

临床路径本体概念框架构建方法与实证研究

朱 晶 吴正荆

(吉林大学公共卫生学院 长春 130021)

〔摘要〕 在总结国内外临床路径发展现状和纸质版临床路径管理弊端基础上,提出基于本体的电子化临床路径的概念框架构建方法,并以卫生部《乳腺癌临床路径(2009 版)》表单为基础,参考医学专业叙词表、专业教科书及相关诊疗指南对乳腺癌临床路径本体概念框架构建进行实证研究。为临床路径电子管理系统建设提供理论参考,有助于医疗信息资源的开发和利用。

〔关键词〕 疾病本体;临床路径;概念框架;本体构建;乳腺癌

Construction Methods and Empirical Research on Clinical Pathway Ontology Conceptual Framework ZHU Jing , WU Zheng-jing, School of Public Health, Changchun 130021, China

〔Abstract〕 Based on the summary of the domestic and international development status of clinical pathway and the disadvantages of paper-based clinical pathway management, the paper proposes measures for constructing ontology-based electronic clinical pathway conceptual framework. With the basis of "Breast Cancer Clinical Pathway (2009 Edition)" promulgated by the Ministry of Health, with the combination of medical subject thesaurus, specialized textbooks and related guidance on breast cancer treatment, the paper also conducts an empirical study on constructing breast cancer clinical pathway ontology conceptual framework, which provides a theoretical reference to the clinical pathway electronic management system construction, it is helpful for medical information resources development and utilization.

〔Keywords〕 Disease ontology; Clinical pathway; Conceptual framework; Ontology construction; Breast cancer

临床路径是医生、护士和其他专业人员针对某个诊断或手术所做的最适当、有顺序性和时间性的照顾计划,以减少康复的延迟与资源的浪费,使服务对象获得最佳的照顾质量^[1],被认为是 20 世纪 90 年代以来医学临床实践的重大变革之一,被称为 20 世纪末最重要的医疗模式。国外已有大量文献报道,临床路径能够以病人为中心指导医疗护理工作,可促进健康恢复,缩短平均住院日,减少住院费用,发达国家的大多数医院已应用了临床路径。

我国目前已组织制定了 22 个专业 112 个病种的临床路径表单,在全国遴选了 73 家医院作为卫生部临床路径管理试点医院,全力推进临床路径在我国的统一开展。

1 纸质版临床路径管理的不足

随着数字化医院的建设和地区间协同医疗的加强,以及国际、国内对医疗标准化的重视和实施,基于纸质版的临床路径管理不利于医疗信息的保存、传递、共享和重用。其不足主要体现在以下几个方面:纸媒介二维和静态的属性使照顾病人信息记录的交流受到时间和空间的限制,医疗小组很难

〔收稿日期〕 2010-11-04

〔作者简介〕 朱晶,硕士,发表论文 3 篇;通讯作者:吴正荆,博士,教授,发表论文 40 余篇。

实时监控整个临床路径的实施及变异过程；纸媒介已不能满足医疗信息交流与共享所必需的信息化、规范化的要求；纸媒介更改的随意性极大降低了临床路径的推广、重用和共享。上述缺点较大地影响了临床路径的整体效力，需要一种基于网络的电子化临床路径（Electronic Clinical Pathway, ECP）管理来实现有效操作和自动监控，以及实现临床路径的标准化，包括诊疗计划和数据交换的标准化^[2]。

2 临床路径电子化研究

2.1 研究内容

临床路径电子化研究主要包括诊疗计划标准化和数据交换标准化两个方面。诊疗计划标准化主要包括诊疗流程标准化、管理方案标准化、实施与评价标准化。数据交换标准化主要包括临床路径管理数据在各医疗机构间传输和交换的标准化，它对于医疗知识组织、共享和重用意义重大。

2.2 国外研究

在国外，加拿大达尔豪西大学的 Abidi^[3]等人提出一个语义 Web 模型，此模型可实现异构领域知识与操作知识的合成，为以病人为中心的面向服务架构的临床路径电子管理奠定了技术基础；Hurley 和 Abidi 将临床路径的实践知识加以抽取并用本体详细表示了临床路径的结构和功能^[4]。在 PARADIGMA^[5]（Participative Approach to Disease Global Management）项目中，研究人员用本体定义了一般临床路径的概念结构，临床路径概念划分为检查、诊断、治疗等，在设计临床路径时可以根据病种细化成具体的路径。

2.3 国内研究

国内的高校和科研院所也在进行积极的探索和研究。李从东等^[6]构建了面向临床路径实施的知识重构整体任务模型、过程模型及其方法模型；王占明等^[7]尝试将临床路径管理与电子病历进行结合，采用半结构化的新型电子病历来促进临床路径的管理；裴振安等^[8]建立了“智能型住院临床路径”，

可实现对病人住院过程中实施及变异情况进行实时统计，将其集成到电子病历中，实现数据库共享，丰富了知识库。但目前有关临床路径电子化管理的标准化和规范化研究还有待加强。

3 基于本体的临床路径概念框架构建方法

3.1 概述

目前的本体大多是根据各自的学科领域和具体工程的不同特点而构建，其构建过程各不相同，尚没有一套经权威标准化机构认可的本体构建方法。当前国内外构建医学领域本体时较多采用的是美国斯坦福大学医学院开发的 7 步法，本文也主要借鉴了 7 步法的思路来构建临床路径本体。临床路径涉及人员复杂，包括临床医学专家、护理专家、药学专家、心理学专家、营养师、检验人员以及行政管理人员等，内容涉及某一特定疾病的预期效果、评估、多学科服务措施、病人及其家属的相关教育、会诊、营养、用药、活动、检验与检查、治疗和出院计划等各个方面，涉及的概念广泛而繁杂。临床路径本体知识获取需要按照时间顺序抽取包括全部或部分上述过程的相关术语并分类，以及对类的属性和关系进行抽取。目前卫生部制定的临床路径表单是以住院时间为横轴，以诊疗护理工作为纵轴的疾病诊疗流程，作为目前最规范和标准的临床路径文本，在实验阶段可先以表单为基础进行概念和关系的抽取。

3.2 临床路径本体专业领域范畴确定及复用的可能性分析

构建临床路径本体先要明确病种，其次是确定疾病的诊疗范围，如《乳腺癌临床路径》表单中限定乳腺癌临床路径仅应用于手术治疗的患者，并且规定了适用对象为确诊乳腺癌的患者。实际上，随着临床路径在应用中的逐渐成熟和完善，临床路径应该包含病人从挂号、诊疗到随诊的全过程，临床路径本体的领域范围也要随之变化。为此，本体需求分析既要维持需求的稳定性和精确性，也要在实施过程中不断地进行动态调整。通过合并现存的本

体结构进行本体重用，应该是获得领域信息的最行之有效的方法。然而，现今的许多本体是用户为了解决一时之需而临时自行构造的，目前可以复用的领域本体少之又少。国内外对于临床路径本体仍主要处在理论研究阶段，目前既没有一个成型的框架可供参考，更没有一个成熟的针对特定疾病的临床路径本体可以复用。

3.3 临床路径本体概念抽取及定义类和类间等级体系

临床路径本体相关概念的抽取，主要以临床路径表单为基础，同时选用相关领域权威性医学书刊等作为参考资料，进行语义分析和概念抽取，使抽取的概念具有一定的科学性、权威性和实用性。抽取概念的方法有自顶向下法、自底向上法和综合法。3 种方法各有优劣，具体采用哪种方法取决于设计者对某疾病诊疗过程的理解程度。考虑到临床路径本体涉及到诊断、治疗、护理等一系列既独立又交叉的医护过程，为避免遗漏，产生过多的细节和反复而导致重复工作，以及分别抽取易产生概念的不一致性，本文倾向于综合法的思想，暂不考虑概念的属性、概念间关系及表达上的重复，将概念按照临床路径的诊疗时间顺序，逐个抽取，列出一份最全的术语清单。从中选取最基本、最常见的概念，并用准确无二义性的术语作为类。本文确定了人物、住院日程、医疗行为、病理分型和临床药物作为核心概念，满足含义精确、互不相交、覆盖整个临床路径领域知识的要求，见图 1。在此基础上定义概念间的父子关系和类——实例关系。

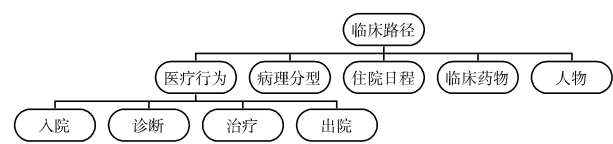


图 1 乳腺癌临床路径本体模型框架结构

3.4 定义类的属性并创建实例

类的属性用来描绘类间的逻辑关系，包括对象属性和数据类型属性两种类型。将个体关联到个体

的属性称为对象属性，通过设置属性的定义域和值域来实现；将个体关联到数据类型的属性称为数据类型属性。本文借鉴中国中医科学院郭玉峰等^[5]翻译的 SNOMED CT 连接词中文版，将抽取的关系词进行标准化处理，并补充了 SNOMED CT 中包含、抽取时遗漏的关键词，以保证关系词的全面和准确。定义属性的同时可直接定义属性的分面，用来描述属性的不同特性，如赋值类型、允许的赋值以及赋值的基数。某一属性的赋值类型主要有布尔型、浮点型、整数型、字符串型、日期型、日期时间型和时间型等。实例是本体中的实际数据。因临床路径本体涉及到的领域较多、概念繁杂，尽量在定义类的同时定义实例，以免去重新理顺概念的反复工作。但在本体编辑工具中（本文以 protégé 作为编辑工具）输入实例前，最好类和属性的结构已经确定，否则以后的改动将可能导致某些信息的丢失。

3.5 利用 protégé 软件编辑临床路径本体

目前常用的编辑工具有：Java Ontology Editor (JOE)、Ontolingua、Chimaera、OntoEdit、OilEd 和 Protégé，多数软件往往不支持中文输入，易出现乱码显示。其中美国斯坦福大学的 Protégé 软件相比之下更加易用，界面友好、性能齐全、具有多种插件、提供多种本体的表现形式，成为国内外众多本体研究机构的首选工具。

4 乳腺癌临床路径本体概念框架构建实证研究

4.1 乳腺癌临床路径本体概念抽取

本文将按照卫生部的《乳腺癌临床路径（2009 版）》表单中规定的操作流程进行乳腺癌临床路径本体构建的实证研究。根据美国国立综合癌症网络 (National Comprehensive Cancer Network, NCCN) 编制的 *NCCN Clinical Practice Guidelines in Oncology*TM (V. 2. 2008)，我国乳腺癌专家组结合中国国情，制订了《2008 年中国版 cNCCN 乳癌治疗指南》（以下简称《指南》）。本文主要结合《指南》，参照卫生

部下发的《乳腺癌临床路径（2009 版）》表单及其他相关文件、人民卫生出版社出版的《外科学》（第七版）、江苏科学技术出版社出版的《乳腺癌综合治疗》及部分学术文献、《医学主题词表 MeSH（2004）》进行乳腺癌临床路径的概念及关系抽取。经过语义分析，共抽取概念 205 个，部分概念集合和层级体系，见表 1。

表 1 乳腺癌临床路径本体概念集合和层级体系（部分）		
内容	名词性概念	动词性概念
病理分型：	非浸润性癌（0 期）、小叶原位癌、导管原位癌、浸润性乳腺癌、早期浸润性癌（Ⅰ、Ⅱ期、部分ⅢA 期）、局部晚期乳腺癌（ⅢC、ⅢB 期、部分ⅢA 期）、转移或复发性乳腺癌（Ⅳ期）	表现、测量、检查、诊断、分析、预测、并发、治疗
人物：	医生、主任医师、副主任医师、主治医师、住院医师、护士、患者	
手术：	乳腺癌根治术、乳腺癌扩大根治术、乳腺癌改良根治术、全乳房切除术、保留乳房的乳腺癌切除术	
住院日程：	住院日、术前准备日、手术日、术后护理日、出院日	
：	：	：

4.2 乳腺癌临床路径本体概念层次关系划分及类间关系建立

按照表单中的时间顺序，将含义明确的概念分别纳入人物、住院日程、病理分型、临床药物和医疗行为 5 个大类中，其中医疗行为大类下按照临床路径诊疗流程分为入院、诊断、治疗和出院 4 个子类，对于含义不明的概念可先做另类处理，待最后综合分析时再做进一步归类。此处划分的概念间等级体系，主要是父子关系和类——实例关系，概念间的其他逻辑关系，可通过属性连接。共抽取 107 个属性（包括 94 个数据属性，13 个对象属性），其中涵盖 SNOMED CT 中抽取的关系词 10 个：发现部位、由于、在……之后、严重程度、所见告知者、直接操作部位、直接形态学、方法、重点是、使用的物质。其中部分属性具有子属性。

4.3 基于 protégé 软件的乳腺癌临床路径本体构建

利用 protégé 软件进行本体编辑。本文采用

protégé3. 3. 1 进行本体构建，在主界面中默认存在 5 个标签：Metadata（元数据）、OWL Classes（OWL 类）、Properties（属性）、Individuals（实例）、Forms（窗体）。分别将类、属性和实例输入在相应的版块中。当大量的实例被赋予相应的属性及相应的值时，就形成了相关领域的知识库。

乳腺癌临床路径本体可视化。信息可视化是信息科学的重要研究领域，并已成为一个研究热点，信息可视化技术的目的是“给信息以形象”，让信息以视觉化的方式表示^[6]。利用 protégé 中的 TGVizTab 插件可对信息进行多维显示及多级显示。以医生的一个实例“张斌”为例，在“Radius”选项中选择 3 级显示，显示的信息为张斌是一位医生，职务是主任医师，擅长乳腺非浸润性癌和浸润性癌的诊断和治疗，同时可以显示医生类和病理分型类的下位类，见图 2。

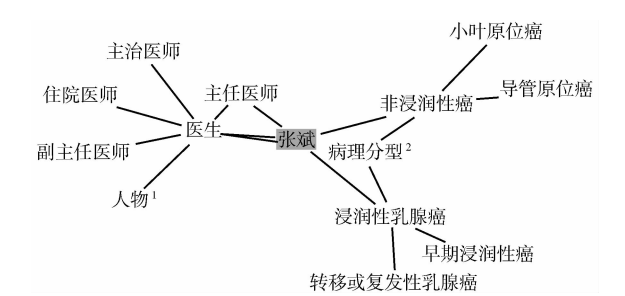


图 2 医生实例“张斌”语义关系的 3 级可视化显示结果

以“乳腺癌扩大根治术”为可视化中心点，在“Radius”选项中选择 1 级显示，见图 3。本图显示的语义为，经诊断为乳腺癌的张敏敏和赵丽两位患者，在手术日当天，由张斌医生为其进行乳腺癌扩大根治术治疗。当然还可以将其显示级数增加，增至 2 级的显示结果，见图 4。级数增加以后显示的信息更加丰富，但也增添了很多无关信息，在实际应用中可根据需要选择适当的显示级数。TGVizTab 可视化的不足之处是，界面只能显示对象，而不能显示数据属性。同时需要将鼠标放在连接线上，方能显示两个类的关系，而且此插件不支持中文属性显示。应继续研究开发支持中文显示的本体编辑工具，同时扩大和强化本体编辑工具的功能，使之更加科学化和人性化。

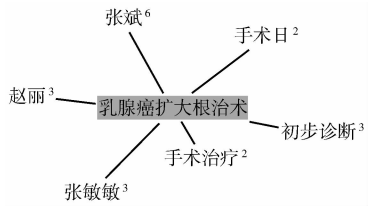


图3 “乳腺癌扩大根治术”语义关系的
1 级可视化显示结果

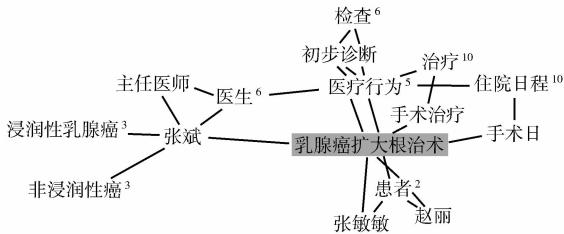


图4 “乳腺癌扩大根治术”语义关系的
2 级可视化显示结果

5 国内的应用与展望

领域本体建设是一项复杂的系统工程，不是轻而易举就能够构建的，需要脚踏实地进行理论与应用试验。首先，要对领域本体构建的理论、方法与技术进行深入研究，建立公认的核心概念和完整的领域本体构建的理论体系，以及具有适用性、针对性的领域本体构建技术体系。其次，面向某一应用领域进行领域本体系统的构建实验。在此基础上，总结经验，寻找规律，将其提升为理性认识，再进一步指导实践，以促进领域本体的实验研究更快、更有效地走向应用。

本文借助人工智能领域的本体方法，构建电子化临床路径的概念框架，为推进临床路径数据交换的标准化，建立电子化临床路径管理系统底层知识概念框架基础提供参考。从病种向病例过渡，才应该是临床路径的重大贡献和最终目标。所设计的临床路径本体是针对某一病种的固定的诊疗章程，而

从病种到病例才应该是临床路径的技术主线。在日后的研究中，临床路径本体也应该结合电子病历对病例进行分型，进一步获取领域本体公理，通过变异分析、本体推理等功能，辅助医生针对特定病人制定出特定临床路径日程表。临床路径本体的实际构建当中，应该包括更加细致全面的疾病诊治方法和住院出院流程，其概念体系应进一步扩展和完善。文中的本体构建过程仍主要以手工为主，应努力通过开发机器学习，以自动或半自动方式获取概念及概念之间的关联，这是临床路径本体进化的关键。

参考文献

- 1 郭锦秋, 黄勇. 标准化临床路径与地域间协同医疗 [J]. 解放军医院管理杂志, 2009, 16 (5): 403 - 404.
- 2 方立珍. 临床路径——全新的临床服务模式 [M]. 长沙: 湖南科学技术出版社, 2002: 7.
- 3 Abidi S S R, Chen H C. Adaptable Personalized Care Planning via a Semantic Web Framework [C] //20th International Congress of the European Federation for Medical Informatics (MIE 2006), 2006.
- 4 Hurley K F, Abidi S S R. Ontology Engineering to Model Clinical Pathways: towards the computerization and execution of clinical pathways [C] //Twentieth IEEE International Symposium on Computer - based Medical Systems (CBMS' 07), 2007.
- 5 Di Leva A, Reyneri C, Sonnessa M. The PARADIGMA Approach for Cooperative Work in the Medical Domain [J]. International Journal of Information Technology, 2005, 1 (4): 168 - 172.
- 6 李从东, 杜红. 面向分布式知识发现的知识网格框架研究 [J]. 情报科学, 2005, 23 (2): 281 - 283.
- 7 王占明, 黄志中, 王景杰. 临床路径管理与电子病历结合应用研究 [J]. 医疗卫生装备, 2008, 29 (8): 50 - 51.
- 8 裴振安, 潘佩光, 潘奔前. 智能型住院临床路径的开发与实施初探 [J]. 中国农村卫生事业管理, 2006, 24 (4): 60 - 61.