

DOI: 10. 13288/j. 11-2166/r. 2022. 03. 016

综述

中医药证据体系及评价要素研究现状分析

苏鹏丽¹, 刘骏¹, 于亚南¹, 廖星¹, 李兵¹, 许海玉², 智恺³, 党海霞³✉

1. 中国中医科学院中医临床基础医学研究所, 北京市东直门内南小街 16 号, 100700; 2. 中国中医科学院中药监管科学研究中心; 3. 中国中医科学院

[摘要] 构建体现中医药特点的证据体系是中医药临床价值评价中的热点与难点问题。通过检索中国知网、万方数据知识服务平台、维普中文科技期刊数据库、PubMed、Google Scholar 及相关网站, 收集整理国际上主要的证据分级体系以及与中医药证据体系构建相关的文献, 梳理目前中医药证据体系中涵盖的评价要素, 并按照中医药理论、人用经验和临床研究三个方面进行归纳和分析, 为中医药证据体系的构建提供参考。

[关键词] 证据体系; 评价要素; 中医药理论; 人用经验; 临床研究

构建体现中医药特点的证据体系, 科学、客观、规范地评价中医药的临床价值, 提高中药新药的研发与创制水平是中医药临床研究与评价中的关键环节。循证医学自 1999 年^[1]引入中医药领域后, 经过 20 余年的发展, 已有多个研究团队发布了依据循证医学原则制定的中医药领域的证据分级体系^[2-9], 但均未得到广泛的推广应用。2019 年 10 月 20 日, 中共中央 国务院发布的《关于促进中医药传承创新发展的意见》^[10]中提出“加快构建中医药理论、人用经验和临床试验相结合的中药注册审评证据体系”。中药新药研制多是在中医药理论指导下, 对医家的临床诊疗用药经验进行总结、提升、归纳而形成相对固定的处方, 最终转化为新药。因此, 构建中药注册审评证据体系, 必须统筹考虑中医药古典医籍、人用经验等独特证据, 并与新药临床试验要求有效衔接, 才能充分、多维度、全面地反映中医药的安全性和有效性。

本研究以“中医药”和“证据”为主题词检索中国知网、万方数据知识服务平台和维普中文科技期刊数据库, 以“evidence or evidence system”为主题词检索 PubMed 和 Google Scholar 数据库及 Grade 工作组 (<http://www.gradeworkinggroup.org/>) 和 Centre for Evidence-Based Medicine 网站 (<https://www.cebm.ox.ac.uk/>), 检索时间为从各数据库建

库至 2020 年 10 月, 去除部分重复和不相关文献, 最终纳入与国内外证据体系构建相关的文献 152 篇, 其中英文文献 49 篇, 中文文献 103 篇。系统梳理了文献中涉及的与中医药证据体系构建相关的评价要素, 分别从中医药理论、人用经验、临床研究三个方面进行整理与分析, 为构建三者相结合的证据体系提供参考。

1 国际主要的证据分级体系

证据分级的概念最早由美国社会学家 Campbell 和 Stanley 于 1963 年提出^[11], 1979 年加拿大定期健康体检工作组 (Canadian Task Force on the Periodic Health Examination, CTFPHE) 提出首个医学领域的证据分级体系^[12]。经过 50 余年的发展, 国际上先后出现了 10 余个主要的证据分级体系。目前国际上常用的证据等级划分标准有牛津大学循证医学中心 (Oxford Centre for Evidence-based Medicine, OCEBM) 发布的证据等级标准^[13]和“推荐分级的评估、制定与评价”工作组推出的 GRADE 标准^[14] (Grading of recommendation assessment, development and evaluation, GRADE)。2001 年牛津大学循证医学中心提出 OCEBM 证据分级标准, 主要用于疾病的病因、诊断、治疗、预防、预后、风险分析和经济学分析领域^[15]; 2011 年的更新版简化了原有的证据分级, 使用者可用该体系流程快速回答临床问题^[13]。2004 年, GRADE 工作组正式发布 GRADE 证据质量分级与推荐强度标准, 将证据质

基金项目: 国家中医药管理局中医药标准化项目 (GZY-FJS-2020-206)

✉ 通讯作者: danghaixia@126.com

量分为高、中、低、极低 4 个等级，并确定 5 个降级因素和 3 个升级因素。GRADE 证据分级将研究设计类型与完成质量相结合共同决定证据质量分级，主要用于疾病的诊断和治疗方面^[14]。与 GRADE 相比，OCEBM 在病因和经济学评价方面处于权威地位，但操作较复杂，缺少简洁的应用工具，令初学者难以掌握，不具应用优势。这也是国内众多学者将国际上的证据分级体系应用于中医药领域时，多采用 GRADE 的原因之一。

2 中医药领域主要的证据分级体系

2007 年，首个基于证据体的中医临床研究证据分级参考建议被提出^[2]，其后经过 10 余年的发展，逐步建立了中医古籍，中医临床和中医文献的证据等级体系（见图 1），在此过程中相继出现了 7 个相对系统的中医药证据分级体系。2009 年，基于辨证论治思想的中医治疗文献证据分级体系^[3]的提出，为筛选能体现中医临床治疗特色的文献资料提供了参考。此后，相继出现了中医临床证据分级与评分体系^[4]及根据 Delphi 法证据分级标准修订的中医文献证据分级标准^[5]。2015 年，有学者^[6]提出将中医临床证据体根据证据来源分为以

统计学与流行病学为基础的研究证据和以中医传承、专家经验或共识证据为主的理论传承证据两大类，并将理论传承证据按照诊断措施和干预措施两方面，各分为 Ia、Ib、II、IIIa、IIIb、IVa、IVb 四级七等。2018 年，有学者^[7]提出构建具有中医药特征的临床评价体系的建议，其中将名医经验和古籍经典证据分为四级，并认为最低等级证据“其他古代医案、个案报道、专家经验集”在实际应用中不推荐使用。2019 年，刘建平团队^[8]对其团队之前发布的 2007 版本^[2]进行更新。在更新的版本中一方面分别建立中医药有效性与安全性评价的证据分级标准，另一方面按照不同的研究设计类型分别提出证据质量评价标准。同年，有学者研制中医古籍防治证据评价分级量表^[9]。这些证据分级体系，依据内容主要分为中医临床研究证据分级体系和中医古籍文献证据分级体系两类。

3 中医药证据体系中的评价要素

综合分析中医药证据体系构建相关的文献，其中涉及的证据评价要素主要涵盖中医药理论、人用经验和临床研究三个方面，具体要素见表 1。

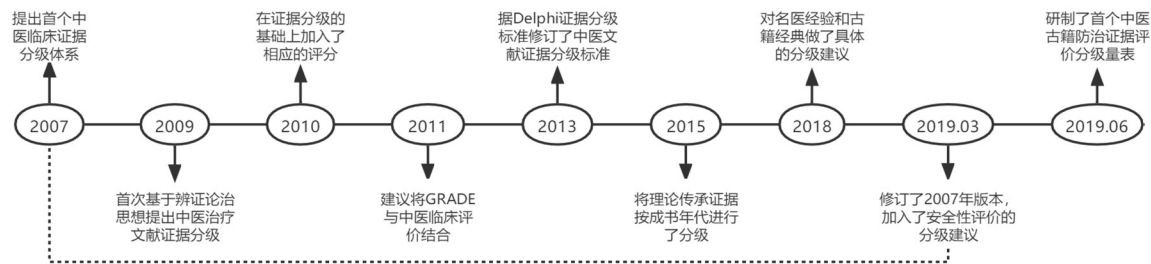


图 1 中医药证据分级体系研究进程

表 1 中医药证据体系的评价要素及特点

证据体来源	评价要素			特点
	中医药理论	人用经验	临床研究	
李敬华 2009 ^[3]	辨证论治	1) 专家经验总结；2) 医案	1) 临床研究类型； 2) 研究设计方案； 3) 临床试验完成质量	1) 将基于辨证论治的研究都列为一级证据 2) 临床研究设计方案及完成质量的评价指标完备 3) 用经验来源，经验的完整性、合理性、疗效，经验的创造性和特色来评价人用经验
王阶等 2010 ^[4]	—	1) 古代经典医籍；2) 名老中医经验；3) 医案医话；4) 国家及行业制定的标准；5) 国家统编教材	研究类型	1) 专家咨询与定量评分相结合构建证据分级标准 2) 采用经验来源、经验自身的系统性和当前使用情况评价人用经验
刘迈兰等 2010 ^[16]	—	1) 中医古籍证据（分为经典理论、临证案例）	—	1) 用经验来源和被后世医家、著作引用的频率评价古籍证据 2) 依据专家共识制订实践经验评价标准对临证案例类古籍证据分级
任玉兰等 2012 ^[17]	疾病病因病机、诊疗思路	1) 古籍载录证据；2) 专家经验；3) 个案报道	临床研究类型	1) 采用经验来源、经验完整性、记录形式、后世的传承性、创造性来评价人用经验

续表1 中医药证据体系的评价要素及特点

证据体来源	评价要素			特点
	中医药理论	人用经验	临床研究	
汪受传等 2013 ^[5]	—	1) 病例报道; 2) 中医专家共识; 3) 专家意见	1) 临床研究类型; 2) 研究质量	1) 基于 2001 年国际感染论坛提出的“Delphi 法证据分级标准”制订中医文献证据分级标准 2) 采用经验来源, 历代引用情况和当代专家意见评价人用经验
袁敬柏 2016 ^[6]	—	1) 病例报道和总结; 2) 临床经验; 3) 专家经验; 4) 专家共识; 5) 经典医籍; 6) 医论, 医案医话; 7) 教材; 8) 行业规范性文件; 9) 近现代医家专著	1) 临床研究类型; 2) 研究质量 (OCEBM 分级 2011 修订版)	1) 将理论传承证据分为诊断和干预措施, 并分别建立分级标准 2) 用成书年代, 经验来源, 被后世医家论述、著作或医案引用的情况评价理论传承证据
吕中茜等 2018 ^[18]	—	1) 名医经验; 2) 经典著作; 3) 医案医话; 4) 专家经验; 5) 个案报道; 6) 专科著作; 7) 古代医籍	1) 研究类型; 2) 研究质量 (GRADE 分级)	1) 针灸证据评级中使用了分层证据评分法 2) 采用经验来源、经验被引量、后世的传承应用评价人用经验
孟庆刚 2018 ^[19]	—	1) 经典文献; 2) 医案医话 (包括古代和现代); 3) 专家经验; 4) 共识意见	1) 研究类型	1) 明确提出研究设计类型不包括平行随机对照试验
钱静华等 2018 ^[7]	—	1) 古籍经典; 2) 名医经验; 3) 个案报道; 4) 医家著作; 5) 专家共识; 6) 中医特色结局指标	1) 研究类型; 2) 研究质量 (GRADE 分级)	1) 提出建立综合安全性评价、卫生经济学评价和对病、证候要素、主要结局指标、患者报告的结局的疗效评价为一体的多维疗效评价体系 2) 用经验的来源和被引频次评价经典医籍医案和名家经验
张磊 2019 ^[9]	疾病的诊疗依据和思路	1) 中医古籍 (分为知识性和案例性)	—	1) 用经验来源、经验自身完整性、被引率、古籍的版本和版本量、知名度评价古籍证据
刘建平 2019 ^[8]	疾病的诊疗信息, 诊疗依据和思路	1) 专家共识; 2) 病例报道; 3) 经验总结	1) 研究类型; 2) 研究质量	1) 分别建立中医药疗效与安全性评价的证据质量分级标准 2) 安全性评价分级中将临床研究数据源自的研究阶段纳入评价 3) 用经验来源评价人用经验
陈耀龙等 2020 ^[20]	—	1) 中医古籍证据; 2) 专家经验; 3) 个案报道	1) 临床研究类型; 2) 研究质量 (GRADE 分级); 3) 临床获益与风险; 4) 卫生经济学; 5) 临床可行性; 6) 临床可接受性; 7) 患者意见	1) 将经典医籍医案、名家经验列为前体证据 (待实践检验的证据), 是分级的补充 2) 证据要素中加入社会学要素

注: “—”代表此项无内容。“临床研究类型”指临床原始研究的类型, 包括干预性研究和观察性研究; “研究类型”指临床原始研究和基于原始研究做的二次研究, 如系统评价、荟萃分析等。

3.1 中医药理论

中医药理论是中医药证据体系构建的关键环节, 涵盖疾病的病因病机、组方理论、药性理论等多个层面的内容。对目前主要的证据体系分析后可以看出, 中医药理论证据的评价要素有辨证论治思想、中药方剂组方理论等。中医药理论证据较少直接用作评价要素, 主要是通过评价其他证据是否体现中医药学理论所要求的思想而作为评价要素出现的。有学者^[17]从理论分析方面评价专家经验中疾病病因病机和诊疗方案部分的分析是否合理; 有学者^[9]对中医古籍中案例类证据评价时, 考虑其是

否有按语或说明诊疗疾病的依据及思路, 包括疾病的病因病机、诊疗过程、治法治则等; 也有学者^[3]将辨证论治思想贯穿到其他证据中来对中医治疗文献进行证据分级, 如将“辨证论治的随机对照试验 (RCT), 设计及完成质量良好”列为最高级证据, 虽然其分级与其他学者的分级相比差别很大, 甚至将辨证论治的个案也列为一级证据, 但其提供的分级思路可以给我们启示; 还有学者^[21]基于中药新药审评审批提出集药方分析、病证分析、药证相符度分析和用药安全性评估为一体的传统中医药组方合理性评价体系, 总结出合理、基本合理但有部分

不合理之处、基本不合理但有部分合理之处、不合理 4 个等级的组方合理性评价结果。

3.2 人用经验

中医药人用经验是在长期临床实践中积累,具有一定规律性、可重复性的关于中医临床诊疗认识的概括总结。中医药人用经验来源广,类型多样,根据历史发展可分为古典医籍和现代人用经验两方面。

3.2.1 古典医籍 对古典医籍证据分级主要有两种方法,一种是先将中医古籍按内容或成书年代分类,然后再制定各自内容的评价指标,如将理论传承证据按成书年代分为东汉及东汉之前、晋代到清代、民国到当代的著作,然后再制定各自具体的分级标准^[6];另一种是直接将古籍证据分级,如构建集“古籍、医家、古籍知识”三者为一体的古籍评价体系^[22]。古典医籍证据评价要素方面,有学者^[2]在“中医临床研究证据分级参考建议”中加入“史料记载的疗法”以体现中医药特色,但在更新版^[8]中删除了此评价要素。有学者^[5]基于 2001 年国际感染论坛提出的“Delphi 法证据分级标准”制订了“中医文献证据分级标准”,该分级进一步细化了专家共识或专家意见的分类及定义,强化了中医专家尤其是古代名家经验在证据分级中的地位,并将“基于古今文献的中医专家共识”列为三级证据。还有学者认为理论传承证据中能体现中医临床传承与发展的经典著作的证据级别不亚于一项 RCT^[6]。

3.2.2 现代人用经验 现代人用经验方面涉及的内容有专家临床经验、病例报道、专家共识、专家观点、医案医话等。临床经验方面,刘建平^[2]在其分级体系中首先提出将临床经验作为证据评价要素,此后各专家对临床经验主要从经验的来源和经验自身考虑。从经验来源看,主要包括经验源自的医籍、医家、记录者三个方面,但对于医家的范围界定则各有不同,如认为名老中医经验指国家中医药管理局确认的名老中医的经验^[4],或认为现代名医经验为国医大师、国医名师、省级名中医、国家级学(分)会会长及副会长的经验^[18]。从经验自身看,包括经验内容是否完整以及是否经过临床研究实践验证,如将未经临床实践验证的经典医籍医案和名家经验定义为前体证据^[23]。此外,有学者^[19]推荐通过德尔菲法、名义组法、兰德公司-美国加利福尼亚大学洛杉矶分校(RAND-UCLA)专家共识法、共识会议法等对专家经验和共识意见进行评价。还有学者^[16]认为可以通过引用率和临床专家共识来评价,确定出认可的、有争议的、淘汰

的 3 个等级,其中有争议的部分再通过多中心、大样本、设计严格的临床试验进一步确证。

除临床经验外,有学者^[2]在其证据分级体系中首次将“病例报道”“专家观点”纳入证据体系内,且将其归入低级别证据组(第四级或第五级),但在更新版^[8]中只保留病例报道和经验总结,加入专家共识,但也都归入低级别证据组。有学者^[3]将基于“辨证论治的专家经验总结,记录者和专家层次高”列为一级 c 类证据,通过对记录者和专家的资质进行评分来确定其层次,将“辨证论治的个案,资料完整,疗效肯定”确认为一级 d 类证据。提出此种分级,是考虑到辨证论治是中医临床的特点,是中医认识和治疗疾病的基本原则和核心内容。

3.3 临床研究

临床研究是现代医学证据分级体系中的主要部分。在中医药临床研究评价中,采用 GRADE 方法学较常见,如建议将 GRADE 与中医药临床评价结合,形成与国际接轨的疗效评价体系^[24],或在“中医临床实践指南制定:证据质量分级”中采用 GRADE 证据分级和升降级标准^[25]等。

临床研究证据的评价要素有研究类型、研究设计和完成质量。刘建平^[8]将研究设计的质量具体化为方法学质量,并对各研究设计类型做出方法学质量评价要素及评分标准的说明,此外还加入上市后研究类型和临床前安全性评价指标,并将基于此形成的中医药临床研究证据体按照是否针对临床核心结局分为高、中、低 3 个证据等级。钱静华^[7]也在其中医药证据评价体系中加入疗效评价结局指标分级、卫生经济学评估、不良反应、安全性评价等包含中医要素在内的中医特色结局指标评价体系。临床试验还应考虑临床可行性、可接受性、患者意见、卫生经济学等因素^[20]。此外,有学者认为研究设计类型不应包括平行随机对照试验,因为此类设计虽符合中医临床个性化诊疗的基本特点,但忽视了疾病转归和干预效应^[19];还有学者^[17]认为研究设计类型不应包括系统评价和 Meta 分析,因为这些研究强调研究的同质性,难以顾及中医辨证论治及复杂干预的核心特点。

4 结语

总体来看,国内学者对中医药证据体系的研究较为分散,或侧重于某一评价要素方面,或提出证据体系构建的框架性建议。在评价要素的研究中,中医药理论证据提及较少,主要体现在辨证论治思想和中药方剂组方理论两方面。人用经验证据涵盖

古籍经典、专家临床经验、病例报道、专家共识与观点、医案医话等方面,主要侧重于对中医古籍、名家经验评价要素的考量,对于能体现中医辨证论治和“个体化”治疗特点的个案报告暂无深入研究,且对各类证据评价要素的分类和界定方面缺乏客观性。临床研究证据的评价要素主要参考国外主流的分级系统,如GRADE、OCEBM,考虑了临床研究设计类型及质量。未来将在中医药理论指导下,充分挖掘并利用中医药的人用经验,建立人用经验的综合评估方法与指南,科学、规范地收集整理与评估人用经验,并将其转化为证据支撑,与中药新药注册审批中的临床试验要求有效衔接,推进经典名方、专家临床经验、院内制剂等人用经验基础丰厚的中药处方的转化与利用,将成为构建中医药理论、人用经验和临床试验相结合的中药新药审评审批证据体系的关键环节。

参考文献

- [1]新华社. 全球首个中医药循证医学中心在中国中医科学院成立[EB/OL]. (2019-03-12)[2021-02-22]. <https://baijiahao.baidu.com/s?id=1627801396105561003&wfr=spider&for=pc>.
- [2]刘建平. 传统医学证据体的构成及证据分级的建议[J]. 中国中西医结合杂志, 2007, 27(12): 1061-1065.
- [3]李敬华. 中医治疗文献质量评价方法及示范应用研究[D]. 北京: 中国中医科学院, 2009.
- [4]王阶, 何庆勇, 姚魁武, 等. 中医临床证据分级与评分体系研究[C]//经济发展方式转变与自主创新: 第十二届中国科学技术协会年会. 福建: 中国科学技术协会学会学术部, 2010: 1-5.
- [5]汪受传, 陈争光, 徐珊, 等. 建立循证中医临床实践指南证据分级体系的构想[J]. 世界科学技术-中医药现代化, 2013, 15(7): 1488-1492.
- [6]袁敬柏. 建立适合中医临床诊疗证据评价方法的建议[J]. 中华中医药杂志, 2016, 31(4): 1146-1148.
- [7]钱静华, 郭志丽. 构建适合中医药特征的临床证据评价体系[J]. 中华中医药杂志, 2018, 33(10): 4302-4304.
- [8]陈薇, 方赛男, 刘建平. 基于证据体的中医药临床证据分级标准建议[J]. 中国中西医结合杂志, 2019, 39(3): 358-364.
- [9]张磊. 中医古籍防治证据评价分级量表的研制及应用[D]. 北京: 中国中医科学院, 2019.
- [10]中华人民共和国中央人民政府. 中共中央 国务院关于促进中医药传承创新发展的意见[EB/OL]. (2019-10-20)[2020-01-20]. http://www.gov.cn/zhengce/2019-10/26/content_5445336.htm.
- [11]陈耀龙, 李幼平, 杜亮, 等. 医学研究中证据分级和推荐强度的演进[J]. 中国循证医学杂志, 2008, 8(2): 127-133.
- [12]The periodic health examination: Canadian Task Force on the Periodic Health Examination[J]. Can Med Assoc J, 1979, 121(9): 1193-1254.
- [13]OCEBM Levels of Evidence Working Group. The Oxford Levels of Evidence 2. Oxford Centre for Evidence-Based Medicine [EB/OL][2022-1-6]. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebml-levels-of-evidence>.
- [14]BROZEK JL, AKL EA, ALONSO-COELLO P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions [J]. Allergy, 2009, 64(5): 669-677.
- [15]Centre for Evidence-Based Medicine. Oxford Centre for Evidence-Based Medicine: Levels of Evidence (March 2009) [EB/OL]. (2009-03-01)[2021-02-22]. <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009>.
- [16]刘迈兰, 吴曦, 任玉兰, 等. 借鉴循证医学实践方法研究中医古籍的思考[J]. 时珍国医国药, 2010, 21(3): 726-727.
- [17]任玉兰, 吴曦, 梁繁荣. 基于循证医学针灸临床研究证据评价体系的初步构建[J]. 辽宁中医杂志, 2012, 39(2): 205-209.
- [18]吕中茜, 郭义, 陈泽林, 等. 针灸临床实践指南制订中证据体的评估方法探索: 分层证据评分法[J]. 中国针灸, 2018, 38(10): 1115-1118.
- [19]孟庆刚. 中医临床疗效评价循证证据体系构建探析[J]. 中华中医药学刊, 2018, 36(5): 1031-1034.
- [20]李承羽, 赵晨, 陈耀龙, 等. 中医药临床指南/共识中推荐意见分级标准的制订[J]. 中医杂志, 2020, 61(6): 486-492.
- [21]金锐, 王宇光, 张冰. 试论完善传统中医药组方合理性评价技术在中药新药审评中的重要性[J]. 中药新药与临床药理, 2020, 31(9): 1117-1122.
- [22]王凤兰. 对中医古籍知识的评价思考[J]. 光明中医, 2019, 34(12): 1775-1777.
- [23]陈昊, 王艳, 陈耀龙, 等. 从GRADE系统方法学看中医药临床证据体系的构建[J]. 中华中医药杂志, 2017, 32(1): 25-28.
- [24]刘为民, 刘保延, 何丽云, 等. 引入GRADE体系制定中医药国际标准的思考与实践探析[J]. 亚太传统医药, 2011, 7(1): 1-5.
- [25]李慧, 陈耀龙, 谢秀丽, 等. 中医(中西医结合)临床实践指南制修订方法: 证据质量分级[J]. 中华中医药杂志, 2016, 31(7): 2652-2656.

(收稿日期: 2021-02-22; 修回日期: 2021-07-16)

[编辑: 柴倩云]