

基于5W1H分析法的 电子健康档案数据治理研究

刘 洪 梁瑞丹

【摘 要】电子健康档案一直是国内外医疗卫生信息化建设的热点,对电子健康档案数据进行治理是提升其价值的必经之路。论文运用5W1H分析法,对电子健康档案数据建设过程中存在的突出问题进行分析与研究,梳理了电子健康档案数据的治理动因、治理内容、治理主体、治理场景与阶段,最后提出细化管理工作、统一数据标准、优化数据质量、推动数据共享、加强数据保护的具体实践策略,以助力电子健康档案数据优质化建设。

【关键词】电子健康档案;EHR数据;数据治理

【作者简介】刘洪,梁瑞丹,安徽大学管理学院。

【原文出处】《北京档案》,2023.11.26~28

电子健康档案(Electronic Health Records, EHR)数据作为健康医疗大数据的重要组成部分,蕴藏着巨大的医疗价值和科研价值,应用范围广泛。但目前存在的各种问题严重削弱了数据可发挥的价值,EHR数据治理的需求进一步深化。依据数据治理的概念,EHR数据治理可以理解为利益相关者以释放数据资源价值为导向,秉承“以人中心”的服务理念,运用相关治理方法、技术与工具,形成全周期数据分析流程,以实现EHR数据在不同领域和层级的创新性应用。

本文运用5W1H分析法,从治理动因(Why)、治理内容(What)、治理场景与阶段(Where&When)、治理主体(Who)和治理策略(How)几个方面入手,对EHR数据治理框架进行剖析,以期充分发挥EHR数据服务社会、服务民生的功效。

一、Why——EHR数据治理的动因

随着大卫生、大健康理念的深入,我国全民健康信息化建设成效显著。在此背景下,EHR数据被赋予了更多的研究和利用价值。EHR数据利用方式的多元化,对其管理水平和质量也提出了更高的要求。

(一)EHR数据治理缺位制约数据价值的发挥

对EHR数据进行挖掘与整合能够实现资源效益的最大化。但当前EHR数据整合能力不足、整体质量低下、保护力度不足等问题较为突出,并未得到相应的重视。EHR数据利用多以内部需求为主,未形成跨行业和跨领域的开放、共享与协同合力。此外,保障机制不到位、管理和技术人才短缺、有关部门的缺位问题仍然存在,严重阻碍了EHR数据的价值释放,亟须构建全面、合理的EHR数据治理体系。

(二)医疗服务的智慧化驱动EHR数据治理

智慧医疗是医疗信息化发展的产物,需要对医院内部的大量数据进行整合、分析和挖掘,以提升人工智能判断能力,从而实现准确、高效的医疗服务。这就要求医疗机构充分利用现有技术,将分散的EHR数据转化为多维度、集成化、高质量的规范化数据,建立EHR数据共享与治理平台,保证不同部门和业务系统之间的数据流动,满足医疗服务的智能化需求。

(三)多层次的数据需求推动EHR数据治理

在公共卫生事件频发的大背景下,人们不只限

于对当前健康状况的监测,收集和利用自身医疗数据的需求日益增强。由于EHR数据是医疗数据中占比较多且价值较大的核心资产,只有保证其采集和存储时的真实性、完整性,才能最大限度地满足用户利用数据进行疾病干预和健康管理的需求。因此,对EHR数据进行全方位治理是满足患者和用户多样化、多层次健康需求的重要举措。

二、What——EHR数据治理的内容

依据EHR数据的特性,并结合上述问题以及国内外数据治理的重点内容(数据标准、数据质量和数据安全),本文总结出EHR数据标准、数据质量、数据共享和数据安全四个方面的治理内容。

(一)数据标准规范

国家卫健委(原国家卫生部)先后发布了90多项有关电子健康档案的卫生健康信息标准,但关注重点仍停留在健康信息本身的标准化工作上,平台交互标准还不够完善。^[1]2021—2022年,仍有32.89%的三级医院缺乏数据标准或产品对标准的依从性差。^[2]各医疗机构的角色定位、应用需求存在差异,国家统一标准规范的使用比例不高,区域“各行其是”的现象在一定程度上阻碍了电子健康档案统一标准规范的进程,制约了EHR数据的互联互通,形成了数据壁垒。

(二)数据质量管理

数据质量不仅指数据的准确性,还包括相关性、及时性、可比性、可获取性、可解释性等若干组成要素。^[3]EHR数据来源不唯一,需要对接的系统数量繁多,数据内容与格式多样,各地医疗信息化建设水平有所差异。因此,当数据管理水平跟不上数据爆炸式增长的状态时,就会出现关键信息缺失、准确度不足、数据不一致和数据冗余等问题,大大降低医疗工作效率。

(三)数据共享范围

目前我国EHR数据所有权、管理权和使用权的法律边界有待进一步明确,数据需求方利益分配具有不规则性,缺乏完善的协商与反馈机制,导致数据的流通与共享力度不足。据《全民健康信息化调查报告(2021)》显示,2019—2021年中部地区省级EHR数据共享的平均值为22万次,远远低于东部地区的

共享平均值。由此可见,法律机制不完善、协商机制不健全、各地区信息化差异较大等问题限制了EHR数据的高效整合与共享,从而导致服务内容与服务方式受限。

(四)数据安全治理

EHR数据具有较强的隐私性和敏感性,《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》《中华人民共和国基本医疗卫生与健康促进法》将EHR数据列为隐私信息,但对EHR数据权属、权责认定与具体惩罚措施没有做进一步的解释。^[4]《中华人民共和国档案法》也没有对非法利用EHR数据所需承担的法律责任和硬性处罚做具体规定。另外,医疗行业总体风险级别较高,各等级医疗机构的专用网络也存在安全风险,加强数据安全的监管与个人隐私的保护已是刻不容缓。

三、Who——EHR数据治理的主体

EHR数据治理要考虑到相关利益主体所发挥的作用,强调各个层级、不同角色的参与性。首先,国家和地方政府作为EHR数据的顶层规划者,应协同制定治理目标和方针,以系统全面的视角审视电子健康档案建设过程中涉及的参与主体和数据要素问题,积极推进EHR数据治理。其次,作为EHR数据的主要利用者,医疗机构应保证数据收集的真实性和原始性,制定数据代码、数据共享接口等关键性标准与规范,注重数据质量与数据安全,积极探索数据长期存储的方式,提高数据的可用性。第三方健康服务机构应依据法律规定,合理利用医疗机构所存储的EHR数据,增强数据利用与分析能力,满足居民的个性化健康服务需求。医学研究中心应加大资源整合的规模和力度,积极挖掘EHR数据价值,将数据转化为具体成果,为政策制定者提供科学依据。最后,作为EHR数据的生产者,公民应努力提升健康素养和信息自治能力,保证健康数据治理结果的有效性。

四、Where&When——EHR数据治理的场景与阶段

从数据流动的角度来看,数据治理要依靠人的主导作用,并没有固定的治理地点一说,因此可以将Where和When看作治理场景与阶段一同来讨论。

和电子文件生命周期一样,电子健康档案的发展也是一个分阶段的运动过程,因此可以将其划分为“系统与平台设计—采集与整合—集成与归档—开发与利用—维护与监管”五个阶段。

(一)系统与平台设计

统一的电子健康档案系统与区域健康信息平台可以解决医疗信息服务效率低下问题。其建设应由医学行业、档案行业以及信息技术领域相关人才共同探讨,使总体架构和功能模块既能满足医疗机构间特定业务的调档需求,又能实现EHR数据在整个生命周期的有序流动。同时该平台应与公共卫生系统和健康门户网站等平台之间实现功能交互,^[5]满足数据的无障碍共享。

(二)采集与整合

电子文件管理的前端控制原则适用于EHR数据的治理过程,即在数据形成和收集时就对其进行质量控制,规范数据来源与格式。采用数据库日志挖掘解析技术、变更数据捕捉技术对EHR数据进行自动采集,辅以人工检查的方式,优化采集流程,提高数据质量。电子健康档案中所包含的生物学数据、诊断型数据应依据其特点进行针对性采集,并对连续变化的信息进行定期记录。

(三)集成与归档

EHR数据集成是对异构数据进行整合使之成为易于理解的结构化健康信息。^[6]该过程可以采用ETL(extract load)技术,^[7]首先将采集到的EHR数据抽取到临时数据库中,根据数据标准对散乱的数据进行校对,并对冗余值、错误值和不完整数据进行过滤;其次对语义异构的数据进行转换,使其具备标识性和唯一性;最后将经过清洗转换后的结构化数据加载到目标存储库中,发挥其作为电子档案的凭证与利用价值。

(四)开发与利用

EHR数据作为反映个人体征状态的数据是不断更新,其价值通常只增不减。但当前我们对EHR数据的挖掘和分析能力相对不足,EHR数据体现的主要是其凭证价值,情报价值大多也仅限于某些医学研究方向。除了诊疗数据本身,健康体检数据和重点人群健康管理记录中的数据都是可供开发的对

象,可在此基础上分析用户行为,利用数据开发技术将隐性知识转变为显性知识,实现EHR数据价值层面的高效治理。

(五)维护与监管

当前EHR数据多存储于计算机等电子介质上,导致其内容与载体具有一定的可分离性,数据被传播与篡改的概率大大提升。通常医疗卫生部门对患者关键信息进行数据匿名和脱敏处理后仍有一定的漏洞,通过匹配其他外部数据等手段,一些数据仍可被辨别,传统加密技术不足以保持数据共享和隐私保护之间的平衡。更可靠的网络架构和技术、风险评估和安全保护制度已成为当前亟待采取的措施。

五、How——EHR数据治理的策略

针对以上所述的数据标准不一、数据质量低下、数据共享不足、数据安全存在隐患等问题,笔者认为应从以下五个方面入手开展EHR数据治理工作。

(一)细化管理工作,确保治理工作的顺利开展

完善的组织架构对EHR数据良性发展有重要的支撑作用,能够尽可能消除由于权责不清、协调不力而导致的医疗风险。建议在宏观层面成立专门的健康数据领导机构或工作小组,负责统筹电子健康档案的总体规划、标准建设以及协调相关方之间的关系,并通过“自下而上”的反馈及时调整发展方向与模式,开展科学的决策。健康数据领导机构或工作小组内部还应包括数据监管部门,定期评估医疗机构和医疗服务行业的运作情况,负责EHR数据的管理、运维和监督,规范数据利用者的行为,平衡数据共享与安全之间的关系。此外,还应明确医疗机构内部各层级调阅和利用EHR数据的权利,设置分级管理程序,确保数据的有序流动。

(二)统一数据标准,确保EHR数据的互联互通

统一的标准规范能够提升信息化管理水平和EHR数据可利用程度。在标准制定方面,鼓励医疗卫生机构、相关行业协会、国家档案局和其他社会团体等参与EHR数据的标准制定工作,^[8]既要实际效应出发,也应从《中华人民共和国档案法》为法律依据,增强标准的合法性和可执行性。在标准实施方面,各级卫生行政机构应加大组织实施力度,保证

电子健康档案标准的落地应用,同时建立监督和评估机制,作为优化EHR数据标准的依据。

(三)优化数据质量,确保EHR数据的价值显现

数据质量是EHR数据治理的核心问题,可以依据全面质量管理中全面过程管理的概念,从前、中、后三个阶段开展治理工作。前期录入阶段,应以执行标准的操作规范为前提,以统一的专用网络为支撑,以标准化录入工具为途径将EHR数据添加到现有信息系统中。中期监控阶段,运用字典映射、信息检索、转换加载、智能识别等各种治理工具开展数据清洗工作,力求从源头减少脏数据,保证后续数据传输和分析的准确性。后期评测阶段,构建衡量指标明确、可量化、操作性强的数据质量评估体系,通过“校验—反馈—优化”的流程使数据质量问题得到根本改善。

(四)推动数据共享,确保EHR数据的多元服务

医疗卫生信息平台建设力度不足限制了EHR数据的利用价值。因此,要建设以业务需求为导向的EHR数据共享与应用平台,在区域医疗联合体内实现居民健康档案、电子病历数据库、全员人口信息数据库三大支撑数据库的互联互通,将不同来源的数据进行规范化处理后上传至统一平台,保持数据的同步更新。这样,数据利用方经过申请和审查就可以一次性调用所需数据。同时对接各医疗机构,实现诊疗结果之间的互认,为居民提供一站式服务,形成完善的“建档—管档—用档”体系。共享范围的扩大化推动着健康服务的多元化,为决策支持、精准医疗、智慧服务等领域提供数据基础。

(五)加强数据保护,确保公民的隐私安全

保护数据安全可从法律层面、技术层面、管理层面三个方面着手。首先是制定专门的健康数据保护法,包括增加公民对自身健康数据的所有权和自主权,提供更全面的个人信息隐私保护规则,落实具体的惩罚措施。其次是融入技术应用,建立完善的信息安全防护体系。可以基于数据的敏感性,采

用分级管理制度,制定并使用授权许可协议,^[9]对数据风险进行及时评估。最后是优化管理流程,深化权责意识。在EHR数据的采集、存储、共享、销毁等阶段制定详细的管理流程,保证EHR数据流通过程可追溯,建立问责和激励机制,落实被授权人员的权利与责任。

六、结语

EHR数据治理是一个复杂的社会性和技术性问题,必须开展跨领域、跨学科的多方协作,不断提升数据质量、促进数据共享、完善服务体系、保障数据安全,才能发挥电子健康档案的长期效益,促进EHR数据全生命周期内的良性循环,推动档案治理现代化,逐步实现健康中国建设的目标。

参考文献:

- [1]中华人民共和国国家卫生健康委员会.卫生健康标准[EB/OL].(2023-10-07)[2023-10-28]. <http://www.nhc.gov.cn/wjw/wsbzxx/wsbz.shtml>.
- [2]中国医院协会.2021-2022年度中国医院信息化状况调查报告[EB/OL].(2023-02-022)[2023-10-28]. <https://www.chima.org.cn/Html/News/Articles/16012.html>.
- [3]余芳东.国外政府统计数据质量管理及最新发展[J].调研世界,2019(11):60-64.
- [4]钟其炎.电子健康档案信息泄露通知制度的国际经验及借鉴——基于美国、欧盟、澳大利亚、中国的比较分析[J].档案与建设,2019(1):39-42;52.
- [5]韦佳雯,袁曦临.基于用户体验的公众健康信息网站可用性研究[J].情报探索,2022(9):56-63.
- [6]吴信东,董丙冰,堵新政,等.数据治理技术[J].软件学报,2019,30(9):2830-2856.
- [7]刘梦苏.ETL技术在数据分析系统中的应用初探[J].信息通信,2018(8):202-203.
- [8]刘璐瑶,曹航.电子健康档案数据治理发展方向分析[J].北京档案,2021(6):14-19.
- [9]刘晓娟,孙镗莉.生命周期视角下科学数据安全分级管理实践与启示[J].情报理论与实践,2023,46(3):68-74.