

临床队列研究的数据标准与共享

岳和欣 湛永乐 边峰 张一方 桂路婷 石英杰 孟耀涵 张娟 江宇

中国医学科学院/北京协和医学院群医学及公共卫生学院,北京 100730

岳和欣和湛永乐对本文有同等贡献

通信作者:江宇, Email:jiangyu@pumc.edu.cn

【摘要】 数据标准在临床队列研究的数据收集、整合及共享过程中发挥着重要作用,并逐渐受到重视。本研究通过对 5 项国际成熟的数据标准模型进行综述,总结其特点和发展现状,并将各模型的数据模块与临床队列通用数据集进行匹配,以探究国际数据标准模型的适用性,为我国临床队列研究数据标准模型的制定与完善提供参考。

【关键词】 临床; 队列研究; 数据; 标准; 共享

基金项目: 中国医学科学院医学与健康科技创新工程(2019-I2M-2-004);食管癌和肝癌专病队列研究

Data standard and data sharing in clinical cohort studies

Yue Hexin, Zhan Yongle, Bian Feng, Zhang Yifang, Gui Luting, Shi Yingjie, Meng Yaohan, Zhang Juan, Jiang Yu

School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

Yue Hexin and Zhan Yongle contributed equally to the article

Corresponding author: Jiang Yu, Email: jiangyu@pumc.edu.cn

【Abstract】 Data standard plays an important role in the process of data collection, integration and sharing in clinical cohort studies, and more attention have been paid to it. This paper summarizes the 5 international proven data standard model, analyze their characteristics and development status, and match their data modules with the general data set of the clinical cohorts to evaluate the international data standard models' applicability and provide reference for the development and improvement of the data standard model for clinical cohort studies in China.

【Key words】 Clinical; Cohort study; Data; Standard; Sharing

Fund programs: Innovation Fund for Medical Sciences of Chinese Academy of Medical Sciences (2019-I2M-2-004); Cohort Study of the Esophageal Cancer and Liver Cancer

临床队列研究在病因和危险因素、预后及其影响因素、防治效果及远期疗效等临床研究领域发挥着重要的作用。即使研究同一临床问题,不同团队开展的队列研究从设计实施、变量定义、数据收集及整理等各个阶段的差异性,导致不同临床队列数据之间存在异质性,阻碍了数据的整合与共享。数据标准模型通过将来自不同卫生信息系统的众多纷杂数据标准化为一种通用格式,有助于数据的规

范化收集。本研究通过梳理现有数据标准模型的现状与特点,明确我国临床队列数据标准发展方向,为今后的临床队列数据的整合与共享提供思路。

一、常用数据标准模型

目前国内外常用的数据标准模型均为国外组织开发,适用于临床队列研究且较成熟的数据标准模型有:开放式电子健康档案(Open Electronic Health Record, OpenEHR)组织开发的开放式电子

DOI: 10.3760/cma.j.cn112338-20200610-00831

收稿日期 2020-06-10 本文编辑 李银鸽

引用本文:岳和欣,湛永乐,边峰,等. 临床队列研究的数据标准与共享[J]. 中华流行病学杂志, 2021, 42(7): 1299-1305. DOI: 10.3760/cma.j.cn112338-20200610-00831.



健康档案规范^[1]、美国卫生信息传输标准(Health Level Seven, HL-7)组织开发的快速医疗互操作资源^[2](Fast Healthcare Interoperability Resources, FHIR)、观察性健康数据科学和信息学(Observational Health Data Sciences and Informatics, OHDSI)协作组开发的通用数据模型^[3](Common Data Model, OMOP CDM)、临床数据交换标准协会(Clinical Data Interchange Standards Consortium, CDISC)开发的临床数据获取协调标准^[4](Clinical Data Acquisition Standards Harmonization, CDASH)、以患者为中心的结果研究所(Patient-Centered Outcomes Research Institute, PCORI)开发的通用数据模型^[5](Common Data Model, PCORnet CDM)。以上 5 种数据标准模型的基本情况见表 1, 其具体的相关模块见表 2。

根据 5 种标准模型的相关数据模块, 结合孙一鑫等^[16]设计的呼吸系统疾病专病队列变量模块, 总结出我国临床队列的通用数据模块, 涵盖入选排除标准、基本信息、生活方式与行为、环境因素、疾病

史、诊断、治疗、随访、卫生经济学、生物样本等 16 个模块。各数据标准模型数据模块与队列数据模块的匹配情况见表 3。目前常用的数据标准模型存在以下特点:

1. 模块丰富, 涵盖多种数据类型: 数据标准模型大多依据病例报告表(CRF)中的问题或研究变量, 将其划分为多个模块, 涵盖纳排标准、人口社会学特征、环境因素、检查与诊断等临床队列常用信息。每个模型的模块丰富且各有特色。如 PCORnet CDM 为功能性分布式研究网络, 可进行多站点临床试验和观察性研究, 主要包括 CONDITION、VITAL、DIAGNOSIS 等 22 种主要模块^[5], OMOP CDM 可规范多源异构的观察性数据的格式和内容, 包括 OBSERVATION、MEASUREMENT、DRUG_EXPOSURE 等 30 种模块^[3], FHIR 模型支持多种文档架构, 提供多种实现的代码库, 包含了 patient、Family Member History、Specimen 等 93 种模块^[17]。另外, 数据标准模型不仅包括结构化数据(如数字、字母、符号等), 还涉及

表 1 国内外常用数据标准模型基本情况

模型名称	开发组织	首版发布时间	最新版本	数据转换过程/技术	辅助工具	特点/优点	不足	应用举例
PCORnet CDM	PCORI	2014	5.1	数据抽取-转换-加载(ETL)	FDA Mini-Sentinel CDM、MSCDM 系统	为功能性分布式研究网络, 可进行多站点临床试验和观察性研究	适用范围相对有限; 数据来源于医疗保健系统, 自费相关记录缺失; 跨医疗系统的数据库(如处方和配药)之间存在匹配率较低的问题	心力衰竭疾病队列数据标准 ^[6] 、临床研究多数据源数据链 ^[7]
CDASH	CDISC	2012	2.1	ETL	XML、REDCap 系统	可实现多源数据整合, 适用于大多数临床试验	部分有所区别的临床变量(如疫苗和药物)共用同一个域, 应用时存在一定的偏差; 干预事件域缺乏灵活性, 不支持方案外的未指定干预	标准化病例报告表、基于系统互操作的公共数据模型 ^[8]
OMOP CDM	OHDSI	2012	6.0	ETL	White-Rabbit、Rabbit-In-a-Hat、Usagi 和 Achilles 数据特征化工具等	OMOP CDM 定义了一种统一的数据标准, 可以规范多源异构的观察性数据的格式和内容	只允许 OMOP 标准概念, 部分研究与公共数据模型词汇表存在对应偏差, 应用时存在不适性	英国临床实践研究数据链 ^[9] 、通用生存分析程序的建立 ^[10]
FHIR	HL-7	2011	4.0.1	StructureMap	XHTML、XML、JSON、异构系统集成技术(RESTful Web-Service)	支持多种文档架构, 提供多种实现的代码库	表达能力受资源定义限制; 不支持构建私有资源; 资源不断维护, 参考模型不太稳定	病案首页信息化模型建立 ^[11] 、患者辅助决策系统 ^[12]
OpenEHR	OpenEHR Organization	1997	1.5.0	ETL	XML、NoSQL、SSIS、EHRServer	用户可参与; 易维护; 支持语义互操作; 领域知识共享	方法语义不严格; 没有以可推理的方式来描述知识, 不足以用于复杂场景; 存在多年但应用率仍不高	电子病历数据信息建模 ^[13] 、多国电子健康档案数据中心标准、冠脉临床影像大数据平台 ^[14] 、LinkEHR 平台 ^[15]

表 2 数据标准模型相关模块

模型名称	模块类别	模型名称
PCORnet CDM	核心表(Core Tables)	人口统计学资料(DEMOGRAPHIC)、注册(ENROLLMENT)、医疗保健事件(ENCOUNTER)、诊断(DIAGNOSIS)、程序(PROCEDURES)、生命体征(VITAL)、药物分发(DISPENSING)、实验室检测结果(LAB_RESULT_CM)、健康状态(CONDITION)、患者报告结果(PRO_CM)、处方(PRESCRIBING)、PCORnet 临床试验(PCORNET_TRIAL)、死亡(DEATH)、死亡原因(DEATH_CAUSE)、用药管理(MED_ADMIN)、提供者(PROVIDER)、临床观察结果(OBS_CLIN)、其他结果(OBS_GEN)、密钥(HASH_TOKEN)、地址纵向历史记录(LDS_ADDRESS_HISTORY)、免疫接种记录(IMMUNIZATION)、数据采集(HARVEST)
	补充表(Supplementary Tables)	个人信息隐私(PRIVATE_DEMOGRAPHIC)、个人地址隐私(PRIVATE_ADDRESS_HISTORY)、地理编码隐私(PRIVATE_ADDRESS_GEOCODE)
CDASH	个体基线水平集(ASSOCIATED PERSONS DOMAINS)	注释(CO)、人口统计学(DM)
	干预集(INTERVENTIONS CLASS DOMAINS)	既往与伴随用药史(CM)、暴露史和前暴露史(EC)、程序(PR)、物质使用(SU)
	事件集(EVENTS CLASS DOMAINS)	不良事件(AE)、临床事件(CE)、实施情况(DS)、方案偏离(DV)、医疗保健事件(HO)、病史(MH)
	结果集(FINDINGS CLASS DOMAINS)	药物分发和回收(DA)、死亡详情(DD)、心电图检测结果(EG)、纳入/排除标准(IE)、实验室检查结果(LB)、微生物样本(MB)、微生物敏感性(MS)、显微镜下发现(MI)、药代动力学采样(PC)、体格检查(PE)、调查问卷、评分和量表(QRS)、生殖系统发现(RP)、疾病反应与临床分类(RS)、受试者特征(SC)、肿瘤/癌前病变识别(TU)、肿瘤/癌前病变结果(TR)、生命体征(VS)、关于事件或干预的其他发现(FA)
OMOP CDM	诊疗信息(CARE)	观察(OBSERVATION)、评估(EVALUATION)、指示(INSTRUCTION)、执行(ACTION)
	管理信息(ADMIN)	患者信息(PERSON)、医院信息(HOSPITAL)、费用分类信息(CHARGE CLASSIFICATION)、医师签名(SIGNATURE)、法律约束(LEGAL CONSTRAINT)
	词汇表 (Standardized Vocabularies)	概念(CONCEPT)、词汇(VOCABULARY)、域(DOMAIN)、概念分类(CONCEPT CLASS)、概念之间的联系(CONCEPT RELATIONSHIP)、相关词汇(RELATIONSHIP)、同义词(CONCEPT SYNONYM)、概念的起源(CONCEPT ANCESTOR)、概念图的资料(SOURCE TO CONCEPT MAP)、药效(DRUG_STRENGTH)
	元数据表(Standardized Metadata)	资源(CDM SOURCE)、元数据(METADATA)
	临床数据表(Standardized Clinical Data Tables)	患者信息(PERSON)、观察周期(OBSERVATION_PERIOD)、入院(VISIT OCCURRENCE)入院详情(VISIT DETAIL)、初始症状(CONDITION OCCURRENCE)、死亡(DEATH)、药物暴露(DRUG EXPOSURE)、进展过程(PROCEDURE OCCURRENCE)、仪器设备暴露(DEVICE EXPOSURE)、检测数据(MEASUREMENT)、记录(NOTE)、NLP 记录(NOTE NLP)、调查(SURVEY CONDUCT)、观察(OBSERVATION)、样本(SPECIMEN)、事实联系(FACT RELATIONSHIP)
	健康系统数据表(Standardized Health System Data Tables)	居住地(LOCATION)、既往居住地(LOCATION_HISTORY)、医疗机构(CARE SITE)、信息提供者(PROVIDER)
	健康经济数据表(Standardized Health Economics Data Tables)	计划付款周期(PAYER PLAN PERIOD)、费用(COST)
FHIR	基础类(foundation)	一致性(conformance)、术语(terminology)、安全性(security)、文档(document)、其他(other)
	基本结构类(base)	个人(individual)、实体(entity)、工作流程(flowsheet)、管理(management)
	临床类(clinical)	概况(Summary)、诊断(Diagnostics)、药物治疗(medications)、护理规定(care provision)需求和反馈(Request & Response)
	经济类(financial)	赞助(sponsor)、发票(Billing)、支付(Payment)、一般情况(General)
OpenEHR	诊疗信息(CARE)	观察(OBSERVATION)、评估(EVALUATION)、指示(INSTRUCTION)、执行(ACTION)
	管理信息(ADMIN)	患者信息(PERSON)、医院信息(HOSPITAL)、费用分类信息(CHARGE CLASSIFICATION)、医师签名(SIGNATURE)、法律约束(LEGAL CONSTRAINT)

非结构化数据(如文档、文本、图片、音视频等)。

2. 同一模型的不同模块间存在交集:虽然各模型对于每个模块均有明确的定义,但在内容上仍存在明显的交集。如 FHIR 中的 Patient 和 Observation 模块是发展的最为成熟和最常使用的重要模块^[18]。Patient 模块中包含了有关接受护理或其他健康相关服务的个人或动物的人口统计信息和其他管理

信息,涵盖有关与健康相关的各种活动的患者和动物的数据,包括策划活动、精神科护理、社会服务、怀孕护理和辅助生活、饮食服务、追踪个人健康和运动数据等内容;Observation 模块是医疗保健的核心要素,用于支持诊断、监测进展、确定基线和模式,甚至捕获人口社会学特征^[17]。二者在人口社会学特征、生活方式等信息方面均存在交集。

表 3 临床队列通用数据模块与各标准的匹配情况

数据模块名称	包含内容	数据标准模型匹配的模块				
		PCORnet CDM	CDASH	OMOP CDM	FHIR	OpenEHR
入选排除标准	研究现场、研究对象招募的纳入排除标准、签署知情同意书与否等内容	-	IE	-	Consent	EVALUATION ACTION
基本信息	编号、纳入时间、通讯电话/地址等	PCORNET_TRIAL LDS_ADDRESS_HISTORY	DM	PERSON OBSERVATION_PERIOD	Patient	EVALUATION ADMIN
人口社会学特征	性别、出生日期、民族、职业、家庭年收入等	DEMOGRAPHIC	DM	PERSON	Patient Observation	EVALUATION
生活方式与行为	吸烟、饮酒、膳食、睡眠、体力活动等	VITAL	SU SC	OBSERVATION	Patient Observation	EVALUATION OBSERVATION
心理状况	抑郁、焦虑等	CONDITION	SC	OBSERVATION	Patient	OBSERVATION
环境因素	职业暴露、空气污染、噪声、绿地、邻里环境等	-	EC	OBSERVATION	Observation	EVALUATION
既往史/现病史	疾病史与基础疾病、用药史等	CONDITION MED_ADMIN IM-MUNIZATION	CM MH	OBSERVATION	Observation	EVALUATION
生育史	初潮年龄、妊娠史等	-	RP	OBSERVATION	Patient	EVALUATION
家族史	一级/二级亲属疾病史、家族遗传病史	-	AP	OBSERVATION	Family Member History	EVALUATION
检查信息	体格检查、实验室检查、影像学检查、其他辅助检查	VITAL LAB_RE-SULT_CM OBS_CLIN OBS_GEN	EG LB PE VS	MEASUREMENT	Observation Diagnostic Report	ACTION OBSERVATION
诊断信息	门诊/入院/出院诊断	DIAGNOSIS	-	PROCEDURE_OCCURRENCE	Diagnostic Report	EVALUATION
治疗情况	药物治疗、手术治疗、康复治疗等	PROCEDURES DISPENSING PRESCRIBING	DA EX PR	DRUG_EXPOSURE DEVICE_EXPOSURE	Medication Administration Detected Issue	EVALUATION INSTRUCTION
疾病进展与转归	不良反应、并发症、五年生存、死亡等	PRO_CM DEATH	AE CE	CONDITION_OCCURRENCE DEATH	Adverse Event	EVALUATION
卫生经济学指标	医疗费用、误工费、医保报销等	ENCOUNTER	-	-	Payment Reconciliation	ADMIN
随访情况	随访时间、失访等	-	DS	VISIT_OCCURRENCE	-	-
生物样本	生物样本采集、储存、检测的数据等	DEMOGRAPHIC	MB	SPECIMEN	Specimen	OBSERVATION

3. 模型设计均基于数据同质化原则:尽管在医疗保健中越来越多地使用标准术语,但临床队列数据在收集过程中由于不同目的、来源和条件限制,仍会导致不同的数据库系统和信息模型采用不同的格式存储数据。这些数据可能未明确捕获到临床研究所需要的数据元素,但各个数据标准模型均基于数据同质化的原则,通过编辑逻辑,创建和共享同类群组,将原始数据整合到一个通用的数据标准模型以充分利用。

4. 国外已成体系,国内起步较晚:目前国外组织开发的数据标准模型大多已发展成熟,并定期更新版本。其中,1987年成立的HL7获得美国国家标准学会(ANSI)认可开发标准,致力于为交换、集

成、共享和检索电子卫生信息提供全面的框架和相关标准,以支持临床实践和健康服务的管理和评估^[19]。OHDSI自2014年成立以来,发展迅速,在其主办的论坛上吸引了来自学术界、医疗相关行业、卫生系统、临床医生、患者等不同利益相关方的2500多名合作者,覆盖计算机科学、流行病学、统计学、生物医学信息学、健康政策和临床科学等多学科^[20]。

纵观我国,自2009年起,原卫生部组织建立城乡居民健康档案、电子病历和区域卫生信息平台等一系列卫生信息标准规范^[21-23],并不断更新完善,使得我国城乡居民的健康信息和诊疗记录走向规范。但关于队列数据标准相关模型和开发组织起

步较晚,目前由北京大学公共卫生学院牵头成立的中国队列共享平台^[24](China Cohort Consortium)正在稳步发展。

二、临床队列研究数据标准模型选择建议

由于临床队列研究方向较多,包含疾病、药物、医疗器械和生物技术产品等,研究目的各有侧重,而不同数据标准模型的应用领域和优势特色也不尽相同。结合表 1~3 中所列的各模型及其模块特点和现有的应用实例来看,不难发现,FHIR 中包含较全面的医疗保健的基本要素,并且涵盖病前期及相关暴露因素的信息收集,所以在研究疾病发生发展和影响因素的临床队列时首推 FHIR 标准;OMOP CDM 最初用于药物评价,其关于药物暴露、进展过程及症状等方面有详细的记录,所以在研究医疗产品安全性和有效性的临床队列则推荐 OMOP CDM 标准;CDASH 适用于治疗领域的大多数临床试验,使通过多个研究收集的数据的定义标准化,且其包含随访模块,可在干预后获得其进展与转归信息,故在临床干预后疾病进展与转归研究中建议选择 CDASH。具体的标准模型选择参考建议见表 4。值得一提的是,一个临床队列研究的问题可以是多方面的,这时候模型的选择建议根据数据标准模型特点进行组合式选择。

三、机遇与挑战

1. 临床队列发展迅猛,亟待数据规范和整合标准:自 2012 年国家临床医学研究中心建设工作启动以来,我国分 3 个批次先后布局建设了 32 家中心,建成 60 余个大型生物样本库、数据库和 143 个临床研究队列,覆盖人群 706.05 万人次,涉及 60 余个病种;并预计到 2021 年底,在主要疾病领域和临床专科统筹建成 100 家左右的中心,开展 20 万~30 万人以上规模的疾病人群队列研究^[31]。然而我国现有的重大疾病队列研究存在诸多问题^[32],包括集中信息系统建设、数据存储及管理、数据标准化、

互操作性接口、术语管理等方面技术水平参差不齐,数据质量问题突出,数据的采集、存储、整合、处理、交换与共享的管理技术缺乏统一标准,进而导致数据的利用率不高,降低临床研究效率。

2. 中文临床医学术语问题仍待解决:2018 年,在国务院发布的《关于促进“互联网+医疗健康”发展的意见》中^[33],明确提出健全统一规范的全国医疗健康数据资源目录与标准体系。结合国外医学术语发展水平,我国中文临床医学术语仍存在术语内容覆盖范围局限,术语标准结构不合理,缺乏体系完整的术语标准等问题^[34],尚未开发出一套可涵盖临床诊断、检查、药物、疾病分类、症状描述等临床实践领域的术语标准。2019 年,中华预防医学会曾就大型人群队列发布一系列关于数据处理技术和数据安全技术等规范^[35-36],在一定程度上有效提高了队列数据的规范性、完整性和准确性等质量要求。我国仍需结合国外经验、国内现有临床队列建设情况,开展合理的医学术语开发规划,进而推动我国临床队列数据标准的发展。

3. 数据整合共享过程伦理法规问题不容忽视:随着医疗大数据的到来,医学伦理问题争议不断。临床队列数据在整合与共享过程中,尤其是对于临床医学数据中最具有价值的患者数据部分,更应加强伦理的关注与考量。我国有学者强调应关注医学大数据中可能出现的伦理问题,如信息安全与个人隐私、知情同意与个人自主性、政府的监测与管理、风险与收益的评估以及医疗大数据建设面临的挑战等^[37-38]。因此,在鼓励临床队列研究数据共享的同时,应加快推进我国相关政策法规的制定,在充分保证知情同意、隐私保护等问题的前提下,保障临床队列研究数据的合理合法利用。

四、总结与展望

当前医疗卫生正处于生物革命和信息革命相融合的全新时代,信息学及其相关技术在临床研究

表 4 临床队列研究数据标准模型建议

研究方向	具体内容	推荐模型
疾病	疾病的发生发展、影响因素研究	FHIR ^[25]
	临床干预后疾病进展与转归研究	CDASH ^[26]
	疾病的预后及死亡研究	PCORnet CDM ^[6]
	高危人群的筛查	FHIR ^[27]
医疗产品	药物、医疗器械或其他医疗产品的有效性和安全性评价	OMOP CDM ^[28]
卫生经济学评价	药物、医疗器械及生物技术产品的成本效益研究	PCORnet CDM ^[7]
临床诊疗辅助决策	智能诊断和导诊、辅助决策的灵敏度和特异度研究	OpenEHR ^[29]
政策评价	健康管理、健康促进方案研究	OpenEHR ^[30]

中得到广泛应用。数据标准模型有助于临床队列数据的有效整合与共享,鉴于现有的临床队列研究领域广泛、问题纷杂,今后在开展专病队列研究时,仍需在通用数据标准的基础上,结合专病队列的特点进行相应的调整和增补。

目前国外对于临床队列研究数据标准的探索与应用较为深入和全面,并取得一定成果,我国数据标准模型的开发尚处于起步阶段,仍需相关科研和技术人员加大投入,充分利用数据标准模型在临床队列研究中的作用,发挥我国临床队列研究数据资源的巨大价值。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] openEHR-Wikipedia[EB/OL]. [2020-07-18]. <https://en.wikipedia.org/wiki/OpenEHR>.
- [2] Fast Healthcare Interoperability Resources-Wikipedia[EB/OL]. [2020-07-19]. https://ja.wikipedia.org/wiki/Fast_Healthcare_Interoperability_Resources.
- [3] OMOP common data model[EB/OL]. [2020-07-18]. <https://ohdsi.github.io/CommonDataModel/>.
- [4] Clinical Data Interchange Standards Consortium (CDISC)[EB/OL]. (2017-01-01) [2020-07-18]. <http://www.cdisc.org>.
- [5] PCORnet CDM[EB/OL]. [2020-07-18]. <https://pcor-net.org/data-driven-common-model/>.
- [6] Tison GH, Chamberlain AM, Pletcher MJ, et al. Identifying heart failure using EMR-based algorithms[J]. *Int J Med Inform*, 2018, 120: 1-7. DOI: 10.1016/j.ijmedinf. 2018.09.016.
- [7] Canterberry M, Kaul AF, Goel S, et al. The patient-centered outcomes research network antibiotics and childhood growth study: implementing patient data linkage[J]. *Popul Health Manag*, 2020, 23(6): 438-444. DOI: 10.1089/pop. 2019.0089.
- [8] Huser V, Sastry C, Breymaier M, et al. Standardizing data exchange for clinical research protocols and case report forms: an assessment of the suitability of the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM) [J]. *J Biomed Informat*, 2015, 57: 88-99. DOI: 10.1016/j.jbi.2015.06.023.
- [9] Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research Datalink (CPRD) [J]. *Int J Epidemiol*, 2015, 44(3): 827-836. DOI: 10.1093/ije/dyv098.
- [10] Hong N, Zhang N, Wu HW, et al. Preliminary exploration of survival analysis using the OHDSI common data model: a case study of intrahepatic cholangiocarcinoma[J]. *BMC Med Inform Decis Mak*, 2018, 18 Suppl 5: 116. DOI: 10.1186/s12911-018-0686-7.
- [11] 王雪梅, 刘敏超, 季磊, 等. 基于 FHIR 的病案首页信息化模型建立 [J]. *中国数字医学*, 2018, 13(6): 43-46. DOI: 10.3969/j.issn.1673-7571.2018.06.015.
Wang XM, Liu MC, Ji L, et al. Establishment of the information model of the first page of medical record based on international medical information standard HL7 FHIR[J]. *China Digit Med*, 2018, 13(6): 43-46. DOI: 10.3969/j.issn.1673-7571.2018.06.015.
- [12] Semenov I, Kopanitsa G, Denisov D, et al. Patients decision aid system based on FHIR profiles[J]. *J Med Syst*, 2018, 42(9): 166. DOI: 10.1007/s10916-018-1016-4.
- [13] 闵令通, 段会龙, 吕旭东. 基于 openEHR 的医疗信息建模方法 [J]. *中华医学图书情报杂志*, 2018, 27(3): 1-4. DOI: 10.3969/j.issn.1671-3982.2018.03.001.
- [14] Min LT, Duan HL, Lv XD. Medical information modeling based on openEHR[J]. *Chin J Med Library Inf Sci*, 2018, 27(3): 1-4. DOI: 10.3969/j.issn.1671-3982.2018.03.001.
- [15] OpenEHR China[EB/OL]. [2020-06-09]. <http://openehr.org.cn/literature.jsp>.
- [16] Link EHR[EB/OL]. [2020-06-09]. <https://linkehr.veratech.es/>.
- [17] 孙一鑫, 裴正存, 詹思延. 呼吸系统疾病专病队列研究的标准制定与数据共享 [J]. *中华流行病学杂志*, 2018, 39(2): 233-239. DOI: 10.3760/cma.j.issn.0254-6450.2018.02.019.
Sun YX, Pei ZC, Zhan SY. Data harmonization and sharing in study cohorts of respiratory diseases[J]. *Chin J Epidemiol*, 2018, 39(2): 233-239. DOI: 10.3760/cma.j.issn.0254-6450.2018.02.019.
- [18] Resource list -FHIR v4.0.1-HL7.org[EB/OL]. [2020-06-09]. <https://www.hl7.org/fhir/resourcelist.html>.
- [19] 崔健, 李俊, 陈先来, 等. FHIR 标准研究现状 [J]. *中国医学物理学杂志*, 2017, 34(9): 924-928. DOI: 10.3969/j.issn.1005-202X.2017.09.013.
- [20] Cui J, Li J, Chen XL, et al. Status quo of fast health interoperable resources standard[J]. *Chin J Med Phys*, 2017, 34(9): 924-928. DOI: 10.3969/j.issn.1005-202X.2017.09.013.
- [21] HL7.org[EB/OL]. [2020-06-09]. <http://www.hl7.org>.
- [22] The Book of OHDSI[EB/OL]. [2020-06-09]. <https://ohdsi.github.io/TheBookOfOhdsi/OhdsiCommunity.html#ohdsi-progress>.
- [23] 中华人民共和国卫生部. WS 365-2011 城乡居民健康档案基本数据集[S]. 北京: 中国标准出版社, 2012.
Ministry of Health, PRC. WS 365-2011 Basic dataset of health record for residents[S]. Beijing: China Standard Press, 2012.
- [24] 中华人民共和国卫生部. WS 370-2012 卫生信息基本数据集编制规范[S]. 北京: 中国标准出版社, 2012.
Ministry of Health, PRC. WS 370-2012 Specification for drafting of health information basic dataset[S]. Beijing: China Standard Press, 2012.
- [25] 中华人民共和国国家卫生和计划生育委员会. WS/T 502-2016 电子健康档案与区域卫生信息平台标准符合性测试规范[S]. 北京: 中国标准出版社, 2017.
State Health and Family Planning Commission of the People's Republic of China. WS/T 502-2016 EHR and regional health information platform standard conformity test specification[S]. Beijing: China Standards Press, 2017.
- [26] 中国队列共享平台[EB/OL]. [2020-07-18]. <http://chinacohort.bjmu.edu.cn>.
- [27] China Cohort Consortium [EB/OL]. [2020-07-18]. <http://chinacohort.bjmu.edu.cn>.
- [28] Bergquist T, Buie RW, Li K, et al. Heart on FHIR: integrating patient generated data into clinical care to reduce 30 day heart failure readmissions (extended abstract) [J]. *AMIA Annu Symp Proc*, 2018, 2017: 2269-2273.
- [29] About CDash[EB/OL]. [2020-07-19]. <https://www.kitware.com/cdash/project/about.html>.
- [30] Rubin L, López NP, Gaiera A, et al. Development, implementation and preliminary results of an electronic reminder for HIV screening using a service oriented architecture[J]. *Stud Health Technol Inform*, 2019, 264: 763-767. DOI: 10.3233/SHT1190326.
- [31] Yang Y, Zhou XF, Gao SQ, et al. Evaluation of electronic healthcare databases for post-marketing drug safety surveillance and pharmacoepidemiology in China[J]. *Drug Saf*, 2018, 41(1): 125-137. DOI: 10.1007/s40264-017-0589-z.
- [32] Spineth M, Rappelsberger A, Adlassnig KP. Achieving interoperability between Arden-Syntax-based clinical decision support and openEHR-based data systems[J]. *Stud Health Technol Inform*, 2018, 248: 338-344.
- [33] Sáez C, Moner D, García-De-León-Chocano R, et al. A standardized and data quality assessed maternal-child care integrated data repository for research and

- monitoring of best practices: a pilot project in Spain[J]. Stud Health Technol Inform, 2017, 235: 539-543. DOI: 10.3233/978-1-61499-753-5-539.
- [31] 中华人民共和国科学技术部. 科技部 国家卫生计生委 军委后勤保障部 食品药品监管总局关于印发《国家临床医学研究中心五年(2017-2021年)发展规划》等3份文件的通知 [EB/OL]. (2017-07-19) [2020-06-09]. http://law.pharmnet.com.cn/laws/detail_4212.html. Ministry of Science and Technology of the People's Republic of China. Notice of three issued documents such as "Five year (2017-2021) development plan of National Clinical Research Center" by Ministry of Science and Technology of the People's Republic of China, National Health and Family Planning Commission of the People's Republic of China, Logistic Support Department under the Central Military Commission and China Food and Drug Administration[EB/OL]. (2017-07-19) [2020-06-09]. http://law.pharmnet.com.cn/laws/detail_4212.html.
- [32] 李伟, 王士泉. 基于专病队列的重大疾病临床样本生命组学数据库建设[J]. 中华医学图书情报杂志, 2017, 26(6): 11-16. DOI:10.3969/j.issn.1671-3982.2017.06.003. Li W, Wang SQ. Development of bio-genomics database for major disease clinical samples based on specific disease cohort[J]. Chin J Med Library Inf Sci, 2017, 26(6): 11-16. DOI:10.3969/j.issn.1671-3982.2017.06.003.
- [33] 中华人民共和国国务院办公厅. 关于促进"互联网+医疗健康"发展的意见[EB/OL]. (2018-04-28) [2020-06-08]. http://www.gov.cn/zhengce/content/2018-04/28/content_5286645.htm. General Office of the State Council, PRC. Opinions on promoting the development of "Internet plus Medical Health"[EB/OL]. (2018-04-28) [2020-06-08]. http://www.gov.cn/zhengce/content/2018-04/28/content_5286645.htm.
- [34] 谢雪娇, 张黎黎, 奈存剑, 等. 国外医学术语标准开发方法及对我国的启示[J]. 中华医学图书情报杂志, 2019, 28(11): 16-21. DOI:10.3969/j.issn.1671-3982.2019.11.003. Xie XJ, Zhang LL, Nai CJ, et al. Development methods of foreign medical terminology standards and its enlightenments to our country[J]. Chin J Med Library Inf Sci, 2019, 28(11): 16-21. DOI: 10.3969/j.issn.1671-3982.2019.11.003.
- [35] 中华预防医学会. 大型人群队列研究数据处理技术规范(T/CPMA 001-2018)[J]. 中华流行病学杂志, 2019, 40(1): 7-11. DOI:10.3760/cma.j.issn.0254-6450.2019.01.003. Chinese Preventive Medicine Association. Technical specification of data processing for large population-based cohort study (T/CPMA 001-2018) [J]. Chin J Epidemiol, 2019, 40(1): 7-11. DOI: 10.3760/cma.j.issn.0254-6450.2019.01.003.
- [36] 中华预防医学会. 大型人群队列研究数据安全技术规范(T/CPMA 002-2018)[J]. 中华流行病学杂志, 2019, 40(1): 12-16. DOI:10.3760/cma.j.issn.0254-6450.2019.01.004. Chinese Preventive Medicine Association. Technical specification of data security for large population-based cohort study (T/CPMA 002-2018) [J]. Chin J Epidemiol, 2019, 40(1): 12-16. DOI: 10.3760/cma.j.issn.0254-6450.2019.01.004.
- [37] 刘辉, 丛亚丽. 临床医学大数据的伦理问题初探[J]. 医学与哲学, 2016, 37(19): 32-36. DOI: 10.12014/j.issn.1002-0772.2016.10a.07. Liu H, Cong YL. Ethical issues of big data in clinical medicine[J]. Med Philos, 2016, 37(19): 32-36. DOI: 10.12014/j.issn.1002-0772.2016.10a.07.
- [38] 刘星, 王晓敏. 医疗大数据建设中的伦理问题[J]. 伦理学研究, 2015, (6): 119-122. DOI: 10.15995/j.cnki.llxyj.2015.06.024. Liu X, Wang XM. Ethical issues in the construction of medical Big data[J]. Stud Eth, 2015, (6): 119-122. DOI: 10.15995/j.cnki.llxyj.2015.06.024.

读者·作者·编者

本刊常用缩略语

本刊对以下较为熟悉的一些常用医学词汇将允许直接用缩写,即在文章中第一次出现时,可以不标注中文和英文全称。

OR	比值比	HBcAg	乙型肝炎核心抗原
RR	相对危险度	HBeAg	乙型肝炎e抗原
CI	可信区间	HBsAg	乙型肝炎表面抗原
P _n	第n百分位数	抗-HBs	乙型肝炎表面抗体
AIDS	艾滋病	抗-HBc	乙型肝炎核心抗体
HIV	艾滋病病毒	抗-HBe	乙型肝炎e抗体
MSM	男男性行为者	ALT	丙氨酸氨基转移酶
STD	性传播疾病	AST	天冬氨酸氨基转移酶
DNA	脱氧核糖核酸	HPV	人乳头瘤病毒
RNA	核糖核酸	DBP	舒张压
PCR	聚合酶链式反应	SBP	收缩压
RT-PCR	反转录聚合酶链式反应	BMI	体质指数
Ct值	每个反应管内荧光信号达到设定的阈值时所经历的循环数	MS	代谢综合征
PAGE	聚丙烯酰胺凝胶电泳	FPG	空腹血糖
PFGE	脉冲场凝胶电泳	HDL-C	高密度脂蛋白胆固醇
ELISA	酶联免疫吸附试验	LDL-C	低密度脂蛋白胆固醇
A值	吸光度值	TC	总胆固醇
GMT	几何平均滴度	TG	甘油三酯
HBV	乙型肝炎病毒	CDC	疾病预防控制中心
HCV	丙型肝炎病毒	WHO	世界卫生组织
HEV	戊型肝炎病毒		