

• 标准与规范 •

## 中医证候诊断量表编制中的问题与思考\*

朱杉杉<sup>1</sup> 刘强<sup>2</sup> 孔静敏<sup>1</sup> 张根明<sup>3</sup>

1. 北京中医药大学第一临床医学院, 北京 100700; 2. 世界中医药学会联合会评价中心, 北京 100101;  
3. 北京中医药大学东直门医院脑病科, 北京 100700

**摘要:** 中医证候诊断量表在量表的功能、条目的表达、条目池的形成、条目的筛选与分级量化、量表的验证与应用方面仍存在尚未解决的问题。如量表的诊断与评价功能相混淆, 条目的表达是否规范合理, 量表缺乏诊断性试验角度的验证等。其中一些问题是可以避免的, 如可以通过加深对诊断与评价二者概念的理解来避免量表功能的混淆; 注意从诊断性试验角度来验证量表。而有些问题目前并无很好的解决方法, 如量表条目的筛选、条目量化等, 还有待于量表研制方法的进一步发展完善。

**关键词:** 中医证候诊断量表; 量表编制; 量化诊断; 量表功能; 条目池; 条目表达; 条目筛选; 量表验证; 量表应用

**DOI:** 10.16368/j.issn.1674-8999.2021.10.447

中图分类号: R241.24 文献标志码: A 文章编号: 1674-8999(2021)10-2138-04

### Key Questions and Thoughts of Scale Development of TCM Syndrome

ZHU Shanshan<sup>1</sup> LIU Qiang<sup>2</sup> KONG Jingmin<sup>1</sup> ZHANG Genming<sup>3</sup>

1. First Clinical Medical College of Beijing University of Chinese Medicine, Beijing, China 100700; 2. Evaluation Center of World Federation of Chinese Medicine Societies, Beijing, China 100101; 3. Encephalopathy Department of Dongzhimen Hospital of Beijing University of Chinese Medicine, Beijing, China 100700

**Abstract:** In TCM Syndrome Diagnostic Scales there are still some unresolved problems in scale functions, the expression of items, the formation of item pools, the selection and grading and quantification of items, and the verification and application of the scales, such as the diagnostic and evaluation functions of the scale are confused, whether the expression of items is standardized and reasonable, the scale lacks verification from the perspective of diagnostic tests, and so on. Some of them can be avoided. For example, the confusion of scale functions can be avoided by deepening the understanding of the concepts of diagnosis and evaluation. Attention should be paid to verifying the scale from the perspective of diagnostic tests. However, there are currently no satisfying solutions to some of the problems mentioned above, such as the selection of scale items, quantification of items, etc., which need to be further developed and perfected in the methodology of scale compiling.

**Key words:** TCM syndrome diagnostic scale; scale compiling; quantitative diagnosis; item pools; scale functions; the selection of items; the verification of the scales; the expression of items

证候是中医学特有的名词,是临床进行辨证论治的基础,准确辨析患者的证候,方证合一,从而提升中医药的临床疗效。证候是综合分析了患者的临

床症状与体征,对疾病过程中所处一定阶段的病位病性等作出的病理性概括,是对疾病当前本质所作的结论<sup>[1]</sup>。王永炎院士指出证候是一个非线性的复杂系统,具有“内实外虚”“动态时空”“多维界面”的特征<sup>[2]</sup>。并且在临床辨证时多依赖于医师自身的临床经验,使得辨证结果具有一定的模糊性与

\* 基金项目: 国家中医药管理局国家中医临床研究基地业务建设科研专项课题项目(JDZX2015050)

不确定性。对中医证候标准化、客观化的研究是中医证候研究的重要部分,中医证候量表的研制进一步促进了中医证候标准化的发展<sup>[3]</sup>。

量表作为一种测量工具,是一种试图确定主观或是抽象概念的定量化测量程序,其主要作用和目的在于精确测量一个较抽象或综合性较强的概念<sup>[4-5]</sup>。早期量表作为心理学评价方法在临床中应用,随着不断发展完善,也应用到了医学的其他方面。前人通过借鉴西医量表的研制方法,基于中医基础理论,研制具有中医特色的证候诊断量表。通过证候诊断量表对证候实行量化诊断,也成为中医证候标准化与客观化研究的重要方式之一<sup>[6]</sup>。随着研究的不断深入,证候诊断量表的研制方法在不断的完善,但在量表的研制关键环节与应用中仍存在一些问题值得思考,详述如下。

## 1 量表功能的区分

量表主要是诊断与评价两方面的功能,依据量表的功能不同可以将量表分为具有诊断功能的诊断量表、具有评价功能的评价量表以及兼具诊断与评价功能的量表。但笔者发现在研制过程中有学者将量表的诊断与评价功能相互混淆,或者混用诊断量表与评价量表的研制方法,导致最终的量表具体功能不明,甚至是无法应用。造成这种情况主要有两方面原因:一方面是因为在最初确定研究目的时未很好地区分“证候诊断”与“证候评价”二者的概念,并且在将二者做出严格区分,因此很容易将证候诊断与证候评价相互混淆;另一方面是因为最初研制中医证候量表时借鉴了西医心理测评量表的方法,并未对诊断量表与评价量表的研制方法进行深入研究及区分,从而出现使用评价量表的研制方法来进行诊断量表研制的现象。

诊断与评价是两个不同的概念,在进行量表编制时应对二者的概念进行明确的界定,且诊断量表与评价量表在研制的核心理论及方法上也有所不同<sup>[7-8]</sup>。评价量表是对目标定量或有序分类的研究,其重点在于对疾病或证候轻重的区分,研制关键在于对目标概念的定义与理解;诊断量表是对目标定性或无序分类的研究,其重点在于对疾病或证候有无可判断,研制关键在于诊断模型的建立。若用评价量表的研制方法来研制诊断量表,或者是二者的研制方法混用,最终将限制量表的诊断能力,甚至是丧失量表的诊断能力。郭盛楠<sup>[9]</sup>以肝郁化热证为例,对证候诊断量表与证候评价量表的研制方法

进行区分,所编制的证候量表具有良好的性能,且通过测评发现在诊断性能上诊断量表优于评价量表,而在评价性能上则是评价量表优于证候量表,表明依据不同的研制方法去构建诊断量表或评价量表能够提升量表的性能。因此,在证候诊断量表研制的初始,就需要明晰诊断与评价二者的概念,严格区分量表的诊断与评价功能,分辨诊断量表与评价量表研制方法的不同之处,选择适合诊断量表的研制方法,从而避免上述情况的发生。

## 2 条目池的形成

条目池由众多与量表相关的条目所组成,即是由患者的临床症状、体征等所组成的集合。文献的梳理研究是目前常用于建立条目池的方法,其他如临床调查、专家咨询、患者访谈等方法也常用于获取四诊信息<sup>[10-11]</sup>。有学者认为由于量表的条目大多来源于中医文献典籍、临床医案、专家咨询等,使得大部分或主要条目有着先入为主的经验框架烙印,在形成条目池时具有较强的主观性,从而使量表的辨证结果产生偏差<sup>[12-13]</sup>。但笔者认为证候诊断量表是在对疾病的专业认识、中医基础理论以及中医对疾病与证候的认识等基础之上,同时结合了专家临床辨证诊断经验来研制的。其中对于疾病、证候的认识是客观的,但专家的临床辨证诊断经验则具有个人主观性。故而在形成条目池时,是一定会存在个人经验,存在主观性,也必然是先入为主的。但在条目池形成中的主观性可以通过多方面的信息来源来解决,比如可以纳入更多的具有代表性的专家来参与这一过程,尽可能地从不角度来形成条目池,从而减少主观性的影响。

## 3 条目的表达

量表由条目组成,每一个条目对应着一个症状或者体征。在临床中,对于症状或体征的描述并不完全相同,但对于量表编制来说,条目不能随意描述,应该有所要求及规范。因此如何对条目进行表达或描述是量表研制过程中的重要环节。目前有不少学者提出应对条目的表达进行规范化、术语化,但是忽视了使用人群的差异对条目表达的影响。量表条目的表达与量表的使用人群息息相关,使用人群不同,量表条目的表达也要有所侧重。

对于中医证候诊断量表,如果其使用者是患者,考虑到患者对于中医相关证候、术语的不了解,在描述条目时需要更加通俗化的语言或者是对条目加以释意以便于其理解。如果使用者是西医学者,在描

述条目时可能需要对一些中医术语进行解释、界定内涵等,或者是使用常见表述来代替中医术语,如“纳呆”可以用“食欲不振”来代替,“不寐”用“失眠”来代替等。如果使用者是中医学者,条目表达时应注重其规范性,可以借鉴相关标准来规范中医术语的表达,如《中医临床常见症状术语规范》《中药新药临床研究指导原则》等,对于量表中存在的一症状多名、多症状一名、一词多义、同名异义、异名同义等现象,可以从症状名称、内涵、外延、独立性等方面对其进行规范以减少偏差<sup>[14]</sup>。也有学者提出可以从明确症名定义、统一规范症名、拆分复合症名、区分症状轻重、避免诊断性症名、纳入客观指征等方面对中医症状进行规范性描述<sup>[15]</sup>。

因此,在条目表达时不仅要考虑到条目描述的准确性、规范性,也需要考虑到使用人群的不同,选择最适合的描述形式,让使用者能够更准确地理解条目的含义,从而避免在使用量表时因对条目理解的歧义而使结果出现偏差。

#### 4 条目筛选及分级量化

条目筛选是量表研制过程中的关键步骤之一,目前常用的筛选方法是主观筛选法、客观筛选法及主客观综合筛选法。主观筛选法是指通过专家咨询等获得信息来对条目进行筛选的方法,包括均值、满分比、变异系数、权威系数等;客观筛选法是指通过对临床数据的统计分析进行筛选,如离散趋势法、区分度分析法、相关系数法、因子分析法、克朗巴赫系数法、Logistic 回归法等<sup>[3,16]</sup>。虽然条目的筛选方法有多种,但每一种统计学方法都有一定的局限性,侧重于从某一角度对条目进行筛选,如离散趋势法从敏感性角度筛选条目,相关系数法考察的则是条目的代表性与独立性等<sup>[17]</sup>。因此更倾向于综合应用多种统计学方法从多角度筛选条目。在量表编制的具体过程中,所选择的筛选方法也不完全相同,主要是遵循重要性大、独立性强、代表性好、敏感性高、确定性好的原则,同时兼顾可操作性及可接受性<sup>[18]</sup>。不过这些原则是否全面、是否需要补充与完善,还有待进一步探究。

四诊条目如何客观合理的量化一直是中医证候量表研制的难点之一,尤其是体征类条目如舌象、脉象等的量化。笔者认为在证候诊断量表的研制中进行症状、体征的量化并非是必须的,但一些症状体征的轻重不同在辨证中有着不同的意义,应该要有所量化区分,并且有些症状体征的表述本身就具有轻重的区分,如舌红、舌绛提示了热象的轻重。在量化

时是否还需要对这些条目进行量化,又该如何量化,还需要进一步探讨。而对于证候等级分化的诊断量表,对各项条目进行客观合理的量化则是研制量表的重点,但目前并没有公认的方法。有学者尝试采用逐步判别分析法<sup>[19]</sup>、综合评价层次分析法<sup>[20]</sup>、症征联合量化的等级计量诊断法<sup>[21]</sup>等方法对四诊条目进行分级量化。

#### 5 量表的验证

量表验证是指在量表成型后考评量表是否具有好的证候诊断或评价能力的过程,是量表编制中非常重要的一个环节。诊断量表与评价量表的编制方法、功能不同,验证量表时也应有所侧重。证候诊断量表研制的目的是用于判断证候的有无,在验证时应侧重于量表诊断能力的评价,主要通过诊断性试验来对量表的灵敏度、特异度及判断准确度进行评价,以此来反映量表诊断能力的优劣。有学者并未通过诊断性试验的方法对诊断量表进行考评,而是采用信度检验、效度检验、反应度等方法对诊断量表进行考评,并不能很好地说明量表的诊断性能如何<sup>[22]</sup>。因为信度、效度、反应度等指标常用于评价量表的验证,而且这些指标并不能体现诊断量表的诊断能力。因此,从诊断性试验角度来验证诊断量表的诊断性能是必不可少的。

诊断量表的验证核心其实是将该量表与已有的诊断标准做对比,从诊断准确率、敏感度等方面来评估量表诊断能力。在量表验证过程中,最好的参照标准是诊断“金标准”,但由于目前尚无统一的中医证候诊断标准,使得证候诊断量表在验证时缺乏最好的参照标准。因此在量表验证时常以专家辨证的结果为参照标准。在缺乏“金标准”的情况下如何对诊断性试验结果进行更加客观的解释与评价是量表研究中的难点。刘强等<sup>[23]</sup>研究表明,多种诊断方法用于同一总体时估计灵敏度与特异度的校正方法,可以对缺血性中风证候要素量表验证的灵敏度及特异度进行校正,但该方法需要早期量表和经验信息数据进行参照,给该方法的推广应用带来一定的困难。缺乏完善“金标准”或无“金标准”情况下的诊断学数据的校正方法,对于中医证候量表的研究具有重要意义,这些方法值得进一步关注与探讨。

#### 6 量表的应用

研制量表的目的是为了能够利用量表来解决一些临床或科研中的诊断与评价方面的问题。但量表在实际应用中能否很好地发挥其作用还需要进一步

观察,从应用中获得反馈来完善量表。目前对于中医诊断量表在临床或科研中的实际应用情况缺乏关注,也少有文献提及这一方面。另外由于各个量表的具体内容、使用人群、应用范围都有所不同,在使用时的要求会有所不同,因此笔者认为应给量表附加一份使用说明,简要说明该量表的适用范围、适用人群、使用时的注意事项等,从而避免因使用不当带来的偏差。

证候要素诊断量表的研制是目前中医量表研究中的热点之一。证候是对疾病当前阶段的病位、病因、病性等所作的病理性概括。王永炎等<sup>[24]</sup>提出证候是由证候要素和证候靶位组合而成。但证候的临床表现并非证候要素和证候要素靶位的简单组合或线性叠加,一个证候所包含的证候要素间有着复杂的相互关系。因此,将证候要素诊断量表运用到临床辨证时所代表的诊断效用有多少?又该如何去认识证候要素诊断量表在临床应用中的定位呢?对此,笔者认为张志斌等<sup>[25]</sup>提出的以证候要素、应证组合为核心的多维多阶的辨证体系有助于在临床上更好地应用证候要素诊断量表。

## 7 结语

中医证候诊断量表在量表的功能、条目的表达、条目池的形成、条目的筛选与分级量化、量表的验证与应用方面仍存在尚未解决的问题,如:量表的诊断与评价功能相混淆,条目的表达是否规范合理,量表缺乏诊断性试验角度的验证等。其中一些问题是可以通过加深对诊断与评价二者概念的理解来避免量表功能的混淆;注意从诊断性试验角度来验证量表。而有些问题目前并无很好的解决方法,如量表条目的筛选、条目量化等,还有待于量表研制方法学的进一步发展完善。对于证候诊断量表的研究一直在进行,相信随着更加深入的研究,其研制思路与方法将会进一步完善。

## 参考文献:

[1]李灿东,吴承玉.中医诊断学[M].3版.北京:中国中医药出版社,2012:4-5.  
[2]田金洲,王永炎,时晶,等.证候的概念及其属性[J].北京中医药大学学报,2005,28(5):6-8.  
[3]李先涛,周旋,方格,等.中医证候诊断量表研究现状[J].辽宁中医杂志,2020,47(1):18-21.  
[4]徐鹏,吕志国,张影,等.中医证候量表研制方法概述[J].长春中医药大学学报,2016,32(3):628-630.  
[5]朱诗兵,裘生梁.中医证候量表的研制进展[J].中外医疗,2012,31(9):182-183.

[6]赵晖,李淳,李志更.证候诊断标准研究的关键问题探讨[J].中华中医药杂志,2017,32(8):3561-3565.  
[7]刘强,高颖.证候诊断量表的理论基础[J].中华中医药杂志,2010,25(7):989-992.  
[8]高颖,马斌,刘强,等.缺血性中风证候要素诊断量表编制及方法学探讨[J].中医杂志,2011,52(24):2097-2101.  
[9]郭盛楠.肝郁化热证诊断与评价量表的研制及其方法学探讨[D].北京:北京中医药大学,2014.  
[10]赵紫薇.重症肌无力中医证候要素诊断量表条目的筛选[D].长春:长春中医药大学,2019.  
[11]孟凡波,杨传华,齐冬梅,等.高血压病肝阳上亢证诊断量表条目筛选的研制[J].南京中医药大学学报,2012,28(3):225-228.  
[12]黄诚,宋炜熙.问卷与量表在中医证候辨证中的研究与应用[J].中华中医药学刊,2011,29(4):889-891.  
[13]严石林,陶怡,曾跃琴,等.从辨证思维探讨中医证候量表存在的问题及对策[J].中医杂志,2013,54(24):2082-2084.  
[14]熊丽枚,甘慧娟.试论阴虚证量表研究中如何对症状进行规范化描述[J].中华中医药杂志,2016,31(7):2672-2674.  
[15]魏佳,夏淑洁,陈锦明,等.论中医“症”描述的准确性与规范化[J].天津中医药,2020,37(5):535-539.  
[16]郑昊,黄嘉华,潘美均,等.近5年中医诊断量表文献分析[J].河南中医,2020,40(11):1703-1708.  
[17]朱羽硕,王宇,李运伦,等.正常高值血压阴虚阳亢证诊断量表条目筛选[J].时珍国医国药,2018,29(7):1777-1779.  
[18]方积乾.生命质量测定方法及应用[M].北京:北京医科大学出版社,2000:11.  
[19]温淑云.脑梗死中医证候量化分级研究[J].中医杂志,2010,51(2):157-159.  
[20]李东涛,王剑,姜洪洋,等.骨质疏松症常见中医定性证候轻重程度量化评价[J].中西医结合学报,2012,10(11):1254-1262.  
[21]曹泽标,周昊,范钊坤,等.基于计量诊断的计量中医模式探讨[J].中医杂志,2015,56(18):1548-1551.  
[22]赵灵燕,毕力夫,陈建新,等.2型糖尿病中医证候要素诊断量表的研制及考评[J].北京中医药大学学报,2018,41(10):848-853.  
[23]刘强,王肖南,王少卿,等.缺血性中风证候要素诊断量表不完善金标准验证结果的校正[J].中医杂志,2016,57(11):949-952.  
[24]王永炎,张启明,张志斌.证候要素及其靶位的提取[J].山东中医药大学学报,2006,30(1):6-7.  
[25]张志斌,王永炎,吕爱平,等.论证候要素与证候靶点应证组合辨证[J].中医杂志,2006,47(7):483-485.

收稿日期:2021-04-20

作者简介:朱杉杉(1994-),女,安徽芜湖人,博士研究生,研究方向:中医药防治脑血管病、中风病诊断量表编制研究。通信作者:张根明(1968-),男,北京人,医学博士,教授,主任医师,研究方向:中医药防治脑血管病临床研究。E-mail:zhanggenming@163.com

编辑:纪彬