

SNOMED CT 术语分类体系设定 学科背景的探讨*

□郭玉峰 刘保延** (中国中医科学院 北京 100700)

尹爱宁 (中国中医科学院中医药信息研究所 北京 100700)

摘要: SNOMED CT (Systemized Nomenclature of Medicine, Clinical Terms, 系统化临床医学术语集) 依据特定的原则, 进行了现代临床医学相关术语的分类设定, 建立起 19 个具有内在关联关系的层级体系 (Hierarchies)。发布 5 年来, SNOMED CT 已得到了 30 余个国家和地区的引进, 广泛应用于现代医学临床科研的重要领域, 说明该术语集具有充分的合理性和实用性, 进行术语分类的原则符合现代西医学的自身规律。本文在回顾 SNOMED 术语顶级分类几次重大调整的基础上, 以剖析 SNOMED CT 顶级分类层级之间的相互关系为切入点, 提出现代西医学本体论是 SNOMED CT 进行术语分类框架体系设定的理论依据。

关键词: SNOMED CT 术语分类 中医临床术语标准化

系统化临床医学术语集——“SNOMED CT”, 由美国病理家学会研发并于 2002 年正式推出, 5 年来已得到包括美、英、德、法、日等发达国家在内的 30 余个国家和地区的引进, 成功地应用于现代医学临床科研的诸多重要领域, 证明该术语集具有较为充分的合理性和实用性。我们通过对 SNOMED CT 构架体系、研发历史、版本调整等进行了较为深入系统的剖析研究后发现, SNOMED CT 作为系统化的现代临床医学标准术语集, 其术语收录范围基本覆盖了临床及科研所需的全部术语种类, 其进行系统内部术语分类框架体系设定的原则符合现代西医学自身的固有规律, 使术语集

构架不仅全面展示了现代西医学对疾病的本体认识观, 而且贴近临床医生的诊疗思维模式、便于科研工作对重点数据的检索获取。本文在回顾 SNOMED CT 研发历史中进行的几次术语分类框架重大调整的基础上, 结合对现代西医学本体论的研讨以及对 SNOMED CT 术语分类框架体系的分析, 总结提出现代西医学本体论是 SNOMED CT 进行术语分类框架体系建设的理论基础。

一、SNOMED CT 的术语分类框架体系

SNOMED 的持续性研究自其前身——1965 年推出的“系统化病理学术语集 (SNOP)”算起, 迄今已经历了 40 余年的历程, 术语集的体系构架逐步由最初以病

组稿日期: 2007-07-28

* 北京市科委重大计划项目 (H020920010130): 中医药防治重大疾病临床个体诊疗评价体系研究, 负责人: 刘保延。世界中医药联合会临床疗效评价专业委员会成立大会暨首届国际学术交流会议论文。

** 联系人: 刘保延, 本刊编委, 主任医师, 国务院政府特殊津贴专家, 中国中医科学院副院长, 主要研究方向: 中医临床研究方法、针灸学、中医信息学, Tel: 010-64014411, E-mail: liuby@tcmcc.com。

理学术语为主体的4轴(Axis)分类模式拓展到目前能够覆盖现代西医学临床科研所需全部术语种类的19层级(Hierarchy)体系的分类模式,见图1。

从最初SNOP采用的4轴分类到SNOMED III的11模块分类再到SNOMED CT的19层级分类的术语体系构架调整过程中可以看出,伴随术语集应用目的和应用环境的改变,收录术语数量和种类的逐渐增加,SNOMED在不断进行着术语分类框架体系的优化和细化。以SNOMED三系列中最为成熟的版本SNOMED3.4与SNOMED CT第一版SNOMED CT2002的术语分类框架为例进行比较(表1),SNOMED CT 2002在SNOMED3.4的术语分类框架的基础上将原有顶级术语分类进行了拆分(如顶级分类“物理因素,活动和力”被拆分形成新的顶级分类“物理对象”、“物理力”)与合并(如“局部解剖学”、“形态学”合并为“身体结构”,“职业”合并入“社会环境”),并适当增加了

新的顶级术语分类(如“物质”、“标本”)。

2007年1月发布的SNOMED CT最新修订版中增加了“Record artifact—医学记录”、“Situation with explicit context—有明确上下文关系的情况”两个新的顶级分类,并废止了原有顶级分类“Context - dependent categories—上下文依赖性种类”,形成了目前具有19个顶级术语分类的术语集构架。

二、现代西医学本体论是SNOMED CT
构架术语分类体系的理论依据

SNOMED CT之所以能够得到世界范围的广泛接受,除去当前美国经济与现代西医学的发展居世界领先水平、SNOMED组织研发能力强、相关标准确立掌握较早等因素外,我们认为SNOMED CT术语分类体系的构建在学术上具有较为坚实,经得起推敲的理论基础是其中一个重要原因。

表1 SNOMED 3.4 与 SNOMED CT2002 的术语分类框架比较

SNOMED3.4 的 11 种术语顶级分类			SNOMED CT 2002 的 18 种术语顶级分类	
1	Topography	局部解剖学	Clinical Finding	临床所见,包括疾病诊断
2	Morphology	形态学	Finding/Dis-ease	
3	Function	功能	Procedures	操作
4	Living Organisms	活有机体	Observable entity	观察对象
5	Chemicals, Drugs, and Bio-logical Products	化学制品、药品和生物制品	Body structure	身体结构
6	Physical Agents, Activities and Forces	物理因素,活动和力	Organism	有机体
7	Occupations	职业	Substance	物质
8	Social Context	社会环境	Pharmaceutical/Biological product	药品/生物制品
9	Diseases/Diagnoses	疾病/诊断	Social context	社会环境
10	Procedures	操作	Linkage concept	连接概念,包括确认连接和
11	General Linkage/Modifiers	连接词/修饰词	Link asser-tion/Attributes	属性连接
12			Specimen	标本
13			Qualifier value	限定值
14			Context - dependent categories	上下文依赖性种类
15			Physical object	物理对象
16			Physical force	物理力
17			Events	事件
18			Environments/geographical locations	环境和地理定位
			Staging and scales	分期与等级
			Special concept	特殊概念

注:表中各项级分类的排序均来自 SNOMED 官方网站

1. 标准化、分类与 SNOMED 术语分类框架体系调整

近 30 年来医学信息学的高速发展,促使与医学信息标准化相关的工作原则逐渐成熟并系统化,知识本体等理论的提出和应用,为术语标准化研究中术语体系构架的设定提供了科学依据。分类是实现信息标准化的重要基础,分类的准则取决于应用目的,所谓分类就是将领域内概念根据属性关系实现概念的有序化,其中属于并行关系的概念分成同一类,而属于包含关系的则进一步分成子类,在分类过程中贯穿始终的序化标准称为轴,分类的过程就是找出领域概念中所有的本质特性即确定轴心,围绕轴心按特性的属性关系,再分为亚目、细目等^[1]。

因此,针对领域概念进行的术语体系框架设定,需要以明确术语集的应用目的和应用环境为前提,对所涉及领域内术语的全部种类进行系统性整理和分析,理清这些术语种类之间的相互关系,将具有最高级别且非重叠的种类设为术语集的顶级分类,每个顶级分类下再按照包含属性逐层设定子类,“纲举目张”,从而实现领域概念的有序化表达。

以 SNOMED 版本更替中术语分类框架体系调整为例,SNOP 建立的 4 轴体系能够满足以病理学术语为主体内容的分类和检索,升级到 SNOMED CT 建立的 19 层级体系能够满足信息化时代临床科研对常用术语的调用,这是根据应用目的和应用环境的变化拓展了术语集内容和体系构架:由于术语集覆盖术语范围的扩大,非重叠的顶级术语种类逐渐增多,需要调整术语集体系构架增加相应的顶级分类,以包容增加的术语种类;由于应用目的从病理学术语的汇集、检索为主转变到以支持临床电子病历书写为主,因此提升了分类“临床所见/疾病”及“操作”的地位,使其成为术语分类体系构架中最主要的两个顶级分类,以满足病

历书写、卫生统计以及临床科研对此两类术语频繁调用的需求,并依据其他术语对这两类术语的支持关系而设定了不同的术语顶级分类。可见,SNOMED 充分应用医学信息标准化的原则,设计了较为全面完整、纲目清晰的现代西医学术语分类体系构架。

2. 现代西医学本体论的探讨

医学本体论是指“关于医学的”最普遍、最一般、最根本、最高的根据、本质、基础的知识或理论,它是医学赖以存在的本质基础或逻辑根据,简言之,是指关于医学在本质上究竟是什么的理论问题^[2]。从医学诞生的一刻起,战胜疾病、解除患者痛苦就是医学存在和发展的最根本目标,因此,对疾病本质的认识在很大程度上决定了医学的本体论。那么现代西医学认为疾病的本质是什么,又是通过什么手段认识疾病的?纵观现代西医学几百年来发展过程不难发现,其间始终贯穿着一条主线,即采用还原论的方法,利用同时期科技进步提供的新手段,以人体解剖学为基础,将寻找人体上实体可见的解剖形态学改变作为认识疾病的方法,人体局部解剖及形态学的改变被视为疾病的本质,造成这种改变的原因就是致病因素,消除这种解剖形态学改变及其诱发因素被作为治疗的方向和疗效评价的重要指标,其中以 19 世纪魏尔啸创立的细胞病理学和巴斯德提出的细菌致病学说标志着现代西医学对疾病本质认识的成熟。

自 16-17 世纪起,显微镜的发明、改良并引入医学研究极大地增强了人类认识微观世界的能力,不仅发现了凭借肉眼难以观察到的众多微生物,更将寻找解剖形态学实体改变的对象由宏观的大体解剖沿着器官、组织、细胞、分子的水平逐步微观化。20 世纪 70 年代兴起的分子生物学,试图从分子水平研究作为生命活动主要物质基础的蛋白质、基因结构与功能,以期阐明生命现象本质。但究其实质,依然是从实体的形态结构变化入手以寻求疾病的证据,只不过是组织细胞水平进一步

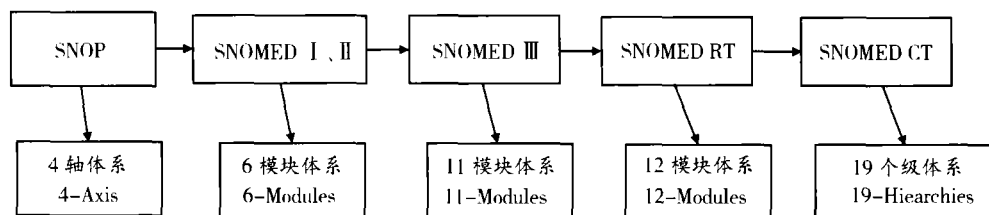


图 1 SNOMED 版本和术语分类框架体系的演变 (1965 年 - 2007 年)

微观化到了生物大分子水平而已。而目前普遍的临床诊疗行为依然遵循着现代西医学疾病本质认识观的指导,将确定解剖形态学改变及捕获致病微生物作为确定疾病诊断和评价疗效的主要标准,并

未出现切实可行、能够指导临床疾病诊疗的突破性理论更新。此外,早在 30 年前恩格尔教授提出“生物医学模式向生物-心理-社会医学模式转变”的观点虽已被广泛认同,但至今也未真正对主流西医学的临床诊疗流程造成本质影响,这从另一侧面证实了以寻找实体的解剖形态学改变作为证据的疾病观仍是现代临床西医学的主流。

3. SNOMED CT 术语体系构架剖析

通过对 SNOMED CT 现有术语顶级分类以及 SNOMED CT 建立起的术语分类间较为标准化的属性关联关系的剖析研究,我们认为 SNOMED CT 术语体系构架设定的理论依据是建立在现代西医学本体论上的疾病观。

图 2 为 SNOMED 官方网站提供的术语集框架体系示意图*,从 2002 年发布的 SNOMED CT 第一版到 2007 年的最新版,其间虽然进行了部分顶级术语分类的调整,但其整体上的术语集体系构建原则并未发生变化。图中列出了术语集分类框架中两个最重要顶级

分类“临床所见/疾病”和“操作”以及与其关系最为密切的几种“支持性术语分类”。属于“临床所见/疾病”、“操作”分类轴的术语,往往在临床数据检索中相关度高、作用大,是临床科研所关注的重点;而属于支持性术语分类轴内的术语,通过特定的属性关系与“临床所见/疾病”、“操作”分类轴下的术语相连接后,能够为其提供具体的值,从而确切地描述临床事件包括其产生的结果,因此被称作支持性术语分类。

图中“临床所见/疾病”、“操作”均通过其下框图中列出的属性关系和中间部分的支持性术语分类建立起关联关系,也就是说,支持性术语分类通过关联关系为两大顶级分类提供了具体的值。例如:通过属性关系“Finding Site—所见的部位”将“临床所见/疾病”与“Normal Anatomy—正常解剖学”连接在一起,形成“临床所见/疾病”+“所见的部位”+“正常解剖学(部位)”的表达模式,能够较为理想地表达“某种临床所见发生的具体部位”的临床意义。再如,通过属性关系“Causative Agent—致病因素”将“临床所见/疾病”

与“物质、有机物、物理对象、物理力”连接起来,为“临床所见/疾病”进行了“致病因素”的赋值,即某种“临床所见/疾病”有可能具有的物质、有机体、物理对象、物理力等方面的致病因素。

将上图中所有经红线连接起的术语分类进行总结,得到表 2 列出的结果。

从表 2 中不难发现,虽然在 SNOMED CT 发布前的 SNOMED 版本中作为最重要顶级分类的“局部解剖学”、“形态学”两种术语分类被降级,已不再是 SNOMED CT 中所称的最主要顶级

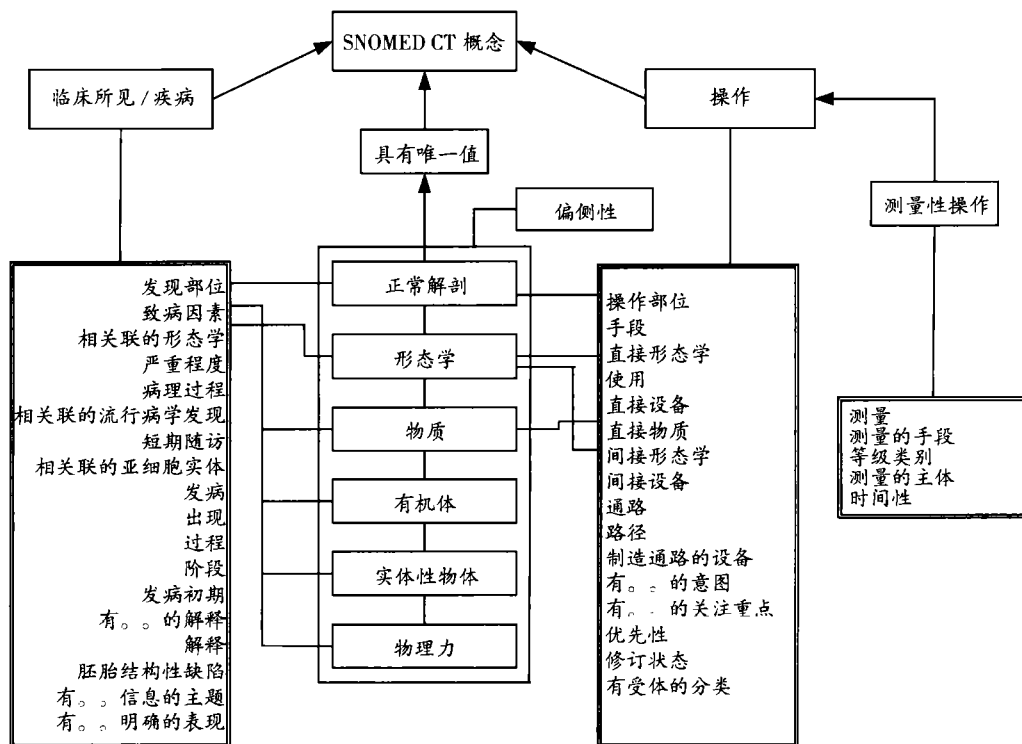


图 2 SNOMED CT 体系构架示意图

* 图片来自 <http://www.snomed.org>

表2 SNOMED CT分类体系示意图解读

顶级分类	属性关系	(顶级)分类
临床所见/疾病	发现部位	正常解剖(属于顶级分类“身体结构”)
	伴随的形态学改变	形态学(属于顶级分类“身体结构”)
	致病因素	物质 有机体 物理对象 物理力
操作	操作部位	正常解剖(属于顶级分类“身体结构”)
	直接、间接的形态学改变	形态学(属于顶级分类“身体结构”)
	直接物质	物质

分类,但两者与“临床所见/疾病”、“操作”仍然具有最为密切的联系,在描述“临床所见/疾病”及“操作”的具体部位,解释疾病的发病原因和病理变化中起到决

定性的支持作用;包括“有机体”在内,几种属于客观致病因素的术语分类则为解释发病原因提供了支持。可见,将支持“临床所见/疾病”的解剖形态学实体变化证据、确定造成解剖形态变化的致病因素等相关术语汇集形成顶级分类,是 SNOMED CT 术语分类框架的构建原则,而这恰恰正是西医学本体论上疾病观的具体体现。

三、小 结

SNOMED CT 以现代西医学本体论疾病认识作为理论依据设计了现代西医临床医学术语框架体系,其术语体系在组织结构、顶级分类设定等方面能够正确地反映现代西医学自身的固有规律,因此得到业界来自临床和科研领域较为广泛的认同和支持,这点值得中医临床术语标准化研究工作借鉴学习。

参考文献

1 贝梅尔编,包含飞译.医学信息学.上海:上海科技出版社,2002.
2 王华生,农乐颂.医学本体论的当代转向—医学教育模式改革的前提性问题.医学与哲学(人文社会医学版),2006,27(5):10-11,29.

Theoretical Basis of SNOMED CT Classification

Guo Yufeng, Liu Baoyan, Yin Aining

(China Academy of Chinese Medical Sciences, Beijing,100700)

SNOMED CT, or Systemized Nomenclature of Medicine, Clinical Terms, makes Western medical terms into 19 hierarchies, using a special classification methodology. The widespread and successful application of SNOMED CT in some 30 countries/regions has confirmed the fact that SNOMED CT is a rational and feasible nomenclature, and its classification methodology is compatible with Western medicine’s intrinsic rules. Authors believe that the ontology of modern Western medicine is the theoretical foundation of SNOMED CT classification, through reviewing the strict regulations of SNOMED CT top-level hierarchies in the past decades, and the mutual relationships between top-level hierarchies of SNOMED CT.

Keywords: SNOMED CT; terms classification; standardization of TCM clinical terms

(责任编辑:王 瑀,责任译审:邹春申)