

· 循证研究与临床转化 ·

网络Meta分析研究进展系列（十八）： 网络Meta分析的GRADE证据分级

武珊珊¹, 田金徽², 杨智荣³, 刘凤琪⁴, 董圣杰⁵, 张天嵩⁶, 孙凤⁴

【摘要】本文主要介绍了GRADE在网络Meta分析中应用的原理和方法，以期通过网络Meta分析结果进行证据质量评价提供参考。

【关键词】网络Meta分析；证据分级；GRADE

【中图分类号】R4

【文献标志码】A

开放科学（源服务）标识码（OSID）



Advances in methodology of network meta-analysis (18): GRADE evidence certainty in Network Meta-Analysis Wu Shanshan*, Tian Jinhui, Yang Zhirong, Liu Fengqi, Dong Shengjie, Zhang Tiansong, Sun Feng.
*Department of Clinical Epidemiology and EBM, National Clinical Research Center for Digestive Diseases, Beijing Friendship Hospital, Capital Medical University, Beijing, 100050, China.

Corresponding authors: Sun Feng, E-mail: sunfeng@bjmu.edu.cn

【Abstract】This paper introduces the principles and methods of The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) system in the application of network Meta-analysis, in order to facilitate and consolidate understanding NMA findings and GRADE certainty of evidence.

【Key words】Network meta-analysis; Evidence certainty; GRADE

网络Meta分析（NMA）扩展了传统Meta分析的方法，可在一个证据体中同时评估针对同一疾病的多个干预措施，在证据一致性的假设下，将直接证据和间接证据合并得到干预措施效果的综合估计，并可获得干预措施间的优劣排序，筛选出最佳干预措施，为医疗保健决策提供信息^[1,2]。相对传统Meta分析，NMA在临床实践中更有参考价值。如未对NMA结果进行证据质量分级，可能会对读者在理解NMA结论的真实性和可靠性方面造成偏颇，甚至误导。因而，对NMA的证据结果进行质量分级，有助于临床医生及决策者了解NMA效应估计值接近真实值的把握度，从而更好地指导临床实践。

随着NMA的制定流程以及方法学的日趋成熟，国际药物经济学和结果研究协会（ISPOR）先后制定了3部专门论述如何制定、解读和评价NMA的手册^[3-5]，但大多仅通过评价纳入研究的偏倚风险来推测总体的证据质量，由于该

方法未考虑到其他偏倚因素，因而存在很大局限性。2014年，推荐分级的评价、制定与评估（GRADE）工作组在British Medical Journal（BMJ）发表了将GRADE方法用于NMA证据质量分级的相关指南^[6]，为NMA制作者基于GRADE方法对相应结果进行证据分级提供了借鉴和指导。随后，希腊大学的Salanti等学者基于NMA自身的特点对上述GRADE分级的主要原则和定义进行了一定的修改和补充，充分考虑到直接比较和间接比较对最终结果的贡献程度，该方法于2014年发表在PLOS ONE杂志^[7]。2019年，瑞士伯尔尼大学社会与预防医学机构基于上述Salanti等学者的理论开发出一款在线的证据质量评价工具——CINeMA（Confidence in Network Meta-Analysis），专门用于评价NMA的证据质量^[8]。该软件的开发，大大提升了NMA证据质量评价的效率。近年来，随着NMA方法学的不断完善，GRADE工作组又连续发表了3篇系列文章，对GRADE方法在NMA中的应用进行了进一步完善和补充^[9-11]。此外，由于NMA综合了多种干预措施直接比较和间接比较的证据，产生的研究结果较多，关于如何报告NMA结果更有助于决策者和临床医生高效做出决策，GRADE工作组也启动了GRADE NMA-SoF（Summary of Findings）表格制订工作，历时7年于2019年正式发布^[12]。2020年，GRADE工作又提出了2种NMA结论形成的方法，即部分背景化框架与最小背景化框架^[13,14]。本文结合GRADE工作组及Salanti等学者的相关系

基金项目：国家自然科学基金(72074011);北京大学教育教学项目(2020YBC02)

作者单位：¹ 100050 北京,首都医科大学附属北京友谊医院临床流行病学与循证医学研究室,国家消化系统疾病临床医学研究中心;² 730000 兰州,兰州大学循证医学中心 甘肃省循证医学与临床转化重点实验室;³ 518055 深圳,中国科学院深圳理工大学计算机科学与控制工程学院;⁴ 100191 北京,北京大学循证医学中心,北京大学公共卫生学院流行病学与卫生统计学系;⁵ 264000 烟台,烟台市烟台山医院骨关节科;⁶ 200040 上海,复旦大学附属静安区中心医院中医科

通讯作者：孙凤, E-mail: sunfeng@bjmu.edu.cn

doi: 10.3969/j.issn.1674-4055.2021.11.02

列文章，对NMA证据分级的方法学进展进行介绍，以期为NMA制定者提供一定的参考。

1 GRADE证据分级系统的基本概念与原理

GRADE工作组开发了一整套科学的证据分级体系用于评价不同干预措施对重要临床结局产生影响的可信程度，即证据质量^[15-21]。其评价的主要原则包括5个降级因素和3个升级因素。其中，5个降级因素分别为：①纳入研究的方法学质量（risk of bias，偏倚风险），②研究关注的人群、干预措施及结局指标的外推性（indirectness，间接性），③不同研究间结果的一致程度（inconsistency，不一致性），④不同研究合并结果的精确程度（imprecision，不精确性），⑤对符合标准研究纳入的全面程度（publication bias，发表偏倚）；3个升级因素分别为：①大效应量（large effect），②剂量效应关系（dose-response gradient），③相关混杂的控制（plausible confounding）。

对于随机对照试验或只纳入随机对照试验的系统评价，证据质量的初始等级即为高级别，若存在5个降级因素中的某一个或多个，则可降级为中级别、低级别或极低级别。对于观察性研究或只纳入观察性研究的系统评价而言，证据质量的初始等级为低级别，若存在5个降级因素中的一个或多个，则可进一步降级至极低级别；反之，若存在任何一个或多个升级因素，则可考虑证据质量的升级。需要注意的是，GRADE证据质量分级不是对单个临床研究或系统评价的质量分级，是针对报告了某个结局指标的证据体的质量分级。

2 GRADE在网络Meta分析中的应用

由于目前NMA主要基于随机对照试验这一临床研究类型，因此无论是GRADE工作组或Salanti等提出的方法学体系，GRADE在NMA证据质量分级中的应用均主要考虑上述5个降级因素。

2.1 第一种方法：GRADE工作组方法学体系 与其他类型系统综述和传统Meta分析相比，NMA有其特殊性，即同时纳入直接比较和间接比较的证据，因此其证据分级过程更为复杂。除要考虑GRADE中5个降级因素之外，还需考虑间接比较中不同组别在人群基线特征、共同对照和结果测量方面的不可传递性，以及直接和间接比较结果的不一致性。2014年GRADE工作组提出了对NMA进行证据质量分级的四步法^[6]，即首先呈现两个干预措施之间直接和间接比较的效应量与可信区间，其次分别对每一组直接比较和间接比较进行证据质量分级，然后呈现NMA的结果，最后评估NMA结果的证据质量。直接比较证据的分级

方法与传统Meta分析GRADE证据分级类似。间接比较证据分级按照就低原则依据其直接比较中证据质量低的组别进行判定。基于混合比较的NMA结果证据质量则依据就高原则将直接和间接比较中证据质量高的组别作为最终等级。

该方法在NMA纳入的干预措施或结局指标数量较少时较为快捷可行，但当NMA中纳入的干预措施或结局指标数目较多时，需逐一针对每个结局指标中的每一个直接比较和间接比较进行证据分级，分级任务繁重耗时。此外，该方法在最终NMA结果证据等级判定时仅考虑了就高原原则，未充分考虑直接比较和间接比较的贡献大小，可能导致结果偏颇。近年来GRADE工作组对GRADE方法在NMA中的应用进行了修订和完善^[9-11]。围绕2018年Brignardello-Petersen等学者制定的分级流程^[9]，王琪等^[22]对更新后的证据分级步骤进行了解读，具体如下（图1）：第1步，将直接比较、间接比较和NMA混合比较的效应量和可信区间分开呈现；第2步，不考虑不精确性因素，对每个比较组的直接比较证据进行质量分级；若直接比较的证据质量等级为高，且对NMA结果的贡献大于等于间接比较，则无需对间接比较证据进行质量分级，最终NMA结果的证据质量依据直接比较的证据质量判定即可，这样既科学地考虑了直接比较和间接比较对最终结果的贡献程度，又提高了NMA结果证据分级的效率；反之，则需进行间接比较的证据质量分级；第3步，基于形成间接证据一阶环路的直接证据质量（不考虑不精确性因素），采取就低原则确定间接证据质量，此外尚需考虑相似性；第4步，基于直接证据和/或间接证据质量等级，考虑连贯性和精确性，最终确定和呈现NMA结果的证据质量。

在更新后的证据分级过程中，需注意以下几点：①在对直接比较进行证据质量分级时，对间接性的降级需谨慎，理论上研究关注的人群、干预措施及结局指标的外推性始终存在，只有存在重大间接性时才会考虑降级，同时给予降级理由的详细说明；②由于NMA合并了直接比较和间接比较的结果，因此在对直接比较和间接比较分别

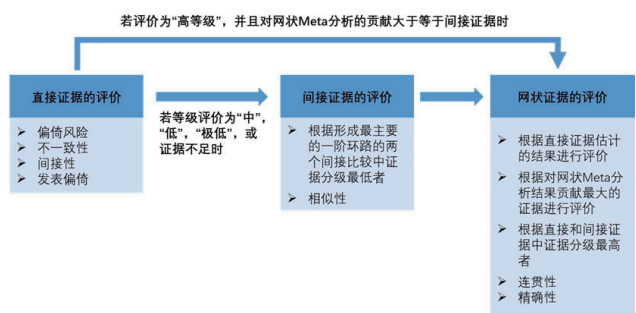


图1 更新后的NMA-GRADE分级流程

进行证据质量分级时无需考虑不精确性, NMA结果的精确性均高于直接和间接证据; ③当直接比较的证据等级为高, 但直接比较对NMA的贡献较间接比较更小时, 需考虑间接比较的证据质量等级; ④在进行间接比较的证据质量分级时应考虑不可传递性, 即不同组别间在人群基线特征、共同对照及结果测量等方面是否存在差异, 若存在较大差异, 需依据不可传递性将该间接证据质量再降低一级; ⑤在对NMA结果进行证据分级时对不一致性的降级需谨慎, 还应考虑直接比较和间接比较对NMA效应量的贡献度; 若NMA效应量主要来自于直接比较或间接比较, 则可认为不一致性对NMA结果影响较小不进行降级; 反之, 若直接和间接比较对NMA的贡献度相当, 则需因不一致性对NMA结果进行降级。

2.2 第二种方法: Salanti等方法学体系 与上述GRADE工作组提出的方法学体系不同, Salanti等学者^[7]基于NMA自身的特点, 将NMA作为一个整体, 结合证据贡献图, 充分考虑了直接比较和间接比较对最终NMA结果的贡献程度, 从6个维度(研究的局限性即偏倚风险、间接性、不精确性、异质性、不一致性和发表偏倚)分别对每个比较的NMA结果及最终排序进行证据质量的分级。每个维度均可根据其严重程度分为不严重(no concern, 不降级)、严重(some concern, 降一级)和非常严重(major concern, 降两级), 最终每个比较的NMA结果及最终排序的证据质量等级与GRADE体系一致, 即高、中、低和极低四个等级。值得注意的是, 当NMA中纳入的干预措施或结局指标数量较多时, 基于该方法进行证据质量分级同样繁重耗时。

采用该方法进行证据等级评价时需注意: ①在评估每个比较NMA结果是否根据研究局限性进行降级时, 需结合每个研究的偏倚风险评价结果来判断NMA中每个直接比较的偏倚风险, 进一步分别为低、中和高偏倚风险的直接比较赋以不同的权重(如分别为低、中和高偏倚风险的直接比较赋以0、-1和-2的权重), 结合证据贡献图, 来判断是否进行降级以及降1级或降2级; ②在从间接性维度进行评估时, 需注意间接性不仅考虑研究人群、干预措施和结果指标的间接性, 还应考虑NMA中间接比较的假设(即相似性)是否成立; ③不精确性的判断可直接通过NMA的可信区间是否跨越无效线和包含最小临床重要差值来判断。若可信区间不包含最小临床重要差值, 则认为不存在不精确性, 无需降级; 若可信区间包含部分最小临床重要差值, 则认为存在一定程度的不精确性, 降一级; 若可信区间完全包含最小临

床重要差值, 则认为存在不精确性, 可降两级; ④异质性的评估需结合每个比较的NMA可信区间及预测区间结果。若NMA可信区间与预测区间结论一致, 则可认为不存在明显的异质性, 无需降级; 反之, 则认为存在较大的异质性, 可对该比较降一级; ⑤不一致性是指直接比较和间接比较结果之间是否存在显著的统计学差异, 目前有多种统计学方法来评估局部不一致性, 如节点拆分法、基于环的不一致性检验等; ⑥可采用校正比较漏斗图或选择模型等进行发表偏倚的评估, 但需注意当纳入研究数目较少时, 无论校正比较漏斗图还是各种模型检验方法的统计学把握度均较低, 因而更应注意NMA检索策略的制定是否足够全面和完善; 若检索策略制定的足够全面, 则即使纳入的研究数目较少, 也可不降级。

3 NMA证据分级在线应用程序—CINeMA

基于上述Salanti等学者提出的NMA证据分级方法学体系, 瑞士伯尔尼大学社会与预防医学机构2019年开发出一款在线的证据质量评价工具—CINeMA, 专门用于评价NMA的证据质量等级, 可通过其官方网站直接访问(<https://cinema.ispm.unibe.ch/>)^[8]。该软件可供用户免费使用, 且无需登录, 半自动化过程大大提升了NMA证据质量评价的效率。

CINeMA通过调用R软件的netmeta程序包, 可绘制网状证据图, 并计算NMA的贡献矩阵, 包括每个研究和每个直接比较对每个比较组NMA结果的贡献度, 从而评价每个比较的NMA结果是否根据研究局限性和间接性两个维度进行降级。此外, CINeMA还可基于频率学框架进行NMA的统计学分析, 使用者可选择数据分析模型(固定效应模型或随机效应模型)和拟分析的干预措施, 可选择的效应指标包括针对二分类结局指标的比值比(OR)、相对风险比(RR)和率差(RD), 以及针对连续型结局指标的均数差(MD)和标准化均数差(SMD)。运行结果中包含NMA中任意两种干预措施相互比较的效应值及95%置信区间、异质性检验结果(每种比较的预测区间)及不一致性检验结果(节点拆分法), 从而帮助使用者从不精确性、异质性及不一致性三个维度来对每个比较的NMA结果进行证据质量的分级, 具体结果示例请见图2。关于CINeMA在线应用程序的使用流程可参见Nikolakopoulou等^[23]于2020年发表在PLOS MEDICINE的相关文章, 国内的王琪等^[24]也对CINeMA在线应用程序的使用以“高血压病患者服用降压药物后糖尿病发生风险的网状Meta分析”为案例进行了具体的详细介绍, 该案例的具体数据可见Elliott于2007年发表在

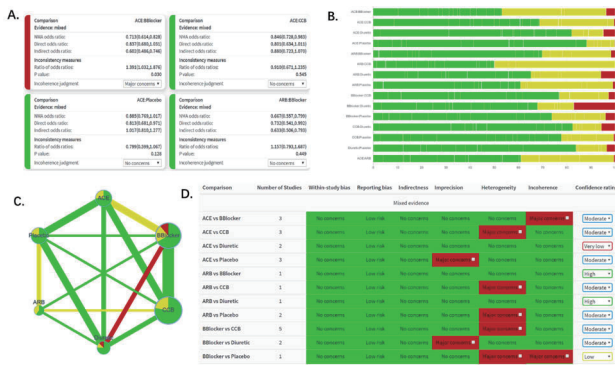


图2 CINeMA在线应用程序分析部分结果示例

注：A：局部不一致性检验；B：不同比较的偏倚风险；C：网状证据图；D：证据质量分级

Lancet的文章^[25]。

虽然使用CINeMA在线应用程序大大优化了NMA证据质量分级的效率，但该工具也存在一定局限性，目前只能实现对单一结局指标的证据分级，若一个NMA中存在多个结局指标，则需重复进行证据分级的过程。此外，在证据分级过程中，诸如间接性、发表偏倚等维度的评价难免存在一定的主观性。

4 GRADE证据总结表

临床医生或决策者在使用证据时，根据不同用途对研究结果的呈现形式有不同的需求。GRADE工作组所提出的结果总结表SoF在传统系统评价中的实用性已经得到了证明，其对证据的总结非常有帮助。NMA综合了直接比较和间接比较的证据，如何报告总结NMA结果更有助于临床决策，GRADE工作组基于质性研究的方法，通过成立专家指导小组、召开头脑风暴会议、使用者测试等对NMA-SoF表格涵盖的内容和呈现形式进行了全面考量和取舍，经过4轮沟通产生了6个候选的NMA-SoF表格并达成一致，建议最终的NMA-SoF表格^[12]应包含以下6方面内容：①临床问题的具体信息（PICO）；②网络关系图的构建，包括研究数量、样本量、干预措施数量及表现形式；③相对效应和绝对效应的估计，包括直接比较效应估计及置信区间和间接比较效应估计及置信区间；④证据的确定性（GRADE证据质量分级结果）；⑤干预措施效应大小的排序（基于累积排序曲线下面积的排序及置信区间）；⑥结果的解释。王巍巍等^[26]也对NMA-SoF表格的主要内容进行了详细介绍，并举例说明SoF表的使用方法和注意事项。

NMA-SoF表于2019年11月发表后迅速成为“高被引论文”，目前Cochrane协作组已推荐使用NMA-SoF表格，其推广和应用不仅可以提高NMA结果报告的规范性，也在一定程度上帮助研

究者开展科学规范的研究设计，提升研究质量。对于指南制订者而言，NMA-SoF不仅提供证据质量分级的信息，也提供了绝对效应值，方便指南制订参考临床决策阈值^[13,14,27,28]进行判断是否给出推荐。对于卫生保健专业人员而言，NMA-SoF可以高效地展示NMA的研究结果，提高证据的实用性，帮助其做出更加明智的决策。

5 基于GRADE呈现NMA结论

为从NMA的大量结果（包括相对效应量、排序及证据可信度等）中正确筛选出最佳干预措施以指导临床实践，2020年GRADE工作组提出了分别基于最小背景化框架^[13]和部分背景化框架^[14]呈现NMA结论的两种方法。背景化指根据特定的阈值或范围对证据可信度进行分级。在政策制定与指南应用的背景下，背景化需考虑当地环境、资源可及性、法律法规、文化、可及性和公平性等因素。GRADE工作组通过对结局呈现的选择、不一致性的考虑和受众群体价值观的重视等方面的考量，将背景化程度分为“完全背景化”、“部分背景化”与“最小背景化”三类。若在结果的分析中不考虑患者的价值观和偏好，则为部分背景化框架；最小背景化则是将所需考虑的背景因素降至最低，最大限度减少价值判断。

部分背景化框架是指在对于干预措施进行分类时，尽可能综合考虑效应的大小、证据可信度及排序。作者首先依据具有结局指标确定好效应阈值和参考组，基于比较每种干预措施与参考组比较的点估计值将干预措施分为无效、较小、中等与较大效应，干预措施可归类为有利或不利两种情况；然后基于GRADE分级将干预措施分为高或中等、低或极低证据等级两类，并保证与排序结果一致，从而找到NMA中的最佳干预措施。最小背景化框架是指针对NMA的某一特定结局，作者首先选择决策阈值和参考组，对每个干预措施与参考组相比效应量的可信区间是否与决策阈值相交做出判断，然后将干预措施进行分类（可分为比参考组更好、更差和无差异）。随后基于GRADE分级在此前分类的基础上将干预措施分为2大类证据等级，并保证SUCRA排序结果同分组结果一致，以找到NMA中最佳的干预措施。朱鸿飞、栗梦婷等^[27,28]分别对部分背景化框架及最小背景化框架的原理进行了解读，并结合实例进行了阐述。

部分背景化及最小背景化框架均综合考虑了效应大小、证据可信度及排序，以寻找NMA中的最佳干预措施，但部分背景化框架评估时不强调研究结果的精确性，侧重于通过点估计值和排序将干预措施分为无效、微小、中等和重大效

应四个类别，同一个类别间的干预措施效应值通常有差异；而最小背景化框架强调研究结果的精确性及证据的可信度，侧重于通过可信区间与决策阈值的比较将干预措施进行分类，同一类别间的干预措施效果不存在差异。

6 结语

本文主要介绍了基于GRADE原则进行NMA证据质量分级两种方法的基本理论，并阐述了GRADE方法在NMA证据质量方面的最新进展，包括在线应用程序CINeMA和NMA结果总结表的使用，应用这些新进展、新工具时应该注意，尽管最大限度地提高研究效率是正确的，但使用这些策略需要仔细衡量。虽然应用GRADE进行NMA的证据质量分级具有一定主观性，但GRADE仍是当前对NMA进行证据分级最客观、最有价值的工具，且可以结合临床决策阈值^[13,14,27,28]给出最终推荐。随着NMA方法学的日益成熟及研究质量的不断提高，GRADE在NMA中的应用将会更加广泛。

参考文献

[1] Abdelhamid AS, Loke YK, Parekh-Bhurke S, et al. Use of indirect comparison methods in systematic reviews: a survey of Cochrane review authors[J]. Res Synth Methods, 2012, 3(2): 71-9.

[2] Gurevitch J, Koricheva J, Nakagawa S, et al. Meta-analysis and the science of research synthesis[J]. nature, 2018, 555(7691): 175-82.

[3] Jansen JP, Fleurence R, Devine B, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1[J]. Value Health, 2011, 14(4): 417-28.

[4] Hoaglin DC, Hawkins N, Jansen JP, et al. Conducting indirect-treatment-comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2[J]. Value Health, 2011, 14(4): 429-37.

[5] Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report[J]. Value Health, 2014, 17(2): 157-73.

[6] Puhan MA, Schunemann HJ, Murad MH, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis[J]. BMJ, 2014, 349: g5630.

[7] Salanti G, Del Giovane C, Chaimani A, et al. Evaluating the quality of evidence from a network meta-analysis[J]. PLoS One, 2014, 9(7): e99682.

[8] Institute of Social and Preventive Medicine. CINeMA confidence in network meta-analysis [Software]. Berne: University of Bern, 2017. Available at: www.cinema.ispm.unibe.ch. [9] Brignardello-Petersen R, Bonner A, Alexander PE, et al. Advances in the GRADE approach

to rate the certainty in estimates from a network meta-analysis[J]. J Clin Epidemiol, 2018, 93: 36-44.

[10] Brignardello-Petersen R, Mustafa RA, Siemieniuk RAC, et al. GRADE approach to rate the certainty from a network meta-analysis: addressing incoherence[J]. J Clin Epidemiol, 2019, 108: 77-85.

[11] Brignardello-Petersen R, Murad MH, Walter SD, et al. GRADE approach to rate the certainty from a network meta-analysis: avoiding spurious judgments of imprecision in sparse networks[J]. J Clin Epidemiol, 2019, 105: 60-7. [12] Yepes-Nuñez JJ, Li S-A, Guyatt G, et al. Development of the summary of findings table for network meta-analysis[J]. J Clin Epidemiol, 2019, 115: 1-13.

[13] Brignardello-Petersen R, Florez ID, Izcovich A, et al. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework[J]. BMJ, 2020, 371: m3900.15

[14] Brignardello-Petersen R, Izcovich A, Rochwerg B, et al. GRADE approach to drawing conclusions from a network meta-analysis using a partially contextualised framework[J]. BMJ, 2020, 371: m3907.

[15] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness[J]. J Clin Epidemiol, 2011, 64(12): 1303-10.

[16] Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence[J]. J Clin Epidemiol, 2011, 64(4): 401-6.

[17] Guyatt GH, Oxman AD, Schünemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology[J]. J Clin Epidemiol, 2011, 64(4): 380-2.

[18] Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias)[J]. J Clin Epidemiol, 2011, 64(4): 407-15.

[19] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency[J]. J Clin Epidemiol, 2011, 64(12): 1294-302.

[20] Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 6. Rating the quality of evidence—imprecision[J]. J Clin Epidemiol, 2011, 64(12): 1283-93.

[21] Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias[J]. J Clin Epidemiol, 2011, 64(12): 1277-82.

[22] 王琪, 王建成, 潘蓓, 等. 网状Meta分析证据质量分级: GRADE方法学前沿与进展[J]. 中国循证医学杂志, 2020, 20(8): 979-85.

[23] Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. CINeMA: An approach for assessing confidence in the results of a network meta-analysis[J]. PLoS Med, 2020, 17(4): e1003082.

[24] 王琪, 王韵华, 赖鸿鹄, 等. 网状Meta分析证据质量分级: CINeMA在线应用程序简介[J]. 中国循证医学杂志, 2020, 20(9): 1111-6.

[25] Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis[J]. The Lancet, 2007, 369(9557): 201-7.

[26] 王巍巍, 杨智荣, 孙凤, 等. 网状Meta分析GRADE证据总结表的制订、解读与应用[J]. 中国循证医学杂志, 2020, 20(12): 1471-6.

[27] 栗梦婷, 朱鸿飞, 后亮瑛, 等. 网状Meta分析结论形成的方法: 最小背景化框架[J]. 中国循证医学杂志, 2021, 21(9): 1102-9.

[28] 朱鸿飞, 栗梦婷, 后亮瑛, 等. 网状Meta分析结论形成的方法: 部分背景化框架[J]. 中国循证医学杂志, 2021, 21(9): 1110-6.

本文编辑: 孙竹