

DOI: 10.13288/j.11-2166/r.2015.07.006

思路与方法

构建中医临床术语标准真实世界 规范化应用技术体系的思考

郭玉峰¹, 谢 琪², 周霞继¹, 曹馨宇³, 宋观礼¹, 胡延杰⁴, 梁爱景¹, 刘保延^{2*}

(1. 中国中医科学院广安门医院, 北京市西城区广安门内大街北线阁 5 号, 100053; 2. 中国中医科学院; 3. 中国中医科学院中医临床基础医学研究所; 4. 北京安博维科技发展有限公司)

摘要 中医临床术语标准真实世界规范化应用技术体系是践行真实世界中医临床科研范式的重要保障, 也是连接中医标准基础研究与应用研究的桥梁。构建该术语技术体系存在缺乏集成性标准和理想的分类框架、应用模式与方法及综合运维体系不健全等诸多难题, 应在标准性、可用性、实用性、便捷性以及动态可更新性、历史延续性等原则的指导下, 以标准术语集成为核心与基础, 建立核心运维小组与终端用户组并共同参与, 建立基于互联网的术语管理平台, 制定并逐步完善相应工作规范、指南。

关键词 中医临床术语; 真实世界研究; 中医临床科研范式; 中医结构化电子病历

近年来, 我们相继开展了一系列中医临床术语标准化的基础性及应用性研究, 旨在将已有标准充分地、规范化地应用到真实世界信息化临床科研工作中, 一方面有效提升中医结构化电子病历的工作效率, 另一方面保障由电子病历系统采集的临床信息的标准化, 确保能够得到高质量、可分析的数据, 为后期跨地域的数据汇交和基于大数据进行的科研工作奠定基础。在此过程中, 系统总结了标准术语规范化应用的关键科学问题与技术问题, 并就解决这些问题进行了探索, 在实践中提出了“中医临床术语标准真实世界规范化应用技术体系”(简称术语技术体系)的理念, 并结合国家中医临床研究基地建设等实际应用的反馈, 逐步明确了技术体系的内涵, 完善了术语技术体系的顶层框架、组成要素及功能、术语集分类框架、术语应用模式等内容, 使术语技术体系得以稳定、成熟。

1 构建术语技术体系的背景

构建术语技术体系是顺应大数据时代中医学发展的客观需求、促进已有标准基础研究成果实现规范化应用、切实发挥标准对真实世界中医临床科研

工作支撑作用的关键。

1.1 践行真实世界中医临床科研范式的保障

真实世界中医临床科研范式的核心是: 以人为中心, 以数据为导向, 以问题为驱动, 医疗实践与科学计算交替, 从临床中来到临床中去的临床科研一体化的科研范式^[1], 是在系统性复杂科学受到高度关注、数字信息技术得到长足发展、大数据时代来临的大背景下产生的。这一科研范式形成的关键靠“数据”“思维”和“技术”3个主要环节, 其中数据产生和数据利用均有赖于技术的支撑, 所以构建实现这一范式的技术体系, 成为范式能否正常实践的首要任务。中国中医科学院牵头组织了中医医疗、科研、标准化、数理统计、信息技术等相关领域的数十位专家参与, 经过 10 余年协同攻关, 成功开发出“中医临床科研信息共享系统”^[2-3]。作为信息共享系统核心之一的结构化电子病历系统, 是真实世界中医临床科研范式所依赖的临床真实数据产生的源泉, 承担着采集临床信息、完成病历撰写并将信息实时转化为可分析数据的重要功能。其中术语是临床信息的主要载体, 也是数据记录、存储与分析的基础, 高水平地发挥结构化电子病历的功能, 必须有与之配套的、在内容上标准、在应用模式上规范的中医临床术语技术体系作为支撑, 方可有效提升电子病历系统在使用上的便捷性, 保障所采集信息的标准性以及将信息转化为数

基金项目: 国家高技术研究发展计划“863”计划(2012AA02A609); 国家中医药管理局中医药行业科研专项(201207001); 中国中医科学院基本科研业务费自主选题(ZZ0708075)

* 通讯作者: cectcm@gmail.com

据的可靠性,从而为后期海量数据的存储、跨地域的数据汇交和基于特定科研目标所进行的大样本数据挖掘工作奠定基础。因此,构建术语技术体系是践行真实世界中医临床科研范式的基础与保障。

1.2 连接中医标准基础研究与应用研究的桥梁

我国中医药领域近 30 年来实施了大量基础性标准化研究,以术语为核心推出了多部行业标准乃至国家标准,对中医药事业的发展起到了重要的促进作用。同时,伴随信息化时代的到来,电子病历/电子健康记录与电子病历报告表正在全国范围内快速推进普及,纸质记录由于在书写、保存、检索以及后期数据挖掘等方面的先天不足,已逐步退向辅助位置。新的载体形式对已有术语标准从体例、应用模式等方面均提出了新的要求,提高目前以纸质版本形式为主体发布的标准的实用性与易用性,使中医临床术语标准真正在信息化临床科研工作中发挥作用,已成为迫切需要解决的科学问题。现有标准是由国家权威部门、行业主管单位以“标准”形式固定下来的术语标准化基础研究成果,体现了相关领域较高的研究水平,但我们调查发现,现有标准并未能理想地应用到临床科研实际工作中^[4]。上述研究提示,术语标准化的基础性研究成果未能充分与临床实际结合是较为普遍存在的客观事实,已成为阻碍真实世界信息化中医药临床科研工作快速有序发展的瓶颈。我们认为,缺乏与标准配套的规范化应用模式与工具是其中主要原因,构建术语技术体系,通过信息化技术与监控机制来规范调用现有标准化术语,能够促进中医标准基础研究与应用研究的有效结合。

2 构建术语技术体系的思考

构建术语技术体系是在工作实践中不断总结各阶段研究成果,逐步提出、明晰并系统完善的。

2.1 标准术语规范化应用的难点分析

实现标准术语的规范化应用存在诸多难题,其中包括并不局限于下列内容。

2.1.1 缺乏集成性标准 即缺乏在收词范围上全面覆盖中医临床科研所需术语的集成性标准。纵观已有中医药学国家/行业标准,如《中医病证分类与代码》《中医临床诊疗术语》《中医基础理论术语》《腧穴名称与定位》《中医病证诊断疗效标准》等,基于特定应用目的,仅以中医知识体系的局部分支如中医诊断术语、中医治法术语为中心,汇总此类术语完成了定名定义,尚无一部能够从内容上全面有机整合各种类的中医术语、现代西医学术语

以及与临床科研工作相关的非医学专有术语的国家或行业标准。集成性标准的缺失,致使实际工作所需术语分散于不同的标准中,各标准间还存在术语的重叠或缺失。来自不同标准、概念上看似相同又有区别的术语加大了实际应用的难度,而标准术语的缺失则导致实际工作中无词可用。

2.1.2 缺乏分类框架 目前标准中缺乏能够理想展示中医学知识本体,在横向分类与纵向层级设定上纲目清晰、具有较强实用性并得到权威部门认定的术语集分类框架。现有标准在术语分类框架上采用的方案截然不同,形成的框架难以相互兼容,为整合已有标准以支撑临床科研造成阻碍;已有标准化研究大多聚焦于中医诊断等分支,很多在实际工作中迫切需求的术语种类(如症状、体征等)尚未完成标准化处理;现有标准在分类框架构建上大多采用了以学科建制为主的方法,以这种框架为基础形成的标准相对更适宜于教学环境,在临床诊疗工作中的实用性则较低^[5]。

2.1.3 应用模式与方法欠成熟 将已有标准与中医结构化电子病历结合,通过高效、高质量的术语调用机制将标准化术语应用到电子病历书写工作中,在形成符合质量控制要求病历的同时将采集到的信息转化为可分析的数据,以及将已有标准与数据挖掘工作结合,提高数据清洗、数据预处理的工作效率,是术语标准在信息化临床科研工作中规范化应用的两大主题。实现上述两种应用的模式与方法以及相应的工作规范与指南,在世界范围相同研究领域均处于探索阶段,并无成熟的经验可以借鉴,需结合实际情况进行创新性研究。

2.1.4 综合运维体系不健全 受到标准制订、修订周期的制约,产生了已有标准与医学快速发展、新生术语涌现、旧术语在概念上发生转换等矛盾,大量临床一线迫切需求的术语无法及时反馈并纳入到标准术语体系中,标准修订调整后的术语也很难及时反馈应用到临床一线。而标准术语,唯有经过临床一线用户结合实际需求并在统一指南下进行规范的本地化再整合,才能更好地发挥对临床科研的支持,故从术语标准集成的研究团队到标准术语应用的各终端单位,均应建立相应人员组织、工作规范,并借助信息化术语管理平台,共同参与到标准术语的基础与应用层面维护工作中。

结合对上述难点的分析与探讨,我们提出了构建术语技术体系的理念。术语技术体系是一个系统化工具,是在完成中医学标准化术语集成的基础上,通过特定的模式与方法将这些标准术语有效而

规范化地应用到信息化临床科研工作中,从而实现结构化电子病历书写以及数据挖掘工作的支撑。术语技术体系应由实体术语集、相关的规范/指南以及基于互联网络的术语管理平台组成。

2.2 术语技术体系的构建原则

构建术语技术体系应遵循标准性、可用性、实用性、便捷性,以及在保持术语体系整体稳定下的动态可更新性、历史延续性等原则。

标准性: 标准性是术语技术体系的核心特质,确保收录术语的标准性是其关键。需建立符合中医学知识本体的术语分类框架,并将已有的标准术语依据分类充实纳入到技术体系中。在采集过程中应按照国家标准 > 行业标准 > 字/词典 > 教材的规则,对已有标准中定义清晰且各标准间不存在概念上重复及差异的术语可直接导入,对于术语在不同标准中具有表述及内涵差异以及标准中缺乏临床迫切所需术语等情况,均需列出详单召集领域专家论证确定后方可纳入体系。

可用性: 术语技术体系应能最大限度地满足实际工作对标准化术语的需求,因此,应全面系统地收录中医术语、现代医学术语以及非医学专有术语,常用、次常用甚至罕用术语都能被方便地检索使用,凭借技术体系内收录的术语即能为临床科研工作提供所需的绝大部分标准化术语。可用性强,术语技术体系才具有存在的价值,故应从扩展标准术语集成的收词容量上确保其可用性。

实用性: 证-治-效等信息是临床与科研工作关注的重点,它们既是病历书写中的主体内容,也是数据挖掘、疗效评价工作所需的关键数据,整合并有针对性地提供描述这些种类信息的术语,能提高术语技术体系的实用性。需在共性的中医学理、法、方、药、术等术语分类基础上,分析临床科研所需的各种关键信息种类,据此建立临床实用性术语分类框架并收录标准术语,为结构化电子病历中需要采集关键信息的段落结构提供术语支持。

便捷性: 对于特定病种下的某种具体疾病的诊断,在临床病历书写与数据挖掘工作中所涉及的术语总量是相对有限的,由收词动辄数十万的术语集为电子病历系统提供标准化术语,会显著降低工作效率。因此,由临床一线专业人员以疾病诊断为中心进行回顾性梳理,筛选出该单位在具体疾病诊断下最为常用的标准术语,建立用于直接支撑电子病历的专病术语集,能够有效精简术语字典的容量、提升便捷性。

稳定与动态可更新性: 术语技术体系中各术语

集的构成、框架、内容、功能均应保持相对稳定,因此,在构建之初应做好顶层设计,并确定框架、术语维护修订的原则与流程。技术体系中的术语,应能够动态更新并及时发布到临床一线;从临床一线工作中捕获的、临床切实所需的新术语,也应能够通过术语技术体系内的标准化认证机制反馈并收录到相应术语集中。

历史延续性: 为保障术语更新后利用原有术语进行数据存储与病历编码的资料的可溯源性,术语技术体系应具有历史延续性,可借助术语编码、术语的状态参数等方式来实现。

此外,术语技术体系由多方人员共同参与构建及维护,其中基础标准术语集成的运维专业化程度高,应由专业团队主持完成;用于支持电子病历的应用性术语集,其构建与使用由终端用户为主体进行,需制定工作指南并加强培训,确保全部参与人员均能够按照统一规范的模式进行术语整理和使用。

2.3 术语技术体系的构建思路与实践

术语技术体系的构建思路是:以标准术语集成(包括实现不同目标的多个实体术语集)为核心与基础,建立核心运维小组与终端用户组并共同参与,建立基于互联网络的术语管理平台,制定并逐步完善相应工作规范、指南。其具体结构与功能见图 1。

2.3.1 建立中医临床标准术语集成

中医临床术语标准集成是术语技术体系的核心与基础,但尚无从中医学本体论角度进行分类框架设计、在容量上能够涵盖全部临床科研工作所需术语的标准。研究实践证明,仅靠一部术语集难以在应用中求得可用性与便捷性的平衡,既要拥有足够收词量以保证临床有术语可用,又要有针对性的汇总一部分最为常用的术语提高临床病历书写与数据挖掘的效率,需要建立具有相同的标准化术语来源、用于实现不同的功能目标、在术语分类框架上各有针对性、在收词量上有所区别的 3 种术语集,这些术语集分属于基础与应用层次,内容同源、相互映射、各司其职,共同完成术语集成工作。

建立基础的《中医临床术语集》,用于保障术语技术体系的标准性与可用性。一方面,《中医临床术语集》收录的术语原则上应来自于已有国际标准、国家标准、行业标准,对已有标准中缺乏但临床上切实需要的术语,应从权威典籍、教材、字典、辞典等正式出版、有据可查的资料中筛选整理,术语在正式纳入前还应遵循相应的规则进行专家咨询论证并保存记录,以确保术语的标准性。另

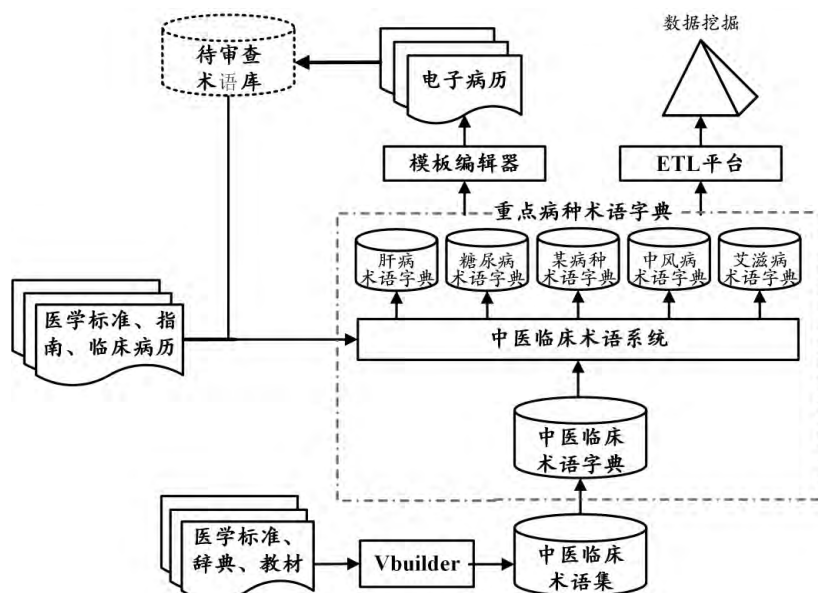


图1 中医临床术语标准真实世界规范的应用技术体系结构与功能示意图

一方面,《中医临床术语集》的分类框架应能够正确地展示中医学临床诊疗本体,在术语分类与术语总量上能够全面覆盖传统中医学所需,并能够与相关现代医学标准术语集接轨,共享现代医学专有术语以及非医学专有术语,以保证有词可用。我们借鉴国际上《临床医学标准术语集》(Systematized Nomenclature of Medicine, Clinical Terms, SNOMED CT)的经验^[6-9],开发具有“原理、经验、临床发现、操作处理”等十二个顶级轴(Hierarchy)分类框架的《中医临床术语集》,且易于同SNOMED CT接轨^[9]。

建立基础应用结合的术语集《中医临床术语字典》,用于保障应用技术体系的实用性。《中医临床术语字典》根据对临床科研工作关注的证-治-效及其时空属性等重要信息节点的分析,并根据对当前中医结构化电子病历中需要重点进行结构化处理的病历段落的逻辑描述结构分析,建立以“症状、体征、其他异常临床所见、中药等27种主体分类以及程度、因素等10种属性分类组成的37轴术语集分类框架,以分类框架为基础全面继承《中医临床术语集》中的标准化术语并加以补充。

建立应用层次的术语集——《专病种临床术语字典》,用于直接支撑电子病历结构化书写,保障病历书写与数据挖掘工作术语调用的便捷性。《专病种临床术语字典》具有与《中医临床术语字典》完全相同的37轴分类框架,不同之处在于《专病种临床术语字典》以病种、诊断的实际需求出发,各分类轴的收词量显著小于《中医临床术语字典》。通过与临床一线人员的反复交流,我们

发现,特定病种用于完成描述证-治-效等主体信息的术语在数量上是相对有限的,如描述病种A常见症状的术语在病种B中的使用几率较为罕见,而描述病种B常用中药的术语在病种A中也较为罕用,证实可由专病种临床工作者为主,以特定疾病诊断为中心,从《中医临床术语字典》中梳理本病种、具体疾病诊断下,每种术语分类(如症状、体征、中药)中最为常用的临床术语,筛选建立《专病种临床术语字典》并装载于电子病历系统中,可有效缩减术语字典容量、提高工作效率。

需要注意的是,《专病种临床术语字典》应装载到电子病历系统中,直接用来支持中医电子病历系统的结构化模板定制与术语调用。由临床一线人员按照模板制作规则,以本病种在临床诊疗中需要重点收集的关键信息为主体,利用电子病历系统中的模板编辑器,从《专病种临床术语字典》中选取标准化术语制作病历模板系统中的基本元素并为其赋值,实现基本元素的命名与赋值的术语标准化,从而最终完成中医结构化电子病历系统对标准化术语的调用。

2.3.2 建立基于互联网的术语管理平台 开发基于互联网的以《中医临床术语字典》为核心的术语管理平台,借助此网络化的术语管理平台,由核心运维小组负责从《中医临床术语集》平台中纳入标准术语,构建维护《中医临床术语字典》,以及增删术语的认证、审核、动态更新。由分布于全国各地的专病种人员结合各自需求,利用《中医临床术语字典》进行检索、查询、加工形成《专病种临床术语字典》,依据权限进行术语导出并应

用到中医结构化电子病历系统。从临床实际工作中采集到的新术语,还可以借助平台通过专家论证审核的方式完成标准化处理,扩充到标准化术语源中并为用户共享。

2.3.3 制定规范与操作指南 针对实体术语集框架与术语的制修订、术语管理平台的操作、术语发布与动态更新、病历模板基本元素制作等工作环节,应制定详细的规范与操作指南并在实践中调整、丰富、完善,以确保术语技术体系的良性运转和可持续性发展。

2.3.4 建立中医术语标准规范化应用研究团队 由专业人员建立核心运维小组,负责主持基础术语集的运行和维护;各终端用户单位建立由临床一线人员、信息化工作人员组成的术语规范化应用小组,核心组编制教材并组织统一培训,实现两组间常规化的交流互动,形成完整的术语应用研究团队。

3 小结

实现临床医学术语标准在电子病历系统中的规范化应用,是当前世界范围的研究热点。具有 50 年研发历程的 SNOMED CT 虽已得到多个国家的认可,但真正实现其在电子病历系统中全面应用的成功案例并不多见^[9]。我们提出的中医临床术语标准真实世界规范化应用技术体系经过 10 余年的探索,已经定型并趋向成熟,为术语标准化基础研究与应用研究建立起了有效桥梁。近年来在国家中医临床研究基地建设中的推广应用,不仅使术语技术体系接受了实际工作的检验,更收获了来自基础、临床、科研、管理等多方面的经验。在今后实践中,

将结合实际应用环节的反馈意见,着重完善术语技术体系的顶层框架及配套规范,使其更好地为大数据时代的真实世界信息化中医临床科研工作服务。

参考文献

- [1] 刘保延. 真实世界的中医临床科研范式 [J]. 中医杂志, 2013, 54(6): 451-455.
- [2] 刘保延, 周雪忠, 李平, 等. 个体诊疗临床信息一体化平台 [J]. 中国数字医学, 2007, 2(6): 31-36.
- [3] Liu B, Zhou X, Wang Y, et al. Data processing and analysis in real-world traditional Chinese medicine clinical data: challenges and approaches [J]. Stat Med, 2012, 31(7): 653-660.
- [4] 郭玉峰, 尹爱宁, 刘保延, 等. 浅谈中医临床术语标准化工作现状及其深化推进 [J]. 中国中医药信息杂志, 2009, 16(11): 3-4.
- [5] 郭玉峰, 刘保延, 周雪忠. 面向中医临床科研需求的术语分类框架研究 [J]. 环球中医药, 2008, 1(2): 9-12.
- [6] 郭玉峰, 刘保延, 崔蒙, 等. 借鉴 SNOMED CT 发展中医临床标准术语集 [J]. 中医药发展与人类健康: 庆祝中国中医研究院成立 50 周年论文集. 北京: 中国中医研究院, 2005: 425-433.
- [7] 郭玉峰, 刘保延, 周雪忠. SNOMED CT2007 的顶级概念分类详解 [J]. 中华中医药学刊, 2008, 26(9): 1928-1932.
- [8] 郭玉峰, 刘保延, 周雪忠. SNOMED CT 的语义关系与连接概念 [J]. 中华中医药学刊, 2008, 26(10): 2206-2209.
- [9] 郭玉峰, 刘保延, 周雪忠. SNOMED CT 2007 的连接词详解 [J]. 中华中医药学刊, 2008, 26(11): 2416-2423.
- [10] Giannangelo K, Fenton SH. SNOMED CT Survey: An Assessment of Implementation in EMR/EHR Applications [J]. Perspect Health Inf Manag, 2008, 5(7): 1-13.

Reflections on the Construction of Standardized Application Technology System of Standard Clinical Terms in Real World of Traditional Chinese Medicine

GUO Yufeng¹, XIE Qi², ZHOU Xiaji¹, CAO Xinyu³, SONG Guanli¹, HU Yanjie⁴, LIANG Aijing¹, LIU Baoyan²
(1. Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, 100053; 2. China Academy of Chinese Medical Sciences, 3. Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences; 4. Beijing Anbowei Technological Development Co., Ltd.)

ABSTRACT Standardized application technology system of standard clinical terms in real world of traditional Chinese medicine (TCM) is an important guarantee to practice the real world of TCM clinical research paradigm, and is also the bridge connecting TCM standard basic research and applied research. The construction of the term technology system may confront many problems such as lack of the integration of the standard, ideal classification framework, plus imperfect comprehensive operation system and so on. Therefore, under the guidance of the principles such as standardization, availability, practicability, convenience, dynamic update, historical continuity, we should take standard terminology integration as the core and the foundation, establish core operation group and the terminal user group and joint participation, devise terminology management platform based on the Internet, and develop and gradually improve relevant working regulations and guidelines.

Keywords clinical terms of traditional Chinese medicine; real world study; paradigm of clinical research of traditional Chinese medicine; structured electronic medical records of traditional Chinese medicine

(收稿日期: 2014-07-10; 修回日期: 2014-12-01)

编辑: 崔京艳