

网状Meta分析偏倚风险评估工具RoB NMA解读



赵文硕^{1,2}, 刘 明^{3,4}, 史纪元⁵, 李 伦^{6,7}, 杨丰文^{8,9}, 薛付忠^{1,2}, 高 亚^{1,2}

1. 山东大学齐鲁医学院公共卫生学院医学数据学系 (济南 250003)
2. 国家健康医疗大数据研究院 (济南 250003)
3. 新加坡南洋理工大学李光前医学院 (新加坡 308232)
4. 兰州大学基础医学院循证医学中心 (兰州 730000)
5. 北京中医药大学护理学院 (北京 100029)
6. 中南大学湘雅二医院普外科 (长沙 410010)
7. 湖南省乳腺疾病临床医学研究中心 (长沙 410011)
8. 天津中医药大学循证医学中心 (天津 301617)
9. 国家药品监督管理局中医药循证评价重点实验室 (天津 301617)

【摘要】 现有偏倚风险评估工具主要针对传统两两比较 Meta 分析, 难以全面评估网状 Meta 分析所特有的偏倚风险来源。为此, Lunny 等研究者研发了针对网状 Meta 分析偏倚风险来源的评估工具 (Risk of Bias in Network Meta-analysis, RoB NMA), 专用于评估网状 Meta 分析在干预选择与网络结构、效应修饰因素以及统计分析等层面可能带来的偏倚。本文旨在解读 RoB NMA 工具的制订背景、条目内容以及评价流程, 为国内研究人员使用该工具提供参考。

【关键词】 系统评价; 网状 Meta 分析; RoB NMA; 偏倚风险; 解读

【中图分类号】 R195.1 **【文献标识码】** A

An introduction to Risk of Bias in Network Meta-analysis (RoB NMA) tool

ZHAO Wenshuo^{1,2}, LIU Ming^{3,4}, SHI Jiyuan⁵, LI Lun^{6,7}, YANG Fengwen^{8,9}, XUE Fuzhong^{1,2}, GAO Ya^{1,2}

1. Department of Medical Dataology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250003, China
2. National Institute of Health Data Science of China, Jinan 250003, China
3. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore
4. Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou 730000, China
5. School of Nursing, Beijing University of Traditional Chinese Medicine, Beijing 100029, China
6. Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha 410011, China
7. Clinical Research Center for Breast of Hunan Province, Changsha 410011, China
8. Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China
9. Key Laboratory of Evidence-Based Evaluation of Traditional Chinese Medicine, National Medical

DOI: [10.12173/j.issn.1004-5511.202507209](https://doi.org/10.12173/j.issn.1004-5511.202507209)

基金项目: 山东大学齐鲁青年学者项目 (202499000161)

通信作者: 高亚, 博士, 研究员, 博士研究生导师, Email: gaoy2021@163.com

Products Administration, Tianjin 301617, China

Corresponding author: GAO Ya, Email: gaoy2021@163.com

【Abstract】 Existing risk of bias assessment tools are primarily designed for traditional pairwise Meta-analysis and fail to comprehensively evaluate the unique sources of bias risk inherent in network Meta-analysis. Therefore, Lunny and others developed a tool the Risk of Bias in Network Meta-analysis (RoB NMA), designed specifically to evaluate potential biases in network Meta-analysis arising from aspects such as intervention selection and network structure, effect modifiers, and statistical analysis. This article aims to interpret the development background, item content, and evaluation process of the RoB NMA tool, providing reference and guidance for domestic researchers of its application.

【Keywords】 Systematic review; Network-Meta-analysis; RoB NMA; Risk of bias; Interpretation

随着临床证据的快速增长, 临床实践与最新研究证据间的脱节问题日益突出。循证医学的问世满足了严格评估和综合现有证据的需要。基于高质量随机对照试验 (randomized controlled trial, RCT) 的系统评价和 Meta 分析被认为是循证医学的最高级别证据, 其中传统的两两比较 Meta 分析已在临床医学、公共卫生等多个领域得到广泛应用, 但其适用范围常受干预措施数量的影响^[1-2]。

网状 Meta 分析 (network Meta-analysis, NMA) 作为传统 Meta 分析方法的拓展, 相较于传统 Meta 分析仅进行两种干预的直接比较, NMA 可以通过整合同一证据网络中的直接与间接证据, 同时对多种干预措施进行比较, 提高结果的可比性^[3-9]。此外, NMA 还可生成干预措施效果的相对排序, 提供有价值的决策信息^[8, 10]。但证据综合过程中往往也会伴随偏倚的产生^[11-12], NMA 自身设计和分析的复杂性在保证证据质量的同时也面临着一定的偏倚风险, 如纳入研究人群或干预组合不合理、效应修饰因素分布不均、统计模型选择不当、选择性报告等, 都可能导致 NMA 结果偏离真实。目前对系统评价和 Meta 分析中偏倚的评估往往局限于系统评价整体层面的偏倚, 如 ROBIS (risk of bias in systematic reviews) 工具和 AMSTAR 2 (a measurement tool to assess systematic reviews 2)^[13-16], 或对所纳入研究的偏倚风险进行评估的 RoB 2 (revised Cochrane risk of bias tool for randomised trials) 工具^[17-18], 及对证据缺失偏倚风险进行评估的 ROB-ME (risk of bias due to missing evidence) 工具、ROB-MEN (risk of bias due to missing evidence in network Meta-analysis) 工具等^[19-22], 都不能完全覆盖

NMA 特有的偏倚来源。

针对上述特有间题, Lunny 等^[23]通过文献综述、潜在需求调查、德尔菲调查和用户调研等方法, 收集并整理了 NMA 潜在偏倚来源, 并经多轮讨论形成条目框架, 于 2025 年制定出针对 NMA 的偏倚风险评估工具 RoB NMA (risk of bias in network Meta-analysis)。本文旨在对 RoB NMA 工具进行系统解读, 以助国内研究者准确应用, 提升 NMA 研究质量。

1 RoB NMA工具简介与评估方法

1.1 评价标准

RoB NMA 工具由干预措施和网络几何结构、效应修饰作用和统计合成 3 个领域共计 17 个条目组成, 具体内容见表 1。每个条目有五种评定结果, 分别为: 是 (true, T), 表示无明显偏倚 (低偏倚风险); 可能是 (probably true, PT)、可能否 (probably false, PF), 表示存在未明确潜在偏倚或偏倚调整不完全、尚不清楚是否对评价产生影响, 但存在一定选择倾向性的情况; 否 (false, F), 表示有明确偏倚问题 (高偏倚风险); 无信息 (no information, NI), 表示根据现有信息尚不能做出倾向性判断。面对信息不足时, RoB NMA 工具鼓励评估者尽量做出倾向性判断 (“可能是” “可能否”), 而非频繁使用 “无信息” 选项, 以增强评估的区分度。

1.2 适用范围

RoB NMA 工具目前针对的是干预性 NMA。对于诊断试验准确性或预后因素的 NMA, 因偏倚来源不同, 不属于 RoB NMA 工具适用范围。RoB NMA 工具评估对象是 “单项 NMA”, 通常对应 NMA 中的一个网络图, 如针对某个结局变

表1 RoB NMA工具条目信息
Table 1. RoB NMA tool item information

领域	条目	选项
1. 干预措施和网状图	1.1 NMA中纳入的所有干预措施及对照，对于整个目标人群而言，均为合理的替代方案	T/PT/PF/F/NI
	1.2 所有的干预措施均已包含在网络中	
	1.3 网络中所有干预措施均被正确划分至相应节点	
	1.4 所有比较的干预措施都通过适当的研究内比较链相互连接	
2. 效应修饰因素	2.1 网络中不同直接比较的结果定义和时间点相似	T/PT/PF/F/NI
	2.2 网络中不同直接比较的受试者特征（效应修饰的患者特征）相似	
	2.3 网络中不同直接比较的研究特征（效应修饰的试验特征）相似	
	2.4 若条目2.1、2.2、2.3被判断为“F”或“PF”，分析是否恰当地考虑了网络中效应修饰因素差异	
3. 统计合成	3.1 分析遵循了各研究内的随机化原则	T/PT/PF/F/NI
	3.2 未发现存在发表偏倚或其他选择性报告偏倚	
	3.3 所有预先计划的分析都已报告，且仅报告了这些分析，或若存在偏离则已给予合理解释	
	3.4 纳入原始研究中的偏倚水平较低或已在综合分析中予以处理	
	3.5 对于多臂试验采用了适当的方法进行处理	
	3.6 在比较组内效应量的同质性或异质性假设方面，研究采用了适当的模型假设	
领域层面风险	3.7 同一效应的直接证据与间接证据间无显著不一致	LR/SC/HR
	3.8 若3.7为“F”或“PF”：直接和间接证据间的不一致性已得到充分处理	
	3.9 若采用贝叶斯分析，所用先验分布设置是合适的	
	对干预措施及网状图中整体偏倚风险的评估	
整体偏倚风险	对效应修饰因素的整体偏倚风险评估	LR/SC/HR
	对统计合成领域整体偏倚风险评估	
整体偏倚风险	NMA 结果的整体偏倚风险	LR/SC/HR
	NMA 结论的整体偏倚风险	

注：T.是；PT.可能是；PF.可能否；F.否；NI.无信息；LR.低风险；SC.有一定风险；HR.高风险。

量的 NMA。如果一个 NMA 对多个结局或亚组进行了多项 NMA，则需要分别对每项 NMA 进行偏倚风险的评估。另外，RoB NMA 工具仅聚焦于 NMA 本身的偏倚风险，而不直接评价整篇研究的其他方面偏倚风险。开发团队建议将 RoB NMA 工具与现有的偏倚风险工具配套使用，可先使用 ROBIS 等工具评估系统评价中的一般偏倚风险^[13]，再用 RoB NMA 评估 NMA 特有偏倚，从而更全面评估系统评价的偏倚风险。

1.3 评估流程

1.3.1 准备工作

采用 RoB NMA 工具对 NMA 进行评估之前，评估者应通读系统评价全文，并预先标注提取 RoB NMA 关注的关键信息，如研究网络构建方法、研究纳入的干预措施、研究所关注的结局指标、研究受试者关键基线特征和研究所采用的统计方法等。

1.3.2 逐条评估

评估者应按照 RoB NMA 所包含的 3 个领域 17 个信号问题对所评估的 NMA 逐条做出评估并

记录理由。开发者建议在填写工具时提供简短文字说明或直接引用原文佐证每个判断，以保证评估透明和可追溯性。此外，应注意评估过程中的逻辑关联问题，部分条目仅在前序相关问题被判断为“F”或“PF”时才需进行后续评估。如在表 1 中 2.1、2.2、2.3 被判断为“F”或“PF”，即网络中存在效应修饰因素差异时，才继续评估 2.4 是否对该差异进行了适当的处理。另外，3.9 仅在 NMA 网络中使用了贝叶斯方法时才适用，否则应标记为“不适用”。

1.3.3 领域层面风险评估

在每个域内信号问题判断完成后，应对每个领域的偏倚风险进行初步的综合判定。具体分为低偏倚风险（low risk of bias，LR）、有一定风险（some concerns，SC）高偏倚风险（high risk of bias，HR）三个等级。判定原则与 RoB 2 工具类似，即仅当该领域所有信号问题均无明显偏倚风险时，可判断为“LR”；若有一项或多项条目显示 NMA 存在潜在的未充分处理偏倚问题，且不明确是否会影响干预效应，则应判断为“SC”；

若某领域有关键问题为“F”，显示出重要的偏倚问题，则应判断为“HR”。

1.3.4 结果和结论偏倚风险评估

NMA 结果偏倚风险评估主要考察 NMA 提供的相对效应估计值是否可信。对分析结果的偏倚风险需综合 RoB NMA 各领域评估结果（如干预选择、效应修饰因素控制）、系统评价质量（如研究筛选偏倚）及 NMA 方法学严谨性（如全局/局部不一致性分析、不可传递性评估）进行权衡。结果偏倚可判定为 LR/SC/HR。

NMA 结论偏倚风险主要考察 NMA 作者在论文中对结果的解释和陈述是否客观公正，是否正确处理了评估中确定的局限性。即使 NMA 结果本身的偏倚风险低，作者的倾向性解读仍然可能引入偏倚。常见的结论偏倚有：过度强调结果的统计学显著性而忽略了实际的临床意义、选择性引用部分结果支持作者立场等。若所比较的干预效果处于高偏倚风险，在得出结论时作者系统处理了这些偏倚风险，或所比较的干预效果处于低偏倚风险，整体结论偏倚风险可判断为“LR”。如果由于 NMA 本身的原因，使结论与原有干预效果偏倚风险产生差异，则应判断为“HR”。

在 RoB NMA 工具的实际应用中，一般先对结果的偏倚风险进行评估，再对结论的偏倚风险进行评估。两者往往具有一定的逻辑关联，若结果的偏倚风险高，结论的偏倚风险通常也高；若结果的偏倚风险低，结论的偏倚风险可能低也可能高。

1.3.5 综合评价

RoB NMA 工具不对文献检索和选择偏倚进行评估，因此建议先使用 ROBIS 工具系统评价整体偏倚风险，再用 RoB NMA 工具聚焦于 NMA 部分的偏倚风险评估。通过综合考量两种工具的评估结果，可全面反映系统评价的可信度，具体操作流程见图 1。若无法独立完成两种工具的评估，在实施 RoB NMA 评估时需特别关注文献检索策略、纳入排除标准等环节的潜在风险，若发现重大缺陷应在 RoB NMA 最终判断时进行考量。

2 RoB NMA工具条目解读与应用

本文对 RoB NMA 工具中每个条目进行解读，并根据发表在 BMJ 杂志上题为《Comparative effects of drug interventions for the acute management of migraine episodes in adults: systematic review and network Meta-analysis》的研究^[24]进行实例解读。

2.1 干预措施和网状图

条目 1: NMA 中纳入的所有干预措施及对照，对于整个目标人群而言，均为合理的替代方案

解读：本条目主要检验 NMA 中传递性假设问题。若报告中提及的干预措施均为同一疾病的可替代方案，即明确限定了受试者的状态或亚组，使所有纳入试验的受试者在理论上都有可能接受每种比较干预，则应判断为“T”。若临床判断患者特征相似，或所有研究均排除了某项禁忌特征，则应判断为“PT”。若尚不明确是否存在部分受试者不满足干预实施条件，则应判断为“PF”。若发现某些干预只适用于特定亚组，而该亚组与其它研究人群不同质且未被剔除，则应判断为

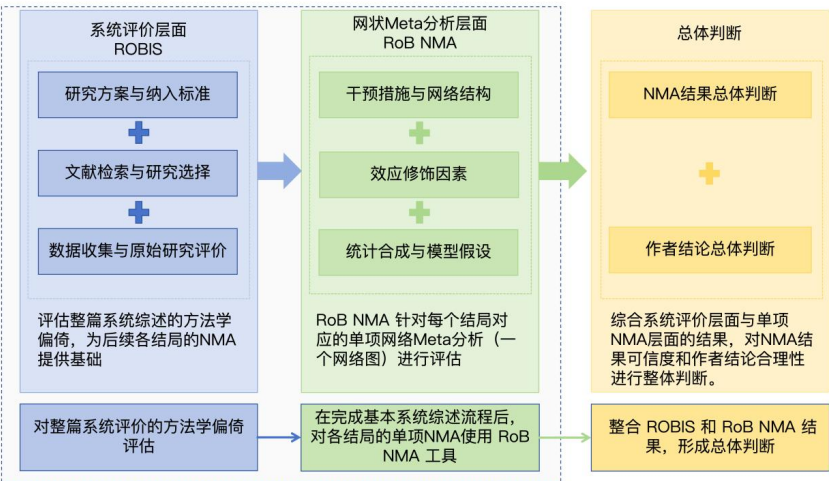


图1 总体ROBIS和RoB NMA评估流程

Figure 1. Overall ROBIS and RoB NMA assessment process

“F”。若报告中信息不能充分判断干预措施是否适用于受试人群，则应判断为“NI”。

实例：研究问题限定为成人门诊急性偏头痛的一线口服单药，比较对象为安慰剂或活性药，并排除了组合用药、非口服制剂与急诊场景，给药途径、处方环境与比较对象在网络内保持一致，故本条目偏倚风险结果判断为“T”。

条目 2：所有的干预措施均已包含在网络中

解读：本条目主要评估 NMA 中的完整性问题，指预先定义的干预措施是否在 NMA 中完整体现，有无基于预期结果排除干预情况。若 NMA 中干预措施无任何不当排除，与现有常见方案一致，则应判断为“T”。若有干预措施被排除，且理由正当，则应判断为“PT”。若有干预措施被排除，但尚不明确排除是否影响，则应判断为“PF”。若有证据表明某些结局相关的干预被不当排除，则应判断为“F”。若报告缺乏是否完全纳入干预的信息，即当前不支持判断 NMA 网络完整性，则应判断为“NI”。

实例：研究将干预措施限定为经监管机构许可的口服单药，作者基于临床实践指南与可传递性假设事先给出两项排除：阿片类用药与急诊场景。排除的干预措施均为事先设定且理由正当，故本条目判断为“PT”。

条目 3：网络中所有干预措施均被正确划分至相应节点

解读：本条目主要评估 NMA 中干预措施分组的合理性，即是否正确将干预措施映射为 NMA 中合适网络节点。若报告中明确提示被合并的干预措施在预期效果中无显著性差异，采取了合理的检验方法，有充足证据支持合并，则应判断为“T”。若存在合并，缺乏充足证据判断合并是否合理，但在临床上判断认为预期效果差异不大，则应判断为“PT”。若合并存在明显不合理情况，将不同预期效果的干预错误合并，则应判断为“F”。若存在合并，缺乏充足证据判断合并是否合理，但在临床上判断认为预期效果可能存在差异，则应判断为“PF”。若报告信息不足以支持判断，则判断为“NI”。

实例：研究主分析以具体口服单药为节点，在许可口服剂量范围内合并不同剂量，后续剂量相关敏感性分析结果未改变主要结论，故本条目判断为“T”。

条目 4：所有比较的干预措施都通过适当的研究内比较链相互连接

解读：本条目主要评估 NMA 中网络结构是否完全连通及连通方式的可靠性。若 NMA 中所有网络均已连通，或有网络节点断开但仅针对连通子网络中的干预措施进行了分析，则应判断为“T”。若为取得感兴趣的间接证据而增加了额外干预或借助外部证据连接了网络，则应判断为“PT”。若作者通过非常规手段连通网络，如将网络与非 RCT 的观察性数据或专家意见等连接，则应判断为“PF”。若存在孤立比较、不连贯比较且未说明处理方法，或 NMA 中存在单臂研究，则应判断为“F”。若现有信息不足以支持判断网络结构是否正确连接，则应判断为“NI”。此外，若多臂研究处理不当也可造成不当连通问题，见后续“统计合成”领域中条目 5 范畴。

实例：研究中报道“所有干预均至少有一项安慰剂对照试验，且多数网络连通性良好”。在主要结局对应的网络中，安慰剂充当了连接度最高的节点，确保不同药物之间存在可追溯的研究内比较链条，故本条目判断为“T”。

2.2 效应修饰因素

条目 1：网络中不同直接比较的结果定义和时间点相似

解读：本条目主要评估 NMA 中采用的结局指标和随访时点是否一致，是否满足比较组间可交换性。若 NMA 明确只分析某个统一的结局及时间点，或比较的结局及时间点相似，差异不足以影响干预效果，则应判断为“T”。若研究所比较的结局及时间点存在一定差异，但有证据显示差异不足以影响干预效果，则应判断为“PT”。若研究所比较的结局及时间点不完全一致存在一定差异，且有证据显示差异可能会影响干预效果，则应判断为“PF”。若研究所比较的结局及时间点存在明显差异，且差异明确会影响干预效果，则应判断为“F”。若现有信息不足以判断研究结局和时间点是否一致，则应判断为“NI”。

实例：研究将主要结局统一为 2 h 无痛与 2~24 h 持续无痛，时间窗与定义在各对比组间一致。在此框架下比较耐受性、安全性等结局，并通过多项敏感性分析未发现违背传递性假设证据，故本条目判断为“T”。

条目 2：网络中不同直接比较的受试者特征

(效应修饰的患者特征)相似

解读: 本条目主要评估 NMA 中受试者基线特征在网络不同比较中的均衡性, 如年龄、性别、疾病严重程度、伴随疾病等。若研究所比较的受试者在关键特征的分布上一致或相似, 则应判断为“T”。若研究所比较受试者在关键特征的分布上有差异, 但经临床判断差异对结局效应影响可能不大, 为合理相似, 则应判断为“PT”。若研究比较受试者在关键特征的分布上有差异, 但经临床判断差异对结局效应可能存在一定影响, 为不合理相似, 则应判断为“PF”。若存在某些已知可能影响结局效应的特征显著不平衡, 则应判断为“F”。若现有受试者特征信息不足以支持判断, 则应判断为“NI”。

实例: 研究纳入群体均为门诊、ICHD 诊断的急性偏头痛成人患者。作者预先将平均年龄、女性比例、基线头痛强度以及是否持续使用预防用药界定为潜在效应修饰因子, 并在各直接比较间比较其分布, 同时采用贝叶斯网络 Meta 回归及多项敏感性分析测试其稳健性, 故本条目判断为“T”。

条目 3: 网络中不同直接比较的研究特征(效应修饰的试验特征)相似

解读: 本条目主要评估 NMA 中研究层面的特征在网络不同比较中的均衡性, 如随机化类型、盲法实施、医疗环境、地区、随访方案等。若各比较试验特征基本均衡, 无显著系统差异, 则判断为“T”。若各比较试验特征存在一定差异, 但经临床判断为合理差异, 可能对结局效应无明显影响, 则应判断为“PT”。若各比较试验特征存在一定差异, 且经临床判断为不合理差异, 可能对结局效应产生明显影响, 则应判断为“PF”。若各比较试验特征存在显著差异并可能影响干预相对效果, 则应判断为“F”。若现有试验特征信息不足以支持判断, 则应判断为“NI”。

实例: 研究对干预进行了前瞻性限定与排除, 两项主要结局均在不允许补救或额外对症用药的条件下评估, 整体减少了共同干预差异。但各试验对预防用药的允许与否仍可能存在残余差异, 提示有一定风险, 故判定为“PT”。

条目 4: 若条目 1、2、3 被判断为“F”或“PF”, 分析是否恰当地考虑了网络中效应修饰因素差异

解读: 本条目为条件性条目, 仅当前面一个

或多个条目被判断为“F”或“PF”时才进行评估, 若条目 1、2、3 均判断为“T”, 代表未发现明显效应修饰差异, 则本条目可判断为“不适用”。若任一条目被判断为“F”或“PF”, 则应重点查看作者在结果部分是否进行 Meta 回归、分层分析等方法调整并报告结果。如果有, 则评估其方法是否合理、是否有足够数据支持。对于患者或研究的特征, 若要在 NMA 中建模一般需要大量试验支撑, 因此如果 NMA 中网络很小, Meta 回归可能无法进行。若作者选用合理的统计方法成功调整了效应修饰因素, 或经敏感性分析显示差异对结局效应无影响, 调整后的结果稳定或改善了其一致性, 则应判断为“T”。若只进行了部分调整仍存在未调整效应修饰因素, 或结局效应分析不充分, 则应判断为“PT”。若作者虽然识别了差异但未进行调整, 只在讨论中提示这一局限, 则偏倚仍存在, 或作者未发现或探索效应修饰因素差异, 则应判断为“PF”或“F”。若现有信息不足以支持判断, 则应判断为“NI”。

实例: 研究在条目 1、2、3 中未被判断为“F”或“PF”, 且作者仍通过网络 Meta 回归与多项敏感性分析探查并处理了潜在修饰差异, 故本条目偏倚风险判断为“不适用”。

2.3 统计合成

条目 1: 分析遵循了各研究内的随机化原则

解读: 本条目主要评估 NMA 分析方法是否尊重各 RCT 内部的随机对照比较, 即在分析框架中对每项研究赋予一个单独的固定效应参数, 以确保每项研究的组内比较不被跨研究直接合并, 从而保留原研究随机化的优势。若研究明确显示 NMA 所用方法能保证每项纳入研究的组间比较均基于研究本身的随机分配, 或使用包含固定效应的模型处理了分组或个体水平的数据, 则应判断为“T”。若作者基于组内对照数据使用了正确的分析方法, 但未能明确所使用模型相关细节。推测应能满足随机化, 但证据略有缺乏, 则应判断为“PT”。若研究使用包含随机效应的模型对分组水平或个体水平数据进行分析, 且不同直接比较之间的平均结果存在差异, 则应判断为“PF”。若作者采用了错误的方法, 如自行合并各试验组数据再进行比较, 没有采用模型分析试验因素, 则应判断为“F”。若现有信息不足以支持判断, 则应判断为“NI”。

实例：研究在频率学随机效应模型的框架内基于研究内相对效应开展网络合并，并处理了多臂试验引入的相关性，对罕见事件采用共同效应 Mantel - Haenszel 方法以保留安全性结局的信息量。上述方法充分尊重研究内随机化，故偏倚风险判断为“T”。

条目 2：未发现存在发表偏倚或其他选择性报告偏倚

解读：本条目评估 NMA 是否存在发表偏倚或选择性报告偏倚。与“干预措施和网状图”领域中条目 2 主要关注预先计划的干预有无遗漏不同，本条目关注过程中有无未纳入的研究或结果，是否进行了发表偏倚分析。若无证据显示存在报告偏倚且对 NMA 中研究检索全面，则应判断为“T”。若有轻微迹象显示进行了排除，但影响较小或与结局效应可能无关，则应判断为“PT”。若 NMA 中存在因报告偏倚或选择性排除而缺失的符合纳入标准的研究且这些缺失研究在结果上存在系统性不同，或缺失数量过大而影响结局效应，则应判断为“PF”。若有确切证据显示 NMA 中存在因发表偏倚或选择性排除，而对结局效应产生了错误估计的情况，则应判断为“F”；若现有信息未显示是否存在可能影响结局效应估计的发表偏倚或选择性排除，则应判断为“NI”。

实例：研究在讨论中明确说明“无法排除部分研究缺失或同一研究被重复计入”的可能性，且未见针对小样本效应或发表偏倚的统计检验结果。但其检索覆盖了试验注册平台与未发表数据，并公开数据集，故本条目判断为“PT”。

条目 3：所有预先计划的分析都已报告，且仅报告了这些分析，或若存在偏离则已予以合理解释

解读：本条目主要评估 NMA 中是否存在报告偏倚，理想情况下评价者应参考该 NMA 的研究方案或 PROSPERO 注册信息进行判断。若方案基本遵循且无遗漏，则应判断为“T”。若 NMA 较原方案有细小偏离但进行了合理解释，则应判断为“PT”。若没有足够的信息显示是否严格遵循原有设计方案，则应判断为“PF”。若明确显示未遵循原有设计方案，或存在明显选择性，则应判断为“F”。若现有信息无法判断是否遵循原因设计方案，则应判断为“NI”。

实例：研究已在 OSF 注册，作者报告了原方案中的主要分析；对于无法开展的亚组或安全性分析也做出合理解释，故应评估为“T”。

条目 4：纳入原始研究中的偏倚水平较低或已在综合分析中予以处理

解读：本条目主要评估组成网络的各项原始研究本身的偏倚问题。可依据作者是否使用 Cochrane RoB 2 或 ROBINS 等工具评价了纳入的原始研究质量。若显示所有试验偏倚风险低，或有少数高风险但进行剔除后不影响结果，则应判断为“T”。若有部分高风险试验，但信息占比不大且经分析证明显示对结局效应影响有限，或进行了统计校正，则应判断为“PT”。若存在相当比例证据显示高偏倚风险且可能会影响结局效应，但作者未进行处理，则应判断为“PF”。若 NMA 中部分重要比较主要靠高偏倚研究支撑且并未进行任何讨论处理，则应判断为“F”。若现有信息无法判断所纳入原始研究质量，则应判断为“NI”。

实例：研究对每项纳入 RCT 的主要结局按研究内层面使用 Cochrane RoB 2 进行风险评估。在证据整合时采用 CINeMA 程序将各研究的偏倚评估结果按贡献比例加权汇总，并对低风险研究开展敏感性分析，故应评估为“T”。

条目 5：对于多臂试验采用了适当的方法进行处理

解读：目前大多数成熟的 NMA 统计软件都会自动处理多臂试验相关性，此条目可着重考察是否有多臂试验或是否自行合并试验数据。若作者明确采用了适当的分析软件或处理策略对多臂试验进行了调整，或无多臂试验，则应判断为“T”。若对于多臂试验处理不详但推断采用了标准模型，则应判断为“PT”。若多臂研究在数据集中未被恰当表示，或所采用的方法未充分考虑随机效应之间的相关性，且未明确说明对该相关性的处理方式，则应判断为“PF”。对于多臂试验若有迹象显示错误处理，如将多臂试验当作独立两臂试验分析，则应判断为“F”。若现有信息不足以支持判断，则应判断为“NI”。

实例：研究方案中声明对多臂相关性作了恰当处理，但因正文章节对实现细节报告有限，本条目判断为“PT”。

条目 6：在比较组内效应量的同质性或异质

性假设方面,研究采用了适当的模型假设

解读: NMA 在统计模型上需要对干预间效应的变异性作出假设,常见假设包括固定效应模型和随机效应模型。若模型选择合理且无证据与数据冲突,则应判断为“T”或“PT”。若模型明显不符合数据,如显示异质性显著却使用了固定效应模型,或模型诊断显示拟合不佳等,则应判断为“PF”或“F”。若现有信息不足以验证是否选择并使用了合适的统计模型,则应判断为“NI”。

实例:研究采用频率学随机效应 NMA 并假定共同异质性,报告 τ^2/I^2 值且以敏感性分析与网络 Meta 回归探查异质性来源。未见与数据冲突,故本条目判断为“T”。

条目 7: 同一效应的直接证据与间接证据间无显著不一致

解读: 本条目主要检查 NMA 一致性假设是否成立,即网络中直接比较结果和通过网络中其他路径得到的间接比较结果是否一致。若网络中不存在或未发现显著不一致,则应判断为“T”。若网络中可能存在不一致,但经合理探索分析未发现,则应判断为“PT”。若发现轻微不一致或检测不一致的测试功效较低,则应判断为“PF”;若发现不一致,则应判断为“F”。若现有信息不足以判断网络中一致性情况,则应判断为“NI”。

实例:研究在全局与局部检验中均提示一些结局在部分比较上存在直接与间接证据的局部不一致,如“2 h 无痛约 8%”等。虽经数据复核与多项敏感性分析后,总体结论较稳健,但不一致性已被观测,故本条目判定为“PF”。

条目 8: 若条目 7 为“F”或“PF”: 直接和间接证据间的不一致性已得到充分处理

解读: 本条目为条目 7 后续条目,只有当条目 7 被判断为“PF”或“F”时才进行评估。当确认存在不一致情况时,作者应当尝试解释和处理不一致。若备选处理措施合理并且解释了不一致来源则应判断为“T”。若进行了处理但仍有残留不一致未解决,则应判断为“PT”。若识别出冲突但未有效处理,如与其他研究进行比较或随机排除原始研究,或未进行处理,则应判断为“PF”或“F”。若现有信息不足以支持判断是否进行了处理,则应判断为“NI”。

实例:研究对已识别的不一致性通过数据复核、机理解释、Meta 回归、限定研究时段及开展多种敏感性分析等方法进行了处理,结论稳健但未证实不一致被完全消除,故本条目判断为“PT”。

条目 9: 若采用贝叶斯分析,所用先验分布设置是合适的

解读: 本条目针对使用贝叶斯统计模型开展 NMA 的情况,要求确认贝叶斯先验是否合理,作者是否进行了先验敏感性分析,且不同先验下结果是否保持稳定。若 NMA 未使用贝叶斯方法则本条目可判断为“不适用”。若使用了信息性先验,且有证据显示先验合理并在敏感性分析中测试了不同先验影响,即先验对结论无实质影响,则应判断为“T”。若使用了现成的、非信息性先验,则应判断为“PT”。若先验设置偏狭窄且无充分理由,即先验明显左右了结果或先验选择欠妥,则应判断为“PF”或“F”。若未报告先验细节,则应判断为“NI”。

实例:研究未使用贝叶斯分析,故本条目不适用。

2.4 领域层面偏倚风险评估

2.4.1 对干预措施及网状图中整体偏倚风险的评估

解读: 结合领域 1 中信号问题答案,对整体偏倚风险进行判断。

实例: 例文所有条目均无明显偏倚,最终领域 1 判定为 LR。

2.4.2 对效应修饰因素的整体偏倚风险评估

解读: 结合领域 2 中信号问题答案,对整体偏倚风险进行判断。

实例: 研究主要结局与时间窗在各比较间较为统一,关键效应修饰因子已事先界定并比较分布,且以网络 Meta 回归与多项敏感性分析验证稳健性。共同干预差异总体可控,对可能的残余差异已通过预设分析与敏感性检验充分探查,领域 2 整体偏倚风险判断为“LR”。

2.4.3 对统计合成领域整体偏倚风险评估

解读: 结合领域 3 中信号问题答案,对整体偏倚风险进行判断。

实例: 领域 3 多数信号问题评定为 T/PT,但研究仍存在未完全消除的局部不一致检验性,因此最终该领域整体偏倚风险判断为“SC”。

2.5 整体偏倚风险评估

2.5.1 NMA结果偏倚风险评估

解读：根据信号问题评估，判断该 NMA 中所得效应估计本身可能受到偏倚影响的程度。主要取决于对领域 1、2、3 的偏倚风险评估：如果任一领域存在高偏倚风险，则整体结果判断为“HR”；若各领域均显示低偏倚风险，则整体结果判断为“LR”；若整体介于中间情况，则判断为“SC”。偏倚风险结果反映的是 NMA 提供的相对效应估计值是否可信。

实例：领域 1 和 2 均提示“LR”，领域 3 提示“SC”，综合判断结果层面偏倚风险为“SC”。

2.5.2 NMA结论偏倚风险评估

解读：若作者在结论中对结果局限性、不确定性和潜在偏倚均做了客观说明，研究中未见明显的倾向性夸大或忽视情况，则结论偏倚风险判断为“LR”；若作者选择性忽视高偏倚因素或关键证据，或做出与证据相矛盾的结论，则结论偏倚风险判断为“HR”；介于两者之间的情况，则判断为“SC”，表明尚存在一定倾向性可能，但并未严重影响结论判断。

实例：研究在讨论部分详细阐述了该 NMA 网络结构、合并方式、面临不一致的处理策略，并承认研究具有部分局限性。此外，研究对短中期不一致也进行了说明，未出现明显过度解读或忽视潜在偏倚的情况。综合判断结论层面偏倚风险为“LR”。

3 结语

RoB NMA 工具为系统评价者和读者提供了一个系统的框架来评估 NMA 中的偏倚风险。但由于方法学文献主要以英文撰写，国内研究者在理解和应用 RoB NMA 工具时可能会遇到一些源于语言和概念理解的挑战。本文旨在通过本土化解读，协助研究者理解并使用该工具完善 NMA 研究的规范度。

在应用 RoB NMA 过程中应避免问题僵化，如有时评估者面对信息不足，可能倾向于频繁使用“NI”。但 RoB NMA 指南强调应尽量做出倾向判断，滥用“NI”会减低评价的区分度。同时评估者应尽力通过合理推理或咨询临床专家来进行判断，切忌因作者未提及相关信息就报告“NI”。对于领域风险判断并非使用“计点法”进行汇总，

而应综合考虑偏倚的严重程度和对结论的影响做出判断。

RoB NMA 工具通过构建的 17 项结构化信号条目，为研究者提供可靠的方法支持，可在研究设计阶段和结果解释环节提供针对性指导，帮助作者充分揭示研究中存在的潜在局限性与不确定性。对于决策者而言，参考 RoB NMA 工具评估结果亦有助于精准识别 NMA 的风险点，在必要时对结论予以慎重保留或进一步核实。但该工具仍需多学科研究者持续验证，国内研究者在使用过程中不应将 RoB NMA 视为一个单纯的评估工具，更应将其作为设计和执行高质量 NMA 研究的蓝图。同时，坚持多学科协作，高质量的 NMA 评估和执行离不开方法学家或统计学家与临床领域专家的紧密合作，并在实践中不断完善其评估效能。建议未来研究者在实践应用中积极反馈使用体验，针对性地提出改进建议，共同推动该工具的迭代升级，使其更好适应 NMA 的评估需求。

伦理声明：不适用

作者贡献：文献检索、清单翻译、论文撰写：赵文硕；文献整理、论文校对：赵文硕、刘明、高亚；论文修订：刘明、史纪元、李伦、杨丰文、薛付忠、高亚；论文审校：高亚

数据获取：不适用

利益冲突声明：无

致谢：不适用

参考文献

- Willis BH, Quigley M. The assessment of the quality of reporting of Meta-analyses in diagnostic research: a systematic review[J]. BMC Med Res Methodol, 2011, 11: 163. DOI: [10.1186/1471-2288-11-163](https://doi.org/10.1186/1471-2288-11-163).
- Zhao H, Hodges JS, Ma H, et al. Hierarchical Bayesian approaches for detecting inconsistency in network Meta-analysis[J]. Stat Med, 2016, 35(20): 3524–3536. DOI: [10.1002/sim.6938](https://doi.org/10.1002/sim.6938).
- Chaimani A, Caldwell DM, Li T, et al. Undertaking network Meta-analyses. Cochrane handbook for systematic reviews of interventions[EB/OL]. [2025-07-01]. DOI: [10.1002/9781119536604.ch11](https://doi.org/10.1002/9781119536604.ch11).
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons[J]. Stat Med, 2004, 23(20): 3105–3124. DOI: [10.1002/sim.1875](https://doi.org/10.1002/sim.1875).
- Mills EJ, Bansback N, Ghement I, et al. Multiple treatment comparison Meta-analyses: a step forward into complexity[J]. Clin Epidemiol, 2011, 3: 193–202. DOI: [10.2147/CLEP.S16526](https://doi.org/10.2147/CLEP.S16526).

- 6 Mills EJ, Kanter S, Thorlund K, et al. The effects of excluding treatments from network Meta-analyses: survey[J]. BMJ, 2013, 347: f5195. DOI: [10.1136/bmj.f5195](https://doi.org/10.1136/bmj.f5195).
- 7 Bafeta A, Trinquart L, Seror R, et al. Reporting of results from network Meta-analyses: methodological systematic review[J]. BMJ, 2014, 348: g1741. DOI: [10.1136/bmj.g1741](https://doi.org/10.1136/bmj.g1741).
- 8 Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence[J]. BMJ, 2005, 331(7521): 897-900. DOI: [10.1136/bmj.331.7521.897](https://doi.org/10.1136/bmj.331.7521.897).
- 9 Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network Meta-analysis[J]. BMJ, 2013, 346: f2914. DOI: [10.1136/bmj.f2914](https://doi.org/10.1136/bmj.f2914).
- 10 Li T, Puhon MA, Vedula SS, et al. Network Meta-analysis—highly attractive but more methodological research is needed[J]. BMC Med, 2011, 9: 79. DOI: [10.1186/1741-7015-9-79](https://doi.org/10.1186/1741-7015-9-79).
- 11 Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment Meta-analysis: an overview and tutorial[J]. J Clin Epidemiol, 2011, 64(2): 163-171. DOI: [10.1016/j.jclinepi.2010.03.016](https://doi.org/10.1016/j.jclinepi.2010.03.016).
- 12 Brunström M, Thomopoulos C, Carlberg B, et al. Methodological aspects of Meta-analyses assessing the effect of blood pressure-lowering treatment on clinical outcomes[J]. Hypertension, 2022, 79(3): 491-504. DOI: [10.1161/HYPERTENSIONAHA.121.18413](https://doi.org/10.1161/HYPERTENSIONAHA.121.18413).
- 13 Whiting P, Savović J, Higgins JP, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed[J]. J Clin Epidemiol, 2016, 69: 225-234. DOI: [10.1016/j.jclinepi.2015.06.005](https://doi.org/10.1016/j.jclinepi.2015.06.005).
- 14 Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both[J]. BMJ, 2017, 358: j4008. DOI: [10.1136/bmj.j4008](https://doi.org/10.1136/bmj.j4008).
- 15 吴琼芳, 丁泓帆, 邓围, 等. ROBIS: 评估系统评价偏倚风险的新工具[J]. 中国循证医学杂志, 2015, 15(12): 1454-1457. [Wu QF, Ding HF, Deng W, et al. ROBIS: a new tool to assess risk of bias in systematic reviews[J]. Chinese Journal of Evidence-Based Medicine, 2015, 15(12): 1454-1457.] DOI: [CNKI:SUN:ZZXZ.0.2015-12-013](https://doi.org/CNKI:SUN:ZZXZ.0.2015-12-013).
- 16 陶欢, 杨乐天, 平安, 等. 随机或非随机防治性研究系统评价的质量评价工具 AMSTAR 2 解读[J]. 中国循证医学杂志, 2018, 18(1): 101-108. [Tao H, Yang LT, Ping A, et al. Interpretation of AMSTAR 2: a critical appraisal tool for systematic reviews that include randomized or non-randomized studies of healthcare interventions[J]. Chinese Journal of Evidence-Based Medicine, 2018, 18(1): 101-108.] DOI: [CNKI:SUN:ZZXZ.0.2018-01-017](https://doi.org/CNKI:SUN:ZZXZ.0.2018-01-017).
- 17 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials[J]. BMJ, 2019, 366: l4898. DOI: [10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898).
- 18 刘津池, 刘畅, 华成舸. 随机对照试验偏倚风险评价工具 RoB2(2019 修订版) 解读[J]. 中国循证医学杂志, 2021, 21(6): 737-744. [Liu JC, Liu C, Hua CG. Risk bias assessment tool RoB2 (revised version 2019) for randomized controlled trial: an interpretation[J]. Chinese Journal of Evidence-Based Medicine, 2021, 21(6): 737-744.] DOI: [10.7507/1672-2531.202011144](https://doi.org/10.7507/1672-2531.202011144).
- 19 Page MJ, Sterne JAC, Boutron I, et al. ROB-ME: a tool for assessing risk of bias due to missing evidence in systematic reviews with Meta-analysis[J]. BMJ, 2023, 383: e076754. DOI: [10.1136/bmj-2023-076754](https://doi.org/10.1136/bmj-2023-076754).
- 20 Chiochia V, Nikolakopoulou A, Higgins JPT, et al. ROB-MEN: a tool to assess risk of bias due to missing evidence in network Meta-analysis[J]. BMC Med, 2021, 19(1): 304. DOI: [10.1186/s12916-021-02166-3](https://doi.org/10.1186/s12916-021-02166-3).
- 21 邓湘金, 胡玉鹏, 高歌, 等. 不报告偏倚风险评价工具 ROB-ME 中文解读[J]. 中国循证医学杂志, 2023, 23(7): 843-848. [Deng XJ, Hu YP, Gao G, et al. Risk of bias due to missing evidence (ROB-ME): a Chinese interpretation[J]. Chinese Journal of Evidence-Based Medicine, 2023, 23(7): 843-848.] DOI: [10.7507/1672-2531.202303107](https://doi.org/10.7507/1672-2531.202303107).
- 22 胡玉鹏, 朱涛, 邓湘金, 等. 网状 Meta 分析证据缺失偏倚风险评价工具 ROB-MEN 中文解读[J]. 中国循证医学杂志, 2024, 24(4): 451-458. [Hu YP, Zhu T, Deng XJ, et al. A Chinese introduction to risk of bias due to missing evidence in network Meta-analysis(ROB-MEN)[J]. Chinese Journal of Evidence-Based Medicine, 2024, 24(4): 451-458.] <https://www.cnki.com.cn/Article/CJFDTotol-ZZXZ202404009.htm>
- 23 Lunny C, Higgins JPT, White IR, et al. Risk of bias in network Meta-analysis (RoB NMA) tool[J]. BMJ, 2025, 388: e079839. DOI: [10.1136/bmj-2024-079839](https://doi.org/10.1136/bmj-2024-079839).
- 24 Karlsson WK, Ostinelli EG, Zhuang ZA, et al. Comparative effects of drug interventions for the acute management of migraine episodes in adults: systematic review and network Meta-analysis[J]. BMJ, 2024, 386: e080107. DOI: [10.1136/bmj-2024-080107](https://doi.org/10.1136/bmj-2024-080107).

收稿日期: 2025 年 07 月 30 日 修回日期: 2025 年 11 月 22 日
本文编辑: 桂裕亮 曹 越

引用本文: 赵文硕, 刘明, 史纪元, 等. 网状Meta分析偏倚风险评估工具RoB NMA解读[J]. 医学新知, 2025, 35(12): 1454-1463. DOI: [10.12173/j.issn.1004-5511.202507209](https://doi.org/10.12173/j.issn.1004-5511.202507209).
Zhao WS, Liu M, Shi JY, et al. An introduction to Risk of Bias in Network Meta-analysis (RoB NMA) tool[J]. Yixue Xinzhi Zazhi, 2025, 35(12): 1454-1463. DOI: [10.12173/j.issn.1004-5511.202507209](https://doi.org/10.12173/j.issn.1004-5511.202507209).