

## · 循证研究与临床转化 ·

## 网络Meta分析研究进展系列之一：概述

孙凤<sup>1</sup>, 杨智荣<sup>2</sup>, 张天嵩<sup>3</sup>, 武珊珊<sup>4</sup>, 田金徽<sup>5</sup>, 董圣杰<sup>6</sup>, 詹思延<sup>1</sup>

【摘要】本文回顾了网络Meta分析(NMA)的步骤,概述了NMA实施过程中潜在的挑战,详细介绍了应对各挑战的最新的的方法学进展,以期帮助我国NMA制定者和使用者更好地理解NMA最新进展及其应用场景。

【关键词】网络Meta分析;方法学研究;系统综述

【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(1): overview** Sun Feng<sup>\*</sup>, Yang Zhirong, Zhang Tiansong, Wu Shanshan, Tian Jinhui, Dong Shengjie, Zhan Siyan. <sup>\*</sup>Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China.

Corresponding authors: Zhan Siyan, E-mail: siyan-zhan@bjmu.edu.cn

【Abstract】This paper reviews the steps of network meta-analysis (NMA), outlines the potential challenges in the implementation of NMA, and introduces in detail the latest methodological advances in response to these challenges, so as to help NMA makers and users better understand the latest progress of NMA and its application scenarios.

【Key words】Network Meta-analysis; Methodological advances; Systematic review

网络Meta分析(NMA)是一种评估多种干预措施效果的方法,它可以充分利用所有直接和间接的证据,与单一的直接或间接估计相比,它可以更精确地估计干预之间的相对效果<sup>[1,2]</sup>。此外,即使有些干预措施从未在随机对照试验(RCT)中进行比较,NMA也可以利用间接证据对这些干预措施之间的相对效果进行估计。NMA可对所有感兴趣的干预措施同时进行比较,并对它们的效果进行相对排名。相比标准Meta分析,NMA通常会提出更广泛的研究问题,从筛选到分析的每个步骤中涉及的原始研究通常也更多,在统计上更复杂,因此需要掌握更多专业知识才能合理使用这种方法,需要更多的时间和资源才能完成NMA。间接比较和NMA要求所纳入的证据具有一致性。在制定研究问题、确定文献纳入标准和评价证据时,需要考虑研究人群和干预措施等方面的可传递性。在解释NMA结果时,尤其是在存在不一致或其他潜在偏倚的情况下,应格外慎重。

近10年来,NMA有一系列的研究进展,其方法学在不断完善。本篇是网络Meta分析研究进展

基金项目:北京大学教学新思路(2019YX007,71009Y0026);北京大学医学部教育教学研究课题(2019YB39)

作者单位:<sup>1</sup> 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与卫生统计学系;<sup>2</sup> CB1 8RN 剑桥,英国剑桥大学临床医学院初级医疗中心;<sup>3</sup> 200040 上海,复旦大学附属静安区中心医院中医科;<sup>4</sup> 100050 北京,首都医科大学附属北京友谊医院国家消化系统疾病临床医学研究中心;<sup>5</sup> 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;<sup>6</sup> 264000 烟台,烟台市烟台山医院骨关节科

通讯作者:詹思延,Email:siyan-zhan@bjmu.edu.cn

doi: 10.3969/j.issn.1674-4055.2020.06.02

系列的第一篇,将在简单回顾NMA步骤和挑战的基础上,重点概述NMA的最新进展,后续的系列文章将对最新进展的每个话题逐一进行理论解读、案例演示和总结展望。

## 1 NMA的步骤

自2002年提出NMA术语发展至今,其实策略和传统系统综述与标准Meta分析基本相同,主要包括以下步骤:

(1) 确定研究问题和纳入标准:清楚定义PICO,所纳入的干预措施应有3种或以上。PICO范围的宽窄依据研究目的而定。

(2) 文献来源和检索:需要广泛多途径搜索所有符合PICO纳入标准的相关研究,并提供检索策略,以便其他研究者能重复执行。

(3) 研究选择:与传统的系统综述与标准Meta分析一样,基于事先制定的明确的纳入标准,在上述初步检索的文献中筛选,选出合格的文献。

(4) 信息提取:研究者应事先制定信息提取表,从合格文献里提取相关的信息。

(5) 质量评估:采用恰当的质量评价工具进行质量评估,主要有两方面:①对所有纳入的合格研究进行偏倚风险评价;②对NMA结果进行证据质量评价,多采用GRADE工具。

(6) 证据整合:分析前需要将基于PICO形成的整个证据体绘制为一个网络关系图。在可能的情况下进行统计分析,异质性和相似性是间接比较的潜在问题,NMA还要重点考虑证据的不一

致性，间接比较和NMA均需要使用不同于标准Meta分析的统计方法，干预措施之间的疗效关系目前多推荐使用贝叶斯框架或Meta回归分析。

## 2 NMA的挑战

(1) 科学实施NMA，需要专业的统计知识，解读复杂模型的方法原理，软件也需要同步的发展。

(2) 假设“网络”中包含的所有干预措施都同样适用于当前证据体所纳入的每一个RCT研究的患者和应用情景，这一假设意味着所有RCT满足相同的目标人群，分析前需要判断该假设是否成立，但实际上很难客观判断。

(3) 与标准Meta分析一样，NMA可能会引入研究选择偏倚，需要检测选择性报告偏倚和发表偏倚。

(4) 作为标准Meta分析的拓展，NMA不仅需要满足同质性假设，还需要满足相似性和一致性假设，即NMA分析中包含的不同研究集之间所有可能影响相关效应的重要因素的分布都是相似的。当不同的信息源（如直接和间接）对某个特定的干预比较不一致时，就会出现不一致性。

(5) NMA仅限于满足基于特定PICO问题的所纳入研究的合并。但在某些情况下，该方法与标准Meta分析相比未提供任何附加价值，尤其当证据体研究数目稀少且干预措施之间联系不紧密时，此时需借助先验知识或补充其他证据丰富证据体以体现NMA的优势。

(6) 经典的NMA关注的结局往往限于一个，而证据体中干预措施的比较如果能同时兼顾多个相关结局从不同层面综合评价（如有效性、安全性、经济学等），能提供更全面的决策信息，因此多个结局的NMA模型近几年得到了较大发展和应用。

(7) 有时NMA也要考虑干预措施/结局自身的特点，比如多个剂量、几个药物大类中分别嵌套多个具体药物、结局指标的多个时点的重复测量等，也需要发展一些新的NMA模型。

(8) NMA所涉及的文献和信息较多，有时候纳入的文献量可达数百篇，需要花费更多的时间和精力对这些文献进行整理，如何避免研究团队之间文献重复整理，搭建共享平台形成有效共享合作机制是很好的倡议。

(9) 软件的可操作性在一定程度上影响着该类型Meta分析的实施，操作便捷的软件和适当的专业培训较为必要。

## 3 NMA进展

近10年网络Meta分析的方法不断更新，从最

开始对随机对照试验的间接比较，发展到今NMA已经可以应用在个体参与者数据、剂量反应数据、“真实世界”观察性研究的证据等方面，可允许对多种形式的汇总数据进行推断，同时也发展了新方法可放宽对异质性、不一致性等假设的严格限制<sup>[3,4]</sup>。主要进展如下：

**3.1 NMA统计模型及模型拟合软件选择** 虽然NMA发展至今已经有多种不同的模型，但当样本量较大时，各种方法产生的结果基本相同，很难区分这些模型孰优孰劣。Efthimiou等<sup>[4]</sup>2016年建议研究者可根据自身偏好，考虑方法的直观性或软件模块的功能特点，来选择统计模型。频率学派模型的一个优点是适用于统计领域的大多数应用程序，这意味着多数使用者可能更容易理解该类方法产生的结果。然而如前所述，贝叶斯方法在很多场景下有更多优势。因此。在实践中Cipriani等<sup>[5]</sup>2018年建议选择一种方法作为主要分析，然后执行另一种方法作为敏感性分析，这种分析策略已变得越来越流行，可以比较这两种方法是否能得出相同结论，如果一致则表明结果是可靠的，如有分歧，则需深入解释可能的原因。

**3.2 NMA发表偏倚的调整** 原始研究质量和发表偏倚会对Meta分析结果真实性造成很大影响。虽然目前已有许多用于评估RCT质量的工具，但并没有定量的方法来评价质量较差的RCT对Meta分析结果的影响，目前的做法主要是咨询专家意见<sup>[6,7]</sup>或参考已发表文献的经验估计<sup>[8]</sup>等探索性的方法。特别是对于NMA来说，这是一个需要进一步研究的问题。针对发表偏倚/小研究效应的直接比较的Meta分析，现在已经有几种调整方法，这些方法也正逐渐被纳入NMA中<sup>[8,9]</sup>。

**3.3 证据结构分层的NMA** NMA是证据综合中常用的方法，当证据体中有多种干预措施或干预措施被细分为多种不同的类别时，会导致不同干预措施之间的RCT数目相对较少，NMA估计的疗效会产生相当大的不确定性，此时很难确定最佳干预措施。

Owen等<sup>[10]</sup>考虑类效应，在现有随机效应NMA工作的基础上，利用贝叶斯马尔科夫链蒙特卡罗方法建立了一个含有三个水平的分层NMA模型，以有限数量的RCT评估广泛的证据体，该模型考虑了同类干预措施之间的互换性以及研究之间残余的异质性。该研究以膀胱过度活跃引起尿失禁发作的例子，演示了这些方法在连续和二分类结局指标中的应用。作者认为贝叶斯三水平的分层NMA不仅保持了个体干预的可解释性，同时也提高了效果估计的精确性。

**3.4 观察性研究的NMA** 真实世界研究旨在评估



干预措施在现实临床实践中是否有效,从而为临床决策提供科学证据,这类研究通常是非随机研究。Efthimiou等<sup>[11]</sup>提出了整合非随机研究证据的新方法,以评估真实世界的治疗效果。作者首先评估了RCT与非随机研究两种证据之间的兼容性,比较了一系列将非随机研究纳入NMA的替代方法(包括原始数据综合、设计调整综合、使用非随机证据作为先验信息以及使用三水平层次模型),并进一步在两个已发表的案例中应用上述方法,分别评估了经皮介入治疗支架内再狭窄和抗精神分裂症药物的治疗效果。经过深入讨论每种方法的优点和局限性得出结论:即纳入来自非随机研究的真实世界证据可证实基于RCT的NMA的结果,提高NMA的准确性。

**3.5 基于目标人群执行校准的NMA** IPD的NMA是Meta分析领域的金标准,除了对证据体中的干预措施数目、有无闭合环以及协变量的类型均不设限制外,还可基于目标人群基线协变量的联合分布进一步校准IPD的NMA结果,从而进行结果的外推,这种理想情况可以避免标准Meta聚合偏倚(生态学偏倚)的问题,Zhang等<sup>[12]</sup>2016年也证实并推荐这种方法可校准目标人群。但获得证据体每一个RCT的IPD是不现实的,因而IPD的NMA通常很难实现。

Phillippo等<sup>[13-16]</sup>2019年开发了一种多水平的网络Meta回归(multi-level network meta-regression, ML-NMR)来执行校准的NMA,该研究显示即便只能获得证据体中部分IPD数据或集合水平的RCT基线特征各水平的组合的结果(如考虑两个因素:性别和年龄,则需要获得每个RCT每对干预在各个性别年龄组的比较结果,即亚组的比较结果),在假定效应修饰因子在研究间是均衡的前提下,ML-NMR也可以将间接比较/NMA的结果进行校准,外推到目标人群中。该ML-NMR模型的优点包括:①治疗的比较仍保留随机分组,避免了指示征混杂;②不需要试验IPD数据;③在IPD的RCT水平上校准效果估计;④可产生药企研发人员、临床医生、临床指南制定者、HTA评估人员等共同关注的结果。这一校准方法为更多疾病治疗措施的评价提供了可行性,有助于把与决策目标群体相关的效果估计应用于决策的制定,今后必将被进一步广泛使用。

**3.6 多结局的多元NMA** NMA综合了多种医疗干预措施大量临床试验的数据,当临床试验缺少一个或多个感兴趣的结局数据,综合这些数据可能导致证据网络断开、不确定性增加和有偏差的估计,从而对决策产生严重影响。为克服这一问题,有学者提出了多结局的多元NMA<sup>[3]</sup>。多元

NMA考虑了多个结局指标之间的相关性,这与单独对每个结局指标进行单变量NMA相比,主要有两个优势:①结局指标之间相关性的引入使得每个结局指标的汇总结果能够利用所有结局指标的数据;②允许纳入那些没有报告所有相关结局指标的研究,从而纳入更多的研究和证据,进而可得出更精确的结论(更窄的置信区间)。此外,多元NMA方法可进一步扩展,当几个药物大类(高水平层次)中分别嵌套多个具体药物(低水平层次)时,采用层次模型借用高水平层次上的同类干预措施之间的信息,在多元NMA模型中加入类效应(高水平层次的药物效应)通常可以提高同类治疗效果估计的精度,有助于临床决策。

**3.7 NMA的证据质量分级** 随着NMA的制定流程以及方法学的日趋成熟,ISPOR组织先后制定了3部专门论述如何制定、解读和评价NMA的手册。但在GRADE之前,很多NMA仅通过评价纳入研究的偏倚风险来推测总体证据质量,这种方法由于未考虑到其他偏倚因素,因而存在很大局限性。如果未对NMA进行证据质量分级,则可能对读者理解NMA结论的真实性和可靠性造成障碍,甚至产生误导。

因此,2014年GRADE工作组<sup>[17]</sup>提出了对NMA进行证据质量分级的四步法,进一步完善了GRADE在NMA中运用的理论基础。但与其他类型系统综述与标准Meta分析相比,NMA中同时纳入直接比较和间接比较证据,因此分级过程会更复杂,除了要考虑GRADE传统的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(intransitivity),以及直接比较和间接比较结果的不一致性(inconsistency)。

2017年瑞士伯尔尼大学社会与预防医学机构最新发布的在线工具CINeMA(Confidence in Network Meta-Analysis),专门用于评价NMA证据质量(Available from cinema.ispm.ch),评价的主要内容仍然基于GRADE五个降级因素,分别为:偏倚风险、间接性、不一致性、不精确性和发表偏倚。评价框架思路主要基于Salanti G等2014年发表于PloS One的文献,该软件可分析并给出直接比较对NMA的贡献,在此基础上评价NMA的可信度(credibility)。该软件的开发,大大提升了NMA的GRADE的评价质量和效率。

Yepes-Nuñez等<sup>[18]</sup>2019年应用了一个基于四轮半结构化访谈的用户体验模型对案例数据进行分析,详细展示并优化了GRADE的结果摘要(SoF)表,以表格形式展示证据分级结果,显示NMA的关键信息。研究组采访了32位进行或

使用Meta分析的利益相关者。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表格应包含以下6方面内容：①临床问题的详细信息（PICO）；②网络关系图的构建；③相对和绝对效应的估计；④证据的确定性；⑤处理的排序；⑥结果的解释。这种NMA-SoF表，有助于理解NMA的主要结果并制定卫生决策。

**3.8 个体患者资料的NMA** 通常NMA的数据主要关注集合水平数据的证据整合。如果可以获得每个原始研究的个体患者数据（Individual patient data, IPD），基于个体水平的协变量进行分析，则可大大提高研究结果的精确度，有助于探讨研究之间的异质性/不一致性或某个亚组病人的成本-效益等从入手。例如，如果要评估年龄对干预效果的影响，传统NMA只能使用平均年龄作为协变量在回归模型中检测交互作用，而IPD的NMA则可直接探讨每个患者的个体年龄等。IPD的NMA模型允许包含患者水平的协变量，可用于二分类结局指标<sup>[19,20]</sup>和时间事件数据<sup>[21,22]</sup>。由于IPD可能无法从所有相关研究中获得，因此有学者对IPD的NMA模型进行了拓展，可同时整合IPD和集合水平数据的证据<sup>[20,22,23]</sup>。

**3.9 诊断试验准确性的NMA** 早期的诊断准确性试验的Meta分析（即Moses-Littenberg合并ROC曲线）允许纳入具有不同诊断阈值的诊断试验进行准确性的Meta分析，虽然考虑到不同诊断阈值是异质性的一个重要来源，但未把这种异质性来源明确地考虑到模型中。随后的双变量模型和分层ROC曲线克服了这一难题，但并未提供确定最佳的诊断阈值的方法。直到2018年，Owen等<sup>[24]</sup>学者开发的诊断性试验NMA才成功解决了这一问题，对同一情景下多种诊断试验同时进行比较，给出到底哪种诊断试验具有最佳的诊断准确性和最佳的诊断阈值。指南制订人员和研究人员如果研究目的是确定哪种试验的诊断阈值最适合相应的临床问题，建议考虑使用诊断性试验NMA，该方法能够帮助确定某种特定疾病下的最佳诊断试验和最佳诊断阈值<sup>[25]</sup>。

**3.10 结局重复测量（多时点）的NMA** NMA纳入的部分研究中某些结局指标可能存在多个时间点的报告数值，且报告的时间点不一致。虽然可以在每个时间点进行单独的NMA分析，但这会降低统计效能。如果充分利用多个时间点的报告数值，把不同时间点之间的相关性纳入到模型，则可得更为精确的结果，从而为临床决策提供支持。Madan等<sup>[26]</sup>2014年已经证实该方法可在二分类结局指标中实现，且该方法还可将特定时间点断开的多个网络连接在一起。此外，针对连续性结局

指标，Pedder等<sup>[27]</sup>2019年开发了考虑时间频次的MBNMAtime包，允许纳入的各研究中包含多个时间点的报告数值。

**3.11 生存数据的NMA** Vickers等<sup>[28]</sup>2016年比较了两种对生存资料的结局进行NMA分析的方法，并对平均生存时间进行估计：①基于研究水平汇总的危险比数据（即基于HR值及其置信区间）拟合NMA；②基于重建的IPD水平的数据拟合NMA。

基于危险比方法估计的NMA结果与共同参照的选择有很大关系。此外，尽管该方法对当前研究纳入的干预措施给出置信区间的预测，但并不一定意味着该方法对其他的干预措施也能做出合理的预测；而且证据体中很有可能有一个或多个干预之间的比较不满足比例风险假定的危险比。Vickers等<sup>[28]</sup>提出使用分数多项式方法生成与生存数据吻合的生存曲线，重建患者个体水平的数据，并进一步调整，以对生存结局给出合理的推断。然而，由于可能会受到发表偏倚影响，该方法只纳入报告了Kaplan-Meier图表的研究，部分研究由于只提供了汇总数据则无法纳入。因此理想情况下，两种方法需同时使用，才能提供一个完整可靠的干预措施之间的相对疗效。

Petit等<sup>[22]</sup>2019年提出应用限制性平均生存时间差异（rmstD）作为一个绝对效应测量指标，在来自个体患者数据（IPD）探讨化疗对鼻咽癌的作用案例中，分别采用NMA比较了两种效应测量（基于HR和基于rmstD）所获得的相对结果，发现基于HR或rmstD的NMA结果在主要结局上的估计有所不同，但在次要结局上没有差异，但考虑到HR对非比例风险的敏感性，作者认为rmstD在Meta分析方法中更值得进一步研究和应用。

**3.12 动态更新的NMA** 随着新的原始研究的发表，现有的证据也会随之更新。在每次更新中，都需要基于最新证据网络下结论并对未来研究的必要性提出建议。在前瞻性Meta分析的背景下，未来的研究计划会作为证据积累的一部分，在解释Meta分析结果时需要考虑多重检验的问题。Cr é quit等2016年扩展了Meta分析序贯检验的概念，为更新NMA提供了一个方法学框架（<http://livenetworkmetaanalysis.com/>）<sup>[29]</sup>，建议在监管环境中使用前瞻性的累积NMA，即使没有新的试验可以直接比较干预的相对效果，也有可能成为结论所基于的证据，因为新的研究提供了间接比较的结果。Nikolakopoulou等<sup>[30]</sup>2018年进一步结合案例证明基于每个证据体成对估计的z分数（效应大小与标准误差之比）和每个研究进入NMA后各自获得的信息，作者构建了证明干预有效和无效



的终止边界；当穿过设定的终止边界时，NMA的分析结果可成为结论性的证据，表明即使没有进一步的直接比较，关于感兴趣的特定比较也可以通过目前的间接证据支持决策。而且，目前技术的发展已经实现了半自动化的制定过程，对前瞻性的累积NMA动态更新提供了技术支持，在这种包含了所有可用的干预措施的“实时”的NMA被称为“实时累积NMA”，可以进一步促进未来研究的优先级和决策的制定，为何时规划新的原始研究，何时更新NMA，何时提供结论性的答案等提供了科学依据，以弥合研究证据和卫生保健实践之间的差距。

但该方法在实践中使用较少，仍需要在一个全面的框架中对方法学进行梳理和总结，同时需要在现实世界的例子中对方法学进行评估，开发便于研究者使用的软件程序和参考材料，从而有助于该方法的推广。

**3.13 证据网络断续的NMA** 尽管NMA对包含的治疗、试验和患者的数量没有限制，但有时根据事先制定的PICO，证据体中各干预措施之间无法形成一个相连的网络，此时无法进行标准的NMA（其要求网络中每对干预措施之间都有路径）。对于这种断开的决策证据体（多个小的网络），目前也有一些方法学上的尝试。英国NICE机构2008年首次推荐允许证据体中纳入与决策无关的干预措施，以协助证据体中各个干预措施之间形成连接的网络；2015年该机构<sup>[31]</sup>又推出可以纳入观察性研究证据来扩充整个证据体；对于二分类结局尤其安全性事件，还可以假定缺少直接比较的某一对/几对干预措施的结局发生数都是0，从而使证据体关联起来。

**3.14 NMA结果的可视化** 尽管NMA的拟合模型多种多样，但对这些模型的结果作直观解释通常是比较困难的，尤其是在许多不同干预措施的大型网络中。Law等<sup>[32]</sup>2019年提出了两种可视化方法，使得NMA模型的结果可以更直观，该方法基于证据体的网络结构分析，可不考虑统计模型或估计方法。作者定义了三种测量网络证据体中干预措施之间距离的指标：①干预措施效果的估计；②标准误差；③对应的P值。通过确定合适的阈值，将一些比较对归类为相近（短距离）。在网络分析理论中，距离相近的干预措施被认为是相互连通的。最后，使用标准的网络方法将干预措施分配到不同的组。由此可以确定网络的哪些干预措施估计具有相似（或不同）的治疗效果。作者还提出了第二种方法，即参数自助抽样（Bootstrapping），在可视化中制作和使用热图。作者推荐使用R软件来制图，并提供了代

码。这种新的可视化方法可让NMA制定者更直观地理解其拟合模型的含义，识别拟合模型的最显著特征。研究者亦可在发表的论文中采用可视化的方法来展示这些图形。

**3.15 NMA干预组随机缺失的处理方法** 由于大多数RCT只比较两种干预措施，因此，NMA中一个典型的数据集可能比较稀疏，就像一个不完整的块结构。Zhang等<sup>[33]</sup>提出了一种基于臂的方法，该方法考虑了同一试验中不同干预措施之间的相关性，并假设缺失的干预组是随机缺失的。然而在RCT中，干预组不可忽视的缺失或非随机的缺失可能是由于研究者在设计阶段有意选择而发生的。此外，进行NMA的研究者可能会选择性地纳入感兴趣的干预措施，这也可能导致非随机的缺失。在这篇论文中，作者基于敏感性分析发展了“选择”模型整合潜在的非随机缺失的证据体，扩展了之前的工作。然后将该方法应用于两个NMA案例，并在完全随机缺失和随机缺失两种假设情况下进行大量的模拟和评价，对常用的基于对比的方法和基于臂的方法进行了全面的比较。此外，目前还缺乏一种可解释和简单的衡量方法来评估非随机缺失带来的偏倚风险，未来的研究应集中于这些问题的阐述及其模型评价、解读和适用性。

**3.16 预测模型的NMA** 临床预测模型或预后评分近年来在临床实践中的应用越来越广泛，通过对个体某些特征指标如实验室检验或症状、体征等的利用来估计现在或未来发生某特定事件的风险，为医疗决策提供一定的科学依据。目前，在很多疾病领域针对同一个研究结局都已发表了为数不少的预测模型或评分，但大多数预测模型长期处于“多数被建立，少数被验证，极少被应用”的情况。这一方面是因为大多数模型缺乏足够的外部验证，导致外推性不高；另一方面是因为模型太多且缺乏对不同模型之间预测准确度的比较研究，导致临床选择困难。为了解决这一问题，Haile等<sup>[33]</sup>学者提出了针对预测模型的NMA，根据每一个原始队列研究中验证比较的模型分为不同的组别，各组之间选择同一个模型作为共同参照，采用Lu等提出的两阶段Meta回归方法，先计算每个组内直接比较Meta分析的结果，再将第一阶段计算出的各组直接比较结果及方差采用Meta回归的方法进行合并，从而得到针对同一结局的所有模型之间预测准确度的比较结果。在预测模型的NMA中，效应指标通常为区分度指标（如受试者特征曲线下面积AUROC），也可为校准度指标或其他评价预测模型准确度的指标如净重新分类指数等指标。

**3.17 基于外部证据确定异质性信息先验的NMA** 当研究纳入的文献数较少时, 研究之间的方差估计可能是不精确的, 此时常推荐采用贝叶斯Meta分析整合有异质性的外部证据(先验信息), 得到更稳定的估计。Rhodes等<sup>[34,35]</sup>提出了一种使用数据扩充(data augmentation)来进行贝叶斯Meta分析的方法, 这种数据扩充可看作是不需要指定先验分布的半贝叶斯分析(semi-Bayesian analysis)。作者用虚拟的伪数据(用其似然来替代异质性参数 $\tau^2$ 的先验分布)表示研究间方差的共轭先验信息, 并使用Meta回归进行估计, 与基于重要性抽样技术(importance sampling techniques)和马尔可夫链蒙特卡罗(MCMC)的完全贝叶斯方法进行了比较, 得到的结果几乎相同。作者将该方法应用于一个实际数据集, 并描述了如何拓展到网络Meta分析。总之, 该方法优化了在文献量较少时的贝叶斯Meta分析, 同时也便于研究人员理解和使用。

**3.18 剂量反应关系的NMA** 在现有的NMA分析中, 剂量反应关系经常作为异质性的一种常见且已知的来源, 但其对结果效应估计的影响, 目前并未引起研究者足够的重视, 大多数研究者在分析过程中, 倾向于直接合并到药物水平组或者将具体剂量作为独立的干预组的处理方式, 缺少对剂量反应关系的深入考虑和调整。实际上, 不同药物由于剂量反应关系的不同、研究时各药物剂量组数不同、剂量组数过多等种种因素, 均有可能影响结局指标效应估计的大小和方向。

Giovane等<sup>[36]</sup>2013年提出, 实施NMA的研究者们有必要重视剂量反应关系的前提假设, 并评价其潜在的影响并采取适当的处理方式。Owen等<sup>[37]</sup>2015年推荐在贝叶斯框架下通过马尔可夫链蒙特卡罗模拟(Markov chain Monte Carlo method, MCMC)的非参数方法建立多水平(层次)的NMA模型来解决剂量反应关系的问题。

Pedder等<sup>[27]</sup>2019年开发了R软件MBNMAdose包可以进行基于模型的网络Meta分析(Model-Based Network Meta-Analysis, MBNMA), 通过考虑证据体中不同药物的多个剂量, 进行剂量-反应关系研究, 该模型中将治疗称为特定的剂量或特定的药剂。该方法还可以通过剂量-反应关系和安慰剂将断开的网络联结起来, 提高估计效应的精度, 并允许根据剂量-反应关系对结局指标的估计进行插补/外推。MBNMA模型2016年由Mawdsley等<sup>[38]</sup>提出, 模型建立和分析均在贝叶斯框架中实现, 遵循Lu和Ades<sup>[39]</sup>2004年提出的标准NMA方法的扩展, 并在JAGS(4.3.0或更高版本)中运行。MBNMAdose提供了一套完整的功

能, 可以执行剂量反应MBNMA、模型检查、预测和绘制许多信息图形。

总之, 不恰当地使用NMA很可能导致错误的结果, 当证据网络结构连接程度较低时统计效能也较低, 而且间接证据始终是观察性的, 应围绕NMA的三个假定谨慎解读。因此, NMA最好由该领域的临床专家、流行病学专家和有经验的统计学家共同实施和评价。此外, 在基于RCT的经典NMA基础上, NMA理论和应用都在快速拓展, 例如预后、预测因素的识别; 通过对NMA和专家意见进行整合改进真实世界中的预后/诊断模型; 单臂研究的NMA在新药、器械疗效评价中的应用等。Cochrane、WHO、NICE等机构也越来越多地推荐使用NMA对多种干预措施进行综合评估, 并进一步通过严格的证据选择、整合和评价方法开发各种指南。

#### 参考文献

- [1] Cooper NJ, Kendrick D, Achana F, et al. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms[J]. *Epidemiol Rev*, 2012, 34:32-45.
- [2] Caldwell DM. An overview of conducting systematic reviews with network meta-analysis[J]. *Syst Rev*, 2014, 3:109.
- [3] Riley RD, Jackson D, Salanti G, et al. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples[J]. *BMJ*, 2017, 358:j3932.
- [4] Efthimiou O, Debray TPA, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology[J]. *Research Synthesis Methods*, 2016, 7(3):236-63.
- [5] Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis[J]. *The Lancet*, 2018, 391(10128):1357-66.
- [6] Turner RM, Spiegelhalter DJ, Smith GC, et al. Bias modelling in evidence synthesis[J]. *J R Stat Soc Ser A Stat Soc*, 2009, 172(1):21-47.
- [7] Welton NJ, Ades AE, Carlin JB, et al. Models for potentially biased evidence in meta-analysis using empirically based priors[J]. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2009, 172(1):119-36.
- [8] Trinquart L, Chatellier G, Ravaud P. Adjustment for reporting bias in network meta-analysis of antidepressant trials[J]. *BMC Med Res Methodol*, 2012, 12:150.
- [9] Mavridis D, Sutton A, Cipriani A, et al. A fully Bayesian application of the Copas selection model for publication bias extended to network meta-analysis[J]. *Statistics in medicine*, 2013, 32(1):51-66.
- [10] Owen RK, Tincello DG, Keith RA. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints[J]. *Value in Health*, 2015, 18(1):116-26.
- [11] Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis[J]. *Statistics in medicine*, 2017, 36(8):1210-26.
- [12] Zhang Z, Nie L, Soon G, et al. New methods for treatment effect calibration, with applications to non-inferiority trials[J]. *Biometrics*, 2016, 72(1):20-9.
- [13] Phillippo D, Ades T, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE[D]. 2016.

- [14] Phillippo DM, Dias S, Ades AE, *et al.* Synthesis of individual and aggregate level data using multilevel network meta-regression: extension to general likelihoods. In: Abstract from Royal Statistics Society International Conference 2019, Belfast, United Kingdom[D]. 2019.
- [15] Phillippo D, Dias S, Ades T, *et al.* Synthesis of individual and aggregate level data using multilevel network meta-regression: extension to general likelihoods, <http://www.iscb2019.info>. In: 40th Annual Conference of the International Society for Clinical Biostatistics—KU Leuven, Leuven, Belgium; 2019.
- [16] Phillippo DM. Calibration of treatment effects in network meta-analysis using individual patient data[D]. University of Bristol 2019.
- [17] 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.
- [18] Yepes-Núñez JJ, Li S-A, Guyatt G, *et al.* Development of the summary of findings table for network meta-analysis[J]. *Journal of Clinical Epidemiology*, 2019, 115: 1–13.
- [19] Donegan S, Williamson P, D'Alessandro U, *et al.* Combining individual patient data and aggregate data in mixed treatment comparison meta-analysis: Individual patient data may be beneficial if only for a subset of trials[J]. *Statistics in medicine*, 2013, 32(6): 914–30.
- [20] Jansen JP. Network meta-analysis of individual and aggregate level data[J]. *Res Synth Methods*, 2012, 3(2): 177–90.
- [21] Saramago P, Chuang LH, Soares MO. Network meta-analysis of (individual patient) time to event data alongside (aggregate) count data[J]. *BMC Med Res Methodol*, 2014, 14: 105.
- [22] Petit C, Blanchard P, Pignon JP, *et al.* Individual patient data network meta-analysis using either restricted mean survival time difference or hazard ratios: is there a difference? A case study on locoregionally advanced nasopharyngeal carcinomas[J]. *Systematic Reviews*, 2019, 8(1): 96.
- [23] Saramago P, Sutton AJ, Cooper NJ, *et al.* Mixed treatment comparisons using aggregate and individual participant level data[J]. *Statistics in medicine*, 2012, 31(28): 3516–36.
- [24] Owen RK, Cooper NJ, Quinn TJ, *et al.* Network meta-analysis of diagnostic test accuracy studies identifies and ranks the optimal diagnostic tests and thresholds for health care policy and decision-making[J]. *J Clin Epidemiol*, 2018, 99: 64–74.
- [25] O'Sullivan JW. Network meta-analysis for diagnostic tests[J]. *BMJ Evid Based Med*, 2019, 24(5): 192–3.
- [26] Madan J, Chen Y F, Aveyard P, *et al.* Synthesis of evidence on heterogeneous interventions with multiple outcomes recorded over multiple follow-up times reported inconsistently: a smoking cessation case-study[J]. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2014, 177(1): 295–314.
- [27] Pedder H, Dias S, Bennetts M, *et al.* Modelling time-course relationships with multiple treatments: Model-based network meta-analysis for continuous summary outcomes[J]. *Res Synth Methods*, 2019, 10(2): 267–86.
- [28] Vickers A. Survival Network Meta-analysis: Hazard Ratios Versus Reconstructed Survival Data[J]. *Value in Health*, 2016, 19(3): A90.
- [29] Créquit P, Trinquart L, Ravaud P. Live cumulative network meta-analysis: protocol for second-line treatments in advanced non-small-cell lung cancer with wild-type or unknown status for epidermal growth factor receptor[J]. *BMJ open*, 2016, 6(8): e011841.
- [30] Nikolakopoulou A, Mavridis D, Egger M, *et al.* Continuously updated network meta-analysis and statistical monitoring for timely decision-making[J]. *Statistical Methods in Medical Research*, 2018, 27(5): 1312–30.
- [31] Faria R, Hernandez Alava M, Manca A, *et al.* NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness for technology appraisal: methods for comparative individual patient data[J]. Sheffield, UK 2015.
- [32] Law M, Alam N, Veroniki AA, *et al.* Two new approaches for the visualisation of models for network meta-analysis[J]. *BMC Med Res Methodol*, 2019, 19(1): 61.
- [33] Zhang J, Chu H, Hong H, *et al.* Bayesian hierarchical models for network meta-analysis incorporating nonignorable missingness[J]. *Statistical methods in medical research*, 2017, 26(5): 2227–43.
- [34] Rhodes KM, Turner RM, White IR, *et al.* Implementing informative priors for heterogeneity in meta-analysis using meta-regression and pseudo data[J]. *Statistics in medicine* 2016, 35(29): 5495–511.
- [35] Turner RM, Jackson D, Wei Y, *et al.* Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis[J]. *Statistics in medicine*, 2015, 34(6): 984–98.
- [36] Del Giovane C, Vacchi L, Mavridis D, *et al.* Network meta-analysis models to account for variability in treatment definitions: application to dose effects[J]. *Stat Med*, 2013, 32(1): 25–39.
- [37] Owen RK, Tincello DG, Keith RA. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints[J]. *Value Health*, 2015, 18(1): 116–26.
- [38] Mawdsley D, Bennetts M, Dias S, *et al.* Model-Based Network Meta-Analysis: A Framework for Evidence Synthesis of Clinical Trial Data[J]. *CPT Pharmacometrics Syst Pharmacol*, 2016, 5(8): 393–401.
- [39] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons[J]. *Statistics in medicine*, 2004, 23(20): 3105–24.

本文编辑：孙竹

欢迎投稿，欢迎订阅！

中国科技核心期刊，科技论文统计源期刊

E-mail: [ebcvm\\_cj@126.com](mailto:ebcvm_cj@126.com); 联系电话: 010-84008479网址: [www.ebcvm.cn](http://www.ebcvm.cn)/[www.ebcvmcj.com](http://www.ebcvmcj.com)/[www.ebcvmcj.org](http://www.ebcvmcj.org)