

屈光手术标准化数据平台构建及应用^{*}

邹昊翰

季书帆 朱天晨

王 雁

(天津市眼科医院/天津医科大学
眼科临床学院 天津 300020)

(北京航空航天大学计算机科学
和工程学院 北京 100191)

(天津市眼科医院/天津医科大学
眼科临床学院 天津 300020)

〔摘要〕 从数据库类型、建立过程、数据标准化等方面介绍屈光手术标准化数据平台构建, 阐述其临床应用, 分析数据共享与安全保护情况, 为高效、准确、合理利用屈光手术大数据奠定基础。

〔关键词〕 人工智能; 大数据平台; 标准化; 屈光手术

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2020.06.014

Building and Application of Standardization Data Platform of Refractive Surgery ZOU Haohan, Tianjin Eye Hospital/Clinical College of Ophthalmology, Tianjin Medical University, Tianjin 300020, China; JI Shufan, ZHU Tianchen, School of Computer Science and Engineering, Beihang University, Beijing 100191, China; WANG Yan, Tianjin Eye Hospital/Clinical College of Ophthalmology, Tianjin Medical University, Tianjin 300020, China

〔Abstract〕 The building of standardization data platform of refractive surgery is presented in terms of database type, building process and data standardization, its clinical application is described in the paper, and the conditions of data sharing and security protection are analyzed, which lays a foundation for efficient, accurate and rational utilization of refractive surgery big data.

〔Keywords〕 Artificial Intelligence (AI); data platform; standardization; refractive surgery

1 引言

近视问题是我国乃至全球公共健康问题^[1-2]。角膜屈光手术作为矫正近视的重要方式之一被广泛接受。据不完全统计, 我国每年完成近视角膜屈光手术约 100 万例以上, 其产生的丰富临床数据如能有效利用将为该领域人工智能发展提供重要支撑和

保证。但临床数据来源多样、格式不一, 尚未建立统一管理标准, 极大限制了应用。规范、高效、安全、合理和有效利用屈光手术数据具有重要意义。为此本文对屈光手术数据平台构建及数据应用相关问题进行阐述。

2 屈光手术数据平台构建

2.1 数据库类型

主要包括关系型数据库、图数据库、键值数据库、分布式数据库等。关系型数据库采用关系模型组织数据, 用于存储二维表格数据及其之间的联系, 贴近现实场景, 便于理解且易于维护, 较适用于临床病历、医学检查报告等数据; 图数据库主要用于存储实体以及实体之间关系, 常应用于知识图

〔修回日期〕 2019-11-17

〔作者简介〕 邹昊翰, 硕士, 发表论文 2 篇; 通讯作者: 王雁, 教授, 主任医师。

〔基金项目〕 国家自然科学基金“基于 AI 技术探究多源临床表征对角膜屈光手术的精准量化及视觉优化机制的研究”(项目编号: 81873684)。

谱、检索等领域，如患者用药指导等；键值数据库用于存储 Key - Value 型数据，其本质是一张哈希表，具有较高检索和处理性能；分布式数据库突破单机数据存储性能瓶颈，实现数据跨空间领域一致性存储。医疗大数据组织模型大多接近于关系模型，即以二维表模式进行采集和存储，少量医疗数据依托于实体和关系抽取与呈现，因此本研究首先构建屈光手术关系型数据库。

2.2 平台建立过程

2.2.1 数据特性 医疗大数据特点之一是具有较强时序性。屈光手术临床数据除包括结构化的病人基本信息、屈光度外，还包括非结构化的用药信息、检查信息、临床记录随访信息等，且个人病历具有时序性，人工或传统方法难以高效地对其检索、处理、分析并利用，现代数据库应用将解决这一难题。

2.2.2 数据库基本结构 本研究使用 MySQL 数据库，搭建包括人眼基本信息、手术信息、术后信息等数据在内的关系型数据库，见图 1。此外将根据第 3 范式设计表结构，避免数据冗余、更新异常、插入异常和删除异常并优化查询效率；采用用户分级权限、共享权限、数据加密技术等安全机制，保障数据合法使用权限及安全性，实现对多源人眼视觉特征大数据的存储与管理，见图 1。

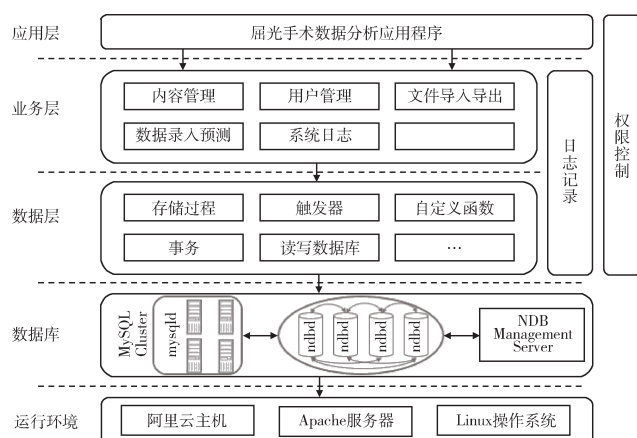


图 1 屈光手术数据平台结构

2.2.3 分布式数据库应用 关系型数据库在数据量大于一定程度情况下单机存储和检索性能会急剧

下降^[3]。当大量眼科数据产生并存于一张表时将可能超过单机数据库表可承受的数据量阈值。采用分布式数据存储系统，将不同种属的分散数据分别储存，既保证各属性下参数独立处理分析，又能满足不同属性间数据的交换和集中处理。人眼相关参数包含形态学和生物学等特性，各属性间既内在联系又相互影响，如角膜地形图所获得的角膜厚度、曲率等数据与角膜硬度、反应速度等力学特性成正相关^[4]。以眼睛整体分析和应用角度出发，从基础参数到各属参数构成分布式数据库，保证数据的完整性和一致性，以此提高整个数据库可用性和工作效率。

2.3 数据标准化

2.3.1 概述 标准化是医学数据库高效管理与应用的重要基础，屈光矫正数据标准化是数据整合与关联的基础。构建屈光手术标准体系是推动该学科数据共享发展的保障。通过建立映射关系寻求多元数据间联系，从而联结孤立数据形成交互关联网，有助于进一步挖掘和发现人眼视觉或生物学规律。

2.3.2 数据库标准 美国《屈光手术杂志》(Journal of Refractive Surgery, JRS) 及《白内障和屈光手术杂志》(Journal of Cataract and Refractive Surgery, JCRS) 倡导建立在报道屈光手术结果时的标准，其中包括作者必须在文章中呈现的基本要素等，方便同行评审以及眼科学领域交流^[5]。不同手术方式、众多参数造成读者在评估和对比不同术式及患者群体时出现困难，标准化数据库应用为提高数据管理和分析效率，实现屈光手术智能化、个性化发展奠定基础。屈光数据库应包括 4 项标准：(1) 基础标准。主要包括基本概念和术语，例如手术前后的视力不仅包括日常裸眼视力，统一为未矫正视力 (Uncorrected Visual Acuity, UCVA)，还应包括最佳矫正视力 (Best Corrected Visual Acuity, BCVA) 等。(2) 数据标准。即数据表述规范、记录无误、明确要保留小数位数、完整性等，如明确视力表示方法统一为对数形式还是分数形式。(3) 技术标准。对应数据库建立的软硬件技术规范，包括软件版本、数据采集、字段定义、约束定义、数据传输、储存、交换、整合等。(4) 管理标准。包

括数据隐私、数据加密、数据安全、用户权限使用等问题。

2.3.3 数据库设计 (1) 邀请码表。记录新用户注册时必须填写的邀请码信息。(2) 用户表。记录用户注册信息及该用户邀请人信息。(3) 授权表。主要存储用户间授权信息。(4) 基础信息表。记录患者基本信息。(5) 术前信息表。记录患者术前检查的体征参数。(6) 手术参数表。主要记录术中由医生和手术平台产生的参数。(7) 术后信息表。记录患者术后各随访时间点的眼部检查指标。(8) 其他信息表。记录病例额外信息, 主要存储该条数据录入时间等。上述设计可对用户和数据本身加以限制, 为数据标准化提供保障。

3 屈光手术数据平台应用

3.1 统一概念

在该数据库建立前期尽可能规范基本概念和专业术语。例如目前屈光手术最新技术为 SMILE (Small Incision Lenticule Extraction), 已在我国普遍开展^[6]。但早期该手术名称混乱不清, 本研究团队将其统一命名为“飞秒激光小切口透镜取出术”并在全国权威学术会议上汇报, 后又制定规范和共识^[7]; 此外还将波前像差等眼科学常用基本概念逐步进行统一。

3.2 信息采集

借助屈光手术标准化数据平台, 根据学科发展特点设计数据库构架, 获取并不断更新人口学基本信息内容, 如姓名、性别、年龄等; 人眼生物学信息, 如屈光度、角膜形态等系列相关参数; 角膜生物力学信息, 如角膜硬度、力反应速度等; 光学基本参数, 如像差分布; 眼球解剖相关参数以及视觉质量评估和视觉矫治参数; 佩戴眼镜种类等结构化或角膜形态图像、病历文本等非结构化数据清洗脱敏收录; 根据各种研究及应用需求扩充完善数据总量和类型。具有高稳定性、高扩展性、高兼容性、易维护性等特点。

3.3 临床应用

为实现临床高效应用, 还需建立有效算法以支撑不同属性参数间的综合分析。本研究借助信息增益算法寻找影响手术效果的可能因素并根据权重筛选纳入, 利用强化决策树算法和多层感知神经网络算法构建手术参数预测模型, 见图 2。满足屈光手术临床病历数据储存管理、相似病例查询、数据分析等基本功能以及手术参数个性化设计、手术结果预测、患者术后管理等高级功能。在术前基于此平台输入患者所需信息, 可自动设计生成手术参数供医生参考使用, 同时满足屈光手术领域临床和科学研究需求。随着人工智能在医学领域的应用和普及, 电子病历建立尤为重要, 该数据库平台将为屈光手术标准化电子病历构建提供思路和基础。该思路已尝试应用于多中心研究和数据库构建。在屈光手术方案设计方面, 经伦理委员会批准, 该标准化数据库形式与内容应用于其他中心, 获得较好效果和收益。



图 2 基于决策树算法挖掘的影响手术设计参数的部分属性

4 数据共享与安全

4.1 数据共享

医学数据只有应用才能产生价值, 共享可打破数据孤岛、扩大应用范围。通过标准化数据库可获取更多数据, 进行集中化分析、个性化应用, 以提高屈光手术精准度。如有更多屈光手术相关机构加入, 将对全国甚至全球范围内数据的更大规模分析提供可能性, 结果将应用于该领域政策制定和临床工作指导, 进一步推动屈光手术行业高质量发展。

4.2 数据安全

流动性是数据共享的基础,将带来安全和伦理问题。医疗数据极具有敏感性,伦理问题是医学健康发展不可回避的问题,在患者隐私、数据权属、知情同意、个体公平等方面面临挑战。对数据库所存储数据应进行严格权限设置,用户申请使用后所上传数据可自定义不同权限:公开、部分公开、保密,且不同用户间可相互授权共享数据。同时数据库使用白名单策略,限定特定用户访问服务器。对数据关键字段脱敏并加密传输,既保护数据信息敏感性又保证其安全。建立防御网络攻击系统,尤其针对分布式拒绝服务(Distributed Denial of Service, DDOS)攻击等。

5 结语

随着医疗信息化和智能化发展日渐成熟,医疗数据产生和利用较为活跃。欧美等发达国家已建立一批高水平健康医疗科学数据平台,如基因序列登记数据库^[8]、世界卫生组织(World Health Organization, WHO)死亡数据库^[9]、美国国家癌症数据库(The National Cancer Data Base, NCDB)^[10]等,但眼科学数据库建设尚处于起步阶段。国内曾报道^[11]有关眼健康大数据平台的构建,主要涉及临床眼病诊疗和应用,而视光和屈光方向尚未涵盖。随着电子病历普及与高分辨率影像学检查结果增多,该领域数据量剧增,有报告显示预计 2020 年将达到 2 314 EB^[12]。我国近视患病人口基数大,总人数超过 6 亿,接受屈光手术患者数量较多,而视觉矫正病症相对单一、数据相对统一,具备充分条件构建标准的屈光手术临床数据库。屈光手术数据库的构建和应用能够改变临床和科研工作耗时费力状态,已初显成果和效益。未来该数据库将不仅局限于视觉矫正,还可拓展到整个眼科学领域乃至大众视觉健康监护和管理。由于已涵盖较多眼部基本参

数和个体指标,可辅助眼部疾病的诊断、临床决策和支持,同时满足科研需求,为解决数据分散、不完整、不标准等问题提供新思路。

参考文献

- 1 李迎新,黄河. 国家卫生与健康大数据平台在医药卫生体制改革中的作用[J]. 国际生物医学工程杂志, 2018, 41 (1): 1-10.
- 2 Holden, B A, Fricke T R, Wilson D A, et al. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050 [J]. Ophthalmology, 2016, 123 (5): 1036-1042.
- 3 朱涛,郭进伟,周欢. 布式数据库中一致性与可用性的关系[J]. 软件学报, 2018, 29 (1): 131-149.
- 4 张耀花,王雁,李浏洋. 角膜地形图筛查疑似圆锥角膜患者角膜生物力学特性的分析研究[J]. 中华眼科杂志, 2019, 55 (6): 442-447.
- 5 Waring G O, Reinstein Dan Z, Dupps William J, et al. Standardized Graphs and Terms for Refractive Surgery Results [J]. J Refract Surg, 2011, 27 (1): 7-9.
- 6 王雁,鲍锡柳. 飞秒激光角膜微小切口基质透镜取出术矫正近视及近视散光的早期临床研究[J]. 中华眼科杂志, 2013, 49 (4): 292-298.
- 7 中华医学会眼科学分会眼视光学组. 我国飞秒激光小切口角膜基质透镜取出手术规范专家共识(2018 年)[J]. 中华眼科杂志, 2018, 54 (10): 729-736.
- 8 Benson D A, I Karsch - Mizrachi, D J Lipman, et al. GenBank [J]. Nucleic Acids Res, 2013, 41 (Database issue): D36-42.
- 9 Say L, Chou Doris, Gemmill Alison, et al. Global Causes of Maternal Death: a WHO systematic analysis [J]. Lancet Glob Health, 2014, 2 (6): e323-333.
- 10 Jessup J M, Menck H R, Winchester D P, et al. The National Cancer Data Base Report on Patterns of Hospital Reporting [J]. Cancer, 1996, 78 (8): 1829-1837.
- 11 周云帆,陈娄,蒋沁. 眼健康大数据平台实践研究[J]. 医学信息学杂志, 2019, 40 (5): 32-35.
- 12 王耿媛. 人工智能在眼科学领域的应用研究进展[J]. 中华实验眼科杂志, 2019, 37 (8): 680-683.