

· 循证研究与临床转化 ·

网络Meta分析研究进展系列（十一）： 观察性研究网络Meta分析

李伦¹，武珊珊²，杨智荣³，董圣杰⁴，张天嵩⁵，孙凤⁶，田金徽⁷

【摘要】观察性研究网络Meta分析越来越受卫生决策人员的欢迎，但观察性网络Meta分析的方法目前尚不成熟。因此，本研究旨在对观察性研究网络Meta分析进行全面的讨论，分析当前观察性研究网络Meta分析的现状、存在的问题和拟解决的问题。基于当前研究现状分析，观察性研究网络Meta分析主要在假设检验评估和合并方法上尚存在不足，仍需进一步探索分析。

【关键词】网络Meta分析；假设检验；数据合并；观察性研究；异质性；一致性

【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 (11): Network meta-analysis of observational studies

Li Lun^{*}, Wu Shanshan, Yang Zhirong, Dong Shengjie, Zhang Tiansong, Sun Feng, Tian Jinhui. ^{*}Department of Breast Surgery, the Second Xiangya Hospital, Central South University.

Corresponding authors: Tian Jinhui, E-mail: tianjh@lzu.edu.cn; Sun Feng, E-mail: sunfeng@bjmu.edu.cn

[Abstract] Network meta-analysis of observational studies was more and more popular in health decision making, but there were several limitations in the methods. Therefore, we aimed to conduct a comprehensive review on network meta-analysis of observational studies, and discuss the research status, potential bias, problems and problems to be solved. Based on the analysis of the current research status, the methods for hypothesis assessment and statistical methods in network meta-analysis of observational studies were insufficient, which needs further exploration and analysis.

[Key words] Network meta-analysis; Hypothesis assessment; Data pooling; Observational study; Heterogeneity; Consistency

1 前言

随着真实世界研究越来越多，更多的观察性研究为当前临床治疗和决策提供重要的证据，可以补充随机对照试验证据的不足。纳入观察性研究的网络Meta分析可更全面地评估当前的临床治疗方案，为临床决策提供更全面的依据。但观察性研究网络Meta分析的方法目前尚不成熟，还存在一些问题。因此，我们对观察性研究网络Meta分析进行全面的讨论，分析当前观察性研究网络Meta分析的现状、存在的问题和拟解决的问题。

2 观察性研究的网络Meta分析

网络Meta分析是一种较为新颖的统计学方

基金项目：甘肃省科技计划项目(20CX4ZA027, 20CX9ZA112); 国家自然科学基金面上项目(72074011)

作者单位：¹ 410011 长沙, 中南大学湘雅二医院乳腺外科; ² 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; ³ CB1 8RN 剑桥, 英国剑桥大学临床医学院初级医疗中心; ⁴ 264000 烟台, 烟台市烟台山医院骨关节科; ⁵ 200040 上海, 复旦大学附属静安区中心医院中医科; ⁶ 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; ⁷ 730000 兰州, 兰州大学循证医学中心 兰州大学基础医学院 甘肃省循证医学与临床转化重点实验室

共同通讯作者：田金徽, E-mail: tianjh@lzu.edu.cn;

孙凤, E-mail: sunfeng@bjmu.edu.cn

doi: 10.3969/j.issn.1674-4055.2021.04.01

法，可以基于多个干预措施的效应量进行合并，并进行排序，进而筛选最佳的治疗措施。在纳入研究的方法学质量较高，且同质性和一致性较高时，网络Meta分析结果相对于单个研究更为精确，样本量更大，结果更为可信。网络Meta分析可相对容易地处理复杂的证据网络，得到不同治疗措施之间的间接比较结果、直接比较结果和混合比较的结果，同时增强统计效能，从而确保临床决策中充分地利用当前所有的证据。

通常而言，干预性网络Meta分析一般只纳入随机对照试验，较少纳入非随机对照试验和观察性研究。主要是考虑观察性研究缺少随机化，不同干预措施组间的基线往往存在差异，导致研究结果可能存在偏倚。但仅纳入随机对照试验的网络Meta分析会存在一定的局限性，因为观察性研究，尤其是大样本队列研究可以提供较多的数据，且证据水平较高。一般来说，设计严格的随机对照试验，内部真实性会比观察性研究的高，但纳入和排除标准严格，其适应范围通常较窄。真实世界数据常基于观察性研究，对人群的选择没有严格的限制，因此其来自于临床实践的真实情况，可以反映真实的临床实践。

评估干预措施的有效性，随机对照试验无疑

是当前最佳的研究设计，其设计严谨，具有严格的纳入和排除标准，结果指标定义清楚，数据前瞻性收集，试验组和对照组随机分配，试验组和对照组在基线上高度相似，是目前评估干预性措施的金标准。观察性研究，一般基于真实世界数据，是对某一人群特定时间长度的随访，比较不同措施之间的实际效果，其人群基于真实临床实践，患者的基础疾病及合并疾病常常较多，更能够反应实际的临床人群，随访时间也更长。无论是随机对照试验还是观察性研究都有一定的劣势。随机对照试验因为其设计严谨，纳入和排除标准严格，因此其纳入人群不能很好地代表全部患者人群，其试验结果的外推性相对受限。对于观察性研究，由于其基于真实世界数据，其混杂因素较多，干预措施、对照措施、患者人群、结果指标、随访时间常常不能够标准化，因此其结果可能存在偏倚。观察性研究数据常具有较大的偏倚，如观察组和对照组基线不一致，在人群、治疗措施上具有较高的异质性。

那么网络Meta分析是否可以纳入观察性研究呢？从统计学方法学角度和理论上讲，纳入观察性研究，合并观察性研究结果是可以的。显然，只纳入随机对照试验的网络Meta分析可以提供最佳的证据，但是这种方法忽略了其他研究设计的数据，这些研究设计的数据也可以作为补充，支持当前的临床决策。因此，网络Meta分析需要观察性研究的数据作为支持和补充。以"network meta analysis" AND (cohort or "case control" or observational)为检索策略在PubMed中进行简单检索，可以发现纳入观察性研究的网络Meta分析最早发表于2011年，近年来逐渐增多（图1）。这体现了近年来卫生决策体系中对纳入观察性研究网络Meta分析的重视。

3 假设评估

需要注意的是，网络Meta分析是一种统计学方法，可以合并多个研究间直接比较和间接比较结果，进而提高统计学精确性。从统计学角度讲，

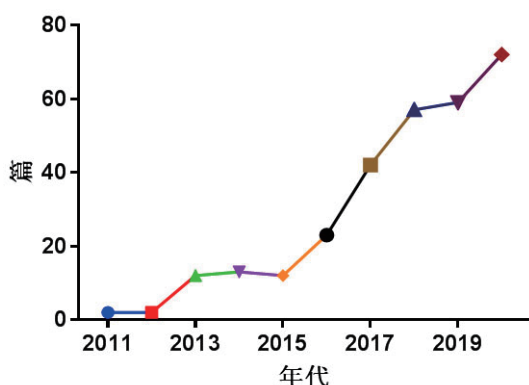


图1 观察性研究网络Meta分析发表数量（篇）的变化

合并观察性研究结果和随机对照试验结果是可行的。但是从假设角度讲，可能会存在问题。一般来讲，网络Meta分析要进行数据合并需要满足三个假设：同质性、相似性和一致性^[1-3]。同质性一般是指所有研究的研究在临床（患者、干预措施、对照措施、结果等）、方法学（研究设计、随访时间等）和统计学上不存在显著的异质性；相似性是指所有研究间以及不同对照组间影响效应量的因素相似。一致性是指直接和间接比较结果一致。纳入观察性研究对于网络Meta分析的最大挑战就是观察性研究是否可以和随机对照试验一样在同质性、相似性和一致性上满足网络Meta分析的假设检验。观察性研究因为研究地点和研究环境的不同，会导致各个研究在患者选择、治疗措施、对照措施、结果指标定义、随访时间等方面存在异质性，进而影响不同治疗措施间研究和各个研究间的相似性。对于一致性检验也存在同样的问题。观察性研究和随机对照试验所产生的直接证据和间接证据是否一致。还有观察性研究和随机对照试验的直接证据是否一致？间接证据是否一致？这都是需要考虑的问题。

如何评估观察性研究网络Meta分析的假设这一直是学术界讨论的问题。网络Meta分析通常仅纳入一种研究设计，因此其假设检验评估通常未考虑不同研究设计对同质性、一致性和相似性的影响。但是对于纳入两种及两种以上研究设计的网络Meta分析，其假设检验评估就需要考虑研究设计对同质性、相似性和一致性的影响。对于同质性评估，通常针对于不同对照措施间的所有研究要在临床、方法和统计上足够相似，没有异质性。对于纳入多种研究设计的网络Meta分析，研究设计本身就是异质性的来源。因此纳入观察性研究的网络Meta分析首先就要评估研究设计对同质性的影响，需要保证在同一对照措施中不同研究设计间研究对象、干预措施、对照措施、结果、随访时间等所有重要的临床和方法学特征同质，且结果不存在显著的异质性。相似性一般针对所有的研究，要求影响结果的重要变量在不同比较组的研究间分布相似。这对纳入的观察性研究的质量提出了很高的要求。一般而言，就要求除了研究设计以外的其他因素必须高度相似。一致性是指直接证据和间接证据之间的一致。在纳入观察性研究和随机对照试验的网络Meta分析中，直接证据和间接证据可以来自于观察性研究，也可以来自于随机对照试验。那么这就需要保证来自于观察性研究和随机对照试验中的直接证据和间接证据无统计学差异。如果两种研究设计的结果一致，但在研究对象、干预措施、对照

措施和结果指标的定义上存在异质性,则两种研究设计也不应该合并^[4]。因此,有学者提出在纳入观察性研究时,研究者应该组成团队评估是否能将观察性研究的数据纳入网络Meta分析中^[5,6]。

4 合并方法

网络Meta分析合并随机对照试验和观察性研究的方法主要有三种:直接合并、作为先验分布信息、等级模型^[7]。此外,Jenkins还采用修正幂验前分布进行网络Meta分析^[8,9]。

4.1 直接合并 直接合并主要有两种形式,第一种是对观察性研究和随机对照试验的结果进行直接合并^[10]。简单合并就是指认为所有研究是一样的,只需要合并所有研究的数据,而不需要考虑研究设计,这是合并不同研究设计的最简单的方法。这种方法不能够对观察性研究的结果进行调整,如降低权重。这种方法假设试验设计之间没有差异,这种方法不需要偏倚调整,也不需要考虑其他不确定因素。一般而言,不推荐此种办法进行观察性研究的网络Meta分析。但是通过这种方法可以检测不一致,如干预措施A和干预措施B之间没有直接随机对照试验,但有观察性研究时,干预措施A和干预措施B之间的间接比较结果就可以和观察性研究的结果进行比较,进而评估间接证据和直接证据之间的不一致。这种情况,网络Meta分析可以直接合并随机对照试验和观察性研究的结果,忽略了研究设计间的异质性,默认观察性研究的数据和随机对照试验的数据一样可信。

第二种方法就是在合并随机对照试验和观察性研究时,对观察性研究的结果进行调整^[11]。对于双臂观察性研究,结果效应量根据偏倚进行调整,因此引入偏倚因子 β_j ,最终的效应量为原始效应量 $\theta_{j,XY}$ 加上偏倚因子 β_j 。方差也进行调整,主要通过引入方差调整因子 w_j ,调整后的方差为 $\frac{S_{j,XY}^2}{w_j}$ ($S_{j,XY}^2$ 是原始效应量的方差)。 β_j 和 w_j 在模型中可以为固定值,也可以为随机值。贝叶斯模型中会赋予 β_j 和 w_j 先验分布,一旦赋予先验值,就无需对 w_j 再次进行降低权重处理。因为赋予先验值后,就意味着在观察性研究间 β_j 和 w_j 为随机变量,且在研究间具有可交换性。如果观察性研究和随机对照试验结果存在不一致性,且 w_j 为随机变量,后验结果就会趋向较小的效应量,进而降低观察性研究对最终效应量的影响。如果将 w_j 定义为0,意味着网络Meta分析仅纳入随机对照试验;如果将 w_j 定义为1,意味着网络Meta分析同时纳入随机对照试验和观察性研究,且该模型直接将观察性研究结果和随机对照试验结果进行合并。一般不建议将 w_j 定义为大于1的值。在双臂

观察性研究中偏倚因子和方差调整因子一般只有一个,但是对于多臂观察性研究,偏倚因子和方差调整因子就为多个,一般为治疗措施数据减去1。对于4臂观察性研究,那么偏倚因子和方差调整因子就应该有3个。这种情况下,方差需要调整,协方差也需要调整。

4.2 观察性研究结果作为先验信息 在网络Meta分析中,可将观察性研究的数据进行合并,将合并结果作为先验分布^[10,11]。有以下三种处理方式。

第一种方法是首先对观察性研究进行数据合并,合并时要对结果进行调整,主要引入偏倚因子 β_j 和方差调整因子 w_j ^[11]。在贝叶斯模型中,需要对偏倚因子 β_j 和方差调整因子 w_j 赋予先验值,但是不需要同时对两个进行赋值。一旦对其随机赋值,其后验数值就会与证据间的一致性相关。偏倚因子 β_j 和方差调整因子 w_j 就会以研究中最小的值为准,降低观察性研究对网络Meta分析结果的影响。如果偏倚因子 $\beta_j=0$ 和方差调整因子 $w_j=1$,就意味着观察性研究的数据非常可靠,在这种情况下,如果采取固定效应模型,就意味着观察性研究和随机对照试验的数据直接合并。对于多臂观察性研究,基本参数之间存在相关性,这种情况下需要使用多维先验分布^[10]。

第二种方法是使用观察性研究的精确似然值^[11]。通过引入似然值的效能函数 α ,达到对观察性研究结果降低权重的目的。一般而言 α 值在0和1之间。第一种方法中 w_j 可以被认为是修正幂验前分布的特殊形式, α 则是针对每一个研究的似然函数结果。在正态分布的数据中, α 和 w_j 值是相等的。

第三种方法是采用混合先验分布^[11]。主要有两部分构成:第一部分是观察性研究结果的先验数值,第二部分是无赋值先验。该方法引入因子 p 控制赋值先验和无赋值先验的比值。

4.3 等级模型 等级模型需要考虑不同研究设计(随机对照试验和观察性研究)间和不同研究间的异质性,常分为三步。第一步先对单个研究进行分析,计算单个研究的效应量。可分为三种方法:

第一种方法,第二步分别计算随机对照试验和观察性研究网络Meta分析的结果。在这种模型中不同研究设计中的异质性与研究设计有关,不同研究设计异质性是固定的。第三步是对不同研究设计中组间比较的网络Meta分析结果进行合并。如首先分别合并观察性研究和随机对照试验获取治疗措施A和治疗措施B的效应量 $RR_{AB随机对照试验}$ 和 $RR_{AB观察性研究}$,接着采取Meta分析的方法对两者合并获得治疗措施A和治疗措施B比较的最终网络meta分析结果。

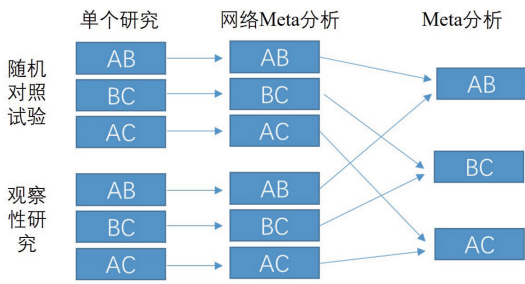


图2 贝叶斯等级模型（方法一）

第二种方法主要是在上述第二步有所不同。其第二步合并观察性研究和随机对照试验结果，进行传统的Meta分析获取所有对照组间的Meta分析结果，该结果并不考虑一致性假设检验，而是应用共同的异质性 τ^2 值。第三步就是在一致性假设检验的条件下，基于第二步采取随机效应模型获取网络Meta分析结果^[11]。

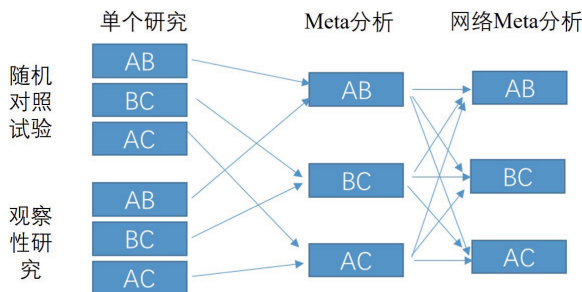


图3 贝叶斯等级模型（方法二）

第三种方法就是在上述第二步时针对不同的研究设计进行网络Meta分析，需要考虑一致性假设检验。在不同的研究设计内部可以认为异质性是一样的（采用固定效应模型），或者不同的研究设计赋予不同的异质性（直接对不同研究设计合并的异质性赋予不同的 τ^2 值）。第三步就是对观察性研究和随机对照试验的网络Meta分析结果再次进行网络Meta分析，这里需要在不同研究设计内和不同研究间分别考虑一致性假设检验，可以将观察性研究网络Meta分析结果和随机对照试验网络Meta分析结果作为一个多臂研究的结果进行合并，并构建方差-协方差矩阵^[11]。

4.4 修正幂验前分布 为了考虑随机对照试验和观察性研究之间的差异，采取修正幂验前分布

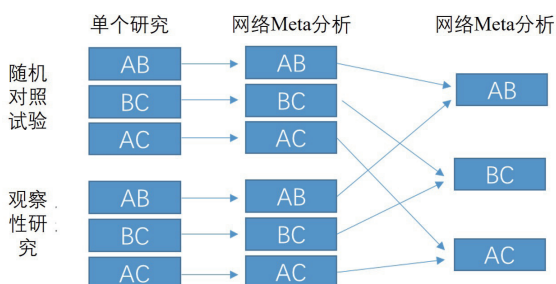


图4 贝叶斯等级模型（方法三）

转换^[8]。这种方法主要对观察性数据进行权重降级，通过引入权重因子 α （其值在0和1之间）来实现。0意味着观察性研究结果在网络Meta分析中的权重为0，1意味着观察性研究的结果可以全部用在网络Meta分析结果的计算中。 $L(\theta|Y)$ 为给定数据Y的似然函数，其主要有两部分组成：随机对照试验和观察性研究的似然函数，并根据权重因子 α 赋予一定的权重，即

$$P(\delta|RCT,OBS) = L(\delta|RCT) \times L(\delta|RCT,OBS)^\alpha P(\delta)$$

（ $P(\delta|RCT,OBS)$ ：最终似然函数； $L(\delta|RCT)$ ：随机对照试验的似然函数； $L(\delta|RCT,OBS)$ ：观察性研究似然函数； $P(\delta)$ ：权重）。

5 异质性分析

贝叶斯模型可以同时纳入随机对照试验和观察性研究的数据，采用复杂的统计学模型进行处理。观察性研究存在两种偏倚，系统性偏倚和非系统性偏倚，因此在合并观察性研究数据的时候必须承认这点。这两种偏倚可以通过似然函数进行调整。以 ξ 表示偏倚，那么非系统性偏倚就遵守 $\xi \sim N(0, t_\xi)$ 这个公式。采用这种方式模拟偏倚，可通过增加方差，降低观察性研究结果的权重，从而调整不精确性。对于系统性偏倚，我们通过给予一个固定值的方式， $\xi = \mu_\xi$ ，那么偏倚就应该是系统性偏倚和非系统性偏倚的加和，即为 $\xi \sim N(\mu_\xi, t_\xi)$ ^[11]。而在实践中，计算偏倚的大小是非常困难的，因此非常有必要采取合适的统计学方法探索和分析异质性。最常用的方法是敏感性分析、亚组分析和Meta回归。在纳入观察性研究的网络Meta分析中，即使观察性研究网络Meta分析和随机对照网络Meta分析结果一致，也应该对影响结果的重要因素，尤其是研究设计进行探索性分析。

6 讨论与展望

在常用的合并观察性研究和随机对照试验结果的统计学方法中，直接合并并不是最理想的数据处理方法，即使对观察性研究的数据进行降权处理。因为在合并观察性研究和随机对照试验，即使赋予了观察性研究结果一定的权重，但是权重的大小仍需要进一步探索。尽管目前研究显示，给予不同的降权因子，并不影响最终网络Meta分析的合并结果。采用观察性研究作为先验分布，就是将观察性研究的合并结果作为先验数据，这在一定程度上忽视了观察性研究证据对网络Meta分析合并结果的贡献程度。目前等级模型是观察性网络Meta分析的最佳合并方式，等级模型存在多种方法。最简单的是首先对随机对照试验和观察性研究分别进行网络Meta分析，接着再

合并观察性研究和随机对照试验的网络Meta分析结果。那么合并观察性研究网络Meta分析结果和随机对照试验网络Meta分析结果的时候是否要对观察性网络Meta分析的结果是否进行调整?是否进行降权处理?这都是未来需要研究的问题。

一致性检验也是观察性研究网络Meta分析的一个关键问题。是单纯评估合并结果中直接证据和间接证据的一致性,还是分别评估随机对照试验和观察性研究网络Meta分析中直接证据和间接证据的一致性,尚无成熟的推荐建议。我们建议评估观察性研究网络Meta分析的一致性,要对所有的直接证据和间接证据进行评估。

随着网络Meta分析方法学的广泛推广,纳入观察性研究的网络Meta分析应用将会越来越广泛。因此解决好观察性网络Meta分析的基本问题将是今后临床流行病学家和循证医学科学工作人员要关注的问题。未来的研究要探讨如何简化和标准化观察性研究网络Meta分析假设检验的过程,进而保证网络Meta分析结果可信。同时未来研究更应该探索和优化观察性研究网络Meta分析的分析方法,进而筛选观察性网络Meta分析最佳的统计分析模型。

参考文献

[1] 张珺,李伦,刘爱萍,等.观察性研究网状Meta分析统计分析方法简介[J].中国药物评价,2016,33(6):461-3.

- [2] 李伦,田金徽,姚亮,等.网状Meta分析的统计学基础、假设和证据质量评估[J].循证医学,2015,15(3):180-3.
- [3] 田金徽,李伦主编.网状Meta分析方法与实践[M].北京:中国医药科技出版社,2017.
- [4] Cameron C, Fireman B, Hutton B, *et al*. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities[J]. *Syst Rev*, 2015, 4: 147.
- [5] 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.
- [6] Thompson S, Ekelund U, Jebb S, *et al*. A proposed method of bias adjustment for meta-analyses of published observational studies[J]. *Int J Epidemiol*, 2011, 40(3): 765-77.
- [7] Zhang K, Arora P, Sati N, *et al*. Characteristics and methods of incorporating randomized and nonrandomized evidence in network meta-analyses: a scoping review[J]. *J Clin Epidemiol*, 2019, 113: 1-10.
- [8] Jenkins D, Bujkiewicz S, Martina R, *et al*. Methods for the inclusion of real world evidence in network meta-analysis. Available from <https://arxiv.org/abs/1805.06839>. Accessed 11 March 2021.
- [9] Jenkins D, Czachorowski M, Bujkiewicz S, *et al*. Evaluation of Methods for the Inclusion of Real World Evidence in Network Meta-Analysis - A Case Study in Multiple Sclerosis[J]. *Value Health*, 2014; 17(7): A576.
- [10] Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model[J]. *Stat Med*, 2013, 32(17): 2935-49.
- [11] Efthimiou O, Mavridis D, Debray TP, *et al*. Combining randomized and non-randomized evidence in network meta-analysis[J]. *Stat Med*, 2017, 36(8): 1210-26.

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; 联系电话: 010-84008479

网址: www.ebcvm.cn/www.ebcvmcj.com/www.ebcvmcj.org