

附表 4 医学二次研究评估工具及其主要组成部分

Supplementary table 4. Major components of the tools for assessing secondary medical studies

A. 系统评价和 meta 分析——评估系统性评价的测量工具 (AMSTAR) -2 (最后修订于 2017 年)			
网址: http://www.amstar.ca/			
主要条目	选项		
1. 研究问题和纳入标准是否包括了 PICO 部分?	是	否	/
2. 是否声明在系统评价实施前确定了系统评价的研究方法? 对于与研究方案不一致处是否进行说明?	是	否	部分是
3. 系统评价作者在纳入文献时是否说明纳入研究的类型?	是	否	/
4. 系统评价作者是否采用了全面的检索策略?	是	否	部分是
5. 是否采用双人重复式文献选择?	是	否	/
6. 是否采用双人重复式数据提取?	是	否	/
7. 系统评价作者是否提供了排除文献清单并说明其原因?	是	否	部分是
8. 系统评价作者是否详细地描述了纳入的研究?	是	否	部分是
9. 系统评价作者是否采用合适工具评估每个纳入研究的偏倚风险?	是	否	部分是
10. 系统评价作者是否报告纳入各个研究的资助来源?	是	否	/
11. 作 meta 分析时, 系统评价作者是否采用了合适的统计方法合并研究结果?	是	否	未进行 meta 分析
12. 作 meta 分析时, 系统评价作者是否评估了每个纳入研究的偏倚风险对 meta 分析结果或其它证据综合结果潜在的影响?	是	否	未进行 meta 分析
13. 系统评价作者解释或讨论每个研究结果时是否考虑纳入研究的偏倚风险?	是	否	/
14. 系统评价作者是否对研究结果的任何异质性进行合理的解释和讨论?	是	否	/
15. 如果系统评价作者进行定量合并, 是否对发表偏倚 (小样本研究偏倚) 进行充分的调查, 并讨论其对结果可能的影响?	是	否	未进行 meta 分析
16. 系统评价作者是否报告了所有潜在利益冲突的来源, 包括所接受的任何用于制作系统评价的资助?	是	否	/
B. 评价系统评价的文献严格评价项目 (CASP) 清单 (最后修订于 2018 年)			
网址: https://casp-uk.net/casp-tools-checklists/			
主要条目	选项		
第一部分: 研究结果可靠吗?			
1. 系统评价是否定义了一个清晰明确的问题?	是	否	不知道
2. 系统评价纳入的研究设计类型合适吗?	是	否	不知道
是否值得继续?			
3. 是否纳入了所有重要的相关文献?	是	否	不知道
4. 系统评价制作者是否对纳入研究的质量进行了充分的评价?	是	否	不知道
5. 如果对纳入研究的结果进行了 Meta 分析, 这样做是否合适?	是	否	不知道
第二部分: 研究结果是什么?			
6. 系统评价的总体结果是什么?			
7. 结果的精确度如何?			
第三部分: 研究结果适用吗?			
8. 研究结果是否可用于当地人群?	是	否	不知道
9. 考虑所有的重要结局了吗?	是	否	不知道
10. 获益是否大于危害和成本?	是	否	不知道
C. 苏格兰国家指南小组 (SIGN) 方法学清单: 系统评价和 meta 分析 (最后修订于 2014 年)			
网址: https://www.sign.ac.uk/what-we-do/methodology/checklists/			
主要条目	选项		

第一部分：内部真实性				
在一项制作良好的系统评价中：		当前研究的情况是否做到？		
1.1 明确定义了研究问题且在文章里面列出了纳入 / 排除标准	是	否	如果否，排除	/
1.2 进行了全面的文献检索	是	否	不适用	如果否，排除
1.3 至少两名研究人员进行研究筛选	是	否	不知道	/
1.4 至少两名研究人员进行资料提取	是	否	不知道	/
1.5 未将研究的发表情况作为纳入标准之一	是	否	/	/
1.6 列出了排除的研究	是	否	/	/
1.7 提供了纳入研究相关的基本特征	是	否	/	/
1.8 评估并报告了纳入研究的科学质量	是	否	/	/
1.9 纳入研究的科学质量适当吗？	是	否	/	/
1.10 采用合适的方法合并纳入单项研究的结果	是	否	不适用	不知道
1.11 合适地评估发表偏倚的可能性	是	否	不适用	/
1.12 声明了利益冲突	是	否	/	/
第二部分：研究的整体评价				
2.1 您对该系统评价 / Meta 分析的整体方法学质量评价是？	高质量 (++)	可接受的 (+)	低质量 (-)	不可接受 - 排除 0
2.2 通过本指南，这项研究的结果可以直接适用于目标病人组？	是	否	/	/
2.3 注释：				
D. 评价系统评价与研究综合的乔安娜·布里格斯学院 (JBI) 清单 (最后修订于 2017 年)				
网址： https://joannabriggs.org/critical-appraisal-tools https://wiki.joannabriggs.org/display/MANUAL/Appendix+10.1+JBI+Critical+Appraisal+Checklist+for+Systematic+reviews+and+Research+Syntheses				
主要条目		选项		
1. 所提出的循证问题是否清晰、明确？	是	否	不清楚	不适用
2. 纳入标准是否适用于提出的循证问题？	是	否	不清楚	不适用
3. 采用的检索策略是否恰当？	是	否	不清楚	不适用
4. 检索研究的资源、来源是否充分？	是	否	不清楚	不适用
5. 评价研究的标准是否恰当？	是	否	不清楚	不适用
6. 是否由 2 名及以上的研究员独立进行评价？	是	否	不清楚	不适用
7. 提取资料时是否采用一定的措施减少误差？	是	否	不清楚	不适用
8. 综合 / 合并研究的方法是否恰当？	是	否	不清楚	不适用
9. 是否评估了发表偏倚的可能性？	是	否	不清楚	不适用
10. 报告的数据是否支持有关政策和 / 或实践的建议？	是	否	不清楚	不适用
11. 针对新研究的特定指令是否合适？	是	否	不清楚	不适用
总体评价：纳入 <input type="checkbox"/> 排除 <input type="checkbox"/> 寻找更多信息 <input type="checkbox"/>				
E. 系统评价和 meta 分析——美国国立卫生研究院 (NIH) 质量评估工具				
网址： https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools				
主要条目		选项		
1. 是否基于充分阐述和描述的重点问题进行综述？	是	否	无法确定 / 不适用 / 未报告	
2. 是否预先确定、并明确了纳入和排除研究的资格标准？	是	否	无法确定 / 不适用 / 未报告	
3. 文献检索策略是否采用全面、系统的方法？	是	否	无法确定 / 不适用 / 未报告	
4. 是否双人独立对文章的标题、摘要和全文进行纳入、排除审查，以尽量减少偏倚？	是	否	无法确定 / 不适用 / 未报告	

5. 每个纳入研究的质量是否由两个或多个研究者使用标准方法独立评估, 以评估其内部真实性?	是	否	无法确定 / 不适用 / 未报告
6. 纳入的研究是否与每个研究的重要特征和结果一起列出?	是	否	无法确定 / 不适用 / 未报告
7. 是否评估了发表偏倚?	是	否	无法确定 / 不适用 / 未报告
8. 是否评估了异质性? (此问题仅适用于 meta 分析)	是	否	无法确定 / 不适用 / 未报告
质量等级	优	中等	差
附加评论 (如果选择“差”, 请说明原因):			
F. 决策支持单元 (DSU) 网状 meta 分析 (NMA) 方法清单 (最新发布于 2012 年 1 月)			
网址: http://nicedsu.org.uk/technical-support-documents/evidence-synthesis-tds-series/			
主要条目			选项
A. 决策问题的定义			
A1. 供决策的目标人群			
A1.1. 是否明确定义了供决策的目标患者群体?	是	否	不适用
A2. 对照组			
A2.1. 决策对照设置: 决策中所有适当的处理方法是否已确定	是	否	不适用
A2.2. 综合对照设置: 综合对照设置中是否有其他治疗, 而决策对照设置中没有其他治疗? 如果是这样, 这是否充分合理?	是	否	不适用
A3. 试验纳入 / 排除			
A3.1. 检索策略在技术上是适当并已适当报告?	是	否	不适用
A3.2. 是否已纳入涉及综合对照设置中至少两种治疗方法的所有试验?	是	否	不适用
A3.3. 是否已包括所有报告相关结果的试验?	是	否	不适用
A3.4. 是否包括其他试验? 如果是这样, 这是否合理?	是	否	不适用
A4. 治疗的定义			
A4.1. 是否所有治疗选择都限于特定的剂量和联合治疗, 还是将不同的剂量和联合治疗“集中”在一起? 如果是后者, 是否有充分理由?	是	否	不适用
A4.2. 还有其他建模假设吗?	是	否	不适用
A5. 综合试验选择的试验结果和量表			
A5.1. 如果有可供选择的结果, 综合中使用的结果度量的选择是否合理?	是	否	不适用
A5.2. 选择标准后的假设是否合理?	是	否	不适用
A6. 患者人群: 与目标人群以外的患者进行的试验			
A6.1. 一些试验是否包括目标人群以外的患者? 如果是这样, 这是否合理?	是	否	不适用
A6.2. 对可能对相对治疗效果产生影响或缺乏影响的假设是什么? 他们有充分的道理吗?	是	否	不适用
A6.3. 是否针对这些差异进行了调整? 如果是这样, 请评论支持此校正的证据是否足够, 以及是否需要进行敏感性分析。	是	否	不适用
A7. 患者人群: 目标人群中的异质性			
A7.1. 是否有关于治疗效果潜在调节剂的文献综述?	是	否	不适用
A7.2. 尽管在目标人群中, 患者群体之间的试验之间是否存在明显或潜在的差异? 如果是这样, 是否已充分考虑到这一点?	是	否	不适用
A8. 偏倚风险			
A8.1. 是否对这些试验或这些试验组合容易受到的偏倚进行讨论?	是	否	不适用
A8.2. 如果确定存在偏倚风险, 是否对分析进行任何调整, 这样是否合理?	是	否	不适用
A9. 数据呈现			
A9.1. 是否有清晰的表格或图表显示基本案例分析中包含哪些数据?	是	否	不适用
A9.2. 是否有清晰的表格或图表显示哪些数据已被排除, 为什么?	是	否	不适用
B. 结果的分析和表示方法			
B1. Meta 分析方法			
B1.1. 统计模型是否清楚描述?	是	否	不适用

B1. 2. 软件操作是否被记录?	是	否	不适用
B2. 相对治疗效果的异质性			
B2. 1. 是否提供了相对治疗效果异质性程度的数值估计?	是	否	不适用
B2. 2. 是否有选择随机或固定效应模型的理由? 是否应考虑敏感性分析?	是	否	不适用
B2. 3. 对异质性有足够的应对吗?	是	否	不适用
B2. 4. 相对治疗效果无法解释的差异是否会威胁到结论的可靠性?	是	否	不适用
B2. 5. 是否讨论过基线之间的统计异质性?	是	否	不适用
B3. 试验结果的基线模型			
B3. 1. 是否在同一模型中估算了基线效应和相对效应? 如果是这样, 这是否合理?	是	否	不适用
B3. 2. 是否已经解释了选择研究作为基线模型的依据?	是	否	不适用
B4. 分析试验数据的结果呈现			
B4. 1. 如果使用固定效应模型, 是否将相对治疗效果 (相对于安慰剂或“标准”对照组) 以及研究之间异质性的度量制成表格?	是	否	不适用
B4. 2. 是否报告了在成本效益分析 (CEA) 中使用的每种处理的绝对效果?	是	否	不适用
B5. 自然历史模型其他部分的综合			
B5. 1. 是否适当描述和证明了选择数据源以告知自然历史模型中的其他参数?	是	否	不适用
B5. 2. 在自然史模型中, 治疗之间的长期差异是否可以通过随机试验结果的差异来解释?	是	否	不适用
C. 网状综合特有的问题			
C1. 有关模型规范和软件实现的信息是否充分			
C2. 多臂试验			
C2. 1. 如果有多组试验, 是否考虑了相对治疗效果之间的相关性?	是	否	不适用
C3. 连接和未连接的网络			
C3. 1. 是否基于随机试验建立了证据网络?	是	否	不适用
C4. 不一致性			
C4. 1. 网络中可能有多少个不一致之处?	是	否	不适用
C4. 2. 是否有先验的理由担心存在不一致的原因, 这是由于比较治疗 A 和 B 的试验患者与比较治疗 A 和 C 的试验患者之间的系统性临床差异所致?	是	否	不适用
C4. 3. 是否进行了足够的非一致检查?	是	否	不适用
C4. 4. 如果检测到不一致, 则对分析进行哪些调整, 以及如何证明其合理性?	是	否	不适用
D. 将合成纳入概率成本效益分析			
D1. 不确定性传播			
D1. 1. 参数估计的不确定性是否已通过 CEA 模型传播?	是	否	不适用
D2. 相关性			
D2. 1. 参数之间有相关性吗? 如果是这样, 相关性是否已通过 CEA 模型传播?	是	否	不适用
G. 系统评价偏倚风险 (ROBIS) 评估工具 (最新发布于 2014 年)			
网址: https://www.bristol.ac.uk/population-health-sciences/projects/robis/robis-tool/			
主要条目		选项	
领域 1: 研究的纳入排除标准			
描述研究的纳入排除标准, 以及是否有证据表明预先规定了目的和纳入排除标准:			
1. 1 系统评价遵循了预先确定的目的和纳入标准吗?	是 / 可能是	否 / 可能否	无信息
1. 2 纳入标准适合系统评价的问题吗?	是 / 可能是	否 / 可能否	无信息
1. 3 纳入标准明确吗?	是 / 可能是	否 / 可能否	无信息
1. 4 纳入标准中所有基于研究特征的限制合适吗 (如: 日期、样本量、研究质量、测量结果)?	是 / 可能是	否 / 可能否	无信息
1. 5 纳入标准中所有与研究来源相关的限制合适吗 (如: 出版物状态或格式、语言、数据可用性)?	是 / 可能是	否 / 可能否	无信息
对纳入标准的描述的偏倚风险程度	高	低	不清楚

评判的理由:			
领域 2: 研究的检索和筛选			
描述了研究检索和筛选的方法 (如系统评价员人数):			
2.1 检索已发表和未发表的研究时所包含的数据库或电子资源的范围合适吗?	是 / 可能是	否 / 可能否	无信息
2.2 使用了除数据库检索以外的其他方法来确定相关研究吗?	是 / 可能是	否 / 可能否	无信息
2.3 检索策略的检索词和结构能尽可能多地检索到符合的研究吗?	是 / 可能是	否 / 可能否	无信息
2.4 基于时间、发表形式、语言的限制合适吗?	是 / 可能是	否 / 可能否	无信息
2.5 研究的筛选中尽可能地减小了误差吗?	是 / 可能是	否 / 可能否	无信息
研究检索和 (或) 筛选所使用方法的偏倚风险程度	高	低	不清楚
评判的理由:			
领域 3: 数据提取和质量评价			
描述数据收集的方法, 从研究中提取或通过其他方法收集了哪些数据, 偏倚风险是如何评估的 (如评估者的数量) 以及用来评估偏倚风险的工具:			
3.1 数据提取中尽可能地减小了误差吗?	是 / 可能是	否 / 可能否	无信息
3.2 系统评价作者和读者能获取足够的研究特征来解读结果吗?	是 / 可能是	否 / 可能否	无信息
3.3 提取了所有相关的研究结果来进行数据合成吗?	是 / 可能是	否 / 可能否	无信息
3.4 使用了合适的工具来正规地评价偏倚风险了吗?	是 / 可能是	否 / 可能否	无信息
3.5 偏倚风险评价中尽可能地减小了误差吗?	是 / 可能是	否 / 可能否	无信息
数据提取和质量评价所使用方法的偏倚风险程度	高	低	不清楚
评判的理由:			
领域 4: 数据合成和结果呈现			
描述合成方法:			
4.1 数据合成包括了所有应该包括的研究吗?	是 / 可能是	否 / 可能否	无信息
4.2 遵循了所有预先确定的分析吗? 未遵循的部分解释了吗?	是 / 可能是	否 / 可能否	无信息
4.3 鉴于纳入研究的研究问题、研究设计和结局指标的性质和相似性, 数据合成方法恰当吗?	是 / 可能是	否 / 可能否	无信息
4.4 数据合成中研究之间的差异 (异质性) 是最小的或者经过处理了吗?	是 / 可能是	否 / 可能否	无信息
4.5 结果是否稳定, 如是否通过漏斗图或敏感性分析来证明?	是 / 可能是	否 / 可能否	无信息
4.6 原始研究的偏倚最小吗? 或者在数据合成中处理了吗?	是 / 可能是	否 / 可能否	无信息
数据合成和结果呈现的偏倚风险程度	高	低	不清楚
评判的理由:			
系统评价的偏倚风险			
描述结论是否得到证据的支持:			
A. 结果解释中处理了领域 1~4 中所有偏倚风险吗?	是 / 可能是	否 / 可能否	无信息
B. 合理地考虑到了纳入研究与系统评价研究问题的相关性吗?	是 / 可能是	否 / 可能否	无信息
C. 评价者避免强调有统计学意义的结果了吗?	是 / 可能是	否 / 可能否	无信息
系统评价的偏倚风险:	高	低	不清楚
风险理由:			
H. 指南研究与评价 II (AGREE II) 工具 (最新发布于 2009 年)			
网址: http://www.agreetrust.org/			
1. 明确阐述指南的总目的	AGREE II 每个条目都是以 7 分制来评价: 1 分, 代表指南“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。条目评分取决于指南报告的完整性和质量, 随着满足更多的标准和考虑因素, 分数会增加。每个条目的“如何评分”部分包括有该项目的评估标准和注意事项的详细信息。		
2. 明确阐述指南所涵盖的卫生问题			
3. 明确阐述指南所要应用的人群 (患者、公众等)			
4. 指南制定组包括所有相关专业的人员			
5. 考虑到目标人群 (患者、公众等) 的观点和选择			
6. 指南的适用者已经明确规定			
7. 用系统的方法检索证据			

8. 明确阐述选择证据的标准	
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	
11. 在形成推荐意见时考虑了对健康的效益、副作用以及风险	
12. 推荐意见和支持证据之间有明确的联系	
13. 指南在发表前经过专家的外部评审	
14. 提供指南更新的过程	
15. 推荐意见明确不含糊	
16. 明确列出针对某一情况或卫生问题不同的选择	
17. 主要的推荐意见清晰易辨	
18. 指南中描述了应用过程中的促进和阻碍因素	
19. 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	
20. 指南考虑了应用推荐建议时潜在的资源投入问题	
21. 指南提供了监控和（或）审计的标准	
22. 赞助单位的观点不影响指南的内容	
23. 指南记录并考虑了制定小组成员的利益冲突	