

OHDSI 通用数据模型及医学术语标准国内应用现状分析 *

侯 丽

洪 娜

李露琪

(中国医学科学院/北京协和医学院 (神州数码医疗科技
医学信息研究所 北京 100020 股份有限公司 北京 100086)

(中国医学科学院/北京协和医学院
医学信息研究所 北京 100020)

唐灵逸

翁春华

蒋国谦

李 姣 钱 庆

(德州大学休斯敦生物医学信 (哥伦比亚大学 (梅奥医学院 (中国医学科学院/北京协和医学院
息学院 美国休斯顿 77030) 美国纽约 10027) 美国罗切斯特 55905) 医学信息研究所 北京 100020)

〔摘要〕 介绍观察性健康医疗数据科学与信息学计划 (OHDSI) 及通用数据模型 (CDM) 工作原理, 选取中美两国从事临床、术语标准研究的相关人员开展问卷调查, 总结国内 CDM 及医学术语标准应用现状, 分析应用难点并提出发展建议。

〔关键词〕 OHDSI; 通用数据模型; 医学术语; 医学术语标准

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2020.02.001

Analysis of Application Status of the Common Data Model and Medical Terminology Standards of OHDSI in China HOU Li, Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China; HONG Na, Digital China Health Technologies Co., Ltd., Beijing 100086, China; LI Luqi, Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China; TANG Lingyi, SBMI of UTHealth, Houston, TX 77030, USA; WENG Chunhua, Columbia University, New York, NY 10027, USA; JIANG Guoqian, Mayo Clinic, Rochester, MN 55905, USA; LI Jiao, QIAN Qing, Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

〔Abstract〕 The paper introduces the Observational Health Data and Informatics (OHDSI) and operating principle of the Common Data Model (CDM). The questionnaire survey and analysis has been carried out on the relevant personnel engaged in clinical and terminology standards study in China and the United States. It summarizes application status of CDM and medical terminology standards in China, analyzes the application difficulties, and puts forward development suggestions.

〔Keywords〕 OHDSI; Common Data Model (CDM); medical term; medical terminology standard

〔收稿日期〕 2020-01-20

〔作者简介〕 侯丽, 副研究员, 博士, 发表论文 50 余篇; 通讯作者: 钱庆, 研究员, 发表论文 80 余篇。

〔基金项目〕 国家重点研发计划“精准医学本体和语义网络构建”(项目编号: 2016YFC0901901); 国家重点研发计划“重大疾病精准医学数据库群”(项目编号: 2016YFC0901602); 中国工程科技知识中心建设项目(项目编号: CKCEST-2020-1-14); 新闻出版业科技与标准重点实验室“医学融合出版知识技术重点实验室”。

1 引言

伴随全球范围内医疗卫生信息化建设进程不断加快,以患者为中心的医疗健康数据迅猛增长,不仅规模巨大,且呈现出类型多样、结构复杂、存储分散等特征,对医疗健康数据的存取、利用、分析和临床循证研究带来巨大挑战。医疗健康数据的有效利用关系到国家乃至全球在疾病防控、临床辅助诊断、药品监督以及精准医疗等多方面的能力,受到各国政府和医疗研究者的极大关注,多维数据采集、多源异构数据互通互联、标准化处理以及开放赋能是促进其合理有效利用,发挥其潜在价值的关键步骤。

美国联邦政府一直努力保障所有相关利益方都能够获得标准格式的医疗健康数据,如健康数据计划(Health Data Initiative, HDI)通过 HealthData.gov 网站发布各种标准化格式的数据(包括各种联邦数据库、临床服务提供者相关信息等),针对开发者提供编程访问接口,以确保数据能够被机器可读及应用程序可下载或访问;英国政府投入建立的国民医疗服务系统(National Health Service, NHS)记录庞大而完备的标准化医疗数据,在保密准则和法律规范的保障下,分类开放以用于公共卫生服务、医药制造等研究中^[1]。而我国在《中共中央 国务院关于深化医药卫生体制改革的意见》中也强调“建立实用共享的医药卫生信息系统。大力推进医药卫生信息化建设,以公共卫生、医疗、医保、药品、财务监管信息化建设为着力点,整合资源,加强信息标准化和公共服务信息平台建设,逐步实现统一高效、互联互通”^[2]。由此可见,研制和应用统一的数据标准和规范是实现卫生信息化互联互通、信息共享的基础。

2 OHDSI 通用数据模型分析

2.1 OHDSI 项目概况

2.1.1 简介 观察性健康医疗数据科学与信息学计划(Observational Health Data and Informatics,

OHDSI)是由美国食品药品监督管理局(Food and Drug Administration, FDA) IMEDS (Innovation in Medical Evidence Development and Surveillance) 计划和 Reagan - Udall 基金会支持,哥伦比亚大学牵头,融合多学科领域专家共同参与,进行跨学科合作研究的全球性非营利科研组织,作为一个世界性的公益型非营利性研究联盟,主要研究全方位医疗大数据分析的开源解决方案,旨在通过大规模数据分析和挖掘来提升临床医疗数据价值,实现跨学科、跨行业的多方合作。其目的是继承并发展前期观察性医学产出合作项目(Observational Medical Outcomes Partnership, OMOP)的核心研究成果,其中 OHDSI 通用数据模型(Common Data Model, CDM)的开发是核心工作。除深入开展术语表研究与应用探索外,OHDSI 还将 OMOP 项目目标、内容、应用领域和组织成员拓展到更为广泛的领域,涉及临床文本数据结构化方法研究、临床数据挖掘和分析方法研究与医疗数据隐私保护研究等。当前 OHDSI 在全世界范围内共有 90 多个跨学科(如临床医学、生物统计学、计算机科学等),分布在学术界、产业界、医疗服务机构以及管理机构等不同领域的合作组织成员,致力于创建和使用大规模数据网络下的开源数据,通过产生新的解决方案来提升人类健康和生活水平^[3]。

2.1.2 关注焦点 OHDSI 项目组在全球范围内关注的焦点集中在以下几个方面:(1)数据标准化。采用 OHDSI CDM 作为数据框架,开发标准接口的开源工具和流程,对数据结构、内容和分析流程进行标准化处理。(2)医药产品安全监控。通过大规模数据分析建立开源证据库,发现并评估潜在的药物影响。(3)个性化风险预测。依据患者人口统计、病史以及健康行为数据挖掘和建模分析进行患者个体水平的健康风险预测。(4)数据质量控制。开发用于数据质量评估和数据库分析的工具,为底层源数据使用和处理提供帮助。(5)医疗服务质量评估。开发开源工具对 CDM 中观察性数据应用服务质量进行度量评估并提供实证依据^[4]。其中,OHDSI 中国组致力于利用数据科学和信息学的方法促进健康医疗数据研究工作,主要内容包括针对中

国的临床数据，扩展和定制 OHDSI 的数据科学方法；通过在中国构建大型观察性健康医疗数据网络，形成对于全球性重要健康医疗问题的调查研究和证据；促进生物医学信息学领域的国际合作与教育培训工作^[5]。

2.1.3 研究开展情况 OHDSI 项目愿景是通过大规模分析来揭示健康数据价值，自组织成立后积极开展各种观察性数据的实证分析和研究工作，在世界各地的临床数据合作伙伴中得到广泛应用。

由 OHDSI 组织成员组建的正式网络协作研究小组目前有 15 个，研究议题涵盖疾病治疗新途径探索、药物效果评估、疾病风险预测以及数据质量研究等多个方面，项目发布与参与者也遍布全球各地，以美国（哥伦比亚大学、斯坦福大学、加州大学）和亚洲（亚洲大学、汉阳大学、香港大学）的研究机构为主，其中多个研究项目已合作完成且产出有效的研究成果。OHDSI 应用案例分析，见表 1。

表 1 OHDSI 应用案例分析

案例详情	研究国家/机构	年份	发表刊物/会议
美国医疗保险和医疗补助服务中心（CMS）通过虚拟研究数据中心（VRDCJ）提供的数据转换为 OHDSI 的通用数据模型，为药物使用的数据分析提供帮助 ^[6]	美国/ 国立医学图书馆	2015	OHDSI 2015 年会
OHDSI 合作小组建立来自 4 个国家的 11 个数据源组成的国际数据网络，通过大规模的数据分析评估二型糖尿病、高血压和抑郁症 3 种慢性病的治疗途径，证明大规模的国际观测研究是可行的 ^[7]	美国、韩国/ 哥伦比亚大学 斯坦福大学 首尔大学 加州大学	2016	《美国科学院院刊》
使用 OHDSI CDM 数据库中的患者病史记录，采用两种特征学习方法为心力衰竭、抑郁症和消化道出血疾病建立风险模型，提高风险模型性能 ^[8]	美国/ 强生公司	2015	OHDSI 2015 年会
利用 OHDSI 开发的开源工具，将 5 个专业护理机构的电子健康数据转换为高质量的数据集合，为家庭护理服务的研究提供支持 ^[9]	美国/ 匹兹堡大学	2016	《改善患者预后的证据和方法》
通过 OHDSI 协作连接多个国家的数据站点，对各国阿片类药物过量率进行监控比较，有助于了解美国阿片类药物使用过量情况，并能够更好地预测全球性危机 ^[10]	美国/ 西奈山伊坎医学院 哥伦比亚大学 纽约长老会医院	2017	《急诊医学年鉴》
在由 8 所德国大学医院组成的联合体内实现基于 OHDSI 的数据整合，评估其在支持数据统一和共享方面的适用性，提出可改进的需求 ^[11]	德国/ 埃朗根－纽伦堡大学 美因茨大学 曼海姆大学 法兰克福歌德大学医院	2015	《临床信息学应用》
基于 OHDSI 框架，使用 490 万患者数据测试分析目前处于一线 的 5 类降压药物的疗效与安全性 ^[12]	美国、澳大利亚/ 加利福尼亚大学 耶鲁大学 哥伦比亚大学 南澳大利亚大学	2019	《柳叶刀》

2.2 CDM 及其工作原理分析

2.2.1 内涵 CDM 是由 OMOP 项目组开发的一套

具有统一标准的数据模型，目的是将不同观察性数据库中的数据转换为通用格式以及统一的表示（术语、术语表、编码方案），然后使用标准化数据库

进行系统分析。CDM 包含标准化术语表、标准化元数据、标准化临床数据表、标准化健康系统数据表、标准化健康经济表和标准化派生元素等 6 大类共 39 张表^[13]。CDM 标准化术语表中的标准化概念包括按照一定的束规则从部分国际通用或专业术语标准中提取的概念，如 SNOMED - CT、RxNorm、LOINC，以及 ICD、MeSH、ETC 等近 50 个国际通用术语表，建立与标准概念的关联。CDM 为每个国际术语标准提供一套详细的规则和映射标准，帮助研究者在需要时转化为满足要求的标准概念以便开展后续大规模的数据分析工作。

2.2.2 工作原理 尽管目前诸多医疗机构建立一系列医疗信息系统，如医院信息系统（Hospital Information System, HIS），电子病历（Electronic Medical Records, EMR），检验信息系统（Laboratory Information System, LIS）等用于存储患者的各种医疗数据，但各个系统构建目的与用途不同导致其逻辑组织方式和数据格式不同，此外用于描述药物信息

和临床表现的术语也因来源不同而表达形式各异。单一的观察性数据来源无法提供患者在接受医疗保健服务时产生的临床数据全面视图，无法满足跨系统、跨医院、跨区域的大规模合作型研究需求，因此需要使用 CDM 所定义的通用数据标准同时评估和分析多个数据源^[14]。CDM 支持不同来源的观察性数据通过数据抽取、转换和加载（Extraction - Transformation - Loading, ETL）过程形成标准化的数据结构：CDM 定义了通用数据结构与规范，以适应各种医疗机构提交主动监测系统必需、期望的关键数据；相关机构可将其本地原始格式的数据转换成 OHDSI CDM 定义的数据结构；将包含所有相关编码字典的概念（如 ICD - 9，CPT - 4 等）映射转换成 OHDSI CDM 标准化的概念。通过 ETL 设计，从数据源提取所需数据，经过数据清洗将源数据集内容转换成符合 CDM 表结构的格式，加载到 CDM 中提供研究分析，其基本原理和设计框架，见图 1。

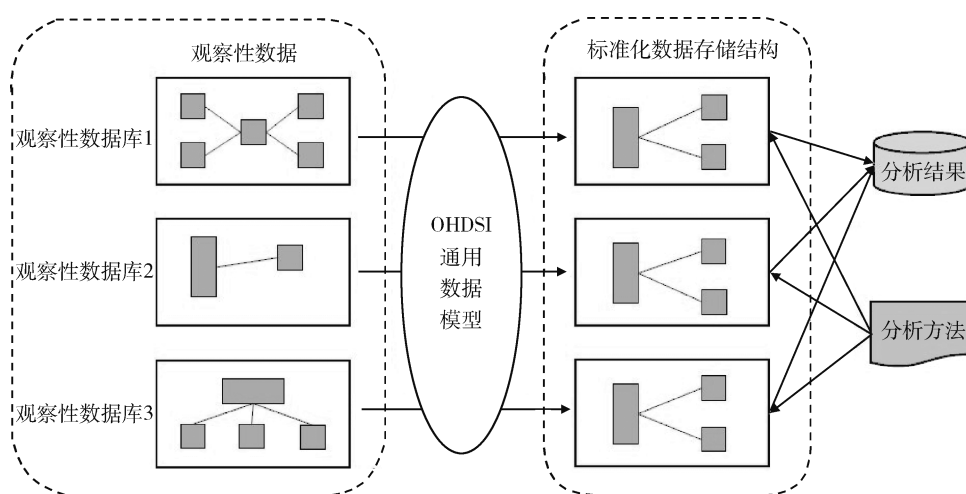


图 1 OHDSI 通用数据模型基本原理和设计框架^[15]

来源网址：<https://www.ohdsi.org/data-standardization/the-common-data-model/>

2.2.3 作用 作为核心工程的 CDM 创建一个通用的结构和框架来组织和标准化观察性数据，对于 OHDSI 组织的发展起到重要作用：一是使从事不同业务、开展不同研究、拥有不同数据特征的机构向 OHDSI 提交的数据更加一致和标准，有利于数据存储和后期的大规模使用；二是基于标准接口允许研究人员开发各种分析方法，在不改变原始数据内容

的前提下可运行在任何标准化后的数据源上；三是有效支持信息共享，包括研究方法、程序、基本分析结果，可以跨数据源实现一致性的应用和解释^[16]。CDM 为 OHDSI 持续性地进行开源分析工具（ATLAS、ACHILLES）开发、分析方法（Cohort Method、Population Level Estimation、Patient Level Prediction）应用，为研究人员进行大规模观察性数

据实证研究奠定基础。

3 医学术语标准在中国的应用现状调查

3.1 概述

OHDSI 中国组致力于利用数据科学和信息学的方法促进健康医疗数据研究工作，针对中国临床数据应用现状将扩展和定制 OHDSI 中的数据科学方法（通用数据模型）作为重点工作之一，其中医学术语标准在临床及科研领域的应用是开展健康医疗数据研究与应用的关键。基于此，选取中美从事临床、术语标准科学研究的从业人员开展问卷调查，以期了解目前我国医学术语标准应用及研究现状，同时通过调研探索隐藏的现实问题，为推进国内相关机构应用通用数据模型奠定基础，也为 OHDSI 工作组进一步开展医学术语标准在中国的落地应用提供思路。

3.2 调查问卷整体情况

本调查发放并最终回收来自中美临床和医学信息研究、应用、教育等人员的共 106 份有效问卷。其中，中国研究人员占大多数（82%），集中在高校、企业和医院；美国研究人员基本来自高校和医疗机构。调查问卷基本信息统计，见表 2。人员分布，见图 2。可以看出随着我国互联网及信息技术的高速发展，不仅有诸多专业的医疗和科研人员参与医学信息学学科建设，同时也有越来越多的高新企业关注医疗大数据这个热门领域的发展，积极与医疗科研机构和医院进行密切合作，从事医学信息相关研究工作。

表 2 调查问卷基本信息统计

类别	选项	百分比（%）
国家	中国	82.1
	美国	17.9
职业分布	高等院校	36.8
	公司	33.0
	医院	26.4
	其他	3.8
是否有使用医学术语标准的经验	是	61.3
	否	38.7

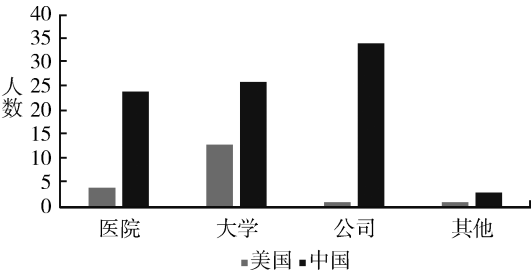


图 2 调查对象国家及职业分布

3.3 我国医学术语标准应用现状调查结果分析

3.3.1 医学术语标准应用场景覆盖医学信息处理的诸多环节 在对医学术语标准应用场景的调查发现其主要用途是数据标准化处理，其次是借助术语标准开展数据分析，再次是用于文献检索与数据采集。由于不同医疗机构的术语使用不一致，故利用现有标准开展数据标准化处理是最广泛的用途。无论是科研人员还是临床医生在数据的收集过程中都涉及不同来源、类型、格式的数据统一存取，需要利用标准编码系统，如 ICD - 10 等将源数据进行合理分类，确定数据收集策略和存储规则。而为了能够大规模使用医疗数据开展数据分析等工作，数据的标准化转换是必不可少的，需要通过术语标准将不同表达方式的数据进行映射编码，确保数据格式的统一，开展数据治理工作。而医学研究人员在进行医学文献检索时需要利用主题词表在数据库内进行文献内容的规范、高效查找，例如由美国国立医学图书馆制定的医学主题词表（MeSH），因此从医学文献检索到医疗数据收集、标准化处理以及分析的各个阶段，诸多研究者都会在实际的科学研究和临床工作中应用医学术语标准，证实医学术语标准在医学信息处理和科学研究中的重要作用。各环节应用分布，见图 3。

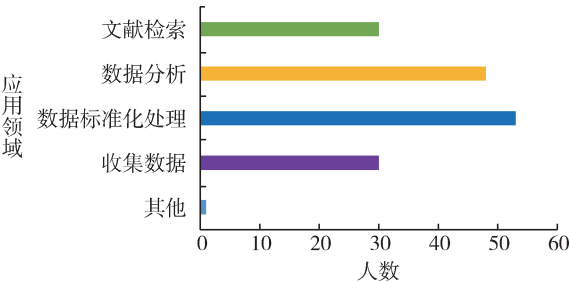


图 3 医学术语标准应用领域分布

3.3.2 中美研究人员对医学术语标准应用存在明显差异 关于使用医学术语标准经验的调查,结果显示中国仅有 60% 左右的研究人员表示使用过医学术语标准,而美国则有 90% 以上的研究人员使用过。可见医学术语标准在国内医疗系统中的推广应用与国外还存在一定差距,推测其原因,国内研究和引进医学术语标准晚于国外多年,在国内大范围的推广和应用存在一定难度,目前众多医疗机构和研究机构内部都有自行定义和使用的术语编码系统,仍有部分医学工作者和研究人员在实际工作中更倾向于使用各自内部自行定义的编码对医学信息进行规范化处理。

3.3.3 国际主流术语标准应用广泛 国际社会对医学术语标准化研究起步较早,各术语标准化组织的共同努力促使国外医学标准及术语集成果丰富,且在全球范围内的应用中取得良好成果。本研究对研究人员具体使用行为的调查结果,见图 4。国内研究人员对一些国际已经被广泛使用的标准存在较高的共识度,包括世界卫生组织 (World Health Organization, WHO) 制定的国际疾病分类与代码 (ICD - 10/ICD - 9),由国际医学术语标准化与研发组织制定的系统化医学术语集——临床术语版 (SNOMED CT),以及 MeSH 等综合性医学术语标准,对观测指标标识符逻辑命名与编码系统 (Logical Observation Identifiers Names and Codes, LOINC)、人类表型本体 (Human Phenotype Ontology, HPO)、临床药品规范化命名表 RxNorm 等专业术语标准有一定的使用;与对美国研究人员的调研结果一致,符合国际主流趋势。而目前国内已有的医学术语标准主要有中文医学主题词表 (CMeSH)、国际疾病分类法中文版 ICD - 10 及其各个修订版等,在成果产出数量和应用效果方面都与国际水平存在较大差距,问卷调查显示仅有少部分中国研究人员 (10%) 偶尔选择使用国内术语标准开展数据分析,证明国内自主开发制定的标准化临床医疗术语标准目前仍处于小范围的应用阶段,影响力远不如国际

主流术语标准,有待进一步完善和推广。

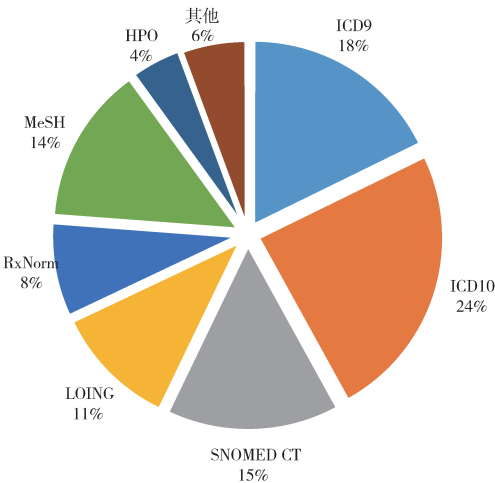


图 4 医学术语标准使用情况

3.4 我国医学术语标准应用现存问题调查结果分析

3.4.1 推广应用存在困难 通过对我国医学术语标准应用现状调查发现大多数研究者都对其应用持积极态度,也肯定了其对于临床工作和科学研究的重要价值,但大部分研究人员普遍认为现阶段医学术语标准在实践应用过程中还存在诸多问题需要解决,包括缺少术语标准相关推广、术语不完备、数据不规范、缺少术语管理工具、采纳何种术语标准更有效,以及实施术语标准存在难度等,见图 5。分析其主要原因可能是面向一线术语应用的培训指导工作欠缺,致使一线工作和研究人员在应用过程中存在术语标准化意识薄弱、操作不规范等问题。同时引进的国际术语标准在进行本土化时存在术语不完备现象,国际术语标准与本地标准的对接还存在一定难度,一些中国特有的疾病并不能及时在汉化版的国际术语标准中出现,例如克山病。此外我国医疗环节中采用的一些本地化术语在国际术语标准中也尚未涵盖。原始数据不规范、缺乏术语管理工具等问题也是致使研究人员无法在实践过程中较好应用术语标准的原因。

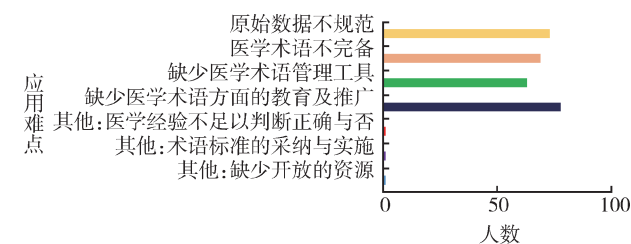


图5 国内术语应用难点分布

3.4.2 共识共通成为最大难题 除了对医学术语标准应用现状进行调查外,本研究还针对我国医学术语标准开发研究工作进行初步调查。调查对象中仅有21%的人参与过医学术语标准开发,且大多数调查对象认为医学术语标准开发困难重重,主要问题集中在术语标准建设的共识共通上。“缺乏协调、各自为政”、“通用性难以估计”、“缺乏正式的基于共识的过程”等都是一线术语开发人员的切实体验。“考虑应用场景和具体需求”、“术语标准的全面性、准确性以及语义关系描述”、“后期的维护支持”等也是术语开发过程中的重点和难点。

3.4.3 国际标准在国内应用发展建议 结合目前国内医学术语标准使用现状,以及对CDM等国际标准模型的应用现状,本研究整理了国内应用CDM及OHDSI相关标准的建议方向,对受访者进行推广医学术语标准建议的问卷调查,见图6。通过分析发现大部分科研人员认为加强CDM及国际术语标准的推广尤为重要,开展国际标准中文版本的建设及整合也较为重要,具体包括国际标准中文版工具平台的开发,而进行不同术语系统间的兼容、整合和映射研究也较为重要,3者协同推进有助于加快CDM及国际术语标准在中国的推广与应用。

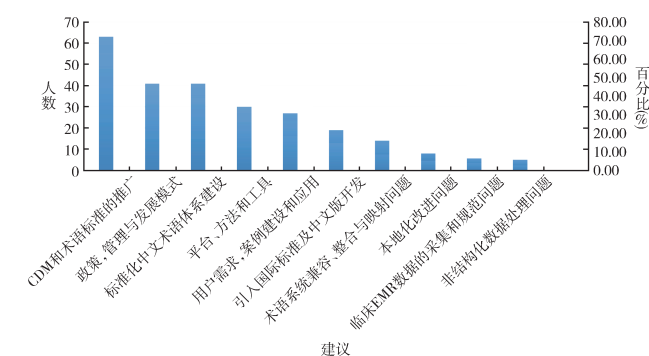


图6 CDM推广与应用建议

4 思考及建议

4.1 应用方面

4.1.1 原始数据结构化程度较低,数据质量难以保证 医疗行业是信息化程度很高但数据化程度很低的领域。目前我国绝大多数医院虽然已实现HIS全覆盖,基于HIS、EMR可以采集到患者身体指标、诊断、治疗、随访等临床数据,但由于患者信息底层逻辑不清晰,医学治疗过程非标准化,大多数治疗数据都是开放式记录,系统存储的患者数据大多为非结构化的文档数据,需要进行大规模、复杂的筛选清洗工作,对医务工作者和研究人员的数 据标准化工作带来极大挑战。

4.1.2 术语标准普及和推广尚需继续加强,完善配套服务 国际医疗卫生组织不仅在医学术语标准的制定方面不断地创新努力,在推动其应用实施方面也不遗余力,高度重视。国内相关政府机构及部门也积极出台一系列相关政策推动医学信息标准化建设,高度重视标准制定和构建,建议在标准实施方面进一步加强政策支持和有效引导,尽快推进医疗标准术语进入实质使用阶段,加速医院之间的互连和信息共享。同时加强对基层医务和科研人员的培训和教育工作,对于已经推广应用的标准建立完善的后续维护、更新机制,帮助已有的标准落地和持续性发展。

4.1.3 标准研究工作不够深入,本土适应性不强 虽然已引进的大量国际术语标准在国内得到较为广泛的应用,但标准的本土化研究仍然不尽人意,除基本的翻译工作外,缺乏对引进术语标准原理机制的深入研究和适应国内医疗环境的改造,当前已有的借鉴国际并翻译生成的本地化术语表往往存在内容不完整、规模较小、质量参差不齐、缺乏长期维护等问题。

4.1.4 医学术语标准研究缺乏统一规划、协同合作 医学信息标准体系是整个医疗卫生信息标准化工作的指南,只有建立科学的国家医学信息学标准体系,将众多医疗卫生信息标准在此体系下协调一致才能充分发挥系统功能,获得良好的系统效

应^[17]。国内相关机构在借鉴和吸收国际先进医学信息标准化成果的基础上启动国家卫生信息标准基础框架,但尚未形成可以统一指导我国医学信息标准化发展的规范体系,诸多研究机构之间缺乏有效沟通协调机制和统一规范,各自为政,致使少数已经形成和应用的标准尚不成熟、不完善、可操作性不强,不能满足医疗卫生事业发展及其信息体系建设的需要。医学术语标准建设与应用是国际医疗卫生组织关注的重点话题,也是当前推进我国医院、社区等公共卫生体系现代化信息建设、实现医疗强国梦的主要瓶颈和障碍。因此基于 OHDSI 项目国际组织及中国组加强建设通用数据模型及术语标准的契机,抓紧制定医学信息管理和技术标准、重视标准贯彻实施、形成与国际标准衔接、具有中国特色的医学信息标准体系至关重要。

4.2 发展建议

4.2.1 规范医学术语标准建设,推动标准应用实施 现阶段我国医学信息标准化建设要以应用为主,强化实施。应采取多项得力措施,从国家战略层面分析我国医学术语标准化的需求特点,明确标准研究和发展重点任务和工作,建立医药卫生信息共享、公用的公共标准和规范。同时加大对医学术语标准研究、推广资金和精力投入,采取强制手段开展医学信息标准化工作,加大基层医疗机构的宣传推广力度,开展大规模持续性的培训教育工作,使医疗机构、团体认识到标准化工作的重要性,从而有效应用在实际工作中,提高医疗服务水平和质量。

4.2.2 学习借鉴国际标准化建设经验,进行本土适应性改进 在我国医学信息标准化水平不高的情况下,要充分借鉴和引用国际先进标准,加强与国际相关组织之间的沟通合作,掌握国际动态和发展趋势,加快我国医学信息标准化建设,制定医学信息标准化发展规划和标准体系,缩小与发达国家的差距。同时要积极开展本土化应用的深入研究,充分结合我国医疗卫生事业发展特点,加强自主创新工作,对引进数据标准进行补充和维护,建设先进、符合国情需要的医学术语标准和相关技术

工具。

4.2.3 促进行业间协同合作,共同发力产出应用成果 加强各科研部门间的沟通配合,对术语标准制定、研究和实施进行统一规划,避免人力物力资源的重复浪费。紧密联系各级卫生医疗机构深入开展调查研究工作,充分了解标准应用实施的具体需求和背景,进行案例建设和应用,借鉴其临床实践经验对医学标准建设提供直接、有效的参考依据。同时与国内具有技术优势的医疗大数据组织或企业合作,借助人工智能、云计算等先进技术开展大规模医疗数据标准化、数据分析、知识发现等工作,充分发挥行业内各方人才、技术优势,将医学标准和数据应用落到实处,产出有力的科研成果,助力我国医学信息标准化建设。

5 结语

本研究对观察性健康医疗数据研究项目 OHDSI 宗旨、开发的通用数据模型 CDM 工作原理展开详细分析,选取中美两国从事医学信息标准研究、应用的相关人员展开问卷调查与结果分析,总结国内外相关机构应用 CDM 及相关术语标准的现状,问卷数量覆盖度有限,后续会考虑进一步扩大问卷调查范围,使调查结果更加科学合理。基于目前的分析,总结国际通用医学术语标准应用难点并提出发展建议,以期为国际医学术语标准在国内的普及提供借鉴。

参考文献

- 1 朱彦,徐俊,朱玲,等.主要发达国家医疗健康大数据政策分析[J].中华医学图书情报杂志,2015,24(10):13-17.
- 2 中华人民共和国卫生健康委员会.中共中央国务院关于深化医药卫生体制改革的意见[EB/OL].[2018-10-09].<http://www.nhfp.gov.cn/tigs/s3576/201309/cc37d909af764f3da261894504d9de9a.shtml>.
- 3 Hripcsak G, Duke J D, Shah N H, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers[J]. Studies in Health Technology & Informatics, 2015 (216): 574-578.

- 4 Observational Health Data Sciences and Informatics. Areas of Focus [EB/OL]. [2018-10-09]. <https://www.ohdsi.org/who-we-are/areas-of-focus/>.
- 5 OHDSI 中国. 观察性健康医疗数据科学与信息学联盟中国工作组介绍 [EB/OL]. [2018-10-09]. <http://ohd-sichina.org/>.
- 6 Fabrício SP Kury, Vojtech Huser. Converting the Data in the U.S. CMS Virtual Research Data Center to the OHDSI Common Data Model version 5 [C]. Washington DC: 1st OHDSI Annual Symposium, 2015.
- 7 Hripsak G, Ryan P B, Duke J D, et al. Characterizing Treatment Pathways at Scale Using the OHDSI Network [J]. Proc Natl Acad Sci USA, 2016, 113 (27): 7329-7336.
- 8 Reys J M, B Ryan P. Discovering the Hidden Risk Factors: an empirical evaluation of incorporating feature-learning methods into a risk model framework using the OHDSI CDM [C]. Washington DC: 1st OHDSI Annual Symposium, 2015.
- 9 Boyce R D, Handler S M, Karp J F, et al. Preparing Nursing Home Data from Multiple Sites for Clinical Research - a case study using Observational Health Data Sciences and Informatics [J]. Egems, 2016, 4 (1): 1252.
- 10 Slovis B H, Averitt A J, Vawdrey D K, et al. 132 Tracking the Opioid Epidemic Through the OHDSI Collaborative [J]. Annals of Emergency Medicine, 2017, 70 (4): S53-S54.
- 11 Maier C, Lang L, Storf H, et al. Towards Implementation of OMOP in a German University Hospital Consortium [J]. Applied Clinical Informatics, 2018, 9 (1): 54-61.
- 12 Marc A S, Martijn J S, Harlan M K, et al. Comprehensive Comparative Effectiveness and Safety of First-line Antihypertensive Drug Classes: a systematic, multinational, large-scale analysis [J]. The Lancet, 2019, 394 (10211): 1816-1826.
- 13 Erica Voss. Data Model Conventions [EB/OL]. [2018-10-09]. <https://github.com/OHDSI/CommonDataModel/wiki/Data-Model-Conventions>.
- 14 Overhage J M, Ryan P B, Reich C G, et al. Validation of a Common Data Model for Active Safety Surveillance Research [J]. J Am Med Inform Assoc, 2012, 19 (1): 54-60.
- 15 王玲. 美国观察医疗结果合作项目中数据组织及通用数据模型的应用研究 [J]. 中国药物警戒, 2015, 12 (6): 341-346.
- 16 OHDSI. OMOP Common Data Model [EB/OL]. [2018-10-09]. <https://www.ohdsi.org/data-standardization/the-common-data-model/>.
- 17 董建成, 周董, 胡新平. 我国医学信息标准化建设存在的主要问题及建议 [J]. 中国医院管理, 2007, 27 (2): 59-60.

2020 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的医学信息学方面的国家级期刊。主管：国家卫生和计划生育委员会；主办：中国医学科学院；承办：中国医学科学院医学信息研究所。中国科技核心期刊（中国科技论文统计源期刊），RCCSE 中国核心学术期刊（武汉大学中国科学评价研究中心，Research Center for Chinese Science Evaluation），美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息技术，医学信息研究，医学信息组织与利用，医学信息教育，动态等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业的师生及广大医教人员。

2020 年《医学信息学杂志》国内外公开发行，每册定价：15 元（月刊），全年 180 元。邮发代号：2-664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号（100020）医科院信息所《医学信息学杂志》编辑部；电话：010-52328673，52328672，52328686，52328687，52328670。

《医学信息学杂志》编辑部