

第六章

临床指南的评价与应用

本章要点

- 临床指南(Clinical Guideline)是以循证医学为基础,收集了最新的研究证据,所有文献均经过严格的科学性评价,分出证据的等级,再结合专家个人经验,在充分讨论的基础上形成的针对具体临床疾病处理的共识意见。
- 临床指南具有科学性、系统性和权威性,是临床实践中更广泛应用的决策依据。
- 临床指南具有时间性、地域性和实用性,需要不断更新,因时因地应用。

教学目的

- 掌握: 临床指南的评价。
- 熟悉: 临床指南的应用。
- 了解: GRADE 证据和推荐分级。

一、指南的定义、形成背景与临床意义

临床实践中,医生每天接触大量的患者,患者就诊时提供给医生很多关于自己病情的信息,随后也会询问医生:我得的是什么病,怎么会得的,如何治疗,能治好吗,会有后遗症吗等。对临床医师来说,那就是诊断、病因、治疗和预后的问题。医生不但要回答患者,还需要给患者合适的诊疗处理。回答这些问题,作出合适的临床决策,医生依据的是什么?

循证医学要求在处理临床问题时,以足够的科学证据作为依据。原始的临床研究,无论是病例对照研究、队列研究或者随机对照临床试验,都是很好的证据。这些原始研究虽有更新及时、能回答的问题多种多样等优点,但大多样本量较小,且人群代表性不强;原始研究虽来源广泛,获得途径容易,但每项研究均需临床医师进一步评价。系统综述和Meta分析虽然克服了上述缺点,但常常是针对患者的某个特殊问题或者特殊治疗进行的评价,缺少对一种疾病或一类疾病的患者包括从诊断到治疗,乃至预防等多方面的综合处理意见,并且缺少权威性。临床指南具有科学性、系统性和权威性,是临床实践中更广泛应用的决策依据。

临床指南(Clinical Guideline)以循证医学为基础,收集了最新多中心大样本随机对照临床试验(RCT)所获得的研究结果或者 RCT 为基础的系统综述,也包括诊断试验评价研究结果、队列研究结果、病例分析与报告等,所有文献均经过严格的科学性评价,分出证据的等级,再结合专家个人经验经过充分讨论形成的针对临床疾病处理的共识意见。临床指南是官方政府机构或学术组织形成的医疗文件,将规范化医疗与个体化医疗相结合,对提高医疗质量起了重要的推动作用,其目的是为了规范医疗范围,提高医疗质量,控制医疗费用。

二、循证指南中的证据及推荐分级

循证指南是遵循科学证据的指南,建立在对原始研究和二次研究的科学的、全面的评价的基础上。近年来的临床指南越来越多地采用了包括 GRADE (Grading of Recommendation Assessment, Development and Evaluation)等在内的证据质量评价、测定和推荐标准用来评估证据分级水平和推荐意见的强度。在临床指南中清楚地呈现纳入证据的质量,并明确给出推荐意见分级,不仅保证指南的科学性,而且更有利于指南的理解、传播和应用。

当前,在各种循证指南及一些循证数据库中,证据的分类和推荐方法存在一些差异。以下是几个例子:

(1) 关于幽门螺杆菌的处理的马斯特里赫特 IV-佛罗伦萨共识报告中提道:“根除幽门螺杆菌可使 1/12 幽门螺杆菌阳性功能消化不良患者获得长期的症状缓解,这一疗效优于其他任何治疗(证据等级: Ia, 推荐级别: A)。”

(2) 美国眼科学会的系列临床指南中,老年性黄斑变性(AMD)的患者教育部分建议“吸烟者戒烟(A: I)”。

(3) 而著名的循证决策数据库 UPTODATE 关于建议 AMD 患者戒烟则给出了(Grade 1A)的提示。

多年来世界各地的学者和各组织提出了数十个证据分级和推荐强度标准。不同的指南采用的证据分级和推荐强度标准各不相同。读者在查阅指南时,需要了解每部指南关于证据方面的分级和分类方法,以免混淆。例如在例(1)中, Ia 等级的证据是指来自基于同质的、方法学质量高 RCT 所成的系统综述; A 推荐为最高的推荐级别。例(2)中 A 为推荐级别,是非常重要的推荐,将显著改善患者的状况; I 为证据力度,至少来源于一个设计、执行良好的 RCT,也包括基于 RCT 的 Meta 分析。而例(3)UPTODATE 采用的是 GRADE 分级, 1A 中的 1 是强推荐,是指在大部分患者中获益显著优于风险和负担; A 则是指高质量的证据。

目前采用证据等级与推荐级别结合的方法,特别是按 GRADE 评价分级来制作的指南日益增多。以下简要介绍一下其证据分级和推荐强度分类。

1. GRADE(2011)证据等级

(1) 高: 我们非常确信效应的真实值与估计值接近。

(2) 中: 对效应估计值我们有中度信心,即真实值可能与估计值接近,但不排除两者大不相同的可能性。

(3) 低: 我们对效应估计值的确信程度有限,即真实值可能与估计值差别较大。

(4) 差：我们对效应估计值没什么信心，即真实值非常有可能与估计值大不相同。

2. GRADE 证据质量评定标准和定义 见表 2-6-1。其中证据质量从高到低分别对应⊕⊕⊕⊕、⊕⊕⊕、⊕⊕和⊕，根据研究设计和偏倚、不一致性、间接性、不精确性、发表偏倚、效应量、剂量-反应梯度以及混杂因素等情况调整⊕的数量，最终评定证据的质量。随机试验初始定义为高质量等级(⊕⊕⊕⊕)，如果存在一些降级条件，则最终可能为中(⊕⊕⊕○)，乃至更低；观察性研究初始质量等级为低(⊕⊕○○)，若符合一些升级条件，最终可能为高(⊕⊕⊕⊕)。

表 2-6-1 GRADE 证据资料评定标准和定义

证据质量等级	如果符合如下条件,升级	如果符合如下条件,降级	研究设计
高(⊕⊕⊕⊕)	效应量大(+1 大,+2 非常大)	偏倚风险(-1 严重,-2 非常严重)	随机试验
中(⊕⊕⊕○)	剂量-反应梯度(+1 有剂量-反应梯度证据)	不一致性(-1 严重,-2 非常严重)	
低(⊕⊕○○)	所有可能的残余混杂因素(+1 可降低所显示的效应,+1 如研究未观察到效应,潜在的混杂因素可提示其为假效应)	间接性(-1 严重,-2 非常严重)	观察性研究
差(⊕○○○)		不精确(-1 严重,-2 非常严重)	
		发表偏倚(-1 可能,-2 非常可能)	

3. GRADE 推荐的方向和强度

推荐(recommendation)有两个维度：方向(支持/反对某项措施)及强度(强推荐/弱推荐)。如果指南作者对某项措施的预期和非预期后果非常有把握，则做“强推荐”；不太有把握的时候则做“弱推荐”。因此，“强推荐”不等同“优先推荐”。指南作者的“强推荐”意味着在获取相关信息后，所有或几乎所有的医生或患者都会选择“支持/反对”某项措施。“弱推荐”则意味着虽有很多人会选择“支持/反对”某项措施，但也有不少例外。在使用指南的“弱推荐”时，医生需要花更多时间与患者进行交流，更多地了解患者的意愿和价值观，以共同作出循证决策。

GRADE 工作组先后在 BMJ 和 Journal of Clinical Epidemiology 上分别发表了 6 篇和 20 篇系列文章作详细介绍，均可在其官方网站 <http://www.gradeworkinggroup.org/publications/index.htm> 上找到其原文链接。读者如有兴趣可以进一步访问及阅读。

三、临床指南的评价

临床指南对各种疾病具体的诊断、治疗和预防等措施提供明确清晰的推荐意见，通过降低临床实践的差异性和随意性，减少不必要的诊断试验和防止采用无效的治疗手段，从而成为提高医疗质量的有用工具。过去数十年中世界各种相关组织发表了上万种指南，虽然指南常在证据基础上产生的，但在临床实践中也需要依据新的证据不断更新，依据不同人群、不同状态选择应用。

好的指南应该具有如下特征：

- (1) 准确：指南推荐应基于高质量的证据。
- (2) 明了：读者或指南使用者能对指南进行评价并应用。
- (3) 便于评价：能够评价使用指南可能对患者的健康、经济费用等方面带来的影响。
- (4) 便于解决争论：能够明确或解决一些不一致，或存在争论的看法/做法。
- (5) 便于应用：指南应可用于具体某个具体的患者。

总之,应用指南前也需要经过科学性和合理性评价。表 2-6-2 列出了临床指南应用评价标准。

表 2-6-2 临床指南应用评价标准

-
1. 检查临床实践指南的设计是否科学
 - (1) 指南是否清楚详细地对每一个问题作了恰如其分的说明
 - (2) 对每一提出的问题有无找到系统综述
 - (3) 系统综述有否应用入选标准
 - (4) 对每一问题是否进行全面的文献检索
 - (5) 是否评价每个问题研究的方法学
 - (6) 对研究的方法学和研究结果是否进行了重复性的评定
 - (7) 在系统综述和指南推荐意见间有无联系
 2. 解释结果是否合理恰当
 - (1) 作者如何划分证据水平
 - (2) 作者如何对推荐意见进行分级
 3. 这项结果能否用于当前患者
 - (1) 患者情况是否与临床指南中患者情况相同
 - (2) 结果包括利与弊,在其应用到临床时能否与具体患者处理相结合
 - (3) 是否包括了重要的亚组
-

四、临床指南应用原则

在临床指南的应用中,与其他循证证据一样,医生需要确定指南的科学性、适用性,也就是对指南进行评价。表 2-6-3 列出了临床指南的应用原则。

表 2-6-3 临床指南应用原则

-
1. 指南的真实性如何?
 - (1) 是否完成了近 12 个月以来的全面的、能重复性的文献综述?
 - (2) 证据是否有推荐分级、类别以及特定的范围?
 2. 如果这个指南是真实的,能应用于当前的患者/医院/社区吗?(4 个 B)
 - (1) 是否疾病的负担(Burden)太低或患者太少,能保证实施吗?
 - (2) 患者/社区的意愿(Beliefs)与指南推荐的相符合吗?
 - (3) 实施这个指南时是否有不利因素(Bargain)?
 - (4) 实施指南是否有些不能克服的障碍(Barriers)(地方的、政府的、合法性等)?
-

五、结语

自 1993 年在 Index Medicus 可以用“实践指南”作为关键词检索到所需要的内容,美国国立卫生研究院公布的临床指南和专家组意见分两个目录收集在 <http://text.nlm.nih.gov>。加拿大医学会 <http://www.cma.ca/cpgs/index/htm> 和澳大利亚医学会 <http://www.mja.com.au/public/guides/guides.html> 也提供了临床指南。中华医学会发布的临床诊治指南虽然没有收集到一起,但中华医学会期刊系列均已全文上网 <http://www.cma.org.cn/qkzz/index.html>,查找原文也非常方便。

临床指南信息量大,收集信息难度高,周期长,更新慢是其缺点。尽管如此,由于临床指南是具有权威性的医疗文件,因此无论制定指南或者应用指南,对其质量评价显得非常重要。一项好的临床指南应具有真实性、可靠性和可重复性。质量不佳的临床指南不仅

对实践循证医学没有好处,有时还会误导临床医生。

(陈世耀 袁源智)

【思考题】

1. 什么是临床指南?与研究原著、系统综述有何不同?
2. 举例说明应用临床指南处理临床问题一般包括哪些过程。
3. 临床指南在应用过程中需要遵循哪些原则?

本篇主要参考文献

王吉耀. 循证医学与临床实践, 第3版. 北京: 科学出版社, 2012.

Kenneth F. Schulz, David A. Grimes. 柳叶刀临床研究基本概念. 王吉耀主译. 北京: 人民卫生出版社, 2010.

Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-based-medicine: How to practice and teach EBM (Fourth Edition). Edinburgh: Churchill Livingstone, London: Elsevier, 2011: 1-13.

李幼平. 循证医学, 第3版. 北京: 高等教育出版社, 2013.

王家良. 循证医学, 第2版. 北京: 人民卫生出版社, 2010.

林果为, 王小钦, 陈世耀. 现代临床流行病学, 第3版. 上海: 复旦大学出版社, 2014.

Hayward RSA, Wilson MC, Tunis SR, et al. Users' guides to the medical literature: VIII. How to use clinical practice guidelines A. Are the recommendations valid? JAMA, 1995, 274(7): 570-574.

Wilson MC, Hayward RSA, Tunis SR, et al. Users' guides to the medical literature: VIII. How to use clinical practice guidelines B. What are the recommendations and will they help you in caring for your patients? JAMA, 1995, 274(20): 1630-1632.

Balshem H, Helfand M, Holger J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of Clinical Epidemiology, 2011, 64: 401-406.