

DOI:10.13288/j.11-2166/r.2025.17.012

中医证候诊断量表文献研究现状

张莹莹¹, 王佳佳^{2,3}, 李建生^{2,3}✉, 冯贞贞^{2,3}

1. 辽宁中医药大学第二临床学院, 辽宁省沈阳市皇姑区崇山东路79号, 110847; 2. 河南中医药大学第一附属医院; 3. 河南中医药大学呼吸疾病中医药防治省部共建协同创新中心/河南省中医药防治呼吸病重点实验室

[摘要] **目的** 分析中医证候诊断量表研究现状及存在问题, 为中医证候的标准化研究提供参考和建议。**方法** 检索中国知网、万方数据库、维普数据库、中国生物医学文献数据库建库至2025年4月22日公开发表的中医证候诊断量表研制及评价相关文献, 对量表名称、疾病名称、研究模式、证候类型、条目池的构建方法、条目筛选方法、条目赋权方法、诊断阈值确定方法、量表评价方法、证候严重程度分级进行统计分析。**结果** 共纳入207篇文献, 涉及164个中医证候诊断量表, 以病证量表138个, 其中中医病证结合量表8个、涉及病种3个, 西医病名结合中医证型量表130个、涉及病种77个, 高血压病相关量表数量最多(18个); 以证统病量表26个, 涉及证候20种。常见证候/复合证候诊断量表91个, 基础证候/证候要素诊断量表73个, 动物类证候诊断量表3个。宏观辨证量表154个, 宏观结合微观辨证量表10个。量表维度可分三类, 一是主观症状、客观体征和实验室指标, 二是复合证候的基础证候和舌脉, 三是证素和舌脉; 条目内容涉及舌脉的量表84个。条目池构建以现代文献研究为主(61.99%, 106/171); 条目筛选多基于临床调查(33.22%, 102/307)、专家咨询(30.29%, 95/307)和文献研究(26.71%, 83/307), 涉及均值、频率、离散趋势法、相关系数法等20余种方法。53.05%(87/164)的量表完成了条目赋权; 40.85%(67/164)的量表确定了诊断阈值; 73.78%(121/164)的量表完成了量表评价, 且均采用专家辨证作为辨证的“金标准”; 32.93%(54/164)的量表只评价了信度等测量学特性, 16.46%(27/164)只进行了诊断试验评价, 21.34%(35/164)同时完成测量学特性和诊断试验评价; 14.02%(23/164)的量表探索了证候严重程度分级, 均采用四分位数法。**结论** 中医证候诊断量表研究取得一定进展, 但在研究模式、证候类型、微观指标、维度和条目、研制方法学等方面研究尚薄弱。

[关键词] 中医证候; 诊断量表; 病证结合量表; 量表评价

证候是对疾病过程中某一阶段的病因、病位、病性等的病理性、整体概括, 是一个非线性的“内实外虚”“动态时空”“多维界面”的复杂系统^[1]。中医辨证易受主观因素干扰, 而量表具有客观化、量化的特点, 故引入量表学的方法构建中医证候诊断量表可为临床辨证提供客观工具^[2-3]。随着证候诊断量表数量的不断增多, 有研究^[4-7]对中医证候诊断量表进行现状分析, 为了解证候诊断量表研究

进展提供了参考。但现有研究多缺乏数据支撑, 或侧重统计方法论述, 或年代较久、缺乏系统性论述, 亟需更新完善。本研究以文献研究为基础, 从研究模式、证候类型、宏观与微观指标、维度和条目、方法学等方面深入分析中医证候诊断量表研究现状及存在问题, 并提出相关建议, 以期证候量化诊断提供参考。

1 资料与方法

1.1 文献检索

以“量表”“问卷”“测试”“证”“证候”“证型”“证素”“诊断”为检索词, 根据检索数据库的特征采用相应检索策略, 分别检索中国知网、万方

基金项目:呼吸疾病中医药防治国家中医药传承创新团队项目(ZYYCXTD-C-202206); 2023年度河南省中医学“双一流”创建科学研究专项(HSRP-DFCTCM-2023-2-21, HSRP-DFCTCM-2023-4-09); 2023年度河南省卫生健康委国家中医药传承创新中心科研专项(2023ZXZX1126)

✉ 通讯作者: li_js8@163.com

数据库、维普数据库、中国生物医学文献数据库。检索时间为各数据库建库至 2025 年 4 月 22 日。

1.2 文献纳入标准

1) 明确提及中医证候诊断量表的名称, 不限定研究对象(人类、动物模型); 2) 文献涉及中医证候诊断量表的研制和/或考核的任一环节。

1.3 文献排除标准

1) 文献综述; 2) 仅有摘要或无法获取数据的文献; 3) 重复发表文献, 排除内容不完整的文献。

1.4 数据提取和分析

由 2 名研究者独立阅读文献题目、摘要, 必要时阅读全文, 根据纳入、排除标准筛选合格文献。对于合格文献, 逐一阅读全文, 提取相关信息, 包括作者、标题、发表年份、文献研究内容、量表名称、条目池建立方法、条目筛选方法、条目赋权方法、诊断阈值确定方式、量表评价指标及证候严重程度分级等信息。以了解发文趋势、确定研究模式、确定量表数量和疾病/证候类型、不同环节分析方法的种类和数量、诊断阈值的有无及判断方法、评价指标的分布及证候严重程度的有无及判定方式。若存在意见分歧, 则讨论解决或由第 3 名研究者判定。

1.5 统计分析

采用 SPSS 25.0 软件, 采用描述性统计方法, 统计文献发表年份、病证结合研究模式分类、中医证候诊断量表数量、条目池构建方法种类、条目筛选方法、条目赋权方法、确定诊断阈值的方法、评价量表的指标及疾病类型分布的频次或频率。

2 结果

2.1 文献检索及筛选结果

共检索 1637 篇文献, 排除重复发表文献 579 篇, 通过阅读标题摘要排除 752 篇, 通过阅读全文排除 99 篇, 最终纳入分析 207 篇。其中, 期刊论文 72 篇, 学位论文 134 篇, 会议论文 1 篇。

2.2 中医证候诊断量表研究模式的分布情况

共纳入分析中医证候诊断量表 164 个, 均为病证结合量表。根据“病”与“证”的主从关系, 西医辨病与中医辨证结合可分为“以病统证”和“以证统病”两种诊疗形式^[8]。1) 以病统证量表 138 个, 占比 84.15% (138/164), 包括中医病名结合中医证型量表 8 个, 涉及不寐、郁证、心悸 3 种疾病; 西医病名结合中医证型量表 130 个, 涉及 77 个病种, 相同疾病的证候诊断量表包含的复合证候不

完全相同。证候诊断量表数量排前三位的分别是高血压病、冠心病、2 型糖尿病(见附表 1, 请扫描本文二维码获取)。其中, 高血压病相关证候诊断量表相关研制与考核研究涵盖高血压病、原发性高血压、高血压颈动脉粥样硬化和老年高血压, 涉及不同证型。2) “以证统病”量表 26 个, 占比 15.85% (26/164), 涉及肺阳虚证、肝火上炎证、肝郁气滞证等 20 种证型(见附表 2, 请扫描本文二维码获取)。

2.3 中医证候诊断量表证候类型的分布情况

目前临床中多采用复合证候或常见证候表述, 复合证候/常见证候诊断量表 91 个, 但复合证候存在表述不明确的问题, 例如肝肾不足、心脾两虚等。基础证由一个主要证素或多个证素组合而成^[9]。基础证候/证候要素诊断量表 73 个。在适用对象方面分为人与动物两类, 动物类证候诊断量表 3 个, 分别是高脂血症痰浊阻遏证动物模型评价量表(2024 年)、肝纤维化大鼠阴虚证评价量表(2019 年)、慢性束缚应激肝郁脾虚证大鼠模型评价量表(2013 年)。

2.4 中医证候诊断量表宏观辨证与微观辨证的分布情况

宏观辨证是指在中医的传统辨证方法基础上, 对收集的四诊信息进行规范化、定量化, 最终以实现计量诊断为目的。微观辨证是相对于宏观辨证而言, 主要是指借助组学研究方法, 从理化、免疫、代谢等方面筛查与中医证型相关的微观指标, 并分析其内在的关联性, 以寻求“证”的本质, 进而据此辨识证候^[10]。164 个中医证候诊断量表中宏观辨证诊断量表 154 个, 宏观结合微观辨证诊断量表 10 个。微观指标涉及血栓素 B₂、前列腺素 E₂、花生四烯酸、D-二聚体、血栓弹力图、内镜下黏膜等, 病种涉及高血压病、冠心病稳定型心绞痛、膝骨关节炎、2 型糖尿病、心房颤动、抑郁障碍、胃癌前病变、慢性心力衰竭。

2.5 中医证候诊断量表维度和条目分类的分布情况

量表由若干领域(也称维度)组成, 领域的下级概念为“方面”, “方面”由若干指标组成, 指标在一定原则下可转换成条目组成^[11]。本研究中各证候诊断量表的维度可归为三类, 一是主观症状、客观体征和实验室指标, 如高血压肝阳上亢证诊断量表^[12], 包括头晕等主观症状、舌脉客观体征、血栓素 B₂ 等实验室指标三个维度; 二是复合证候的基础证候和舌脉, 如高血压病肾阴阳两虚证诊断

量表,包括肾阴虚、肾阳虚、舌脉三个维度^[13];三是证素和舌脉,如高血压痰瘀互结证诊断量表,包括痰湿、血瘀、舌脉三个维度^[14]。所有量表均未设置“方面”。条目内容包括各疾病或证候常见症状和体征,共84个量表保留了舌象和/或脉象。

2.6 中医证候诊断量表研制各阶段方法学构成分布情况

量表研制包括条目池的构建、条目筛选、量表评价等关键环节^[5]。164个中医证候诊断量表的研制涉及条目池构建、条目筛选、条目赋权、诊断阈值确定、量表评价、证候严重程度分级共6个重要环节。表1示,1)条目池构建共8个来源(共171次),以现代文献研究为主(61.99%,106/171)。2)条目筛选共5种方式(共307次),以临床调查(33.22%,102/307)、专家咨询(30.29%,95/307)、文献研究(26.71%,83/307)为主。筛选条目的方法在临床调查时以离散趋势法、相关系数法为主,专家咨询时以均值为主,文献研究时以频次、频率为主。3)87个(53.05%,87/164)量表进行了条目赋权,赋权方法涉及主观赋权(17.24%,15/87)、客观赋权(49.43%,43/87)、主客观联合

赋权(33.33%,29/87)三种。明确条目赋权方法19种(87次),前三位依次是Logistic回归分析法(27.59%,24/87)、专家咨询满分比联合因子分析法(27.59%,24/87)、Fisher判别法(11.49%,10/87)。2个量表提及赋权但未明确方法。采用2种方法联合赋权的量表29个(33.33%,29/87),1种方法赋权的量表58个(66.67%,58/87)。4)明确诊断阈值的量表63个(40.85%,67/164),62个(37.80%,62/164)量表采用了受试者工作特征曲线(ROC),1个自定义了诊断阈值。5)123个(73.78%,121/164)量表完成了评价,54个(32.93%,54/164)量表只完成了信度、效度、可行性等测量学特性评价,27个(16.46%,27/164)只完成了诊断试验评价,35个(21.34%,35/164)同时完成测量学特性评价和诊断试验评价。诊断性试验中需要比较新标准与“金标准”之间灵敏度、特异度和辨证准确度。测量学特性评价中常选择认可度高的量表作为校标。除1个量表采用Fried衰弱量表作为“金标准”外,其余均选择专家经验作为“金标准”。6)14.02%(23/164)的量表探索了证候严重程度分级,均采用四分位数法。

表1 中医证候诊断量表文献中研制各阶段方法学构成分布情况

Table 1 Distribution of methodological components across different development stages in literature on TCM syndrome diagnostic scales

阶段	方法	频次
条目池构建	现代文献研究	106
	古籍结合现代文献研究	17
	医案研究	6
	医生和/或患者访谈	13
	专家咨询法	24
	头脑风暴	3
	个人经验	1
	临床病例回顾	1
条目筛选	文献研究(频次、频率、关联规则、复杂网络分析)	82
	专家咨询(重要性评分均值、变异系数、满分比、界值法、共识度)	93
	临床调查(频率、离散趋势法、相关系数法、因子分析法、克朗巴赫系数法、系统聚类、主成分分析法、判别分析法、区分度分析法、逐步回归法、偏相关的Spearman规则)	102
	核心小组讨论	10
	预调查	20
条目赋权	Logistic回归分析法	24
	专家咨询(满分比)联合因子分析	24
	Fisher判别法	10
	层次分析法	5
	因子分析法	4
	专家咨询百分权重法	4
	专家咨询(均值、满分比、变异系数、等级和)	3
	专家咨询(均值)联合因子分析	2
	专家咨询(均值归一化)	2

续表 1 中医证候诊断量表文献中研制各阶段方法学构成分布情况
Table 1 (continued) Distribution of methodological components across different development stages in literature on TCM syndrome diagnostic scales

阶段	方法	频次
诊断阈值确定	复杂熵聚堆	1
	关联规则置信度	1
	Logistic 回归法联合专家咨询法	1
	层次分析法联合因子分析法	1
	专家咨询（条目得分与总的条目均值之比）联合因子分析	1
	专家咨询（众数、均数）	1
	最大似然判别法	1
	MCMC 算法	1
	百分化法	1
	ROC 曲线	62
量表评价	自定义	1
	可行性	26
	信度（内部一致性、重测信度、分半信度、评分者信度）	89
	效度（结构效度、校标效度、内容效度、区分效度）	86
	反应度	16
	灵敏度	63
	特异度	63
	准确度	50
	阳性似然比	21
	阴性似然比	21
	漏诊率	1
	误诊率	1
	独立样本 <i>t</i> 检验	1
	Kappa 值	33
严重程度分级	四分位数法	23

注：MCMC，马尔可夫链蒙特卡罗；ROC，受试者工作特征。

3 讨论

本研究中医证候诊断量表共涉及 77 个病种，研制和评价环节相关技术方法应用多样，但尚存在以证统病模式下量表较少、部分疾病缺乏证候诊断量表、量表研制涉及关键环节方法尚无统一标准等问题。

3.1 强化以证统病与基础证候的中医证候诊断量表研究

本研究中量表均基于病证结合模式，其中以病统证占 84.15%（138/164），以证统病仅占 15.85%（26/164），表明当前量表研制主要基于以病统证形式，以证统病量表研究较少。以证统病不仅对于西医诊断不清或无法确诊的疾病可实现西医诊断逐渐明确与中医辨证治疗实施并举，比较符合中医临床实际，并且可横向综合总结不同疾病的中医诊疗理论及其特点以补充同病类证类治的不足，有利于辨证论治基本理论研究的深入^[8, 15-17]，弥补以病统证的不足。因此，需加强以证统病证候诊断量表研

究，完善证候诊断量表研究的理论体系。本研究量表涉及复合证候和基础证候。复合证候临床应用较多，但往往内涵重复，彼此之间界限模糊^[18]。目前针对基础证候研制的诊断量表较少，而基础证候介于证素和临床常见证候之间，具有界限清晰、更加实用、灵活等优势^[9, 18]。本研究中存在相同疾病证候诊断量表包含的复合证候不同、复合证候表述不明确的问题，例如肝肾不足、心脾两虚等，这可能与缺乏证候规范表述有关，也可能与各辨证体系相互交错、逻辑复杂以及病情多变有关。完善和规范证候术语的使用，可有效避免出现复合证候意思相近但表述不同的情况。此外，以基础证候为切入点研制证候诊断量表，可使证候表述更明确，也有利于证候的组合和临床应用。

3.2 加强疾病证候诊断量表研制和同一疾病不同证候诊断量表对比研究

本研究发现：1）部分疾病尚未开展证候诊断量表研究。运动系统仅 2 个量表，免疫系统疾病 3

个量表, 呼吸系统疾病中尚未见慢性阻塞性肺疾病、支气管扩张等疾病的证候诊断量表, 建议继续加强疾病的证候诊断量表研究。2) 疾病存在不同亚型的证候诊断量表。如高血压病, 有原发性高血压、正常高值血压和老年高血压的证候诊断量表, 这类量表的研制表明不同亚型疾病也可能存在差异, 未来可基于人群特征、疾病分级和分期等因素探讨同一疾病的证候差异, 进而研制相关诊断量表。3) 同一疾病存在多个证候诊断量表, 但量表间缺乏对比研究。不同团队对同一疾病开展证候诊断量表研究, 如出血性中风证候要素诊断量表^[19]和出血性中风证候要素诊断量表^[20], 两个量表在证候覆盖范围上有交集, 但后者包含证候更多。此类量表虽为证候诊断提供了量化工具, 但在临床应用时选择困难。因此, 有必要将同一疾病的证候诊断量表进行对比评价, 并参考其他权威评价规范, 制定证候诊断量表相关评价规范, 为临床应用提供参考。

3.3 丰富条目池来源、优化条目筛选并推进微观指标的应用

条目质量影响诊断效能和量表质量。本研究发现: 1) 在建立条目池时纳入古籍文献和医案的研究较少。古籍和医案蕴含历代医家丰富诊疗经验, 可为疾病诊治提供思路与方法, 应将其纳入条目池来源。2) 在条目筛选时: (1) 仅 1 个量表^[21]文献研究阶段选择关联分析, 表明对条目与证候的相关性分析的关注度不足, 需加强条目与证候的关联分析, 减少主观偏差。(2) 专家咨询采用了专家评分均值、满分比和界值法等多种方法, 其中界值法能有效减少主观差异, 快速高效筛选出核心指标, 通过明确的阈值提升标准化程度, 可作为未来研究中的主要筛选方法。(3) 临床调查数据分析方法在具体标准方面尚未统一, 如离散趋势法筛选标准差大于 1 还是 0.9, 多种统计方法联合应用时需要满足几种方法时保留该条目。这可能与数据分布及缺乏量表研制标准有关, 建议从重要性、确定性、敏感性、代表性、区分能力、稳定性和内部一致性等角度选择条目分析方法。(4) 现代测量理论在证候诊断量表中应用较少, 仅 1 个量表^[22]应用项目反应理论筛选条目。项目反应理论可从难度、区分度方面优化条目, 提升条目质量, 有必要探索不同测量理论及其联合指导下的条目筛选方法在中医证候诊断量表中的适宜性应用。(5) 部分量表缺少舌象和脉象, 这主要源于舌象和脉象难以量化, 但舌象与脉象具有中医特色, 对证候诊断具有一定价值, 量

表不包含舌脉可能导致诊断准确性下降。建议根据研究情况, 将舌脉按照“有”“无”进行分类, 分别记为“1”和“0”; 或利用舌诊仪、脉诊仪等现代科学技术实现对舌脉的客观量化。

本研究发现目前证候诊断量表多为宏观辨证。宏观辨证将中医四诊信息进行规范化、量化分析, 临床适用性强; 微观辨证从细胞、分子、基因等微观层面探索中医证候诊断的生物标志物, 有助于提高辨证的可信度、客观性和可重复性^[10, 23]。本研究仅 10 个证候诊断量表含有实验室/影像学指标, 这可能与证候尚无统一公认的实验室诊断标准^[24]有关。未来需推进分析微观指标和证候之间的相关性, 将微观辨证指标融入诊断量表中, 以提高中医辨证的客观性和国际认可度。

3.4 以诊断效能为导向优化赋权方法

条目赋权是诊断的关键环节之一。诊断条目彼此相关联系, 1 个症状与 2 个或多个症状共同出现时, 其辨证价值会大于原先构成症状, 甚至会具备新的辨证价值^[25]。本研究发现 53.05% (87/164) 量表实施了赋权, 其中 33.33% (29/87) 采用 2 种方式联合使用确定权重, 赋权方式多样, 包括主观赋权、客观赋权与主客观联合赋权。主观赋权包括专家评分法、层次分析法等; 客观赋权包括 Fisher 判别法、Logistic 判别、最大似然判别法、复杂熵聚堆等。主客观联合赋权目前常用专家咨询联合因子分析。不同赋权方式各具特点, 赋权方法的选择直接影响量表的诊断效能, 现有赋权方法如何选择是亟需解决的关键问题。有学者通过对比 Fisher 判别法和 Logistic 判别, 选择后者对肝郁化热证诊断量表^[26]、慢性心力衰竭气虚辨证量表^[27]赋权。诊断效能是评价诊断量表的核心指标之一, 以诊断效能为导向, 主客观赋权相结合, 对比不同赋权方法, 选择适宜性较高的方法, 有利于提升量表质量。此外, 当前量表赋权多针对条目, 对于不同维度是否需要赋权以及如何赋权尚无探讨。

3.5 明确诊断阈值和证候严重程度分级

本研究显示 40.85% (67/164) 的量表明确了诊断阈值, 后续研究需推进诊断阈值的确定以便于临床应用。确定诊断阈值的主要方法是 ROC 曲线, 这可能与 ROC 曲线直观易懂、不依赖于具体的类别分布有关。但 ROC 曲线无法体现阳性个体占比的影响, 在实际应用中可能存在严重的误诊问题, 对患者和医生以及医疗资源均有一定不良影响^[28]。未来仍需探索新的诊断阈值确定方法, 例如查准

率-查全率、决策曲线分析法在证候诊断标准中应用适宜性。应当从灵敏度、特异度、稳定性等多个维度探索新方法。值得注意的是,老年高血压病肾气亏虚证诊断量表^[29]、心悸(室性早搏)阴虚火旺证诊断量表^[30]采用项目组合法结合诊断阈值进行证候诊断,将定量与定性相结合,为证候诊断量表的临床应用提供了新方式。

本研究显示 14.63% (23/164) 的量表对证候严重程度进行了分级,且均采用四分位数法,如高血压病肝阳上亢证宏观信息融合诊断量表^[31]、糖尿病肾脏病肾精亏虚证诊断量表^[32]。证候严重程度分级可为疗效评价和临床用药提供依据与参考。证候严重程度的评估不仅与条目(症状/体征)出现的数量有关,也与各条目的严重程度有关,建议在证候诊断量表研制时建立条目分级体系,通过综合评分实现证候严重程度的客观分级。

3.6 以权威证候诊断标准为“金标准”进行验证

中医证候诊断量表评价需兼顾量表属性和诊断属性。1) 需从量表学角度分析信度、效度、反应度等测量学特性。本研究发现 32.93% (54/164) 的量表仅评价测量学特性。如冠心病痰浊证中医诊断量表舌脉维度的内部一致性欠佳,结构效度不同于预设理论框架,这可能与样本量小、代表性差以及舌脉的解读差异有关。2) 需从诊断角度分析准确率、灵敏度、特异度等诊断效能。本研究发现 16.46% (27/164) 的量表仅评价诊断效能,且多选用专家经验辨证为评价时的“金标准”。通过规范化研制出的证候诊断标准更加权威、客观,建议将高质量、广泛应用的中医证候诊断标准作为“金标准”,同时实验室等微观指标也可提供重要参考价值。本研究中 21.34% (35/164) 的量表同时完成了测量学特性和诊断试验评价。仅评价测量学特性可能存在诊断敏感性不足,而仅评价诊断效能则可能存在结构效度等缺陷。因此,量表研制须进行测量学特性和诊断试验双重评价,以确保量表质量和诊断应用价值。

综上所述,本研究建议深化以证统病与基础证候研究,推进微观辨证和舌脉客观化研究,探索应用现代测量理论筛选条目,注重对比多种赋权方法选择最佳赋权方案,采用 ROC 曲线明确诊断阈值并探索新方法,将应用广泛的证候诊断标准作为评价量表诊断性能的“金标准”,开展测量学特性及诊断性能双重评价,参照国际研制规范结合中医证候特点制定中医证候诊断量表研制规范/指南,明确证候程度分级以指导临床诊疗。

参考文献

- [1]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.
- [2]刘晓雨,刘志华,李京璠,等.浊毒证诊断量表的研制与权重赋值:基于德尔菲法及层次分析法[J].中医杂志,2024,65(7):684-690.
- [3]庞湃,王斌,吴深涛,等.2型糖尿病浊毒内蕴证中医诊断量表的研制及验证[J].中医杂志,2024,65(2):159-170.
- [4]李乐,郭家娟.探讨中医证候诊断量表的研究现状[J].中文科技期刊数据库(全文版)医药卫生,2023(5):168-171.
- [5]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.
- [6]孙玉娇,陈璇,尹朱丹.中医诊断量表的研究现状及文献分析[J].中医药导报,2017,23(16):111-114.
- [7]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.
- [8]李建生.基于病证结合模式重视以证统病的诊疗形式[J].中医杂志,2011,52(7):558-561.
- [9]李建生,余学庆,王至婉.病证结合模式下证候诊断标准建立的关键环节[J].中医杂志,2013,54(15):1261-1264.
- [10]曹月红,刘凤斌,王妍.中医证候量化诊断研究思路探析[J].时珍国医国药,2015,26(9):2305-2306.
- [11]刘保延.患者报告结局的测量:原理、方法与应用[M].北京:人民卫生出版社,2011:71-74.
- [12]孟凡波,李运伦,王苗苗,等.原发性高血压肝阳上亢证诊断量表信度与效度检验[J].山东中医药大学学报,2013,37(5):378-380,391.
- [13]沈真真.高血压病肾阴阳两虚证证候诊断量表的研究[D].济南:山东中医药大学,2015.
- [14]李帅.高血压病痰瘀互结证诊断量表的研制[D].济南:山东中医药大学,2015.
- [15]李涵,孙杨,张晓雨,等.以证统病:中医临床思维的回归与创新[J].世界中医药,2019,14(10):2552-2556,2562.
- [16]杨勤军,童佳兵,王传博,等.基于肺气虚证四级分度理论探讨“以证统病”的肺病管理模式[J].中华中医药杂志,2022,37(9):5467-5470.
- [17]李建生,姚建平.重视异病类证与异病类治的诊疗思维方式[J].中医学报,2011,26(9):1056-1057.
- [18]岳振松,潘战宇,于春泉,等.以单证为基础兼及复合证候和证候要素的证候命名分类新系统[J].辽宁中医杂志,2015,42(4):694-696.
- [19]徐鹏.出血性中风证候要素诊断量表条目筛选的研究[D].长春:长春中医药大学,2016.
- [20]张琳婷.出血性中风证候要素诊断量表编制研究

- [D]. 北京:北京中医药大学, 2020.
- [21] 寿鑫甜, 王禹萌, 钟森杰, 等. 慢性心力衰竭不同病理因素中西医结合诊断量表研制[J]. 湖南中医药大学学报, 2021, 41(1):7-13.
- [22] 曹月红. 多维项目反应理论框架下功能性消化不良中医辨证量表研究[D]. 广州:广州中医药大学, 2015.
- [23] 陶文娟, 陈启亮, 靖媛, 等. 传统及中西医“病证结合”的同与异[J]. 天津中医药, 2021, 38(4):409-413.
- [24] 方肇勤. 中医辨证标准制定的若干关键问题[J]. 上海中医药杂志, 2006, 40(1):11-13.
- [25] 春柳, 冯贞贞, 李建生, 等. 基于症状间关系及其对证候的贡献度初步建立新冠肺炎常见证候诊断依据[J]. 世界科学技术-中医药现代化, 2021, 23(3):874-882.
- [26] 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.
- [27] 章轶立. 慢性心力衰竭气虚证辨证与评价量表研制及其常模的初步构建[D]. 北京:北京中医药大学, 2021.
- [28] 朱映璇, 李杨, 吴疏桐, 等. ROC 曲线和 PR 曲线在临床诊断试验评估中的应用[J]. 中华预防医学杂志, 2022, 56(9):1341-1347.
- [29] 戴霞, 郭伟星. 老年高血压病肾气亏虚证规范化诊断标准的建立方法探讨[J]. 时珍国医国药, 2011, 22(6):1490-1492.
- [30] 杜雪翠. 心悸(室性早搏)阴虚火旺证证候诊断量表的研究[D]. 济南:山东中医药大学, 2013.
- [31] 李坤, 李运伦, 朱羽硕. 高血压病肝阳上亢证宏观信息融合诊断量表阈值的建立[J]. 中华中医药杂志, 2022, 37(2):693-697.
- [32] 柳红芳, 赵雪杉, 安至超, 等. 糖尿病肾脏病肾精亏虚证诊断量表的制定与验证[J]. 环球中医药, 2023, 16(6):1127-1132.

Current Status of Literature Research on Traditional Chinese Medicine Syndrome Diagnostic Scales

ZHANG Yingying¹, WANG Jiajia^{2,3}, LI Jiansheng^{2,3}, FENG Zhenzhen^{2,3}

1. The Second Clinical Medical College, Liaoning University of Traditional Chinese Medicine, Shenyang, 110847; 2. The First Affiliated Hospital of Henan University of Chinese Medicine; 3. Co-construction Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan and Education Ministry of P. R. China/ Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Henan University of Chinese Medicine

ABSTRACT Objective To analyze the current status and existing problems of research on traditional Chinese medicine (TCM) syndrome diagnostic scales, and to provide references and suggestions for the standardization of TCM syndrome research. **Methods** Literature related to the development and evaluation of TCM syndrome diagnostic scales published up to April 22, 2025, was retrieved from CNKI, Wanfang Database, VIP Database, and Chinese Biomedical Literature Database. Statistical analysis was conducted on scale names, disease names, research models, syndrome types, item pool construction methods, item screening methods, item weighting methods, diagnostic threshold determination methods, scale evaluation methods, and syndrome severity grading. **Results** A total of 207 articles were included, involving 164 TCM syndrome diagnostic scales. Among them, 138 were disease-based syndrome scales, including 8 combined TCM disease-syndrome scales (covering 3 diseases) and 130 Western medicine disease-TCM syndrome scales (covering 77 diseases), with hypertension-related scales being the most numerous (18). There were 26 syndrome-based disease scales involving 20 syndromes. In total, 91 scales focused on common or compound syndrome diagnosis, 73 on basic syndrome or syndrome element diagnosis, and 3 on animal syndrome diagnosis. Macroscopic syndrome differentiation scales numbered 154, while macroscopic-microscopic combined differentiation scales numbered 10. Scale dimensions fell into three categories: (1) subjective symptoms, objective signs, and laboratory indicators; (2) basic syndromes of compound syndromes and tongue-pulse diagnosis; (3) syndrome elements and tongue-pulse diagnosis. Tongue-pulse indicators were included in 84 scales. Item pool construction was mainly based on modern literature research (61.99%, 106/171). Item screening was often based on clinical surveys (33.22%, 102/307), expert consultation (30.29%, 95/307), and literature review (26.71%, 83/307), using over 20 methods such as mean, frequency, dispersion trend, and correlation coefficient methods. Of the scales, 53.05% (87/164) had completed item weighting, 40.85% (67/164) had determined diagnostic thresholds, and 73.78% (121/164) had undergone scale evaluation, all using expert syndrome differentiation as the "gold standard". Among these, 32.93% (54/164) only evaluated psychometric properties such as reliability, 16.46% (27/164) only performed diagnostic test evaluations, and 21.34% (35/164) evaluated both psychometric properties and diagnostic tests. A total of 14.02% (23/164) explored syndrome severity grading, all using the quartile method. **Conclusion**

Research on TCM syndrome diagnostic scales has made certain progress, but remains weak in aspects such as research models, syndrome types, microscopic indicators, scale dimensions and items, and development methodology.

Keywords traditional Chinese medicine syndrome; diagnostic scale; disease-syndrome combined scale; scale evaluation

(收稿日期: 2024-11-21; 修回日期: 2025-05-19)

[编辑: 柴倩云]