

· 述 评 ·

关于中医药循证临床实践指南的制定和质量评价

蒋跃绒 陈可冀

摘要 中医药循证临床实践指南(evidence-based clinical practice guidelines, ECPGs)的制定已引起行业内的重视,AGREE II 等国际通用的指南评价工具也被逐渐应用于中医药 ECPGs 的质量评价。目前制定的部分中医药 ECPGs 存在一定的方法学缺陷,应用性质量低,更新缓慢。建议建立符合中医特色和 国际通用规范的中医药循证指南制定技术规范,加强中医药循证指南质量的评价,重视中医药循证指南的定期更新和推广应用。

关键词 中医药;循证临床实践指南;质量评价

Development and Quality Evaluation of Evidence-based Clinical Practice Guidelines of Chinese Medicine JIANG Yue-rong and CHEN Ke-ji Cardiovascular Diseases Centre, Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing (100091)

ABSTRACT More attentions have been paid to the development of evidence-based clinical practice guidelines (ECPGs) of Chinese medicine (CM). International guideline evaluation instruments such as Appraisal of Guidelines for Research and Evaluation (AGREE II) has been gradually applied in ECPGs quality evaluation of CM. Nowadays, there are some certain methodological defects in partial ECPGs of Chinese medicine, with relatively low applicability and slowly update. It is suggested to establish technical specifications of CM-ECPGs in accordance with the characteristics of CM and international general specification, strengthen the quality evaluation of CM-ECPGs, attach great importance to the regularly update as well as popularization and application of CM-ECPGs.

KEYWORDS Chinese medicine; evidence-based clinical practice guidelines; quality evaluation

采用循证方法制定指南已经成为国际上临床实践指南制定的主流趋势与共识。指南的潜在获益取决于指南本身质量的高低。在指南开发过程中,恰当的方法学和严格的制定策略对于推荐和建议的成功实施是非常重要的^[1]。指南的质量可能参差不齐,有时甚至达不到通用标准。为提高指南制定的质量,指南研究与评价工具也应运而生。

1 指南研究与评价工具 2003 年指南研究与评价工具 [The Appraisal of Guidelines for Research and Evaluation (AGREE) Instrument]^[2] 制定发布,其中指南的质量定义为“对指南制定的潜在偏倚得以充分考虑,以及对指南推荐意见具有内部真实性、外部真实性和实施可行性的信心”^[3]。2005 年国内学者对 AGREE 工具进行了翻译,将其正式引

入中国^[4]。为进一步提高 AGREE 的科学性及可行性,由 AGREE 协作网的部分成员组建的 AGREE Next Steps 协会对 AGREE 工具开展了修订工作,并于 2009 年发布了 AGREE II^[5],其与 AGREE 的比较见表 1。2012 年,国内学者谢利民等^[6]对 AGREE II 的条目进行了翻译。

AGREE II 由 1 个用户手册、6 个领域(23 个条目)和 2 个总体评估条目组成。每个领域针对指南质量评价的一个特定问题。其适用对象包括:卫生保健提供者、指南制定者、卫生决策者和相关教育工作者。AGREE II 中推荐评价指南的人数至少为 2 人,最好为 4 人。AGREE II 每个条目的评分为 1~7 分,1 分表示指南完全不符合该条目,7 分代表指南完全符合该条目,2~6 分代表指南不完全符合该条目,得分越高说明该条目符合程度越高。

AGREE 自发布以来,已得到多个卫生保健机构的认可,成为国际通用的评价工具。但也存在以下问题:(1) 没有衡量不同领域和条目所占的权重;(2) 更

作者单位:中国中医科学院西苑医院心血管病中心(北京 100091)

通讯作者:蒋跃绒, Tel: 010-62835342, E-mail: jiang_yuerong@163.com

DOI: 10.7661/CJIM.2016.01.0011

表 1 AGREE 与 AGREE II 条目比较

AGREE 条目	AGREE II 条目
领域一 范围和目的	
1. 明确阐述指南的总目的	无变化
2. 明确阐述指南所涵盖的临床问题(clinical questions)	明确阐述指南所涵盖的卫生问题(health questions)
3. 明确阐述指南所要应用的患者	明确阐述指南所要应用的人群(患者、公众等)
领域二 利益相关人员	
4. 指南制定小组包括所有相关专业的人员	无变化
5. 考虑到患者的观点和选择	考虑到目标人群(患者、公众等)的观点和选择
6. 明确规定指南的适用者	无变化
7. 指南指导最终用户的使用	删除该条目,合并入条目 19
领域三 制定的严谨性	
8. 用系统的方法检索证据	内容无变化,调整序号为 7
9. 明确阐述选择证据的标准	内容无变化,调整序号为 8 新条目 9. 清楚地描述证据体的优势和不足
10. 明确阐述形成推荐意见的方法	无变化
11. 形成推荐意见时考虑了健康获益、副作用及风险	无变化
12. 推荐意见和支持证据之间有明确的联系	无变化
13. 指南在发表前经过专家的外部评审	无变化
14. 提供指南更新的流程	
领域四 表达的明晰性	
15. 推荐建议明确不含糊	无变化
16. 明确列出针对某一情况的不同选择	明确列出针对某一情况或卫生问题的不同选择
17. 主要推荐意见明晰易辨	无变化
领域五 应用性	
18. 指南提供了应用的配套工具	指南提供了推荐意见如何应用于实践的建议和(或)配套工具(调整序号为 19)
19. 讨论了可能阻碍推荐意见应用的因素	指南描述了促进或阻碍应用的因素(调整序号为 18)
20. 考虑了应用推荐建议时潜在的成本投入	考虑了应用推荐建议时潜在的资源投入
21. 指南提供了用于监测和(或)审计目的的关键评价标准	指南提供了监测和(或)审计标准
领域六 编辑独立性	
22. 指南在编辑上独立于赞助单位	赞助单位的观点不影响指南内容
23. 指南记录了制定小组成员的利益冲突	指南记录并考虑了制定小组成员的利益冲突

像是临床指南制定的框架;(3)只评估了不同科目的报告与否而没有评估建议内容的有效性。因此,应该研发一种比 AGREE 更好的评价工具或对 AGREE 进行进一步完善。

2 循证临床实践指南制定的一般流程

国际著名的指南开发平台如英国国家卫生与临床优化研究所(National Institute for Health and Clinical Excellence, NICE)和苏格兰学院间指南网络(Scottish Intercollegiate Guidelines Network, SIGN)分别发布了指南制订手册(the Guideline Manual)^[7]和指南制订者手册(SIGN50:a Guideline Developer's Handbook)^[8]。循证临床实践指南的制定具有严格程序,主要包括:(1)确定指南的主题和目的;(2)成立指南制定小组;(3)严格制定指南(系统检索文献、评价证据、根据证据的级别和强度提出建议);(4)撰写指南;(5)临床指南的修改、评审及更新。世界卫生组织(WHO)也发布了指南制定手册(WHO Handbook for Guideline Development)^[9],对指南制定流程包括指南设计、指南小组的设立、利益声明和管理、问题提出和结局选择、证据检索和合成、

证据评价、推荐建议的形成、指南开发和发表、实施和评价等均进行了详细说明。

指南制定者可参考或采用推荐等级的评估、制定与评价(Grades of Recommendation, Assessment, Development, and Evaluation, GRADE)工作组制定的 GRADE 标准,制定统一的证据质量级别与推荐强度的划分评级标准。SIGN 指出,为确保指南的质量,指南在发表前还要反复听取使用者和患者的意见,以及公认的有权威性的机构(学会)和专家的意见并得到认可,在此基础上完成对指南的修改和定稿工作^[8]。

3 中医药循证临床实践指南制定和评价的现状

3.1 中医药循证临床实践指南制定的概况 近年来,中医药循证临床实践指南的制定逐渐引起行业内的重视。2007 年,中国中医科学院与 WHO 西太区达成合作意向,开发一套中医药循证临床实践指南,涉及 28 种有中医诊疗特色和优势的疾病。2011 年,由中国中医科学院和 WHO 西太区合作开发的第一批中医药循证临床实践指南先后出版,分别为《中医循证临床实践指南·中医内科》^[10]、《中医循证临床实践

指南·专科专病》^[11]和《中医循证临床实践指南·针灸》^[12]。其中中医内科分册包括原发性支气管肺癌、慢性稳定性心绞痛、脑梗死、2 型糖尿病、感冒、类风湿关节炎、偏头痛、失眠症、原发性骨质疏松症、慢性胃炎、血管性痴呆、高血压病、慢性阻塞性肺疾病、慢性乙型肝炎、艾滋病、IgA 肾病、再生障碍性贫血、单纯性肥胖、抑郁症和甲型 H1N1 20 种内科病的中医药循证临床实践指南。专科专病分册介绍了年龄相关性黄斑变性(湿性)、特发性皮炎、寻常型银屑病、神经根型颈椎病、慢性前列腺炎、慢性盆腔炎、小儿肺炎、桡骨远端骨折 8 种疾病的中医药循证临床实践指南。针灸分册则介绍了针灸循证性诊疗指南制订的方法以及带状疱疹、贝尔面瘫、抑郁症、中风假性球麻痹、偏头痛 5 种疾病的循证性针灸临床实践指南。2008 年国家科技部立项的“中药新药临床评价研究技术平台(南京)建设”课题发布了“小儿急性上呼吸道感染中医诊疗指南”^[13]、“小儿病毒性肺炎中医诊疗指南”^[14]。2014 年世界中医药联合会心血管病专业委员会发布了急性心肌梗死(真心痛)中医临床实践指南^[15]。这些指南多采用了国际上普遍采用的循证性临床实践指南的制定方法,对目标人群、指南制定小组的组成、文献的检索和评价、证据分级和推荐意见的形成、指南的起草、指南形成的审定和专家评价等循证性指南制定的核心内容进行了详细描述。

3.2 中医药循证临床实践指南的质量评价现状
近年来,AGREE 工具已被逐渐应用于中医临床实践指南的质量评价。

宇文亚等^[16]采用 AGREE 工具评价国内第一批 28 个中医药循证临床实践指南的质量,全面评估显示,强烈推荐使用的指南仅有 9 个,推荐(补充或改进)的指南有 5 个,11 个指南不能确定是否推荐使用。其编辑独立性质量最高,平均得分 81.46%;其次为制定的严谨性,平均得分 80.95%;范围与目的平均得分 79.96%;清晰性与可读性平均得分 70.88%;参与人员平均得分 61.28%;应用性质量最低,有 23 个指南的应用性得分是 0 分,平均得分 27.09%。说明中医药循证临床实践指南的应用性质量很低,在以后的指南研制中,要注重该方面质量的提高。

而采用 AGREE II 工具对抑郁症、带状疱疹、中风假性延髓麻痹、偏头痛、贝尔面瘫 5 个针灸循证临床实践指南的质量进行评价发现,不同评价小组得出的结论有所不同。如房繁恭等^[17]的评价表明针灸临床实践指南的数量不多,整体质量较好,其中编辑独立性质量最高,平均得分 97.9%;其次为清晰性与可读性,平

均得分 83.3%;参与人员平均得分 78.1%;制定的严谨性平均得分 75.6%;范围与目的平均得分 68.1%;应用性质量平均得分 62.5%;5 个指南的总体评价分数均为 6;全面评估结果均为推荐。而陈昊等^[18]的评价则认为当前针灸领域的专科循证实践指南存在一定的方法学缺陷,指南质量不高,5 个指南 AGREE II 的平均得分情况为:范围与目的 55%,参与人员 27%,制定的严谨性 4%,表达的明晰性 55%,应用性 4%,编辑独立性 1%。对相同指南的质量采用相同评价工具得出截然不同结论的现象,推测与不同评价者对指南内容及 AGREE 工具的理解不同有关。提示一方面应加强对 AGREE II 条目和用户手册的培训,另一方面应组织相关专家参考国际通用指南评价工具(如 A-GREE),结合中医药特色,制定出符合中医药特点的临床指南评价工具,对现有的循证中医药临床指南进行系统评价。且评价中医药循证指南时,注意与指南制定者沟通,加深对指南内容的理解。

4 问题与建议

4.1 建立符合中医特色和国际通用规范的中医药循证指南制定技术规范
中医药循证临床实践指南的制定,一方面要求符合中医学整体观和个体化辨证论治的学术特色、体现中医临床最新进展和实际诊疗水平,另一方面要符合国际临床实践指南制定的规范,使之便于与国际接轨。

(1)在指南针对的临床问题方面,要注意中医辨病和辨证的关系。应充分考虑到中医临床整体观和个体化辨证论治的诊疗特色。可根据提出的临床问题,采用病证结合的诊治模式、辨病为主的诊治模式及辨证为主的诊治模式。如针对冠心病、高血压病等现代疾病的中医和中西医结合指南,可采用病证结合的诊治模式;针对西医生使用的中成药指南,可采用辨病为主、辨证为辅的诊治模式;针对某一中医病证的指南,可采用辨证为主的模式。

(2)在证据的合成和评价方面,要建立符合中医特色的证据质量评价和分级以及推荐意见的分级体系。证据的检索与评价、推荐意见的形成与推荐强度的确定是循证临床指南制定的重要步骤^[8]。中医药循证证据的检索与评价也应通过制定科学的检索策略进行全面文献检索,并对检索到的文献进行筛选以形成支持指南的临床研究证据,然后在此基础上根据不同的研究类型形成具有不同强度的推荐意见。中医学在长期临床实践中形成了重视经典古籍和经验传承的特色,大量古籍文献、医案医话、名家经验中蕴含了丰富的诊治经验,但这些文献归属于专家经验,循证证据

级别较低。近年来国内中文期刊发表的大量临床研究报告中,高质量的系统评价和随机对照临床研究相对匮乏,证据级别相对较低。刘建平教授^[19]2007 年提出了关于传统医学证据体的构成及证据分级的建议,其中将专家经验按是否经系统临床研究验证和长期广泛应用赋予不同的证据分级。可通过专家共识等进一步优化中医临床证据分级体系。在形成推荐意见和确定推荐强度时,应加强中医古籍的甄别整理,对古籍中记载、广泛应用至今且当代专家达成共识者,可适当提高证据级别和推荐强度。

(3) 提高中医临床研究的质量,增加高水平证据的来源。针对国内大量临床研究报告质量差,能够真正进入指南推荐的高级别证据少的问题,中国医师协会中西医结合医师分会和中国中西医结合学会循证医学专业委员会制定了《中医药与中西医结合临床研究方法指南》^[20],以期进一步规范中医临床试验设计并加以实施,提高中医药临床研究的质量。同时,还应重视中医证候的规范化,加强对中医治疗措施安全性的评价等。

(4) 注重指南的实用性,提出证据不足时形成指南建议的方法。中医药循证临床指南既要符合循证方法学要求,又要体现辨证论治的特色和优势,突出实用性。对个体化辨证论治的临床疗效评价,可考虑纳入实用性随机对照试验、单病例随机对照试验、队列研究、真实世界研究等。在证据不足时,可采用基于循证证据和专家共识相结合的方法,客观、科学地提取专家经验,提出符合临床实际的建议。

4.2 加强中医药循证指南质量的评价 指南的质量决定了指南的科学性和实用性。目前的中医药循证指南在指南的范围和目的、参与人员、制定的严谨性、清晰性与可读性、编辑独立等方面质量较高,但是指南应用性质量普遍很低,26 个指南中有 23 个指南的应用性得分是 0 分^[16]。AGREE II 工具只针对指南制定的方法和框架进行评价,对中医药循证指南的质量评价,除重视对考虑指南制定的方法学评价之外,还要考虑推荐建议的真实性和可行性、以及这些推荐建议的可能的获益、危害和花费等。对指南使用者和评价者要进行培训,加强对指南相关知识的学习,有利于指南评价者深入理解指南,便于临床应用评价的开展。

4.3 重视中医药循证指南的定期更新 对于指南的更新,澳大利亚国家卫生与医学研究委员会(National Health and Medical Research Council, NHMRC)和英国国家卫生与临床优化研究所^[7]的指南手册均明确要求更新周期为 3 年。临床上比较重要的

国际指南更新周期则一般为 2 年。如果评价周期过长,则影响指南的更新和临床指导价值。目前,我国中医药循证指南存在制定过程时间较长,更新缓慢等问题,需要加以改进。应清楚说明指南更新的程序,包括更新时间表、接受定期更新的文献检索模板、更新指南相关人员职责、更新方法等。对同一种疾病存在多个中医指南的情况,应对各指南推荐建议进行对比分析,开展比较研究,对其中推荐建议存在争议的领域,在原指南的基础上,不断纳入新证据,对指南进行更新修订。

4.4 加强中医药循证指南的推广应用 我国在循证中医药临床实践指南研制中往往只注重制定,而推广应用重视不足。指南只有应用于临床实践才能体现其价值,因此,应加强指南应用性方面的研究。指南在发布前,应进行试行研究。对已制定的质量较好的中医药循证指南,可通过学术会议、专业杂志、网络、举办基层培训班等推广应用,请资深专家对指南进行推广讲解,将指南与日常临床工作结合,用于医疗质量控制和临床路径管理,在实践中发现不足,并不断修订、完善。

参 考 文 献

- [1] Burgers JS, Fervers B, Haugh M, et al. International assessment of the quality of clinical practice guidelines in oncology using the Appraisal of Guidelines and Research and Evaluation Instrument[J]. J Clin Oncol, 2004, 22(10): 2000-2007.
- [2] AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project[J]. Qual Saf Health Care, 2003, 2(1): 18-23.
- [3] AGREE History. <http://www.agreetrust.org/about-agree/agree-history> [OL]. 2013-08-01.
- [4] 詹思延. 临床指南研究与评价工具简介[J]. 中国循证儿科杂志, 2007, 2(5): 375-377.
- [5] 10 Years of AGREE (2003-2013). <http://www.agreetrust.org/2013/10/10-years-of-agree-2003-2013/> [OL]. 2014-09-21.
- [6] Xie LM, Wang WY. A brief introduction to Appraisal of Guidelines for Research and Evaluation II [J]. J Chin Integr Med, 2012, 10(2): 160-165.
- [7] NICE. The Guideline Manual. National Institute for Health and Clinical Excellence, 2009 [OL]. <http://www.nice.org.uk>.
- [8] SIGN50. A Guideline Developer's Handbook. Scottish Intercollegiate Guidelines Network, 2008

- (收稿:2014-11-30 修回:2015-11-04)

(以上名单按姓氏笔画为序)