

DOI: 10.13288/j.11-2166/r.2016.11.006

思路与方法

中医临床常用术语的本体构建技术探讨

谢琪¹, 刘亮亮^{2*}, 曹馨宇³, 郭玉峰⁴, 刘保延¹, 周洪伟³

(1. 中国中医科学院, 北京市东城区东直门内南小街 16 号, 100700; 2. 上海对外经贸大学商务信息学院; 3. 中国中医科学院中医临床基础医学研究所; 4. 中国中医科学院广安门医院)

摘要 临床数据是开展临床研究不可或缺的资源, 中医临床常用术语是临床数据的载体, 其完整性与规范性是获得高质量临床数据的重要前提。分析中医临床常用术语质量的基本要求及中医临床常用术语在完整性、规范性、概念关系等方面的常见问题, 提出利用本体技术建立中医临床常用术语本体, 并总结中医临床常用术语本体构建的关键技术环节, 包括构建临床常用术语应用技术体系、建立中医临床术语标准规范、获取中医临床常用术语、建立中医临床常用术语本体关系、进行本体验证与可视化。从本体的角度对中医临床常用术语进行概念识别与描述, 建立概念之间的关联关系, 可消除中医概念及其关联关系的不明确性。

关键词 临床术语规范化; 临床数据; 概念关系

注重临床实践是中医学的显著特征, 中医临床诊疗过程所形成的数据是临床研究的基础, 掌握大数据的技术和方法、充分利用海量临床数据是现代医学研究发展的必然趋势。近年来, 中医临床研究领域注重单病例系统评价、观察性研究、效果比较研究、真实世界临床研究、临床病例注册登记等多种利用医疗卫生与健康数据的研究, 都更直接、更紧密地以临床实践为基础进行数据采集, 从整体、动态和个体化层次进行健康状态与诊治规律的研究, 更有研究者提出中医临床科研的发展就是以数据为导向、以问题为驱动的医学实践与科学计算交替的过程^[1]。中医临床常用术语是临床数据的一种表现形式, 是中医数据的一种基础资源, 对电子病历的结构化、中医数据挖掘等中医信息化起到决定性的作用。

中医临床常用术语是临床数据的具体表现形式, 需适应临床数据采集所要解决的多样化的需求。同时医学数据是富知识型的, 不结合领域知识的数据分析是不可靠的。数据中的概念不可能是孤立的而是需要语义关系进行概念的界定和鉴别, 大

规模语义关系是术语本体的主要内涵, 是领域知识的主要体现, 数据挖掘需要结合知识推理形成创新的知识发现结果, 因此, 可靠和创新的数据挖掘发现更依赖于高质量的具有概念化定义的医学数据。目前中医临床常用术语的完整性不足, 没有形成高质量的语义关系, 极大地影响了有意义的数据挖掘分析结果的产生。

1 中医临床常用术语质量基本要求

中医临床数据的质量应以完整性、规范性、真实性、准确性进行衡量, 其中真实性和准确性是对数据质量的基本诉求。中医临床数据是医生在临床诊疗过程中所采集的与医疗活动相关的患者健康信息, 以及医生对健康与疾病状态做出的诊断及采取的诊疗措施, 是医生诊疗工作的客观记录。我国卫生主管部门建立了全面的医疗管理制度, 通过规范与监督医疗行为, 从临床诊疗数据的时效性、可溯源等方面保证了数据的真实性和准确性^[2-3], 其完整性及规范性具体要求如下。

1.1 中医临床常用术语的完整性

中医临床数据的完整性包含两个层面。一方面, 解决临床研究问题所需要的各项数据都具备, 即实体完整性 (entity integrity), 由数据模式中的隐含约束所确定。中医临床研究问题可概括为 5 个方面: 1) 发病与临床特征: 揭示疾病的危险因素

基金项目: 国家“十二五”科技支撑计划 (2013BAI02B10); 国家自然科学基金 (81403281); 国家中医药管理局中医药行业科研专项 (201407001-13)

* 通讯作者: lingyun79626@126.com, 15952904020

与病因,分析疾病与证候的临床症状特征,调查疾病的流行病学特征;2) 疾病诊断与证候分型:建立疾病诊断模型,探索、建立中医证候分类识别、诊断预测模型,揭示证候的生物学基础;3) 疗效评价研究:探索疗效影响因素,寻找疾病治疗时间窗,治疗方案的经济性评价;4) 预后与转归研究:基于临床观察,分析研究相关影响因素,进行继发病/并发症发生危险的预测分析;5) 药物研究:分析中药及临床用药规律,进行药物临床应用特征和安全性评价。从其内在约束性分析,中医辨证论治是以患者状态调整为导向,“证-治-效”紧密结合的个体化、动态化的复杂干预过程^[4]。基于中医诊疗模型的研究所要求的核心数据必然包括临床表现、证候诊断、中医临床疗法(如图1),进而可以确定中医临床数据的完整结构(如图2)。

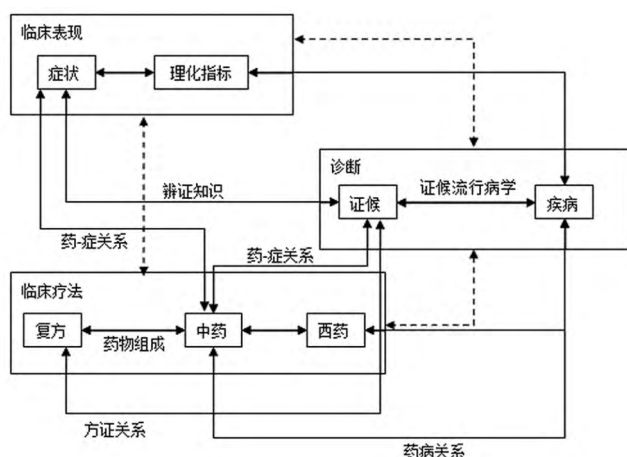


图1 中医临床诊疗模型核心数据构成图

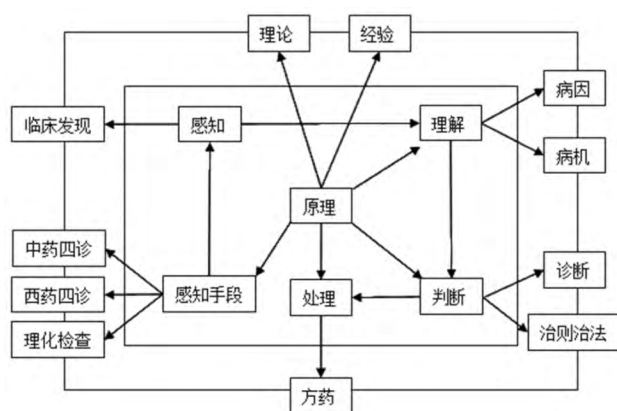


图2 基于中医临床核心数据的术语模型图

另一方面,中医临床数据的内容或赋值必须满足某种特定的数据类型或约束,即域完整性(domain integrity),依赖于数据语义与应用的显式约束。《国际疾病分类》《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》《中医药学名词》等相关中医药学术语标准的制订和颁布为中医临床数据规范确定了标准。通常前或后结构化数据采集系统,如结构化电子病历系统、临床病例注册登记系统、临床科研数据管理系统,嵌入式地应用标准术语集,通过限定性的数据采集来满足临床数据赋值的域完整性,进而实现临床数据的规范化。中医临床常用术语表达多样,如证候“痰热”有痰热内蕴、痰热蕴结、痰热壅盛、痰热互结、痰热互阻、痰热内生等众多表达,故诊断性实际数据中的临床常用术语大大超过了标准术语的规模。遵循临床医生的辨证论治思维特征,并能够在医疗实践中高效、便捷地记录下来的诊疗数据,蕴含在丰富的临床常用术语中,难以进行数据整理和分析,使临床数据缺乏完整性。

1.2 中医临床常用术语的规范性

数据的规范性要求中医临床常用术语具有逻辑的一致性、正确性、有效性和相容性。临床医生作为医疗活动的主体,以患者作为研究对象的客体,疾病随时间发展将呈现不同的临床表现和规律,患者表现出外在的临床表征信息,医生(主体)通过四诊和医学检查获取到患者(客体)信息-分析处理-赋予其“证”的特定中医诊断-形成治疗策略,在这一过程中所产生的中医临床常用术语表达了客观事物在人脑中反映的抽象概念。语义可以简单地看作是现实世界中的事物所代表的概念的含义以及这些含义之间的关系,是数据在某个领域上的解释和逻辑表示。因此,从语义层次上理解概念的含义及含义之间的关系,分析中医临床常用术语的特点和结构,将促进中医临床数据的共享和互操作性,体现中医临床数据的应用价值。

1.3 中医临床常用术语概念关系

中医临床常用术语兼具临床实践的科学性、具体性和语义表达上的哲学抽象性、宽泛性。中医临床常用术语包含中医内涵特征,需要以中医药基础理论为指导,故中医临床常用术语概念关系具有以下4个特征:1) 注重整体观念指导下脏腑功能的相互作用。整体观念是中医学的基本特点,以此为指导,各脏腑的功能活动都是整体活动的组成,生理功能上相互制约、相互依存、相互为用,以经络为联系通道,气血津液输布周身,形成协调统一的整体,中医的概念关系更注意相互作用的状态,如营卫不和、水火既济。2) 注重功能关系的表达,而物理、时间、空间概念关系较少。中医学涉及形

态结构的藏象、经络的描述同样密切地结合了生理与病理功能,如“脾阳根于肾阳”,生理上是后天与先天的关系,相互资助、相互促进,在病理上亦常互为因果、相互影响。3) 中医学的概念关系相对稳定,表达形式具有文学色彩。《黄帝内经》《伤寒杂病论》等古汉语承载着概念和概念关系,也因此具有相对稳定的结构,而相对于现代医学从相对严密的逻辑概念出发的语言描述不同,中医学概念和概念关系具有文学、艺术性的语言,应用拟人、隐喻、类比、象征的表达形式,如“肝喜条达”的“喜”及“心恶热”的“恶”。4) 具有中医特色的概念关系。如在脏、在体、在音、在声、在变动等,基于五行相同属性事物之间的类比是概念相关与功能相关的复合,难以用其任何一个下位关系表述说明,具有整体性、方向性和层次性,是中医概念关系的独特价值。

体现中医理论及临床实践经验的中医临床常用术语,需要以中医学理论为指导,利用计算机技术进行解读,从语义层进行概念描述与知识表达,使中医临床常用术语所蕴含的丰富知识能够得到高效、准确地采集、存储与利用。本体是概念模型的明确的规范说明^[4],因此,从本体的角度对中医临床常用术语进行概念识别与描述,建立概念之间的关联关系,消除中医概念及其关联关系的不明确性,是中医药领域迫切需要解决的关键问题。

2 中医临床常用术语本体构建的关键环节

利用本体技术建立中医临床常用术语本体能明确概念之间的语义关系,从而消除中医概念的不确定性,确保中医常用术语的完整性、规范性、真实性、准确性。构建中医临床常用术语本体的关键环节包括以下 5 个部分。

2.1 构建临床常用术语应用技术体系

中医临床术语应用体系由中医临床术语集、中医临床术语字典、临床诊疗模板及相关的技术与管理规范构成,用以保障临床研究数据的可用性。其中,中医临床术语集在中医基础理论的指导下借鉴医学系统命名法-临床术语(SNOMEDCT)的方法,发展了原理与结构,从本体论的角度分析中医辨证论治的诊疗过程,构成了中医临床术语分类框架,整合了现有的中医学及医学领域的标准术语;中医临床术语字典面向临床诊疗数据采集以及数据挖掘工作,重构中医临床术语的分类框架,建立与中医临床术语集框架的映射关系,以疾病为中心,用人

工方式遴选常用的、更丰富的临床术语;临床诊疗模板实现了中医电子病历文档的结构化,建立起与中医临床术语字典的连接,通过规范化术语的应用支持数据挖掘分析,全面、准确地展现了中医临床诊疗知识本体,能够辅助医生遵循中医诊疗思路进行电子病历书写,促进具有中医内涵的诊疗信息获取。目前为支持中医临床术语字典,全国 25 家中医临床科研单位 220 位临床科研人员历时 2 年,通过人工方式整理了中医临床常用术语 188 629 条,用其支持创建了临床诊疗电子病历模板 6200 种(国家中医药管理局中医药行业科研专项“全国中医医疗与临床科研信息共享的关键技术及示范应用研究”研究组 2014 年度总结报告数据)。

2.2 建立中医临床术语标准规范

目前中医临床常用术语规范化研究仍采用人工方法,技术问题突出表现在以下 5 个方面: 1) 术语的人工规范化研究的效率极低,如乙型肝炎相关性肝衰竭需要广泛选取关于研究病种 1949 - 2012 年出版的所有中医内科教材与权威性著述、近 10 年国内核心期刊公开发表的中医文献、近 1 年的临床研究病历,人工阅读量极大^[5];此外,上述全国 25 家中医临床科研单位开展的大规模中医临床常用术语研究虽初具规模,但离满足应用需求还有较远的距离。2) 标准术语之间由于缺少概念关系而没有互操作性,如关于症状,《中医诊断学》记载症状信息 1373 条,《中医症状鉴别诊断学》记载 623 条,《中医药学名词》记载了 456 条^[6],相互之间没有联系。3) 术语的规范化研究本身缺乏技术规范,临床常用术语规范化研究的主要步骤包括选取并确定术语来源和范围→确定术语的分类框架→遵循科技术语的命名原则即单义性、简明性、约定性、系统性、国际性,以《标准化法》《确立术语的一般原则与方法》《中医药学名词审定原则与方法》等为依据进行规范化整理加工→专家论证与修订,而规范化的术语加工标准主要由主观评估所确定,即便通过专业培训也难以完全统一,使之缺少质量控制而造成质量差异。4) 中医临床常用术语规范化为应用基础研究,缺少有效的技术方法而主要由临床医生进行术语的收集与整理,与临床医生的主要职责有错位。5) 缺乏对中医临床常用术语内部概念关系的梳理,临床常用术语与标准术语难以建立映射关系,虽然实现了数据的多样性但遗留了数据规范性的问题。

经过多年研究,中医药领域已经形成了中医临

床术语的相关标准规范,尤其是中医临床术语集,其借鉴了代表国际先进水平的临床医学术语集 SNOMEDCT 的模式并结合国内外标准化术语集构建思路和方法,以中医自身独特理论为核心,组织了国内 10 多家参研单位近百名专家,以现有的中医药国家标准、行业标准以及各种辞典、教材等为术语源,采用成熟的本体论工程和术语知识库构建方法,历时近 10 年研制完成并动态更新。中医临床术语集包括分类框架、概念、术语、编码、属性(时间、程度)、调用方式、状态、语义关联等内容,收录概念 20 余万条,实现了对已有中医标准的整合,可作为中医临床术语规范化应用体系的术语来源。需要注意的是,临床常用术语规范化、规范化术语的概念化、概念化术语的临床应用、从临床应用过程进行常用术语的补充,是实现临床常用术语规范化应用的不同阶段,特别是临床常用术语的补充阶段,能够对整体工作起到验证和完善的作用,应重视并加强。

2.3 获取中医临床常用术语

中医临床术语的自动抽取有以下 3 个局限性:

1) 目前自动抽取的方法主要基于统计的方法,而基于统计的方法则需要大规模人工标注的训练语料;2) 由于中医术语的多样性、历史性,其不但包含古代汉语,而且包含了大量的现代汉语的表达法,成为中医术语自动抽取的一个难点;3) 目前中医术语的识别都是基于经典中医著作中术语的识别,不是从临床应用的语料中进行识别,而大量新的术语都出现在临床应用中,因此,需要从临床病历语料库中获取临床常用术语。

随着中医信息化的发展和大数据时代的到来,在临床病历中出现了大量新的术语,医生在临床病历的撰写中使用大量临床常用术语,通过传统的人工方法来收集术语显然已经无法满足中医学发展的需要,因此,建议基于已有的中医临床术语集,利用规则和统计的方法从临床病历语料库中识别新的术语从而实现中医临床术语的自动获取。主要包括以下 4 个方面:1) 收集临床病历数据,建立临床病历语料库。因临床病历一般都是非结构化的自由文本,所以需要对其进行数据处理,包括数据清洗、分词、标注等。2) 对中医临床标准术语集中已有术语进行特征分析,学习中医临床常用术语的表达模式和构词规则。3) 对病历语料进行统计,建立统计模型,利用规则和统计相结合的方法来识别临床病历语料中的新术语。4) 对新识别的术语

进行自动验证。

2.4 建立中医临床常用术语本体关系

术语间的语义关系存在于抽象的类型或具体的概念之间,用来表达某个领域中概念之间的重要关系,完整的语义关系可以反映该领域的知识结构^[8],具体见图 3。目前各个语言系统的语义关系均以同时期其他系统语义关系为基础,根据自身研发需要,人工添加新的语义关系。中医临床术语集给出术语的 10 种属性关系,即程度、因素、特征、频次、时间、量的变化、性质、分型和分期、部位、阴性判断,主要停留在术语所表达的概念自身特性上,在语义关系体系中为“属性关系”(图 3),而在中医学领域中存在大量两个术语之间的关系,如“心合小肠”“心气通于舌”“风热犯肺”“气能生津”“寒极生热”等。

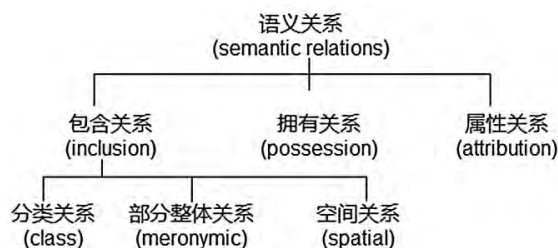


图 3 中医临床常用术语语义关系分类框架图

如何从一定规模的语料中抽取得出反映某一领域或者某一学科的两两词语间的概念关系即术语关系尤为重要。国内外术语关系的获取技术大体有 3 种^[9]:1) 基于人工获取方法。采用专家参与的方式,对术语进行标注,从语法、句法等所需方面进行术语关系抽取。该方法人工工作量大、耗时相对较长、效率相对较低^[10-12],优点是获取的关系可靠。2) 紧密地将计算机语言与术语相结合,基于模板匹配的方法,通过已知文本等建立一个共性的模板,通过该模板进行语料库的检索匹配,以获得大量的进一步信息,经反验证后得到最终的术语关系。该方法对模板表示的依赖性非常强,并且要求模板是无歧义的,抽取效果明显受到模板规模和质量的制约^[13]。3) 基于机器学习方法,将关系抽取转化为分类问题,通过构造候选关系,利用机器学习得到的分类器来标注这些候选关系属于哪一类预定义关系^[14]。机器学习依赖于一些属性和变量,这些属性和变量用来描述学习模型中的实体,来自语境信息语句或语句部分的实体类型、词、短语、词性、语串及标签等^[15]。

在以上 3 种不同术语关系的提取方法的基础

上,进行中医临床常用术语的关系抽取方法如下:第一步,确定中医临床常用术语关系抽取的语料来源;第二步,提取中医临床常用术语概念的关系词;第三步,进行中医临床常用术语概念关系词的语义分析;第四步,确定中医临床常用术语关系,对所得到的术语关系进行审核,建立中医临床常用术语本体关系。

2.5 进行本体验证与可视化

随着本体研究的发展,本体构建与本体验证工具也大量出现,完整的本体构建工具能够完成对本体的创建、解析、存储和验证等工作,但是目前还没有支持中医本体构建和验证的工具和系统,而且当前本体构建工具在对中文进行可视化、推理、验证等方面存在很多问题。可视化是反映本体的直观简洁的方式,可以通过图形展示的方法清晰地反映中医本体中概念的层次关系、属性关系等内容,并且可以直观、简便地进行正确性验证等检查工作。因此,本体验证与可视化也是目前中医本体研究的一个关键因素。

3 结论与展望

通过建立中医临床常用术语本体,建立概念之间的关联关系,消除中医概念及其关联关系的不明确性,是中医药领域迫切需要解决的关键问题。中医临床常用术语的自动获取以及术语关系的自动获取又是中医临床常用术语建立的关键环节。因此,在中医临床术语集的基础上,利用中医临床常用术语的构词特点,利用人工智能中自然语言处理的方法和模型以及数据挖掘的方法,识别临床数据中新的术语,从而实现中医临床标准术语集的扩展,同时建立和获取临床常用术语之间的语义关系,以支持中医临床病历模板的构建等中医信息化建设。具体应包括以下 4 个方面: 1) 从中医临床术语集出发,通过对已有术语的特征分析,自动生成中医术语的表达模式和构词规则、启发式规则等,利用统计和规则相结合的模型实现中医临床常用术语的识别; 2) 从中医文本的研究角度出发,分析临床术语之间的语义关系,建立和获取临床术语之间的关系; 3) 利用本体论的方法形成中医临床常用术语知识库,以支持中医临床病历模板的构建、中医电子病历的书写和数据分析等中医信息化建设,为中医诊疗信息化平台如临床病历模板的构建等提供基础性

的支持; 4) 建立术语自动识别与关系获取平台,为中医研究和中医信息化建设提供方法和工具。

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] 邵明义, 魏明, 段俊国, 等. 电子化数据管理条件下的临床试验质量控制模式 [J]. 中国新药与临床杂志, 2010, 13(9): 1-2.
- [3] 邵明义, 姜小帆, 段俊国, 等. 临床研究数据获取过程中动态质控方法 [J]. 中药药理与临床, 2009, 13(9): 1-2.
- [4] 刘保延, 王永炎. 证候、证、症的概念及其关系的研究 [J]. 中医杂志, 2007, 48(4): 293-296, 298.
- [5] GRUBER TR. A translation approach to portable ontology specifications [J]. Knowledge Acquisition, 1993, 5(2): 199-220.
- [6] 张茜茜, 张涛, 陈斌, 等. 肝衰竭黄疸症中医临床术语集建立方法探讨 [J]. 世界科学技术-中医药现代化, 2013, 15(5): 843-846.
- [7] 张启明. 中医症状学研究 [M]. 北京: 中医古籍出版社, 2012: 2.
- [8] WINSTON ME, CHAFFIN R, HERRMANN D. A taxonomy of part-whole relations [J]. Cogn Sci, 1987, 11(4): 417-444.
- [9] 孙霞, 王小凤, 董乐红, 等. 术语关系自动抽取方法研究 [J]. 计算机科学, 2010, 37(2): 189-191, 215.
- [10] 谢琪. 基于本体方法构建中医药概念信息模型的方法学示范研究 [J]. 北京: 中国中医科学院, 2011: 61-63.
- [11] APPELT DE. Introduction to information extraction [J]. AI Commun, 1999, 12(3): 161-172.
- [12] AONE C, RAMOS M, REES S. A large-scale relation and event extraction system [C] // Proceedings of the 6th Applied Natural Language Processing Conference. New York: Acm Press, 2000: 76-83.
- [13] YU H, HATZIVASSILOPOULOS V, FRIEDMAN C, et al. Automatic extraction of gene and protein synonyms from MEDLINE and journal articles [J]. Proc AMIA Symp, 2002: 919-923. PMID: 12463959.
- [14] 刘克彬, 李芳, 刘磊, 等. 基于核函数中文关系自动抽取系统的实现 [J]. 计算机研究与发展, 2007, 44(8): 1406-1411.
- [15] MIAO C, LIU XZ, QIN J. 从社会性标签中进行语义关系抽取: 一种元数据生成方法 [J]. 现代图书情报技术, 2009(3): 38-45.

(下转第 937 页)

Clinical Observation on *Tongdu Ningshen Decoction* (通督宁神方) Treating 48 Cases of Insomnia of Governor Channel Discord Type

MA Jie¹, YAN Hua², XU Yan², YAN Shiyun³

(1. Department of Hospital Administration, Shanghai University of Chinese Medicine, Shanghai, 201203; 2. Shuguang Hospital, Shanghai University of Chinese Medicine; 3. Shanghai University of Chinese Medicine)

ABSTRACT Objective To observe the clinical effect of *Tongdu Ningshen Decoction* (通督宁神方) treating insomnia of governor channel discord type. **Methods** Patients with insomnia of governor channel discord type were randomized into the treatment group and the control group, 48 cases in each group. The treatment group was given *Tongdu Ningshen Decoction*, one dose every day. The control group was given alprazolam tablet, 0.4 mg each time, once before sleep at a draught. Both groups were treated for 4 weeks. Insomnia Severity Index (ISI) scores, Pittsburgh Sleep Quality Index (PSQI) scores, Self-Rating Scale of Sleep (SRSS) scores and Chinese symptoms scores in both groups were compared before and after treatment. **Results** After treatment, ISI, PSQI, SRSS scores and scores of symptoms in both groups, such as difficult to fall asleep, waking in the night and early morning, dreaminess, oppression in chest, palpitation, depression, night sweating, dizziness, headache, weakness, amnesia, cervical rigidity, dysphoria, etc. decreased ($P < 0.01$). After treatment, ISI, PSQI, SRSS scores and scores of waking in the night and early morning, dreaminess, oppression in chest, palpitation, depression, night sweating, dizziness, headache, weakness, amnesia, cervical rigidity and dysphoria in the treatment group were lower than those in the control group ($P < 0.01$). **Conclusion** *Tongdu Ningshen Decoction* might improve insomnia symptoms of governor channel discord type and relieve secondary symptoms of insomnia.

Keywords insomnia; governor channel discord; *Tongdu Ningshen Decoction*

(收稿日期: 2015 - 09 - 16; 修回日期: 2016 - 01 - 05)

编辑: 王利芬

(上接第 923 页)

Research on Ontology Construction Technique of Commonly Used Chinese Medicine Terms

XIE Qi¹, LIU Liangliang², CAO Xinyu³, GUO Yufeng⁴, LIU Baoyan¹, ZHOU Hongwei³

(1. China Academy of Chinese Medical Sciences, Beijing, 100700; 2. Business Information School, Shanghai University of International Business and Economics; 3. Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences; 4. Guang'anmen Hospital, China Academy of Chinese Medical Sciences)

ABSTRACT Clinical data is a dispensable source to conduct clinical research. Commonly-used terms of clinical Chinese medicine are carrier of clinical data. Its completeness and normalization are the important precondition of acquiring high-quality clinical data. This paper analyzed the fundamental quality demand of commonly-used clinical Chinese medicine terms, and the common problems in aspects of completeness, normalization and conceptual relation. The authors raised to establish ontology of commonly-used terms of clinical Chinese medicine using technique of ontology. Additionally, they concluded the key technology steps of constructing ontology of commonly-used clinical Chinese medicine terms, including establishing application technology system of clinical commonly-used terms, constructing standard of clinical Chinese medicine terms, acquiring commonly-used terms of clinical Chinese medicine, establishing ontology relationship of frequently-used terms of clinical Chinese medicine, and conducting testing and visualization of ontology. Conducting concept identification and description of commonly-used clinical Chinese medicine terms from the aspect of ontology, and establishing correlation between concepts, can eliminate the uncertainty of Chinese medical concepts and uncertainty of their correlation.

Keywords clinical terms normalization; clinical data; conceptual correction

(收稿日期: 2015 - 07 - 23; 修回日期: 2016 - 01 - 08)

编辑: 崔京艳