

医疗大数据的应用及其隐私保护

黄尤江^{1,2}, 贺莲², 苏焕群², 郑利荣²

[摘要] 大数据为临床研究、生物科技发展及患者个性化治疗提供了有力的数据支持,但在医患关系紧张的医疗环境下,医学伦理及数据隐私是必须考虑的问题。分析了当前医疗行业大数据存在的利与弊,针对其平衡发展及隐私保护提出了相关策略,以平衡技术创新带来的利益与风险。

[关键词] 大数据; 医疗; 隐私

[中图分类号] R-058; R197

[文献标志码] A

[文章编号] 1671-3982(2015)09-0043-03

Application of big data in medical care and their privacy protection

HUANG You-jiang^{1,2}, HE Lian², SU Huan-qun², ZHENG Li-rong²

(1. Wuhan University Information Management School, Wuhan 430072, Hubei Province, China; 2. Guangdong Medical Academic Exchange Center, Guangzhou 510180, Guangdong Province, China)

[Abstract] Big data can provide effective data support for clinical research, biomedical technology progress, and individualized treatment of patients. However, medical ethics and data privacy must be considered due to the poor relation between patients and medical workers. The advantages and disadvantages of big data in medical practice were thus analyzed with certain strategies proposed for their development and privacy protection, and for the balance between the benefits and risks of technological innovations.

[Key words] Big data; Medical care; Privacy

随着数据挖掘技术、云技术等新兴技术的发展,从大量繁冗的信息中挖掘并揭示隐藏在其中的规律和具有重要价值的知识,以及不同数据或数据集间的相互关系已成为可能。个人数据是一种非常重要的资源,然而有些人基于大数据技术非法使用个人数据进行不正当行为,导致个人信息安全漏洞的出现。在医疗界,医院合法利用患者基本信息、病情、用药以及生物科技、医疗器械、基因组序列等方面的大数据,对医师指导患者治疗和健康事业发展有重要贡献。反之则可能违背医学伦理,恶化医患关系。

1 大数据概述及隐私顾虑

大数据(big data),或称巨量资料,是需要新处

理模式才能具有更强的决策力、洞察力和流程优化能力的海量、高增长率和多样化的信息资产^[1]。它与云技术的发展密切相关,是数据分析的前沿技术,与新的数据储存和处理技术不同,大数据是基于传统的数据库技术。大数据革命促使我们在处理复杂事件、捕获在线数据、研发产品、提供服务,以及做出更佳决策方面的能力不断提升,同时也提出一些我们应该做什么以及如何做的问题,有学者将之称为实质伦理问题与程序伦理问题^[2]。因此,有必要鉴定这些伦理性问题,并寻求解决方案。

隐私顾虑是指个人将其相关数据提供给组织后,对组织如何利用和保护其信息的顾虑。Murdoch 研究分析了医疗领域大数据应用的隐私顾虑问题,认为大数据用于医疗保健是一种必然趋势,而病人知情尤为重要^[3]。相关数据表明,我国民众隐私顾虑存在收集方信用、本人知情、事后救济 3 个维度^[4]。美国 Target 百货公司曾因通过上线的客户分

[作者单位] 1. 武汉大学信息管理学院,湖北 武汉 430072; 2. 广东省医学学术交流中心,广东 广州 510180

[作者简介] 黄尤江(1986-),男,重庆市人,本科,助理研究员。

析软件,对顾客购买记录进行挖掘分析,并向顾客推销相关产品,泄露了公民的隐私而被投诉。

2 医疗大数据的应用

大数据分析目标可以总结为以下几类:获得知识与推测趋势,分析掌握个性化特征,通过分析辨识真相^[5]。以临床诊疗为例,大数据的应用是基于患者信息等数据通过挖掘技术及合理推测指导患

者个性化治疗,具体流程如图 1 所示。患者到医院就诊,初诊后会产生体格检查、实验室检查及影像学检查等数据,数据经传送至云端或其他存储器,进而分散存储到移动设备、个人计算机,通过数据分析、数据挖掘、可视化分析等则形成最终报告,医生或专家则通过分析报告指导临床决策,实现患者个性化治疗。

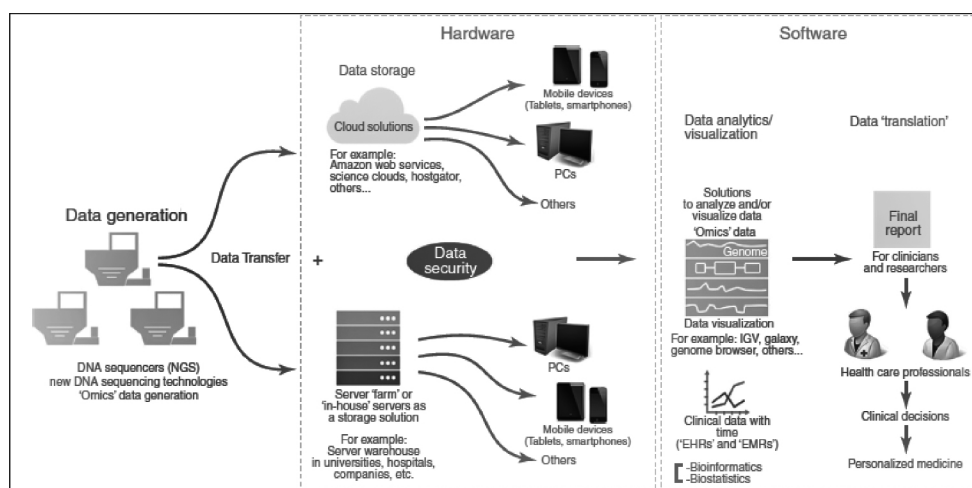


图 1 个体化治疗诊疗流程

利用大数据技术,医生除指导患者治疗外,通过预测模型,也可为疾病的预防提供数据支持。如大数据技术可用于临床随访,已出院但仍需长期治疗者可通过终端移动设备(如基于 Android 系统的 App)与医生进行沟通,医生可通过患者提供的信息为患者预后康复提出指导意见,这是远程医疗的一种升级模式。鉴于我国医院存在“看病难、看病贵”、医患关系紧张的问题,基于大数据技术识别和管理高风险、高成本患者不失为一种有效方法^[6]。为了更好地实现个体化治疗,个人的基因组可添加到电子病历^[7],但是个人基因组是私有的,而基因序列可能推测出很多隐私问题,那么隐私保护成为重点。国外已有公司建立了基因组数据库,如布什就签署了遗传信息无歧视法案(GINA)。

从医院发展战略的角度看,整合医院人力资源、医疗、科研、教学、学科建设、院科两级综合目标管理、后勤与设备等信息资源,构建基于云技术平台的一体化现代医院综合管理系统,可提高医院综合管理水平和竞争力,促进医疗数据统计的自动化,消除

信息孤岛,降低运营成本,提高医院绩效,提供决策支持^[8]。总之,大数据技术的应用能提高医院的综合实力,取得更好的经济效益和社会效益。

3 隐私保护

大数据时代,隐私的泄露极易发生。如果不解决隐私或伦理问题,就难以实现信息共享、消除信息孤岛。无论是 Target 公司案例,还是患者诊疗图中“Data Security”,都强调了隐私保护的重要性。现从以下方面探讨大数据医疗中的隐私保护策略。

3.1 遵循大数据利用的基本原则

大数据技术,乃至更大范围的信息通讯技术的创新、研发和应用,其初衷都是在合法、合乎伦理和非歧视性的基础上,提高人类生活质量。大数据方面的任何行动应根据不伤害人和有益于人的伦理原则给予评价,以此作为权衡预期受益和可能风险的基础。同时也应合理平衡个体与公共的利益。在为了公共利益而限制个人权利和利益时,这种限制应该是必要的、相称的和最低限度的^[9]。在应用大数据的同时应尊重他人隐私,例如在进行病理流行病

学分析时,可通过患者资料推断其生活规律、职业特点,那么作为医疗工作者就应该尊重并保护其隐私,更不能利用这些信息谋利。

3.2 利用更先进的网络安全技术

网络安全技术是保障信息安全的重要手段。采用更好的安全系统、先进的加密系统、优越的算法及存储方案都能有效防止信息泄露,保护患者隐私。匿名保护技术的应用可有效抵抗某些推测类型的攻击,数据发布匿名保护是实现隐私保护的核心关键技术^[5]。刘小波将隐私信息访问策略和信息匿名化作为隐私保护的研究重点,提出了一种隐私保护框架^[10]。Cheng 等人^[11]提出了基于多级别安全模型的风险自适应访问控制解决方案;杨维嘉^[12]提出了转换隐私信息的数据变换方法和整合多个数据源的算法,消除了恶意合谋对原始数据的威胁以及一个能够定制隐私保护程度的方法,同时还提出了一种适应大规模参与者的隐私保护方法;Ahlbrandt 等^[13]则基于 IT 构建采用分散存储方案实现患者数据隐私与大数据需求间的平衡。

3.3 健全法律法规

大数据的发展是全球趋势,也是一个长期过程,针对大数据立法以保证个人信息安全是大势所趋。我国目前已初步建立了与个人信息和隐私权保护有关的法律体系,包括民事、刑事和行政法律体系,但这类法律体系缺乏全面性、系统性和专门性,需要设立有针对性的个人信息保护法来平衡信息自由流动和个人信息保护。苹果和 IBM 合作布局大健康数据时代,健康类移动应用搜集的大部分数据都不在美国隐私保护法规健康保险可携带和责任法案(HIPAA)范围内。美国联邦贸易委员会(FTC)强调,此类数据仍具有高度敏感性,希望调查数据的共享、交换和保护情况。HIPAA 限制了数据分享的方式,Terry 等^[14]认为 HIPAA 法案对保护健康隐私困难但不是不可能。有学者^[15]认为大数据立法不可急于求成,有关个人信息保护法的制定,需要界定个人信息保护主体的义务,确立诸如目的明确、利益平衡等个人信息保护的基本原则,规定信息主体的权利。大数据时代数据的边界模糊,信息量庞大,存在的问题更为复杂,因此保护网络和信息安全的难度也大幅度提升。

4 结语

大数据在医疗中应用甚为广泛,如临床诊疗、医院管理、试剂研发、流行传染病控制、医疗器械、祖国医学发扬等。在享受大数据带来的好处的同时,应考虑到它涉及的医学伦理、个人隐私、数据安全等问题,通过技术手段、道德规范、法律法规健全等合理应用大数据,为医疗行业的蓬勃发展增添技术力量。

【参考文献】

- [1] 杨旭. 数据科学导论[M]. 北京: 北京理工大学出版社, 2014.
- [2] Kord Davis. Ethics of Big Data – Balancing Risk and Innovation [M]. USA: O'Reilly Media, 2012.
- [3] Murdoch TB, Detsky AS. The Inevitable Application of Big Data to Health Care [J]. JAMA, 2013, 309(13): 1351–1352.
- [4] 王忠, 赵惠. 大数据时代个人数据的隐私顾虑研究: 基于调研数据的分析[J]. 情报理论与实践, 2014, 37(11): 26–29.
- [5] 冯登国, 张敏, 李昊. 大数据安全与隐私保护[J]. 计算机学报, 2014, 37(1): 246–258.
- [6] Bates DW, Saria S, Ohno-Machado L, et al. Big data in Health Care: Using Analytics to Identify and Manage High-risk and High-cost Patients [J]. Health Aff (Millwood), 2014, 33(7): 1123–31.
- [7] Erdmann J. As Personal Genomes Join Big Data Will Privacy and Access Shrink? [J]. Chem Biol, 2013, 20(1): 1–2.
- [8] 全筱筱, 郭岱琦, 熊文举, 等. 基于云平台的现代化医院综合管理系统的构建[J]. 中国医院, 2013(8): 43–45.
- [9] 邱仁宗, 黄雯, 翟晓梅. 大数据技术的伦理问题[J]. 科学与社会, 2014, 4(1): 36–48.
- [10] 刘小波. 一种敏感信息共享隐私保护框架[J]. 图书情报工作, 2011, 55(13): 116–120.
- [11] Cheng PC, Rohatgi P, Keser C, et al. Fuzzy Multi-Level Security: An Experiment on Quantified Risk-Adaptive Access Control [J]. IBM Corporation, 2007: 222–230.
- [12] 杨维嘉. 在数据挖掘中保护隐私信息的研究[D]. 上海: 上海交通大学, 2008.
- [13] Ahlbrandt J, Brammen, Majeed RW, et al. Balancing the Need for Big Data and Patient Data Privacy – an IT Infrastructure for a Decentralized Emergency Care Research Database [J]. Stud Health Technol Inform, 2014(205): 750–754.
- [14] Terry N. Health Privacy is Difficult but not Impossible in a Post-HIPAA Data-driven World [J]. Chest, 2014, 146(3): 835–840.
- [15] 张紫. 大数据时代个人隐私危机亟待法律破解[J]. 计算机与网络, 2014(3): 8–9.

[收稿日期: 2015-06-23]

[本文编辑: 刘娜]