

临床科研共享系统质量控制体系升级^{*}

胡彭博 卢程 陈南

(湖南中医药大学第一附属医院 长沙 410007)

〔摘要〕 以临床科研共享系统升级过程中建立的研究平台为基础,进行科研数据结构优化,分析原有数据质量的问题,对原有存储方式进行优化,建立新的数据质量控制体系,使临床科研共享系统与医疗大数据时代、精准医疗的未来发展方向接驳。

〔关键词〕 临床科研共享系统;结构化数据;数据质量控制

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2017.08.008

Upgrading of the Quality Control System of the Clinical Scientific Research Sharing System HU Peng-bo, LU Cheng, CHEN Nan, The First Affiliated Hospital, Hunan University of Chinese Medicine, Changsha 410007, China

〔Abstract〕 Based on the research platform constructed in the process of upgrading of the clinical scientific research sharing system, the paper optimizes the scientific research data structure, analyzes the problem of the original data quality, optimizes the original storage mode, and establishes a new data quality control system, so as to connect the clinical scientific research sharing system with the future development orientation of the medical big data era and precision medicine.

〔Keywords〕 Clinical research sharing system; Structured scientific research data; Data quality control

1 引言

临床科研共享系统是国家临床研究基地业务建设的重要内容之一,是应用临床实际诊疗数据开展

真实世界临床研究的关键技术平台,是开展真实世界研究的前提和基础。在临床科研中,实时连续地采集并整理包含病历资料、处方、临床检查等具有临床价值的数据,发挥大数据的作用,是中西医精准医疗发展的当务之急。湖南中医药大学第一附属医院作为国家中医临床研究基地之一,承担着大量的基础研究和科研攻关任务,2012年部署实施临床科研共享系统,对解决医疗信息孤岛,实现临床、科研、教学和社会的资源共享发挥了重要作用。但随着新的业务与需求不断涌现,其原有的数据质量控制体系和不够完善的数据结构,已经无法满足临床研究基地的建设要求和日常科研需要,同时也严重制约着与其他科研单位的交互合作。在临床科研共享系统应用过程中,基地科研人员面临的重大问题不是缺少临床科研数据,而是面对太多的日常数

〔修回日期〕 2017-05-25

〔作者简介〕 胡彭博,硕士,助理工程师,发表论文1篇;通讯作者:陈南。

〔基金项目〕 中医药部门公共卫生专项中医药防治重大疑难疾病临床服务能力建设项目(项目编号:2100409);湖南中医药大学第一附属医院国家中医临床研究基地中医药防治重症肝炎临床防治中心开放基金项目资助(项目编号:ZZGY12)。

据，缺乏工具软件来筛选、检查和提取关键数据，不知如何挖掘数据之间的精确关联性^[1]。而且，由于中医科研数据较西医科研数据更加复杂，歧义性大，碎片化大，而基地目前体量的科研数据已经超出研究人员的理解和处理能力，因此必须研究如何通过共享系统的整体升级来尝试解决这些问题。

为使系统更好地服务于临床和支持科研，保证临床研究基地科研数据、工具和水准的先进性，在临床科研共享系统升级工作中，必须研究如何从底层中西医数据的结构上对共享系统进行升级优化，以达到系统支持基于第 4 范式的数据挖掘和科研演绎推理体系的目的。同时，尝试整合无法用常规电子病历软件进行捕捉、管理和处理的以中医特色为辅的医疗数据集合，建立支持更强决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的基于大数据的新科研模式。此外，尝试从医疗数据的准确性、医学术语的信息规范化、属性选项的标准化等方面着手强化对数据质量的控制^[2]，优化临床科研共享系统的表结构和字段结构，将数据属性精细化，逐步建立一对多以及多对多的标签化属性来增强医疗数据属性的可复用性和逻辑关联性，为质控体系的优化升级提供支撑。

2 优化数据质量控制体系

2.1 原有数据质量问题分析

原有数据存在如下问题：（1）大量关键核心数据为医护人员自由录入，病案书写方式随意，医疗数据格式缺乏规范，空白率和错误率得不到有效控制。（2）模板设计欠完善，不用模板或没有对模板必须包含的元素进行审核，没有将所有有用的元素用标签属性体现并层次化存储，模板元素属性之间没有建立关联，属性字典缺乏标准化约束^[3]。（3）医疗数据颗粒度大，数据与数据、系统与系统之间没有建立联系，无法及时发现数据不一致或者数据元素之间的逻辑错误。（4）缺乏有效的质控纠错机制，仅仅靠定期抽查纸质文档或电子文档的方式并

不能保证数据质量，仍然存在大量误填、漏填及逻辑冲突等错误。受系统功能限制，无法对底层数据信息元素的准确性和完整性以及是否按照字段结构存储进行检查。（5）对科研数据采集人员而言，数据的采集格式和要素属性设定自主度不够，智能性缺乏，软件开发者既无法充分理解临床科研应用的拓展性思维，同时其开发的程序也无法充分满足临床科研的需求。（6）原有对科研数据的采集和加工模块以西医标准为主，没有兼顾到中医。随着中医药信息化的发展，以及中医数据各种字典库和标准的建立，对中医科研数据的采集应用已迫在眉睫。

2.2 原有数据质量存储方式优化

原有中西医科研数据以稀疏矩阵 3 元组的形式存储在数据库中，具有高维、稀疏、噪音、不完整等特点。就医学研究而言，无法有效度量和利用实体间的语义联系；同样就科研数据本身而言，存在事实不完备和关系不平衡性、条目缺失、事实缺失和知识框架缺失等问题，导致在数据基础上的科研无法顺利推进。同时，中医科研数据经常面临信息描述模糊、歧义性、含干扰因素等情况，而目前研究者对中医知识的系统化归纳较西医知识而言存在更大的局限性。对中西医科研数据的存储优化，通过改造数据库结构实现保存初步的实体属性语义关联，见图 1。将语义关联分为中医和西医两种方式，着重于保存同义词关系、语义范围不同的包含关系、属性与发病时间或疾病发展阶段的关联关系。

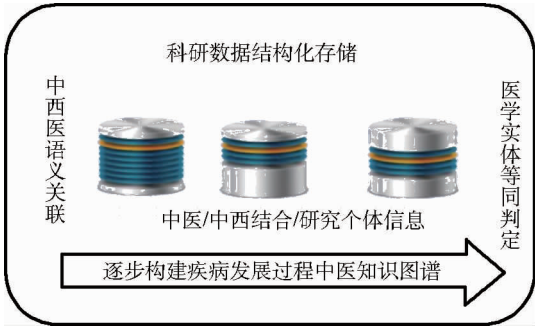


图 1 数据存储优化

2.3 建立数据质量控制体系 (图2)

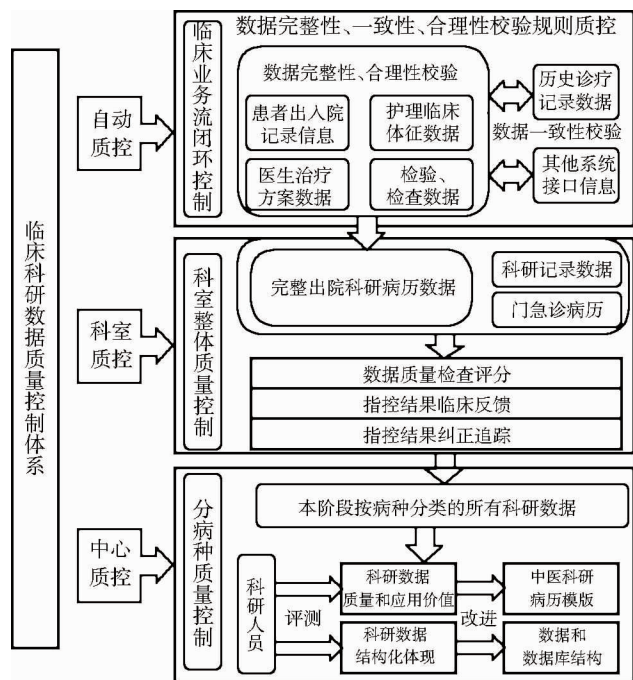


图2 临床科研数据质量控制体系

2.3.3.1 自动质控 临床科研共享系统须具有科研数据质量自动控制功能，能进行科研数据的筛查和审核，保障科研数据采集质量。通过建立规则，不断优化调整质控规则，以保证数据的完整性、一致性和合理性。中医科研数据质量自动控制手段有：（1）质量监控类型包括时限监控、书写次数监控、必填项监控、文书删除监控。（2）内容监控类型包括用语规范，复制、雷同的限制，前后一致性约束。自动质控基于临床业务流的闭环管理体系质量评价，其范围包括：患者出入院记录信息，包含中医特色护理措施和操作在内的临床护理信息和体征数据，手术信息，中医特色治疗手段，中医临床诊断库应用，中医病案首页，中医特色体格检查，中医自制药应用，中草药应用，中医辨证诊疗记录，中医出院调养建议在内的医生治疗方案数据，检验、检查数据。

数据的完整性审核及控制，其主要目的是保证临床科研共享系统能完整获取目前或者将来要应用到的所有科研数据的完整性，以自动质控为主，自主反馈^[4]。对不同病种，如基地的核心研究病种

——重症肝病，需要重点观察和研究的疾病中医要素有“皮肤颜色”、“舌苔”、“肝掌”等中医望、闻、问、切等辨证属性，在建立模板时，必须与之后的科研数据录入提取建立关联的质控机制，为每个病种设置一个科研要素字典，将每个要素的重要性分为3级。第1级：必须存在，在自动质控时根据住诊断类型判断，如果是科研病种，则关联到关键科研要素字段，当检测到医生书写的模板中不存该元素时，根据质控结果提醒医生改进。第2级：必须存在，对填写方式有要求，对取值也存在区域性限制。第3级：必须存在，对填写方式有要求，且存在逻辑性验证检查需求，要求与信息来源进行自动同步。其实现包括：（1）完整性校验规则定义。（2）强制控制，警戒控制。

数据的一致性审核及控制，必须保证历史诊疗产生数据的基础信息的一致性，通过修正手段保证病情发展数据的连贯一致性，同时还必须保证与包含门诊病历、住院管理系统、检验检查信息在内的其他临床医疗系统数据的一致性。其实现包括：（1）一致性校验规则定义。（2）一致性过程管理。数据的合理性校验，其主要目的是纠正科研数据产生过程中因医生或护士疏忽产生的一系列逻辑错误，如患者身份证号码校验、诊断或体征与性别冲突等问题。其实现包括：（1）合理性校验规则定义。（2）合理性过程管理。

2.3.3.2 科室质控 以人工逻辑审核为主,由科室的临床科研质控人员审查科研数据是否存在自动控制无法发现的质量问题^[5],督促临床科研人员补充完整中医科研数据,修正中医科研数据中存在的不足或错误。同时导入并审查门诊病历或被研究者提供的其他科研数据,发现并提醒数据中存在的风险点。科室质控的主要任务和计划实现方式如下:

- (1) 智能化提醒、控制规则定义。
- (2) 病历质量检查评分。
- (3) 质控结果临床反馈。
- (4) 质控结果纠正追踪。

2.2.3.3 中心质控 从临床科研数据的研究应用出发,以基于模板的分病种数据审查为主,偏重于预防和教育警示、引导改进。由科研人员评测科研病种数据的质量和 应用价值,是否需要从科研模板层

面进行改进,哪些科研要素需要修改。同时由数据挖掘人员评测产生的底层科研数据元素的合理性,是否需要在数据结构层面进行改进。评测人员将结果反馈到质控中心,在可行性验证及临床科研人员沟通同意后改进。中心质控的主要目的是指导临床科研人员,提高医务人员整体的医疗和科研水平,提升科研数据的整体质量。

2.4 升级全结构化模板

为提高临床科研数据的结构化水平,在系统升级过程中需要对中医科研模板进行全结构化改造升级。具体包括:(1)改进院级、科室级中医科研病历模板,支持并按科室病种提供模板。(2)提供病历数据元及数据集管理功能。(3)可组织管理模板的各段落、元素等结构,并设计其格式。(4)提供术语集标准模板。(5)能完成病历模板设计、模板管理、模板预览等功能。(6)可将病历模板文件导入导出管理。(7)简化“引用”操作步骤,引用系统接口属性保证数据一致性。

2.5 数据质量优化升级准备

2.5.1 质控工作体系完善 (1)角色设置:按照质控流程要求设置科室质控员、中心质控人员,明确人员职责。(2)模板梳理:对现有中医科研模板进行梳理,规避模板本身质量问题。(3)规范相关业务流程:包含中医科研模板制作审核、中医科研病历书写、3级检诊、数据归档、数据质控流程。(4)质控点设置:科研基地按照质控工作业务流程,制定自身质控监控点。(5)报表部分:结合质控关注点制定报表,辅助对基地中医科研数据质控情况的把控。(6)质控管理方面:在绩效考核中加入质控情况,将质控工作落实到位。(7)建立基于临床科研共享系统的全质控体系:在跨系统之上,建立医疗全过程质控功能。

2.5.2 实现结构化病历 (1)实现结构化病历的关键是保证模板质量,必须取消个人模板,严格限制医生不经医务处审批制作个人模板的行为。

(2)对全院通用模板进行梳理,由专门的模板制作人员对通用模板进行检查,检查点包括模板的编码、格式、内容等并进行修正。(3)梳理科室模板,各科室模板维护人员对各科室模板进行检查,对于缺陷较少的模板进行改造,对于严格不符合模板制作要求的模板给予取消,组织人员对各科室模板进行抽检。(4)建立模板审批流程,临床科研人员提出增加或修改模板的需求,由质量控制中心审批,各科室模板制作人员根据审批的内容进行模板的制作,制作好的模板由质量控制中心审批通过之后方可使用。

3 结语

升级临床科研共享系统的数据质量控制体系,从科研人员应用角度出发,可为临床科研和教学科研提供结构化数据支撑,为国家中医馆云平台的建立奠定坚实基础。优质的临床科研数据同时将支撑北京大学信息工程学院、深圳市云计算关键技术与应用重点实验室等科研教学机构医疗大数据研究工作,促进内科学知识图谱的完善工作,推动医学推理技术和精准医疗发展。

参考文献

- 1 陈南,陈臻,周博翔.系统优化升级理论研究及其在临床科研信息共享系统的应用[J].中国数字医学,2017,12(3):49-51.
- 2 周霞继,谢琪,刘保延,等.中医医疗与临床科研信息共享系统《中医临床术语字典》的构建[J].中国数字医学,2016,11(1):103-105.
- 3 茅伟强.基于结构化电子病历的临床路径系统设计应用[J].信息系统工程,2017,(1):80-81.
- 4 刘保延,周雪忠,张润顺,等.“医疗与临床科研信息共享系统”中医电子病历系统信息基本要求[J].中国数字医学,2012,7(10):57-60.
- 5 卢敬泰,吕晓娟,程钦安,等.临床信息标准化与电子病历深化应用的研究[J].中国数字医学,2017,12(3):12-14.