

# 中医药学术语系统研究概述

贾李蓉, 李海燕\*, 刘静, 董燕, 刘丽红

中国中医科学院中医药信息研究所, 北京 100700

**摘要:** 中医药术语规范化是信息化的基础。随着中医药信息化的发展, 术语系统也随之发展起来。中国中医科学院中医药信息研究所在术语规范化和标准化的研究方面一直走在行业前列, 该所研制的中国中医药学主题词表、中医药学语言系统和中医临床术语系统都已具有一定规模, 其应用与推广工作也在进行中。

**关键词:** 中医药; 术语系统; 术语规范化; 中医药学主题词表; 中医药学语言系统; 中医临床术语系统

doi: 10.3969/j.issn.2095-5707.2015.05.002

## Overview of the Research on the Three Terminological Systems of Traditional Chinese Medicine

JIA Li-rong, LI Hai-yan\*, LIU Jing, DONG Yan, LIU Li-hong

(Information Institute of Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China)

**Abstract:** The standardization of TCM terms is the basis of informatization. The terminological systems have developed with the development of the TCM informatization. Institute of Information on Traditional Chinese Medicine of China Academy of Chinese Medical Sciences has been at the forefront of research on terminological normalization and standardization. The three TCM terminological systems: Traditional Chinese Medicine and Materia Medica Subject Headings, Traditional Chinese Medicine Language System and Traditional Chinese Medicine Clinical Terminological System developed by the institute have had a certain scale. The application and promotion of the three systems are under way.

**Key words:** traditional Chinese medicine (TCM); terminological system; terminological normalization; Traditional Chinese Medicine and Materia Medica Subject Headings; Traditional Chinese Medicine Language System; Traditional Chinese Medicine Clinical Terminological System

自 20 世纪中叶以来, 科技术语的数目呈现出指数增长态势, 尤以生命科学和信息科学最为突出<sup>[1]</sup>。术语激增, 带给人们理解上的困难, 随之而来的是使用混乱和交流障碍。对于中医药学科而言, 信息化的发展, 对于其术语概念的规范要求越来越高。术语系统则是实现名词术语规范化的有效手段<sup>[2]</sup>。中国中医科学院中医药信息研究所自 20 世纪 80 年代起开始研制中医药学主题词表, 在中医药学科的术语系统构建中迈出了探索的一步。到 21 世纪初,

我所又借鉴国际先进技术和方法, 分别针对不同应用环境, 研制了中医药学语言系统和中医临床术语系统。目前, 这三大术语系统都已具有一定规模, 其应用与推广工作也在进行中。

### 1 中国中医药学主题词表

《中国中医药学主题词表》<sup>[3]</sup>是我国第一部中医药专业词表, 被国内外医学及中医药学信息界广泛采用。本词表以其科学性、适用性以及 MeSH 词表的兼容性获得使用者好评。它的诞生标志着中医药学信息处理和信息服务跨入计算机信息时代, 并以其独特的学术内涵和广泛的兼容性为中医药信息产品在国内外推广应用创造重要条件, 提供技术保障。

#### 1.1 研究历程

中医药学主题词表的研究起步于 20 世纪 70 年代, 发展于 80 年代。早在 1976 年, 国家科学技术委员会领导的“748 工程”《汉语主题词表》中,

基金项目: 中国中医科学院基本科研业务费自主选题 (ZZ070301); 中国中医科学院科技创新团队项目 (PY1306)

第一作者: 贾李蓉, 助理研究员, 研究方向为中医药信息化与标准化。E-mail: jialr@163.com

\*通讯作者: 李海燕, 研究员, 研究方向为中医药信息标准。

E-mail: lihy@mail.cintcm.ac.cn

已将中医药学内容置于重要位置。1984年,原中国中医研究院中医药图书情报中心(现中国中医科学院中医药信息研究所)开始着手进行中医药学专业主题词表的编纂。经过反复修改与论证,《中医药学主题词表》于1987年问世。此后经过多年的实践与应用,于1996年以机读版和印刷版的形式出版了修订版,更名为《中国中医药学主题词表》。2004年12月,国家中医药管理局将修订1996年版《中国中医药学主题词表》列为标准化研究项目,经过广泛调研国内外医学主题词表的理论、方法和经验,进行了大量的词频分析及用户反馈意见分析后,出版了《中国中医药学主题词表》2008版。2013年,我所设立课题项目《中国中医药学主题词表网络版研制与修订》,基于近5年中医药学文献主题词标引和检索实践,在词频统计<sup>[4]</sup>、共词分析<sup>[5]</sup>以及用户反馈意见分析等基础上进行选词与修订,建立中医药学主题词表管理系统,实现网络发布,同时培养中医药学主题词表维护的专门人才。2015年底将完成该词表网络版的发布,正式投入使用。

## 1.2 词表结构

《中国中医药学主题词表》由字顺表、树形结构表、副主题词表、出版类型表、医学家姓名附表、索引表6个部分构成。

1.2.1 字顺表 又称为主表。指将全部主题词及入口词按汉语拼音顺序排列而生成的印刷版或电子版表。主题词款目结构为汉语拼音、主题词名称、主题词英译名、树型结构号、主题词定义、标引及编目注释及参加项。

1.2.2 树形结构表 又称范畴表。指将主题词按中医药学学科分类划分为15个子类目及二级子类目,列出隶属于该类目的全部主题词,按属分关系逐级展开,呈树形结构,每个主题词均有双字母数字标识以显示主题词的级别;该词表的分类方式可以与MeSH词表的分类相兼容。

1.2.3 副主题词表 包括1个专题副主题词表及3个编目副主题词表。专题副主题词表收录了副主题词93个,其中MeSH副主题词83个,中医药学副主题词10个,分别为中医药疗法、中西医结合疗法、针灸疗法、按摩疗法、穴位疗法、气功疗法、中医病机、针灸效应、气功效应、生产和制备。在标引和检索时用副主题词对主题词进行限定,使主题方面更加专指,每个副主题词均有明确的定义和范围,对其允许组配的主题词类目有严格的规定。

1.2.4 出版类型表 收录MeSH词表中出版类型44个,供标引与检索使用。

1.2.5 医学家姓名附表 收录医学家姓名59条,按汉语拼音顺序排列。本表供书本式检索工具书编制索引及书籍主题编目使用,在数据库的标引及检索时做主题姓名标引和检索的参考。

1.2.6 索引表 为帮助用户多途径快速检查主题词,本词表编制了3种主题词索引表,包括汉语拼音索引、汉字笔画索引、英(拉丁)汉对照索引。

## 1.3 词表应用

《中国中医药学主题词表》是中医药文献数据库建设与检索的重要工具,自1984年研制以来,主要应用于中医药文献数据库的建设与检索,采用与MeSH表联合使用的方法,进行中医药学文献的主题标引。主要应用于中国生物医学文献服务系统(SinoMed)、中国中医药期刊文献数据库、万方医学网等数据库的标引与检索。

## 2 中医药学语言系统

由于历史沿革等原因,中医药学术语存在不规范、不统一现象,同物异名、同名异物现象很多,给中医药数字化和信息化提出了不小的难题。在中医药各领域、各类型数据库的发展过程中,亟需将描述同物同类的概念整合、规范,从而使中医药学语言系统建设成为中医药信息化发展的必然基础。2002年,中国中医科学院中医药信息研究所开始研制中医药学语言系统。

### 2.1 研制历程

中医药学语言系统的研制经历了建设初始、迅速发展、国际标准化3个发展阶段。

2.1.1 建设初始阶段 主要是对中医药学语言系统的构建进行方法学研究。方青<sup>[6]</sup>用本体论的方法提出了建立中医药语言系统的方案。尹爱宁等<sup>[7-8]</sup>从中医药学语言系统复杂度的约定方法、多学科交融与碰撞方法及本体论与诠释论关联应用的方法进行研究,为建立中医药学语言系统提供新思路、新方法。

2.1.2 迅速发展阶段 对中医药学语言系统进行大规模的术语数据采集和加工,包括现代中医药术语以及古代中医药术语,崔蒙等<sup>[9]</sup>从语义类型和语义关系两个角度出发,对系统中的低质量数据进行数据清洗,提出基于语义网络的数据清洗策略,并对语义关系进行分析整理,重点增加了中医特有的语义关系,如“开窍于……”“与……相表里”等。朱玲等<sup>[10]</sup>根据古籍特点,从中医古籍分类、古籍概

念定义、与现代系统兼容等方面探讨中医古籍语言系统构建的关键问题与对策。但系统中的术语数据仍存在很多质量问题。

2.1.3 国际标准化阶段 2008年以后,李海燕等<sup>[11]</sup>通过参加国际标准化组织健康信息技术委员会(ISO/TC215)的工作,推进中医药学语言系统语义网络的国际标准化发展,《ISO/TS 17938:2014 中医药学语言系统语义网络框架》国际标准以中医药学语言系统为基础,重新定义了中医药学语言系统的语义类型、语义概念及它们之间的语义关系。该标准不仅规范和支持了中医药学语言系统的建设,还为中医药学术语系统和本体创建提供了语义标准,为中医药学语言系统和统一的医学语言系统的映射提供了支持,对于中医药学术语信息的交换具有重要意义。

## 2.2 主要组成

2.2.1 基础词库 是整个中医药学语言系统的核心组成部分,是在对现存各种主题词表、分类表、数据库、工具书中有关词汇进行分析、选择和组织的基础上产生的一个大型词库。本系统基础词库涵盖可控词表与中医药学科系统及与中医药学科相关联的生物、植物、化工等自然与人文科学专业词汇。基础词库建立在全面、广泛、准确、严谨的基础上,最终形成超级基础词库。词条属性符合中医药学科自然语言与可控词的双重属性。用准确、完整的方式揭示与表达中医药学科及相关术语。这个系统可以相当于统一医学语言系统(UMLS)的超级叙词表和专家词典。

2.2.2 语义网络 是构建语言系统概念词相互关系的权威规则。中医药学语义网络综合各类可控词表的特征,以中医药学科及相关学科的概念为主干,同时对照 UMLS 结构和特点,设计中医药学语言系统网络结构。整体结构符合中医药学结构特点,同时能满足现有中医药信息数字化需求,以中医药学为基础,涉及与其相关联的学科,如医学、生物学、制药工艺等。语义网络是通过语义类型,为出现在中医药学语言系统基础词库中的所有概念提供一种目录组织结构。目前,中医药学语义网络是由 127 种语义类型和 58 种语义关系组成。

## 2.3 应用

中医药学语言系统作为中医药领域收词范围最广的一个术语系统,其应用范围很广。目前该系统主要应用于数据处理、自然语言处理、知识检索、专业词表编制、本体建设等方面。

在数据处理方面主要应用于数据库建设支持。中医药语言系统被用于数据库建设中信息抽取的标准化、规范化、结构化处理。在知识检索方面主要应用于文献检索。我所开发了基于中医药学语言系统的文献检索服务平台,该平台以中医药学语言系统为基础,提供中医药文献检索、术语检索及相关检索等多种功能,将中医药文献进行关联检索及关联性研究,为使用者开扩思路<sup>[12]</sup>。在本体建设方面,朱玲等<sup>[13]</sup>利用中医药学语言系统构建传统针灸知识本体是有益的尝试。基于中医药学语言系统针灸部分术语体系,构建了传统针灸知识本体,揭示了针灸概念之间的隐性关系。

## 3 中医临床术语系统

中医临床术语系统依据中医临床特色,确定了中医临床术语的概念、术语间的语义关系,并参考医学系统命名法-临床术语(SNOMED CT)的结构构建了大型临床术语系统。目前该系统收录概念词 11 万多条,术语 30 万多个,每个概念包括概念词、概念属性、概念定义、相关概念、概念状态等信息。为解决中医临床术语缺乏统一标准,规范化、系统化程度较低,机读效能低下等问题提供了帮助。

### 3.1 构建方法与原则

中医临床术语系统的研制借鉴了国际先进的医学临床术语集 SNOMED CT 的建设模式,并与中医临床、语言学和信息学专家深入研究中医学特有理论体系和临床诊疗思路,采用本体论的方法,构建中医临床术语系统。

系统构建需遵循 4 项原则。(1)中医原则:以中医临床诊疗理论为中医临床标准术语系统研制的指导思想;(2)共识原则:争取得到中医临床医生的普遍认同;(3)质量原则:通过设置多级审查、交叉互审的控制环节,以保证术语的质量;(4)实用原则:强调术语的实用性,以临床用户的实用为术语系统建设的基本原则,为系统最终用户提供最大的便利性。

### 3.2 概念框架

中医临床术语系统的概念框架是参考 SNOMED CT 的结构模式,并根据中医临床诊疗特点构建。2002 年最初构建中医临床术语系统时,采用了 12 个轴的主体框架。经过 10 余年的发展,国内研究者对中医临床术语顶层概念分类的探索也一直在进行,并于 2011 年向国际标准化组织中医药技术委员会(ISO/TC 249)提交了“中医临床术语分类结构”国际标准项目建议。2013 年,根据

ISO/TC249/WG5 专家意见, 名称调整为“中医临床术语系统分类结构”(Categories of TCM Clinical Terminological System)。2014年1月, 该项目通过立项投票。该标准针对中医临床术语提出分类规则, 并确定分类结构, 将中医临床术语顶层概念分为17个大类, 见表1。

表1 中医临床术语系统分类结构17个顶层概念

序号	中文名称	英文名称
1	症状体征	syndrome and sign
2	四诊对象	four examination objects
3	病证	disease and pattern
4	中医操作/方法	TCM operational approach
5	病因病机	cause and mechanism of disease
6	原理和经验	protocol and guideline
7	治则治法	principle and method of treatment
8	中药	Chinese medicinal
9	机体形态	body system
10	分期与传变	staging and transmission/transmutation
11	中医体内物质	internal substance in TCM
12	中医环境和地理定位	environment or geographic location in TCM
13	中医器械和设备	TCM equipment and device
14	中医计量单位和量词	unit of measurement and qualifier value in TCM
15	连接概念	linkage concept
16	医案结构	structure of medical case record
17	短语	phrase

### 3.3 应用

中医药信息化发展涉及信息平台、电子健康档案、远程医疗系统、电子病历等建设, 其核心内容是要实现医疗信息的交流与共享。实现医疗信息系统的互通互用, 需要一系列语义交换标准, 而中医临床术语则是中医医疗信息系统的基础部分。该系统可支持中医临床信息建设、中医临床数据利用。

支持中医临床信息建设, 主要体现在临床信息系统的开发应用, 如电子病历、临床数据数字化、护士工作站等。支持中医临床术语的应用, 主要体现在可支持中医临床数据的数理统计与数据挖掘方面。该系统更加便于对临床数据进行有效的科学研究, 以促进临床科研一体化发展。未来还可应用于卫生管理系统、医疗保险系统等。

### 4 小结

我所在术语规范化研究中做了很多工作, 中医

药三大术语系统是这些工作成果的部分展现。《中国中医药学主题词表》是中医药文献控制和组织的重要方法之一, 它从主题的角度揭示中医药学及其相关学科文献内容, 是中医药文献主题标引和检索的重要手段; 中医药学语言系统则是集合中医药所有术语及其相关学科的大而全的术语系统, 为各术语系统间的术语建立统一的映射标准, 用于支持各个术语系统和中医药信息系统交换; 中医临床术语系统是针对中医临床诊疗特点所设计的, 其涵盖和应用范围主要侧重于临床电子病历规范化和临床诊疗信息标准化。

上述三大术语系统在中医药领域的不同方面都发挥了重要的作用。我们将继续发展和完善三大术语系统, 使之能在中医药信息化的发展过程中起到更好的作用。

### 参考文献

- [1] 李芸. 信息科学和信息技术术语概念体系研究[D]. 北京: 北京语言文化大学, 2003.
- [2] 贾李蓉, 杨硕, 董燕, 等. 中医药学名词术语规范化现状[J]. 中国数字医学, 2012, 7(4): 2-4.
- [3] 吴兰成. 中国中医药学主题词表[M]. 北京: 中医古籍出版社, 1996.
- [4] 温先荣, 张晶, 刘静, 等. 基于文献标引词频统计的《中国中医药学主题词表》选词研究[J]. 中国中医药信息杂志, 2013, 20(10): 16-18.
- [5] 董燕, 刘静, 朱玲, 等. 《中国中医药学主题词表》中医药学及其相关学科类主题词修订选词研究[J]. 中国中医药信息杂志, 2015, 22(7): 40-42.
- [6] 方青. 基于本体论的中医药一体化语言系统[D]. 杭州: 浙江大学, 2004.
- [7] 尹爱宁, 张汝恩. 《中医药一体化语言系统》技术标准[J]. 中国中医药信息杂志, 2003, 10(7): 92-94.
- [8] 尹爱宁, 张汝恩. 建立中医药学一体化语言系统的方法学研究[J]. 中国中医药信息杂志, 2003, 10(9): 90-92.
- [9] 崔蒙, 胡雪琴. 中医药语言系统中的数据清洗策略研究[J]. 中国数字医学, 2010, 5(5): 17-19.
- [10] 朱玲, 尹爱宁, 崔蒙. 中医古籍语言系统构建的关键问题与对策[J]. 中国中医药信息杂志, 2010, 17(4): 98-99.
- [11] 李海燕, 崔蒙, 任冠华, 等. ISO/TC215 传统医学信息标准化工作进展[J]. 国际中医中药杂志, 2011, 33(3): 193-195.
- [12] 贾李蓉, 刘丽红. 基于中医药学语言系统的文献检索服务平台[J]. 医学信息学杂志, 2012, 33(1): 54-56.
- [13] 朱玲, 刘静, 董燕, 等. 基于传统针灸知识本体的语义检索研究[J]. 中国数字医学, 2014, 9(7): 67-69.

(收稿日期: 2015-09-10; 编辑: 魏民)