

中国护理临床实践指南质量评价标准的信效度评价



周英凤¹, 王强², 胡雁¹, 邢唯杰¹, 章孟星¹, 钟婕¹, 傅亮¹

1. 复旦大学护理学院, 复旦大学JBI循证护理合作中心, 上海市循证护理中心(上海 200032)
2. 国家卫生健康委医疗管理服务指导中心(北京 100044)

【摘要】 目的 评价中国护理临床实践指南质量评价标准的信度和效度。方法 由不同评价者采用中国护理领域临床实践指南质量评价标准和 AGREE II 工具对两份护理领域内的临床实践指南分别进行质量评价, 采用组内相关系数测量评价工具的信度, 采用一致性检验测量评价工具的效度。结果 5名评价者采用中国护理领域临床实践指南质量评价标准对艾滋病指南及静脉输液指南进行评价, 组内相关系数(ICC)分别为 0.979 及 0.890, 采用 AGREE II 对艾滋病指南及静脉输液指南进行质量评价, 组内相关系数(ICC)分别为 0.957 及 0.887。评价者采用中国护理领域临床实践指南质量评价标准及 AGREE II 对艾滋病指南评价的一致性系数为 0.880, 呈高度一致性, 对静脉输液指南评价的一致性系数为 0.886, 也呈高度一致性。结论 中国护理领域临床实践指南质量评价标准具有较好的信度和效度, 可作为护理领域内指南的质量评价工具, 在使用该工具前, 应对该用具进行详细解读, 或进行统一培训, 以减少评价者偏倚。

【关键词】 临床实践指南; 循证护理; 质量评价标准; 信效度评价

Evaluation of reliability and validity on quality criteria for clinical nursing practice guidelines in China

ZHOU Yingfeng¹, WANG Qiang², HU Yan¹, XING Weijie¹, ZHANG Mengxing¹, ZHONG Jie¹, FU Liang¹

1. School of Nursing, Fudan University; Centre for Joanna Briggs Institute Centre of Excellence of Evidence-based Nursing, Fudan University; Shanghai Evidence-based Nursing Center, Shanghai, 200032, P.R.China

2. National Center for Medical Service Administration, National Health Commission, Beijing, 100044, P.R.China

Corresponding author: WANG Qiang, Email: jason2019@sina.cn; HU Yan, Email: huyan@fudan.edu.cn

【Abstract】 **Objective** To evaluate the reliability and validity on quality criteria for clinical nursing practice guidelines. **Method** Five appraisers were included to assess two clinical nursing practice guidelines using the quality criteria for clinical nursing practice guidelines and AGREE II. ICC and coefficient of consistency were adopted to assess the reliability and validity of the appraisal tools. **Results** ICC for two guidelines of AIDS and infusion by the quality criteria for clinical nursing practice guidelines were 0.979 and 0.890. ICC for two guidelines of AIDS and infusion by AGREE II were 0.957 and 0.887. The coefficient of consistency for guidelines of AIDS and infusion by the two appraisal tools were 0.880 and 0.886, both were high consistent. **Conclusion** The quality criteria for clinical nursing practice guidelines had good reliability and validity and can be the appraisal tool for quality evaluation of clinical nursing practice guidelines. It is strongly recommended that the users should understand the appraisal tool or should be trained to decrease the bias.

【Key words】 Clinical practice guideline; Evidence-based nursing; Quality criteria; Reliability and validity

20 世纪 80 年代临床实践指南制订工作开始在全球范围开展, 其作为指导医疗护理实践活动的重

要工具和决策依据, 受到了研究者、实践者和管理者的极大关注, 国际国内临床实践指南的数量不断增加^[1, 2]。随着循证医学的发展及对指南构建方法学的影响, 美国医学研究所(institute of medicine, IOM)于 2011 年更新了指南的定义: 综合考虑基于系统评价的证据并权衡不同干预措施的利弊, 形

DOI: 10.7507/1672-2531.201807027

通信作者: 王强, Email: jason2019@sina.cn; 胡雁, Email: huyan@fudan.edu.cn

成为患者提供最佳保健服务的推荐意见^[3]。基于循证证据的临床实践指南也被视为无偏倚的卫生标准,能够规范医疗护理行为,降低医疗费用,提高医疗服务质量^[4]。然而,现有指南的制订因存在方法学缺陷导致指南质量参差不齐^[5,6]。为了提高临床实践指南的质量,推动指南的发表和传播,国际上推出了指南质量评价工具(The Appraisal of Guidelines for Research & Evaluation, AGREE II),但该工具缺乏对指南内容真实性的评价,各条目权重不清及个别条目不完全适用于中国指南^[7]。因此,2017年9月受国家卫生健康委医疗管理服务指导中心的委托,由复旦大学循证护理中心牵头联合全国护理行业专家,在复旦大学王吉耀教授团队制作的“中国医疗领域临床实践指南质量评价标准(草案)”的基础上,结合护理学科的特点和循证护理实践发展要求,制定了我国护理领域临床实践指南质量评价标准。本研究旨在评估该指南质量评价标准的有效性和可靠性,为我国护理领域内指南的质量管理提供标准和评价依据。

1 对象与方法

1.1 评价对象

本研究选取护理领域内艾滋病临床护理实践指南及静脉输液临床实践指南作为评价对象。艾滋病临床护理实践指南由上海市公共卫生临床中心与复旦大学 JBI 循证护理合作中心共同构建,并于2014年3月推出,其指南涵盖了艾滋病高效联合抗反转录病毒治疗依从性及艾滋病相关的12大症状方面护理的证据和推荐意见^[8]。静脉输液临床实践指南是由美国静脉输液协会于2016年推出的,涵盖了静脉输液相关人员、血管通路装置、输液治疗及相关并发症等9大方面的证据和推荐意见^[9]。

1.2 评价人员

采用目的抽样法,遵循专业性和科学性原则,选取相关领域专家组成质量评价小组。遴选标准:具有硕士及以上学历、受过循证方法学专业教育、参与或主持过护理领域临床实践指南的制定。评价小组共包括5位成员,平均年龄 $34.8+10.1$ 岁,3位具有博士学位,2位具有硕士学位,1位具有高级职称,1位具有副高职称。所有专家均参与过护理领域临床实践指南的制定,具有临床实践指南质量评价的研究经历。

1.3 评价工具

1.3.1 中国护理领域临床实践指南质量评价标准

该标准经过两轮团体焦点访谈法,包括5个维度的15个条目,从指南制定过程的科学性/严谨性、推荐意见有效性/安全性、经济性、指南的可用性/可行性、指南制定过程的利益冲突5大方面对护理领域临床实践指南质量进行评价。每个条目均采用0~5分制进行评分(0分为完全不符合,5分为完全符合),并考虑到不同条目对指南质量的影响程度设置了不同的权重,总分100分,60分以上为强烈推荐,31~59分为弱推荐,30分以下为不推荐。我国护理领域临床实践指南质量评价标准与AGREE II相比:①在评价条目上,由23条精简到15条,实际评价中更简单高效;②在内容上考虑到临床实践者更关注有效性和安全性,因此将其作为单独的评价领域,并强调并纳入中国本土化研究证据及考虑护理领域证据的多元性;③在评分上由7分制改为5分制,根据不同条目的重要性不同设置了相应的权重;④在评分标准上,采用逐级评分的方法,将定性描述和定量评价相结合,具有较好的可操作性^[10,11]。

1.3.2 AGREE II AGREE II由英国AGREE国际协作组织于2003年制定,在2009年发布了第二版^[12],包括6个领域23个条目,内容涵盖范围和目的、参与人员、制定的严谨性、清晰性与可读性、适用性、及编撰的独立性。每个条目均采用1~7分进行评价(1分为非常不同意,7分为非常同意),然后计算每个领域的得分,并进行标准化处理。最后确定是推荐、修改后推荐或不推荐。

1.4 评价方法

将“艾滋病临床护理实践指南”、“静脉输液临床实践指南”与两个指南评价工具一起发给评价者,为避免AGREE II对中国护理领域临床实践指南质量评价标准的影响,要求所有成员首先采用中国护理领域临床实践指南质量评价标准对两份指南进行质量评价,然后再采用AGREE II对两份指南进行质量评价。所有成员均独立完成2种质量评价标准对指南的评价。

1.5 统计方法

采用SPSS 22.0进行统计分析,采用均数、标准差、变异系统来描述各评价条目在评价人员之间的集中和离散趋势。中国护理领域临床实践指南质量评价标准的信度采用组内相关系数(intraclass correlation coefficient, ICC)对同一份指南在不同评价人员间的内部一致性进行测定,当ICC值 >0.70 时代表评价者间具有非常好的一致性。中国护理领域临床实践指南质量评价标准的效度采用与

AGREE II 比较的一致性检验进行测定, 由于中国护理领域临床实践指南质量评价标准和 AGREE II 在维度、条目数及各条目的评分标准均不同, 因此, 借鉴 AGREE I 标准化处理方式, 即对每位评价者的评分进行标准化处理: 标准化得分率=(评价的实际得分-最小可能得分)/(最大可能得分-最小可能得分)×100%, 然后进行一致性检验, 判断两个评价工具之间的一致性, 当一致性系数>0.75 代表两个评价工具间具有非常好的一致性^[13,14]。

2 结果

2.1 两个评价工具对两份指南的评价情况

首先采用中国护理领域临床实践指南质量评价标准对艾滋病指南进行评价, 结果显示, 评价者对 3 个条目评价的不一致性较高, 变异系数>0.3; 对静脉输液指南的评价结果显示, 评价者对 10 个条目评价的不一致性较高, 变异系数>0.3。不一致性主要集中在证据的质量评价、证据体的分级、形成推荐意见的方法、推荐等级、推荐方案的有效性、安全性及推荐意见的经济性等条目上。评价者再采用 AGREE II 对两份指南进行质量评价, 结果显示, 在艾滋病指南的评价中, 评价者对 5 个条目

的评价存在不一致, 变异系数>0.3; 在静脉输液指南中, 评价者对 13 个条目评价的不一致性较高, 变异系数>0.3。具体见表 1。

2.2 中国护理领域临床实践指南质量评价标准的信效度测定

5 位评价者采用中国护理领域临床实践指南质量评价标准对艾滋病指南及静脉输液指南进行评价, 评价者组内相关系数 (ICC) 分别为 0.979 及 0.890。采用 AGREE II 对艾滋病指南及静脉输液指南进行质量评价, ICC 分别为 0.957 及 0.887。采用中国护理领域临床实践指南质量评价标准及 AGREE II 对艾滋病指南评价的标准化得分率分别为 83%、83.2%、80%、77%、71.8% 及 81.9%、87.7%、79.7%、73.2%、63%, 两个评价工具用于评价该指南之间的一致性系数为 0.880, 呈高度一致性。采用中国护理领域临床实践指南质量评价标准及 AGREE II 对静脉输液指南评价的标准化得分率分别为 68.6%、56.4%、57.4%、43.4%、46.8% 及 75.4%、50.0%、66.7%、51.4%、47.8%, 两个评价工具用于评价该指南之间的一致性系数为 0.886, 亦呈高度一致性。

表 1 两个评价工具对两份指南的评价结果

评价工具及评价条目	艾滋病指南			静脉输液指南		
	均数	标准差	变异系数	均数	标准差	变异系数
中国护理领域临床实践指南质量评价标准						
科学性/严谨性						
1. 指南制定小组由相关的多学科团队组成。	4.80	0.45	0.09	4.20	0.45	0.11
2. 说明了指南的背景、目的和应用对象	5.00	0.00	0.00	4.60	0.55	0.12
3. 考虑证据的多元性, 列出正确的文献检索策略, 开展全面的证据检索	9.20	1.10	0.12	7.60	1.67	0.22
4. 对检索到的证据进行质量评价, 对证据/证据体进行分级	10.00	0.00	0.00	8.40	2.61	0.31
5. 说明了从证据到形成推荐意见的方法	7.20	1.10	0.15	1.60	0.55	0.34
6. 列出了推荐意见的推荐等级	5.00	0.00	0.00	1.20	0.45	0.38
7. 指南发表前经过了临床试点和外部专家评议	4.40	0.55	0.12	2.00	0.71	0.36
8. 提供了指南的更新计划	1.68	0.33	0.20	0.56	0.22	0.39
有效性/安全性						
9. 推荐方案的有效性: 同一临床问题, 如有备选方案, 应列出。必要时列出效应大小的具体数据	6.00	1.41	0.24	5.20	2.28	0.44
10. 推荐方案的安全性: 推荐意见考虑了方案的不良作用, 列出安全性相关数据	4.40	2.19	0.50	4.80	2.68	0.56
经济性						
11. 推荐意见考虑了卫生经济学问题	1.00	0.71	0.71	0.80	0.45	0.56
可用性/可行性						
12. 推荐意见明确不含糊, 容易理解	4.80	0.45	0.09	4.80	0.45	0.09
13. 指南容易推广	8.00	0.00	0.00	7.20	2.28	0.32
14. 指南检索和评估了中国研究的证据	2.52	0.78	0.31	0.00	0	—
利益冲突						
15. 指南制定过程有“利益冲突声明”	5.00	0.00	0.00	1.60	0.55	0.34

续表 1

评价工具及评价条目	艾滋病指南			静脉输液指南		
	均数	标准差	变异系数	均数	标准差	变异系数
AGREE II						
范围和目的						
1. 明确描述指南的总目的	6.80	0.45	0.07	5.60	1.34	0.24
2. 明确描述指南涵盖的卫生问题	5.80	1.10	0.19	5.40	1.67	0.31
3. 明确描述指南适用的人群(患者、公众等)	6.80	0.45	0.07	5.80	1.30	0.23
参与人员						
4. 指南开发小组包括了所有相关专业的人员	6.40	1.34	0.21	5.20	1.79	0.34
5. 收集目标人群(患者、公众等)的观点和选择意愿	6.00	1.00	0.17	0.80	0.45	0.56
6. 明确规定指南的使用者	7.00	0.00	0.00	6.20	1.10	0.18
严谨性						
7. 应用系统方法检索证据	5.80	0.84	0.14	4.80	0.84	0.17
8. 清楚描述选择证据的标准	6.00	1.00	0.17	3.80	0.84	0.22
9. 清楚描述证据的强度和局限性	5.20	2.05	0.39	3.00	2.35	0.78
10. 清楚描述形成推荐建议的方法	4.80	1.30	0.27	3.00	2.00	0.67
11. 形成推荐建议时考虑了对健康的益处、副作用以及危险	4.80	1.64	0.34	4.40	1.14	0.26
12. 推荐建议和支持证据之间有明确的联系	6.80	0.45	0.07	5.20	2.49	0.48
13. 指南在发布前经过外部专家评审	7.00	0.00	0.00	6.00	2.24	0.37
14. 提供指南更新的步骤	4.00	0.71	0.18	2.80	0.84	0.30
清晰性						
15. 推荐建议明确, 不含糊	6.40	0.55	0.09	6.20	0.84	0.13
16. 明确列出不同的选择或卫生问题	5.60	1.67	0.30	6.20	0.84	0.13
17. 容易识别重要的推荐建议	5.80	1.64	0.28	6.80	0.45	0.07
适用性						
18. 指南描述了应用时的促进和阻碍因素	3.60	1.14	0.32	2.20	1.30	0.59
19. 指南提供应用推荐建议的意见和/或工具	6.00	1.22	0.20	4.80	1.79	0.37
20. 指南考虑了推荐建议应用时潜在的相关资源	4.80	1.30	0.27	3.80	0.84	0.22
21. 指南提供了监督和/或审计标准	6.00	1.73	0.29	6.20	0.84	0.14
独立性						
22. 赞助单位的观点不影响指南的内容	6.80	0.45	0.07	2.40	0.89	0.37
23. 指南开发小组成员的利益冲突要记载并公布	1.20	0.45	0.38	2.80	1.30	0.46

3 讨论

3.1 中国护理领域临床实践指南质量评价标准具有较好的信效度

临床实践指南作为医务人员临床决策的指导性工具, 有助于医护人员了解当前的最佳证据及指南的推荐意见是否能帮助其解决所面临的具体临床问题^[15]。因此, 临床实践指南质量的高低对医护人员做出科学、合理的临床决策至关重要, 这就迫切需要一个帮助医护人员对临床实践指南进行快速有效判断的评价工具, 以判断该指南的质量和效用。目前国际上已发布了多个指南评价工具, 但又研究指出这些评价工具存在较大的差异性^[16, 17], 虽然这些工具日趋全面和实用, 但仍然缺乏对指南内容的评价, 也无法很好地应用于中国指南^[7]。国际上应用最广泛的 AGREE II 工具作为唯一使用数字

评分且经过验证的工具, 仍然存在条目众多、缺乏对指南内容真实性的评价及条目权重不清等问题^[7]。而本研究所构建的中国护理领域临床实践指南质量评价标准, 包括 5 个维度、15 个条目, 秉持了 AGREE II 使用数字评分的方法, 但同时增加了对指南推荐意见的安全性、有效性及经济性的评价, 且对每个条目设置了权重, 弥补了 AGREE II 的缺陷。此外, AGREE II 推荐每个指南至少有 2 名、最好有 4 名评价人员进行质量评价^[18], 本研究在充分考虑评价人员的资质及专业背景的情况下, 遴选了 5 位评价者对 2 篇护理领域内的临床实践指南采用上述两种标准进行质量评价, 结果表明, 中国护理领域临床实践指南质量评价标准在不同评价者之间的具有较好的一致性, 且对同 1 篇指南, 中国护理领域临床实践指南质量评价标准与 AGREE II 的评价结果具有高度一致性。因此, 中国护理领域临

床实践指南质量评价标准具有较好的信度和效度, 可以作为护理研究者、管理者及实践者评价护理领域内临床实践指南质量的工具。

3.2 评价者使用该指南质量评价工具时应对该工具详细解读

尽管中国护理领域临床实践指南质量评价标准具有较好的信度和效度, 但是从本研究的评价结果看, 不同评价者对同一条目的评价仍存在较大的不一致性, 主要集中在证据的质量评价、形成推荐意见的方法、推荐方案的有效性、推荐意见的经济学问题等条目, 不同评价者之间的变异系数 >0.3 。特别是对静脉输液指南, 不同评价者之间不一致的条目较多。首先, 这可能与本研究未对评价者进行统一培训有关。本研究为避免培训对评价者带来的影响, 未对5位评价者在指南评价前进行统一培训。其次, 静脉输液指南中对上述条目的描述比较笼统, 导致不同评价者对其的判断和评价存在较大差异。尽管静脉输液指南已经由中华护理学会翻译汉化并发布, 但该指南仍更适用于美国静脉输液环境, 这也反映了中国护理领域临床实践指南质量评价标准对中国本土化指南具有更好的适用性。但上述不一致的条目仍反映了指南内容的真实性及推荐意见可否解决临床具体问题, 故对这些条目评价的准确性至关重要。因此, 建议护理指南评价者在使用该指南质量评价标准前, 应了解循证护理的理论, 熟悉指南制定的方法学, 并认真解读该指南质量评价标准, 理解评价条目的内涵及评分标准。若有多位评价者使用该工具对多份指南进行质量评价, 建议在评价前对评价者进行统一培训, 减少不同评价者对指南质量评价不一致导致的偏倚。此外, 本研究仅选取了两篇护理领域内的临床实践指南对中国护理领域临床实践指南质量评价标准的信效度进行预评价, 尚未进行大样本验证和评价, 因此, 该指南质量评价标准随着其在更大范

围的推广和使用, 仍需持续更新和完善。

参考文献

- 1 王小钦, 王吉耀. 循证临床实践指南的制定与实施. 北京: 人民卫生出版社, 2016:1-50
- 2 Available at: <https://www.g-i-n.net/library/international-guidelines-library>.
- 3 IOM. Clinical practice guidelines we can trust. Washington DC. The national Academies Press, 2011.
- 4 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*, 1993, 342(8883): 1317-1322.
- 5 Lannone P, Montano N, Minardi M, et al. Wrong guidelines: why and how often they occur. *Evid Based Med*, 2017, 22(1): 1-3.
- 6 李楠, 姚亮, 吴琼芳, 等. 2012~2013年中国大陆期刊发表临床实践指南质量评价. 中国循证医学杂志, 2013, 13(6): 621-623.
- 7 王琪. 中国临床实践指南质量评价研究. 兰州: 兰州大学, 2017.
- 8 上海市公卫卫生临床中心, 复旦大学 JBI 循证护理合作中心. 艾滋病临床护理实践指南. 2014.
- 9 Infusion nurses society. Infusion therapy standards of practice. *J Infus Nurs*, 2016, 39(1S): S1-S113.
- 10 王吉耀, 王强, 王小钦, 等. 中国临床实践指南评价体系的制定与初步验证. *中华医学杂志*, 2018, 98(20): 1544-1548.
- 11 周英凤, 王强, 胡雁, 等. 中国护理领域临床实践指南质量评价标准的构建. 中国循证医学杂志, 2018, 18(9): 1-5.
- 12 Brouwers MC, Kho ME, Browman GP, et al. The Global Rating Scale complements the AGREE II in advancing the quality of practice guidelines. *J Clin Epidemiol*, 2012, 65(5): 526-534.
- 13 方积乾, 陆盈. 现代医学统计学. 北京: 人民卫生出版社, 2002: 80-87.
- 14 马斌荣. 医学科研中的统计方法. 北京: 北京科学出版社, 2005: 3.
- 15 陈耀龙, 杨克虎, 田金徽, 等. 循证实践指南的制定: 国际经验与中国实践. 兰州大学学报(医学版), 2016, 42(1): 29-35.
- 16 sierung U, Eikermann M, Hausner E, et al. Appraisal tools for clinical practice guideline: a systematic review. *PLoS one*, 2013, 8(12): e 82915.
- 17 张悦, 张士靖, 周志超, 等. 国际临床指南评价工具的发展与启示. *中华医学图书情报杂志*, 2015, 24(1): 11-16.
- 18 胡雁, 郝玉芳. 循证护理学. 北京: 人民卫生出版社, 2018: 216.

收稿日期: 2018-07-06 修回日期: 2018-09-11

本文编辑: 樊斯斯