

失眠症中医药信息平台构建的思路与方法

蔡晓鸿 游秋云 王 平*

(湖北中医药大学老年医学研究所, 湖北省武汉市洪山区黄家湖西路 1 号, 430065)

基金项目: 湖北省科技厅 2010 年研究与开发计划项目(2010BCB004); 2011 年湖北中医药大学科研课题

* 通讯作者: pwang54@yahoo.com.cn, (027)68890008

[摘要] 中医药对失眠症有较好的疗效, 为了更好地找出中医诊治失眠症的有关内涵和规律, 提出将文献、实验和临床等信息, 通过构建的失眠症本体把三者联系起来。对于文献采用网络爬虫、全文检索、中文分词等技术实现信息智能检索; 对于实验可以全程记录、监控, 并结合数据分析工具对实验结果进行处理; 此外, 在医院电子病历系统的基础上生成临床科研数据库。借助信息技术将三者融合、分析、挖掘, 探索一条中医现代化研究的思路和方法。

[关键词] 失眠症; 中医药疗法; 信息平台; 数据挖掘

近年来, 失眠症的发病率急剧上升, 对失眠的研究日趋成为提高国民生存质量的重要问题^[1]。中医药治疗失眠具有副作用小、消除伴随症状快、便于减少镇静安眠西药剂量等优势, 对失眠症有较好的疗效。中医认为, 失眠症发病原因较多, 思虑劳倦、内伤心脾、心肾不交、阴虚火旺、脾胃不和等均可影响心神而导致不寐。本文拟就失眠症中医药信息平台的构建进行探讨, 从而探索一条以病种为对象的综合研究方法。

1 研究背景

随着现代科学技术的发展, 许多先进的技术被引入临床研究中, 如动物实验、临床药物试验、基因组学、代谢组学、蛋白质组学等为人类认识疾病、应对疾病做出了巨大贡献。中医学和现代技术之间相辅相成, 互补互用。

李梢等^[2]将病与证的关系置于生物分子网络的背景下进行比较分析, 将某疾病的大量文献和实验数据结合起来, 在此基础上研究发现病与证存在结合点。同时希望在宏观的证候表型与微观的生物学指标之间建立有机的联系, 并从中阐发证候的内在整体特性, 从而为证候生物学基础研究提供了一种可行的思路与方法^[3]。王映辉等^[4]提出, 名老中医经验的继承应遵循“临床信息采集—挖掘提取经验—临床应用验证—机理机制研究—理论指导临床”的模式。上述两种模式都是以数据为载体, 对数据进行处理分析。随着国家中医临床研究基地建设的深入, 各基地医院越来越注重临床信息的采集, 临床科研一体化平台较好地解决了临床数据的采集、整理、集成、分析和挖掘等工作。该类平台在于实现以

患者为中心的医疗数据的整合和信息的共享, 在中医临床诊疗的同时采集实际发生的临床数据, 希望从海量的中医临床诊疗数据中发现中医临床的经验、规律和问题, 验证中医临床疗效, 发现中医临床疗效相关的规律。中医的现代化研究应该将临床、实验和文献等信息综合在一起, 通过领域本体建立相互关系, 在此基础上进行数据分析和挖掘, 才能真正地达到继承和发展中医的目的。

2 失眠症中医药信息平台构建思路

失眠是常见疾病之一, 但其与大多数疾病不一样的特点是, 它一般没有机体或功能的损伤, 但它给患者的身心、学习及工作带来极大的危害。中医在很早以前就开始关注失眠的研究, 有许多医家对失眠的病因病机和临床经验进行了总结。由于其病因复杂, 难以采取单一的方法进行治疗。所以许多经验也难以直接应用于患者, 需依靠临床经验非常丰富的医者才能灵活地进行运用。将前人诊疗失眠的经验、现代的实验技术和临床诊疗数据等信息综合在一起进行研究, 才能更好地为患者服务。但这些信息不是简单地堆砌在一个数据库中, 而是让数据有机地结合起来。

数据的融合是信息平台建设的关键, 实际上文献、实验和临床三种信息之间存在紧密的关系。文献是前人的经验总结, 因为每种经验都是从某个角度或方面去阐述疾病的问题, 疾病的核心是一致的。医学实验提供对人体发育成长过程的深入了解和对各种疾病病因和发病机理的可靠知识。可以说, 实验方法的引进对现代医学科学的建立和发展起着决定性的作用。医学实验都是有目的地围绕目标建

模,进行实验,得到实验数据。有关失眠的实验也不例外。各种实验数据正呈指数式的增长,各种共享数据库也在蓬勃发展。随着医院信息化建设的飞速发展,电子病历的应用也越来越广泛,许多临床诊疗信息也已电子化。如何把文献、实验和临床等信息关联起来,为疾病研究构筑综合平台?

本体方法是中医药领域知识分析与组织的可行方法。通过对文献和病种资料的研究,引进本体技术,创建失眠症本体知识库,由本体语义的规范性及其之间的关系重构所有的信息资源。大部分学者公认的本体定义是 Gruber^[5] 在 1995 年提出的:本体是关于共享概念的一致约定。共享概念包括用来对领域知识进行建模的概念框架、需要互操作的主体之间用于交互的与内容相关的协议,以及用于表示特定领域理论的共同约定。在知识共享的情况下,本体的形式特化为具有代表性的词汇的定义。一种最简单的形式是一种层次结构,用来详细描述类和它们之间的包含关系。遵循中医基础理论,通过本体建模解释中医概念及其关联关系,对以中国传统哲学为指导、以整体观念和辨证论治为特色的中医药概念知识体系,应用现代信息技术进行还原与诠释,将为中医药信息组织与数据挖掘分析利用奠定基础,对中医药领域知识的表达、中医药信息的交换与重用必将有着积极和促进作用^[6]。

失眠症中医药信息平台的构建思路为首先在领域专家的指导下建立失眠症本体知识库;其次开发文献资源管理与检索系统,实现文献的采集、分类、智能检索;建立实验数据管理系统;从医院电子病历系统生成临床科研数据库。文献资源管理系统、实验数据管理系统和临床科研数据库都应按本体知识库中的本体作为元数据组织资源。最后运用数理统计、计算机模拟、数据仓库和数据挖掘等技术实现信息的综合分析及运用。总体结构如图 1 所示。

3 失眠症中医药信息平台各组成部分的构建方法

3.1 构建失眠症本体知识库

失眠是一种持续一段时间的睡眠质量令人不满意的状态,常表现为难以入睡、维持睡眠困难、过早或间歇性觉醒而导致睡眠不足,醒后出现不适感、疲乏、白天困倦等症状,是典型的心身疾病。中医药通过调整人体脏腑气血的功能,常能明显改善睡眠状况,且不引起药物依赖,更不会引起医源性疾患,因而受到患者的欢迎。

中医本体的构建是通过对中医概念的提取、关

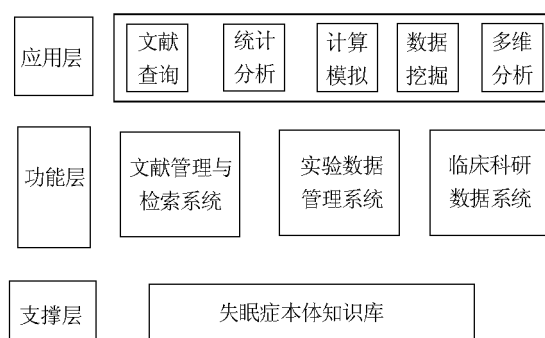


图 1 失眠症中医药信息平台总体结构图

系的分析,将中医知识明确、规范地表达,以便于知识的共享和传承。本体的构建一般遵循明确和客观性(即本体应该用自然语言对所定义术语给出明确的、客观的语义定义)、完全性(即所给出的定义是完整的,完全能表达所描述术语的含义)、一致性(即由术语得出的推论与术语本身的含义是相容的,不会产生矛盾)、最大单调可扩展性(即向本体中添加通用或专用术语时,不需要修改已有的内容)和最小承诺(即对待建模对象给出尽可能少的约束)等五项规则^[5]。

领域本体构建常用的方法是“骨架”法,其一般流程是:①确定本体的应用目的与应用范围。确定本体的研究目的和用户的意图;确定本体的研究范围,领域越大,所建立的本体就越大,因此需限制研究的范围。②本体的分析。定义本体所有术语的意义及其之间的相关关系,该步骤必须有领域专家的参与,专家对该领域了解得越深入,所建立的本体就越完善。③本体的表示。用语义模型表示本体。④本体的检验。按照本体构建的规则进行检验。⑤本体的建立。对所建立的本体按上述标准进行检验,符合要求时,可以建立本体,否则转“②本体的分析”。

根据前述的目标,失眠症本体的应用范围包括文献、实验和临床等方面,为失眠症的各类研究人员提供支持。失眠症本体构建以世界卫生组织(WHO)发布的《基于失眠症的传统医学临床实践指南》^[7]作为基本规范性文件,辅以失眠症的中医经典文献来抽取失眠症的概念、关系等。又以国家标准《中医基础理论术语》^[8]框架为基础,将失眠症本体顶层概念分为失眠症的疾病分类、症状、证候、病因病机、诊断、治则治法、方药、护理与康复、实验、病例等大类。如在方药类下有药和方两类,方类又可分为经方和验方等子类,具体的方可与证候和治则治法关联。本体的构建应用已有的中医学本体和国

内外生物医学领域本体,不能单独自创体系。本体的建立过程是一个反复、不断完善的过程。

3.2 构建文献管理与检索系统

中医药文献历史悠久,古代文献数量较大,其明显有别于现代文献类型,经过多年的中医药文献数字化处理之后,大量文献以电子化格式的形式存储起来。中医古籍中蕴含着大量的中医临床经验的记载,然而由于临床各科内容没有严格的界限,其记载的内容涉及十分广泛,各种内容庞杂而纠集在一起^[9]。所以这种形式的电子化文档可用性还较差。由于科学的分工和细化,现代文献内容更专更精,已基本实现了数字化存储,故现代文献更能有效利用。如何合适地将古今文献结合起来成为有机的整体,是科学研究中非常突出的问题。依据失眠症本体,本学科的概念术语就可以关联。引进先进的搜索技术,如全文检索、中文分词、网络爬虫等技术应用于文献库,可实现文献的自动收集、入库、自动分类、智能查询等工作,提高文献的利用效率。为更好地为研究人员提供文献服务,本文提出将基于本体的垂直搜索技术应用于失眠症的文献研究中。垂直搜索是相对通用搜索引擎的信息量大、查询不准确、深度不够等提出来的新的搜索引擎服务模式,通过针对某一特定领域、某一特定人群或某一特定需求提供的有一定价值的信息和相关服务。其特点就是“专、精、深”,且具有行业色彩,相比较通用搜索引擎的海量信息无序化,垂直搜索则显得更加专注、具体和深入。构建基于失眠症的垂直搜索可满足研究者高质量的信息检索需求。

3.3 构建实验数据管理系统

信息技术不只是实验结果的分析工具,而应是自始至终伴随实验进行,为实验提供知识上的支持。目前已有的医学生物学的数据库(如序列数据库)已经无法再将医学生物学自身复杂的知识整合进去。这是数据库建模思想和技术上的缺陷^[10]。通过本体去规范词汇、字典和操作过程,使研究人员和实验室之间充分共享授权数据。

实验数据管理系统应可实现数据记录、处理、保存和备份等功能,使研究人员高效管理实验室动物种群信息和动物实验数据。并可实现数据报告、自动记录、自动提醒、绘图和分析工具等功能,以提高实验室工作效率。

3.4 构建临床科研数据系统

通过对失眠症患者长期诊疗的观察资料的分

析,可为治疗方案评价、疾病病理演变、诊疗指南修订等提供真实可靠的证据。目前以电子病历为核心的医院临床信息系统在国内医院普遍运行,而将临床数据转化为科研数据则是下一步医院信息化发展的重点。在临床医疗电子病历系统的基础上进行临床科研数据的收集和提取是可行的,临床科研数据库中的信息可以整合患者所有医疗流程所产生的信息。但是由于我国卫生医疗标准制定的落后,不同的电子病历系统数据标准不一。失眠症临床科研数据库应是按失眠症本体建立的,因此,需要数据转换工具将电子病历系统中的信息转换成临床科研数据库中的格式。

3.5 平台应用构想

以失眠症本体为纲,将文献、实验和临床等信息贯穿起来。由于各类信息的结构格式不一,研究人员的科研目标也不一样,所以需要设计数据管理工具完成数据的抽取、转换和清洗等功能。建设信息平台的目的除了对信息的规范和采集信息外,更主要的是能对这些已经做过规范处理的信息进行客观分析。利用计算机技术可开展文献智能查询、统计分析、计算模拟、数据挖掘和多维分析等应用。

4 结语

中医理论作为一套理论体系,必然有其稳定、核心的内容,中医对疾病的认识必然存在一定的共性规律。共性规律有利于理论形成和推广,但缺乏具体运用时针对性;个性差异体现了特色和疗效优势,但难于总结和提炼。通过运用本体的技术,可解决生物分子—证候—文献、生物分子—证候—方药—文献等多元关系的表示。根据失眠症本体知识库可对某一证候进行智能搜索,构建此证候表象网络,如建立心脾两虚证失眠症的症状和体征网络。实验数据经过聚类分析、主成分分析构建数据网络,如分子生物学的实验数据可构成生物分子网络。上述不同的网络间通过病证关系关联起来形成巨复杂系统。临床信息包括了理、法、方、药,可以对表象网络、数据网络进行验证。重视信息资源的纵贯性,信息资源应包括基因组、蛋白质组、生化、病理、临床等多方面多维度的数据。模型的构建不仅要考虑到分子层面,还要涵盖临床治疗。通过这些方法和技术,希望能将失眠症的共性规律和诊治的个性差异结合起来进行研究,为中医现代化作出有益的探索。

(下转第 772 页)

阴中求阳,则阳得阴助而生化无穷;善补阴者,必于阳中求阴,则阴得阳生而泉源不竭。”以薏苡仁、肉桂、椒目为佐,温肾健脾利湿。中焦湿土得健,升降运化正常,则湿气得除。以三棱、莪术、鳖甲、泽兰为使,软坚散结,活血利水。综观全方君、臣、佐、使配伍,药专力宏,水寒得温,肝郁疏解,脾虚得健,水湿得除,则肝经乙木得升,宗筋经络气血通畅,性功能得以改善。

中药药理研究显示,附子、淫羊藿、肉桂、菟丝子等组方的温阳药能增强下丘脑—垂体—性腺轴的分泌功能,亦能明显提高性机能。其中淫羊藿能显著增加血浆 T 的含量及睾丸、提肛肌的重量,具有明显的促性机能作用,其作用强度与肌注 T 无显著性差异,并且促进睾丸组织增生与分泌^[10]。另有研究发现,淫羊藿煎剂能有效地修复大鼠睾丸间质细胞的损伤,支持睾丸曲精管上皮正常生精周期^[11]。章振保等^[12]以环磷酸胺致生殖系统受损、雄激素部分缺乏大鼠模型,给予菟丝子提取物 200 mg/kg 灌胃,能升高大鼠血清 T 值、增加睾丸重量,并可减缓睾丸组织的退行性变化。说明菟丝子具有雄激素样或促雄激素样作用。

本研究结果提示,温肾软肝清化汤能促进男性肝硬化患者血清 T 复常,明显改善患者性功能减退、乏力、纳差、肝区不适症状,并有较持久的治疗效果,使患者生活质量得以提高。且使肝功能得以明显改善。

参考文献

- [1] 石胜利,聂玉强. 脂联素与非酒精性脂肪肝病[J]. 世界华人消化杂志,2007,15(3):272-275.
- [2] 余剑英. 男性肝硬化患者血清性激素与垂体激素水平变化及意义[J]. 实用预防医学,2005,12(5):1104-1105.
- [3] 郝家明. 男性肝硬化患者血清性激素水平变化[J]. 放射免疫学杂志,2009,22(4):340-341.
- [4] 苏庆军,赵燕,陈建国,等. 男性肝硬化患者血清性激素水平的分析[J]. 实用医技杂志,2006,13(22):3923-3924.
- [5] 姚光弼,赵慧娟,聂青和,等. 临床肝脏病学[M]. 2 版. 上海:上海科学技术出版社,2011:319.
- [6] 中华医学会传染病与寄生虫病学分会、肝病学分会. 病毒性肝炎防治方案[J]. 中华肝脏病杂志,2000,8(6):324-329.
- [7] 中国中西医结合学会消化系统疾病专业委员会. 肝硬化中西医结合诊疗共识[J]. 中国中西医结合消化杂志,2011,19(4):277-279.
- [8] 国家中医药管理局. 中医病证诊断疗效标准[S]. 南京:南京大学出版社,1994:45.
- [9] 陈从新,周永兴,陈琍曲,等. 慢性乙型肝炎和肝炎后肝硬化患者性功能和血清性激素水平[J]. 中华传染病杂志,1995,13(1):46-48.
- [10] 牛锐. 淫羊藿炮制前后对小鼠血浆睾酮及附性器官的影响[J]. 中国中药杂志,1989,14(9):18-20,62.
- [11] 沈映君,徐秋萍,陈奇,等. 中药药理学[M]. 北京:人民卫生出版社,2000:904.
- [12] 章振保,杨庆涛,杨镜秋,等. 淫羊藿甙、菟丝子提取物对雄激素部分缺乏大鼠生殖保护作用的比较研究[J]. 中国老年学杂志,2006,26(10):1389-1391.

(收稿日期:2012-10-25;修回日期:2013-01-03)

[编辑:王利芬]

(上接第 749 页)

参考文献

- [1] 黄攀攀,王平,李贵海,等. 老年阴虚失眠动物模型的建立与评价[J]. 中华中医药学刊,2010,28(8):1719-1723.
- [2] Li S, Wu LJ, Zhang ZQ. Constructing biological networks through combined literature mining and micro array analysis: a LMMA approach [J]. Bioinformatics, 2006, 22(17):2143-2150.
- [3] 李梢. 中医证候生物分子网络标志的构想与研究[J]. 中医杂志,2009,50(9):773-776.
- [4] 王映辉,张润顺,刘保延,等. 名老中医经验传承研究模式探索研究[J]. 环球中医药,2008,1(2):13-14.
- [5] Gruber TR. Toward principles for the design of ontologies used for knowledge sharing[J]. Int J Hum Comput Stud,1995,43(5-6):907-928.

- [6] 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[D]. 北京:中国中医科学院,2011:85.
- [7] 基于失眠症的传统医学临床实践指南[S]. 世界卫生组织·西太区,2007.
- [8] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会. 中医基础理论术语[S]. 北京:中国标准出版社,2006.
- [9] 裴丽. 中医古籍信息利用障碍与数字化信息平台构建[D]. 哈尔滨:黑龙江中医药大学,2006:94-95.
- [10] 张其鹏,孙冬泳,卢铭,等. 医学生物信息学在心血管研究领域的研究应用—数据与知识[J]. 生理科学进展,2005,36(2):119-124.

(收稿日期:2012-09-10;修回日期:2012-11-19)

[编辑:崔京艳]