查明原因；验收资产时，核实资产的名称、型号是否和实物相符；资产转移、报废时，及时办理手续，做到账、卡、物三相符。

三、实施效果

1. 以资产主数据为基础，实现资产全生命周期管理。

全院资产统一建账、统一调配，优化资产布局，掌控全局资产状态，盘活闲置资产，合理配置资产，实现医院内资产资源平衡，降低了因资产及备件采购带来的资金占用；规范资产管理领域的业务流程，有效地建立资产全生命周期管理体系；

2、以设备设施运维为基础，防控资产运行风险

规范、标准化现场点巡检及检维修作业机制，确保设备异常、隐患可以提前获知，并匹配合理的检修方式，实现异常或隐患的及时消缺，强化设备管理，提高设备利用率，降低因设备发生故障而给医院运营带来的风险，保障产品检修计划落实及设备运行安全性，控制检修过程中的人力、物力消耗，提高运营管理能力。

3、实时感知资产运行情况，深度挖掘和运用固定资产的数据价值，

在医院资产管理系统中，根据资产重要程度建立设备资产的差异化管控模式，对设备现场运行情况进行全面监控，对重点、核心设备实现精细化的测量项目定义及监控，在降低维修维护成本的基础上，控制管理成本的上升，从而可以实现资产管理综合成本的最优化。

4、科技赋能，提速增效，降低资产管理成本。

通过 RFID 技术应用，将传统手工盘点方式或接触式扫描盘点方式转变为非接触式盘点，简化盘点流程，提高盘点效率，大幅降低资产盘点和查找、复盘工作量，大大提升医院盘点管理效率。实现医院固定资产管理工作的实时、动态、精准的精细化管理模式。单件资产盘点时间从 1-2 分钟缩短到 10s 内，效率提升 80%，准确率 100%。同时解决了账实不符、账账不符、系统缺失、协同不畅、效率低下、生命周期管理难、数据价值难以挖掘等问题，实现资产保值增值，进一步降低成本、提质增效，助力医院高质量发展。

四、主要经验

（一）完善固定资产管理制度。

公立医院在实际发展过程中应充分认识到完善的管理制度是推动固定资产管理成效提升的重要基础保障。为此，在实际工作中应注意积极依托于现行制度文件以及自身实际情况，及时优化、革新医院资产管理相关制度，并确保其落到实处。通过合理分工以及责任划分，切实保障医院固定资产管理各项工作高效高质有序开展。

（二）加强固定资产日常管理。

公立医院要建立固定资产的绩效评估制度，并以此为依据，制定科学、合理的评估指标，并将其纳入评估系统，对各部门资产运行状况进行全面的分析和评估； 同时，将评估结果与各科室的业绩相结合，确保医疗器械的正常使用，避免盲目的招标，从而减少医院的经营费用。同时，要加强对现有医疗器械的成本和效益的分析，提高医疗器械利用率，增加医院的经济效益与社会效益。

（三）加快信息化管理系统建设。

构建资产管理系统，可以动态反应设备存量、使用、存放等情况，确保各部门在实际工作过程中可以及时掌握资产变化信息，同时为大型医疗设备效益分析提供必要数据支持。

结语

固定资产是公立医院经营活动的基础，是其最重要的资产。从内部经营视角分析，强化体制建设，提升管理水平。从实际发展角度进行分析，固定资产直接关系到公立医院各项业务开展，是公立医院生存发展的重要物质基础。通过积极优化医院固定资产管理体系，可以切实保障医院发展过

程中各项物资充足，相关医疗设备完善，在完善的制度体系支持下，医院整体管理水平也会随之大幅提升。从目前来看，RFID 技术在医疗设备管理方面起到很大作用。但是随着精细化管理要求的提升，业务管理电子化，流程化已不能满足运营管理要求，还是有需要进一步优化改进的工作：

（一）、持续完善的固定资产管理制度，积极推动资产增置、资产报废等工作流程规范性。

（二）、充分利用现代信息技术优势，为推动公立医院固定资产管理水平提升提供有力保障，实现医院高质量发展。

PU-280

DRG 支付模式下医院精细化管理的思路与愿景

张远林 蔡文君 *卢熙 李俊锋 黄茜
湖南省人民医院

随着人们对医疗健康的重视，对医疗服务的需求也越来越高。传统的按项目付费的医保支付方式的弊端日益突显，医保基金支付过多，不合理使用造成资源浪费。个人和国家层面都期盼能用更少的钱得到更优的服务，在这一现实需求下，国家出台了一系列政策来推进医保改革。2021 年 11 月，国家医保局发布了《DRG/DIP 支付方式改革三年行动计划》，按照文件要求，到 2025 年底，DRG/DIP 支付方式覆盖所有符合条件的开展住院服务的医疗机构[1]。

疾病诊断相关分组(Diagnosis Related Groups,DRGs)付费和基于大数据的病种的 DIP(Big Data Diagnosis-Intervention Packet,DIP)付费是一种比较科学、精细化的医保支付方式，通过大数据分析，科学的决策，对各类疾病诊疗进行分组/分值，医生在给患者治疗之前就可以了解该病种可报销的费用。这种精细化的支付方式能有效的提高医保基金的使用效率、提高医疗效率和医疗质量、降低老百姓看病负担等[2]。从传统的按项目付费向病组/分值付费方式的转变，对医院管理的思维方式、行为方式以及管理方法等方面都会产生深远的影响。医保改革关乎医院的切身利益，只有从原来的粗放式管理向精细化管理转变[3-5]、才能实现医院的长足发展。

本文结合医院精细化管理的实际过程，从管理制度、业务流程和借助基于 DRG 管理系统的管理方法上为其他医院有效应对 DRG 支付改革、推动医院的高质量发展提供借鉴思路和方法。

1 DRG 对医院的影响

1.1 对思维方式的影响

传统的支付方式考虑的是从收入规模上进行扩张，多项目会造成不必要的资源浪费。而在 DRG 付费方式下，考虑的则是成本、效益、效果三者的最优，尽可能减少成本的支出，达到同样的或更好的治疗效果。

1.2 对行为方式的影响

思维方式的转变将引导行为方式的改变，一方面医院管理层需要根据医改制定与之相适应的管理制度，包括绩效考核制度和相应的奖惩制度等，另一方面需要引导医生个人规范自己的诊疗行为。

1.3 对管理方法的影响

传统的付费方式，其管理相对简单，而在新医改政策下，医院需要进一步优化流程，进行系统改造，并借助基于 DRG 的信息平台提供运营管理数据，以帮助医院进行决策分析和学科建设发展。

1.4 对管理水平的影响

随着医改政策的落实，医院的管理水平向高质量方向发展。在强调“成本结余”的 DRG 支付模式下，一方面通过发展专科诊疗特色、主控制运营成本、提高医疗服务能力、提升医疗服务效率

等提升医院的核心竞争力，另一方面对内部管理体系进行重塑，定期监管病案首页质量和相关数据，提升医院的病案质量，保障数据精准规范。一

2 医院 DRG 支付模式下精细化管理思路

在 DRG 支付方式改革下，医院应围绕“医保支付方式改革”与“医院高质量发展”两条主线，建立医院多部门联动与管理机制，以信息技术[6]为支撑，按照“提升数据质量、建立 CHS-DRG 成本核算机制、建立医保管理新能力、建立考核评价体系、建立薪酬绩效分配机制、建立学科发展规划”六步走的应对方法来重塑医院的业务和流程，积极应对支付改革给医院带来的变化，实现医院的高质量发展。本院从医院管理制度、业务流程和借助基于 DRG 的管理方法上共同来推进管理实践。

2.1 制定相应的管理制度

DRG 付费改革其实就是按医疗服务项目支付方式向预付制方式的一种改变，是一种新型的付费方式。要实现成本和效率的统一，需要医院多部门共同努力。医院成立了以分管医疗保险部的副院长为组长的项目组[7]，项目成员包括：医务部、护理部、医疗保险部、财务部、病案室、质控办及信息中心等多部门组成。项目组明确了各小组成员的职责并在各职责范围内制定相应的管理制度和方法。实行医院、科室两级主要领导责任制、管理责任制，医院行政科室、临床医疗单元、护理单元、财务、信息等多个职能部门的协同管理。如医务部、病案科负责对临床医生进行疾病诊断和手术操作规范等专题培训；医疗保险部负责对 DRG 付费原理、实施办法等进行培训，并负责医保的监督工作；财务部制定相应的绩效考核办法；信息中心建立相配套的 DRG 监管考核平台[8]，与省医保平台对接，结合临床医技需求，优化业务流程。

2.2 流程优化

构建以病人和医院管理为中心的服务型文化，把 DRG 的服务效率指标体现为优化医疗服务、优化医院业务流程。比如通过放射、B 超、心电等医技检查预约等流程改造，一方面能大大节省病人的检查和诊治时间，这样医生能及时地对患者病情做出处理，缩短病人的住院天数，加快床位的周转，提高效率，另一方面通过检查预约的电子化流程，病人无论是从线上还是线下，都无需医生再开具纸质检查申请单，既方便了患者就医，又大大节省了纸张的支出，也为医院节省了耗材成本，从根本上节省了病人和医院的成本。

2.3 建立基于 DRG 的管理系统

在深化医改的大环境下，建立基于 DRG 系统的是医院精细化管理的关键。医院采取事前、事中和事后的全方位管理，如下图 1 所示。事前做好对临床科室的 DRG 相关培训，提高疾病分组的准确性。事中通过系统数据、规则和基于医院历史数据挖掘的诊疗路径实现对医疗行为的监控，对违规行为进行识别，并对费用情况进行预警。事后以科室为单元，对每月、每季度、每年的病组进行盈亏额、盈亏原因等方面进行总结分析，制定对应的措施或计划，并通过激励约束机制来规范医生行为，提升效能。在设计考核体系时，不仅要考虑综合收入、成本、风险因素等，还要考虑医疗服务、医疗效率和医疗安全[9]三方面的各项指标。

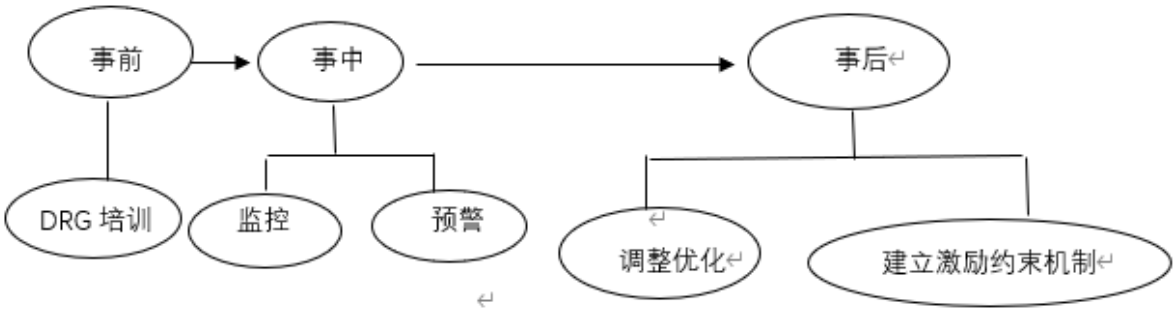


图 1 管理过程

2.3.1 加强数据质量管理

数据质量是影响 DRG 实施的首要因素，疾病诊断和手术操作编码是否有缺失，或填选是否正确，将直接影响医保付费情况，因此病案管理部门、医保部门和临床医生都要提高认识，一方面，加强对医院编码员和临床医生编码培训。另一方面，建立病案首页、医保结算清单、医保智能审核规则、物价收费规则及医疗临床用药规则为一体的知识库体系,系统内置规则对上传的病案首页数据和疾病编码进行校验和智能判断，定位临床病案书写及编码中存在的问题，对不符合上传要求的返回临床医生进行修改，并对临床医生进行相关的培训、指导。

2.3.2 对运营数据的监控

通过系统知识库的医保限制规则、物价规则及医疗临床用药规则、临床诊断依据知识库，对数据进行智能审核，实时发现、精准定位存在的疑似医疗违规行为，例如药品超限制使用、超年龄、与所属病组治疗不符、诊断与医疗过程和病历文书不符等违规行为。

2.3.3 数据分析预警

围绕医疗服务能力、医疗服务效率、医疗安全、费用控制四大数据指标，采用象限分析、趋势分析、奇异点分析等方法，给出分析建议。从全院、科室、医生三个维度分析。医院层面，对全院的病组、预警病例占比、费用结构分析，以及平均住院日、出院带药、高低倍率病例占比等；科室层面，对科室内病组、结算费用、费用结构分析及 DRG 工作评价等[10]。医生层面，综合病人诊断、手术、住院天数、病情等因素，通过 DRG 分组器，测算 DRG 入组情况，对标预算标准，全程监控诊疗过程中的费用超支、结余情况，及时预警提醒。

2.3.4 基于病种分组的成本分析与监控

将国家医疗保障疾病诊断相关分组（CHS-DRG）中的核心疾病诊断相关分组（ADRG）的疾病诊断和手术操作亚目再进行分类，形成比 DRGs 更细的细分组；再基于医院的历史数据，挖掘每个 DRGs 细分组的费用成本及费用结构组成，形成每个 DRGs 细分组历史参考费用成本。对每一位在院患者进行预入组后，对其费用消耗的进度情况与历史参考费用成本进行对比，实时监控每一位在院患者的费用消耗情况，自动分析出总额偏离较大、费用消耗进度异常的原因，实时提醒临床医生进行医疗行为过程的调整，在保证医疗质量的前提下，全面控制疾病治疗的费用成本。

2.4 调整优化

针对各项指标和 DRG 盈亏进行分析，定期开展科内、医疗组内的分析会，分别从科室和医生两方面，通过从成本构成及临床路径上进行分析，以问题为导向精准发现问题、解决问题，形成统一的管理流程，提升决策效率，不断总结修正，促进诊疗过程优化，形成良性循环，促进医保基金的合理使用。

2.5 建立激励和约束机制

医院采用了最新的以资源消耗为基础的相对价值比率(RBRVS)绩效分配模式[11]，将医院的收费项目进行分类，将其划分为手术、检查和检验等若干类型，并为每类确定一个绩效费率。其中手术被细分为四级手术，每级手术赋予不同的绩效费率，反映不同手术的难度级别。护理则用护理时数来解决不同病区护理工作的相对价值。同时，根据医、技、护岗位的不同,设置不同的考核标准。根据 DIP 的精细化指标 CMI（即病例组合指数，用于评价收治病例的平均技术难度，是医院服务能力和水平的体现），结合 DRG 绩效评估体系，制定基于 CMI、RW（Right Weight）值提高情况的新绩效分配方案和科室绩效二次分配方案[12]，主动提高违规行为的辨识度，重点针对诊疗规范性、分解住院、合理用药[13]等方面落实管理制度。一是将医保违规行为及质量安全纳入绩效考核，督促医生做到合理检查、合理用药、合理治疗、安全诊疗。二是与医院的评先评优、进职称等相结合，做的好的加分，达不到标准的扣分。

2.6 加强学科建设管理

通过对全院病组进行盈亏分析，了解各学科的专病专治程度，技术诊疗水平以及医疗服务效率，使各学科开展更符合自身成本和技术优势的项目[14]。运用波士顿矩阵，对盈利类病组，重点分析其价值效益，进一步将学科做大做强，形成核心竞争力。对于亏损类病组，从诊疗、护理、手

术、耗材等成本构成上进行分类分析，讨论可改进优化的措施，来提高诊疗科室的服务水平和医保基金的使用效率。科室可通过院内学习交流讨论方式来调整优化诊疗方法，或向外院有开展更优的诊疗技术的单位进行研修学习，进一步优化，来提升自身的服务能力水平。

3 建设成效与分析

医院在 2020 年 10 月份开始上线 DRG 项目，以 2021 年全年的数据与 2020 年全年的数据进行了统计分析，在出院患者增加 9.14%的情况下，医疗总费用下降了 24881.2 万元，同比下降 7.16%，病人的平均住院天数同比下降 8.47%，住院均次费用下降 6.03%，患者大型检查等待时间平均下降 21%。通过对服务能力、医疗效率、医疗安全维度的比较，发现应用系统上线后，出院病例、四级手术、平均住院日等都有了明显改进，患者的医疗费用、次均费用、药占比以及耗占比都有不同程度的下降或者降低，详见下表 1 服务能力表、表 2 医疗效率表。

DRG 医保支付项目的上线，医院通过数据的分析和对比，不断优化就诊流程，进行一系列措施的改进与实施，使临床科室、医技科室的工作人员都有了成本效益意识，服务意识变被动为主动，医疗服务能力得到了促进和提升，特别是大型检查（如核磁），由原来的平均等待 4--5 天才能做，降为 3 天内出结果，得到了病患的认同，切实减轻了老百姓就医的经济负担。

表 1 服务能力

服务能力	本期(系统应用后)	同期（系统应用前）	同比
出院病例数	146689	134399	9.14%↑
四级手术人数	11447	9612	19.09%↑
C、D 型病例人数	97369	87430	11.37%↑

表 2 医疗效率

医疗效率	本期(系统应用后)	同期（系统应用前）	同比
平均住院日（天）	8.43	9.21	8.47%↓
次均费用（元）	21319.34	22689.11	6.03%↓
药占比	26.71%	31.53%	4.82%↓
耗占比(分母不包括药品收入)	34.85%	41.23%	6.42%↓

4 结语

目前，部分省份的医疗机构已经进行了 DRG 医保支付的试点，在试点过程中，各定点医疗机构也遇到了一些问题，比如暂时无法涵盖多种合并病种的疾病的病组，由此给患者和医疗机构都会带来一系列影响。总之，改革的路上不会一蹴而就，需要各医疗机构跟医保管理部门保持沟通，及时反馈实际过程中遇到的问题。医保管理部门充分调研并结合医疗机构在实际运行中的数据、效果，进行动态调整，进一步完善相关政策和方法手段，做好考核与激励机制，使医保改革更加科学合理。在深化医改的大环境下，其本质目标是通过支付手段的激励机制引导医疗行为意识的优化升级，实现合理诊疗，提高医保基金的使用效率。坚持从病种的成本控制和诊疗能力两个维度出发，实施成本控制[15]、充分利用信息化手段，运行监管、考核评价、绩效分配等工作，促使临床专家主动挖掘成本效益更高的治疗手段，使群众能享受到价廉质优的医疗服务。同时引导医疗机构优化业务流程和资源配置，促使医院主动提高医疗服务水平和患者的满意度，从而实现医保管理部门、医疗定点机构和患者三方共赢。

PU-281

基于微信公众号平台的公立医院病案复印申请与邮寄流程创新研究——以上海市 L 医院为例

侯娜娜[1]，沈亮[2]，刘清海[2]
1.复旦大学附属眼耳鼻喉科医院
2.上海中医药大学附属龙华医院

近年来，随着我国医疗信息化的相关政策相继出台，国家大力支持各实体医院发展互联网医疗服务，基于信息化的医院服务流程创新研究涉及医院方方面面。目前来看，门诊、住院或医技检查流程这方面的研究发展迅速，研究成果颇为丰富[1]。国内现有的线下病案复印场景已经在信息化方面有了很大的提升[2]，包括病案翻拍打印、病案流转位置示踪[3-4]。然而，大部分医院未能实现基于信息化的病案复印流程再造，更有部分医院仍未意识到其重要性。目前大部分医院的病案复印仍然存在人员聚集、排队等待时间较长、病人或委托人往返医院、复印工作效率低、复印工作量无法准确统计等情况，传统的出院病案现场复印流程已无法满足医院的发展和患者的需求，各医院病案科及医院发展受到越来越严峻的考验[5]。随着国家大力推进深化医疗服务改革，提倡各个医疗积极加强信息引导，普及微信和小程序等便民应用，越来越多的医疗相关服务提至线上[5]。从门诊预约到检验报告和各类票据的查阅，无不展现微信公众号所体现的“互联网+医疗”的时代号召[6]。基于上述现状，有必要将病案复印的部分工作流程移至线上，病案复印工作流程的再造迎来新的契机。

上海市 L 医院作为当地一家大型三级综合公立医院，连续三年住院人次超过 11 万人次，其中外地来源病人超过 40%，本市非本区来源病人超过 50%，仅 10%不到的病人来源于 L 医院所在区。毫无疑问，非本区来源的病人到 L 医院进行病案复印的现场申请，将严重影响 L 医院的保卫、门诊、车辆等各方面管理，而且造成这部分申请者带来时间与金钱上的相对付出较多。据统计，L 医院病案复印工作时间日均接待人次超过 70 人次，平均一小时接待超过 10 人次。面对如此繁忙的现场病案复印接待工作，L 医院传统病案复印流程方面的问题尤为突出。如何基于互联网创新复印流程，提高工作效率，降低人力与服务成本，成为该院一项亟需解决的技术与管理难题。

1 上海市 L 医院复印流程现状及流程创新的必要性分析

1.1 医院现有复印流程调查

本研究采用问卷调查法和专家咨询法对上海 L 医院的复印流程进行现场调查，具体了解该院目前病案复印流程的现状，系统分析病案复印流程中存在的问题，为流程创新提供依据。

1.1.1 病案复印满意度问卷设计

基于客户满意度理论，参照国家三级医院的评审标准，设计病案复印满意度调查问卷。问卷分为四个维度，包括复印环境、复印流程、服务质量、整体评价。四个维度下包括若干题项，采用李克特五分量表法，从 5 分的“非常满意”到 1 分的“非常不满意”等五个分数值供调查对象选择，测量分数越高，表明被调查对象病案复印服务满意度越高；相反，如果获得分数越低，则表示被调查者满意度越低。分数 3 分以上的题项，本研究定义为对该题项满意。分数 3 分以下的题项，本研究定义为对该题项不满意。经问卷预测试后，收集问卷预测试的数据用 SPSS 23.0 统计软件进行信度分析，以此判断该量表的准确性和可靠性。经分析，α 系数为 0.728，0.8>α 系数≥0.7，说明信度较好，可用于后续研究。

1.1.2 病案复印满意度问卷调查

本次问卷以病案复印流程为主线，调查上市 L 医院住院患者及委托人的满意度。随机选取 L 医院病案科登记的 2019 年 5 月 20 到 2020 年 5 月 20 的来院复印过病案的住院患者或家属共 500 人进行随访，并通过问卷进行复印满意度调查，所有问卷通过微信发放填写。截止问卷填写结束日期，共收得问卷总计 458 份，回收率达到 91.6%，样本数量满足 80%以上的抽样方案和标准。根据设定好的是否满意分值标准，得到满意和不满意的总人数，得然后根据统计学分析方法中的百分比构成，分析 L 医院各项指标的患者满意度，结果如表 1-1 所示。

从表 1-1 可以看出，现场复印人员对 L 医院的现场复印的环境和复印流程都表现出不高的满意度除了对复印审核要求满意度为 83.2%以外，其余项目的满意度均未达到 50.0%以上；现场复印

人员对该院病案复印时的服务质量满意度较高，对复印接待人员的复印态度和仔细程度较认可，达到 82.1%和 84.7%。而对于现场复印接待时间的满意程度则较低，这可能和复印需求人员需要到场复印的时间和其工作生活时间有冲突。患者满意度调查问卷结果提示 L 医院的复印流程较为繁琐，如要提高患者的复印满意度，减少排队次数、缩短无效等候时间、减少来院次数是关键。

1.2 专家咨询

就 L 医院当前病案复印流程存在的问题召开专家会议，收集其对 L 医院病案复印流程存在问题的意见和建议。专家人选共 28 人，包括 L 医院的院领导以及医院病案管理委员会、临床科主任、护士长、参与流程的病案工作人员等。经汇总整理，各专家对 L 医院复印程存在的问题达成了以下四点主要共识：第一，复印流程繁琐。患者在 L 医院复印过程中，需要经过来院—排队—问需—举证—审核—查询—复印—交费—离院等多个环节，患者需要多次问询确认各个环节，大大增加了患者及其委托人的复印时间成本[7]。第二，信息化程度不高。L 医院的复印登记流程依然较为落后，复印信息输入纯手工，对于患者复印的材料的分类与数量、各科室的复印需求等难以做到有效追踪，失去了很多具有管理学价值的统计数据。第三，人力资源配置不合理。 根据 L 医院出院病历 48 小时回收率可知，医护人员工作强度大，经常无法按时完成病历的书写与运行功能，大大增加了复印工作人员的无效工作时间和患者的无效往返时间[7]。第四，医患矛盾突出。信息系统不完善降低工作效率，且长时间处于高强度的复印工作压力之下，沟通时间较短，医患矛盾时有发生。

1.3 基于微信公众号平台的流程创新的必要性分析

结合 L 医院的总体情况、病案复印满意度问卷调查结果和专家咨询的汇总资料，可以得出目前病案复印流程繁琐是 L 医院复印流程中存在的主要问题，而这些主要问题都体现在复印流程的始末。患者往返不便、排队等候混乱、身份材料不全、付款方式单一、复印时间不符、复印环境较差等，都影响着复印流程的顺利进行[8]。将这些问题承载于互联网的浪潮，分流于信息技术的生产线上，成为解决这一系列难题的良方。

L 医院基于信息化进行流程创新，依托微信公众平台进行病案线上复印与邮寄，不仅可以优化病案复印工作流程，还可实现病案复印的便捷应用、高效服务，提升患者就医体验及满意度，改进医院环境秩序，完善医院一站式医疗服务。因此，基于微信公众号平台的流程创新势在必行。

2 基于微信平台的公立医院病案复印申请与邮寄流程创新方案

2.1 方案的设计原则 1

2.1.1 以患者为中心

国家卫生健康委员会高度重视公立医院的医疗改革，随着医疗改革的深入开展，特别强调要以患者为中心，改善患者的就医体验。 L 医院目前复印流程的设计未完全以患者为中心，进行复印流程的优化应该围绕患者的需求，充分运用信息化手段重新设计科学合理、高效便捷的复印流程，减少复印过程中的非增值环节，提升复印服务质量，提高患者或者其委托人的复印体验和满意度，方便病案复印的同时也能提高病案科的工作效率，减少各方面矛盾的产生。

2.1.2 以信息技术为主要手段

信息化的完善和互联网、移动互联网的应用在流程的优化和创新中起到至关重要的作用。L 医院在病案复印流程创新的过程中应结合自身实际情况，对病案信息系统做出整体规划，选择可靠稳定、能提高良好售后服务的信息产品服务企业，在保证数据安全的情况下，通过建立由点及面的管理平台，为病案复印疗流程的创新保驾护航；同时结合互联网及移动互联网的应用，实现复印流程智能化，为患者提供更高效、便捷的病案复印服务，降低患者的复印成本，提高科室及医院的运营效率。

2.2 基于微信平台的公立医院病案复印与邮寄流程方案设计

本研究所指的的 L 医院的复印流程是指由该院病案科提供的，以患者为服务对象并与其发生互动的病案复印相关服务过程，包括病案复印前的身份审核、病案复印内容、复印费用、复印后的确认环节等。进行基于微信平台的公立医院病案复印与邮寄流程创新，需要完善医院的信息化服务平台和微信公众号服务模块[9]。

2.2.1 系统流程设计

基于微信公众号平台的病案复印邮寄服务是指用户通过微信公众号预约申请与付费即可实现病案的线上预约与邮寄服务。主要包括复印申请、线上支付、相关提醒和后台业务管理等全流程。全

面提升用户体验，提高病案复印服务品质。根据病案复印预约流程，将整个系统分为线上和线下两部分，实现线上申请和线下审核与复印，客户端提示申请进度，申请管理端提示复印申请处理进度，方便快捷无接触，满足患者需求。流程如图 1 所示。

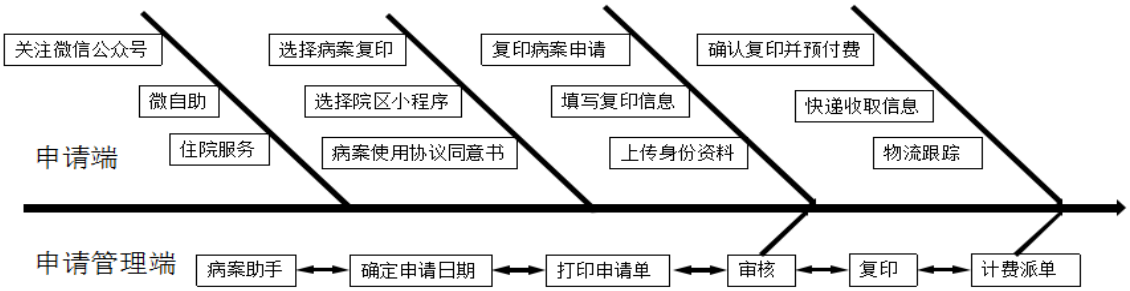


图 1 病案复印申请与邮寄流程

2.2.2 申请端线上流程设计

申请端主要通过关注上海市 L 医院微信公众号方式向患者提供复印申请、进度查询以及一些常见问题查询功能。为了保障患者能顺利拿到复印材料，在结合医院病历完成与回收率的情况，设置 2 周为住院患者出院后申请病案复印材料的期限，按提示上传证件及其他材料便可成功申请，真正做到以患者为中心的原则。病案复印小程序分为两个模块，复印病案申请和我的复印病案。在结合现场复印所需证件、用途及其费用的基础上，设定患者为住院病人，所需用途为报销凭证、就诊、转诊、工伤鉴定及其他。复印内容则设定为根据《医疗事故处理条例》中规定的允许复印的客观病历部分，如需复印主观病历则须持有关证件与材料到医院现场复印。复印内容按照 L 医院病案的保存分类，供患者或者委托人选择。再根据是否为患者本人申请复印，提示申请人上传所需证件等照片。而后填写快递信息，并预付一定的复印费，多退少补。在病案复印工作人员审核申请之前，申请人可自行撤销。如过审核不通过，申请订单将会退回至申请者，并显示注明的退回原因。申请者可根据注明的原因，更改信息后重新提交申请。微信申请端可显示申请订单的受理情况，实时掌握申请进度与物流查询。

2.2.3 申请管理端线上流程设计

申请管理端分为两大功能模块：申请管理和查询管理，满足线上复印所需预约申请审核、一键派件、订单查询、新增预约、预约设置及工作量统计等化需求。申请管理端为网页形式，可于电脑与手机上操作，随时跟踪患者申请情况。具体流程为，复印人员在受理复印申请时，打印当天提交的申请，挨个审核申请者是否提供了完整的身份资料，所需复印的病案是否在库，复印内容是否符合该病人出院的具体病案内容等。对于一些不符合复印申请要求的订单，申请审核为拒绝的，必须注明拒绝原因，避免患者产生误会。有些涉及到纠纷、死亡等病案，则需注明告知患者或者委托人到现场复印。采用微信线上申请流程后，病案工作人员可于申请管理后台查询期间申请单所涉及的各项可量化数据，包括复印量、用途及其具体复印内容等。

3 病案复印流程创新的保障措施

基于微信公众号平台的 L 医院病案复印申请与邮寄流程创新方案，涉及到医院的每个临床科室和患者利益，是病案复印的创新，更是搭载“互联网+医疗”健康快车的新方向。这是一个不断完善的过程，由医院领导层牵头，组建流程创新方案工作小组，协调各部门积极配合，动员全体医务人员积极完善病历书写周期全过程管理。通过调查和研究，对 L 医院基于微信公众号平台的 L 医院病案复印与邮寄流程创新提出以下四点保障措施。第一，提升“互联网+医疗”背景下病案复印流程创新的认知。第二，协调多部门，促进流程落地。第三，充足的资金投入和人员配比。第四，加强医务人员对该流程的的应用宣教，同时收集医护人员的意见和建议，综合运用 PDCA 理论完成运行质控与反馈[10]，了解在实际工作中创新方案流程的不足，持续改进完善。

4 结论

本研究以上海市 L 医院为例，通过问卷调查法收集患者现场复印满意度相关数据和专家咨询结果，梳理医院现行现场复印流程中存在的主要问题，分析进行流程创新的必要性，提出基于微信公

众号平台的 L 医院病案复印申请与邮寄流程创新方案。该创新方案以患者为中心，强调数据安全的大前提下[11]，创造智慧医院延伸段，不仅优化该院的病案管理工作流程，提高病案科以及医院整体运营效率、服务质量以及工作绩效，改善患者的就诊服务体验，降低医疗成本，实现医疗资源的有效配置，具有较高的实践价值，符合科室与医院长远发展的利益，同时为公立医院基于“互联网+医疗”进行各方面流程的创新提供借鉴。

PU-282

智能医疗背景下医患交流的类型、特征与优化策略

范美玉 赖昕
广州中医药大学第二附属医院/广东省中医院

1 引言

伴随着云计算、大数据、5G、移动互联网、人工智能、数字孪生等数字化技术的飞速发展，人们的生活越来越智能。与人体生命健康息息相关的医疗也不例外，“指尖医疗”“云端医疗”“智慧医院”等智能医疗新形态不断重塑诊疗流程，提高了医院的服务效率和运营水平，改善了患者的就医体验，也推动着医患关系由传统的“以医生为中心”的交流模式走向新的“以患者为中心”的交流模式[1]。本文从医患智能交流的类型、特征来剖析智能医疗是如何影响医患交流，继而针对医患智能交流中的难点痛点，给出对应的优化策略，以供参考。

2 医患智能交流的类型

2.1 按服务过程

（1）智能的问诊与导诊

问诊导诊是医患交往的最初环节。传统的问诊导诊由人工来完成，增加了患者排队时间，也容易出现挂错号而致使患者在该环节反复的现象，带来不好的诊前体验。智能问诊导诊相关工具，如人工智能导诊移动应用、智能导诊服务助理、智能导诊小程序、导诊机器人，不仅能帮助到院就医患者梳理活动路线，避免因反复流动产生无序，简化和畅通原本复杂的流程，还能够让医生精准问诊，也能够让患者自查疾病知识、健康知识和卫生知识，让问诊导诊体验更加高效便捷，也让患者能够有更加多样化的就医选择和主动权。

（2）智能就诊

在人们就医的历史上，“三长一短”问题曾长期存在，即：挂号排队长，就诊等候时间长，取药时间长，但看病却时间短。这往往不利于和谐医患关系的构建。在数字技术引入医疗场景后，智慧就医逐渐向着“就诊全流程全覆盖，线上线下就诊一体化”的方向努力，“让信息多跑路，让群众少跑腿”。常见的是“指尖看病”，我国多家医院在微信小程序上实现了“码”上就医，借助智能手机和微信应用，线上完成各类患者（门统、门慢、门规）预约挂号、缴费（甚至还能实现医保结算报销）、取药等流程，减少了排队的次数，信息呈现透明和更新及时。此外，“云诊室”“互联网医院”也可以满足就医不便的患者足不出户看病的需求，在线问诊、就诊，通过在线的医患交流和上传检查检验报告，就可以享受到不限地域不限医院的医疗资源，得到医生的诊断结果、治疗建议和用药指导，甚至在线购买药品。

（3）智能随访

随访能够延续患者服务、支撑满意度调查、辅助慢病管理、加强专病质量控制，是医疗活动的重要组成部分和医患沟通的重要桥梁。传统的随访大多是由人工完成，一般工作量非常大、流程繁琐，容易致使形式化的随访出现，也常常给患者带来厌恶感。智能化随访结合互联网、人工智能等数字化技术，基于专业的随访知识库，以微信小程序、公众号等形式呈现给患者，提供以医患沟通、院前、院中、院后随访和健康宣教为核心的服务，实现医疗机构和患者保持长期联系。人工智能随访助手不仅能够帮助医院对大批量门诊、出院患者实现精准的自动化随访管理和自动生成随访报告，大幅节省医务人员的精力，还可以完成职工、患者的满意度调查。在慢病管理方面，智能化随访平台可以实现从院内到院外的高效综合管理慢病患者，患者可通过微信小程序、公众号等方式获取康复指导、复诊提醒、在线沟通、健康宣教等多种医疗服务。在专科专病方面，智能化随访平

台能够帮助医护人员更加全面、高效、连续地管理患者并收集专病数据，降低医护人员工作量，持续提升医疗服务质量、科研水平和医院的品牌价值。

2.2 按交互工具

（1）短信

短信具有高效、低成本、快速简便和安全等特点，非常适合与服务对象延伸互动和个性化精准服务。在智慧医疗场景中，以电子病历为临床信息源，借助信息化技术优化和再造业务流程，通过短信触发为患者提供就医过程中的温馨提醒，改善患者“看病难”“看病烦”困境，让数字化有温度，减少医患纠纷[2]。新兴的 5G 消息技术具有多媒体化、轻量化、交互化的特点，正逐渐推动短信服务的智慧化，患者在家里即可收到来自医院的含有图文并茂的各种检验检查报告的手机短信，也可以在线咨询，更快更好地实现医患间的有效沟通。

（2）聊天机器人

医疗聊天机器人主要利用人工智能等数字化技术，模拟人类对话，为患者答疑指引，能够有效改进传统医患沟通中的实践限制、专业术语理解等难题。人工智能语音识别技术、自然语言处理技术等是医疗聊天机器人的关键实现技术。医疗聊天机器人可以安装在医疗网站、移动医疗应用程序、新媒体程序、医院自助服务设备上。ChatGPT 是目前智能对话技术的代表，也衍生出了医疗 ChatGPT。实践表明，医疗 ChatGPT 能在医患间沟通发挥积极作用，如实时线上对话能够解决患者看病难、就医不便的问题，也更加凸显以患者为中心的沟通模式，全面专业、多语言交互、简明易懂的解释能够让患者更加了解自身疾病，从而增进医患间的相互理解与信任，提高治疗依从性和疗效[3,4]。

（3）网站类

在线问诊平台、在线医疗健康社区等是在“互联网+医疗”良好生态下，得以大规模推广应用、模式较成熟的一类健康医疗服务，核心本质是医患交流，为患者提供寻医问药的线上平台，是医生在线疾病诊治和健康咨询的空间。春雨医生、好大夫在线、寻医问药、丁香园等国内远程医患交互平台为患者提供医院和医生信息的查询，还支持医患间的在线沟通等。医疗机构也逐渐意识到医患沟通需要多样化，开通了官方网站上的医患在线交流，增加医患沟通时间。

（4）应用软件类

除了一些常规的医患互动的应用软件之外，QQ、微信等新媒体应用程序也为医患交流提供了更加便捷的途径。医疗机构通过微信公众号和视频号、微博账号等其他新媒体账号，开展健康知识传播、在线咨询和指导。微信群、QQ 群等也成为科室和医生管理病人的有效渠道，使得医患交流的方式更加灵活，信息共享和传播更加便捷。

2.3 按应用场景

（1）互联网医院

互联网医院为医患线上互动提供了平台，主要表现在能够为患者提供在线咨询、视频问诊。对患者来说，可以随时随地在系统上查看医生信息，学习健康宣教知识，发起咨询，自主选择合适的医生，并与医生及时进行交流，了解自己的身体状况和诊疗方案，提升了就医可及性和透明度。对于医生来说，随时随地接诊延伸了医疗服务的范围，提高了服务效率，拉近了与患者的距离。近年来，智能客服、虚拟护士等在互联网医院的应用，更是增进了患者便捷就医、有温度的智慧化体验，特别是一些适老化的创新举措，方便了老年患者的核心诉求，使得老年群体也享受到更多的就医获得感和幸福感。

（2）远程医疗

远程医疗中的视频会诊是促进医患沟通的重要场景。借助 5G、人工智能等信息化手段，推动优质医疗资源下沉，患者在家门口就可以享受到远程实时会诊，与专家们通过实时视频、图文等方式在线交流病情、共享病历，实现线上疾病咨询和诊断等，有效缓解优质医疗资源总量相对不足、分布不均衡的问题，推动分级诊疗的落实，提高基层医疗服务能力，减轻病人负担。

（3）智慧医院

我国医疗服务正在经历数字化转型，向着智慧化迈进。智慧医院的建设和评估已经取得了一定成果，“指尖医疗”开始普及。智慧医院借助短信、人工智能、5G 等数字化技术，创造覆盖诊前、诊中、诊后的全流程“智慧就医”生态，让患者看病更舒心、医疗服务更体贴。线上复诊省时省力、

院内就诊动态实时通知、诊疗和费用信息更加透明等无一不体现着医患沟通的提升。在这样的环境里，患者的就医疑问和痛点少了，医生的诊疗空间和时间多了，医患间的互动变得更加便捷、多样化、有温度。

3 医患智能交流的特征

3.1 更加灵活和多样化

从医患交流的类型变化可以看出，智能医疗时代医患交流不再局限于面对面的物理诊室内的交流，开始变得更加灵活和多样化。这不仅是因为数字化技术在医疗领域的广泛应用，也是因为医患群体思维和生活习惯的更新。线上的医患交流能够减轻病人的经济负担，让患者与医生的交流能随时随地开展，增加了医患交流的频次和深度。不同的人群可以根据自己的实际情况[5]，自主选择线上线下的交流方式。

3.2 更加绿色高效

有研究揭示，医患交互的不充分和低效率是造成现阶段医患矛盾重要原因[6]。智能医疗时代的医患交流由于越来越少受限于时空因素，能够让信息更加及时、透明，平衡医患双方医学信息不对称，节约不必要的搜索选择医疗资源、排队等待时间和经济支出等沟通成本，增加一对一的交互时间，让彼此更多理解和信任，从而让每一次医患交互尽可能地有效，让诊疗和就医的体验尽可能地是绿色良好的，最终，提高患者满意度、医嘱遵从率、疾病治愈率等诊疗结果[7]。

3.3 以患者为中心和共同参与决策

随着人们的健康意识越来越强，高质量的医疗服务需求被越来越多地提出。加上，互联网和社交媒体为患者选择医疗资源和了解医学知识提供了越来越多的机会。患者自主选择的时代到来了[8]。掌握更多的医学知识、愿意和敢于同医生讨论的知识性患者越来越多，医患关系模式逐渐向“医患共同参与型”转变。然而，这种转变对医生来说，一定程度上产生了压力，甚至会反感和压制患者的这种主动参与交互行为，不利于和谐医患关系的创建。

3.4 数字共情不可或缺

医学本质上是人学，现代医患关系往往是建立在陌生人和陌生人之间，就医过程中，医生和患者各自承担不同的压力，容易缺乏情感互动和双向理解。而患者身患疾病时，对医生的期望不仅是来自身体康复诊疗，也体现在精神和心理的抚慰。这就需要医生对患者共情。而线上的交流，尤其是智能客服，往往不具备生物人医生的人文思维[9]，且看不着、摸不到，难以精准捕捉彼此的情绪变化，不容易产生医患共情。将线下的共情引入到线上，创建医患数字共情，有利于缓解患者疾病发生时的消极情绪，促进和谐在线医患关系和在线问诊的良性发展[10]。

3.5 虚假信息的误导容易破坏医患信任

网络信息良莠不齐，容易被有利益纠葛的企业、医疗机构和个人利用，从而传播一些误导患者的虚假信息[11]。而不同背景的患者对真假信息的辨别能力各异，无法估量虚假信息背后的风险，轻则是短暂的信息冲突[12]，重则可能是暴力伤医、杀医事件。当这些不良就医事件传播时，医患信任的建立就更加艰难，防御性医疗也就不难理解了，患者“看病难”的内涵就更深了。

3.6 智能交互支撑技术存在一定的医学伦理风险

医疗聊天机器人、5G、移动互联网、远程会诊等各种智能交互支撑技术均需要存储、处理和传播大量的医学数据和隐私信息，存在来自平台、政策、网络黑客等的信息泄露风险[13,14,15]。此外，以 ChatGPT 为代表的智能对话技术的算法可解释性有待考证，如何界定人工智能的法律责任是一个重要的伦理问题[4]。而且这类算法依赖于真实世界的训练数据，是否会具有一定的倾向性且难以顾及到个体化差异等也是应该认真考虑和避免的医学伦理问题。

4 医患智能交流的优化策略

4.1 转变临床决策的模式

医患智能交流模式的转变反映出，随着数字生活的发展，科技知识普及，医生的权威和主导决策地位都在发生改变，患者决策能力得到极大的提升，并开始积极参与到临床决策中。为应对这种变革，医方应该认识到线上医患交流对健康促进的积极作用和巨大潜能，主动利用智能医疗的各种技术推动平行化医学，为患者参与临床决策创造更多有利条件。而患者也应该增强权利意识，积极参与决策，提高自身的网络健康素养，为自我健康、参与诊疗方案的制定赢得更多主动权。

4.2 医患共同营造数字共情

研究发现，受到医生共情的患者更愿意向医生讲述更多病情，态度也会更加信赖，语言表达也更轻松活泼，而收到患者感谢信和数字礼物等积极反馈的医生，也更加愿意主动回应患者的诊疗需求，通过这样的一个积极互动，医患信任更加容易建立和巩固[10]。因此，一方面，参与到智能诊疗中的医生可以练习数字共情的技巧，通过友好的在线会话态度、主动的在线健康宣教等“软服务”来打造高质量的有效沟通，拉近与患者的距离。另一方面，患者需要了解在线诊疗的局限性，调整自己在线诊疗的期望，在沟通过程中，约束自己的沟通行为，对医生的付出给予公正合理的反馈。

4.3 提升医患的网络健康信息素养

首先，医务人员应该积极转变观念，适应时代变革，并帮助完善平行医生的培养，积极参与到智能医患交流中来，提高网络健康信息素养和沟通能力，妥善处理知识型患者的大胆质疑，邀请患者参与诊疗决策，推动医患共同体转向积极共生。其次，提高患者及其他人群的网络健康信息素养，借助智能医疗的技术使原本就处于信息弱势的患者群体能够更加深入的了解自身疾病状况，参与到诊疗决策，透明化就医过程，增进医患信任。主要措施如医生、医疗机构、卫生管理部门等将信息搜索、信息辨别、信息利用、信息安全等知识宣教给人群，加入适老化的调整，吸引老年群体主动接触和学习智能医疗服务，让数字化变得有温度。

4.4 保障信息安全

首先是政府部门需要加强网络安全监管，整顿恶意传播虚假医疗健康信息的网站和个人账号，制定个人隐私数据的分级分类管理办法。其次，医疗机构和医药行业组织应该及时对错误的医疗和健康信息进行辟谣，帮助人们分辨真假网络信息，同时在开展智能医疗服务过程中，要借助 5G、区块链等创新网络安全技术保障电子病历数据、处方数据、检查检验数据共享的效率、安全和可信[16]。最后，提高人群的网络信息安全意识，鼓励人群主动学习网络安全和个人信息保护的相关法律法规、信息搜索和利用的安全知识，多与专业的医护人员沟通交流，避免因处于信息弱势而被误导。

PU-283

病例智能分组与规范化诊疗质控管理系统的实践与探索

秦金华① 田晓琼① 欧成川① 刘超毅①Δ

秦金华 田晓琼 欧成川 刘超毅*

湖南省肿瘤医院

0 引言

传统的医疗系统，存在流程不规范、收费不透明、费用虚高等问题，不能满足长期持续的健康保险报销要求，无法准确评估医疗服务的质量和效率，影响患者就医体验以及医疗行业的健康发展。近年来，疾病诊断相关分组支付因其收费标准规范、收费透明、费用合理等优势，成为了医疗领域普遍应用的医疗服务收费模式。疾病诊断相关组(Diagnosis Related Groups)DRG 是根据病人的年龄、性别、住院天数、临床诊断、病症、手术、疾病严重程度，合并症与并发症及转归等因素，把病人分入 500-600 个诊断相关组，决定应该给医院多少补偿[1]。

湖南省肿瘤医院（以下简称“本院”）是一所肿瘤专科医院，近年来，本院也面临费用管控的问题，一直在寻求一种标准化的服务计费和管理方式，经过多次多渠道院内外综合调研分析[2-5]，决定搭建一套适合本院的病例智能分组与规范化诊疗质控管理系统，实现医疗质量规范化管理。通过规范医疗实施路径，减少不必要的医疗开支，规避不正当的医疗行为，构建医疗行业健康可持续发展的应对策略，对推动多科室深入协作，医院诊疗规范化提供新的科学实践依据。

1 我院传统系统流程及存在的问题

1.1 传统系统流程

传统的医疗系统流程包括医院同步物价标准、医生开立医疗服务项目、患者缴纳费用、医保审核报销、病案归档等几个步骤：

（1）医院同步物价标准：严格按照上级主管部分和物价部门的收费标准，规定医疗服务项目价格。

（2）医生开立医疗服务项目：医生根据患者的病情，开立相应服务项目（如住院、门诊、化验、检查、手术等）。

（3）患者缴纳费用：患者向医院缴纳相应的费用（如医疗服务费、药品费、卫生材料费等。）

（4）医保审核报销：医院向医保机构提交相关的医保费用清单，医保机构审核费用的真实性和合理性，并予以报销。

（5）病案归档：医院将患者的病历和费用清单进行整理和归档，并存档于医院档案中。

1.2 存在的问题

传统的医疗系统在运行过程中，逐渐暴露出诸多问题，从不同主体层面分析有：

（1）医生层面：容易存在病例分类主观性大的问题，医生在诊断和治疗时难以做到全面、客观和科学，从而导致医疗质量不高。

（2）患者层面：存在着计费方式不透明的问题，医疗费用的不确定性和不透明性，导致患者难以承受高额的医疗费用。

（3）医保层面：存在着医保报销标准不统一的问题，导致对医疗资源的使用和浪费。在传统医疗系统中，医保费用的报销标准各个地区和单位之间存在很大差异，导致了资源的浪费和无法统一管理。

（4）病案层面：存在着病案管理不完善的问题，病历记录不规范，标准不统一，容易出现记录病历不真实、记录不全、记录错漏等问题，影响医疗质量[6]。

（5）管理层面：管理手段比较滞后，处理医疗问题缺乏科学、系统的方法，管理方式也比较僵化，容易引发经济效益指标过分强调的问题，面对复杂、多变的医疗环境表现不佳。

（6）控费层面：存在着控费手段不力的问题，医生处方过度开药、检查数量过多、检查费用过高等问题，导致医疗费用过高，影响了医疗质量和医疗服务的普及程度。

2 系统建设分析

以解决医院现存问题为导向，通过信息技术，建立病例智能分组与规范化诊疗质控管理系统（以下简称“本系统”），主要分为三个阶段来进行本系统的设计：（1）建立临床数据质控系统，解决数据质量问题；（2）开展医院智能分组支付测算，解决入组控费问题；（3）多维度分析监测，推进规范化诊疗院内质控管理系统落地实施。

2.1 建立临床数据质控系统

该部分主要是将数据质控前移，从源头护航支付结果。主要完成以下内容：（1）诊疗过程实时质控与盈亏预判，形式到内涵质控（诊断查漏+手术查漏）；（2）病案端专业编码缺陷纠错，临床沟通培训；（3）风险病案排查，医保结算清单质控。（4）建立院内数据质控模型，编码知识图谱。见图 1。

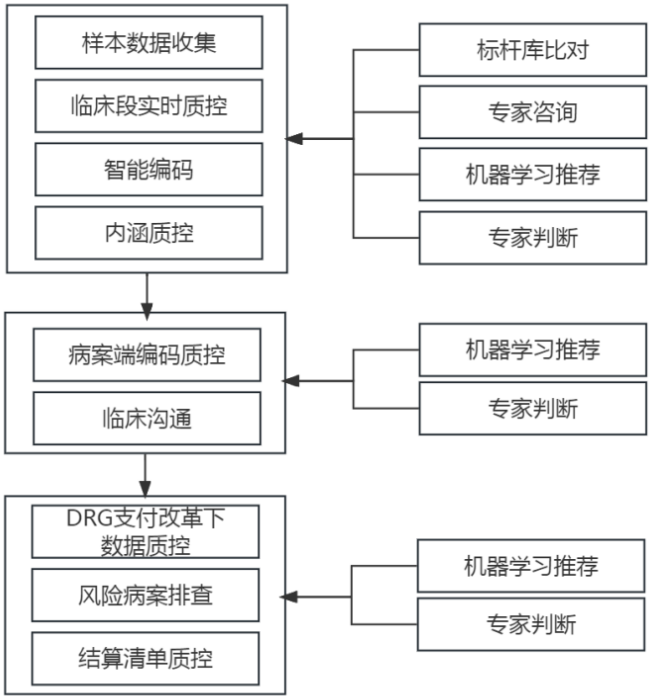


图 1 临床数据质控系统设计

2.2 医院智能分组支付测算平台

该部分主要建立与本地医保局相一致的支付测算平台。主要完成以下内容：（1）本院 DRG 支付预判，模拟分组，编码推荐；（2）基于本区域数据池建立标杆对比。（3）DRG 分组付费分析，对 DRG 模型使用前后的盈亏进行对比分析。

2.3 规范化诊疗院内质控管理系统流程设计

该部分主要建立医保申诉与成本管控体系；颗粒度细化至科室-医师-病种的多维度指标分析监测；透视症结，反哺深层管理与机制建立；建立以院长为中心的规范化诊疗院内质控管理工作组。主要完成以下内容：（1）医保申诉，医保管控；（2）搭建精细颗粒度的院内绩效评价，对标确立医院等级评审目标；（3）大数据支撑下的临床路径管控；（4）建立与支付方式改革相一致的薪酬绩效分配体系；（5）关键指标追踪管理，高效调控支付与绩效运营。见图 2。

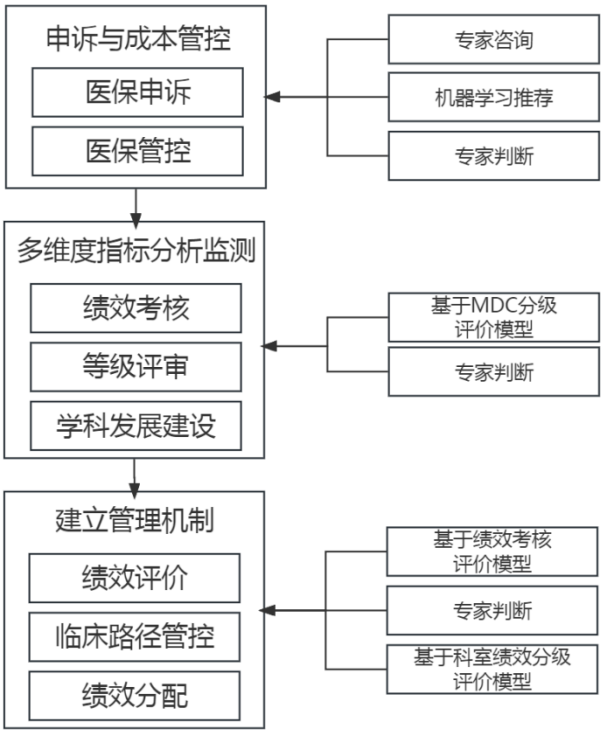


图 2 规范化诊疗院内质控管理系统设计

3 系统设计与实现

3.1 设计思路

围绕 3 项主要工作进行病例智能分组与规范化诊疗质控管理系统的设计：智能分组方法设计、建立四级质控流程、规范化诊疗院内质控管理流程设计。

3.1.1 智能分组方法设计

完备的医学知识和智能化规则是实现病例智能分组的核心模块。首先获取患者的病案数据和结算数据，利用人工智能 AI 技术自动获取数据的特征集（包括年龄、性别、住院时间、医疗总费用、主诊断、主手术操作级别和 5 天内转院或死亡等），并利用裁剪算法对特征集数据进行处理；然后根据主诊断分类 MDC，结合主要手术操作，进行细分获得 ADRG；在 ADRG 的基础上结合并发症、伴随病将 ADRG 进一步细分成三组；从特征集中选择核心特征因子，对细分三组的病案数据进行聚类得到 DRG；最后利用决策森林对 DRG 组进行处理，合并部分分组，获得最终 DRG 分组推荐[7]。

肿瘤会持续增生，肿瘤患者在诊疗过程中需要长期监测病情，选择主诊和主操作时需要对患者的历次病历调阅进行深入分析。特别是年龄大、体质弱、基础疾病较多的的恶性肿瘤患者，经手术治疗或化疗后免疫力下降，治疗的预后效果不佳等原因导致入组难度提升[8]。以肿瘤专科常见乳腺恶性肿瘤为例，原来的诊断为乳腺恶性肿瘤，手术操作为静脉注射化疗药物只能入歧义组；根据病例智能分组推荐，将诊断更改为 1、手术后恶性肿瘤化学治疗，2、乳腺恶性肿瘤，手术操作改为静脉注射化学药物，就能入组 RE1 恶性增生性疾患的化学治疗和/或其他治疗。

3.1.2 建立数据四级质控流程

建立临床医生质控—科室质控专员质控—病案室质控—医保科质控四级质控流程。实现质控的全面覆盖，具体质控流程见下图 3。

1、医生端实时质控，根据病人诊疗过程的诊断、检查、检验、手术情况，通过 DRG 智能编码推荐，对 DRG 预分组与权重进行提示；通过与标杆值进行对比，提示并指导医生用药与耗材等费用管控；同时对“手术/诊断”编码从形式到内涵进行质控提示，规避因“手术/诊断漏编、主要诊断选择错误以及其他诊断多编导致的 DRG 支付损失”；还能对编码缺陷智能提示，纠正关键轴心问题；对非编码缺陷智能提示，从数据源头保证首页填全填准。

2、医生填写完病历进行提交，到达科室质控专员质控阶段，科室质控专员对整个科室的病历质量进行把控，然后将病历提交给病案室。

3、病案室编码员检查病案编码，对存疑病案一键反馈临床待查待办，描述关键疑问。

4、最后是医保科质控，医保科进行风险病案排查，结算清单质控。首先根据常见风险病案库对五类风险病案逐一排查，批量处理；然后根据医保局提供的数据结构，自动生成医保结算清单；保持与医保局上报标准一致的质控原则，对医保结算清单质控纠错，对风险问题进行批注，形成风险清单；最后根据风险清单批注的问题及提示分析，对可能影响 DRG 分组的问题指派相关专业人员来解决处理；通过优化清单问题，持续跟踪优化效果，进行统计分析，开展具体到科室的追踪。

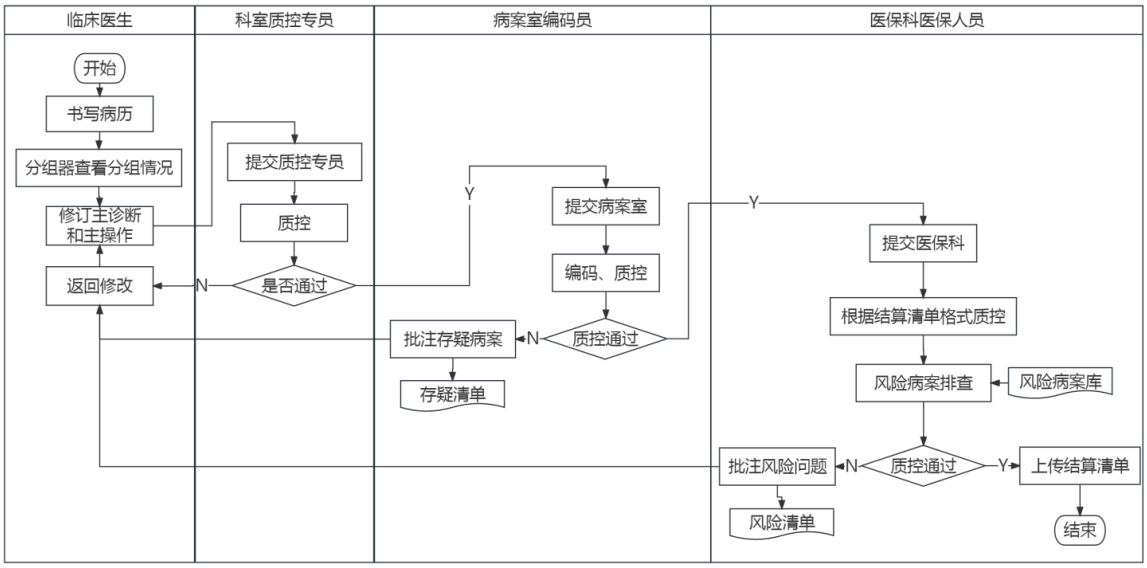


图 3 四级质控流程

3.1.3 规范化诊疗院内质控管理系统

规范化诊疗质控管理通过实时监测+质疑沟通，对潜在风险进行预警。根据本地 DRG 结算规则，定位潜在问题病案和单议病案；根据二八定律，按照亏损大小前 20 的病组排序，分轻重缓急对影响亏损最大的病组优先进行控费；通过核查可能存在主诊断与主手术不匹配、主诊断或主手术选择错误的情况；核查“有手术收费入内科组”的病例，例如患者同时做了化疗、靶向或免疫治疗，但结算清单中主操作将靶向、免疫放在化疗之前，分组时错误进入靶向组、免疫组，造成医院医保基金亏损；又如对给予静脉注射化疗药物的患者，漏填该操作为主要操作，导致误入“RU1 恶性增生性疾患化学及或靶向、生物治疗”内科组[9]。最后筛选可申诉病案，建立院内医保结算申诉反馈机制；对常见申诉原因提示，自动生成申诉报告。

对关键指标追踪管理，高效调控成本与绩效运营。首先，根据医保支付测算的结果，预测本院的整体盈亏情况，把握收支管控方向；对“亏损风险”较大的科室进行追踪管理，明确规划管控范围；然后汇总医院各项病种的平均费用，通过模拟点值，评估各个科室的亏损情况，指导临床科

室；如果结果差异不大，就维持现状，如果超过实际发生的费用 110%，就需要设定控费方案（主要是从耗材、药品等方面着手）；根据各个科室的收入变化情况，来定点主要集中科室，重点加大力度管控这些收入变化大的科室；对收入变化进行持续性监测，把控院内管理整体成效。

通过监测本院服务能力指标（CMI 值，DRG 组数，已开展 MDC 种类），比对行业标杆值，明确医院整体能力水平在行业的位置；监测全院各个科室指标（CMI 值，DRG 组数，时间消耗指数，费用消耗指数，低风险组死亡率等），导向科室差异化发展策略；纵览全院整体的五个 DRG 核心指标在行业分布情况（高于同行业率，低于同行业率），来定位医院发展管控空间，建立以院长为中心的规范化诊疗院内质控管理工作组。

3.2 技术架构

本系统的总体架构分为数据库层、数据处理层、功能层、和用户层，见图 4。

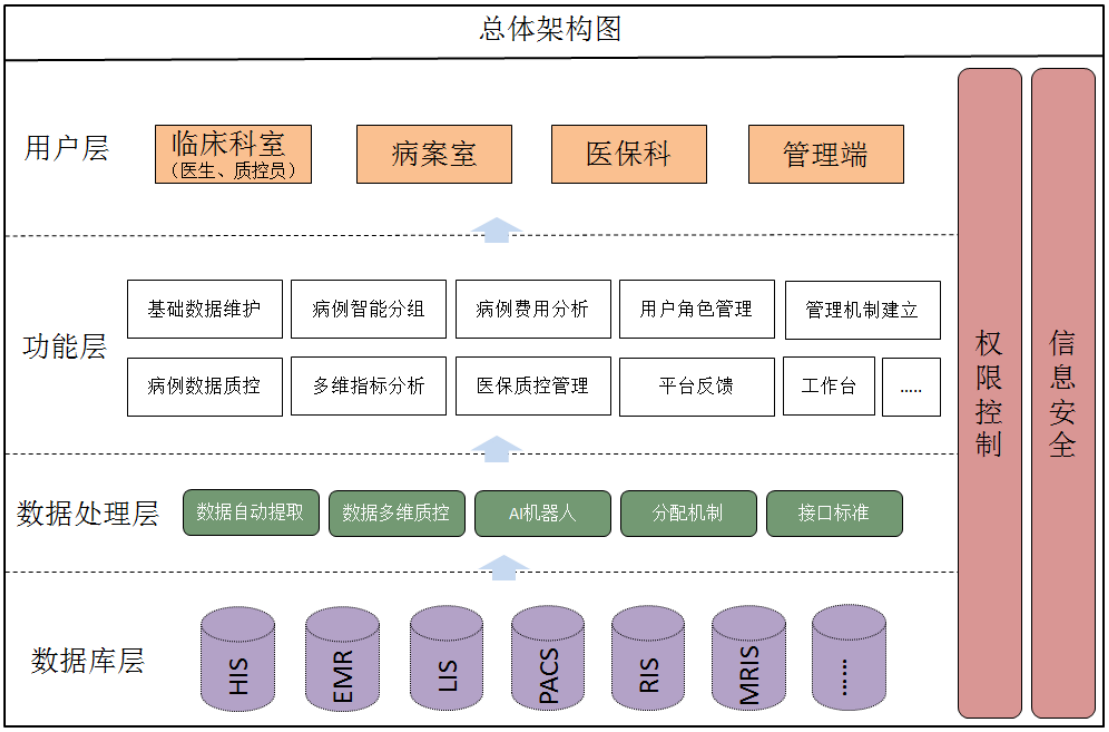


图 4 系统总体架构图

3.2.1 数据库层

本系统采用 B/S 架构，主要采用前后端分离技术，前端负责页面展现和用户交互，后端负责处理业务逻辑，两者侧重点不一样，有利于前后端各自开发体系仅关注自己的重点，有利于解决前后端职责不清的问题，解决页面难以维护的弊端，更加关注组件性能优化响应用户需求，有助于项目高效，互不干扰的完成。患者数据采集对接医院 HIS、LIS、PACS、电子病历等系统，以建立准确的患者档案，使用 Oracle 进行数据存储。

3.2.2 数据处理层

原数据是驱动决策的基础，数据来源越全面，标杆库建立越全面。采用 ETL 数据治理对过去三年数据清洗分析，对比本地标杆，建立本院的多角度基础标杆库，供质控各阶段目标对比，透视差距。对常见编码缺陷进行收集，建立编码缺陷库；对五类常见风险病案（歧义病案、编码未识别病案、费用极高病案、费用极低病案、低风险组死亡病案）收集分析，建立风险病案库。

3.2.3 功能层

包括基础数据维护、病例数据质控、病例智能分组、病例费用分析、多维指标分析、医保质控管理、用户角色管理、管理机制建立等功能。

3.2.4 用户层

包括临床科室（医生、质控员端）、病案室质控端、医保科质控端和管理层决策管理端，通过分层分域管理机制实现权限管控和信息安全保障。

3.3 功能模块

系统主要包括病例数据质控、病例分组、病例费用分析、基础数据管理、系统管理、接口需求等模块。此外，系统预留标准化数据接口，方便接入新的设备和外部数据库；预留编程接口，方便跟新的系统做对接。

3.3.1 病例数据质控模块

病例数据质控包含病例数据质控概览、病案首页校验、结算清单校验。 数据质控概览从数据质控规则、问题字段分布，展示病案首页/结算清单数据质控问题，并描述其质控问题趋势。病案首页校验包含病案首页查询、校验规则配置、校验结果查询、病案质量统计。结算清单校验包含结算清单查询、质量校验规则配置、结算清单质量分析、校验结果查询、结算清单质量统计。

3.3.2 病例分组模块

病例分组包含病例分组概览、在院分组、DRG 智能分组管理。 病例分组概览展示归档病案/结算清单的入组情况，对未入组原因情况分析。在院分组每天对在院患者数据进行 DRG 分组，展示当前病例的数据分组、分组标准费用、住院天数等，并可查看在院患者住院期间的历史分组情况。DRG 分组管理包含分组方案配置、诊断分组查询、手术分组查询、CHS 入组预测、DRG-CHS 分组、入组情况查询。

3.3.3 病例费用分析模块

病例费用分析包含 DRG 费用分析和 DRG 指标管理。 DRG 费用分析包含 DRG 综合分析、DRG 盈亏分析、DRG 科室统计、DRG 医生统计、DRG 医疗组统计。DRG 指标管理包含数据质量分析、医疗服务能力、医疗服务质量、医疗费用控制、DRG 科室绩效排名、考核评估细则。

3.3.4 基础数据管理

基础数据管理包含疾病目录维护、手术目录维护、权重点数维护、数据字典管理、计划任务维护、计划任务结果管理、医保政策管理、医保政策查询、医院组织架构管理、医生信息维护、医院信息维护、科室信息维护。

3.3.5 系统管理和接口管理模块

系统管理包含用户管理、角色管理、模块管理。接口管理包含基础业务数据采集、结算清单数据采集接口、病案首页数据采集接口、结算清单数据质量校验接口、病案首页数据质量校验接口、病案首页 CHS-DRG 分组接口、结算清单 CHS-DRG 分组接口、DRG 在院预分组接口。

4 建设成效与应用分析

医院在 2023 年 1 月份开始正式上线本系统， 截至 2023 年 06 月 30 日共上传清单 73609 人次。 2023 年 01 月至 2023 年 06 月住院总费用为 107384.82 余万元，基本统筹支出 39456.54 万元，经测算 DRG 支付费用为 45373.58 余万元，预计盈利 4411.37 万元。住院次均费用为 14594.69 元，CMI 值为 1.45。通过对服务能力、医疗效率、医疗安全维度的比较，发现本系统上线后，病例入组率、平均住院天数等都有了明显改进，患者的医疗费用、次均费用、药占比以及耗占比都有不同程度的下降。

4.1 诊疗过程实时质控，有效提升入组率

本系统上线后，通过诊疗过程实时质控，运用高拟态预分组技术模拟分组，编码推荐，智能分组实时提示，充分释放临床医生编码时间，规避漏编风险，提升了编码质控效率，有效提升入组率。在 DRG 结算患者增加 7.71%的情况下，2023 年上半年正常入组 73042 人次，入组率达到 99.23%，同比增长 86.45%；DRG 组数增加为 379 组，同比增长 12.80%；CMI 值虽有所下降，但总权重同比提升 114.30%。见表 1.

表 1 医疗服务能力

时间	DRG 结算人次	正常入组	入组率	DRG 组数	总权重值	CMI
2022 年 1-6 月	68341	36371	53.22%	336	49439.8306	1.59
2023 年 1-6 月	73609	73042	99.23%	379	105948.5194	1.45
同比	7.71%↑		86.45%↑	12.80%↑	114.30%↑	8.8%↓

4.2 提升医疗服务效率，让诊疗更安全

本系统上线，医院通过数据的分析和对比，不断优化就诊流程，进行一系列措施的改进与实施，使临床科室、医技科室的工作人员都有了成本效益意识，服务意识变被动为主动，医疗服务效率得到了促进和提升。见医疗服务效率表 2。

2023 年上半年医疗总费用同比增长 6.55%，总盈亏同比增长 152.52%，患者次均费用同比下降 5.28%；费用消耗指数、时间消耗指数同比均大幅下降，药耗占比、手术费用占比呈现不同程度下降；患者的平均住院天数同比下降 1.62%，但床日产出率同比增长 102.36%。同时通过各风险组的病例死亡死亡率分析，本系统能保证更好的医疗安全。见表 3。

表 2 医疗服务效率

时间	总费用（万元）	次均费用（元）	总盈亏（万元）	平均住院日	床日产出率
2022 年 1-6 月	100669.52	15408.44	-8399.16	7.41	0.0976
2023 年 1-6 月	107384.82	14594.69	4411.37	7.29	0.1975
同比	6.55%↑	5.28%↓	152.52%↑	1.62%↓	102.36%↑

时间	费用消耗指数	时间消耗指数	药占比	耗占比	四级手术占比	手术费占比
2022 年 1-6 月	2.02	1.86	34.94%	16.33%	55.76%	7.58%
2023 年	0.94	0.83	30.49	15.	58.13%	6.97%

1-6 月			%	56%		
同比	53.47 %↓	55.38 %↓	12.74 %↓	4.7 1%↓	4.25%↑	8.05% ↓

表 3 医疗服务质量

时间	低风险组	中低风险组	中风险组	中高风险组	低风险组死亡率	中高风险组死亡率
2022 年 1-6 月	12974	8192	9753	5452	0.01%	0.11%
2023 年 1-6 月	27680	2705 1	9216	9095	0.00%	0.09%
同比					100%↓	18.18% ↓

4.3 多维度指标监测，有效提升医保管控能力，精准预警潜在“亏损”

本系统上线后，能从全院-科室-医生-DRG 病组四个层面实时监测，精准预警潜在 DRG 支付风险。通过医院当期盈亏将风险拆解至风险病案，指导追踪；通过科室当期盈亏对各个科室盈亏情况一目了然，并下钻至医生，提前对潜在“亏损”进行风险预警；通过医生盈亏制定差异化的医生成本管控能力培养策略，将有实力医生重点培养为科室标杆，带动低效医生；通过 DRG 病组盈亏分析，优先管控差距值较大的病组。

表 5 为科室盈亏情况（2023 年上半年）

科室盈亏情况	科室数	病例数	病例数占比(%)	平均盈亏(元)	总盈亏(万元)
亏损科室	14	11020	14.98%	-2849.28	-3139.90
盈余科室	29	62589	85.02%	1207.08	7551.27
合计	43	73609	100%	599.55	4411.37

5 结论

病例智能分组与规范化诊疗质控管理系统的搭建与应用，构建了医疗行业健康可持续发展模式，解决了当今医保管控难题。充分释放了临床医生编码时间，规避了漏编风险，提升了编码质控效率，有效提升了入组率；规避了不正当的医疗行为，提升了医疗服务效率，让诊疗更安全；减少了不必要的医疗开支，能够智能预警潜在风险，高效赋能医保管控，高效调控运营成本。为推动多科室深入协作，规范医院诊疗，提高医院管理效能提供了新的科学实践依据。

PU-284

科学与技术交叉视角下智能医学领域研究热点与技术预测研究

殷策 张云秋*
吉林大学公共卫生学院

0 引言

智能医学（Intelligent Medicine）是融合脑认知、智能生物材料、机械外骨骼、微纳机器人、医疗大数据等前沿技术, 以现代医学理论为基础, 研究生命的本质, 探究疾病发生、发展与演进的客观规律、发展新型疾病诊疗方案与人机协同临床应用的新兴交叉学科[1]。智能医学所覆盖的技术范

围包含医学人工智能、手术机器人、远程通信、医学大数据、5G 手术、医学 3 D 打印、可穿戴诊疗设备共 7 个技术领域[2-4]。

在大数据时代，科学与技术的互动融合日益紧密，而较为重要的创新也往往发源于互动关系较强的科学研究和重点技术[5]。目前，国内学者[6-7]从专利与论文关联的角度识别前沿技术主题，挖掘潜在研究热点的研究已经开始显现，而国外则少见相关研究。

1 基于科学与技术交叉的技术前沿识别模型构建

本研究在技术创新相关研究[7-9]的指导下，从科学与技术交互融合的视角出发，以科学和技术两个维度，构建基于科学与技术交叉的技术前沿识别模型如图 1 所示。

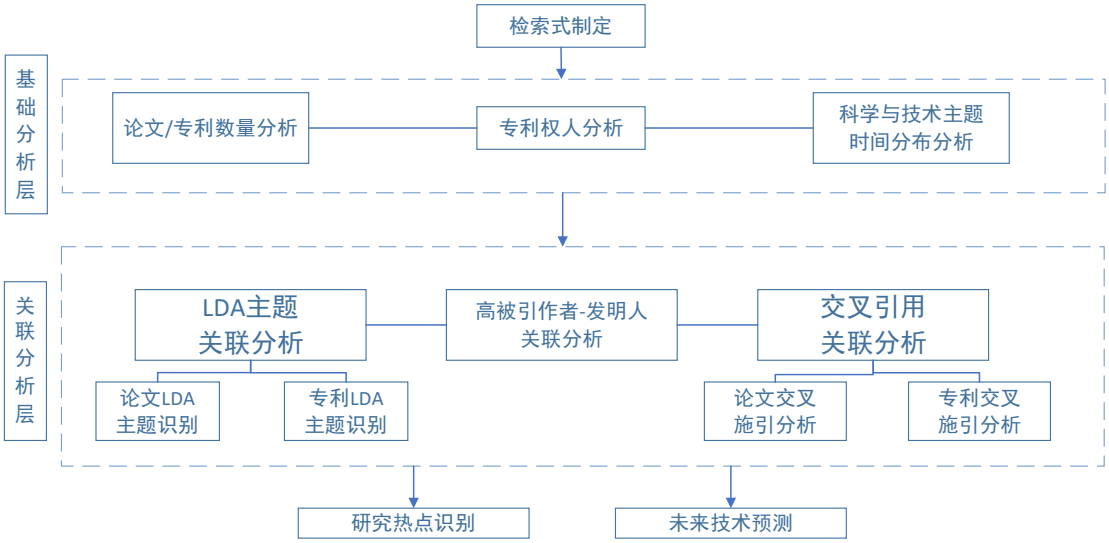


图 1：基于科学与技术交叉的技术前沿识别模型

该模型将本研究分为基础分析层与关联分析层两个层次，其中基础分析层包括论文/专利数量分析、专利权人分析和科学与技术主题时间分布分析三个子分析，意在探究研究领域的计量学特征。关联分析层包括 LDA 主题关联分析、高被引作者-发明人关联分析和交叉引用关联分析三个子分析，意在从科学与技术融合的角度，通过探寻研究领域内科学与技术的交织来达到研究热点识别与未来技术预测的目的。

2 数据与方法

2.1 数据来源

德温特数据库（Derwent Innovation Index，DII）是全球最权威的专利情报和科技情报数据库之一，因此在 DII 数据库中对专利文献进行检索。检索时间截止到 2023 年 9 月 2 日，共检出相关专利 91645 件，去除重复、核心字段缺失的专利后共纳入研究 91583 件。

Web of Science（WOS）是科睿唯安旗下的重要产品，本研究在 WOS 核心合集数据库中检索英文文献，检索时间截止到 2023 年 9 月 2 日，共检出相关文献 32028 篇，去除重复、核心字段缺失的文献后共纳入研究 32022 篇。

2.2 研究方法

本研究使用文献调研法、引文分析法、专利分析法等方法，利用 Python3.10 自主编程，并结合 CiteSpace 6.2R3 等软件对数据进行分析，对分析结果进行可视化展示。

3 研究结果

3.1 基础分析层

3.1.1 论文/专利发表数量分析

智能医学领域的论文发表数量前 20 名国家如表 1 所示。可以观察到，美国和中国是在该领域发表论文数量最多的两个国家，这表明美国和中国在智能医学领域具有突出的学术研究实力和贡献。

表 1：智能医学领域论文发表数量前 20 名国家表

国家	论文数量
美国	8897
中国	5551
德国	2449
意大利	2310
英国	2228
印度	1759
日本	1634
韩国	1599
加拿大	1475
法国	1240
澳大利亚	1060
西班牙	952
荷兰	876
瑞士	636
比利时	580
土耳其	539
新加坡	478
沙特阿拉伯	462
巴西	435
澳大利亚	419

在对不同国家发表专利情况进行分析时，有必要考虑到同族专利。同族专利是一组基于同一技术或理念的专利，其中某些权利可能已在其他国家或地区授予。因此，在评估不同国家的专利发表情况时，需要综合考虑同族专利的数量和质量。智能医学领域的专利（包括简单同族专利）发表数量前 20 名国家如表 2 所示。

表 2：智能医学领域专利发表数量前 20 名国家表

国家	专利数量（含简单同族）
美国	56730
中国	50974
日本	33833
韩国	17063
加拿大	15997
德国	15506
澳大利亚	15320
奥地利	14743
印度	9541
巴西	7738
西班牙	5101
以色列	4429
墨西哥	3860
英国	3237
俄罗斯	3090
新加坡	2415
丹麦	2119
法国	1731
南非	1615
波兰	1495

3.1.2 专利权人分析

智能医学领域发表专利数量排名前 10 的专利权人如表 3 所示，专利所承载的技术价值通常由企业专利权人的实力、专利施引/引用、同族专利数量等多个维度评价[10]，本研究从专利被引及专利维护两个角度评价专利质量，全球申请智能医学领域专利数量前 10 名的公司专利质量最高的前两个核心专利如表 4 所示。

表 3：智能医学领域发表专利数量前 10 名专利权人表

专利权人	所属国家	专利数量
直觉外科手术操作公司	美国	5096
柯惠 LP 公司	美国	2718
西拉格国际有限公司	瑞士	2086
奥林巴斯株式会社	日本	1263
CMR 外科有限公司	法国	1169
奥瑞斯健康公司	美国	1157
爱惜康有限责任公司	美国	1111
皇家飞利浦有限公司	瑞典	985
马可外科公司	奥地利	827
威博外科公司	巴西	817

表 4：智能医学领域前 10 名专利权人核心专利表

专利权人	核心专利 1	专利号	被引频次	核心专利 2	专利号	被引频次
直觉外科手术操作公司	用于手术机器人的平台腕状连接装置	US7691098B2	2564	手术机器人安装臂中的力和扭矩传感装置	US9895813B2	829
柯惠 LP 公司	用于机器人手术系统的腕部和颚部组件	US9937626B2	799	用于确定手术器械中线性运动参数的方法和装置	US9913644B2	742
西拉格国际有限公司	外科固定钉发射装置中包含多个故障状态保险丝的点火组件	US20180168610A1	1438	外科手术机械臂中包含离合器的轴组件	US20180168592A1	1420
奥林巴斯株式会社	远程医疗的支持系统	US7386730B2	1202	包含从臂、轴组合的手术机械装置	US9788902B2	729
CMR 外科有限公司	对称排列的手术机械关节	US20180360549A1	305	用于康复的手术机械关节	US10813705B2	261
奥瑞斯健康公司	用于机器人辅助腔内手术的系统和相关方法	US9763741B2	382	手术插入式器械工作通道补偿方法	US10299870B2	287
爱惜康有限责任公司	具有间隙开口的外科缝合器械	US9844379B2	949	手术机器人工具交换装置	US20180049824A1	496
皇家飞利浦有限公司	基于网状网络的远程中风监测系统	US20090318779A1	566	用于血管可视化和监测的光纤传感器导航/引导装置	US10194801B2	500
马可外科公司	基于图像的机器人手术系统	US10028788B2	171	在手术工具移动期间控制手术工具的系统和方法	US9937014B2	165
威博外科公司	将手术机器人机械臂连接到手术台的耦合器	US20180271604A1	476	用于机器人手术的沉浸式 3D 显示器	US10786327B2	137

智能医学领域发表专利数量排名前 10 的专利权人涉及多个国家和地区，其中美国企业居多。这些企业在智能医学领域投入了大量维护成本，致力于保护和推动其核心领域的技术创新。

3.1.3 科学与技术主题时间分布分析

对 2013 年至今的论文数据和专利数据进行数据清洗和格式转换后，将数据导入 CiteSpace 6.2R3 软件，将 Node Type 设置为“Abstract”、“Keywords”，以 Timeline 方式制作基于摘要和关键词的主题时间线图。

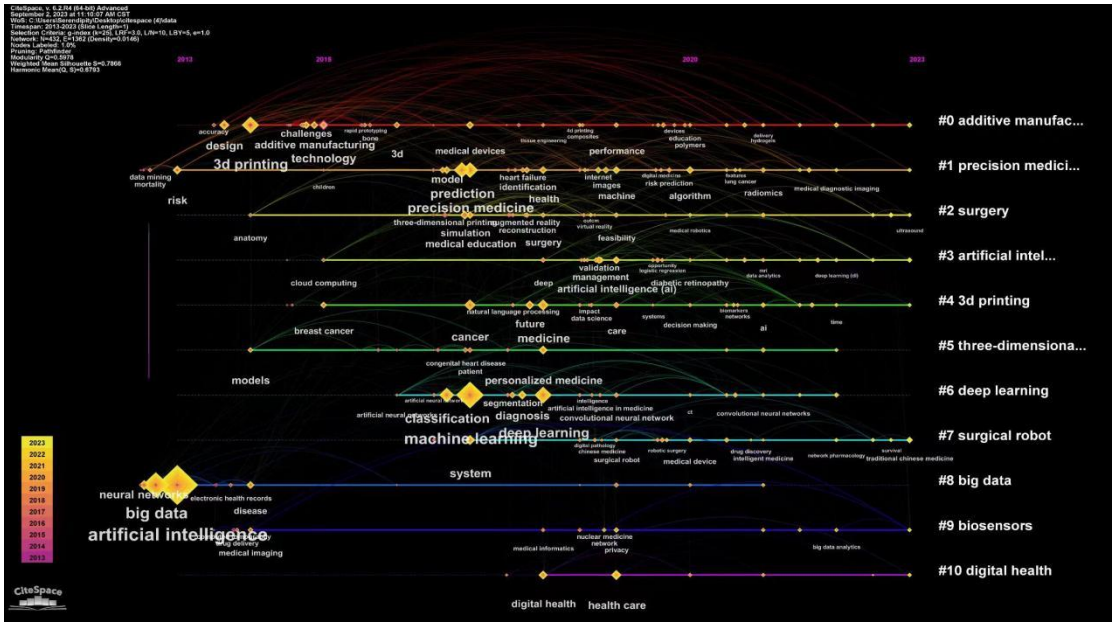


图 2：科学与技术主题时间分布分析图

如图 2 所示，智能医学领域的发展历程涵盖了 10 个主题簇。其中，“airtificial intelligence”（人工智能）作为贯穿始终的核心主题，为智能医学技术的发展提供了基础。近年来，智能医学领域出现了炙手可热的研究领域，其中“big data”（大数据）成为重要的关注点。大数据的引入为智能医学提供了更多的数据支持和分析能力，进而推动医疗诊断和治疗的准确性和个性化水平[11]。在 2017 年，出现了“precision medicing”（精准医疗）的主题簇。精准医疗技术的发展将个体基因信息、生物样本和临床数据结合起来，为每个患者提供个性化治疗方案，从而提高医疗效果。随着时间的推移，2018 年前后出现了“deep learning”（深度学习）的主题簇。深度学习作为一种新兴前沿技术，利用多层神经网络模拟人脑的运作方式，为智能医学领域的资源管理和深度神经网络的应用奠定了基础。“ultrasound”（超声）、“traditional chinese medicine”（中医）是 2023 年新出现的主题，这说明智能医学领域可能在传统中医和超声两个领域产生越来越多的科学与技术的交融。

3.2 关联分析层

3.2.1 LDA 主题关联分析

使用 Python3.10 自主编程，对获取的论文和专利数据进行数据清洗后，建立 LDA 主题识别数据库。

（1）论文 LDA 主题识别

根据论文数据的困惑度/一致性与主题数的关系，本研究最终选择主题数为 7 进行 LDA 主题分析。

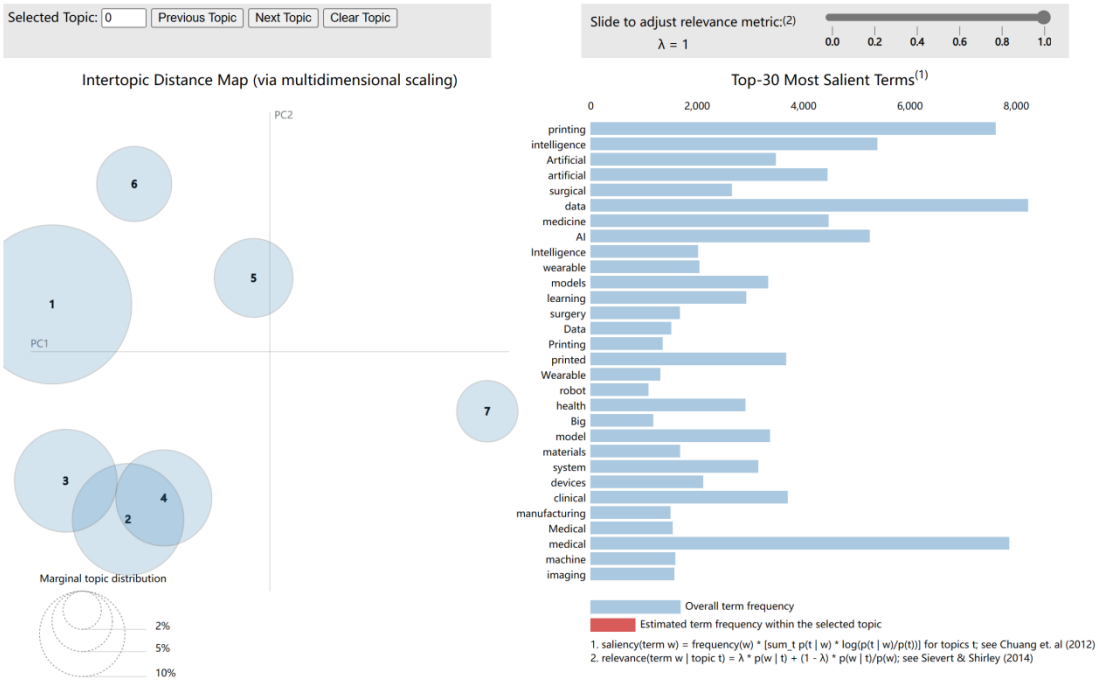


图 3：论文 LDA 主题分析

如图 3 所示，左侧的 7 个蓝色气泡代表论文数据聚类形成的 7 个不同主题，分别为“人工智能与大数据在医学领域的应用”、“智能医学方法与模型”、“医学影像操作设备与应用”、“3D 医学影像”、“可穿戴设备与远程通信技术”、“基于临床数据的人工智能算法”、“手术机器人”。根据论文数据集的特征词数量和主题集的分布，数据与人工智能是目前智能医学领域的研究热点。

（2）专利 LDA 主题识别

根据专利数据的困惑度/一致性与主题数的关系，本研究最终选择主题数为 7 进行 LDA 主题分析。

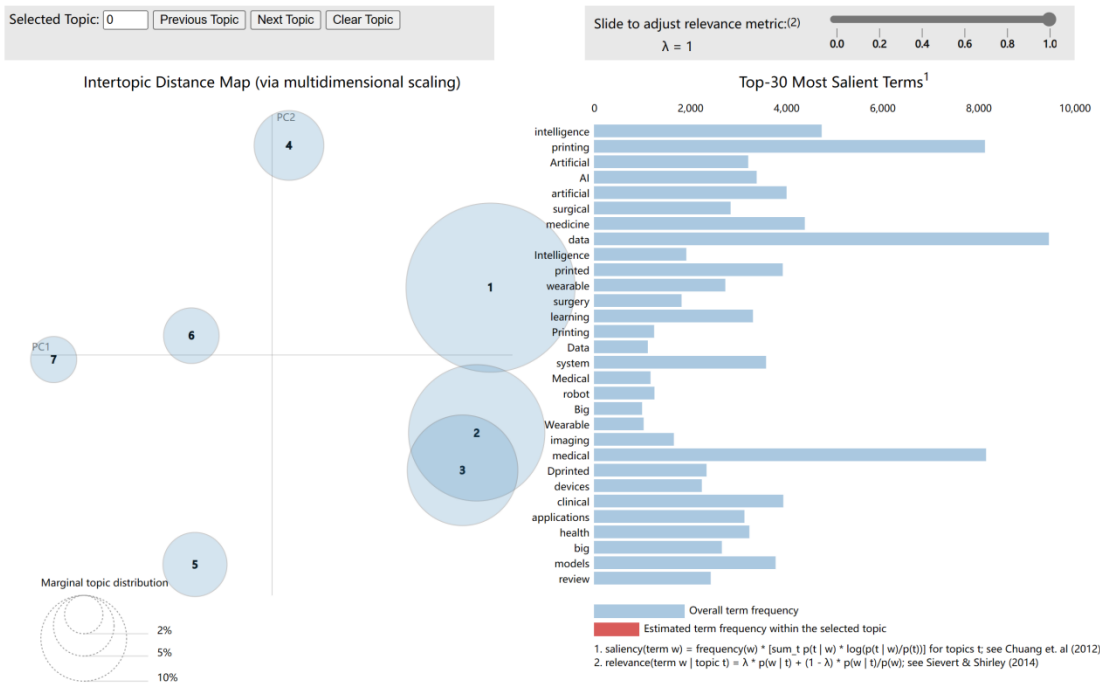


图 4：专利 LDA 主题分析

如图 3 所示，左侧的 7 个蓝色气泡代表专利数据聚类形成的 7 个不同主题，分别为“操作临床数据的技术”、“医学影像构建技术、模型”、“可穿戴设备与遥感技术”、“人工智能算法”、“手术机器

人”、“生物学结构的 3D 成像技术”、“医疗大数据分析系统”。根据专利数据集的特征词数量和主题集的分布，手术机器人、临床数据的处理和分析是目前智能医学领域的技术前沿。

3.2.2 高被引作者-发明人关联分析

本研究以 Boyack 提出的姓名匹配方法为基础，采用“姓名+机构+成果时间”相结合的方法进行高被引作者-发明人的识别[12]。本研究检索了智能医学领域被引频次前 100 的论文作者的姓名、机构等信息，并对其可能的录入名进行补充后，使用 DII 数据库的检索规则构建检索式，寻找高被引作者-发明人。最终找到的六位高被引作者-发明人如表 5 所示。

表 5：高被引作者-发明人表

序号	作者-发明人	所属机构	所属国家	论文数量	专利数量
1	Peppas, Nicholas A.	University of Texas Austin	美国	479	1
2	Michael Mitchell	University of Pennsylvania	美国	101	1
3	Takao Someya	The University of Tokyo	日本	507	16
4	Ru Cheng	Soochow University	中国	66	1
5	Azaria, Asaph	Massachusetts Institute of Technology	美国	5	1
6	Rajkomar, Alvin	Google Incorporated	美国	32	34

这些作者-发明人来自美国、日本和中国，抽取出这些学者的智能医学领域高被引论文和专利数据分别如表 6、表 7 所示。

表 6：高被引作者-发明人智能医学领域高被引论文表

序号	作者-发明人	领域内高被引论文	被引频次	发表年
AIL1	Peppas, Nicholas A.	Hydrogels in biology and medicine: From molecular principles to bionanotechnology	3045	2006
AIL2	Michael Mitchell	Engineering precision nanoparticles for drug delivery	1991	2021
AIL3	Takao Someya	A large-area, flexible pressure sensor matrix with organic field-effect transistors for artificial skin applications	1558	2004
AIL4	Ru Cheng	Glutathione-responsive nano-vehicles as a promising platform for targeted intracellular drug and gene delivery	1089	2011
AIL5	Azaria, Asaph	MedRec: Using Blockchain for Medical Data Access and Permission Management	941	2016
AIL6	Rajkomar, Alvin	Scalable and accurate deep learning with electronic health records	969	2018

表 7：高被引作者-发明人智能医学领域专利表

序号	作者-发明人	专利号	领域内专利	专利权人	申请年
AIP1	Peppas, Nicholas A.	WO2016115234A1	用于递送治疗化合物的水凝胶	德克萨斯大学董事会	2016
AIP2	Michael Mitchell	WO2023056418A1	可用于微型胶囊的脂质纳米颗粒(LNP) 组合物	宾夕法尼亚大学理事会	2021
AIP3	Takao Someya	WO2017026393A1	血流障碍检测仪	SEKINO MASAKI 等	2016
AIP4	Ru Cheng	WO2011009478A1	一种用于生物装置的可被改性的聚合物及其制备方法	SSENS 有限责任公司	2009
AIP5	Azaria, Asaph	US62252476P0	一种手口强迫行为监测装置	Azaria, Asaph 等	2015
AIP6	Azaria, Asaph	US62653238P0	使用深度学习和声音理解生成诊断健康信息的系统和方法	谷歌	2018
AIP7	Azaria, Asaph	EP3634204A1	用于从电子健康记录中预测和总结医疗事件的系统和方法	谷歌	2017
AIP8	Azaria, Asaph	US62778833P0	使用神经网络从临床记录中预测未来患者的健康状况	谷歌	2018
AIP9	Azaria, Asaph	CN109697283A	用于生成患者-健康护理提供者对话的文字记录的方法和系统	谷歌	2018

这些论文和专利的主题表明，当前智能医学领域的科学研究和技术研究都集中在开发高效的药物递送系统、研究生物传感与医疗设备以及探索基于区块链的医疗数据管理等方面。

3.2.3 交叉引用关联分析

本研究使用 Lens.org 网站的 PatCite 功能检索论文与专利交叉施引信息，从此前建立的论文与专利数据库中抽取出论文的 DOI 号和专利的专利号，对论文与专利的交叉施引信息建立数据库。

（1）论文交叉施引分析

从论文交叉施引数据库中抽取出被论文所引用的专利的名称和摘要字段，识别被引专利的技术领域并制作被论文引用的专利 IPC 分类树状图。

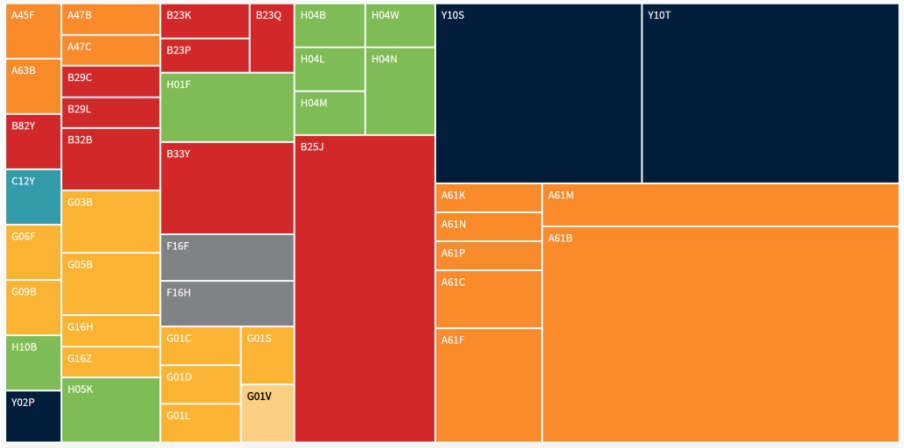


图 5：论文引用专利的 IPC 分类号树状图

如图 5 所示，在智能医学领域的论文中引用的专利中，除了医学相关的 IPC 分类号（A61 类），高比例地引用了 Y10S、Y10T 和 B25J 这三个分类号。智能医学领域是一个快速发展和技术创新的领域，与现代科技和工程技术密切相关。Y10S、Y10T 和 B25J 这些 IPC 分类号涵盖了不同领域的专利技术，包括特殊技术、技术交叉引用注册和机器人技术。

（3）专利交叉施引分析

从专利交叉施引数据库中抽取出被专利所引用的论文的名称和摘要字段，识别被引专利的技术领域并制作被专利引用的论文所属领域饼图。

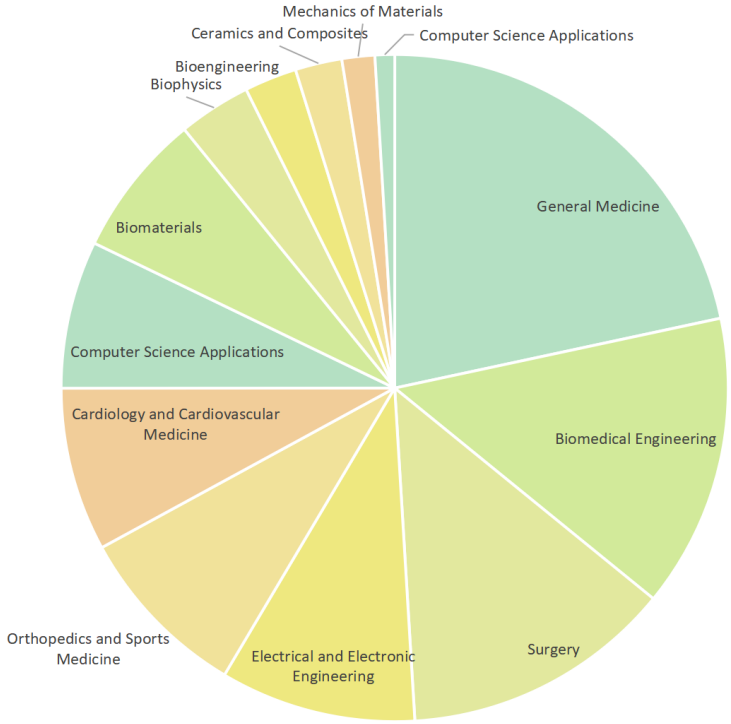


图 6：被专利所引用的论文所属领域图

如图 6 所示，根据引用数据，我们可以看到总体医学、生物医学工程、外科学和电气与电子工程等领域在智能医学的技术发展中处于关键地位。这意味着智能医学需要结合医学、工程、外科手术等领域的理论知识来推动技术的发展和创新。

4 研究结论

4.1 智能医学领域研究热点与技术前沿

基于基础分析层和关联分析层的研究，我们可以得出结论：智能医学领域的研究热点为人工智能、大数据、药物递送系统和生物传感。这不仅体现了智能医学领域的技术前沿，也预示着智能医学的发展将借助多领域知识的交叉融合。随着人工智能的迅猛发展，智能医学将能够更有效地处理和分析医疗数据，提供个性化的诊断和治疗方案。大数据的引入为智能医学提供了更多的分析能力，可以挖掘出隐藏在海量数据中的宝贵信息，为医疗决策提供科学依据。药物递送系统和生物传感技术的进步，将改变传统治疗方法的方式，使药物能够更加精准地作用于病灶部位并实现个体化治疗。

4.2 智能医学领域未来技术预测

本研究将上文中作者-发明人所著的高被引论文、智能医学领域专利作为实体集，其中高被引论文为 AIL 序列、专利为 AIP 序列；将科学与技术主题时空分布分析中的热点主题、交叉引用关联分析中科学与技术交融较强的交融主题作为环境集，其中热点主题用蓝色椭圆形表示，交融主题用

黄色三角形表示，按时间线梳理结果集与环境集的实体因子与环境因子，得到智能医学领域未来技术预测图。

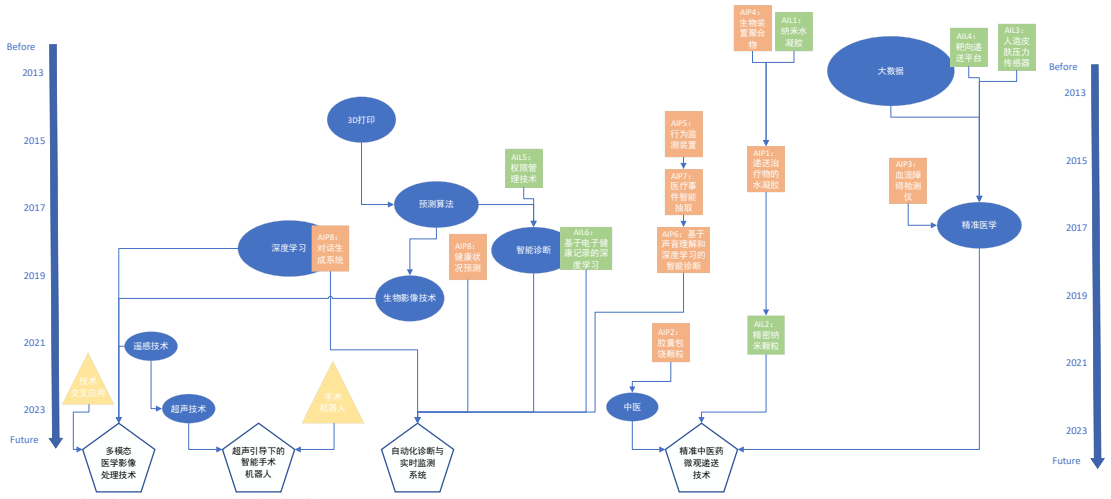


图 7：智能医学领域未来技术预测图

如图 7 所示，本研究认为智能医学领域在未来可能出现的四个技术为：多模态医学影像处理技术、超声引导下的智能手术机器人、自动化诊断与实时监测技术、精准中医药微观递送技术。

PU-285

构建适合我国国情的医疗 AI 技术评估体系的探讨

何达 [1] 李嫣然 [2] 苏锦英 [3]
1. 上海市卫生和健康发展研究中心
2. 华东政法大学
3. 上海中医药大学

伴随着医疗 AI 技术的发展，以深度学习为代表的人工智能技术产品已经广泛运用于医疗领域，但权威或系统性的医疗 AI 技术评估体系尚缺失，虽然存在散在的相关探索，但研究角度多元、评估维度不一、指标各异，这不利于医疗 AI 产品的完善发展和技术的快速创新。本研究基于对传统卫生技术评估体系的总结，综合考虑医疗 AI 技术有别于传统医疗技术的显著特点，为构建一个适合我国国情的医疗 AI 技术评估体系提出前沿性、实用性的相关建议。

1 传统卫生技术评估体系

卫生技术评估（Health technology assessment，HTA）体系的具体评估内容，在全世界范围内还未达成统一的共识。学者徐佩佩、李汶睿、曾力楠根据文献检索以及搜集 26 个国家的政府及学术组织的 83 个官方网站的卫生技术评估指南，发现英国 HTA 的评估内容包括最为完善，并总结出六个一级指标，即技术特性、有效性、安全性、经济性、社会和伦理适应性。

中国国内对于卫生技术评估的具体评估内容也未达成统一的共识。陈洁等认为卫生技术评估包含技术特性、临床安全性、有效性、经济学特性及社会适应性。陈英耀等认为卫生技术评估是对卫生技术的技术性能、安全性、有效性、成本、成本效果、伦理、法律和社会影响等进行系统评价的多学科活动。金春林等认为卫生技术评估内容包括医疗卫生技术的安全性、有效性、成本和效益（效果）及社会影响（政治、法律、伦理及道德等）。肖月等等认为卫生技术评估，尤其是医用设备的核心评估维度，应当包括技术特性、安全性、质量、效果、公平可及性和服务体系效益。

2 医疗 AI 技术的特征及风险因素

医疗 AI 技术区别于传统医疗的显著特点为算法、大数据、算力，算法是核心，大数据与算力是基础。医疗 AI 技术的这些特点给整个医疗过程带来了新的风险。

第一，医疗 AI 算法的不透明度和隐蔽性给医疗过程带来新的风险。医疗 AI 算法本身不在并不在执业医师的控制之下，其决策过程往往缺乏透明度，即“黑箱”效应。此外，算法的隐蔽性也使人们很难直观地发现其问题。如果智能医疗算法存在隐蔽性的风险与问题，可能直到健康损害实际发生这些风险和问题才能被识别出来。结合智能医疗算法具有可复制性且应用广泛的特点，一旦智能医疗算法决策被用于危害健康的活动，其危害可能导致大范围、多点源的公共健康危机。

第二，医疗大数据所涉及的数据安全与隐私问题、数据输入单一问题以及数据质量和完整性问题也给医疗过程带来新的风险。针对安全与隐私问题，医疗大数据包括患者的生理指标、基因信息、疾病史、药物使用历史等，如果这些信息被非法获取和使用，可能会对个人的隐私权造成极大的侵犯。针对数据输入单一问题，目前医疗数据来源相对单一，这些训练数据会导致医疗 AI 产品的广泛适用性与公平性问题。针对数据质量和完整性问题，数据的偏差、噪声或遗漏等问题都可能会对 AI 的决策能力产生不良影响，甚至可能导致误诊。

第三，医疗 AI 的算力所涵盖的数据处理速度和效率同样也会给医疗过程带来风险。医疗 AI 需要处理和分析大量级的数据，且数据处理和分析的复杂度也远高于普通数据，这些数据的规模和复杂度都需要极高的计算能力。如果医疗 AI 的算力不足，可能会导致数据处理速度过慢，这将直接影响到临床决策的效率。在紧急救治、监护、手术辅助等需要快速响应的医疗场景中，如果数据处理速度和效率不足，则将影响救治效果、增加手术风险或者影响病人的治疗和康复。

3 传统卫生技术评估体系的适用困难

算法、大数据依赖、算力及其风险因素属于医疗 AI 技术所独有的技术特性，因此，首先应当考虑传统卫生技术评估体系的评价指标中技术特性的评价指标能否对医疗 AI 技术进行直接适用。但根据文献情况，传统的卫生技术评估体系中各评价维度的评估侧重点均无法对算法、大数据、算力及存在的风险因素进行客观、完整、全面地评估。

4 构建适合我国国情的医疗 AI 技术评估体系的建议

4.1 以传统卫生技术评估体系为基础框架

当前，中国已经建立了一个全面、系统的传统卫生技术评估体系，这套体系通过运用国际公认的科学方法——例如随机对照试验（RCT）、科学文献综述、荟萃分析等，可以对药品和传统医疗设备进行全面而深入的评估。评估的内容不仅包括卫生技术的技术特性、有效性和安全性，还包括这些技术在实际应用中的实施情况和产生的社会经济影响，以及这些技术对公众健康、社区福祉和整个医疗系统运行的影响。

许多医疗 AI 技术及其应用被视为一种新型的医疗设备，传统的卫生技术评估体系可以为医疗 AI 技术的评估提供基础性框架。这种基础框架可以提供对医疗 AI 技术作为医疗技术的一种，其常规的特性。

4.2 新增与医疗 AI 技术特征相关的相关指标

现有的传统卫生技术评估体系无法针对医疗 AI 技术独有的技术特性，即算法、算力、大数据及存在的风险因素进行评估，故而以传统卫生技术评估体系为基础框架，增添新的有关医疗 AI 技术的评估内容势在必行。医疗 AI 技术评估体系有两种可能的构建路径：

其一，不改变传统 HTA 评估维度，对原框架内的相关指标进行扩大解释。比如对安全性进行扩大解释，将数据安全及隐私列为安全性的次级指标和评估内容，使其与使用该医疗设备后发生不良反应或者意外伤害的概率及严重程度的评估内容进行并列。该路径的优点在于可以与以往传统的卫生技术评估体系保持一致性和稳定性，避免整个体系过于庞大和复杂，并且不必为每一种新的情况或情形都构建新的体系。但是其缺点在于，会使某些一级指标的含义，比如安全性的下位概念，变得模糊和不确定，也无法明显地突出构建的新体系的重点。

其二，增添新的评价维度。在保持原有基础框架的基础上，针对算法、大数据、算力及其风险因素，增添新的评价维度，比如新增 AI 技术核心特点评估指标，亦或是 AI 技术的可信赖性评估指标。在该维度下，次级指标展开为算法、大数据、算力，其对应的风险因素为具体的评估视角。该路径的优点在于突出构建的新体系的重点，并且可以通过更加准确地反映所评价对象的特性和表现从而提高评估结果的准确性。缺点是该路径会增加评估体系的复杂性，也有可能有部分新增指标中的评估内容会和原有框架中的评估内容重复，导致评估结果不够准确。

本研究认为，构建新的医疗 AI 技术评估体系最好采取第二种途径，即增添新的一级指标。新的体系中需要包含对于算法、大数据、算力及其风险因素的评估，评估内容繁多，只采用第一种的扩张解释方法会使得大部分一级指标的概念模糊，且无法准确地反映所评价对象的特性和表现。

4.3 参考相关领域的评估指标和方法以调整完善医疗 AI 技术评估体系

医疗 AI 技术评估体系的构建是一个深度结合理论与实践的过程，而不能仅停留在理论设想的阶段。新建评估体系的评价维度、评价指标及其评估内容都是基于对整个医疗 AI 领域深度理解的宏观概念，这些指标需要被进一步的细化与量化，从而可以在实际评估过程中发挥其应有的作用。要实现这一目标，我们必须广泛参考相关领域的评估指标，比如信息技术、机器学习等工程学的评估指标和方法纳入构建过程。例如，算法的透明度是一个相当宏观的评估内容，如果从工程学角度来看，算法的透明度可以包含模型复杂性、可视化、模型行为的一致性等，这些可以被纳入 AI 特有的评价指标中。此外，工程学领域经常使用如精度、复杂性、可扩展性、鲁棒性、效率等具体度量，也可以被纳入 AI 特有的评价指标中。

PU-286

基于患者特异性 CCTA 的左前降支多尺度建模——0D-1D 耦合模型与 0D-3D 耦合模型的比较

刘杉峰、李伍各、吴辉群*
南通大学医学院医学信息学系

引言

近年来，心血管疾病一直是全球范围内的主要健康挑战之一。随着生活水平的提高以及人口老龄化的加速，中国心血管患病率呈持续上升趋势。中国心血管健康与疾病编写组发布的《2022 年中国心血管健康与疾病报告概要》指出我国心血管疾病的发病率以及死亡率仍在升高，在城乡居民疾病死亡构成比中心血管疾病为首位[错误!未找到引用源。]。

深入了解心血管的形态与功能，以及其在疾病发展中的作用，对于提高患者生活质量以及制定更有效的治疗方案至关重要。在医疗领域的快速发展过程中，计算机辅助技术对疾病进行诊断、治疗和预测已成为一项重要的任务，计算流体力学已广泛应用于模拟大血管的血流[错误!未找到引用源。,错误!未找到引用源。]。

心血管降阶模型对于深入理解患者心血管的病理状态以及制定个性化的治疗方案具有重要意义。本研究旨在探讨一种基于患者特异性 CTA 的方法，通过将 0D-1D 耦合模型与 0D-3D 耦合模型进行比较，以更全面地了解不同建模策略对于血管仿真模拟的影响。

动脉血流模拟中，出口边界条件代表下游血管系统，包括小动脉、毛细血管及小静脉。全血液循环范围巨大、建模复杂，忽略下游影响导致 CFD 模型对速度与压力预测不准确，因主动脉间流量分布未知。若出口压力相同或零，流体分流仅由感兴趣区域血管模型分支的阻力决定，忽略了下游血管床的阻力。解三维 CFD 模型需适当边界条件模拟血流，通常用速度入口与压力出口，但在模拟人体动脉波传播时并不完全适用。更佳策略是使用降阶血管模型对下游血管建模，并为动脉求解的三维 CFD 模型提供边界条件，这也是进行 0D-3D 耦合血管模型实验的原因。

心血管系统单尺度模型可以分为三类：0D、1D 以及 3D 血管模型，其中 0D 和 1D 是降阶模型，3D 则是高维模型。0D 也称为集总参数模型[错误!未找到引用源。]（Lumped Parameter,LP），它通常忽略具体的空间信息。Windkessel 模型[错误!未找到引用源。]是目前描述血流动力学最经典的 LP 模型。一维模型也称为分布参数模型，考虑沿血管长度维度的空间信息，将血流、血压、血管截面积等参数描述为相对于时间和血管长度坐标的偏微分方程。因此，0D-1D 模型既具有简单、灵活的优点。考虑 0 维-3 维耦合模型通过在不同层次上进行建模，将心脏和大血管作为 3 维模型，而将小血管和微血管作为 0 维模型，从而在整体和局部上都能够准确地预测血流的动态行为。低维模型为高维模型提供边界条件[错误!未找到引用源。,错误!未找到引用源。]在多尺度建模中具有很大潜力，能够对整个心血管系统进行全面的生物力学分析。

本文将对这两种模型进行详细比较，包括其建模方法、实现过程以及结果解释。通过比较不同模型的优缺点，旨在为动脉系统的生物力学研究提供更全面的视角，为临床实践中疾病的预防和治疗提供更可靠的依据。此外，我们还将探讨这些模型在不同临床应用中的潜在价值，并提出未来研究的方向。

1 材料与方法

1.1 使用软件

Simvascular2022.07.20[错误!未找到引用源。]。

1.2 患者数据及三维重建

数据来源于南通大学附属医院放射科的患者冠脉 CTA 数据，在进行处理影像数据之前，患者的个人信息被匿名化。

从冠脉 CTA 图像中重建三维血管模型的过程包括从图像中沿着感兴趣血管的走行方向标记中心点并连接而成路径线进行分割，感兴趣动脉的横截面积在图像上进行分割[错误!未找到引用源。]。在本实验中我们只研究左前降支以及旁侧的对角支，左前降支存在狭窄。

图 1 给出总体流程图，介绍了血管 3D 模型以及 1D 模型的创建以及与 0D 集总模型的耦合。在耦合模型中，我们需要考虑到心脏左室的影响，描述左心室动力学的方程见式（1.3），该模型用于导出与实体冠状动脉相关联的主动脉流量、压力、左心室压力[错误!未找到引用源。,错误!未找到引用源。]，用于 CFD 分析。

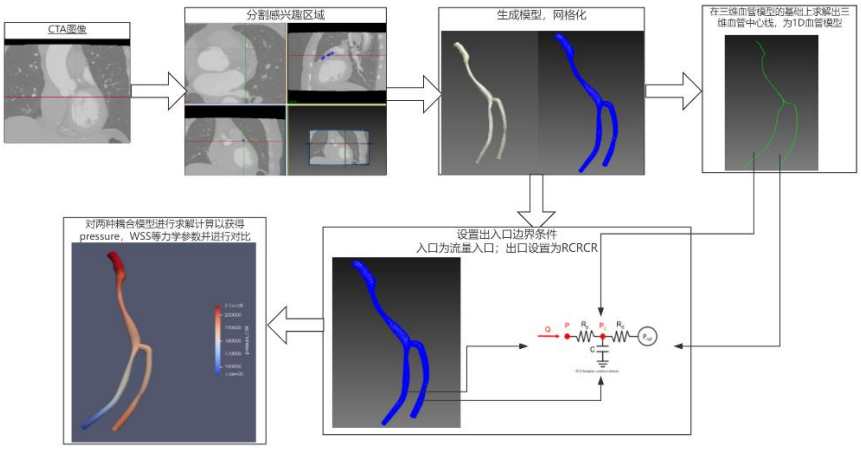


图 15 多尺度耦合左前降支模型的工作流程示意图以及流出边界条件表示

Figure 1 The Workflow Diagram and Effluent Boundary Conditions Representation of the Multi-Scale Coupled Left Anterior Descending Coronary Artery Model

1.3 血管参数

在三维血管 CFD 模型的模拟上，我们假设血管可以被描述为密度为 1060kg/m^3[错误!未找到引用源。]，流体的动态粘度为 0.004dynes/cm^2 的牛顿流体，些材料参数的取值在生理上都是合理的。

1.4 控制方程与边界条件

1.4.1 边界条件

所有流体为刚性表面，在左前降支入口以及其分支出口处指定无滑移的流入和流出 BC，应用流量入口（流速曲线见图 2），集总参数模型应用于血管模型的出口。集总参数模型的 RCRCR 参数（R_a，C_a，R_{amicro}，C_{im}，R_{vmicro}）通过总冠状动脉阻力来计算，总电阻和总电容按其半径之比进行分割分布在出口上[错误!未找到引用源。]，两种耦合模型的边界条件相同。因无真实测量的流量数据，实验中参数取自文献[错误!未找到引用源。,错误!未找到引用源。]。时间步数为 1000，时间步长设置为 0.001，残差控制为 0.001。

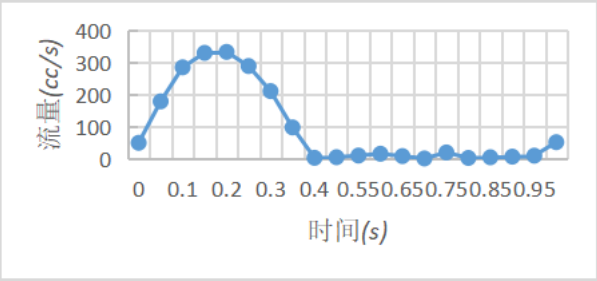


图 2 多尺度血管耦合模型流量入口

Figure 2 Multidimensional vascular coupling model flow inlet

1.4.2 控制方程

(1)0D 模型

0D 模型[错误!未找到引用源。]的构成模块有电感、电阻以及电容，类似于电路。其模型中的流速对应电路中的电流。在心血管建模的背景下，电阻模拟了血流的粘性效应，电容模拟了血管壁的弹性形变能力。这些元素的控制方程见公式(1.1);控制整个 0D 模型的微分代数方程见式(1.2)

$$\Delta P = RQ, Q = C\Delta\dot{P}, \Delta P = L\dot{Q} \tag{1.1}$$

$$E(\boldsymbol{y},t)\cdot\boldsymbol{y} + F(\boldsymbol{y},t)\cdot\boldsymbol{y} + \boldsymbol{c}(\boldsymbol{y},t) = \boldsymbol{0} \tag{1.2}$$

在式(1.1)中 P为压力， R为电阻， Q为流量， C为电容， L为电感；在式(1.2)中 $\boldsymbol{y}$为求解变量的全局向量，包括了 0D 模型流量以及压力， E和 F是它们的关联系数， $\boldsymbol{c}$是常数向量。

左心室动力学方程为:

$$\dot{P}_{lv} = \frac{1}{C_{lv}(t)}\left[U_{mi}\times\frac{P_{la}-P_{lv}}{R_{mi}} - U_{ao}\times\frac{P_{lv}-P_{sa}}{R_{ao}} - \dot{C}_{lv}(t)P_{lv}\right] \tag{1.3}$$

在式(1.3)中， P_{la}、 P_{lv}、 P_{sa}分别为左心房、左心室、全身压力变量， R_{mi}、 R_{ao}为二尖瓣和主动脉瓣的瓣膜阻力。 C_{lv}为左心室顺顺应性， U_{mi}, U_{ao}是模拟各自心脏瓣膜打开或者关闭的功能[错误!未找到引用源。]。

(2)3D 模型

血液被认为是层流、不可压缩的流体。其控制方程包括连续性方程与 Navier-Stokes 方程[错误!未找到引用源。]。连续性方程是指质量定律在流体中的表达方式，可见公式(2.1)。Navier-Stokes 方程是描述粘性不可压缩流体动量守恒的方程，可见公式(2.2)。

$$\nabla\cdot v = 0 \tag{2.1}$$

$$\rho_f\left(\frac{\partial v}{\partial t} + (v - v_g)\cdot\nabla v\right) = -\nabla p + \nabla\cdot T \tag{2.2}$$

在公式（2.2）中v是流体速度矢量， v_g动网格移动速度，p是血液压力， ρ_f是血液密度，T是应力张量。

2 结果

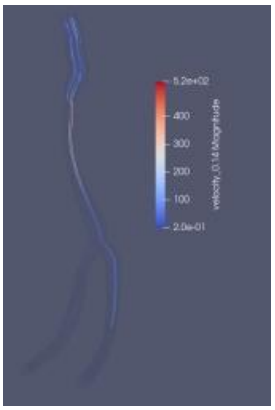
本研究求解出同一狭窄血管模型的两种不同多维度耦合模型，得出多种力学参数以及仿真结果（见图 3），比较了 0D-1D 耦合模型与 0D-3D 耦合模型的性能，两种耦合模型的求解时间见表 1。

表 1 两种耦合模型求解结果对比

耦合方式	0D-1D 耦合模型	0D-3D 耦合模型
时间（*min*）	0.033	137

2.1 血流速度分布分析对比

血管耦合模型在狭窄处附近形成了血流速度的最大值区域，左前降支狭窄区域的血流动力学特征通过峰值时刻的内部血流速度流线图（见图 3）得以呈现。在血液通过狭窄区域之前，左前降支内部的血流呈现均匀分布。然而，当血液到达分叉部位的狭窄区域时，我们观察到了明显的流动分离现象，导致血液的稳定层遭受破坏，形成了复杂的流态分布。0D-1D 耦合模型相比 0D-3D 耦合模型，都是狭窄处出现血流速度变大，但左前降支远端的流线出现缺失。

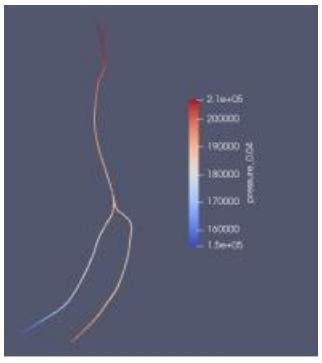


a

图 3 速度流线分布图：a0D-1D 耦合模型；b0D-3D 耦合模型

2.2 血流压力分布分析对比

左前降支压力分布（见图 4）的结果显示，在血管狭窄区域出现压力变高。随着狭窄程度的增加，狭窄区域的壁面压力波动幅度增大，压力值逐渐减小。左前降支狭窄会影响远端血管的血液流动，狭窄区域的低压力导致远端血流动力不足，甚至引发血液回流现象，可能导致脑部供血不足，从而增加脑卒中发生的风险。两种多维度耦合模型相似度较高

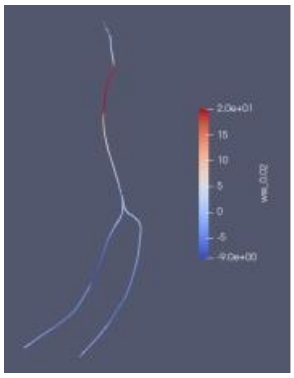


a

图 4 血管压力分布图：a0D-1D 耦合模型；b0D-3D 耦合模型

2.2 壁面剪切应力分布分析对比

图 5 中的壁面切应力分布图显示了左前降支狭窄区域的壁面切应力情况。结果显示，在狭窄处壁面切应力明显增高，高壁面切应力区域的面积也相应扩大。随着左前降支狭窄率的增加，狭窄区域内过高的壁面切应力可能导致血管断裂，同时也增加斑块脱落的风险，从而可能引发缺血性脑卒中的发生。高壁面切应力会导致红细胞聚集，增加斑块破裂和脱落导致血管堵塞的风险。在两种耦合模型的对比中，0D-3D 耦合在局部壁面剪切应力的显示上更加精细。



a

图 4 壁面剪切应力分布图：a0D-1D 耦合模型；b0D-3D 耦合模型

3 总结以及讨论

基于患者特异性 CCTA 的左前降支多尺度建模是冠心病诊疗研究的重要方向之一。本实验比较了 0D-1D 耦合模型和 0D-3D 耦合模型，探讨了它们在模拟患者冠状动脉血流动力学方面的优劣。0D-1D 模型在整体血流动力学模拟方面具有优势，其简洁的数学描述提高了计算效率，可提供准确的心输出量、压力阻力等参数，为临床评估提供可靠依据。

0D-3D 模型在局部血流动力学特性精细建模方面更具优势。通过将 0D 模型与高精度的 3D 模型结合，能更准确地模拟冠状动脉内部细节和血流分布，特别对于狭窄或异常区域的评估更为精确，对于个体化治疗和手术规划至关重要。综上所述，两种模型各有优势，临床可根据具体情况选择应用。期待随着技术发展，它们将相互补充，为冠心病患者提供更可靠的个体化治疗决策支持。

本研究同样存在一些局限性需予以考虑。首先，由于缺乏真实测量边界条件，我们在模型构建中使用了估算值，可能导致模拟结果的一定程度不准确。边界条件的准确性对模拟可靠性至关重要，故未来研究应争取获取更多真实测量数据以提升模型精度。

其次，受技术与资源限制，我们未能完全重建出全冠脉模型，而是以左前降支为重点建模。这或许限制了对整体冠状动脉系统血流动力学特性的全面评估能力。未来研究应致力于全冠脉模型的建立，以提供更全面的冠状动脉血流动力学信息，从而更准确地指导临床决策。

为进一步提升模型的准确性与可靠性，未来研究可通过获取更多真实测量数据以确保准确的边界条件。同时，致力于完全重建全冠脉模型，实现对整个冠状动脉系统的全面评估。随着技术发展，引入先进的计算方法与模拟技术将进一步提升模型精度与稳定性。

此外，在模型构建中，由于缺乏流固耦合实验[错误!未找到引用源。]，我们未能准确地观察血管壁的状态及其在血流作用下的变化。这一方面限制了我们对于动脉病变的全面理解，另一方面也影响了对于治疗方案的精准制定。因此，在未来的研究中，我们将争取获得更为全面的数据支持，以进一步提升模型的精度和可靠性，使其在临床实践中发挥更大的作用。

PU-287

深度学习在冠脉 CTA 影像中的应用研究进展

李瑾、孟艺菲、周彤彤、栾晨曦、吴辉群*
南通大学医学院医学信息学系

1 引言

心血管疾病（Cardiovascular Disease，CVD）一直是人类健康的严重威胁，它被认为是慢性疾病中的重要一员，也是许多国家高致命率的疾病之一[错误!未找到引用源。,错误!未找到引用源。]。CVD 通常在早期没有明显的临床症状，这导致大多数患者在就医时已经进入了中晚期，此时的治疗变得更加复杂，预后也相对较差。因此，早期筛查和诊断对于防治心血管疾病至关重要。

目前，临床上的诊断主要依赖于影像学检查，冠脉 CTA（Coronary CT angiography, CCTA）检查在鉴别心血管疾病方面具有较高的准确性、灵敏度和特异性，为临床治疗提供了有力的诊断依据[错误!未找到引用源。,错误!未找到引用源。]。通过 CCTA 全面、准确、综合地评估冠状动脉的解剖和功能特征，有助于筛查高危患者，从而减少不良事件的发生，这对于心血管疾病的预防和治疗具有重要意义。然而，传统的 CCTA 图像后处理和解读仍然完全依赖于医生的人工操作。这种方式不仅工作量大，过程繁琐，而且耗时较长。此外，评估冠脉狭窄和斑块性质的判断容易受到医生的临床经验和水平的限制，从而降低了 CCTA 的特异性和阳性预测值[错误!未找到引用源。]。因此，引入基于深度学习(deep learning, DL)的图像处理技术已成为现代医学影像学领域的不可避免趋势。这些 DL 方法能够协助医生更快速、更准确地进行疾病诊断。

本文将综述 DL 在 CCTA 图像诊断领域的研究进展，并探讨这些方法在临床实践中的潜在价值。这将为心血管疾病的早期诊断和管理提供新的洞察和方向，有望进一步提高患者的生活质量和预后。

2 DL 在 CCTA 血管分割中的应用

冠状动脉血管的准确分割对于冠状动脉的研究和诊断至关重要。同时，这项技术还为其他冠脉相关应用提供了重要的信息基础。然而，对于细小的冠状动脉，如何实现精确且高效的分割一直是一项挑战。U-Net [错误!未找到引用源。]是一种广泛应用于图像分割的卷积神经网络架构，在冠状动脉 CTA 图像中，U-Net 被成功应用于识别和分割冠脉血管，有效地将血管从周围组织中分离出来。一些研究者进一步改进了 U-Net 架构，Sadid 等[错误!未找到引用源。]提出的 CoronarySegNet，它引入了通道感知注意力块和深度残差块，以增强特征图的独立性，最终达到 0.78 的 DICE 指标。冯等[错误!未找到引用源。]则使用了带有注意力机制的 U-Net 网络，成功地实现了冠状动脉的分割，取得了 0.0213 的损失函数、0.9626 的准确率和 0.9999 的 Dice 指标，尤其在 CTA 图像的切片级别上表现出色。尽管已经有多种方法支持了冠状动脉的分割，由于冠脉血管的复杂性和周边组织的影响，分割时边界的清晰度也是一个难题。邵等[错误!未找到引用源。]通过增强边界特征来改善冠脉边界的清晰度，提高网络对分形特征的描述能力，解决了特征不明显的问题。另一方面，刘等[错误!未找到引用源。]提出的方法通过分水岭算法对预处理后的 CTA 图像进行冠状动脉子区域的提取，然后使用区域生长方法进行区域合并，最终实现了冠状动脉的三维分割，有效地消除了周围组织的干扰，实现了精确的分割结果。除了直接使用卷积神经网络进行分割，一些研究还探索了自动提取冠状动脉中心线的方法，并根据冠状动脉的结构构建数据结构来辅助分割。Kong 等[错误!未找到引用源。]采用预分割提取中心线，然后构建树结构数据，并在 FCN 结构下引入 TreeConvGRU 模块对构建的树结构数据进行分割处理。此外，基于 DL 回归方法的 Polar1DMLP 模型结合了中心线先验信息，取得了较好的分割效果[错误!未找到引用源。]，Du 等[错误!未找到引用源。]则开发了一种多目标聚类和环形模型引导跟踪方法，用于自动分割冠状动脉树。

DL 在冠脉 CTA 图像的血管分割和中心线提取中具有巨大的潜力，这些方法为医疗领域带来了前所未有的机会，有望提高医学影像的自动化程度和准确性，从而改善心血管疾病的早期诊断和治疗。

3 DL 在冠脉血管狭窄和斑块辅助诊断中的应用

CCTA 可用于评估冠状动脉狭窄的严重程度，在临床实践中，准确判断冠状动脉狭窄程度对于诊断和治疗冠心病至关重要。DL 技术已经在辅助诊断冠状动脉狭窄方面展现出巨大的潜力，它能

够提供与专家医生判断高度一致的狭窄程度测量结果，并且显著缩短了诊断所需的时间[错误!未找到引用源。-错误!未找到引用源。]。韩等[错误!未找到引用源。]采用卷积神经网络进行训练，成功地实现了对斑块类型和冠状动脉狭窄程度的识别，在区分不同程度的冠状动脉狭窄时，CCTA-AI 的诊断准确率优于传统的人工方法。此外，冠状动脉斑块的性质对于预测疾病也具有重要影响，其中冠状动脉钙化被认为是最强的心血管事件预测因素之一[错误!未找到引用源。]。Fischer 等[错误!未找到引用源。]受到传统的钙化斑块检测方法启发，提出了一种利用冠状动脉中心线来约束钙化斑块位置的方法。除了单独关注狭窄或斑块的研究，Zreik 等[错误!未找到引用源。]使用 3D 卷积神经网络来同时检测和分类冠状动脉斑块和狭窄程度。他们的方法基于多尺度卷积神经网络进行 DL 分析，实现了对患者是否存在功能性冠状动脉管腔狭窄的无创性判断，以及对斑块的自动检测和表征。与以往的方法不同，这种方法只需要输入 CTA 图像中的冠状动脉中心线，从而大大简化了分析过程。在最新的研究中，Li 等[错误!未找到引用源。]开发了一个计算机辅助诊断（computer Aided Diagnosis, CAD）自动诊断任务，包括血管分割和识别任务。他们首先通过血管分割来提取感兴趣的区域，然后进行识别任务，通过 3DNet 模型输出提供≥50%的冠状动脉狭窄的诊断结果。这种方法显著提高了 CCTA 图像分析的效率，并减轻了放射科医生的工作负担。此外，冠状动脉钙化积分（Coronary Artery Calcification Score, CACS）也是诊断心血管疾病的重要指标，可用于预测不良心血管事件。一些研究人员采用卷积神经网络对 CCTA 图像进行 CAC 自动评分，甚至实现了风险分类。有研究尝试利用卷积网络对 CCTA 图像进行 CAC 自动评分[错误!未找到引用源。]，MU 等[错误!未找到引用源。]利用 DL 进行自动 CAC 评分的同时更是实现了风险分类。

DL 在冠状动脉 CTA 图像分析中为早期诊断和治疗心血管疾病提供了强大的工具，提高了诊断准确性、效率和个体化治疗的可能性，这些方法有望改善患者的生活质量并降低不良心血管事件的风险。

4 总结与展望

DL 在冠状动脉 CTA 诊断方面展现出显著优势，它能够自动识别血管，进行血管分段的标识，并准确评估血管狭窄和斑块性质。这一技术的引入大幅度缩短了图像后处理的时间，提高了诊断效率，让影像诊断医生从繁琐的处理工作中解放出来，更加专注于提高诊断的准确性，有助于心血管疾病的诊治标准化。此外，DL 能够同时完成更高精度的图像分割、分类和识别工作，不仅有力地克服了影像医生测量所带来的主观性问题，还提供了多任务并行的方式，能够带来更迅速、更深层、更标准的量化分析。

然而，目前基于 DL 的诊断仍存在一些缺点，这需要更多的数据用于学习和改进，同时也需要确保分段标注的准确性，以使人工智能在临床应用中更加可靠。DL 在心血管领域中的潜力巨大，但仍需要进行更深层次的疾病智能化和精准化研究。如何更好地利用 DL 有效提高基于 CCTA 对心血管疾病的诊断效能，同时优化 DL 模型以适应临床需求将是研究趋势。积累更多数据集、整合人工智能到临床实践、推动个性化治疗、处理监管和道德问题将是未来的挑战和机遇。

PU-288

智能医学的发展与应用

吕倩

郑州大学第一附属医院

智能医学是指应用人工智能技术对医学数据进行挖掘和分析从而实现对疾病进行诊断、治疗和预防的医学分支。智能医学的发展历程可以追溯到 20 世纪 80 年代，当时人工智能技术开始应用于医学领域。随着计算机技术的不断进步和大数据技术的广泛应用，智能医学得到了快速发展随着医学技术的不断进步和人工智能的快速发展，智能医学成为医学领域的一个新的发展方向。智能医学结合了人工智能技术和医学领域的专业知识，旨在提高医学诊断、治疗和预防的水平和效率。随着人工智能技术的不断进步和医学数据的不断积累，智能医学将会得到更加广泛的应用。

1. 医学数据

挖掘医学数据挖掘[2]利用是一种技术，通过挖掘大量医学数据，提取其中有价值的部分，以支持医学实践的多个方面[3]。

1.1 增加产生新知识的速度：通过计算机技术分析电子病历等包含的非结构化数据，能够更好地进行自动化的数据采集。

1.2 帮助知识传播：对于大量的研究，知识的转化可能非常困难，大部分临床医生难以跟上最新的证据来指导临床实践。通过分析现有的电子病历，可以产生一个仪表板来指导临床决策。

1.3 支持个性化医疗计划：通过整合系统生物学与电子病历数据，医学数据挖掘可以为个性化医疗计划转变为临床实践提供机会。

1.4 转化卫生保健知识：通过直接向病人提供信息，让病人在治疗过程中发挥更积极的作用。

1.5 医学数据挖掘是智能医学的另一个重要应用方向。通过对医学数据进行挖掘和分析，可以发现疾病的规律和特征，为疾病预防和治疗提供重要依据。例如，通过对大规模病例数据进行挖掘，可以识别出心脏病、癌症等疾病的危险因素。

2. 医学智能辅助治疗

2.1 医疗影像分析[4]是智能化和自动化的诊断和治疗的重要领域之一。

通过图像处理技术和机器学习算法，可以对医学影像进行自动分析，如 X 光片、CT 扫描和 MRI 等，以检测异常病变，判光片、CT 扫描和 MRI 等，以检测异常病变，判断病情的严重程度和发现肿瘤等病变。这种技术提高了诊断的准确性和效率，减少了医生阅读影像的时间和工作量。

2.2 智能药物管理

智能药物管理是智能化和自动化的诊断和治疗的一个重要应用。通过 AI 算法和机器学习技术，可以对药物的作用和副作用进行预测和分析，为医生选择最佳治疗方案提供参考。同时，智能药物管理系统[5]还可以对药品的使用情况进行实时监控和管理，确保药品的安全性和有效性。

2.3 病理学辅助诊断

病理学辅助诊断是医学智能辅助治疗的另一个重要方向。通过利用人工智能算法对病理图像进行分析，可以自动识别和分类病变组织，提高病理诊断的准确性和一致性。例如，一些智能辅助诊断系统可以通过分析病理图像，自动识别肿瘤细胞的组织类型和分级，为医生提供参考意见。

2.4 智能决策支持系统

智能决策支持系统是医学智能辅助治疗的核心组成部分。通过利用人工智能算法和数据分析技术，可以对患者的病情和治疗方案进行综合分析，为医生提供最佳的决策建议。例如，一些智能决策支持系统可以根据患者的病情和治疗历史，推荐合适的治疗方案，帮助医生制定个性化的治疗方案。

2.5 病人数据实时监控

病人数据实时监控是智能化和自动化的诊断和治疗的另一个重要方面。通过物联网技术和大数据分析，可以对病人的生命体征、生理参数等数据进行实时监测和分析，及时发现病情的变化和异常情况，为及时干预和治疗提供支持。同时，病人数据实时监控还可以为医疗数据的共享和管理提供便利，实现了跨科室、跨机构的医疗协作。

3. 个性化诊疗

3.1 随着科技的不断发展，个性化诊疗[6]的智能化和自动化的诊断和治疗已经成为医疗领域的重要趋势。这些技术的广泛应用正在改变着医疗行业，提高诊断和治疗的准确性和效率，同时也为医生和病人带来了许多便利。

3.2 个性化诊疗是智能医学的一个新兴应用方向。通过利用人工智能技术对患者的基因、代谢等信息进行分析，可以为患者提供更加个性化的诊疗方案。例如，基于基因检测的个性化诊疗方案，可以根据患者的基因信息，自动生成最适合的治疗方案更加智能化和自动化的诊断和治疗。

3.3 电子健康记录是指将个体的健康数据进行电子化记录和管理，方便医生和患者随时随地查看和使用。电子健康记录可以包含个体的基本信息、病史、检查报告、诊断结果等方面的信息为医生和患者提供全面的健康数据支持。

3.4 具体而言，电子健康记录可以通过电子病历、健康管理 APP 等方式进行记录和管理。同时，电子健康记录还可以通过对数据的分析和挖掘，为个体提供个性化的健康指导和建议。此外，

电子健康记录还可以实现跨机构、跨地区的医疗信息共享促进医疗资源的优化配置和社会健康管理的协同发展。

3.5 个性化诊疗是多学科结合的结果，主要依赖于大数据挖掘和人工智能技术。通过对医疗数据的分析和挖掘，可以获取患者的疾病信息、生活习惯、遗传信息等，从而实现个体化治疗。同时，人工智能技术可以为个性化诊疗提供技术支持，如机器学习算法可以实现疾病的预测和诊断，自动化系统可以实现治疗方案的制定和执行。个性化诊疗在癌症治疗、遗传疾病等领域已经得到了广泛的应用，并取得了良好的效果。然而，个性化诊疗也面临着一些挑战，如数据质量、隐私保护、技术标准等问题。

4. 智能预防

4.1 使用智能设备进行健康监测，及时发现潜在的健康问题。

智能设备如可穿戴设备、移动设备等可以实时监测个体的生理指标，如心率、血压，血糖，及时发现潜在的健康问题。个体可以定期将这些数据输入智能系统，以获取个性化的健康管理建议。

4.2 预测性健康管理

预测性健康管理是智能化和自动化的诊断和治疗的重要领域之一。通过基因检测、大数据分析和机器学习技术，可以对疾病的风险进行预测和管理，为个体化健康管理和精准预防提供支持。这种技术可以帮助人们更好地了解自己的健康状况，采取积极措施预防和治疗疾病。

4.3 医学健康智能预防

医学健康智能预防主要依赖于生物标志物检测、健康数据采集和分析、风险评估和预测等方法。其中，生物标志物检测是关键环节，可以实现对疾病早期的发现和预防；健康数据采集和分析则是通过各种传感器和移动设备获取个体的健康数据，并通过大数据技术和人工智能算法进行分析和预测。医学健康智能预防在心血管疾病、糖尿病、癌症等慢性疾病管理中具有显著的应用效果，可以有效地提高患者的生存质量和降低医疗成本。

5. 远程医疗

5.1 远程诊断与会诊

远程诊断与会诊[7]是智能远程医疗的重要应用之一。通过视频会议、医学影像传输等技术手段，医生可以在远程对患者进行诊断和治疗。这种服务模式不仅可以提高医疗效率，降低医疗成本，还可以缓解医疗资源紧缺的问题。例如，偏远地区的患者可以通过远程诊断与会诊服务，接受城市中心医院专家的诊断和治疗，避免了长途奔波和高额的医疗费用。

5.2 远程监护

远程监护[7]是利用各种传感器和监控设备，对患者的健康状况进行实时监测。这些设备可以包括心电监护仪、血压计、血糖仪等，可以将患者的生理数据实时传输给医生，以便及时发现病情的变化。例如，心脏病患者可以在家中接受远程监护，一旦出现异常情况，医生可以及时采取措施，避免病情恶化。

5.3 远程医学培训

远程医学培训是通过网络技术，为医生提供远程的医学教育和培训服务。这种服务模式可以打破地域限制，让医生随时随地接受高质量的医学培训，提高医疗水平。同时，远程医学培训还可以实现医学资源的共享，降低培训成本，提高培训效率。例如，偏远地区的医生可以通过远程医学培训，接受城市中心医院专家的医学教育和培训，提高自身的医疗水平。

5.4 远程医疗信息服务

远程医疗信息服务是通过移动设备、电话等方式，为患者提供远程的医疗咨询和服务。这种服务模式可以方便患者获取专业的医疗建议和服务，提高医疗服务的可及性。例如，患者可以通过电话咨询医生关于药物的使用方法、疾病的预防等问题，方便快捷地获取专业的医疗建议。

然而，远程智能医疗也存在一些挑战和问题如信息技术的安全性和稳定性问题、医疗服务的质量和标准化问题等。同时，由于医疗服务的地域性和个体性差异，远程智能医疗也难以完全替代传统的医疗服务。

结论：

未来，智能医学将会更加注重个性化诊疗和预防医学。通过对个体进行精准的评估和预测可以为个体提供更加个性化的健康管理和预防方案。智能医学是医学领域的一个新的发展方向，具有广

阔的应用前景和重要的现实意义。未来，随着人工智能技术的不断进步和医学数据的不断积累，智能医学将会得到更加广泛的应用，为人类健康事业做出更加重要的贡献。

PU-289

医疗数据创新延伸——“数据胶囊”的建设思考

杨雪*
哈尔滨医科大学附属肿瘤医院

中共中央、国务院印发的《“健康中国 2030”规划纲要》中，对提升医疗服务水平和质量提出了明确的要求：建立与国际接轨、体现中国特色的医疗质量管理与控制体系，基本健全覆盖主要的国家、省、市三级医疗质量控制组织，推出一批国际化标准规范。建设医疗质量管理与控制信息化平台，实现全行业全方位精准、实时管理与控制，持续改进医疗质量和医疗安全，提升医疗服务同质化程度，再住院率、抗菌药物使用率等主要医疗服务质量指标达到或接近世界先进水平。“数据胶囊”是满足医院信息化建设与应用所需基本功能的重要平台，包含惠民服务、医疗业务、医疗管理、运营管理。医疗协同、数据应用、移动医疗、信息安全、信息平台基础等 9 类 122 项功能，能够解决全国医院信息化建设中面临的系统名称不一致、系统功能不一致等现象，为逐步实现全国医院信息化标准规范奠定了坚实基础。

一、数据胶囊的概念

数据胶囊是一种新型的医疗数据管理与应用模式，旨在通过高效、安全的方式对医疗数据进行收集、处理和分析，从而为医疗行业提供更加精准、个性化的服务。数据胶囊可以将患者的医疗数据进行集中管理，包括病历、检验报告、影像资料等[1]。通过构建统一的数据平台，实现数据的标准化、结构化和互联互通，从而为医疗工作者提供更加便捷、高效的工作环境。在数据胶囊的支持下，患者仅凭身份信息就能查看包含检查检验结果、用药差异对比及分析在内的各项数据信息，为患者带来更加精准、个性化的医疗服务。同时，数据胶囊收集了患者的生理、生活等多个方面的数据，有助于医生全面了解患者的健康状况，为患者制定更精准的治疗方案。

二、数据胶囊的技术架构与实现

（一）数据胶囊的技术架构

数据胶囊的技术架构主要包括四个部分：数据采集模块、数据处理模块、数据存储模块和数据访问模块。

其一，数据采集模块。负责从各种医疗设备中采集医疗数据，包括病历、检查报告、诊断结果等。其二，数据处理模块。负责对采集到的医疗数据进行清洗、转换和标准化处理，使其符合数据胶囊的存储要求[2]。其三，数据存储模块。负责将处理后的医疗数据存储到数据胶囊中，保证数据的安全性和完整性。其四，数据访问模块。负责提供数据胶囊的访问接口，使医疗人员可以方便地查询和使用数据胶囊中的医疗数据。

（二）数据胶囊的关键技术

1.人工智能技术

人工智能技术在数据胶囊的建设中起到了至关重要的作用。通过使用机器学习和深度学习算法，可以从大量的医疗数据中提取有价值的信息，帮助医生做出更准确的诊断和治疗决策。同时，人工智能技术还可以用于自动化数据分析和处理，进一步提高数据胶囊的效率和准确性。

2.大数据技术

大数据技术在医疗数据创新延伸中扮演着举足轻重的角色。数据胶囊需要处理和分析海量的医疗数据，包括患者的病历、诊断结果、治疗方案等。通过使用大数据技术，可以快速地处理这些数据，提取出对医疗保健有意义的信息[3]。同时，大数据技术还可以用于预测疾病趋势和优化医疗资源分配，从而提高医疗系统的整体效率。

3.云计算技术

云计算技术为数据胶囊提供了强大的计算能力和存储能力。通过云计算平台，可以实现医疗数据的快速传输、存储和分析。同时，云计算技术还可以帮助医疗机构之间实现数据共享，促进医疗研究的合作和进展。

4.物联网技术

物联网技术在医疗数据创新延伸中的应用，使得医疗数据的获取和传输变得更加便捷。通过物联网设备，如智能健康监测器和远程医疗设备，可以实时收集患者的健康数据，并将其传输至数据胶囊。这有助于医生实时了解患者的状况，为患者提供更加精准和个性化的医疗服务。

（三）数据胶囊的实现流程

1.数据采集

数据胶囊的建设首先需要进行数据采集。这一步骤涉及对医疗机构内部各种医疗数据的收集，包括患者的基本信息、病历、诊断结果、治疗方案、药物处方等。数据采集的方式有多种，可以通过医疗机构的电子病历系统、医疗影像系统等自动获取，也可以通过人工录入的方式进行[4]。在这一步骤中，要保证数据的准确性和完整性，以便后续的数据存储和利用。

2.数据存储与管理

数据胶囊的第二步是数据存储与管理。收集的医疗数据需要进行安全、高效地存储和管理，以便后续的数据利用。数据存储可以采用分布式存储系统，通过数据的切分和分布式处理，提高数据的存储效率和安全性。数据管理则需要建立完善的数据管理体系，包括数据的分类、归档、检索等功能，以便于数据的查找和使用。

3.数据利用

数据胶囊的最终目标是实现医疗数据的创新利用。通过对医疗数据的深度分析和挖掘，可以实现对疾病的早期预测、个性化治疗方案的推荐、医疗资源的优化配置等功能。在这一步骤中，需要利用大数据分析、人工智能等技术，对医疗数据进行深入地分析和挖掘，以实现数据的增值。

三、数据胶囊在智慧医疗系统中的应用

（一）医疗资源的优化配置

医疗资源优化配置是智慧医疗系统中的重要环节，而数据胶囊技术在实现这一目标方面具有巨大潜力。

数据胶囊有助于实现医疗资源的合理分配。通过分析患者的就诊历史、病历资料、治疗方案等信息，数据胶囊可以挖掘出医疗资源的使用规律，进而为医疗机构提供有针对性的资源配置建议。例如，数据胶囊可以提示医疗机构在某段时间内需要增加某个科室的医生数量，或者建议将某个设备分配到更加需要的科室等。数据胶囊还可以支持医疗资源的精细化管理[5]。通过对医疗设备、药品、耗材等资源的使用情况进行实时监控和分析，数据胶囊可以帮助医疗机构发现资源浪费、损耗等问题，并采取相应的措施进行优化。这不仅可以降低医疗机构的运营成本，还有助于提高资源利用效率，为患者提供更好的医疗服务。

（二）患者管理与健康干预

患者管理是医疗系统中的关键环节，通过对患者的全面了解和有效管理，可以提高医疗服务的质量和效率。在智慧医疗系统中，数据胶囊作为一种新型的数据管理工具，可以帮助医疗机构更好地实现患者管理。

首先，通过数据胶囊，医疗系统可以全面收集患者的个人信息、病历资料、诊断结果、治疗方案等数据。这些数据可以帮助医疗机构建立完整的患者档案，为患者提供个性化的医疗服务。同时，医疗机构可以通过对患者数据的分析，发现患者的潜在需求和健康问题，提前采取干预措施，降低患者的患病风险。其次，数据胶囊可以实现患者健康数据的实时监测和管理。在智慧医疗系统中，患者可以佩戴健康监测设备，如智能手环、智能血压计等，实时收集自己的健康数据[6]。这些数据会通过数据胶囊实时传输到医疗系统中，便于医生及时了解患者的健康状况，对患者的健康问题进行及时干预。最后，数据胶囊还可以用于患者的健康干预。通过对患者数据的分析，医疗系统可以制定个性化的健康干预方案，如饮食建议、运动计划、药物治疗等。这些干预方案可以通过数据胶囊实时推送给患者，帮助患者更好地管理自己的健康。

（三）医疗服务优化

随着医疗数据创新的不断延伸，数据胶囊作为一种新型的医疗数据管理工具，正逐渐在智慧医疗系统中发挥重要作用。

首先，数据胶囊可以提升医疗服务的个性化水平。通过对患者的医疗数据进行整合和分析，数据胶囊能够更加准确地评估患者的健康状况，从而为患者提供更加精准、个性化的医疗建议。例如，数据胶囊可以分析患者的病历数据和药物过敏史，为患者推荐最适合的药物和治疗方案。数据胶囊还可以通过分析患者的健康数据，为患者提供预防疾病发生的生活建议，从而帮助患者更好地管理自己的健康。其次，数据胶囊可以优化医疗服务的连续性。通过将患者的医疗数据进行集中管理，数据胶囊能够为医生提供全面、连续的患者信息，从而使医疗服务更加连贯。例如，当患者在不同的医疗机构就诊时，数据胶囊可以确保医生能够随时查看患者的完整病历，从而避免重复检查和误诊[7]。此外，数据胶囊还可以实现患者与医生之间的实时沟通，使患者能够更加便捷地获取医疗服务。最后，数据胶囊可以提高医疗服务的效率。通过对医疗数据的分析和处理，数据胶囊能够帮助医生快速识别患者的病情，从而节省医生的诊断时间。例如，数据胶囊可以通过图像识别技术辅助医生识别病症，从而提高癌症等疾病的诊断速度。此外，数据胶囊还可以通过自动化处理技术，实现医疗数据的自动整理和归档，从而减轻医生的工作负担，提高医疗服务效率。

（四）智能诊断与治疗方案的推荐

医疗体系中，医生往往需要根据患者的症状、病史、检查结果等多种信息进行综合分析，才能给出诊断和治疗方案。然而，由于医疗数据的复杂性和医生经验的局限性，误诊和误治的情况仍然时有发生。数据胶囊作为一种新型的医疗数据处理工具，能够有效地整合患者的医疗数据，并利用人工智能技术进行智能诊断和治疗方案推荐。

首先，数据胶囊可以对患者的医疗数据进行深度挖掘和分析，发现其中的规律和特征。通过运用机器学习和自然语言处理等技术，帮助医生精准诊断患者的病情，并给出准确的诊断结果，提高诊断的准确性的同时降低误诊率，还可以大幅缩短医生的诊断时间，提高医疗效率[8]。其次，数据胶囊可以根据患者的诊断结果，推荐合适的治疗方案。数据胶囊中存储了大量的医学知识，包括各种疾病的治疗方法、药物副作用、手术风险等信息。通过对这些知识的深度分析，数据胶囊可以结合患者的具体病情，推荐最为合适的治疗方案。这不仅可以避免因医生经验不足而导致的误治，还可以为医生提供更多的治疗选择，提高治疗效果。最后，数据胶囊还可以通过持续学习和更新，不断提高自身的诊断和治疗推荐能力。随着医疗数据的不断积累和医学研究的不断进展，数据胶囊可以实时地更新自身的知识库，以适应医学领域的发展。这将使得数据胶囊在智能诊断和治疗方案推荐方面的能力越来越强大，为患者提供更加精准和个性化的医疗服务。

四、结语

总的来说，数据胶囊作为一种新型的医疗数据管理与应用模式，具有很大的发展潜力和广泛的应用前景，既充满了机遇，也面临着挑战。随着我国医疗信息化建设的深入推进，我们有理由相信，数据胶囊将会在优化医疗资源配置、提升医疗服务质量、推动医疗科技创新等方面发挥越来越重要的作用。然而，要实现这一目标，需要进一步加强医疗数据的基础研究和应用研究，推动数据胶囊的技术创新和产业发展。只有通过持续的研究和创新，我们才能真正发挥出数据胶囊的潜力，推动我国医疗事业的持续发展。

PU-290

运用 PDCA 循环提升儿科门诊辅助诊断决策支持系统应用效果

尚于娟 葛小玲*
复旦大学附属儿科医院

目的 探索人工智能产品在临床实际应用难点及推广模式，提升医生对儿科门诊辅助诊断决策支持系统（Clinical Decision Support System，CDSS）接受度及系统应用效能。

方法 基于已建成的儿科门诊辅助诊断决策支持系统，在某单中心儿童医院门诊业务流程中实际推广应用，采用 PDCA 循环作为质量改进管理工具。选取 2022 年 6 月-8 月系统使用数据作为干预前

组，对医生应用推广及病历质控实施 PDCA 循环管理。选取实施管理后的 2022 年 9 月-11 月数据作为干预后组；通过采集客观使用数据进行系统应用综合效能评价。

结果 儿科门诊辅助诊断决策支持系统在该单中心门急诊上线后，发热门诊、皮肤科、口腔科、内分泌科、呼吸科使用量高，占全院使用量的 85%。在医生层面，住院医师使用量占比近 90%，青年医生接受度高。经 PDCA 循环工具干预后，推荐应用科室中医生使用率提升 5.26%，智能诊断采纳率提升 3.72%，医生接受度明显提升。全院诊断结果一致率提升 1.93%，疑似无依据诊断率降低 1.8%，病历质量呈现不断改进趋势。

结论 针对人工智能产品在临床实际应用管理工作中实施 PDCA 管理工具，可有效提升医生接受度及系统应用效能，增强医生提升书写高质量病历的意识，本研究为系统未来在基层医疗机构应用提供推广模式参考。

PU-291

智能医学应用现状和创新发展

丁立旻 [1,2,3] 渠鸿竹 [1,2,3] 方向东 [1,2,3]

1. 中国科学院北京基因组研究所（国家生物信息中心） 中国科学院基因组科学与信息重点实验室 北京 100101
2. 中国科学院大学 北京 100049
3. 基因组与精准医学检测技术北京市重点实验室 北京 100101

1 引言

在当代科技飞速发展的背景下，大数据、云计算和人工智能等新兴技术为多个学科带来了深远的影响。尤其是在医学、生命科学和工程科学领域，这些技术的应用推动了跨学科的交叉融合。随之而来，传统的医学模式经历了从工程医学到智能医学的转变。智能医学，一个结合了人工智能与医学的领域，旨在通过大数据、人工智能和以及其他先进技术实现更加精确的个体化医疗服务，为患者提供更加高效和精准的疾病预防、诊断、治疗和管理。随着这些技术的进一步发展和完善，智能医学有望引领医学界进入一个全新的时代，其中人工智能不仅作为工具，更是为医学实践提供重要的策略和方向。在此背景下，本综述将探讨智能医学的应用现状和发展趋势，以及它为现代医学带来的机遇和挑战。

2 智能医学发展现状

随着技术的飞速发展，智能医学已逐渐崭露头角，为传统医疗带来前所未有的创新。利用人工智能等技术，现今的医学实践更为个性化，能精准地调整药物剂量和治疗方案。智能穿戴设备和生物传感器的结合为患者提供了实时、连续的健康数据，而大型语言模型如 ChatGPT 则展示了 AI 在医学沟通和文档处理上的巨大潜力。这些进展不仅提高了治疗效果，还带动了医疗行业朝着更加智能、便捷的方向发展。

2.1 人工智能与疾病诊断

人工智能（AI）主要应用于医学影像识别和组学数据的挖掘，从而用于疾病的诊断筛查。以癌症为例，肿瘤的检测主要涉及到计算机辅助检测（CADe），旨在识别放射图像中的关注对象[1]。在这一模式下，将具有可疑成像特征的区域高亮显示，并呈现给读取者。CADe 已被用作辅助工具，帮助检测低剂量 CT 筛查中可能错过的癌症[2]，改善 MRI 中脑转移的检测时间[3]，还能在乳腺筛查中定位微钙化团块，作为早期乳腺癌的指示器[4]。

研究者已经鉴定出有许多不同类型癌症的潜在生物标志物，儒如利用液体活检分析循环肿瘤 DNA，为监测疾病的进展或回退提供了一种途径，还允许近乎实时地监测目标癌症突变的出现。然

而，这些标志物仍需要在大量人群中得到验证。多组学数据结合如放射组学等数据时，可以克服传统病理学的局限性，提高疾病诊断准确性。放射组学特征在定义癌症亚型方面显示出前景[5,6]，并可能在肿瘤分类的精准医学背景下，作为主要组学数据的补充或替代[7,8]。此外，生活方式、环境因素等其他因素也可以整合到分析中，为其增添新的维度。由于机器学习单独不能解决因果推断的问题[9]，因此需要临床医生、生物学家和计算分析师的综合努力。其中，由医疗人员收集的主要样本由生物学家进行实验性分析，然后由分析师进行计算分析。经生物学家验证后，提取的信息再次反馈给医疗人员。在这样的工作流程中，为了多组学分析的有效结果，跨学科知识的流畅分享是至关重要的。

2.2 人工智能与药物开发

在现代智能医学的背景下，人工智能为药物研发过程带来了革命性的突破和创新方向。起初，人工智能被应用于分子生物学和化学领域，旨在预测二级蛋白结构的排序[10]。随后，神经网络因其识别模式的能力而被药物工业所采纳，如预测癌症药物的作用机制[11]。

随着技术的进步，基于机器学习和 AI 的完整分子设计方法已展现出解决问题和应对新情境的效能，如随机森林（RF）[11]和支持向量机（SVM）[12]。这些算法被用于准确预测分子的重要特征。而深度学习技术则进一步深化了对毒性、生物活性和分子结构间复杂关联的研究[13]。同时，机器学习已增强了多种决策和预测方法，如 QSAR[14]。其中，有监督的学习方法如决策树、SVM 等，在药物发现的多个应用中都有所体现。无监督学习算法如主成分分析（PCA）则被用于虚拟筛选、数据可视化和分子模式识别[14]。

在肿瘤领域，机器学习的强大预测能力已得到了充分体现。例如，研究人员通过对近 1000 种癌症细胞系进行药物筛选，评估基因表达、DNA 甲基化和体细胞突变如何影响药物反应[15]。深度学习，作为机器学习的一个分支，使用了多层处理的人工神经网络，为药物的溶解性、毒性和药物再利用提供了更准确的预测[16]。随着技术的不断进步，AI 在药物发现和设计领域的应用展现出巨大的潜力。这不仅为药物研发提供了全新的方法和视角，也为未来的医学领域带来了巨大的机遇。

2.3 人工智能与个性化医疗

在现代医学实践中，确保疗效和安全性是至关重要的。然而，如何为每位患者精确地选择和调整药物剂量长久以来一直是一个挑战。特别是在考虑联合疗法时，存在着几乎无限的剂量参数空间，使得精确的治疗方案选择变得更加困难[17]。虽然有人认为正确的药物剂量主要是为了提高疗效和安全性，但实际上，正确的剂量选择对于最初选择用于治疗的药物也有直接影响。

人工智能和机器学习技术为这一领域的发展带来了新的可能性。通过这些技术，研究者现在能够在连续的模式下确定药物给药参数，甚至能够从大量的候选化合物中同时确定最佳的药物和剂量，以实现最佳的联合治疗疗效和安全性[18–20]。例如，AI 基于的二次型表型优化平台（QPOP）能够在无需基于协同效应的预测或复杂的疾病机制信息的情况下，从肿瘤药物候选物中确定新的药物组合[19]。此外，QPOP 还被用于从体外到临床前开发阶段重新优化药物组合的剂量参数，而这些人工智能发现的组合在很大程度上优于临床护理标准[19]。

值得注意的是，在联合治疗中，最佳的药物剂量比例和多药治疗的最佳药物选择往往被忽视，尤其是在采用剂量递增和最大耐受剂量时[17,19,21,22]。然而，最近的前瞻性试验表明，基于表型个性化医疗（PPM）的新技术平台成功地个性化了移植后的免疫抑制[23]。此外，还对儿童急性淋巴细胞性白血病的联合疗法方案、对维持 HIV 患者病毒载量不可检测的改进联合疗法、以及免疫抑制剂量的优化进行了优化分析[24]。

2.4 智能穿戴设备

在智能医学的快速进展中，穿戴式医疗设备与生物传感器的融合与发展日益受到重视。这些新型设备通过高度精确地捕捉生物标志物，能够为个性化治疗提供持续且强大的生物数据读出，使得医疗治疗能够更加精准和即时[25]。例如，经过批准的 MiniMed 670G（Medtronic）作为一个封闭回路的人工胰腺，便能够通过算法驱动的方式为 1 型糖尿病患者控制血糖水平，实际应用中也已显示出能有效增加目标血糖范围内的时间，同时降低高低血糖的发生率[26]。

而如今，已有技术将电极嵌入到硅微纤维管中，用于血压、血管刚度和心率的监测，这为动脉

硬化、静脉溃疡的非侵入性诊断和监测提供了可能，甚至可以与绷带整合，为各种应用提供便捷性[27]。通过这样的非侵入性多路监测技术，未来可能能够确定个性化的生物标志物特征[28]。

穿戴设备的数据采集不仅使患者及其家庭能够主动参与自己的健康管理，还避免了频繁地前往医生处使用传统方法进行生物标志物分析。随着这些数据宝库的持续增长，我们可以期待其不仅对个性化医学做出重大贡献，更是为精准医学的实践打开了新的治疗思路和选择，为众多患者带来希望。

2.5 大语言模型为基础的智能医疗助手

随着技术的发展，AI-driven 的聊天机器人在医学应用领域的价值日益凸显。聊天机器人这一概念并非新鲜事物。它是一个利用 AI 和自然语言处理技术来理解问题并自动回应，模拟人类对话的计算机程序。早在 1964 到 1966 年，Joseph Weizenbaum 就在麻省理工学院的人工智能实验室开发了一个非常早期的医学聊天机器人——ELIZA。然而，时至今日，随着计算能力的增强和几乎无限的互联网数据，语言模型已经拥有数千亿的参数，这使得它们具有生成新文本的能力。例如，由 OpenAI 训练的 ChatGPT 在 2022 年 11 月向公众介绍后，展示了 AI 驱动的机器与人交互的新方式[29]。

新一代的聊天机器人在医学应用中展现了其作为助手和指导的潜力，但也存在着一些关键的问题需要解决。ChatGPT 的开发者在其发布时描述了许多这样的问题，特别是当它用于医学领域时，需要特别注意。目前，这类聊天机器人可以帮助解决医学文档问题，并回答有助于鉴别诊断的关键问题。但要判断所提供的答案是否基于合适的事实是困难的[30]。因此，临床医生需要校对聊天机器人的工作成果，就像他们需要校对自己口述的临床笔记。然而，这种校对可能超出了使用者的专业范围。对于一个没有临床经验的用户，他可能很难区分事实与虚构。因此，使用聊天机器人进行诊断或推荐治疗是一个既有巨大潜力又备受关注的应用领域[30]。

3 总结

随着大数据、云计算和人工智能等新兴技术的飞速发展，智能医学在医疗领域掀起了一场革命。这一领域的应用潜力巨大，已经为患者提供了更加精准和高效的医疗服务。从医学图像分析到药物开发，从疾病分类到生物标志物鉴定，人工智能展现出了卓越的能力。多组学数据的整合和穿戴式医疗设备的普及进一步推动了个性化医疗的实现。而基于大语言模型的智能医疗助手也为医学实践带来了全新的可能性。然而，随之而来的挑战需要我们继续不断探索，跨学科合作和知识分享将在智能医学的道路上扮演关键角色。智能医学有望引领医学界进入一个全新的时代，为患者的健康和医疗保健带来更多希望和机遇。

PU-292

基于“互联网+”背景下国内外慢病管理研究热点的比较分析

孙雪超、鲁兴胜、赵玉虹*
中国医科大学健康管理学院

1 引言

慢性非传染性疾病（non-communicable chronic diseases，NCDS）简称“慢病”，是对一类起病过程隐匿、潜伏时间长、病程长且缓慢、病因复杂、缺乏可靠传染性生物学病因证据等特点的疾病的统称，常见的慢病主要包括糖尿病、心脑血管疾病、慢性阻塞性肺疾病（chronic obstructive pulmonary disease，COPD）和恶性肿瘤等。世卫组织数据表明，2019 年全球慢病平均死亡率为 73.6%[2]。在我国，慢病导致死亡人数占全国总死亡人数的 88%，慢病疾病负担占我国疾病总负担的 70%[3]，这无疑加重了国家、社会以及个人的经济负担，慢病已成为当前威胁人类生命健康的重要公共卫生问题。因此，慢病管理及防控得到了更多重视。

随着新技术、新模式、新产业不断涌现，2015 年 3 月 5 日十二届全国人大三次会议上，李克强总理在政府工作报告中首次提出“互联网+”概念，这个新概念的提出为破解慢病管理难题提供

了新的方法路径。在“互联网+”时代大背景下，慢病管理的相关研究也得到迅速发展。在国外，基于“互联网+”的慢病管理模式已广泛应用于包括高血压、糖尿病、癌症等的多种慢性病之中，其中包括开发血糖自我管理应用和智能化临床信息系统等慢病管理系统。近些年来，我国“健康中国”战略部署同样推动了国内慢病管理新模式的发展。从 2017 年起，国务院就发布了“健康中国”战略的纲领性文件，也是首个国家级计划《中国慢性病防治中长期规划（2017—2025年）》，规划中提出要促进互联网与健康产业融合，发展智慧健康产业，探索慢病健康管理服务新模式。推动移动互联网、云计算、大数据、物联网与健康相关产业的深度融合，充分利用信息技术丰富慢性病防治手段和工作内容，从而实现慢性病的全面防治[4]。

目前，国内外有关“互联网+”背景下慢病管理的研究既存在一致性也存在差异性，为了解国内外慢病管理领域在“互联网+”背景下的研究现状和前沿动态。本研究以 WOS 核心数据库和万方数据库为基础对其近十年发表的相关研究进行文献计量学分析和可视化分析。探索近十年该领域国内外的研究热点、发展趋势以及差别，从而为我国未来该领域相关科研人员及医疗人员选择研究方向提供依据。

2 材料与方法

2.1 数据来源

本研究以 Web of Science™ 核心集合和万方数据库作为主要的数据来源，检索互联网背景下慢病管理相关的文献，检索策略见表 1。

表 3 文献检索策略

检索策略	国外	国内
检索数据库	Web of Science™ 核心集合（SCI）	万方数据库
检索式	(ALL=(chronic disease OR chronic illness OR Chronic non-communicable diseases OR chronic illnesses OR Disease,chronic OR illness,Chronic)) AND ALL=(intelligence OR artificial intelligence OR Intelligent Medicine OR smart medicine)	（（主题：（慢性病） or 主题：（慢病） or 主题：（慢性非传染性疾病）） and （主题：（智能） or 主题：（互联网）））
时间跨度	2014 年 1 月 1 日至 2023 年 9 月 12 日	2014 年 1 月 1 日至 2023 年 9 月 12 日
检索时间	2023 年 9 月 12 日	2023 年 9 月 12 日
纳入条件	文献类型：article	排除会议摘要、学位论文、政策性文献等
检索结果	2629 篇	2443 篇

2.2 研究方法和步骤

①Web of Science™ 核心集合文献以纯文本格式导出相关信息，万方数据库文献以 refworks 格式导出相关信息。每条记录均包含作者、关键词、摘要等信息。②使用 BICOMB2.0 分别提取中、英文文献发表的年代、关键词，进行可视化分析并对比；③对关键词进行略缩词还原、同/近义词合并等预处理，根据普赖斯计算公式 $M = 0.749\sqrt{N_{max}}$（其中 M 为高频词阈值，N_{max} 表示区间学术论文被引频次最高值）[5]，计算高频词阈值，得到中文文献高频词阈值为 14，英文文献高频词阈值为 13。使用 VOSviewer 软件进行数据分析，在软件设置中分别导入国内文献和国外文献数据，中文关键词频次设置为“14”，英文关键词频次设置为“13”，统计方式选择“Full counting”，生成可视化图谱。

3 结果

3.1 国内外年度发文数量趋势

在近十年间，“互联网+”在慢病管理研究领域中的英文发文量总体呈逐年递增的趋势，中文发文量则在 2021 年达到高峰（2014 年为 401 篇），近两年略有下降。截至 2023-9-12，2023 年英文发文量为 459 篇，中文发文量为 213 篇，见图 1。

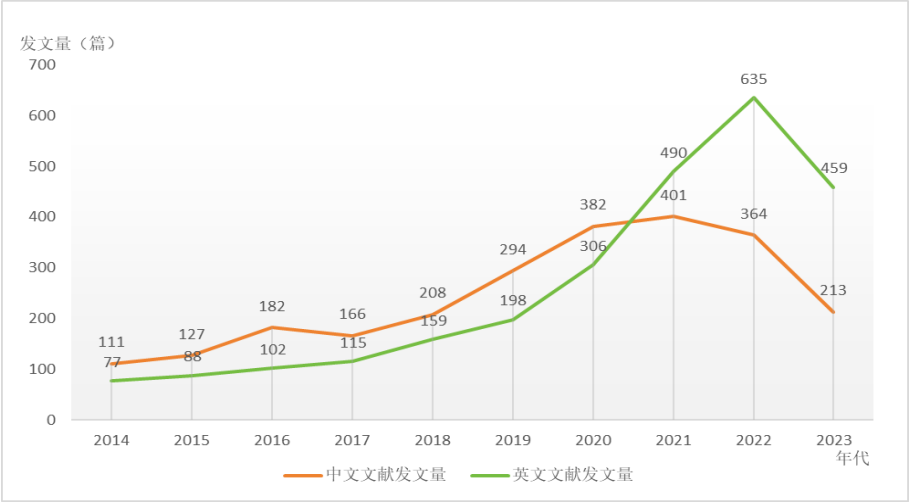


图 1 “互联网+”背景下慢病管理研究领域国内外年度发文量数量趋势

3.2 国内基于“互联网+”背景下慢病管理领域的高频词及聚类结果

应用 VOSviewer 软件对中文文献中出现频次≥14 次的关键词进行关键词聚类分析，共生成 6 个聚类，见图 2。每种颜色代表一个类别，每一类别代表一个研究热点。

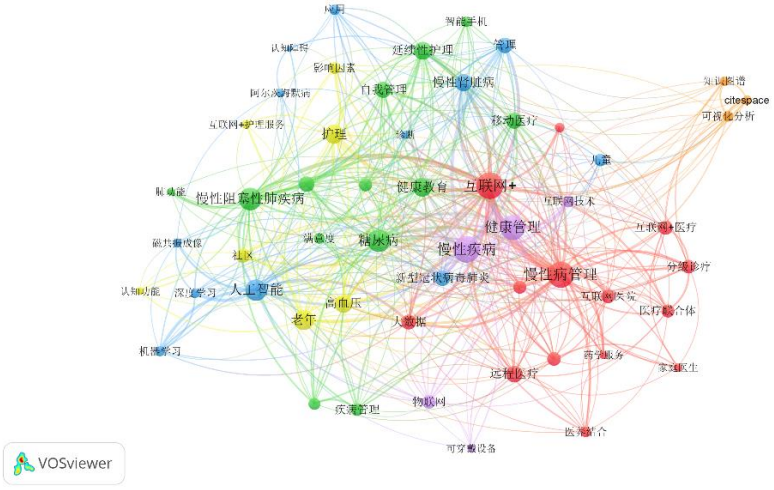


图 2 国内“互联网+”背景下慢病管理研究领域关键词聚类可视化图谱

中文总关键词数 4226 个，出现频次≥14 次的关键词共计 54 个，出现频次排在前 10 位的关键词依次为互联网+（370 次）、慢性疾病（254 次）、慢性病管理（213 次）、人工智能（121 次）、健康管理（117 次）、慢性阻塞性肺疾病（117 次）、糖尿病（92 次）、老年（66 次）、护理（58 次）、大数据（55 次）、高血压（54 次）。

3.3 国外基于“互联网+”背景下慢病管理领域的高频词及聚类结果

应用 VOSviewer 软件对英文文献中出现频次≥13 次的关键词进行关键词聚类分析，共生成 6 个聚类，见图 3。

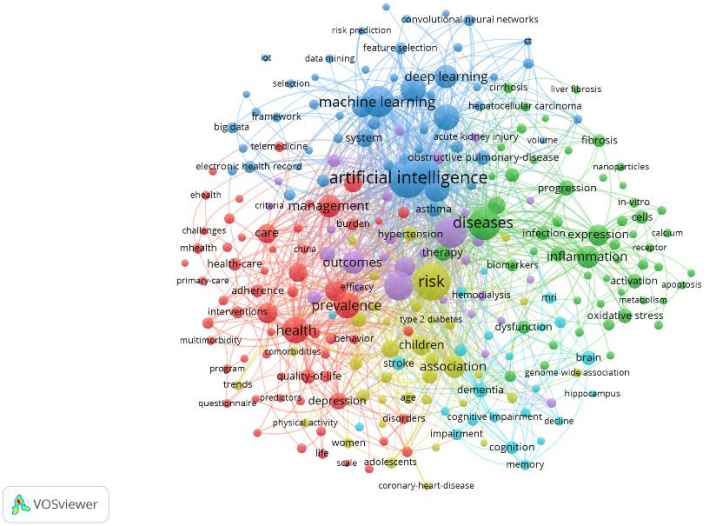


图 3 国外“互联网+”背景下慢病管理研究领域关键词聚类可视化图谱

英文总关键词数 7821 个，出现频次≥13 次的关键词共计 258 个，出现频次排在前 10 位的关键词依次为 artificial intelligence（332 次）、risk（302 次）、diseases（233 次）、chronic kidney disease（213 次）、machine learning（213 次）、mortality（189 次）、health（156 次）、prevalence（144 次）、diagnosis（144 次）、classification（144 次）、copd（138 次）、models（135 次）、deep learning（123 次）。

3 讨论

本研究通过文献计量学分析方法，进行了年度发文数量分析、高频关键词分析、高频关键词聚类的可视化分析，全面展示、梳理了国内外“互联网+”背景下慢病管理领域的相关研究成果，有助于帮助研究者了解国内外该领域研究的现状、趋势和热点。

3.1 国内外年度发文数量趋势分析

通过中文文献发文量的趋势来看，自 2014 年以来，“互联网”背景下的慢病管理领域的年度发文量整体呈逐年上升趋势，但在 2015 年，发文量激增，这与当年出台的《全国医疗卫生服务体系规划纲要（2015—2020 年）》以及《国务院关于积极推进“互联网+”行动的指导意见》密切相关，国家从战略层面高度重视“互联网+”与健康产业的深度融合。在互联网的时代背景下，要将新兴的信息技术与慢病防治的老问题结合起来，建立慢病管理新模式，以提高我国的慢病管理水平。 从 2015—2020 年，发文量稳步上升。2020 年突发的全球性新型冠状病毒肺炎（新冠），对全世界都造成了巨大冲击，但新冠疫情也推动了互联网慢病管理的纵深发展，借助互联网技术的慢病管理新模式愈发成熟，我国互联网背景下慢病管理领域的文献发文量略有下降。

通过英文文献发文量的趋势来看，2014—2020 年，国外相关研究论文量增幅低于国内，主要原因在于国外早已针对各国自身需求完成设计并实施多种慢病管理模式，例如适应发达国家的慢性病照护模式（CCM），针对资源紧缺的中低收入国家的创新型慢性病照护框架（ICCCF），以及以社区和家庭为依托的慢性病自我管理计划（CDSMP），其慢病管理体系已经趋于成熟[6]。根据《The Lancet》2020 年 10 月 15 日发表报告显示，从 2020 年起，受新冠疫情的影响，患有慢性病的人群出现重症死亡的风险较高，各国为抵御 COVID-19 大流行的影响采取很多紧急措施，这使得国外文章发文量呈现快速增长态势，且增长速度大于国内。

3.2 国内基于“互联网+”背景下慢病管理研究领域的高频词聚类结果分析

国内高频关键词聚类主要聚成六类，第一类包含互联网+、慢性病管理、互联网医院、医疗联合体、分级诊疗、家庭医生等词。这一类别是我国现如今的研究重点，“互联网 +”慢病管理模式为患者提供优质、连续、全程的慢病管理服务。为了解决传统慢病管理模式带来的诸多问题，“互联网 + 医疗”模式迅速发展，通过智慧医疗助力慢病管理，提高慢病管理服务的工作效率。推行分级诊疗制度是我国慢病管理模式发展的新方向，在新医改的大背景下，应在重视居民的需求的基础上，借助医保等政策引导，推行建立分级诊疗制度，这是我国医疗卫生事业的必经之路。分级诊疗

制度在推进过程中，切实落实到地，有助于提升我国慢病管理的工作质量[7]；第二类包含糖尿病、慢性阻塞性肺疾病、慢性心力衰竭、肺功能；智能手机、移动医疗；以及自我管理、生活质量、满意度等词。其中包含类别有常见的慢性病、智能手机中的移动医疗应用以及通过自我管理模式管控慢性病。这一类可总结为借助日常化手段预防和控制常见慢性病；第三类包含人工智能、深度学习、机器学习、磁共振成像、应用、诊断等词。该类专注于应用人工智能、机器学习等方法诊断并预测慢性病；第四类包含老年、护理服务、社区、认知功能、高血压等词。聚焦于通过社区卫生服务，为社区内老年慢病患者做好分层分类的长期照护，重点关注老年人常发易发的慢性病，比如高血压、认知障碍等，进行重点筛查，有效防治；第五类包含互联网技术、物联网、可穿戴设备、健康管理等词。借助可穿戴设备、物联网、IOT 等互联网新兴技术，实现数据存储、共享与智能辅助决策，构建了新型医生在线慢性病防控培训平台。同时，语音机器人可为慢性病患者提供健康随访、效果评估和统计分析等健康管理服务[8]；第六类包含 Citespace、可视化分析、知识图谱。是对互联网背景下慢病管理的研究方向和发展趋势进行研究，便于梳理该领域浩如烟海的文献。

3.3 国外基于“互联网+”背景下慢病管理研究领域的高频词聚类结果分析

国外高频关键词聚类主要聚成六类，第一类包含 health、prevalence、management、impact、care、validation、chronic disease 等词。主要研究的是慢病管理相关护理干预，慢性病的护理干预方式呈现多元化的趋势，涵盖延续性护理、健康管理服务、药学服务等多学科联合等，探究依从性、自我效能、生活质量等领域，这与聚类高频词分布基本一致。例如 Nguyen,V 和 Park,J 的研究 [9]；第二类包含 diseases、inflammation、expression、cancer、progression 等词。主要研究的是慢性病的发病机制和病因学研究。既能通过多组学检测方法研究微观病因学[10]，又通过研究血压、血脂、慢性压力等慢性病相关的危险因素探究与慢性病的关联[11]；第三类包含 artificial intelligence、machine learning、deep learning、diagnosis、classification、models 等词。聚焦点在于人工智能技术。借助可穿戴医疗设备、物联网、大数据和云计算等新兴技术以及深度学习、机器学习等方法建模，对慢性病防治控进行研究[12]；第四类包含 risk、children、association、obesity 等词。主要研究的是有关于青少年体重超重带来的风险。据估计，由于超重和肥胖、不健康饮食和缺乏身体活动的增加导致很多青少年患有 2 型糖尿病。在过去几年里，由于超重带来的青少年糖尿病患病率持续上升，最为严重的是，不仅是血糖升高，还有并发症和其他影响，如智力下降、心脏病发作、中风等，因此关注青少年的身体健康就显得尤为重要[13]；第五类包含 chronic kidney disease、mortality、outcomes、epidemiology 等词。关注点在于对慢性肾病的研究，其中包括慢性肾病流行病学研究[14]和治疗方案；第六类包含 brain、dementia、memory、cognitive function 等词。主要研究的是认知功能障碍引发的疾病[15]。

4 讨论

随着科技的发展，人工智能正被广泛应用于医疗健康领域，为慢性病的管理提供了新思路。目前，国内外在“互联网+”背景下对慢病管理的研究集中在人工智能技术在慢病管理的应用[16, 17]。同时，由于医疗水平的提高，人均寿命得到了延长，加上人口增长率[18]的下降，我国正处于老龄化社会的加速期，老年人的人口比重快速增加。因此，我国尤为关注老年人慢性病的预防、治疗和照护相关研究。而经过多年的探索和实践，西方发达国家在慢病管理领域已取得诸多显著成效[19-21]，所以国外更专注于慢性非传染性疾病流行病学领域的研究。

本文采用可视化方法对 2014—2023 年国内外“互联网+”背景下慢病管理领域发表的相关文献进行分析。根据对高频关键词聚类的结果和关键词分布特征分别总结了国内外“互联网+”背景下慢病管理领域的研究热点和趋势，并比较了两者的一致性和差异性。

随着互联网新兴技术与医疗领域的深度融合，我国慢性病信息化管理水平正在逐步提高，逐渐形成了一套独具中国特色的慢病管理模式。然而，国内在慢病管理领域的研究和实践与国际先进水平还存在差距，一些问题亟待解决。比如，国家政策在基层的落实不够到位，社区卫生资源明显不足，疾病预防重要性认识不够等问题依然存在[22]。为了进一步完善我国的慢病管理服务体系，我们需要继续推进深化医疗体制改革，在学习借鉴国外成功经验的同时，结合中国国情，调动社会各界力量，共同探索一条基于“互联网+”时代背景下符合中国实际的慢病管理之路。

PU-293

合理用药系统在医院中的应用与管理

周千惠[1] 彭晨辉[2]

[1]首都医科大学附属北京儿童医院

[2]黑龙江医院哈尔滨医科大学附属第六医院

1 引言

随着社会持续发展、科技的不断进步，医院信息化建设已经成为了现代医疗行业的重要趋势。医院信息化是指医院通过数字化手段，先对数据进行收集、并储存到对应服务器、再传输到相应系统，达到数据互通的目的；满足所有授权用户的功能需求，提升用户的工作效率，使医院提高服务质量。而合理用药系统则是医院信息化中的一个重要组成部分，它利用信息技术，对医生开出的药品医嘱进行合理化监测和评估，以保障患者的用药安全和合理。

本文将详细阐述合理用药系统的概念及重要性，并通过具体案例分析其在临床治疗中的实际应用意义，最后对未来发展趋势进行展望。

2 合理用药系统的介绍

由于全球范围内存在的药品剂量开立错误、不结合患者实际情况开立药品，以及患者未能正确使用药物的问题数量逐步增多，使得药物不合理使用已成为危害人类健康的全球性问题。随着计算机技术和互联网的发展，合理用药系统逐渐得到了重视，自 2000 年起，一些地区的卫生行政部门开始推广合理用药信息化建设，通过政策引导和资金扶持等手段以及结合医院临床经验，推动合理用药系统的完善。

合理用药的定义最早是在 1985 年，由世界卫生组织（WHO）在会议上提出[1]，它是以现代计算机技术水平为基础，结合医药学及其相关学科知识，并以服务患者为核心的系统。合理用药系统以数据库中的大量数据为依托，从药品的用量、用法、适应症等数据入手分析，确保药品的药物剂量、用法妥当，辅助医生为用药对象开立适合的药品，预防药品相关不良事件的发生，用最小的治疗风险让患者获得最大的治疗效果。

3 合理用药系统的实施与管理

3.1 合理用药系统的实施

2009 年 8 月 18 日，医院、药品相关监督、管理部门发布三份关于药物管理办法的文件并于同年 11 月份出台配套措施，相继又颁发了药物临床使用说明等文件，标志着国家基本药物数据化、合理化，文件的颁发推进合理用药系统的实施。结合本院合理用药系统部署前的准备工作，总结出以下几个方面实施时需注意的事项：

3.1.1 合理用药系统的人员权限分配

系统部署后，需由信息科人员与药学部负责人取得联系，及时对使用合理用药系统的药师进行权限分配。主要分为主任（最高权限人员）、负责人（可登录所属模块，可进行权限调整）、药师（只可登录负责模块）、医师（可使用系统提交审核）。

3.1.2 开展系统的使用培训

需由医院信息科人员或系统相关讲师对药师、医务人员进行系统培训，培训后及时为老师们解答听讲过程中不理解的问题。

3.1.3 完善药品基本数据

由药学部负责人组织安排科员进行药品基本数据的完善，包括但不限于药品的名称、剂量、基本单位。

3.1.4 定期审核

加强对医务人员的药品监管，实施定期的药物使用评估和审核，及时与临床进行沟通，发现问题及时进行改正。

3.2 合理用药系统的管理

3.2.1 建立并完善合理用药系统功能

医院信息化已广泛投入使用，越来越多的医务工作者把目光看向合理用药系统。合理用药系统的使用，是信息技术与临床优秀的衍生作品，把工医进行结合，通过院内药师在系统中已维护的药品基本数据、提前设置的药品限制，实现对临床医生开医嘱进行审查，进而可以及时发现潜在的不规范用药问题。

3.2.2 合理用药监管委员会的建立

合理用药监管委员会是由医学、药学、卫生管理等方面的专家组成的。该委员会的建立可制定和更新临床药学使用指南，及时与临床进行沟通，以提高临床合理用药准确性。

3.2.3 建立临床药师制度

接受过专业培训，具有一定的药学专业知识和以及积累相当的临床经验，并通过资格认证的临床药师，主要负责审核医生提交的待审核医嘱，协助医生进行合理用药。

3.2.4 合理用药政策制定

政府部门应制定和完善合理用药政策，加强对医疗机构合理用药情况的监督和管理[2]，强化药物不良反应监测和通报，促进合理用药水平的提高[3]。

3.2.5 宣讲制度

可由合理用药监管委员会定期组织院内医务人员培训，加强医护人员对临床药学知识的了解，减少临床出错率。

4 合理用药系统的功能与应用场景

4.1 合理用药系统的基本功能

4.1.1 限制用药

用于设置机构、科室或个人限制用药。当药学部负责人在系统中设置用药权限后，机构下科室或医生开出限制药品时，将给予对应配置级别的提示。

4.1.2 药品排他互斥

设置药品排他互斥。当医生开出配置互斥的药品时，将给予对应配置级别的提示。

4.1.3 药品剂量限制管理

用于设置药品剂量限制，系统中药库自带说明书的特定剂量限制，某些特殊剂量限制可自行配置，配置后当医生开出超剂量药品时，将给予对应配置级别的提示。

4.1.4 诊断限制用药

用于设置诊断限制用药。当医生开出诊断限制的药品时，将给予对应配置级别的提示。

4.1.5 门诊处方审核

用于审核门诊处方，当医生提交一张待审门诊处方时，系统将根据定向算法计算当前在线药师的审核工作量权重，再将门诊处方进行派发，以均衡各药师之间的工作压力，为避免药师审核压力过大。药师会结合患者诊断以及药品相关数据对医生提交的处方进行审核，通过或拒绝。

4.1.6 住院医嘱审核

用于审核住院医嘱，当医生提交一份待审住院医嘱时，系统将根据药学科室负责人在系统中的设置，结合患者的诊断，对医嘱数据进行判断，如不符合规则，由科室医生通过系统向药师发出申

请。

4.2 合理用药系统在医院中的应用场景

下面通过一则具体案例来论证合理用药系统在临床治疗中的实际应用意义。

【案例】

一位患者因皮肤病就诊于某医院，看诊后，医生根据患者的病情，选择了适合的药物进行治疗。然而，在医生开具处方过程中，加大了患者的用药量，出现过度医疗的行为。

合理用药系统自动拦截，无法签发药品，给出相应错误提示：“某药已超过最大使用剂量，预审发生错误，请提交药师审核”。如需继续开立医嘱，请输入申请理由，由药师进行审核，审核通过才可签发医嘱。

在这个案例中，医生没有遵循合理用药系统的原则，但系统及时拦截，降低药物不良事件的发生率。通过这个案例，我们可以看出合理用药系统在临床治疗中的实际应用意义。

5 合理用药系统的未来展望

5.1 合理用药系统的未来发展趋势

5.1.1 数字化

随着人工智能、大数据等技术的快速发展，未来的合理用药系统将会更加数字化。通过基因检测、生物标志物检测等手段，可以更好地了解患者的病因和发病机制，为患者提供更加精准的药物治疗方案，提高治疗效果和减少不良反应。

5.1.2 智能化

随着医疗技术的进步，未来的合理用药系统将会更加智能化[4]。通过智能化用药系统，可以实现对患者数据的实时监测，根据患者的病情和个体差异，自动推荐合适的药物和剂量，从而提高用药的准确性和安全性。

5.1.3 监管化

合理用药系统的未来发展会受到监管和政策的影响。相关监管单位会根据政府的医疗政策、药品政策等对各使用单位的合理用药系统的实施进行定期审查，监管机构需要对药品供应链、处方开具、药品销售等进行全面监管，确保药品的安全和合理使用。

5.1.4 大众化

未来的合理用药系统可趋近生活化、大众化，与群众紧密关联。使用者可通过系统模糊搜索到与自己症状相类似的病症，再根据系统提供的资料（例如：药物适应症、成分等药物信息或药学相关教育材料），帮助患者更好地了解自己的病情和用药方案，提高患者对药物治疗的依从性和参与度。

5.1.5 跨学科化

合理用药系统的未来发展需要多学科的合作。药物治疗不仅涉及到医学领域，还涉及到生物学、化学、心理学等多个学科。跨学科合作可以帮助解决合理用药系统中的各种问题，例如药物相互作用、心理因素对药物治疗效果的影响等。

5.1.6 数据共享化

合理用药系统的发展需要实现数据的共享和安全。通过建立统一的数据标准和接口，实现不同医疗机构之间的数据共享，有助于提高医疗服务的质量和效率。同时，要加强数据安全保护，防止数据泄露和滥用。

5.2 针对合理用药系统在医院中的改进和优化建议

合理用药系统不仅适用于医院，也建议在未来应用于社区医疗中心、药店、家庭医疗等场所。通过普及合理用药知识，提高医务人员的用药水平，可以有效降低患者的用药风险，减少医疗费

用，保障公众健康。

6 结语

合理用药系统在医院中的应用与管理，对于提升医疗质量、降低药物浪费、保障患者安全、优化资源配置、促进临床决策、提升治疗效果以及实施电子处方管理等方面都具有重要意义。通过不断加强合理用药系统的建设和应用，医院可以为患者提供更加优质、高效的医疗服务。

PU-294

智慧医疗安全防护之道——安全“脱管”不如安全托管

苏子仁[1] 刘阳[2,3] 曹港华[1]
1.海南医学院第二附属医院
2.中南大学湘雅医学院附属海口医院
3.海口市人民医院

1. 背景

随着信息技术的飞速发展，信息化和网络化建设已经成为我国社会经济活动的基础保障，网络安全建设工作上升为国家战略高度[1]。在医疗健康领域，随着医院信息化的普及和深入，信息系统成为医院运营核心技术支撑，特别是深化医药卫生体制改革以来，医院业务从院内走向院外，从封闭走向开放，医院信息系统面临着前所未有的新风险。

与此同时，医院合规监管不断增强，监管压力越来越大。特别是《网络安全法》实施以来，医院行业面临的信息安全合规和监管压力极大增强。在国家新型网络空间安全形势和背景下，国家卫建委相继出台了《关于印发全国医院信息化建设标准与规范（试行）的通知》[2]、《关于促进“互联网+医疗健康”发展的意见》[3]、《国家健康医疗大数据标准、安全和服务管理办法（试行）》[4]、《关于印发全国基层医疗卫生机构信息化建设标准与规范（试行）的通知》[5]、《关于落实卫生健康行业网络信息与数据安全责任的通知》、《中华人民共和国基本医疗卫生与健康促进法》[6]等政策文件，全力推进我国医疗行业网络安全工作的的部署。并且连续开展医院网络安全检查工作，目的是通过常态化的安全检查，对医院信息系统进行有效验证，以查促建、以查促管、以查促改、以查促防。

面对黑客攻击、未经授权入侵和其它威胁，医院许多的系统管理人员常常忙于解决日常运维过程中出现的琐碎问题，以维持信息系统的持续可用，而没有更多精力定期检查信息系统中是否存在安全隐患、跟踪并获得相应的漏洞补丁、及时修复信息系统安全问题。

根据调研，很多医疗机构网络安全管理工作处于传统的安全管理模式——“脱管”状态[7]。一方面，由于大部分医院由于无法定期对系统进行安全风险评估，识别资产存在的安全隐患，导致“重建设、轻运营”尤为明显。另一方面，医疗机构在网络安全建设中，虽然采购并部署了合规的安全设备，但是，在部署后缺乏专人运维管理，导致发生安全事件时不能及时发现和动态防护。总体而言，受限于人力、技术资源等隐私，医院安全运营经验不足，安全“脱管”普遍存在。

2. 信息系统安全建设挑战

2.1 资产底数不清晰

（1）不知道网络有多少资产（业务系统、中间件、数据库、虚拟机、服务器、网络设备、IP、端口等等）；

（2）资产频繁的上线、下线，存在大量影子资产、废弃资产，给黑客留下大量机会；

（3）缺乏精细化的资产管理手段，导致安全运营工作无法落到实处。

2.2 安全漏洞多，通报压力大

（1）医院业务系统多，安全漏洞多而杂，面对通报压力，漏洞修复难，无从下手，无法判断漏洞修复的优先级，缺少漏洞生命周期追踪管理机制；
（2）漏洞修复难，缺乏行之有效的漏洞修复方案；
（3）面对最新漏洞响应滞后，不知道哪些关键资产，可能造成哪些危害。

2.3 防守边界不明确，防守策略无效

（1）医院业务交互复杂，暴露面过多，防护边界模糊，端口映射混乱等等，给黑客留下大量机会；
（2）设备上线之后缺乏对安全策略的有效性验证及优化，无法适配当前威胁情况；
（3）采购了大量安全设备，没有最大化产品本身具备的安全能力。更多是基于事后的被动响应，当出事了安全事件时后知后觉，安全攻击窗口期过长，出现安全事件后疲于应付。导致即使部署了大量的设备，但对安全现状没有很好改善。

2.4 运营压力大，处置不及时&不彻底

（1）据调查显示，90%的医院没有专业的安全运维人员，都是兼职，仅 10%的医院配置了专职的安全人员；医院信息化人员非常紧缺，维护设备和处理医院事务已经耗费的大量的精力，没办法把网络安全的工作做好；
（2）缺乏运营支撑流程、运营支撑工具，发生安全事件时，没有相应的运营能力进行闭环处置。

2.5 网络安全投入有限

（1）随着信息化的发展和业务量的增长，部分设备可能面临性能、防护功能瓶颈，购买的设备将会被淘汰，造成资源浪费。
（2）医院认为信息系统只是一个管理工具，“烧钱”却不创收，信息化满足日常基本管理就行，没必要在网络安全维度投入过多的预算和精力；

3. 智慧医疗网络安全防护之道

新一代的数据中心安全建设采用“组件+人员+云端”的模式，该模式通过专家技术和平台能力的共享实现小投入大产出，具体而言：该模式为医院信息系统提供威胁的检测、响应、处置能力，通过安全托管实现以持续有效为目标的安全运营体系。

网络安全托管服务以“人机共智”为核心。其中，“人”是指安全专家，为了有效应对威胁，设置了不同级别（T1、T2、T3）专家团队；“机”则是指 AI 安全运营平台、基于威胁情报打造的脆弱性管理平台、标准化服务监控平台。人机共智的创新模式，通过 “人”“机”两者相互配合，更能发挥出彼此的优势，为用户提供更能保障安全效果的持续性标准化安全服务，最终实现 7*24H 可持续安全运营。

图 1 安全托管服务组织架构图

安全托管服务核心运转情况如图 2 所示，该服务结合各种安全平台及工具，整合厂商中安全专家的安全知识与能力、代理商和医院本地处理及服务能力形成了“三级联动”，并以三级联动为核心应对各项风险，整合各方力量共建安全网络空间。经济角度，医院通过购买服务、租赁设备的商业模式获取安全平台及工具，这有助于减少初期一次性投入；功能角度，该服务整合了厂商安全专家、安全平台及工具、代理商和医院运维团队的力量，显著提高医院信息系统安全运维能力，使各医院信息系统安全运维不再单打独斗。

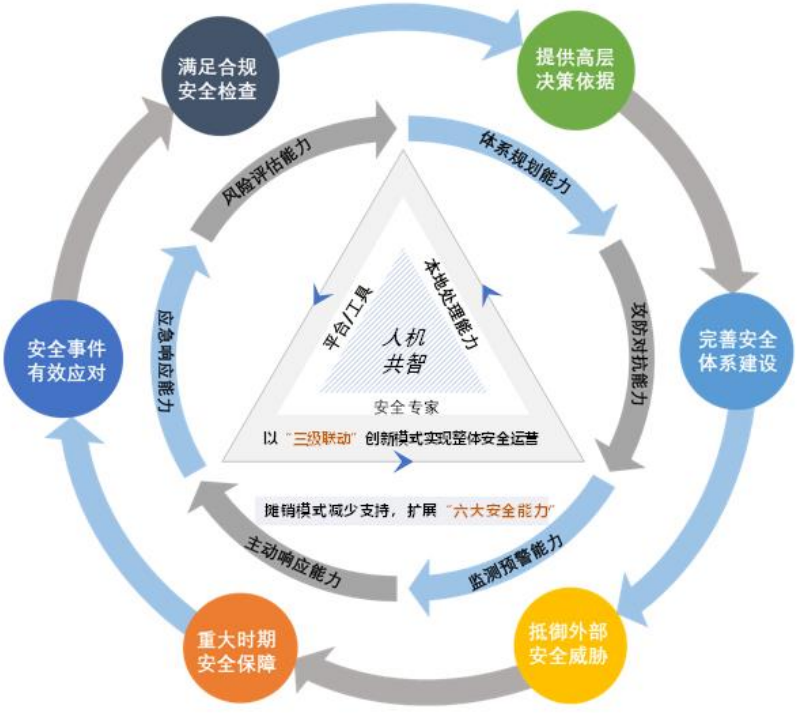


图 2 安全托管服务运转情况

三级联动配合示例如图 3 所示。本地平台是医院信息系统安全运营的第一道防线，当发现病毒攻击时，本地平台自动发出处置工单，厂商安全专家首先对系统报警进行审核，然后将特定的问题交由本地运维工程师和厂商安全专家进行处理，并在处理完成后对问题进行复核，然后归档。安全托管服务中规范化的事件处理流程在整合资源的同时，提高了安全事件处理效率。

图 3 三级联动配合示例图

图 4 展示了安全托管服务的运转机制。如图所示，在安全托管模式中，本地服务商向安全厂商订购安全平台等产品，同时向厂商获取交付能力、安全运维能力以及云端的安全响应服务能力；医院每年支付一笔服务费用，向本地服务商租赁安全平台等产品，并获取本地化服务以及厂商云端服务能力；本地服务商承担医院日常的运维服务，以及安全漏洞、威胁、事件等的处理能力；安全厂商提供产品及云端安全服务，通过设备与安全专家实时监测医院整体安全状况，一旦发现安全漏洞、威胁、事件等，可在获得授权情况下进行处置，或提出最佳处置建议；医院运维人员或本地服务商承担最终处置角色，收到厂商提供的处置建议后可根据要求进行处置；处理完成后由厂商后台输出处置报告及运营报告，运营闭环。

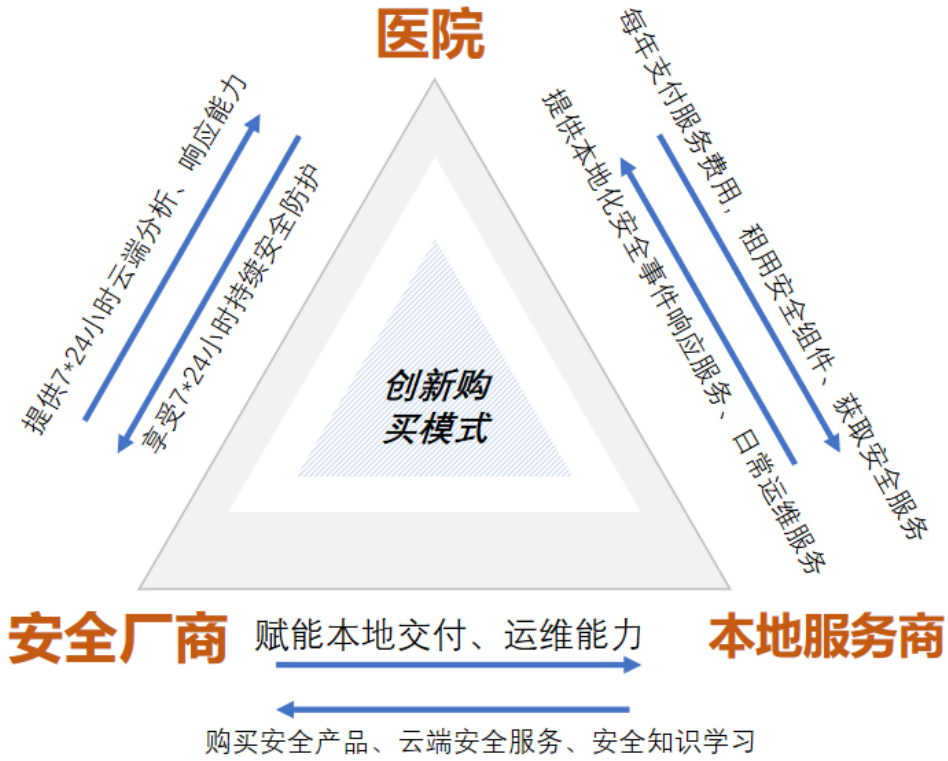


图 4 安全托管服务的运转机制

4. 安全托管服务的优势

4.1 低投入，高性价比

安全托管服务建设模式可避免一次性投资过大[10]，缓解资金压力。用户可向服务商订阅安全服务能力，由服务商结合专业的安全厂商产品和专业的安全人员为用户提供服务，达到比一次性硬件投入等同的安全效果。

4.2 高度先进的安全运营中心

安全运营服务是基于“人机共智”模式，应用了高度先进 AI 技术的自动化平台，使用“产品+平台+专家”的模式为医疗用户提供专业的安全运营服务。

4.3 7*24 小时持续保护

持续服务的好处是可以贯穿安全事件的生命周期，为防御手段提供参考。利用安全运营中心（SOC）帮助客户持续进行安全监测，时刻洞悉网络的事件根因，在威胁未发生之前实现最大化精准预警，并进行安全策略调整；对已确定的安全威胁持续监测，并进一步验证策略的有效性；从而实现主动快速响应，精准拦截黑客攻击，保障网络安全的目标。

4.4 精准预警与极速响应

单一时刻的安全日志并无法描述整个攻击行为的全过程，持续地监测有助于安全专家了解攻击行为当前所处的状态以及下阶段采取的攻击方式，实现精确预警[7]。同时基于持续的服务，安全专家可以清晰了解用户的网络安全情况以及资产所面临的潜在安全威胁[8]。并在后续持续的监测中发现新的安全威胁，当攻击行为发生质变或量变时，安全专家第一时间响应并对安全威胁的根因进行排查与处置。

4.5 服务监控大屏保障服务过程可视化

以往安全服务通常存在服务过程不可视的问题[9]，作为服务购买方来说，无法实时掌握第三方服务过程，导致服务效果无法把控。为了解决以上问题，安全托管服务使用服务监控大屏使得整个服务过程可视化，让服务购买方及时了解服务过程中的各个阶段以评估服务效果。同时，配备安全专家服务 KPI 考核平台对线上安全专家的服务进行全面监控，保障服务的及时性以及有效性。

5. 结语

随着医院物联网和人工智能技术的发展，医院网络安全面临新的挑战。信息安全建设从传统的信息中心范围，拓展到医院全空间。互联网医疗、医联体和以患者为中心的医疗服务模式创新，使网络安全突破了医药围墙的限制。信息在有效支撑医院的医疗服务、医学教育、医学研究和疾病预防任务工作的同时，带来了网络安全新风险。

医院网络安全建设升级既需要内部的力量，也需要借用外力，安全托管建设模式必定成为未来智慧医疗安全建设的趋势。该模式不需要组织投入大量的人力、物力去建设，真正实现了按需使用，随时随用的目的，并且将组织宝贵的 IT 人员从繁琐的工作中解放出来，将更多的精力聚焦在更有价值的工作上，为安全能力云化和安全服务云模式的发展画下浓墨重彩的一笔。

PU-295

从查新项目统计分析广东省医疗卫生领域人工智能应用现状

李翠霞 郑利荣
广东省医学学术交流中心

1 引言

从《国务院关于印发新一代人工智能发展规划的通知》（ 国发〔2017〕35 号）和《进一步改善医疗服务行动计划（2018-2020 年）》来看，我国已经把人工智能作为一个国策，将人工智能提升到国家战略水平，为医疗事业下一步发展指明了方向，并提出三步战略目标，为智能医疗的发展设定智慧医院建设、手术机器人、智能诊疗、智能生理监测、影像识别、新药研发、医药监管以及流行病智能监测与防控等方向[1]。广东省医学学术交流中心查新咨询部是卫生部第一批定点查新单位之一，是广东省医学查新主要查新点，为各级医药卫生科技项目立项(包括省卫建委、科技厅、中医药管理局，及各地市卫监局和科技局等项目)和成果鉴定提供文献学新颖性评价。现对查新中心近年来在“人工智能”方面的立项和成果进行统计分析，以了解其在广东省的发展现状。

2 资料来源与方法

2．1 资料来源

主要资料来源于广东省医学学术交流中心查新咨询部科技查新项目管理信息系统，包括“人工智能”相关课题的查新咨询委托单、查新立项申请书及查新相关成果资料、查新咨询报告。

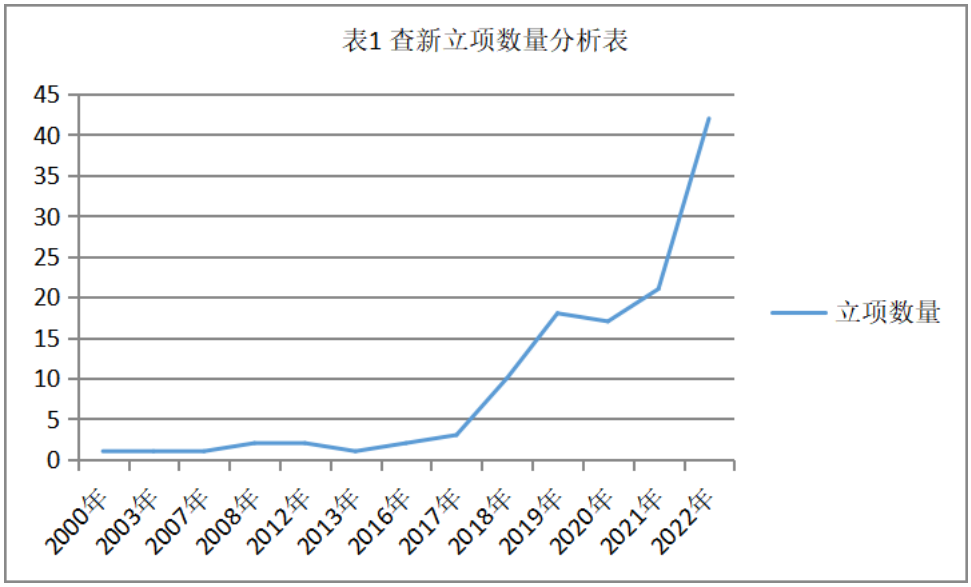
2．2 方法

在查新项目管理信息系统的报告管理—报告检索—查新项目中文名或检索用词字段，用“人工智能”或“AI”进行检索，将检索结果的各项数据输入 Microsoft Excel，并进行统计分析。主要针对查新系统查新项目数量、查新项目学科分类、及查新项目地区进行统计分析。

3 结果与分析

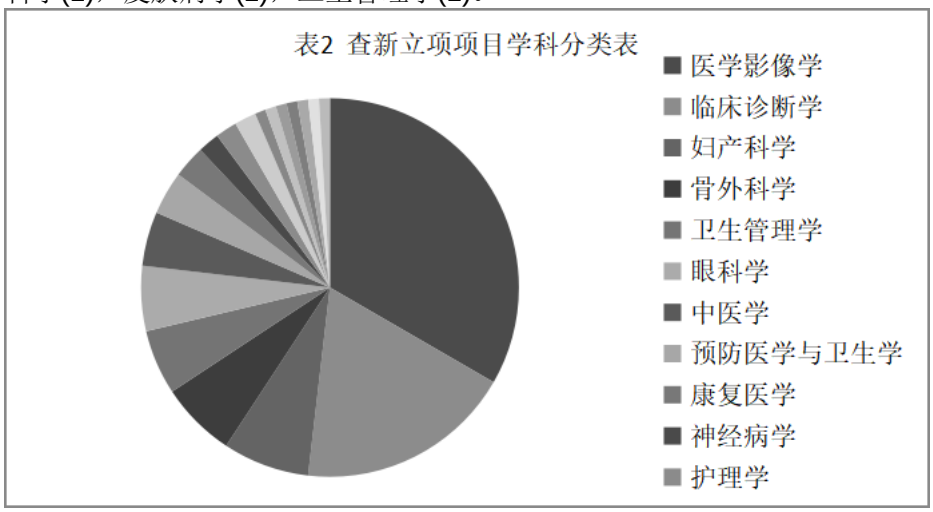
3．1 查新项目数量分析

共检索到人工智能相关项目 177 项，排除引文分析及外省相关项目，得到广东省科研立项查新 122 项，成果鉴定查新 13 项。122 项科研立项查新的时间分布为 2000 年、2003 年、2007 年及 2013 各 1 项，2008 年、2012 年和 2016 年各 2 项，2017 年 3 项，2018 年、2019 年、2020 年、2021 年、2022 年分别为 10、18、17、21、42 项。从立项查新的数量上来看，2000 年-2017 年为低速增长阶段。2018 年到 2022 年为高速增长阶段，自此查新数量明显上升并逐年增长。见表 1。13 项科研成果查新的时间分布为 2017 年 1 项，2020 年 2 项，2021 年 3 项，2022 年 7 项。



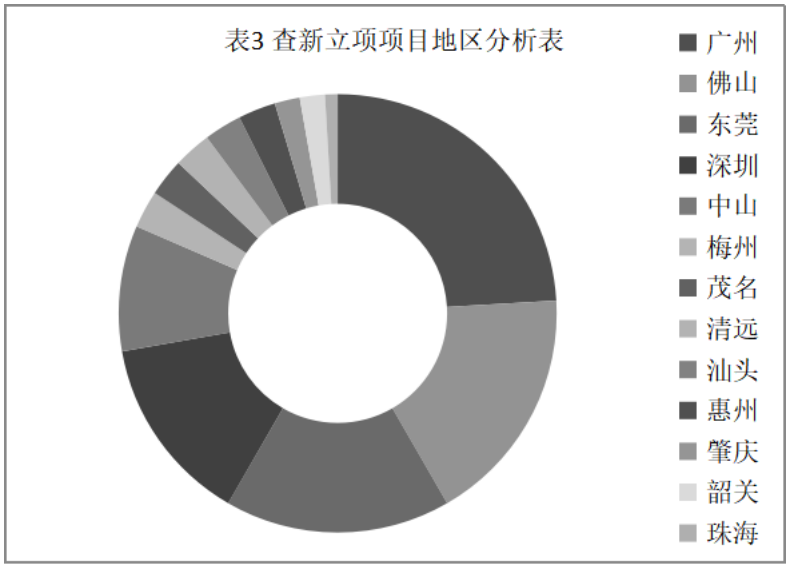
3．2 查新项目学科分类分析

根据我中心使用的“中华人民共和国国家标准学科分类(医类)”，对 2018-2022 年的 108 项查新项目进行学科分类分析。人工智能在医学科技领域的涵盖面丰富，共涉及到 19 个学科分类，既包括各类疾病的诊断和治疗，也包括疾病的预防及中医学等，充分反映了 AI 在医疗领域应用的丰富性和广泛性。学科分类的具体分布情况见表 2。其中查新项目不少于 5 项的学科分类有：医学影像学(36)，临床诊断学(20)，妇产科学(8)，骨外科学(7)，卫生管理学(6)，眼科学(6)，中医学(5)。13 项科研成果的学科分类及查新数量包括：眼科学(4)，呼吸病学(3)，肿瘤学(2)，口腔医学(1)，骨外科学(1)，皮肤病学(1)，卫生管理学(1)。



3．3 查新项目地区分析

对 2018-2022 年的查新项目的所属地区进行统计分析。共涉及到 13 个地区，包括：广州(26)，佛山(19)，东莞(18)，深圳(15)，中山(10)，梅州、茂名、清远、汕头、惠州各 3 项，肇庆、韶关各 2 项，珠海 1 项。具体分布情况见表 3。13 项科研成果地区分布 12 项为广州，1 项为汕头，具体单位分布为中山大学中山眼科中心(2 项)，中山大学肿瘤防治中心(2 项)，广州呼吸健康研究院(2 项)，中山大学孙逸仙纪念医院、中山大学附属第三医院、中山大学附属口腔医院、广州医科大学附属第一医院、广东省人民医院、广州市妇女儿童医疗中心、汕头大学·香港中文大学联合汕头国际眼科中心各 1 项。



3 讨论

人工智能的发源时间是从 20 世纪中叶开始，在 1956 年的达特茅斯学院会议上“人工智能”这个词被正式提出，科学家也是从此开始真正踏上智能研究的道路，而医学领域进行 AI 探索的尝试从 20 世纪 50 年代开始[2]。而广东省人工智能医学查新报告从 2000 年才开始出现，并且在 2017 年之前数量较少，从侧面反映出广东省在该方面研究的滞后性。然后从 2018 年至今，相关立项研究快速增长，说明由于政府的引导及自动化技术的提升，交叉学科的应用，使得人工智能在广东省医学领域的研究热潮达到高点。

人工智能在医疗健康领域应用广泛，从全流程辅助诊疗、临床决策支持、健康管理和生物医药产业升级等方面都有人工智能的广泛应用。未来，人工智能与医疗卫生将会进一步融合，重点将在新药研发、基因测序、可穿戴医疗设备、远程医疗、健康管理等方面，深刻改变目前的医疗现状和人们的生活[2]。但从查新报告的学科分类分析，广东省人工智能相关学科主要集中在临床诊断学和医学影像学，占比为 52%(63/108)，而在其他学科中的应用却较少。广东省应拓宽人工智能的研究范围，针对发展智能化医疗配套能力，严格规范行业监督管理、加强数据管理、重视培育复合型人才、健全法律法规和伦理规范，从而促进和深化人工智能的全方面发展[3]。

广东省为中医药大省，近年更是出台一系列重大措施，强力推进中医药强省建设。我国早在 1982 年就提出中医学与人工智能的关系，表明我国极早在医疗领域就开始探索 AI，而且是与中医学相关的辨证理论[4]。而广东省在中医方面的人工智能相关研究只有 5 项，主要集中在疾病诊断学方面，无论在研究范围还是数量和质量上都有待提高。广东省应进入新时期的新阶段，抓住“智能升级”挑战与“弯道超车”快速发展的机遇，实现高科技人工智能技术与传统中医模式的深度融合发展，强化中医文化传承、规范中医诊疗发展、提升中医诊疗的全面发展[5]。

广东省人工智能方面科研立项查新(122 项)与成果鉴定查新(13 项)的数量极不平衡。即科研立项与成果产出数量差别巨大。广东省应加提升研究成果转化为实际应用的能力。医院参与研发过程，不仅仅是提供数据支持，使本身的研发创新能力得到提高，切实促进广东省智能医疗走上新台阶。

4 结论

广东省在人工智能医学领域相关立项和成果鉴定中的项目，无论是在数量还是质量都有待提高，应扩大相关领域的研究，加强相关复合型人才的培养，提升成果转化和实际应用的能力，切实促进广东省智慧医疗的发展，提升人民的医疗水平。

PU-296

医院信息化建设存在的问题及思考

王春杰
郑州大学第一附属医院

0 引言

医院信息化建设是指利用先进的计算机技术实现医院各方面的综合性管理[1]。在当前医学水平不断进步、医院规模持续扩大的趋势下，医院信息系统建设的重要性日渐凸显。随着我国各方面改革的不断深入，医疗方面也进入了更加深层次的发展阶段，并且在当前我国医院的管理效率和整体的服务质量等方面都可以看到较为显著的改善。但是从整体的角度来看，当前信息化发展在我国医院信息化建设方面的应用并不是一帆风顺的，还是存在着一些固有的问题，并且带来了一些新的风险[2]。为了促进信息化建设在我国医院中更好的应用，促进医疗质量的提升，有必要对当前的问题进行分析以促进其发展。因此，对当前我国医院信息化设过程中面对的难题进行分析，并提出发展对策，具有显著的现实意义。

1 医院信息化建设的重要价值和作用

1.1 医院信息化在医院发展中的价值体现

医院通过信息化建设，实现了大量数据的信息共享，大大降低了医院的运营成本，提高了运营效率。另外信息化的应用为患者的预约挂号提供了便利，不但明显地降低了患者排队等号的时间，而且降低了人为因素引起的差错，这也为医院的运营降低了成本[3]。

通过信息化建设，医院在临床治疗中大大降低了差错的发生，并且有效改善了临床护理的程序。在数字信息平台的帮助下，可以将患者的诊断信息、检验信息、影像信息等进行详细的储存，这样不但提高医生诊断的效率和准确性，而且大大缩短了患者的诊疗实践，实现了患者和医院的双赢。

通过对信息技术的使用，医院在运营过程中大大提升了医院利益群体的满意度，其中包括政府、保险机构、患者等等对医院的评价。医院信息化对医院的服务流程进行了有效的改善，促使医院的战略目标快速达成[4]。另外医院通过系统平台可以对各项工作进行集中管理，对各项收费信息实现及时传递和公示，在很大程度上提高了患者的满意度。

1.2 医院信息化建设在医院发展中的作用

医院信息化建设有利于提升医院核心竞争力。各行各业的发展要想在领域内形成一定的影响力，就必须拥有特色的核心竞争力。医院作为社会的医疗服务保障部门，必须充分利用现代化技术，实现信息化建设，不断提高自身的核心竞争力。医院信息化建设，是医院在管理、服务、宣传等各个环节的工作中，通过利用计算机和网络技术提高医院运营水平的过程。通过医院信息化的实现，可以有效运用数字平台对医院各个部门、各个科室的运营和患者数据，进行有效的采集和分析，在对数据进行汇总的基础上，形成医院独特的优势，进而提高医院的核心竞争力[5]。

医院信息化建设有利于构建强大数据平台。随着信息化建设的深入，医院信息化的力度也在不断加大，在此过程中医院的信息数据平台的建设也得到了加强。

通过医院信息数据平台，可以对患者从入院诊断到入院治疗的所有档案，进行精确的电子化处理，并把这些数据并入医院的大数据平台，充分实现数据共享，为医院各个部门的工作提供全面的信息数据[6]。通过信息化数据平台，还可以提高医院各类数据的准确性，大大提高了医院的工作效率。

医院信息化建设有利于提高医院运营效率。随着医院信息化建设的加深，信息化系统平台已经成为医院进行日常工作管理的重要手段。医院的信息化建设彻底改变了医院在以往管理中的传统理念，为患者的就诊提供一站式的医疗服务。因此，医院通过对信息化平台的使用，不但规范了医院

的管理工作、提高工作效率，而且拉近了医患之间的关系，降低了医患矛盾。

医院信息化建设有利于提升医生专业水平。医生在对患者进行诊断治疗时，可以通过医院信息平台的数据库，对患者的诊断信息、治疗信息进行查询、汇总、分析，通过对数据库中患者的数据进行分析，可以为医生的诊断提供一定的参考依据，提高医生诊断的准确性[7]。同时医院信息平台的建立，可以把相似患者的病例进行总结和比较，为医生的治疗提供学习和借鉴的机会，进一步提高医生的专业水平。

2 医院信息化建设存在的问题

2.1 部分医疗人员对信息化建设认识不足，专业人才缺乏

现阶段在我国各级医院进行信息化建设过程中，面临的首要难题就是医院内部的相关工作人员和管理人员，对于信息化建设的认识存在偏见和不足。这主要是由于当前我国各级医院进行信息化建设的过程中存在着参差不齐的现象。一些医院现有的信息化建设完成程度相对较低，而信息化整体建设过程又是一个时间相对较长的工程，不仅仅周期长，而且在相对较短的时期内难以达到十分显著的良性效果。这就导致医院内部的管理人员和工作人员难以及时的认识到信息化建设的好处和长远意义，因而，对信息化建设在医院运行过程中的应用产生了一定的抵抗和抗拒心理。这是一个不得不重视的问题，因为医院内部的管理人员和工作人员缺乏认识，会导致信息化建设在我国各级医院建设中面临较大的阻力，因而必须加以及时的解决和改善。

医学信息化建设不是一个单独的学科[8]。因此医院信息化管理人才既需要熟练掌握计算机的相关知识，具有较高的信息化处理水平，同时还要具有一定的医学知识，但是在实际工作中，这样的复合型人才非常少见。目前我国在医院信息方面的工作人员主要有两个来源，一个是计算机专业毕业的人员，由于计算机专业的人员在知识结构上比较单一，并且对于医学知识和医院的管理方面的理论了解的比较少，在进行医院信息化建设过程中，往往不能关注到医疗行业的特殊性；另一个是医院其他岗位转岗的人员，这一部分人员虽然具备一定的医学知识，但是对于信息化建设的理解不够深入，在技术上不够熟练。因此，现代医院的信息工作人员并不能满足医院的信息化发展需要。

2.2 医院信息网络存在安全隐患，隐私保护不足

医院信息网络担任维持医院业务的有序运转的重要角色，它的安全与否关系着医院工作能够正常运转。医院管理者的各个决策以及各管理系统的操作安全都直接关系着医院信息网络安全[9]。目前许多医院的多数信息系统仍采用单机运行模式，缺乏合理的备份机制，使得系统的安全性难以得到有效保障。此外，操作系统漏洞也加大了医院信息平台的危险性，一旦受到攻击发生系统瘫痪或数据丢失等情况，会对医院造成难以估量的损失。

当前阶段，信息化建设在我国各级各类医院建设过程中的应用，还面临着隐私保护不足的问题。这是由于随着网络的发展和信息技术的不断进步，当前我国各级各类医院都在更多地采用电子病历。同时，会有医院对相关的具体诊疗过程合并进行保存。为了更好地发挥医疗信息的作用，在很多医院的技术部门中，会采用现代化的信息系统对相关信息数据进行收集，为后续的相关医学实验和科室的医学研究提供更多的便利和更加充分并且科学的数据。这一过程的确有利于我国各级各类医院信息化管理水平的提升以及总体医疗质量的改善。虽然信息系统在发展过程中有这些优势，但也无法掩盖其由于信息保管不完善、流通渠道存在漏洞等导致的信息泄露隐私保护不足的问题。我国各级各类医院在信息化建设过程中所面临隐私保护不足的问题是十分重要的，这不仅仅会对医院发展产生重大威胁，而且会大大提升公民个人信息的风险系数，因此需要尽快改善，加强技术维护。

2.3 信息化建设经费不足

医院信息化建设是一个长期的发展过程，需要不断的调整、更新和开发，这不仅需要知识和技术支持，更需要大量的资金支持。目前大部分医院未开展信息化管理的重要因素是缺乏资金[10]；同

时有部分医院采用信息化建设初期，给予大量资金支持，在建设到一定规模，消减资金投入，严重制约了信息化建设的深入、全面发展。医院信息化建设不能给医院带来直接的经济效益也一定程度的影响了资金投入。此外，信息化建设的重要性和长久性未得到充分的认识，在一定程度上也限制了医院信息化建设的发展。目前我国医疗行业信息化建设资金投入与国外相比仍存在很大差距。发达国家医疗行业信息化建设资金投入占比 3% ～ 5%[11]，而我国医院还不到 1%。

2.4 医院信息化建设过程中缺乏信息公开、共享

当前，我国各级医院所拥有的各类医疗信息，并不是医院私有，也并不是单个医院可以私自处理的，它是一种十分显著的公共资源。不仅仅需要在我国相应的卫生行政部门进行备案，不同等级的医院还需要在信息化建设过程中做到向社会共享一些信息。但是，在各级医院实际运行和发展的过程中，很多医院为了维护自身的既得利益，并没有向社会大众公开、共享这些信息。这就使得很多需要就医的患者在自身拥有医保卡的同时，还是需要前往多家医院办理其医院提供的诊疗卡或者居民健康卡，并且还可能会在检测和康复的过程中进行多次的同种检查。这种信息化建设无法共享互通引发的问题，不仅仅损害了我国到医院就医的消费者的个人利益，而且会阻碍信息化建设的发展和深入运行。当前阶段，我国各级各类医院之间的信息无法实现共享、互联互通，这会导致信息化建设受到较大的阻碍，因此需要加以重视。

2.5 医院信息化建设缺乏高层次的设计与规划

目前我国医院信息化建设工作主要集中在地方层面，缺乏较高层次的领导和规划。使得医院信息化建设在全国范围内表现出在区域发展不平衡，发达城市的医院信息化建设要远远高于偏远地区的医院水平。另外在基层医疗机构信息化发展的水平也各不相同，大部分的基层医疗机构受到资金和人才因素的制约，导致信息化建设水平偏低[12]。

3 医院信息化建设的发展策略

3.1 加强宣传信息化建设的重要性，培养专业人才

为了改善当前医院内部管理人员和工作人员存在的对信息化建设的偏见，要加强宣传，发挥信息宣传的作用。这不仅有利于改善医院内部相关工作人员和管理人员对于信息化的固有理念，促进医院内部管理方式的转变，而且有利于实现医院发展过程中以患者为中心的目标。加强信息公开和信息披露不仅仅有利于促进医院信息化建设的深入开展，而且有利于整体管理水平的提升，促进医院更好的发展。

要想实现良好的信息管理，医院必须组建专业的信息管理团队，不断提高信息管理工作人员的专业水平。为此，医院必须重视培训工作的开展，要定期就前沿技术和知识，对信息管理工作人员进行培训，提高信息管理工作者的专业水平和创新意识，针对培训效果，也要进行相应的确认。与此同时，医院也要根据信息化管理的实践工作，定期组织总结回顾，通过不断总结，持续探索信息化发展的长久道路。例如，在对档案进行信息化建设时，除了要完善档案管理平台，更要就档案管理人员的信息化管理技能进行培训，使其具有相应水平的操作技能和知识储备，同时也要对档案管理人员的保密意识进行培训。在此基础上，通过先进的信息管理技术不断提高档案信息管理平台的安全性和保密性，进一步完善、提高医院日常工作中档案信息的管理水平。

3.2 加强医院信息网络安全管理，保护病人隐私

加强医院信息网络安全管理关键在于三个方面：第一，医院负责人要加强对医院信息网络安全的重视程度，从软硬件配备、人才配备等各方面及时完善医院信息网络安全管理机制；第二，医院应重视对医护人员的网络安全知识培训，培养他们及时查杀病毒、设置高强度密码、不随意打开不

明电子邮件等习惯；第三，做好数据备份工作，这样一旦计算机被黑客入侵，医院也可以及时恢复其数据系统，确保工作能够正常运行。

为了解决当前我国各级各类医院进行信息化建设过程中面临的隐私威胁，需要采取相对科学并且有效的措施促进信息保护。在确定相关患者病历信息的准确性之后，相关医院的技术部门可以将患者的具体资料运用一定的技术进行匿名保护；在医院发布的公开数据中，可以将一些敏感信息或者隐私信息标识进行有选择性的公开，并且进行匿名处理。促进患者以及信息化建设过程中的隐私保护，不仅仅有利于促进医院的信息化发展，而且有利于保障公民个人信息安全，促进社会平稳运行。

3.3 加大信息化发展的资金投入，调整投资结构

信息化发展对于医院管理和日常工作的价值往往需要一定的时间才能体现，因此医院管理者必须正确认识到信息化发展的意义和对医院长期发展的价值。在此基础上，根据医院的长期发展规划，制定相应的投资计划，加大对于医院信息化发展的资金投入，做好信息化发展的保障。伴随信息科学技术的发展，医院的信息系统更新会更加频繁，因此需要加强资金的保障。另外在投资结构上，医院信息化建设既要注重购置硬件设施，同时也要发展应用软件和培养信息化人才。

3.4 各级医院信息化建设发展改进需要公开、共享医疗信息

为了解决当前我国各级医院进行信息化建设过程中所遇到的难题，十分重要的一个方面，就是要促进当前各级医院各类医疗信息的公开和共享。要想实现这一目标，就需要各级医院加强对于患者医疗信息的共享力度和管理力度。我国各级各类的相关政府医疗卫生部门应该促进建立地区统一信息公开、共享规范，将这些医疗信息以相对公开并且透明的方式披露给大众。具体的内容不仅仅需要包括相关各级各类医院的医疗费用，还需要将医院信息化建设过程中的流程等进行公示共享，保障相关居民在这一方面的知情权，同时还可以促进公民更好地实施自身在医院信息化建设方面的监督权。促进我国各级各类医院建设过程中公开共享医疗信息，不仅仅有利于打破当前我国各级各类医院形成的灰色利益格局，而且有利于促进统一强制性标准的建设，加快推进各级医院信息化的步伐。

3.5 制定国家层面的医院信息化发展规划

根据现代医院信息化发展的情况来看，只有从国家层面对信息化建设进行统一的制定，才有助于全国医院信息化的发展，各级医院的信息系统数据共享的程度也就越高。因此，国家相关部门应尽快出台医院信息化建设的规划和指导意见，指导各级各类医院的信息化建设，为医院的发展提供一定的理论依据。

4 总结

医院信息化建设是一个长期的动态过程，通过对医院信息化系统调整、优化，实现动态发展和应用，使医院朝着现代化、系统化的方向发展，不断提高医院的综合实力，为患者提供最优质的服务，可有效促进医院长期稳定的发展。

PU-297

基于 BERT 模型的乳腺癌影像学报告中 TNM 分期实体识别预测方法

唐现策[1]，许延杰[2]，燕青[2]，刘长鹏[1]，韩枫[1*]
1.郑州大学附属肿瘤医院
2. 健康数据（北京）科技有限公司

肿瘤分期是肿瘤诊治的重要环节，是临床医生对患者制定治疗方案的评估依据。乳腺癌治疗前评估中最常用的标准是由美国癌症联合委员会（AJCC）出版的基于肿瘤-淋巴结-转移（TNM）状态来评估的乳腺癌 TNM 临床分期[1]。一般诊断信息来自乳腺癌影像学检查，如乳腺 B 超、乳腺钼靶以及乳腺 MRI 等影像学报告。由于 TNM 分期依赖大量不同来源的影像学信息，医生实际决策时间较有限，很难准确提取与分期相关信息，因此识别和提取指示病变是一项关键的临床任务。计算机技术的使用可以为医生决策提供必要的分期信息，从而提高肿瘤临床分期的准确性。然而，大多数 TNM 分期信息以自然语言的形式存在于各种影像报告中，不能被计算机直接利用，或许人工评价对于小规模任务可能是可行的，但在大样本量的任务中，人工方法非常耗时、容易出错且成本高昂。因此，从影像学报告中快速自动提取 TNM 分期信息对于治疗方案推荐、预后评估等方面是非常有意义。

目前，虽然有一些用于乳腺癌诊断的 NLP 研究，如从乳腺 X 线报告中提取 BI-RADS 分类和相关影像学特征[2、3]. 然而，很少有研究调查基于自由格式文本的乳腺影像学数据中的提取 TNM 相关实体信息。本研究的目的是探索运用基于 Bert 模型的乳腺癌 TNM 分期相关实体识别的有效性，使用深度学习模型，主要以 Transformer 机制的 Bidirectional Encoder Representation from Transformers(BERT) 模型及改良的 A Lite Bidirectional Encoder Representations from Transformers (ALBERT)模型分别与 Global Pointer(GP)、Bidrectional Long Short-Term Memory Conditional Random Field(BiLSTM‐CRF)相结合, 从乳腺影像学报告的自由文本格式中准确提取 TNM 相关病变实体及其相应成像特征。

1 资料与方法

1.1 数据来源

河南省某三级甲等肿瘤专科医院 2019 年至 2021 年所有乳腺影像学报告自由文本记录，按照以下纳排标准选取乳腺 B 超报告、乳腺钼靶报告、乳腺 MRI 报告各 500 份共 1500 份报告，我们其中的 1200 份（乳腺 B 超、乳腺钼靶、乳腺 MRI 报告各 400 份）来进行 Bert 模型的训练，剩余 300 份数据进行模型效果评估。

所有患者均符合以下纳入标准：1）乳腺癌确诊患者在本院治疗前影像学报告；2）影像学报告需描述肿瘤、淋巴结相关特征信息。排除标准：1）术后影像学报告；2）B 超引导下穿刺记录等操作记录。

1.2 实体定义

根据 AJCC 第 7 版乳腺癌 TNM 临床分期标准，我们将分期内容进行元素信息化提取，并整理形成无重复的影像学特征集合。所有的影像特征实体和值均由乳腺疾病领域经验丰富的专家根据 TNM 分期系统的数据项名称分别在乳腺 B 超、乳腺钼靶及乳腺 MRI 报告中分别制定实体类别及其含义（表 1-表 3）。

表 1 乳腺 B 超文本相关实体命名

实体类别	实体含义	报告中举例
否定	否定词	未见、否
肿瘤	肿物、结节等	肿物、病灶、肿块、结节、囊肿、卫星灶、
位置	解剖部位	乳腺外上限、左乳内上至外侧、12 点钟、*乳、皮肤、*侧锁骨上下、颈部双侧、*侧腋窝、*侧乳腺、*侧胸壁

形态	肿物性状或异常边缘	转移、异常、肿大、病变、栓子声像、血栓、占位、缺如、融合
大小	肿物或淋巴结大小	1.0×12mm、5cm、19mm×16mm×14mm、17mm×6mm、44x20mm、5㎜×3㎜
诊断分级	影像学诊断分级	BIRADS 分类 4A、BI-RADS Ⅵ、BI--RADS 分级Ⅵ级、
诊断	临床诊断	*术后改变、乳腺癌、癌、手术切除、双侧乳腺局切术后、脂肪瘤、治疗后改变、
淋巴结	淋巴结	淋巴结

表 2 乳腺钼靶文本相关实体命名

实体类别	**实体含义**	**报告中举例**
否定	否定词	未见
肿瘤	肿物、结节等	肿物、病灶、肿块、团块、占位
位置	解剖部位	右乳、左乳、双侧腋下、皮肤
形态	肿物性状或异常边缘	毛刺、毛糙、异常、肿大、恶性、恶性征象、恶性钙化、高密度、密度增高、密度局限增高、融合、泥沙样钙化、段样*钙化、水肿
大小	肿物或淋巴结大小	1.2*1.1cm、
诊断分级	影像学诊断分级	BI-RADS 分类：右 6；左 4A、BI-RADS 分类：左侧 1 类，右侧 6 类 、BI-RADS：2 类、BI-RADS：4c、BI-RADS 2 类
诊断	临床诊断	乳腺癌、良性、术后改变、术后、治疗后改变、多灶癌、副乳
淋巴结	淋巴结	淋巴结

表 3 乳腺 MRI 文本相关实体命名

实体类别	**实体含义**	**报告中举例**
否定	否定词	未见、非、不、
肿瘤	肿物、结节等	结节、病灶、肿块、瘤、团块、
位置	解剖部位	右乳腺、右乳外上象限、三点钟
形态	肿物性状或异常边缘	转移、异常、肿大、占位征象、恶性病变、消失、病变、水肿、肿胀、融合、毛刺、增大、侵犯、
大小	肿物或淋巴结大小	13.8mm×13.4mm×11.9mm
诊断分级	影像学诊断分级	BI-RADS 6 类、BI-RADS 2、
诊断	临床诊断	乳腺癌、良性
淋巴结	淋巴结	淋巴结

1.3 研究方法

实验设计方案如图 1 所示，由两名专业医师进行人工双盲标注乳腺 B 超报告 500 份、乳腺钼靶报告 500 份、乳腺 MRI 报告 500 份共计 1500 份，并交叉核验标注结果，如有发现标注结果之间的任何差异均由第三位独立审查员进行评估，然后分别选取每类报告中 400 份用于模型训练，每类剩余的 100 份用于模型训练效果评估。使用的训练模型选择了 Bert-CRF、Bert-GlobalPointer、AlBert-BiLstm-CRF 三种不同的 Bert 算法模型，用于对比不同模型训练后的性能。

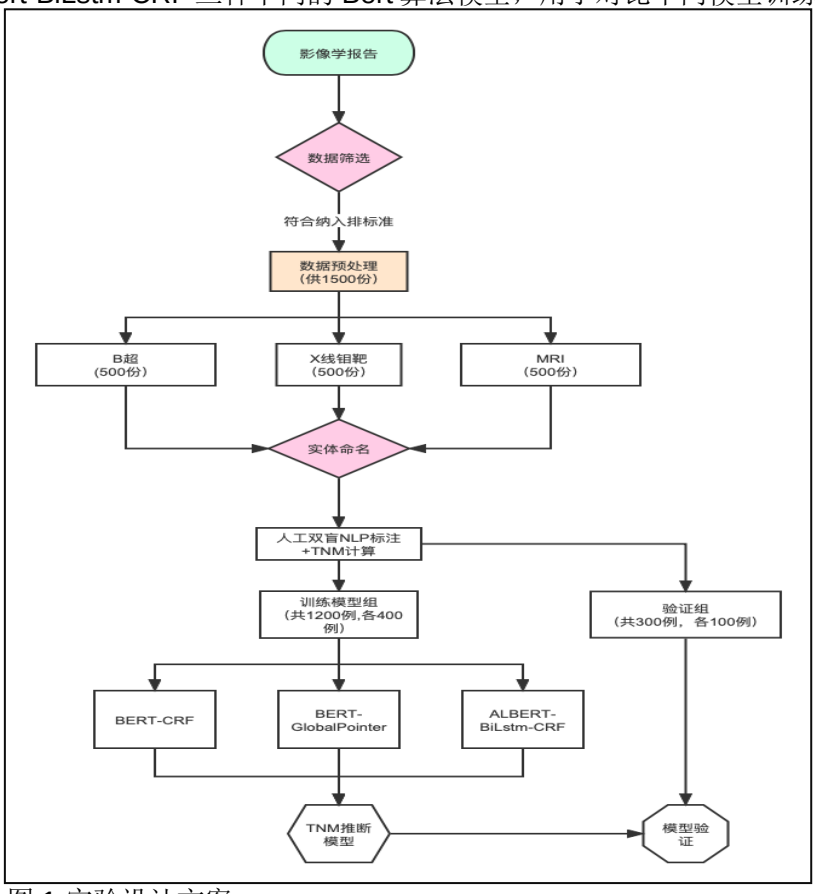


图 1 实验设计方案

2 结果

用 Bert-CRF、Bert-GlobalPointer 和 AlBert-BiLstm-CRF 三种算法对乳腺 B 超报告、乳腺钼靶报告和乳腺 MRI 报告三种不同类型影像学报告中的 TNM 分期实体识别的结果如表 4-表 6 所示，可以发现三种算法训练的结果都非常好，其中 Bert-GlobalPointer 算法训练出效果是最好的，在三种不同的类型报告 TNM 分期实体识别中的 F1 值分别达到了 99.55%，98.74%，97.79%。与此同时，我们比较了基于 Bert 模型的 NLP 程序和人工审查之间的 TNM 分期实体提取时间，发现每位人工审查员提取每份报告的指示病变和相应特征的平均时间分别为 4.4±0.95 分钟（95%置信区间：4.393-4.547）和 4.3±0.88 分钟（95%置信区间：4.480-4.640），而 NLP 程序在不到 2 秒的时间内就提取到同样结果，效率非常显著。

表 4 3 种不同算法对乳腺 B 超报告中 TNM 分期实体识别结果

模型名称	F1 值(%)	Precision(%)	Recall（%）
Bert-CRF	99.41	99.43	99.39
Bert-GlobalPointer	99.55	99.48	99.62
AlBert-BiLstm-CRF	97.57	97.64	97.50

表 5 3 种不同算法对乳腺钼靶报告中 TNM 分期实体识别结果

模型名称	F1 值(%)	Precision(%)	recall (%)
Bert-CRF	98.32	98.26	98.38
Bert-GlobalPointer	98.74	98.30	99.19
AlBert-BiLstm-CRF	96.08	95.86	96.30

表 6 3 种不同算法对乳腺 MRI 报告中 TNM 分期实体识别结果

模型名称	F1 值(%)	Precision(%)	recall (%)
Bert-CRF	96.82	96.00	97.65
Bert-GlobalPointer	97.79	97.35	98.22
AlBert-BiLstm-CRF	94.99	94.28	95.70

3 讨论

我们的研究结果表明，深度学习和其他机器学习方法能够准确识别绝大多数非结构化乳腺放射学报告中的临床命名实体，并且 Bert-GlobalPointer 模型在整体实体识别中表现最好，但并非在每个实体中表现最佳。其在肿瘤状态分类方面表现显著。NLP 分类结果的准确性与临床专业人员分类相当。

3.1 乳腺癌影像学特征自动提取应用潜力

由于其强大的软组织对比能力，乳腺 MRI 对于恶性肿瘤的检测和表征非常敏感。指标病变是最关键的因素，因为它们直接决定治疗和预后。许多回顾性研究需要从指标病变中提取 MRI 信息，以分析病变特征与临床结果之间的关联[4-5]。本文采用基于 Bert 模型的 NLP 方法从乳腺影像学报告中识别指标病变和相应特征。我们基于规则的系统使用概念匹配和否定检测，将自由格式文本报告中的图像特征映射到相应的类，并实现了极高的性能。根据指标病变占影像特征最多的规则，NLP 程序自动从每份报告中选择指标病变，用于乳腺癌的随访、靶向治疗和病变注释等各种工作。例如，乳腺癌的 10 年生存率超过 70%，局部病灶的生存率约为 89%[6].因此，随访对于乳腺癌患者来说非常普遍。随访的目的是检测指标病变的短期变化或评估治疗效果。如果 NLP 程序能够在随访时精确提取甚至比较指标病变的差异，将大大省去人工提取的耗时过程。

3.2 基于 BERT 模型的 NLP 方法优化方向

在本文的研究过程中，负面案例的数量太少，这可能会影响方法的性能。我们将在未来添加更多的负面案例来验证我们的基于 Bert 模型的 NLP 方法的性能并持续优化。此外，本文的研究是基于索引病灶占影像特征数量最多的假设来提取索引病灶，该假设是人为定义的，并非直接提取指标病变。将来，可以考虑根据影像特征直接从影像中提取指标病变。

4 结论

本研究主要关注基于中国乳腺影像学报告自动提取肿瘤 TNM 分期相关临床信息。从 3 个常见不同类型的乳腺影像学报告（乳腺 B 超、乳腺钼靶及乳腺 MRI）中识别肿瘤、解剖位置、形态特征、肿物大小、淋巴结及诊断等实体。为了使机器能够智能地处理临床信息，应用基于字典的匹配，Bert-CRF、Bert-GlobalPointer 及 AlBert-BiLstm-CRF 模型来提取临床命名实体。实验表明，机器学习方法能够自动从乳腺影像学报告中提取乳腺肿瘤状态相关实体，Bert-GlobalPointer 模型在严格和宽松的整体 F1 值中表现最好，分别达到 99.55%、98.74%和 97.79%，其准确性可与专家相媲美。本研究有助于临床命名实体自动提取，有利于加速临床非结构化数据的二次应用。我们的目标是为大规模乳腺癌大数据挖掘研究建立一个有用的工具，以探索乳腺癌发病率、诊治方面的新因素。

摘要交流

ZY-001

浅谈智慧医院背景下的病案数据质量提升

魏梦迪
郑州大学第一附属医院

目的 近年来，随着医疗信息化和智慧病案行业的快速发展，病案数据的价值在医院管理乃至公共卫生领域越发重要，随之而来对病案数据的要求也越来越高。
方法 本文首先指出传统的病案管理办法很难适应智慧医院形势下的新需求，甚至一定程度上出现“脱钩”，病案数据质量的缺失正限制着整个信息化体系的发展。传统意义上的病案质量控制主要是在病案书写及病案归档过程中的内涵质控，而随着时间推移以及智慧医院概念的提出和发展，医院希望通过病案质控提升医院管理效率，甚至为医院提升经济效益。由此，智慧病案解决方案开始从内涵质控转向全生命周期解决方案。
结果 讨论传统病案管理问题产生的三个原因，即一，专业管理人员不足；二，重视程度不够；三，病案管理信息化手段缺乏，点明病案质量控制内容主要集中在四个方面，分别是完整及规范性、逻辑判断、时限质控和内涵质控。
结论 建议建立三级六类质控系统，即将医院病案管理需求分为三类，包括高频错误、低频问题和临床管理问题。最后提出智慧医院背景下的病案数据质量提升，一方面体现在病案管理的流程上，另一方面，提供病案首页“一键生成”功能。此外，还应当支持专门的定制化服务及提供 DRG 预入组、诊断排序建议、智能编码等智能化服务。

ZY-002

全结构化电子病历的发展及困境

王奥钰
郑州大学第一附属医院

目的 简述全结构化电子病历的背景，指出目前全结构化电子病历的困境及完善结构化电子病历的必要性。
方法 分析全结构化电子病历的优、缺点，及其在病人就医、医保结算、科学研究、社会经济等方面的作用及仍需完善之处。
结果 全结构化电子病历是临床信息化建议和发展的需要。不仅能服务于临床诊疗，更满足二次利用的需求，为决策支持、医学研究、卫生行政、统计机构及其他实体提供临床信息。
结论 目前仍有许多因素制约全结构化电子病历进一步发展，仍需投入大量人力物力解决。

ZY-003

浅谈系统方法在现在化病案管理的应用

李芳
郑州大学第一附属医院

目的 本文着重从系统方法的有序化原则出发，探讨系统有序化原则在现代化病案管理的具体应用，以期为我国现代化病案管理建设提供科学的指导。
方法 系统方法是随着系统工程、一般系统论产生的一种崭新的科学方法，是二十世纪以来自然科学方法论研究所产生重大成果之一。所谓的系统方法指的就是，将研究对象放在整个系统中加以考察的一种方法，具体来说就是从系统的角度出发，综合考虑系统与外部环境之间、系统与要素之间

以及要素与要素之间的联系，从而准确掌握系统的本质和其运动规律，以达到最佳处理办法。[1]系统方法将研究对象当做一个完整、统一的整体，并着重于研究系统的整体功能，该方法是建立在物质统一性、系统整体性和层次秩序性等原理之上的一种新的科学研究方法。

结果 系统方法使人类的思维方式发生了深刻的变化，为解决问题提供了新的思路，对科学研究和发展产生了深远影响。

结论 系统方法应用到医学图书馆管理中应从整体出发，从部分和整体的联系中合理协调其中的因素和结构，此外，还应注意与系统其它的原理、方法和原则相结合，使病案管理系统以一个有机整体的最快的速度、最高的质量、最少的人、财、物耗费，将最多的最有价值的医院病案信息资料提供给用户，获取最大的社会效益和经济效益。

ZY-004

基于电子健康卡的急诊预检分诊管理系统的设计

唐跃涛
永州市中心医院

目的 急诊科是医院面向社会的“窗口”，直接反映了医院的医疗、护理、管理水平。以往的急诊科就诊流程复杂，普遍存在着看病难、流量多、时间慢等问题。而护理工作也完全依靠手工浪费大量的护士大量的时间，从而无法提供更高质量的护理；电子健康卡是随着移动通讯技术的发展，在手机终端虚拟实际健康卡的一种形式。电子健康卡是由国家卫生健康部门统一规划，统一标准，面向全国居民发放。将电子健康卡应用于急诊科的预检分诊系统不仅符合国家推行的全民健康信息化的要求，而且能帮助我院急诊科优化就医流程，改善病人来院就医环境。

方法 基于电子健康卡的应用场景，结合永州市中心医院急诊科医护人员在日常工作中碰到的问题，我院自主设计并研发了基于电子健康卡的急诊预检分诊管理系统，本文详细说明了系统的设计思路，关键技术，实现方式及运行效果。

结果 2.1 精准了解病人就医前病史记录，实现精准分诊就医

随着区域卫生信息化建设的加强，省互联互通系统平台的建设，电子健康档案全程记录了居民从出生到死亡的相关医疗保健和医疗机构就医记录；随着电子健康卡的应用，居民电子健康档案查询的开放，极大的方便医院急诊科工作人员，了解病人以前的身体状态，并结合本次发病的症状，可以精准的给病友提供医疗服务。

2.2 优化就医流程，改善就医环境

我院急诊科来院就医病人较多，病种复查，医护人员工作强度高，传统的手工分诊模式工作量大，出错率高；研发基于电子健康卡的急诊预检分诊系统是服务全民健康的要求，极大的简化了病人的就医流程，系统全程跟踪急诊患者的就医状态，大大减少了护士的工作量与出错率，病友的医疗数据在各部门之间高速传输，使得急诊就医高效便捷，为病人就医提供更加优越的环境。

结论 基于电子健康卡的急诊预检分诊系统结合了互联网技术并对预检分诊流程的改正，有效弥补了传统急诊预检分诊制度的不足，与传统医护人员通过手工记录分诊相比，本系统具有强大的功能和可操作性，依靠信息化的手段根据急诊的病人的症状进行分级评估，区别对待病情的轻重缓急，让危重病人得到及时的诊疗和抢救，系统简化了预检分诊流程，更加关注病人就诊信息的实时采集，并与院内其他系统进行无缝对接，实现医疗数据的实时共享，通过预检分诊系统的应用，科室管理变得更加规范，我院急诊科医护人员的工作效率有了明显提高，来院就医的病人对科室的满意度有了显著的提升；展望未来，随着人类科技的进步，希望将人工智能，大数据挖掘应用于急诊就医,，为急诊科医护人员的提供更多的辅助支持。

ZY-005

基于病案无纸化管理的管理实践

雷保仓*
漯河市第一人民医院

目的 随着信息化的不断推进，很多医院以病案无纸化为建设目标，积极推进病案无纸化管理模式，以提高病案管理质量的标准规范化程度和相关业务人员工作效率。其是机遇，亦是挑战，是重大改革创新点。
方法 该研究通过总结病案无纸化管理优化的意义和现状不足的基础上，结合数据时代背景，从技术和制度两个维度出发，提出了八条对病案无纸化管理的研究新思路，为医院提高病案无纸化管理质量提供参考。
结果 随着时代的不断演变，为适应医疗卫生行业的大环境，病案无纸化管理已逐步成为医院的重点关注和建设对象。持续推进和不断优化病案无纸化管理，可以给医院带来众多效益，通过本文研究，提出两个层次、八条新思路，为病案无纸化领域提供参考，其中，技术支撑层面：一可引入可信的电子病案管理系统，二可引入云存储技术，三可建设病案远程利用平台，四可加强基于病案的数据挖掘。再者，制度保障层面：一可加强病案管理人员的专业程度和综合素质水平，二可创新病案管理制度，三可建立完善的质控体系，四可严格实施病案考核和奖惩体系。基于大数据时代下的病案无纸化管理新思路，将共同促进病案管理工作的快速高质量发展。
结论 随着时代的不断演变，为适应医疗卫生行业的大环境，病案无纸化管理已逐步成为医院的重点关注和建设对象。持续推进和不断优化病案无纸化管理，可以给医院带来众多效益，通过本文研究，提出两个层次、八条新思路，为病案无纸化领域提供参考，其中，技术支撑层面：一可引入可信的电子病案管理系统，二可引入云存储技术，三可建设病案远程利用平台，四可加强基于病案的数据挖掘。再者，制度保障层面：一可加强病案管理人员的专业程度和综合素质水平，二可创新病案管理制度，三可建立完善的质控体系，四可严格实施病案考核和奖惩体系。基于大数据时代下的病案无纸化管理新思路，将共同促进病案管理工作的快速高质量发展。

ZY-006

ICU 患者医院感染直接经济损失评价

刘新元 尚小平 杨林朋 吕倩 刘新奎*
郑州大学第一附属医院

目的 医院感染是影响患者安全最为突出的公共卫生问题之一，重症监护病房(Intensive Care Unit, ICU)患者的医院感染发病率、致死率和经济损失均明显高于其他科室，因此对 ICU 患者医院感染的预防与控制具有十分重要的现实意义。本研究首先对郑大一附院 ICU 患者的医院感染直接经济损失进行测算；其次应用 AMOS25.0 软件构建递归系统模型，分析匹配成功后医院感染患者直接经济损失各影响因素的直接、间接贡献，并计算总贡献，为未来更好控制医院感染经济损失提供实践经验。
方法 基于医院感染实时监控系统，收集 2021 年 1 月 1 日～2022 年 12 月 31 日郑大一附院使用该系统满 1 年的所有 ICU 患者的相关病历信息，采用病例对照研究设计，发生医院感染的患者为病例组，未发生医院感染的患者为对照组。运用 SPSS 软件，按照医院、性别、年龄组、入院科室、入院病情严重程度、医疗保 险类型进行 1:1 匹配，采用病例对照研究设计，配对秩和检验(Wilcoxon Signed Ranks Test)比较病例组与对照组住院费用和住院天数的差异。应用 AMOS25.0 软件构建递归系统模型分析 患者医院感染直接经济损失各影响因素的直接与间接贡献，并计算总贡献。

结果 本研究符合纳入排除标准的共 516 例 ICU 患者。ICU 患者医院感染直接经济损失为 101652.56 元。采用递归系统模型发现：抗生素使用天数为关键因素，其对患者医院感染直接经济损失的影响不仅存在直接贡献而且存在间接贡献。
结论 抗生素使用天数是影响 ICU 患者医院感染直接经济损失的关键因素。为河南地区更加深入科学地探讨医院感染直接经济损失影响因素的内在联系以及直接、间接和总的贡献提供了实践经验。

ZY-007

DRGs 实施背景下公立医院高质量发展的思考

韩慧欣
郑州大学

目的 促进公立医院在 DRG 实施背景下找准定位，为推动公立医院高质量发展蓄力。
方法 本文拟介绍医保支付方式改革及公立医院高质量发展的背景、DRGs 实施背景下公立医院高质量发展现状及探讨其存在的问题，提出部分建议，从而推动公立医院高质量发展。
结果 医保支付改革背景下公立医院高质量发展存在：病案质控管理较薄弱、医疗质量安全管理建设不足、医院信息化水平不高等问题。
结论 DRGs 改革系国家医疗保障制度深化改革的有效尝试，公立医院应抓住机遇，努力由规模扩张转向提质增效，谋求公立医院高质量发展的改革之路。公立医院运用 DRGs 等工具完善运营精细化管理，不断完善和持续改进，使 DRGs 改革成为优化医院高质量发展的有力抓手，从而推动公立医院高质量发展。

ZY-008

关于麻醉临床信息系统启用手术 ICD 编码的思考

任田
郑州大学第三附属医院

目的 随着国家电子病历评级、互联互通、智慧医院等政策标准的相继实施，对医院各信息系统提出了更高的要求和标准，同时，依托相关政策开发的医院信息系统也规范了临床操作规范，使得病历档案趋于标准化。在此形势下，麻醉临床信息系统中的手术名称等相关数据也面临标准化。
方法 人员字典表与手术字典表数据统一。临床手术名称与标准手术名称对照。临床书写与标准化的冲突。
结果 手术名称及 ICD 编码标准化
结论 标准化 ICD 编码的推动工作虽然不易，且还有很多细节工作需要相关科室共同参与，但标准化落地实施后，将给医院信息化带来本质变化，同时也会反哺临床，给临床科研工作带来潜移默化的优势。

ZY-009

论医学信息标准研究与应用

智鹏
解放军医学院

目的 本研究旨在深入探讨医学信息标准的研究与应用，以强调其在医疗领域中的关键作用，并提出如何更有效地推广和应用这些标准以改进医疗信息管理和卫生服务的质量。
方法 通过对多个医疗信息系统的案例进行研究，以了解标准在实际医疗环境中的应用效果，分析了医学信息标准的发展历程、不同领域中的应用情况以及标准制定的过程。开展一项定性研究，通过采访医疗专业人员和信息技术专家，以获取他们的观点和建议。
结果 研究发现，医学信息标准在医疗信息管理中起着关键作用。这些标准不仅有助于实现医疗数据的互操作性，还提供了一个共同的语言，促进了卫生信息的共享和协调。案例研究表明，在临床实践、电子健康记录、医疗设备互联网、医疗图像管理等领域，标准的应用已经显著改善了医疗服务的质量和效率。定性研究的结果显示，医疗专业人员和信息技术专家普遍认为标准的制定和推广需要更多的资源和培训，以提高各方的接受度和遵守度。他们也提出了一些建议，包括制定更全面的标准、建立更有效的培训计划以及加强监督和执行。
结论 医学信息标准的研究与应用对于提高医疗信息管理和卫生服务的质量至关重要。本研究强调了标准在医疗领域中的积极影响，但也指出了在标准推广和应用方面仍然存在的挑战。通过更广泛的合作、更多的资源投入以及更加全面的培训，我们可以更好地利用医学信息标准，以实现更高水平的医疗信息互通和卫生服务协调。这将对提高患者护理的效果和医疗体验产生积极影响。

ZY-010

中国大学生对突发公共卫生事件健康信息需求的定性研究

黄丹蓝[1] 赖雨婷[1] 周旭毓[2] 谢瑶[1] 刘靖榆[1] 严波[1]
1.广州新华学院
2.中山大学

目的 本研究旨在通过了解中国大学生关于 COVID-19 的健康信息获取行为，以及对类似突发公共卫生事件的健康信息诉求，总结其对突发公共卫生事件的健康信息需求，为国内健康传播者针对大学生制定健康信息传播策略提供参考。
方法 本研究采用定性研究法，对广州某高校 10 名大学生进行面对面半结构化访谈，并通过 colaizzi 现象学资料 7 步分析法分析数据。
结果 对反复出现的观点进行编码，最终提炼为 2 个主题，包括“对 COVID-19 的信息获取行为”以及“对未来类似突发公共卫生事件的信息获取诉求”。中国大学生主要通过人际交往、新媒体等途径获取信息，其中社交媒体是大学生首选的信息获取渠道，在线视频是其认为最直观易懂、最便捷的传播的渠道；此外，受访大学生追求健康信息的真实性和权威性，会通过官方渠道来获取真实权威的信息，亦会通过关注社交媒体上其他用户评论等方式，甄别信息。对未来突发公共卫生事件的信息获取诉求，主要包括：希望从新媒体渠道获取信息；获取来自政府机构或专家发布的信息；期望获取的健康信息内容包括：疾病基本情况（健康科普信息）、防控方式、疾病后遗症处理、与日常生活有关的健康信息（如不同地区的防控政策）。
结论 面对突发公共卫生事件，中国大学生存在对信息来源真实性和权威性的需求，可呼吁我国建立由专业机构统一管理的平台，专门发布相关信息，同时引导大学生成为准确健康信息的传播者。从信息内容层面，应根据公共卫生事件的不同阶段，及时调整相关信息内容。

ZY-011

基于表型网络的白化病共表型分析与预测研究

郜杨芳* 蔡月琴
山西医科大学

目的 本研究旨在挖掘白化病领域相关文献的摘要数据，通过分析表型共现情况及规律，为白化病早期诊断和治疗提供依据。
方法 采用基于词典的命名实体识别技术，提取白化病文献摘要中的表型信息，并构建表型共现网络。通过社会网络分析方法对网络结构进行深入研究，此外，还运用关联规则方法和链路预测技术，挖掘共现规律并预测未来可能共现的表型组合。
结果 在与白化病相关的表型网络中，眼球震颤、皮肤色素减退和中央凹发育不良显示出重要地位。同时，该网络呈现出小世界特性和凝聚子群特征。通过设置最小支持度和最小置信度，挖掘到16 条表型共现强关联关系，其中与眼球震颤共现的表型最为显著。预测出未来可能出现与白化病共表型及共表型共现等多种表型组合。
结论 分析和预测结果表明，涉及多种表型共现，但目前相关研究仍较少。未来的研究应着重关注多种表型异常共同存在的情况，以帮助不同白化病患者确定最佳治疗方案和进行预后评估。

ZY-012

我国短视频应用研究现状和热点分析

孙伟伟
蚌埠医学院

目的 从年度发文量、期刊来源、文献作者、关键词等内容进行量化分析，阐述国内短视频应用研究的现状和研究热点。
方法 采用文献分析法。

文献资料来源于中国知网数据库（CNKI），时间范围为 2000 年到 2023 年，为保证来源文献的质量，期刊限定为 SCI、北大核心和 CSSCI 三种类型。

检索式为：（篇关摘=短视频）AND（主题=短视频）AND（全文=短视频）；时间：2000-2023；期刊类型为 SCI、北大核心和 CSSCI。初步检索得到相关文献 2907 篇。首先通过辨别篇名、摘要及研究领域等信息剔除与主题内容无关的文献 1576 篇，初次筛选得到文献 1331 篇。然后，通过阅读文献内容，剔除与短视频应用研究内容无关的文献 190 篇，最终得到有效文献 1141 篇。
结果 （1）短视频应用相关的发文数量呈持续增长趋势。
（2）短视频应用研究相关文献主要分布在图书情报领域的期刊，新的社会媒体的应用及社会信息的传播仍然是图书情报领域研究的关注点
（3）短视频研究文献的核心作者为黄楚新、张志安、李静、喻国明、叶铁桥和胡泳。
（4）短视频、抖音、主流媒体、新媒体、内容生产、媒体融合和传播策略等内容是短视频应用研究的热点词汇。
结论 当前短视频应用研究处于上升期，相关研究成果主要集中于图书情报领域，研究内容涉及短视频内容、传播媒介和传播策略等方面，通过关键词分析发现短视频与媒体融合及国家的政策要求息息相关。但在改善短视频传播内容同质化、加强政府监管等相关策略研究方面仍需要加强。

ZY-013

知识产权在创新企业中的发展趋势

张宁雅
中南大学湘雅二医院

目的 随着世界经济步入知识经济时代，让一切的劳动、知识、科技技术等都开始有了自己独特的价值。越来越多的人开始重视这一项“无形资产”，开始重视知识产权的重要性。中国入世已 10 余年，中国企业在国际竞争中仍面临着严峻形势。一个国家经济的发展，不单单只体现在物质经济，知识产权保护更是发展经济不可或缺的成分，也成为了经济上的一种财富资源，在此过程中，知识产权制度起着至关重要的作用。
方法 以创新企业腾讯为例，阐述我国典型企业在知识产权中面临的问题，以及其未来的发展趋势。
结果 对于创新企业而言，创新技术成果是其核心竞争力，创新成果的保护显得尤为重要。今年 6 月，华为就发布了创新和知识产权白皮书，华为表示：知识产权是创新的基础，华为每年将收入的 10%到 15%投入研发，若不尊重和保护知识产权，创新型企业将面临重大的风险。
（1）搭建知识产权管理体系，建立知识产权管理制度（2）建立知识产权风险监控机制，将知识产权管理与企业创新相结合
结论 对于知识产权在创新企业中的发展趋势，笔者有几点猜想：
（1）企业的知识产权投入会有较大的增长
（2）互联网全面介入知识产权行业的各个业务环节
（3）法律领域的人工智能产品将会出现

ZY-014

国内外医学院校图书馆研究文献的可视化分析

蔡汉虹
武汉大学中南医院

目的 了解国内外医学院校图书馆研究的现状及热点，未后续发展提供参考
方法 运用 citespace 软件
结果 我国医学图书馆及医院图书馆领域研究仍有较大发展空间,未来应以服务为导向引入企业管理中品牌建设、构建信息传播等渠道的理念。
结论 计算机检索共获得中文文献 426 篇，英文文 献 166 条。通过共现分析的研究结果显示，总体上逐年增长,美国、英国、中国、澳大利亚是该领域的重要国家。 中文文献主要关键词有图书馆、信息服务、医学院校、阅读推广、医院、知识服务、医学生、健康素养；英文文献关键词包括：libraries、medical library association、medical students、medical、medical library、education、information literacy、health information 等。国外 health、africa、covid-19 pandemic、health library、consumer health information、evidence-based practice、information science、bests practices 是目前研究热点，国内服务、重塑、活动设计、阅读推广、路径、信息服务、应急管理等是目前研究热点。

ZY-015

浅谈元宇宙与图书馆的融合共生

马子龙 王馨
大连医科大学图书馆

目的 本文旨在探讨元宇宙技术在图书馆未来发展中的应用，分析其优势和挑战，并提出相关建议。
方法 观察法、实证研究法
结果 新时代已经到来，图书馆也已经由传统的藏书空间，演变为融藏书、资源、服务、空间共享为一体的知识场所，要致力于为全民信息素养与全民阅读创设更为具体化、更具有置身体验的平台与场景，实现现实空间与虚拟空间的深度融合。
结论 元宇宙是图书馆实现转型的强大助力，元宇宙赋能智慧图书馆是科技助力文化发展的新路径，从传统图书馆、数字图书馆发展而来的智慧图书馆与元宇宙在概念上有许多相似之处，元宇宙与图书馆有着共同追求与发展的“双向奔赴”的关系。图书馆必须主动融入元宇宙浪潮之中，勇于尝试和创新提升用户体验，开创图书馆发展新局面。

ZY-016

努力建设一流的生命医学图书馆

王俊
武汉大学

目的 在国民经济和社会发展第十四个五年规划中期，随着综合性高校、生命医学类院校学科专业和医院事业的持续稳定发展与进步，对作为教学医疗研究支持中心的生命医学图书馆的建设提出了更高的要求。我们需要顺应这一要求，也应该抓紧这一机会，努力建设一流的生命医学图书馆，为生命医学领域师生读者和医务工作者服好务，为高校和医院事业发展提供支持、支撑和推动力，为国民经济和社会发展作出新的更大贡献。质言之，建设一流的生命医学图书馆是建设一流生命医学学科专业的需要，是建设一流大学一流医院的需要，是提高国家战略能力与综合国力的需。
方法 建设一流的生命医学图书馆的方法和途径，一是大力增强硬实力建设，通过建设一流的硬件与丰富的文献资源，并积极推进智能图书馆建设来实施。二是着手提升软实力建设，包括对外服务与内部服务的高质量发展。
结果 从现实上讲，一流的生命医学图书馆通过大力增强硬实力以及着手提升软实力建设，对生命医学领域内读者必然具有全方位服务的功能，同时又具有科学、文化及教育方面全面塑造的功能。一流的生命医学图书馆会塑造出一流的生命医学读者，而一流的生命医学读者群体也必将成就一流的生命医学图书馆，因此实现了馆读的交互、共同成长与相互生成。
结论 随着当代生物学、生命科学及医学学科的迅猛发展，称二十一世纪为生命医学的世纪毫不为过，这有助于我们紧跟时代潮流，在关键领域、核心领域奋力赶超国际上发达国家，而建设一流的生命医学图书馆则是这种赶超的坚强后盾和强大支撑力量，同时建设一流的生命医学图书馆也是这种赶超的标志之一。从理论上讲，建设一流的生命医学图书馆是建设一流生命医学学科专业及一流大学一流医院的需要，也是提高国家战略能力与综合国力的需要。从现实上讲，一流的生命医学图书馆通过大力增强硬实力以及着手提升软实力，对领域内读者必然具有全方位服务的功能，同时又具有科学、文化及教育方面全面塑造的功能。因此，通过回答为什么要建设一流的生命医学图书馆、怎样建设一流的生命医学图书馆，以及建设一流的生命医学图书馆的现实意义，笔者提出以上拙见，敬请方家批评指正。

ZY-017

医院信息化数据平台的应用分析探究

付宇
太和县人民医院

目的 医院信息化在现代医疗领域扮演着至关重要的角色。随着计算机技术的发展和医疗信息化进程的推进，医院信息化已成为现代医疗体系的重要组成部分。医疗数据平台作为医院信息化的核心，集成并管理了各类医疗数据，具有广泛的应用前景。本文分析医院信息化中医疗数据平台的实际应用情况，并探讨其面临的挑战。

方法 2.1 医疗服务

医疗数据平台整合了患者的电子病历、影像数据、实验室检查结果等多种数据源，医务人员可以通过平台获取全面的患者信息。这些信息帮助医生进行准确诊断和制定个性化治疗方案，提高医疗服务的质量和效率。

2.2 临床决策

医疗数据平台通过数据挖掘和分析技术，对大量的医疗数据进行深入研究。基于平台提供的数据支持，医务人员可以做出更准确的临床决策，例如预测疾病风险、选择合适的治疗方案等。平台还可以提供临床指南、知识库等辅助工具，帮助医务人员进行决策。

2.3 研究发展

医疗数据平台为医学研究提供了丰富的数据资源。研究人员可以利用平台中的数据进行临床研究、药物研发、流行病学调查等。平台还支持数据共享和协作，促进多中心研究和跨学科合作，推动医学科学的进步。

结果 医疗数据平台在医院信息化中发挥着重要作用，对医疗服务、临床决策和研究发展等方面产生了积极的影响。仍需解决一些挑战。未来，随着技术的进步和经验的积累，医疗数据平台将发挥更大的作用，给医院各部门和患者提供更为满意和细致的服务，进而加强医患之间的沟通与交流，提升工作质量、效率，增强管理水平。数据平台的建设和运营面临着数据安全性、隐私保护、数据质量和互操作性等一系列挑战。通过采取适当的解决方案只有克服这些挑战，我们才能实现医疗数据的有效利用，并为医疗研究、决策支持和医疗服务的改进提供更好的支持。

结论 医疗数据平台不仅为临床实践提供了支持，还为医学研究提供了丰富的数据资源。研究人员可以利用平台中的数据进行各种类型的研究，包括临床研究、药物研发、流行病学调查等。通过平台中的临床数据可以被用于开展临床研究。研究人员可以分析大规模的患者数据，探索疾病的发病机制、治疗效果和预后因素等。通过对这些临床数据的深入研究，研究人员可以获得新的见解，并为临床实践提供更好的依据。医疗数据平台也为药物研发提供了重要的数据支持。研究人员可以利用平台中的临床试验数据、药物剂量信息和不良事件报告等，评估药物的安全性和有效性。这些数据有助于筛选候选药物、优化药物剂量方案，并加速药物的研发过程。

ZY-018

智慧运营管理助力医院高质量发展

唐宇
河南大学淮河医院

为深入贯彻《国务院办公厅关于推动公立医院高质量发展的意见》,巩固 2020-2021 年“公立医疗机构经济管理年”活动成果，落实 2022-2023 年经济管理年活动继续围绕“规范管理、提质增效、强化监管”的主题，着力推动“以业财融合为重点的运营管理建设，助力提高医疗服务质量、提升资源配置效率效益”。河南大学淮河医院结合自身情况在原有基础上进行智慧运营管理建设，推动人财物一体化管理，提升运营管理质量与效率。

一、国家医改对于运营管理提出新要求

医疗体制改革是国家改革浪潮中最为艰难的一项，而医疗体制改革的深水区是公立医院的公益性改革。新的医改政策要求公立医院承担起提供“安全、有效、方便”医疗服务的社会责任，确保生命安全，提高医疗质量；合理控制医疗费用，提高医院资源使用效率，保障医院的长久持续发展，如下几方面的挑战已经摆在医院管理者的面前：

1.支付制度改革倒逼医院管理重心调整

在全民医保下，医保基金支付已构成公立医院补偿的主要来源。增强医院管理内涵，将离不开全面控制运营成本、实行科学绩效考评、建立更为合理的激励和约束机制、充分挖掘人的潜能、增强整体管控手段等重要措施。

2.政府监管促使医院强化全面管控

作为社会公益事业，“外加压力，内增活力”，既表明政府主管部门对医院发展的期望，也表示出主管机构必将加强行业监管。各级医院要在政府的指导下，以建立现代医院管理制度为目标，规范法人治理结构，提高医院效益和竞争能力。政府监管措施，对医院的管理改革，既是挑战，也是机遇。

二、多举措推进运营管理能力提升

通过深入调研，对医院现状进行梳理，发现医院管理存在业务协同差、风险控制难、精细管理低、机制不健全、管理不规范、反馈机制慢、控制不健全、决策分析慢、战略执行差、成本消耗高等问题。为了满足医院持续发展所需要的运营管理的要求，解决医院的管理难题，制定了以下措施：

1.以咨询顾问导向，营造更加规范健全的制度流程，提高医院经营管理能力

整体思路遵循“目标——风险——控制”的六字方针,从医院运营管理目标出发，围绕医院运营管理的各种风险点，优化完善医院预算、收支、采购、合同、资产、绩效等内部管理制度和流程，实现运营管理工作的规范化、体系化、科学化的目标。

2.以预算为主线，实现事前计划、事中控制、事后核算的全过程管理体系

建立预算分析体系。将医院的资金从源头管控起来，形成一个完整的闭环式管理。让每一个预算指标、每一笔经费支出都可以溯本逐源，让每一项资金的收入都可以追踪到最终的用途与效益真正做到将管控端口前移，深化全面预算管理，推进科室成本的精细化管理，科室预算目标实际执行情况及科室精准成本为绩效考核提供财务维度考核的强大有效数据支撑。

3.以智慧财务为主线，实现完整的账务管理体系，账务的视角贯穿医院运营管理的整体过程中统一账务标准，建立以会计科目为核心的分类标准体系。建立物资分类、固定资产分类等业务管理分类与财务科目之间明确的对应关系；建立统一的信息标准，建立标准的凭证记账规范。

4.实现采购、合同、资产、物资业务流程一体化、全过程管理

对采购、合同、项目、资产、物资等相关制度、业务流程进行梳理诊断，优化业务流程，针对业务关键环节通过大数据分析技术嵌入审批要点、风险防范措施以及制度依据等固定化到信息化系统，实现工作自动找人的工作模式。

5.资产管理全过程跟踪，实现资产全生命周期管理

在固定资产管理方面，建立对固定资产的条码化管理理和全生命周期管理和全过程跟踪。固定资产的购置预算、资产购置、资产使用、资产处置、日常管理、资产清查等业务过程，实现资产变动申请、审批到财务入账的全过程一体化管理。

6.规避履职风险，降低医疗成本，保障医疗安全

围绕预算、支出、采购、合同、资产等经济活动开展运营规范建设，明确按职责权限办事和按流程办事的原则，建立全面规范、公开透明、安全有效的经济活动管理制度，让业务信息流动起来，做到让领导放心、决策有力度，财务人员舒心、监管有尺度，业务人员省心，执行便捷高效。

三、智慧运营管理助力医院高质量发展

通过 1-2 年时间的建设，遵循“管理制度化、制度流程化、流程岗位化、岗位职责化、职责表单化、表单信息化”的思路，我院的智慧运营管理建设取得了卓越成果，已经实现以“预算和费用管控为核心，构建整体多维管控方式”，整体提升财务管理水平；实现医院预算对经济业务活动的

约束，严格执行政府采购规定，降低经济风险；推动经济活动运行的电子化办公，实现医院经济活动业财融合、精细化管理。

通过智慧运营管理建设，我院不仅实现业务信息的自动化、智能化、移动化，提升了核算、控制等流程的效率，提高信息时效性和关联性，实现业务信息的有效公开与共享，为医院开展经济决策提供数据分析。还能够将大量人、财、物的运营数据进行统一，并进行高效分析以及可视化展示，为管理层决策提供数据的支撑，全力支撑“医院高质量发展”的目标。

ZY-019

移动医疗在化疗患者管理骨髓抑制中应用的范围综述

张琳[1,2] 王晓华[1] 杨莉[2]
1.遵义医科大学
2.遵义医科大学附属医院

目的 本研究旨在收集关于癌症化疗患者使用移动应用程序和系统来管理骨髓抑制的研究并进行范围综述，以期为化疗患者管理骨髓抑制不良反应的信息化发展提供借鉴。
方法 在 Pubmed、Scopus、EMBASE、CINAHL、中国知网、万方、维普数据库中筛选相关研究，检索时间为建库到 2023 年 5 月 31 日，提取的数据包括癌症类型、样本量、方法设计、研究重点、测量工具、干预时长、干预类型（应用程序或系统）、测量的定性和定量数据等。
结果 结果 14 项研究的重点；最常见的衡量标准是应用程序为交互式（n = 11），其次是可用性/可行性（n = 13）和患者通过移动医疗报告的健康数据（n = 16）。
结论 事实证明，在化疗期间合理使用移动医疗对于支持监测复查血常规依从性、控制细胞毒性或为患者提供骨髓抑制自我护理的过程至关重要，交互式系统可增强患者参与度，以此增强基于移动医疗干预措施的设计和有效性和增强患者使用移动医疗设备的依从性，值得进一步推广使用，建议未来的研究应包括增强自我效能的移动健康干预计划，以促进患者的骨髓抑制不良反应的自我管理。

ZY-020

“智慧医院”下的医疗装备全生命周期管理探讨

张利中
山西医科大学第一医院

目的 探讨“智慧医院”范围中面向医院管理的“智慧管理”，重点探讨医疗装备的全生命周期管理。
方法 分析所在省市三甲医院医疗装备面临的规模、现状、人员和管理体系等问题，总结医疗设备全生命周期管理的过程和阶段，探讨采用全生命周期管理的必要性和初步分阶段实施的内容。
结果 医疗设备全生命周期管理可以针对不同时期的工作重点实施标准化管理，能够使医院设备管理工作始终集中统一，按计划、按步骤、按规律、按标准有条不紊地进行，便于更好的发挥管理职能，提高设备管理工作的效率和效果。
结论 构建规范化的医疗设备全生命周期管理体系，能够提高医院医疗设备的利用率，对医疗设备的管理水平，降低成本，切实提升医院的整体实力和竞争力。

ZY-021

智慧医疗、智慧服务、智慧管理 三位一体智慧医院建设——智慧医疗在医院高质量发展中的应用

郝心歌
郑州市第一人民医院

目的 随着社会环境发展，人们对医疗的需求和要求越来越高。同时从医院自身长远发展考虑，要求管理、诊疗、服务更加规范化、精细化、智能化，从而催生了医院发展的几个趋势，体现在：医院将越来越依赖信息技术来提高医疗服务质量和效率，包括电子病历、远程医疗、医疗大数据等；人工智能在医疗领域的应用将会越来越广泛，包括辅助诊断、智能手术、药物研发等方面；医院将更注重特色专科发展，提供个性化的医疗服务，例如儿童医院、心脏中心、整形美容等，满足不同人群的医疗需求；同时随着社会老龄化的到来，医院将与养老院等机构进行合作，提供康复、护理等服务，满足老年人的医疗需求；医院会更加注重强化科研创新，加强科研能力，推动医学科技创新，提供更先进的诊治手段和药物，从而形成良性循环，不断提升整体医疗水平。
而在这些发展方向中，信息技术的全面应用成为大势所趋，发挥着无可替代的关键性作用。而其中利用信息技术建设智慧医疗体系，则成为了规范诊疗行为、提升医疗质量、降低医疗差错发生的有效途径
方法 统计分析、列举
结果 智慧医疗发展前景广阔，对医疗水平的快速提升起着至关重要的作用
结论 智慧医疗在当今医疗服务中的应用非常广泛，可以帮助医疗机构提高医疗服务质量和效率，减少医疗资源的浪费，为患者提供更加全面、准确的医疗服务

ZY-022

山西省肿瘤医院规范化诊疗系统的探索与应用

高翔* 田园
山西省肿瘤医院

目的 1、通过肿瘤患者的数据治理，构建肿瘤患者就诊数据画像
基于医院现有信息系统的建设，通过肿瘤大数据的治理与信息集成数据，建立肿瘤患者数据主索引，将患者全病程多次就诊的诊疗信息能够汇总集合，经过数据治理与管理，形成完整准确的患者诊疗数据画像。
2、构建肿瘤诊疗规范化知识库，为诊疗规范化提供业务支撑
基于肿瘤专科的诊疗规范化解读，借助肿瘤单病种专家委员会的密切指导，由经验丰富的肿瘤科临床经验的临床医生完成质控指标的医学转化，形成契合临床，可执行、可衡量的规范解读，并将规范解读转换为临床行为规则，应用到临床实践中来。
3、构建肿瘤诊疗规范化及费用合理化平台，实现诊疗规范提升和费用合理使用
基于肿瘤标准化规范知识库，实现诊疗风险预警、医保费用实时管控、肿瘤诊疗方案智能推荐，从而推进肿瘤全程诊疗规范化管理。
方法 1、研究对象
本研究选取中国医学科学院肿瘤医院山西医院住院患者，主诊断为 17 个癌种（乳腺癌、肺癌、肝癌、结直肠癌、宫颈癌、胃癌、卵巢癌、食管癌、甲状腺癌、肾癌、前列腺癌、膀胱癌、鼻咽癌、喉癌、胰腺癌、淋巴瘤、黑色素瘤）的住院患者为样本，确定肿瘤患者诊疗过程中质量异常值或改进重点，发现该类患者的诊疗合理性可能存在可疑的地方，为肿瘤单病种诊疗质量督导检查、质量改进策略的完善提供参考。

2、数据采集和统计学方法

肿瘤诊疗规范化及费用合理化系统内置数据标准的生产流程，提供标准定义、数据接入、数据标化治理、数据加工和增强、数据质控、数据应用六个流程的基础数据介绍，包括其各个流程的数据概览。

在数据汇聚方面，本系统使用数据库同步技术和 ETL 技术，实现在不干扰医院原始生产业务数据库的情况下，完成对院内业务数据的抽取、同步、汇集，实现针对医院多源异构数据的采集和汇聚。系统在实现院内诊疗数据的结构化方面利用基于字典映射识别技术、NLP 智能语义识别技术实现对文本结构化处理，提供基于 NLP 人工智能识别技术的肿瘤单病种质控的文本结构化：支持现病史、既往史、辅助检查、病理报告、影像报告、出院小结等肿瘤单病种质控相关需求的文本结构化；支持诊断名称的 TNM 分期结构化；支持文本提取手术名称、药物名称、诊断名称、临床分期、TNM 分期、病理诊断、药物类别、药物治疗方案、检查项目名称、放疗、不良反应等信息；支持从文本中提取时间、否定词、灰度词等属性信息。支持诊断归一化、手术归一化、检验归一化、检查归一化、药品活性成分归一化；支持 NLP 实体识别结果的归一化。

3、技术路线

肿瘤诊疗规范化总体建设架构将依平台内置的诊疗规则库、DRG\DIP 分组器，以及智慧医院管理指标体系，通过 AI 大数据算法引擎将院内患者诊疗数据与规则引擎进行患者特征重构，形成针对于特定应用场景的患者质控规则体系。

结果 依托质控规则体系，系统从临床与管理两方面来赋能智慧医院建设。从临床方面，凭借 AI 大数据算法引擎所构建的 17 类癌种的诊疗特征图谱，实现对院内患者诊疗数据的分析，并构建患者的全息数据画像，从而实现医生的诊疗行为、诊疗数据、费用数据等发生变化时，触发相应的质控规则，诊疗过程中为临床医生提供实时提醒、诊疗推荐等功能内容。从管理方面，大数据分析能力将满足医院各层级的职能科室与临床科室的管理需求。对于院级领导，系统通过数据看板的功能，为领导提供全院级的质控分析与费用合理化统计，为管理者指定决策提供完善的数据支撑与决策支持。对于院级各职能部门，系统将提供完善的 PDCA 闭环管理体系，为各部门对质控问题等具体管理问题进行下钻统计与责任回溯提供技术手段与数据支撑，从而推动我院从整体上提升管理水平与规范化管理能力。

结论 1、 提高肿瘤诊疗规范化水平是提升我国肿瘤诊疗能力的重要手段

国家卫健委等 3 部门于 2021 年 9 月印发《关于印发肿瘤诊疗质量提升行动计划的通知》，要求自 2021 年起开展“肿瘤诊疗质量提升行动”，通过加强单病种管理、加强抗肿瘤药物临床应用管理等，推动肿瘤诊疗质量和诊疗规范化水平进一步提升。

2、依托信息化手段构建肿瘤诊疗规范化智能体系

《关于印发健康中国行动--癌症防治实施方案（2019—2022 年）的通知》中也明确规定：实施癌症诊疗规范化行动，探索建立规范化诊治辅助系统，利用信息化手段对医生诊治方式进行实时规范。

目前，全院完成医院管理信息系统 （HIS）、医院临床信息系统（CIS）、LIS 系统、PACS 系统、RIS 系统、病理检查系统等的建设工作，医院在全面应用管理信息系统的基础上，重点突出临床信息系统的建设，近年来又建立起临床辅助决策系统，全院网络运行平稳，系统管理软件已趋于完善，人群数据和临床病例资料丰富、完善，能够很好保证该课题完成质量，这些均为本项目的执行提供了强大的基础数据。

ZY-023

复合闭环下临床用血管理系统的构建

康宁
厦门医学院附属第二医院

目的 输血是抢救危急重症患者重要手段之一。如何有效管理临床用血，一直是医院一大管理难点。我院基于主流用血管理流程——输血申请、发血、输注等，针对重要节点执行不到位、缺少监管拦截等问题，构建并完善临床用血管理系统，实现临床用血过程全闭环管理，保障患者用血安全。

方法 以常规异体输血为例，我院基于临床用血基础大闭环（图 1）设计单一循环的临床用血管理系统，出现以下问题：①用血执行不到位；②转科患者血袋交接不规范（短时间内患者多次转科发生一起血袋漏输注的不良事件）；③输血不良反应上报率不高，与实际不符合。针对问题，不断梳理、细化流程，优化系统，实现了“大环套小环、环环相扣”的复合闭环临床用血管理模式，即在临床用血基础大闭环（图 1）上建立 3 个相关子闭环（图 2）： ①输血不良反应上报子闭环：实现护士上报、医生处置、输血科分析、医务科审核的完整的输血不良反应上报闭环的管理；②转科血袋流转子闭环：实现异体血血袋跨病区流转的可追溯；③检验数据辅助决策子闭环：采集用血相关数据得以建模、分析结果，输血科医师得出用血、用药建议，辅助临床用血。

系统架构

将复合闭环下的临床用血管理模式信息化，构建的临床用血管理系统与检验信息系统、移动护理管理系统等多个信息系统互联互通，系统架构见图 3。

功能设计

针对问题，基于临床用血基础大闭环对应的系统功能模块（如用血申请和审核、配血和发血管理、签收和输注、血袋回收等），设计以下功能：

输血不良反应上报子闭环：结合护士巡视录入生命体征的操作，将输血知识与护理知识等规则智能联动，增加输血不良反应上报自动提醒功能（如体温上升超过 38℃、血压降幅较大时等生命体征变化自动识别和预警），及时提醒护士上报，由医生处置后，输血科分析情况，最后医务部审核相关信息。

转科血袋流转子闭环：增加血袋余量交接的控制。血袋签收后到输注结束过程中，若血袋随转科患者跨病区流转，需由交接人员共同完成交接血袋余量信息登记；若未完成，系统提示交接完成方可操作下一步骤。

检验数据辅助用血决策子闭环：通过规范用血诊疗和护理行为得到翔实的用血数据，得以实现数据建模，由此，输血科医师运用数据和模型（如凝血异常疾病干预的模型）得出患者出凝血的预判结果，为医生在用药、用血上提供合理的医疗决策参考和建议。

输血前评估智能审核功能：系统中配置完善输血前评估等智能审核判定规则，如患者 PLT 结果≥50×109/L 时申请血小板等情况，将备血申请审核前置到备血申请时，节约输血科医师审核申请单时间。

结果 收集 2019 年到 2022 年上半年我院全部异体血用血数据进行分析。

临床用血基础大闭环完成率：通过用血执行报表，对输血护理工作持续监督改进。全院临床用血基础大闭环完成率（输注开始、输注结束、输注中巡视、血袋预回收操作均完成）由 5.85%（2019 年上半年），大幅度提升到 66.92%（2020 年上半年），再上升至 90. 30%（2022 年上半年）。

血袋余量交接执行率：针对异体血输注上线血袋余量交接功能，血袋余量交接的执行率从 37. 50%（2019 年），显著提升到 91. 84%（2022 年的第一季度），直至 100%（2022 年第二季），实现血袋流转全过程的跟踪管理，我院未再发生血袋丢失或漏输注的不良事件。

输血不良反应上报率：生命体征变化自动识别功能和预警功能，极大避免了人为主观漏报输血不良反应的情况，使我院输血不良反应上报率逐步上升至 1. 7%（2022 年），远超国内的报道及中国血液预警联盟发布的 0. 2%。

临床用血决策支持：输血科医师根据用血相关数据和据此建立的出血性疾病智能循证干预模型，为临床诊疗决策提供用血决策。以产科为例，我院输血科规范和指导产科对孕妇产后出血的预防和处理，使我院产科近 4 年来产后出血的产妇大幅减少，产后出血的发生率由 2. 3%（2019 年）下降到 1. 3%（2022 年）。

红细胞用量、出院患者数、手术例数的增长对比：2019 年—2022 年上半年新冠疫情防控期间，患者病种非正常变化（部分病房关闭、重症患者比例增多），导致我院红细胞用量（单位 U）、出院患者数和手术例数环比增长率（图 4），在 2020 年第四季度至 2021 年第四季度间数值波动较大，其余时间，前者与后两者均成正比，且增速低于后两者，整体趋势延缓，表明通过信息化手段，规范了临床用血管理，全流程的执行到位、质量监督有效、改进有针对性，整体提升了我院科学、合理、安全用血水平。

结论 使用复合闭环下的临床用血管理系统，规范临床用血行为，实现临床用血全闭环管理，大幅度提升临床合理用血能力，通过分析用血数据，为进一步完善和丰富用血辅助决策知识库打下基础。通过以问题为导向的项目管理，培养了具备临床思维的信息工程师，临床能更好参与信息化建设，实现临床和信息的双向融合。

ZY-024

基于物联网的智慧医疗信息化关键技术探究

汤友喜
河南省儿童医院郑州儿童医院

目的 在进入到信息化时代后，信息技术已经成为推动社会发展的核心力量，尤其在我国政府提出“互联网+”战略后，各个行业和领域都在探索以信息技术为支撑的新型发展模式，在此背景下，“智慧医疗”这一概念也应运而生，推动了我国医疗事业的信息化以及智能化发展。本文主要针对基于物联网下的智慧医疗信息化关键技术进行分析和探究，希望给予医疗结构以些许参考和借鉴。

方法 针对物联网下的智慧医疗信息化关键技术开展分析和探究

结果 1、物联网与智慧医疗

物联网是一种以信息技术、通讯技术、大数据技术为支撑的一种新型技术。在我国医疗服务领域，物联网技术的应用也更加广泛，其能够实现物与物、人与物、人与人之间的充分联系，提升数据信息的传输和共享程度，更好的为患者提供医疗服务。智慧医疗是近些年出现的新名词，其以大数据、人工智能、物联网等技术为支撑，与传统的医疗服务不同，智慧医疗具有智能化以及信息化特征，可以更好的治疗和预防疾病，在临床医学、医学科研中都获得广泛应用。

2、物联网下智慧医疗信息化关键技术的应用价值

2.1 合理分配医疗资源

我国属于发展中国家，人口基数较大，随着城市规模的扩大，大量居民向城市聚集，造成城市医疗资源较为紧张。尤其在市场经济体制下，人们生活水平不断提升，对医疗服务的需求也持续增加，进一步加剧了医疗资源紧张的现状。通过发展智慧医疗能够为患者提供远程医疗服务，不仅能够节约患者时间，还能够实现医疗资源的优化配置，以满足所有居民的生活需求。

2.2 提升医院服务效率

医院是为广大基层民众提供医疗服务的重要机构，其具有较强的社会职能，而随着我国医疗事业的发展，其服务水平和服务质量不断提升，但是依然难以满足所有患者的问诊需求，通过发展智慧医疗能够提升医院的服务效率。例如通过远程问诊和网络挂号，可以节约患者的问诊时间，对就诊过程进行优化。

2.3 实现资源公平分布

当前我国医疗资源主要集中在城市，偏远山区的居民想要求医问药依然面临着一定的困难，而医疗资源分布不均衡，也对社会稳定带来负面影响。通过发展智慧医疗能够实现医疗资源公平分布。

3、基于物联网下的智慧医疗信息化关键技术分析

3.1 信息融合技术
信息已经成为社会发展的宝贵资源，在开展医疗服务中，每天都会产生大量的信息，信息的交互和传递，能够更好的患者提供优质服务。信息融合技术是一项信息共享的重要技术，例如在一些健康检测设备中，其信息传输方式呈现多元化，包括 WIFI 传输、蓝牙传输以及互联网传输等，可以实现设备之间、科室之间的数据交互和数据共享。
3.2 大数据分析技术
大数据分析技术是物联网的核心技术，其通过数据分析可以为各项医疗工作的开展提供支撑，为医院制定各项决策提供参考依据。尤其在信息化背景下，医院运行每年都会产生海量的信息，通过应用大数据技术能够将其中有价值的信息提炼出来，通过数据整合分析，发挥数据的潜在价值，通过智能化分析能够实现信息资源的高度汇聚与整合，降低医院运行成本。
3.3 远程医疗服务技术
远程医疗是医疗事业的主要发展方向，其通过远程的方式，以信息技术和通讯技术为支撑，实现患者的在线诊治，患者不需要亲自来到医院即可完成看病。在物联网技术的支撑下，远程医疗服务技术目前已经在医疗事业中广泛应用，并且取得显著的应用效果，为患者提供了巨大便捷。而随着信息技术、人工智能技术的发展，远程医疗服务技术的应用范围更加广泛。
3.4 健康数据共享技术
随着我国城市化建设进程的加速，对医疗数据共享也提出新要求，基于物联网下的健康数据共享技术，可以实现医疗数据在全国范围内高度共享，医生通过查询健康数据库，能够获悉患者的健康情况，从而制定针对性治疗方案。在物联网技术、大数据技术以及信息技术的支撑下，患者健康数据共享的程度不断提升，为实现跨地域协作医疗提供了重要的技术保障。
结论 当前我国正处于经济建设和社会转型的关键阶段，城市化建设进程的加速也推动了医疗事业的发展。合理应用智慧医疗关键技术，能够优化医疗资源、方便患者就诊以及降低医疗成本。因此，医疗机构要结合自身实际，积极探索和应用新兴技术，以技术为支撑推动我国医疗事业的稳定以及可持续发展。1

ZY-025

智慧医院建设中的信息技术应用与挑战分析

许赞
开封市人民医院

随着信息技术的迅猛发展，医疗领域正经历着深刻的转型，其中智慧医院作为这一变革的核心，已逐渐成为全球关注的焦点。本文系统地探讨了智慧医院的定义、主要特点以及信息技术在其中的关键应用，如电子病历、远程医疗、人工智能等。同时，文中深入分析了智慧医院在实际建设和运营中可能遇到的技术、管理和法律伦理等挑战，并通过实际案例进行了验证。研究表明，智慧医疗的成功不仅依赖于先进的技术，更需要综合的管理策略和对法律伦理问题的深入考虑。
1、引言
随着全球信息技术的飞速进步，医疗领域也正在经历前所未有的变革。传统的医疗模式正逐渐向数字化、智能化转变，其中，智慧医院作为这一变革的先锋，已经在全球范围内受到了广泛关注。智慧医院不仅仅是医疗信息化的体现，更是通过技术手段，实现医疗资源的优化配置，提高医疗服务质量，降低医疗成本的有效途径。鉴于此，本文旨在深入探讨智慧医院的定义和核心特点，详细分析信息技术在智慧医院中的具体应用，并针对其在实际建设和运营过程中可能遇到的技术、管理和法律伦理等挑战进行深入研究，以期为智慧医院的建设和发展提供有益的参考。
2、智慧医院的定义与特点
2.1 定义

智慧医院是一个综合应用现代信息技术、通信技术和医疗健康科技的医疗机构，旨在提高医疗服务的质量和效率，同时降低运营成本。其核心是通过数据驱动的方法，实现医疗资源的优化配置和服务的个性化。

2.2 特点

（1）信息化基础设施。

（2）优化的医疗流程。

（3）患者中心的服务模式。

（4）绿色环保。

（5）安全与合规。

3、信息技术在智慧医院中的应用

随着信息技术的快速发展，其在医疗领域的应用也日益广泛。智慧医院作为现代医疗的代表，充分利用了这些技术来提高医疗服务的质量和效率。以下是几种主要的应用场景。

3.1 电子病历系统

3.2 远程医疗

3.3 人工智能与机器学习

3.4 物联网技术

3.5 大数据分析

3.6 互联网应用

4、智慧医院面临的挑战

随着智慧医院的快速发展，其所面临的挑战也日益凸显。这些挑战不仅涉及技术层面，还包括管理和法律伦理等多个方面。

4.1 技术挑战

（1）数据安全与隐私保护。

（2）系统兼容性与集成。

（3）技术更新速度与医疗机构的适应能力之间的差距。

4.2 管理挑战

（1）人员培训与接受度。

（2）技术更新与维护。

4.3 法律与伦理挑战

（1）数据使用与共享。

（2）人工智能决策。

5.1 开封市人民医院

开封市人民医院是开封市政府举办的集医疗、科研、教学、预防、康复、保健等功能于一体的大型综合性三级公立医院。2016 年，医院新的领导班子组建后，高度重视智能化、信息化在医疗服务中的实施与应用，坚持以人民健康为中心，着力破解“八需八难”之就医难题，科学谋划医院信息化建设发展规划。经过 5 年的不懈努力，目前已完成 HIS、LIS、EMR 升级换代，引入 CDR、BI、CA、CDSS 等系统，同时，医院完成了“三大远程中心”和互联网医院建设。

具体应用：远程医疗：经建设成为集远程心电、远程影像、远程会诊、互联网医院于一体的“互联网+健康”综合服务平台，可横向对接优质三甲医疗资源纵向促进优质医疗资源下沉基层；大数据分析：通过集成平台完成如病历、检查检验等历史医疗数据的整合，为医院沉淀保留的医疗数据资产，为未来大数据的利用奠定基础；互联网应用：全面实现“通卡就医”，目前患者可通过社保卡、身份证、电子健康卡、电子医保卡在我院进行就诊缴费等的全流程就医，有效得解决了人民群众就医过程中存在的一院一卡、重复发卡、互不通用的堵点问题。

信息化建设存在的问题：信息专业技术人员紧缺，招录高水平信息专业技术人员极为困难，人才培养周期长；信息建设受区域影响，全市没有形成医院信息化建设统一发展规划，医疗信息数据碎片化，医疗机构间未实现互联互通及信息互认，形成了信息孤岛；政府对公立医院投入不足，医院信息化建设资金缺口巨大。

本文深入探索了智慧医院的概念、核心特点及其在现代医疗领域中的重要性。通过对信息技术在智慧医院中的应用进行详细分析，我们认识到了电子病历、远程医疗、人工智能、物联网和大数据等技术如何助力医疗服务的提质增效。同时，我们也对智慧医院在建设和运营过程中可能遇到的技术、管理和法律伦理等挑战进行了深入的剖析。随着技术的持续进步和医疗需求的日益增长，智慧医院的建设和发展将面临更多的机遇。未来，我们预见到更加先进的技术，如量子计算、增强现实和虚拟现实等，将进一步融入医疗领域，为患者提供更加个性化、高效和安全的医疗服务。同时，随着全球对数据隐私和安全的日益关注，智慧医院也需要在确保技术创新的同时，加强对数据安全和伦理问题的研究和管理。

ZY-026

关于移动报账系统在医院的应用

李腾 刘珂* 姜博
郑州人民医院

1、引言

医院现有的财务管理系统仍存在诸多需要改善提高的环节，例：业财系统不互通；线下报销业务审核周期长，核验发票量大，费用无法实时核验是否超标等。

因此亟需在财务管理系统的基础上，建立智慧报账平台，以报销平台为核心，实现线上费用自助报销，报销中自动查询发票真伪和是否重复使用，同时按照报销人级别所享有的报销标准判断是否合理，实现报销单移动审批，提高审批效率和实现报销流程的规范化。

2、现状及目标

医院采购卫生耗材、药品、后勤物资、信息耗材等，从物资采购到支付款项完成，周期较长，每月都需制定付款计划。专项会计付款时，需要手工整理出各物资系统月报表与付款计划进行核对，确认哪些账期的款项还未支付，核对无误之后再进行审批流程，中间过程较为繁琐。

因此在经过详细的项目论证后，确立了智慧报销系统的建设目标：

1、以集团化智慧财务管理运营平台为基础，建立智慧报账系统，内嵌财务政策及标准，实现报销流程规范化、智能化、自动化。

2、实现银行和医院收付款联动，通过银医直连系统进行报销付款，及时将报销款项支付给职工或供应商。

3、通过费用报销系统和全面预算系统集成，实现预算业务的事前和事中控制。

方法：3、建设内容

3.1 建设思路

通过建设报销申请、审批、报账、发票识别验伪等服务，同时作为业务管控的端口进行费用预算控制，业务审批完成后自动对接财务核算系统。上线后，职工只需三步（填单-传票-交票），即可完成报销而且可随时查看差旅报销标准，跟踪报销进度。实现以业务为核心，提升财务智能化能力，加强财务管控。

3.2 建设内容

3.2.1 报销制度和标准在线查询

系统内提供了在线查询报销制度及各地市住宿标准功能，根据出差地点和职务职称即可定位出差地住宿标准。

3.2.2 事前申请

提供了“公务外出申请”、“外出参会/培训申请”、“举办会议申请”、“日常费用申请”填报入口，用户在报销系统中直接填写申请单。

3.2.3 事后报销

提供了“公务外出报销”、“外出参会/培训报销”、“举办会议报销”、“日常费用报销”、“外出进修报销”、“劳务费报销”填报入口。

3.2.4 发票智能采集
可通过拍照识别、微信卡包发票导入、PDF 电子发票上传等多种归集方式，将票据与报销流程结合，并根据相关规定自动生成税额。
3.2.5 智能稽核
可灵活设置票据、单据的稽核校验规则。票据识别后，可自动进行票据合规性校验，包含且不限于发票真伪性、电子发票重复报销、发票抬头、税号错误
3.2.5 快捷支付报销款
通过与银企联云服务的深度融合，可以方便、快捷实现医院在线支付职工报销款和对公业务报销款。
3.2.6 费用预算控制
根据全面预算控制的颗粒度灵活配置预算编制的维度，如按照单据类型（申请单、报销单等）、组织、部门、人员、费用类型等维度自由组合维度进行费用控制，做到事前、事中预算预警、预算强管控，并提供了预算数、发生数、可用额度实时分析查询。
3.2.7 自助收单接收报销资料
便捷院区人员进行报销资料提交，引入了智能收单机器人，可实现申请单和报销单封皮打印、纸质发票与上传发票信息校验、纸质资料投递及自动装订功能。
3.2.7 临床人员外出进修报销由教学部代办调整为进修人员自主报销
3.2.8 职工实时掌握报销进度
报销人员自主查看审批流进程，实时掌握报销业务进展。
3.2.9 领导审批
审批人员不受时间、地点的限制，在钉钉内随时随地进行报销业务审批。
4、讨论
4.1 通过建设智慧报账系统，重塑内控流程，降低了财务风险。
通过建设智慧报账系统，满足医院全员报销应用的需求，内嵌财务政策、预算标准和报销标准，减少手工核算、规范报销流程，简化手工报销方式。
4.2 建立资金管控平台，降低财务支付风险
通过智慧报销系统和银行支付系统的连接，审批完成的报销单据通过银医直连接口，快速、安全的支付职工报销款和对公业务报销付款，节省人力成本，降低支付风险，保障资金安全。
4.3 预算管控通过智慧报账实现事前控制
医院全面预算系统和智慧报销系统相结合，通过智慧报销过程中的预算控制，实现全面预算的事前控制，支撑医院战略目标的实现。
4.4 智慧报销系统和财务管理系统无缝集成，保障财务会计信息质量
智慧报销系统帮助提高财务会计信息质量以及财务分析水平，通过系统自动归集和智能核算，将财务人员从日常基础性事务解放出来优化资源配置，实现财务转型。
5、结语
智慧报账系统存在的问题和困难都可以通过技术改进、管理优化等多种手段解决，只要有足够的准备和技术支持，就能实现更高效的报销业务管理。
从未来来看，智慧报销业务将会继续和院内其他管理系统进行互联互通，满足更多的业务应用。

ZY-027

智慧医疗对患者就诊，医护工作的优势分析

马欣钰
哈尔滨医科大学附属第六医院

目的 大数据时代，信息化系统非常重要。信息系统是由计算机硬件、网络和通信设备、计算机软件、信息资源、用户、规章制度组成的以处理信息流为目的的人机一体化系统。本文主要介绍了智慧医院信息管理系统包含的具体内容以及发展史，分析了一些信息管理系统建立的重要性与必要性。信息化系统在医院的发展在近几年非常迅速，主导其飞速发展的因素都有哪些？随着人工智能和计算机水平的飞速发展，临床工作现在几乎已经脱离不了信息化系统，比如 His（Hospital Information system）系统，Lis（Laboratory Information System)系统，Pacs(Picture archiving and communication system)系统等。智慧医疗目前已经覆盖了大多数医院。为什么智慧医疗能受到医疗界的如此好评和依赖。到底什么是智慧医疗？本文详细阐述了智慧医疗建立后所包涵的模块与建立前形成了哪些对比，说明一些板块建立的意义和优势，为患者就诊，医护日常工作带来了哪些实质性的方便。哪些板块是非智慧医疗模式达不到的，如果达不到会对临床工作带来哪些不便。说明智慧医疗帮助医疗模式从根本上解决了哪些问题，优化了哪些步骤，具体为医疗体系带来了哪些飞速的发展。除了医护工作者，对于医疗管理岗位，信息化系统也提供了很大的帮助，包括一些临床数据的统计，病案系统的管理在未来，我们还有哪些更高的期望，想要去实现哪些功能。
方法 大数据，人工智能，云数据
结果 未来，我认为人工智能帮助医疗界能做到的事情远不止这些，我们现在就可以了解到有些医院已经开始用机器人来做手术，这就是一个良好的开端。人工智能智慧医疗势必是以后的发展趋势，我们甚至可以利用它在线上帮助患者分析疾病及诊疗方案，让患者更加放心，远离焦虑情绪。智慧医疗是深化我国医疗改革的重要手段，也是我国新型城镇化，特别是智慧城市建设的重要组成部分。如能从多个维度推动智慧医疗建设，将加快健全基本医疗卫生制度，不断提高医疗卫生水平，早日实现健康中国的战略目标。
结论 如能从多个维度推动智慧医疗建设，将加快健全基本医疗卫生制度，不断提高医疗卫生水平，早日实现健康中国的战略目标。

ZY-028

统一预约系统助力传染病医院提升就医体验

张现辉
郑州市第六人民医院

目的 1.1 贯彻落实改善医疗服务、便民惠民利民的政策背景。　　为贯彻落实深化医药卫生体制改革和公立医院服务管理模式的创新改革，河南省委改革办、河南省卫生健康委、河南省医疗保障局推行“便民就医少跑腿”七大举措相关要求，以人民群众看病就医感受为出发点，立足解决群众看病就医过程就诊关键问题，特别是对门诊与住院患者开展预约诊疗服务，提升就医体验。
1.2 传统预约模式在传染病医院弊端太多。
普通疾病的患者与传染病患者的检查队列重合，时空重合，易造成交叉感染、患者精神紧张甚至投诉。
对于急诊患者队列和普通病人混在一起造成医患工作被动。
传统的预约模式，为在就医体验、医院资源利用、医院管理等方面的弊端。
患者重复排队、登记。传统情况下，检查科室比较分散,检查时间不固定，各种排队在患者就诊环节占 60％以上的时间。

人员利用率低。患者反复排队工作人员则需为同一患者服务多次，降低了工作效率；此外，由于患者在医院停留时间过长，医院还要专门安排人手负责维持秩序，增加了人力消耗。
管理滞后，解决问题不及时。由于检查预约分布在各科室，无法做到实时调控患者检查时间，尤其在需要做多项检查时，经常出现假性的“忙闲不均”情况，且“闲”并非意味着真没有检查患者。
其他影响预约效率问题。如患者不符合检查条件只能重新预约；预约时间长造成“无效预约”，系统流程原因，等等。
所以建立医院统一预约系统势在必行。
方法 统一医技预约管理平台是改善预约诊疗服务的良好手段。
2.1 传染病医院统一医技预约平台的实践方法
推行全场景预约流程，将放射、超声、胃镜、CT、MR、网络心脑电等科室资源进行整合，不仅仅实现一个信息系统，更重是要打造一个支撑预约业务持续优化的工具，为患者省去 2 次排队等待的环节。
2.1.1 实现精细化预约
系统实现可支持自定义预约策略，根据各种来源的病人以及实际检查量设置预约策略，协助医院服务中心进行统一的精细预约管理，传染病疾病患者、急诊患者、普通病人均是按精细化预约思维进行登记安排。
2.1.2 实现可视化管理
设计可视化的报到单，清晰显示检查室、排队号、具体检查时间、排队人数；可视化队列显示，候诊大厅和检查区实现二次分诊。
2.1.3 协助实现医院业务可持续优化改进
设置互斥项目、自动安排不同项目的就诊时间段。
实现预约提醒功能。预约完成后，在合适时间将通过短信提醒患者。
2.2 优化医技预约检查流程，整合医院检查资源
医技预约管理系统应用全面服务于医院门诊、住院、医技部门等科室，按照各项业务的特点灵活配置符合医院的预约流程，尽可能的减轻患者预约医技项目的难度，优化院内业务衔接流程。
2.2.1 门诊医技预约流程
医技部门设置号源池，预约工作由门诊医师在开单后直接按号源预约，门诊医师和患者沟通检查事项，提前了解预期检查时间点。
2.2.2 住院预约流程
医技部门设置号源池，预约工作由病房护士审核完医师开立的检查医嘱后直接按号源预约；预约完成后短信推送至患者手机，病房护士可直接打印预约单给患者。
2.2.3 特殊医技项目预约流程
需医技部门评估后方可预约的检查项目，如无痛胃肠镜、pet-ct 等检查项目，由于检查的特殊性，由医技科室自行维护号源，医生为患者开单后，患者需到医技部门沟通宣教检查事项，评估患者状态后才能预约检查。
2.3 安全设计
医院预约管理系统与医院内部信息系统对接，涉及信息包括：病人的个人信息、诊疗数据、电子病历、住院信息等。这些业务信息遭到破坏或失窃程度表现为严重损害，较大范围的不良影响等。因此本系统承载的业务信息数据安全保护等级不低于第二级。
结果 通过统一预约管理平台建设，以“预约资源统一化、管理资源科学化、资源接口开放化”为目标，合理开放医院资源给患者使用，实现从“单模块预约”到“全预约”、减少患者排队、提高医护工作效率、提高科室设备资源的使用率。
结论 结论与效益分析
河南省传染病（郑州市第六人民医院）的统一预约系统已经稳定运行 3 个月，系统全面覆盖了医院门诊、住院等业务环节，解决了各类预约项目、排班、规则的维护，系统功能延伸至医生站、护士站、医技科室、综合服务台等需要使用该系统的科室，整合所有患者在检查过程中使用的医疗资源，改善了医技人员手工记录预约信息效率低下，无法快速处理患者改约，爽约，降低科室工作效率，改善了患者无法自主选择合适的检查时间，被动接受安排，遗漏相关检查注意事项的问题，缓

解了患者长时间在检查科室等待的问题，更好的实现了“以患者为中心的”服务理念，为患者计算最优检查次序和检查路线，同时也使医技科室的医疗资源得到合理分配让医技科室的工作负荷更加合理与均衡。
在医院未来的规划中，统一预约系统将在院内所有检查科室推广，进一步推进便民就医政策的落地。

ZY-029

论述智能医学的应用现状与创新发展

张嘉成
无锡市人民医院

目的 智能医学是医学重要发展方向与技术前沿，能够促进医学进步，提高医疗水平，研究人的生命和疾病现象的本质及其规律。
方法 提炼分析智能医学的现状与发展趋势、发展与应用前景，运用 SWOT 分析法分析智能医学的优劣势、存在的机会与挑战，运用 GE 行业吸引力矩阵分析法，进一步分析智能医学的发展潜力。
结果 智能医学通过技术的不断提升与进步，有望为医疗领域带来更加精准、高效和个性化的服务，改善人们的健康状况和生活质量。
结论 智能医学通过技术的不断提升与进步，有望为医疗领域带来更加精准、高效和个性化的服务，改善人们的健康状况和生活质量。

ZY-030

医院智能随访系统的探讨

张骏
无锡市人民医院

目的 本文对医院智能随访系统的发展、系统架构流程和应用进行了详细探讨。
方法 医院智能随访系统的发展从早期的电子健康记录到现代智能虚拟助手，医院智能随访领域的演进标志着科技和医疗实践的不断结合和互相推进。随访系统的架构流程涵盖了数据收集、处理、分析和患者与医疗专业人员的交互。这一流程以生物参数传感器、患者报告和医疗记录的数据收集为起点，通过数据传输和云存储将数据传输到服务器，随后进行数据处理和分析，运用机器学习等技术提取有用的特征，最终以虚拟助手、应用程序等方式进行患者和医疗专业人员的交互，同时通过自动化提醒和提醒功能增强了患者的自我管理。在应用领域方面，它涵盖了慢性病管理、老年护理、预防性医疗、健康宣教等多个领域，为各种类型的患者提供了更好的医疗支持。
结果 医院智能随访系统在提高医疗效率、个性化医疗护理和患者参与方面发挥了积极作用
结论 等级医院考核和医院管理要求对患者进行随访和追踪，在《医院智慧服务分级评估标准体系》、《医院分级管理标准》等规范和标准中，对随访体系的建设也有相关的指南和规定，通过建设以患者服务为核心的智能随访系统，实现医疗机构与患者的长期联系，以及提高患者的依从性和就医体验。同时帮助医护人员更全面、高效、持续地管理患者，收集特殊疾病数据，减轻医护人员工作量，不断提升医疗质量和科研水平，帮助医院完成智能化服务水平评估和电子病历评估，为了改善患者医疗体验提供了新的机会和可能性。

ZY-031

人工智能在医疗信息服务中的应用研究

朱佳棋
无锡市人民医院

目的 人工智能在医疗、金融、教育和国防等多个垂直领域有着广泛的应用。①一方面来说人工智能技术的迅速发展，能够为医疗服务机构实现信息化管理和服务的改革提供更加完善的技术支持。另外，现在人们都希望得到更好的医疗信息服务，医疗信息服务的效率和质量变得愈加重要。本文的研究目的有以下几方面：
（一）阐述人工智能技术特点和应用价值。
（二）从综合医院系统、社区卫生系统、家庭健康系统三方面，阐述人工智能在具体场景中的应用情况。
（三）根据人工智能在不同医疗信息服务中的具体应用，就其中存在的问题展开分析。
（四）针对存在的问题，提出人工智能在医疗信息服务中的应用策略。
方法 （一）文献分析法
文献分析法主要是通过知网等数据库平台，检索和搜集同“人工智能”“AI” “医疗服务”“信息服务”“智慧医疗”等关键词相关的文献，进行阅读和分析。通过对此类文献的研究，划分出国内和国外的文献，进行分类研究，对其进行归纳总结，提炼其核心思想，总结现有研究的结果，在此基础上，进行深度探索和发掘并得出结论。
（二）比较分析法
本文通过比较国内外相关研究，在应用方面，对综合医院系统、社区卫生系统、家庭健康系统的信息服务进行比较，对具体应用场景中人工智能存在的优势和不足进行分析，探究医疗信息服务领域可以高效、高质量应用人工智能的举措。
（三）案例分析法
本文在医疗信息服务不同应用场景中列举案例，对人工智能在综合医院系统的辅助诊断信息服务、医学影像信息服务等应用，以社区卫生系统的慢病管理信息服务、疾病转诊信息服务等用，以家庭健康系统的自诊导诊信息服务、健康管理信息服务等应用结合案例进行分析。
结果 智慧医助理在社区卫生系统的重要辅助作用，智医助理能够作为医生搭档为医生提供疾病诊断信息，为社区患者进行疾病转诊提供及时的医疗信息服务。
结论 本文研究了人工智能在医疗信息服务中的实际应用、面临的挑战以及相应的优化策略。首先，我们从社区卫生系统的角度探讨了医疗信息服务的具体应用场景。然后，针对人工智能在医疗信息服务中存在的诸多问题，如智能管理缺乏灵活性、影像数据获取难度大、智能辅助诊断有偏差、隐私数据存在泄露风险、慢病管理操作难度大、转诊前后信息被割裂、智能自诊诊断有偏差、算法分析预测不透明以及健康数据采集精度低等，进行了深入的分析。最后，我们提出了提升管理的灵活性、增加数据获取来源、优化辅助诊断结果、加强隐私数据保护、优化技术简易操作、创新信息交互服务、优化自诊数据处理、建立算法管理标准以及提高数据采集精度的优化策略，以期推动人工智能在医疗信息服务中的应用和发展。

ZY-032

健康管理中心智能导检方案的研究与应用

熊佳佳
重庆医科大学附属第一医院

目的 人们的健康意识逐步提升，对健康体检的需求迅速增大。但是完全依靠传统的人工排队等候的方式已经不能满足迅速扩大的健康体检的需求。

1　传统的人工排队方式存在的问题
第一排队等待时间过长，
第二现场秩序混乱，
第三人力成本大，
第四体检者满意度低。
方法 2　智能导检方案的设计
为解决体检中心体检排队问题，改善目前健康管理中心体检排队秩序混乱、需要大量人工导诊的现状，智能导检系统是根据健康管理中心特殊情况研发的智能导检系统，制定合理的排队流程，与体检管理系统的无缝连接，为体检客户提供“全科室全流程的智能导诊”[2]　服务，解决插队、医生指定、医生权限的控制。
2.1 时间优先
体检客户的等待时间是判断智能导检方案最有效的指标。时间优先是指根据医院实际情况设置楼层、检区、项目时长、空腹等各种权重，计算出体检者的最少等待时间的最优队列，指引体检者前往最少等待时间的科室进行体检[3]　。当体检者完成该科室的检查后，再次根据剩余项目的权重等计算出剩余项目中最少等待时间的最优队列，以此类推，直到体检者完成最后一个科室的检查。
2.2 项目依赖
根据医院运行实际情况设置体检项目之间的依赖关系，以实现优先项目和滞后项目之间的配置管理（或者某些科室检查优先），比如动脉硬化检查前需完成身高体重血压检查等，避免体检者在同一个科室多次排队的情况。
2.3 区域优先
根据医院运行的实际情况，将某些就近诊室划分为同一区域。根据科室间的两两距离，计算体检者所做项目的总时间[4]，优先选择同一区域内的项目检查。
2.4 项目互斥
在体检过程中，部分项目之间存在着互斥关系，比如做了妇科 B 超就不应该做阴道 B 超，做了肝功 3 项就不应该做肝功 4 项。在智能导检系统中设置这些项目间的互斥关系就可以达到这种效果。
2.5 预留体检
对于特别贵宾或是临时需要插队的体检者，可通过授权取预留号（灵活配置）的方式来实现。该方式可以进行灵活配置，目前系统支持两种预留号形式：第一种可支持逢号预留，逢尾号为几预留，如逢尾号 3 为预留号；第二种支持隔几位为预留号，如每隔 2 位为预留号。可根据医院运行实际情况选择配置其中一种预留号形式。
结果 智能导检系统从 2022 年使用至今，实现了体检者全流程的“智慧导诊”服务。通过各种权重设置、智能排队算法、时间优先、区域优先等规则[9] 合理安排体检者的检查科室，让体检者以最少等候时间完成体检项目[10]　，同时可通过平板端、自助机、微信端等方式查询体检者的体检排队情况，为体检者提供更方便、更快捷的体检服务。
4.1 缩减体检者排队时间
健康管理中心在未使用智能导检[10] 系统前采用的是传统的人工导检模式。体检者来到健康管理中心开单缴费完成后，根据现场指引人员的人工指引去到相应的诊室排队检查。当体检者完成该诊室的检查后再询问现场指引人员需要去到的下个排队的诊室。根据统计，使用智能导检系统前体检者的平均每个诊室的排队时间在 23.8 分钟，而使用智能导检系统后体检者的平均每个诊室的排队时间在 12.9 分钟。智能导检系统大大减少了体检者的排队等候时间，也就缩短了体检者完成所有体检项目的时间，从而大大提升了体检者的满意度。
4.2 保证体检现场井然有序
传统导检模式下的健康管理中心都是依靠体检者自己排队，无论在哪个诊室外面都是排队等候的体检者，造成一种“人声鼎沸”的现象。而智能导检系统可以消除这一现象，智能导检系统可以给体检者提供多种渠道的排队查询方式，同时也可以在即将到达该体检者时进行提醒，从而避免体检者都在所有诊室外面堆积。同时智能导检系统可以消除体检者人工排队的心急、心烦、心慌等情绪，减少了现场的嘈杂因素，保证了体检现场的井然有序。

结论 智能导检系统自使用以来取得了良好的效果，实现了全流程的排队叫号。根据“时间优先、项目依赖、区域优先、项目互斥、预留体检”等多种规则设计智能导检流程，根据诊室排队情况实时调整体检者体检路径，缩短了体检者排队等候时间，提升工作人员的工作效率以及服务质量。

ZY-033

肺腺癌预后相关转录组分子免疫分型及其特征综合分析

谢戴鹏[1] 尤培蒙[2,3] 周海榆[1,2,3]
1.广东省人民医院（广东省医学科学院）
2.江西省肿瘤医院
3.江西省肺癌研究所

目的 非小细胞肺癌作是世界范围内造成重大疾病负担的癌症，免疫治疗作为癌症治疗的具有巨大前景的治疗模式，其治疗疗效与肿瘤免疫微环境（Tumor immune microenvironment，TIME）存在密切相关的关系，肿瘤的个体异质性和瘤内异质性给 TIME 的评估带来极大难度，导致缺乏能够进行免疫治疗疗效及预后预测的生物标志物。在前期研究中我们已经发现在肺腺癌患者中存在一组与预后相关的 TIME 分子分型，本研究将针对肺腺癌患者借助转录组数据探索具备一定临床价值的 TIME 分子分型及探索该分型的特征。
方法 本研究基于 TCGA 数据库和 GEO 数据库中非小细胞肺腺癌（Lung Adenocarcinoma，LUAD）患者的转录组学数据。借助 xCELL 算法对 LUAD 患者的免疫微环境进行分子学评估，并通过使用无监督聚类（K-means 聚类）机器学习方法对通过对免疫微环境组分状态进行非先验性分类。利用 Kaplan-meier 进行生存分析。基于 RNA 转录组数据通过差异基因表达、GO/KEGG 分析、GSEA 分析探索通过无监督聚类获得与肿瘤预后相关的分子分型在基因组表达与通路富集特征，通过 STRING 数据库构建通路差异基因表达蛋白质互作网路可视化功能蛋白调控模式，同时通过体细胞突变分析确定并比较免疫新分型肿瘤突变负荷差异与驱动基因突变谱。通过 TIDE 分析模拟 NSCLC 患者接受免疫治疗预测反应及计算肿瘤免疫浸润状态，评估不同分组的差异。
结果 本研究通过将 TCGA 队列中的 LUAD 患者队列通过 xCELL 算法利用 RNAseq 数据进行免疫微环境浸润计算，并通过无监督 K-mean 聚类获得具备明确生存差异的两组病人（Cluster1 具有较好生存）。对这两类存在明显生存差异的患者进行了免疫微环境景观分析，研究发现了在 DC 细胞类群如（aDC 细胞、cDC 细胞、DC 细胞、iDC 细胞和 pDC 细胞）、细胞免疫相关的细胞（如 CD8+ 与 CD4+T 细胞）、固有免疫相关细胞（B 细胞等），巨噬细胞等在 Cluster1 上明显富集程度更高。在检查点的表达上（如 PDCD1、CD274、CTLA4、CD28、CD40 等）在 Cluster1 中也存在较为明显的高表达。通过对两组患者的差异基因表达及通路分析得到差异基因 3362 个，其中上调基因为 1665 个，下调基因为 1697 个，GO 通路分析中提示，在生物学过程通路中主要富集在体液免疫相关及免疫介导激活等相关通路；细胞成分定位在细胞膜表面，主要为免疫识别和免疫激活功能蛋白；分子功能上主要表现为抗原抗体结合激活通路。KEGG 通路分析中主要体现在受体-配体结合和信号传递通路。进一步通过 GSEA 分析两个分组之间的生物学行为差异得到主要影响了细胞代谢和免疫活动激活上。在体细胞突变层次上研究发现了 Cluster1 主要突变谱系集中在 TP53（40%）、TTN（38%）、CSMD3（34%）、MUC16（33%），cluster2 主要突变谱系集中 TP53（52%）、TTN（49%）、MUC16（43%）、RYR2（39%）、CSMD3（39%）。研究进一步通过 TIDE 分析工具利用转录组数据模拟免疫治疗疗效，结果显示在两个分组中，Cluster1 的 T 细胞浸润程度较好，Cluster2 较差。在免疫细胞排斥上 Cluster1 排斥程度较 Cluster2 好。在免疫检查点结果分析上，CD274、CD8、IFN-γ 均是 Cluster1 表达较高。研究借助差异基因进行免疫-预后相关模型构建，通过随机森林进行特征工程，纳入%IncMSE top 20 的基因进行逐步回归筛选并进行 logistics 回归获得最终模型，在 TCGA 队列的训练组及测试组中 AUC 分别可以达到 0.98[0.968,0.992]、0.966[0.946,0.986]，该模型利用 GSE30219（n=85），GSE31210（n=204），GSE37745（n=106），GSE50081（n=128）共 523 例非小细胞肺腺癌患者形成的队列作为外部验证组进行模型

验证，结果提示被区分的两组患者在 OS 上存在明确差异，并呈现出与无监督聚类获得的分组具有相类似的微环境状态。

结论 通过对 LUAD 免疫微环境进行分析，并借助该特异性分组的差异基因提取与识别模型构建，研究实现了对 LUAD 患者免疫治疗疗效及预后的识别。研究提出的具备明确免疫微环境差异的两组患者，主要体现在体液免疫相关通路的激活上，作为细胞免疫反应的重要补充，这与先前的研究结果相符。而体液免疫对 LUAD 患者的免疫治疗效果之间的关系现阶段阐述较少，仍需要进一步的研究明确两者的相关关系。该研究获得的基于免疫微环境差异的 5 基因预后及免疫治疗疗效识别模型给该群患者的预后识别提供了潜在手段。

中华医学会
CHINESE MEDICAL ASSOCIATION
1915

中华医学会医学信息学分会
CHINESE SOCIETY OF MEDICAL INFORMATICS
CSMI
1993