

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

· 专题 ·

## 基于 openEHR 的医疗信息建模方法

闵令通,段会龙,吕旭东

[摘要] 健康医疗大数据的发展和有效利用对医疗信息建模提出了挑战。提出了一种基于 openEHR 的医疗信息建模方法,利用协同方式构建可共享、重用的医疗信息模型,并通过电子病历信息建模验证该方法的可行性。结合 openEHR 规范和协同信息建模方法,能够提高医疗信息模型的扩展能力和互操作能力,促进健康医疗大数据的发展和有效利用。

[关键词] 健康医疗大数据;医疗信息建模;openEHR

[中图分类号] R197.1;TP311.13

[文献标志码] A

[文章编号] 1671-3982(2018)03-0001-04

### Medical information modeling based on openEHR

MIN Ling-tong, DUAN Hui-long, LÜ Xu-dong

(Zhejiang University Biomedical Engineering and Instrumental Science School, Hangzhou 310027, China)

Corresponding author: LÜ Xu-dong

[Abstract] The development and effective use of healthcare big data are facing the challenge of medical information modeling. The openEHR-based medical information modeling was thus proposed, a sharable and reusable medical information model was established with collaborative method which was verified by electronic medical information modeling. Combined openEHR specifications and collaborative modeling can improve the scalability and interoperability of medical information models and promote the development and reuse of healthcare big data.

[Key words] Healthcare big data; Medical information modeling; openEHR

随着生物医学和信息技术的不断发展,健康医疗大数据的发展和有效利用已经成为生物医学领域的研究热点<sup>[1]</sup>。健康医疗大数据主要包括临床诊

疗数据、基因组学数据、体检数据、健康监测数据等与人体健康相关的数据。健康医疗大数据的有效利用可以为精准医疗研究提供基础,加速医学模式的变革,使其展现出预防性、预测性、个体化、参与性和精准性的“5P”医学特征<sup>[2]</sup>。然而,健康医疗大数据的有效利用仍存在诸多挑战,而支持大数据有效利用的医疗信息模型构建就是挑战之一。医疗信息建模不仅可以促进信息融合,而且有利于大数据的获取、存储、管理和有效应用。随着生物医学领域的不断发展,健康医疗大数据的类型、关联和规模正在以前所未有的速度增长,导致医疗信息模型需要具备灵活的扩展能力以适应领域的变化。与此同时,语义互操作也是影响健康医疗大数据融合、共享和利用的重要因素之一。因此,健康医疗大数据的有效利用要求医疗信息模型既要能够适应领域的变化,

[基金项目] 国家重点研发计划-精准医学研究专项-精准医学大数据处理和利用的标准化技术体系建设项目“疾病表型数据标准化技术体系建设”(2016YFC0901703)

[作者单位] 浙江大学生物医学工程与仪器科学学院,浙江 杭州 310027

[作者简介] 闵令通(1987-),男,山东鱼台人,在读博士研究生,研究方向为医学信息学研究。

[通讯作者] 吕旭东(1973-),男,浙江绍兴人,教授,博士生导师,荷兰艾因霍芬科技大学访问研究教授,中国医疗器械行业协会医用软件分会秘书长,中国医疗器械行业协会智慧与移动医疗分会副秘书长,中国医院协会信息管理专业委员会委员,中华医学会数字医学分会青年委员,发表论文 102 篇。

又要支持一定的语义互操作能力。

## 1 openEHR 介绍

openEHR 是国际标准组织认同的一系列用于促进健康数据的电子化和通用互操作的规范<sup>[3]</sup>。openEHR 方法是一种两层建模方法,为了实现信息和知识分离将信息模型分为参考模型和原型模型。参考模型是一个稳定的对象模型,定义了领域内用于信息表达和交互的不变语义和通用属性,包括标识、数据类型、数据结构和设计模式等;原型模型通过原型和模板实现对领域语义的表达,原型和模板都可以由领域专家进行定义。原型是领域知识的表达<sup>[4]</sup>,是临床内容的形式定义,是对应临床概念的通用的、可重用的、可组合的最大数据集。原型的种类包括 Composition、Section、Observation、Instruction、Action、Evaluation 和 Admin\_entry。模板是面向具体应用实现的,往往对应着系统的显示界面、打印报告、数据交互过程的数据块和消息模块等。模板的定义可以由领域专家通过对一系列原型的组装和约束完成。

作为一个开放标准,openEHR 方法已被多国采用<sup>[5]</sup>,对发展中国家具有吸引力<sup>[6]</sup>。基于 openEHR 规范采用原型驱动方法开发的信息系统或者数据利用有以下优势。第一,用户可参与。软件工程师可以基于参考模型进行系统的实现,无需充分学习和理解领域知识,领域专家可通过编辑原型和模板实现对信息系统或功能模块的控制。第二,易维护。原型驱动的方法可以通过动态加载原型的方式满足由领域知识变化而带来的需求变更。第三,支持语义互操作。稳定的参考模型可以支持信息系统的句法互操作,原型的共享和重用可提供领域语义的表达,二者结合实现语义互操作<sup>[7]</sup>。第四,领域知识共享。原型作为一种可共享的、计算机可处理的领域知识既能够促进厂商和临床专家之间的沟通交流,又能作为一个领域知识源为领域的发展提供知识服务,如标准的制定、指南的编写、信息模型的构建等。

openEHR 的两层模型以及开放式协同建模机制具备解决健康医疗大数据信息建模挑战的潜力。本文提出一种基于 openEHR 的医疗信息建模方法,使其既能满足大数据应用的灵活性需求又支持语义互操作。

## 2 医疗信息建模方法

虽然领域专家或其他利益相关者可以自行构建原型,但是自由的、分散的原型构建模式可能会导致原型定义质量参差不齐、语义重叠等妨碍语义互操作的问题。整合领域资源进行协同建模是提升原型定义质量、促进原型共享和重用的一种途径。本文的医疗信息建模方法包括:角色划分、原型构建和原型管理 3 个主要部分,其中原型构建又包含需求收集、概念表达、原型映射、原型定义 4 个步骤。

### 2.1 角色划分

原型定义过程中需要领域知识、信息技术知识和 openEHR 知识 3 种知识,其中领域知识是原型定义的来源和评估的依据,信息技术知识是原型定义的媒介和评估的依据,openEHR 知识是原型定义、评估和使用的指导和规则。针对 3 种知识来源,本方法中定义了领域专家、信息专家和建模工程师 3 种角色。每种身份都由拥有对应知识的人员所扮演,并且允许一人分饰多个角色。

### 2.2 原型构建

#### 2.2.1 需求收集

原型是一种计算机和人都能理解的领域知识表达方式,也是信息需求的一种表现形式。信息建模需求采集是原型构建的开始,同时也决定了原型构建的质量。信息需求的来源多样,包括现有信息系统需求文档、系统功能、临床指南、专家知识、发表的文章等。为了促进原型的构建,在需求采集阶段要求将众多来源的需求整理成为一个语义清晰的数据元素集合。数据元素的元数据包括数据元素名称、数据类型、数据元描述、术语绑定、数据元素来源。需求采集可以由 3 种身份中的任意一种身份完成。

#### 2.2.2 概念表达

每个原型都对应 1 个领域概念,因此概念表达是原型定义过程中重要的一环。概念的表达形式有多种,包括自然语言、思维导图和其他计算机化的表达方式。为了促进原型的协同构建,本方法规定概念的元数据包括概念名称、概念描述、概念包含的属性及其关系。在需求采集阶段的数据元素对应着概念中的属性。为了促进多角色人员的沟通与协作,概念的表达最好由领域专家完成或确认,采用思维导图作为可视化媒介。

### 2.2.3 原型映射

原型映射是对现有原型的重用。首先是依据领域概念表达对现有原型进行检索,其次是基于检索结果决定对应的原型映射操作。原型的映射类型包括 5 种类型。直接采用:领域概念的语义被原型完全表达,领域概念的属性被原型数据元素全部覆盖,概念元数据被原型元数据完全表达。原型修正:领域概念的语义被原型完全表达,领域概念的属性被原型的数据元素全部覆盖,概念元数据被原型元数据部分表达。原型扩展:领域概念的语义被原型完全表达,领域概念的属性被原型的数据元素部分覆盖。原型特例化:领域概念的语义是原型表达语义的一种特例。不兼容性修改(新建):领域概念的语义与原型表达语义不一致、不兼容。

### 2.2.4 原型定义

当领域概念未被现有原型表达或者需要进行不兼容修改时,需要构建新的原型。首先依据“问题解决逻辑”和领域概念的语义确定原型所属类型,然后依据领域概念元数据定义原型的元数据。参照领域概念的属性及属性间关系定义原型的数据元素及数据元素对应的约束。原型的定义可以由建模工程师完成。

### 2.3 原型管理

本文定义了一个管理流程,并开发部署了一个开放的原型协作平台——健康建模联盟(HMC)<sup>[8]</sup>来促进原型的协同构建、共享和重用,以获得更好的语义互操作能力。

HMC 实现了原型管理流程,并提供角色的定义、原型编辑、原型查看、原型检索、原型下载等功能。其管理流程如图 1 所示。

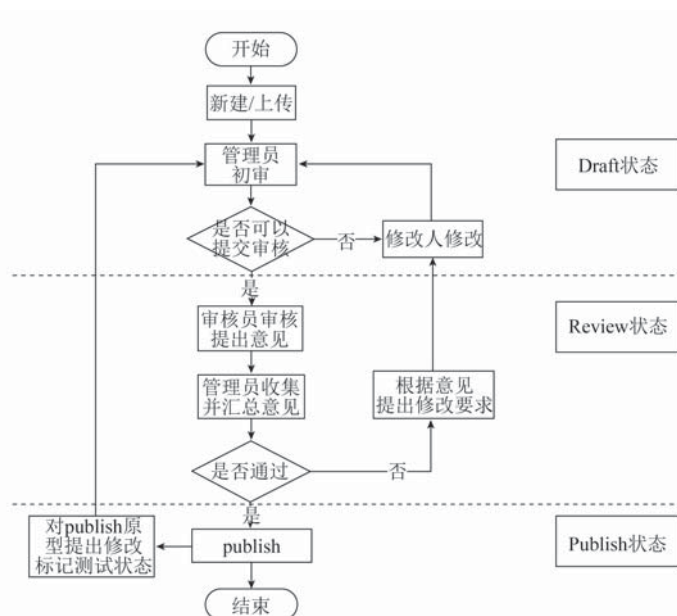


图 1 HMC 原型管理流程

## 3 案例实践

为了验证本方法的可行性,本文利用提出的医疗信息建模方法开展了针对我国电子病历数据的信息建模。针对建模需求,我们组建了一个电子病历信息建模团队,包括 2 名临床专家、1 名建模工程师和 2 名信息专家。在需求采集阶段,为了确定需求的采集范围,建模工程师分析现有的 2 个电子病历系统的数据库文档并提取字段信息作为初始需求来

源,然后利用《电子病历基本数据集》《卫生信息数据元目录》和《卫生信息数据元值域代码》3 个标准对数据元素集合进行补充和标准化,进而形成了包含 964 个数据元素定义的数据元素集合。在提取现有电子病历数据库文档字段信息的过程中,具有相同语义的不同数据元素仅保留其中一个以避免语义重叠。在数据元素补充和标准化过程中,建模工程师通过分析和比对标准中的数据元素语义和数据库



文档中提取的数据元素的语义决定后续操作,包括利用标准中数据元素补充数据元素集合和利用标准中数据元素定义替代现有的数据元素定义。在概念表达阶段,建模工程师参考现有电子病历系统数据库文档和《电子病历基本数据集》中的数据组织形式抽象出初始的临床概念,再根据临床专家和信息专家的意见对初始概念进行修正,最终归纳出 42 个临床概念。案例实践过程中,HMC 作为原型协作平台所提供的原型检索、原型编辑和原型仓库等服务有效地促进了原型的映射、定义和管理。建模团队依据归纳的领域概念基于 HMC 分别进行了原型的检索、映射、重用、新建和审核工作,共构建了 64 个原型,其中直接重用 35 个、修正 2 个、扩展 18 个、特例化 1 个、新建原型 8 个。在整个原型构建过程中,建模工程师负责需求收集、概念表达、原型映射和原型构建,而临床专家和信息专家由于工作繁忙仅负责概念表达、原型构建和原型管理的审核工作。目前,本文所构建的原型已经在 3 家三级甲等医院被用于构建临床数据中心,能够满足医院内部的医疗数据应用和医学研究的数据服务需求。

#### 4 讨论和总结

openEHR 信息模型的抽象方式是面向问题解决的,所以表达的内容不仅仅局限于电子病历或者电子健康档案等内容,也可以表达基因、兽医、建筑等领域内容。因此,基于 openEHR 方法构建的信息系统能够适应领域知识的变化。本文通过电子病历数据信息建模的案例实践证明了协同医疗信息建模方法的可行性。虽然医疗信息建模方法中领域专家和专家可以全程参与原型构建工作,但在现实情景中临床专家和专家由于工作繁忙更倾向参与审核工作。这意味着领域专家单独负责医疗信息建模的可行性较低,也从另一个角度说明了协同建模的必要性。

虽然基于 openEHR 提供的开放信息建模框架利益相关者可以通过独立的、分散的原型构建实现对信息系统功能或服务的控制,但是这种构建模式会导致原型的构建质量存在问题。一方面,利益相关者因为缺乏切实可行的原型构建方法而难以定义完全符合 openEHR 规范的原型;另一方面,在分散的、私有的原型构建模式下,建模人员可能会因为缺

乏足够领域知识而定义出不恰当、不完整、不合理的原型,难以在领域内达成一致。协同构建原型可以提高原型的定义质量,促进领域共享和重用,进而提升语义互操作能力。原型在领域内的共享和重用程度越大,基于其构建的信息系统或者数据应用的语义互操作能力越强。本文提出的协同医疗信息建模方法通过多角色、多资源的参与和协同可以提升原型定义的质量,促进原型的共享和重用,使其基于其定义的医疗信息模型的信息系统和数据应用既容易扩展,又具备语义互操作能力。

健康医疗大数据可以为精准医疗提供数据基础,其涵盖的数据种类和数据类型正以前所未有的速度增长,这要求医疗大数据的信息模型能够适应这种增长。此外,健康医疗大数据的管理和有效利用需要语义互操作能力的支持。openEHR 规范的两层建模方法能够保证构建的信息模型具有灵活扩展的能力,而基于协同建模方法构建的一套可共享和重用的原型使其具有一定的语义互操作能力。因此,本文所提出的基于 openEHR 的医疗信息建模方法可以成为应对健康医疗大数据信息建模挑战的潜在解决方案。

#### 【参考文献】

- [1] Martin-Sanchez F, Verspoor K. Big data in medicine is driving big changes[J]. Yearbook of medical informatics, 2014, 9(1): 14.
- [2] 黄小龙, 罗旭, 汪鹏, 等. 健康医疗大数据驱动下的精准医疗实施进展[J]. 医学信息学杂志, 2017, 38(9): 17-21.
- [3] What is openEHR? [EB/OL]. [2018-01-10]. [https://www.openehr.org/what\\_is\\_openehr](https://www.openehr.org/what_is_openehr).
- [4] Lopez DM, BG Blobel. A development framework for semantically interoperable health information systems[J]. International journal of medical informatics, 2009, 78(2): 83-103.
- [5] Christensen B, Ellingsen G. Evaluating model-driven development for large-scale EHRs through the openEHR approach[J]. International journal of medical informatics, 2016, 89(5): 43-54.
- [6] Pahl C, Zare M, Nilashi M, et al. Role of OpenEHR as an open source solution for the regional modelling of patient data in obstetrics[J]. Journal of biomedical informatics, 2015(55): 174-187.
- [7] Eguzkiza A, Trigo JD, Martínez-Esproncada M, et al. Formalize clinical processes into electronic health information systems: modelling a screening service for diabetic retinopathy[J]. Journal of biomedical informatics, 2015(56): 112-126.
- [8] Healthcare Modeling Collaboration[EB/OL]. [2018-01-10]. <http://openehr.org.cn/hmc>.

[收稿日期:2018-02-23]

[本文编辑:刘娜]