

· 循证研究与临床转化 ·

网络Meta分析研究进展系列（十四）： 动态网络Meta分析简介

许建国¹, 高亚¹, 孙凤², 武珊珊³, 杨智荣⁴, 董圣杰⁵, 张天嵩⁶, 田金徽¹

【摘要】本文介绍了一种用于更新网络Meta分析的新方法即动态网络Meta分析, 阐述了该方法产生的背景、统计方法、制作步骤和可行性分析, 并展望了未来的发展和应用前景, 以期为研究人员了解和开展动态网络Meta分析提供参考。

【关键词】动态网络Meta分析; 网络Meta分析更新; 方法学进展

【中图分类号】R4

【文献标志码】A

开放科学(源服务)标识码(OSID)



Introduction to the Living Network Meta-Analysis Xu Jianguo*, Gao Ya, Sun Feng, Wu Shanshan, Yang Zhirong, Dong Shengjie, Zhang Tiansong, Tian Jinhui. *Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University; Key Laboratory of Evidence-based Medicine and Knowledge Translation of Gansu Province, Lanzhou 730000, China.

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

[Abstract] This paper briefly introduces a new method for updating network meta-analysis, that is living network meta-analysis. The background, statistical methodology, production framework, and feasibility analysis of this method were described, and the future development and application prospects were discussed, so as to provide a reference for researchers to understand and carry out living network meta-analysis.

[Key words] Living Network Meta-analysis; Updating Network Meta-analysis; Methodological progress

卫生保健专业人员和其他负责为患者提供安全和高质量医疗服务的决策者需要获得有关诊断、治疗和预后的最新研究证据。通常对于某一特定条件的医疗场景, 包含多种相互竞争的干预措施, 传统Meta分析(MA)不能系统呈现所有干预措施全貌, 网络Meta分析(NMA)可同时比较特定条件下多种干预措施, 为解决这一临床问题提供了可行方案^[1]。在证据使用者需求下, 证据生产者应该定期更新证据, 但是, 有研究显示50%的系统评价(SR)在完成最后一次检索14个月后才得以发表, 7%的SR在发表时已经过时^[2-4]。Cochrane系统评价手册建议SR应每两年更新一次, 由于SR的更新费时费力, 随着更多新的SR出现, Cochrane工作组也不得不根据临床需要和优先级对未来研究的重点进行调整^[5]。不完整、过时的证据综合必然带来资源浪费,

并对证据使用者产生广泛的影响。基于这一研究背景, 有学者提出动态系统评价(LSR)的概念, 通过对大量数据库全面、系统检索, 若有新的研究可纳入, 应立即更新SR^[6]。对于多种干预措施相互比较的NMA进行即时更新的动态网络Meta分析(LNMA), 又称动态累积网络Meta分析(LCNMA), 可视为传统LSR的拓展, 可获得关于某一医疗健康问题所有干预措施的所有现有证据的最新综合^[7]。本文在梳理介绍LNMA的基础上, 介绍了LNMA统计方法、制作步骤和现阶段制作可行性等关键内容, 同时展望了LNMA未来发展和应用前景, 以期为研究人员了解和开展LNMA提供参考。

1 LNMA简介

2014年, Elliott等提出了LSR的概念, 提出对原始研究进行持续跟踪, 补充代替不定期间断性更新模式作为持续更新MA的新框架^[6,8]。LSR可确定某一特定干预措施的证据是如何随着时间的推移而积累, 并为指南的制定提供最新证据。近十年, NMA由于克服了两种干预措施之间没有直接比较, 可通过间接证据支持临床决策而大受欢迎^[9], 但也有学者对其时效性表示担忧。2016年, Cr é quit等回顾性分析了关于晚期非小细胞肺癌二线治疗的NMA, 通过构建随机对照试验的累积NMA证据, 发现原有NMA并不能提供一套完整和最新的证据, 造成了研究资源的

基金项目: 北京大学教学新思路(2019YX007, 71009Y0026), 甘肃省科技计划资助项目(20CX4ZA027, 20CX9ZA112)

作者单位: ¹ 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; ² 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; ³ 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; ⁴ CB1 8RN 剑桥, 英国剑桥大学临床医学院初级医疗中心; 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

共同通讯作者: 孙凤, E-mail: sunfeng@bjmu.edu.cn;

田金徽, E-mail: tianjh@lzu.edu.cn

doi: 10.3969/j.issn.1674-4055.2021.07.01

浪费^[10]。该研究中Cr é quit等拓展了LSR的应用范围, 提出LNMA方法学框架, 为比较特定疾病所有干预措施提供了全面且最新证据综合和动态更新的方法。2018年, Nikolakopoulou等进一步研究发现LNMA通过直接比较和/或间接比较, 可比传统MA早2~6个月得出有效性比较的可靠结论, 指出LNMA在促进动态指南的及时发布、减少研究浪费并在新药上市评价中具有良好应用前景^[11]。此外, 国际权威决策机构, 如英国NICE(国家卫生保健卓越研究所)和WHO(世界卫生组织)等机构指出, 适用于NMA结果的证据分级方法即GRADE方法仍然适合于对LNMA结果进行分级, 进而为临床指南推荐建议的形成提供参考^[12]。

2016年, Cr é quit等在BMJ Open发表第一篇LNMA计划书, 2017年10月在BMC Medicine发表全文^[13,14]。Cr é quit等通过搭建在线网站(<http://livenetworkmetaanalysis.com>)对结果进行更新, 2018年12月和2019年8月通过该网站完成LNMA的2次更新, 原定于2020年2月的第3次更新截止2021年5月暂未见发表。2020年1月, Cochrane皮肤组将一篇2017年发表全文的NMA在第一次更新时转化为LNMA, 并描述计划每月进行一次文献筛选, 最近一次适时更新时间为2021年4月^[15,16]。截止2021年5月, 全球共发表LNMA7篇, 在研计划书不超过20篇。

2 LNMA的统计学方法

LSR与累积MA统计方法相同, 每当有新的研究纳入并更新MA结果时, 证据的合并效应量和精确性都会被重新计算, 更容易导致I类错误(假阳性)出现^[17]。序贯MA已被开发用于传统MA, 2018年, Nikolakopoulou等扩展了序贯MA为LNMA计算累积效应量提供了思路^[18]。在LNMA更新过程中, 考虑到不同研究间存在异质性, 一般采用随机效应模型对数据进行分析, 并考虑研究间方差作为异质性大小的先验信息。假定新纳入研究同样满足研究特征相似性假设, 即研究间效应量大小的传递性经评估被认为是合理的。一致性假设即不同干预措施间合并后的直接效应与间接效应一致, 是不同来源证据间相似性的统计学表现, 当干预措施间形成闭环时即可对不一致性进行检验。

假定LNMA中任意干预措施Y与干预措施X在纳入第*i*项($i=1, \dots, n$)研究之后, 效应量 $\hat{\mu}_{i,XY}^N$ 近似服从正态分布 $\hat{\mu}_{i,XY}^N \sim N(\mu_{i,XY}^N, (\sigma_{i,XY}^N)^2)$, 并假定方差 $(\sigma_{i,XY}^N)^2$ 已知且等于抽样方差 $se(\hat{\mu}_{i,XY}^N)^2$, 则接受零假设的Z指数(效应量大小与标准误差之比)计算公式为: $Z_{i,XY}^N = \frac{\hat{\mu}_{i,XY}^N}{se(\hat{\mu}_{i,XY}^N)} \sim N(0,1)$ 。这种计算方法所得到的Z指数与NMA中合并效应量的精度直接

相关, 并由累积样本量所决定。根据此公式, LNMA中每一个比较(Y vs X)所包含总信息量可表示为 $I_{i,XY}^N = \frac{1}{se(\hat{\mu}_{i,XY}^N)^2}$ 。绘制每次更新时Z指数与总信息量的成对关系, 以图表的形式呈现两种干预措施间证据积累情况, 并可构建证明干预有效和无效的终止边界。

在LNMA过程中, 要进行多次重复显著性检验, 干预有效的终止边界遵循一致性假设下的 α 分割界值原则。成组序贯试验方法的应用使其在期中分析的研究数量和时间上具有灵活性, 当穿过设定的终止边界时, NMA的结果可成为结论性的证据。Nikolakopoulou等在LNMA统计方法中, 有效性终止边界的构建采用类似O'Brien-Flemming提出的单调递减边界^[18,19], 连续更新下显著性水平 $\alpha(t_{i,XY}^N)$ 应满足 α 损耗函数: $\alpha(t_{i,XY}^N) = 2(1 - \Phi(z_{1-\alpha/2} \sqrt{I_{i,XY}^N}))$, 其中 Φ 是标准正态分布函数。 α 损耗函数 $\alpha(t)$ 表示在每次期中分析中损耗占总显著性水平的比例, 信息分数*t*表示已经积累的信息所占的比例。

α 损耗函数用来计算每次更新时检验的名义显著性 $\alpha(t_{i,XY}^N)$, 有效性终止边界 $E(t_{i,XY}^N)$ 对应 $\alpha(t_{i,XY}^N)/2$ 。当Z值穿过边界时, 意味着对于干预措施Y与X的比较, NMA已得出结论性的答案。图1展示了 $(Z_{i,XY}^N, I_{i,XY}^N)$ 成对关系图中一个有效性终止边界的假设框架, 假定按序列顺序依次合成9项研究, 在综合纳入第8项研究之后, 得出确定性结论($P < 0.05$)。需要注意, 即使NMA效果估算值是结论性的, 如果网络中的直接比较没有足够的证据来推断其结论性, 则间接证据仍可作为判断有效性终止边界的重要来源。同样地, LNMA效果估算值或许会在更新中仅通过间接比较可得到确切的结论。当效果估算值在下次LNMA更新中出现与正常预期相反的情况而变得不确定时, 可能是由于使用随机效应模型合并数据时, 异质性或不一致性的显著增加导致了较低的精确效应估计。在这种情况下, 需要对可变性的来源进行充分探索和解释, 才能将新的证据纳入NMA进行更新。无效边界的确定同样类似传统累积MA, 大致有4种主要方法被推荐用于确定无

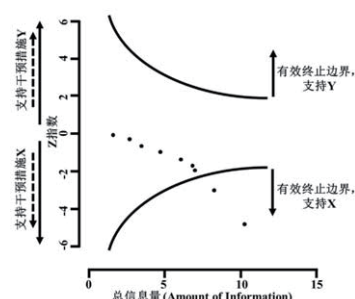


图1 Z指数与总信息量($Z_{i,XY}^N, I_{i,XY}^N$)成对关系图

效停止边界：条件效能法、预测效能法、序贯概率比检验及 β 损耗函数^[18,20]。

LNMA中所使用的序贯MA存在诸多争议^[21]。有效性终止边界的设定可减少I类错误（假阳性）发生的几率，并判断当前所获得的证据是否足够。无效边界的设定可以提示适时停止原始研究的时机，但过早结束试验可能会增加II类错误（假阴性）的发生，即便效果估算值穿过了无效性终止边界，LNMA研究结果也不能决定未来开展更多原始研究的临床意义和价值。LNMA中使用的序贯分析方法鼓励基于统计显著性的推断，其次研究结果会影响未来试验设计的决定。Cochrane工作组认为当进一步证据不太可能挑战目前的结论时，新的试验是不必要、昂贵且不道德^[22]。LNMA的结果不是用来决定是否进行进一步的研究并确定其研究设计，而是为未来研究提供建议，找出现有证据库中的潜在差距，并讨论LNMA结果对未来研究的影响，必要时仍建议开展新临床试验。

3 制作LNMA的步骤

LNMA并不是一种全新的NMA，而是用来更新NMA的新模式，强调其区别于传统不定期间歇性更新的模式^[23]。制作LNMA第一步是制作一篇纳入最全最新证据的NMA，以作为今后更新的初始版本。在LNMA的初始版本或先前发布的计划书中，应介绍接下来的更新团队人员分工、更新计划和步骤、明确描述检索方法、各来源数据库的检索频率、未来可纳入研究整合到证据体的决策框架和时机，并对相关统计方法进行描述。在初始版本的基础上，定期重复迭代，以保证LNMA的时效性。LNMA进行更新原则主要有：NMA研究领域进展相对较快且不断有新的研究发表（如COVID-19相关进展较快研究领域）；临床实践指南中基于NMA的推荐意见有必要持续更新，以获得最新证据（动态指南将基于LSR进行更新）；基于现有证据精确性不高，纳入新的研究极有可能改变NMA结果^[8]。已发表的NMA，可以评估动态更新的必要性，决定是否转换为一篇LNMA。

与任何新兴事物的发展规律相同，LNMA总是在摸索中不断发展，其制定步骤也在不断完善，及时纳入原始研究、实时更新NMA结果，是LNMA的根本体现和根本要求。LNMA每次迭代与传统SR制作流程类似，包括以下6个步骤^[10]（图2）：①定期适时检索；②原始研究的筛选和纳入；③数据提取；④偏倚风险评估；⑤NMA结果更新；⑥发表和传播。在迭代更新流程中，可使用检索平台的定期推送功能进行检索，如经

筛选没有发现可纳入的新研究，则回到第一步继续进行定期检索，记录说明相关检索与筛选文献结果^[24]；如经筛选后有符合纳入标准的文献，则对文献数量和对应的工作量进行评估，决定将新文献整合到已有证据体并更新NMA结果。

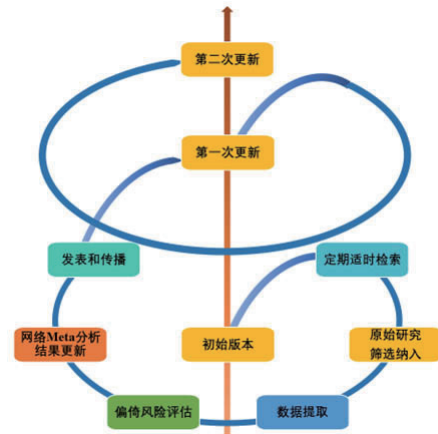


图2 LNMA更新步骤

每当有新的原始研究可供纳入时，需要对新纳入原始研究进行独立的偏倚风险评估，并重新计算结局指标合并效应量，重新使用GRADE评估证据质量。规范且高标准的制作SR，需要充足的人力资源和时间保障，由于LNMA本身的复杂性，需要加大工作中的成本投入同时，需要一个稳定的团队。鉴于LNMA需要一个平台实时更新其最新进展，传统纸质媒介并不适合定期报告此类实时合成证据的结果。在线发布全文并更新最新检索结果说明可增加其透明性，促进LNMA传播。

4 制作LNMA的可行性分析

尽管LNMA的概念和框架于2016年提出，但考虑到承担LNMA所需付出的巨大工作量，在更新纳入研究数据分析诸如同质性检验、相似性检验、一致性检验，以及何时停止更新等诸多问题上存在挑战，且在开始制作LNMA之前，作者团队必须先做出承诺来保障未来一段时间内对NMA的实时更新，这使得很多团队望而却步。LNMA对制作团队的其中一个要求是能够保证在一段时间内对NMA进行维护并更新，研究人员需定期检索与筛选文献，根据文献量适时更新NMA结果，但现实情况是这一承诺没有被很好的执行。目前缺失一套有效的监管制度来保证LNMA的时效性，仅凭作者团队的前期承诺保障显然是不够的。另外对于LNMA应该何时停止动态更新仍没有统一的标准^[8]，如在药效评价的NMA中，不断有新的药物或新的原始研究出现，将使得证据网络图不断发生变化。

关于LNMA可行性的问题，关键是确定检索和更新NMA结果的频率和时机。连续更新方案的

实现仍然具有挑战性,并且难以确定先验最优更新频率,证据产生的速度是每个临床问题所特有的,必须具体临床问题具体分析。Créquit等为了评估LNMA的可行性,系统分析了不同更新频率(4、6和12个月)下,每次更新时所需纳入的新研究数和更新时工作量相对于初始制作NMA版本的工作量的百分比^[25]。考虑不同医学领域证据生产速度不同,研究共评估了涵盖17个不同医疗领域的77篇NMA。结果显示,平均有78%的NMA每年可纳入新研究不超过4篇,只有7%的NMA年更新纳入文献量超过了7篇。在更新频率为4、6和12个月的情况下,更新纳入文献量中位数(四分位数)分别为0(0,1)、1(0,2)和2(1,4)篇,平均每次更新对应的工作量分别为4%、5%和11%。根据不同医疗领域原始研究更新的频率不同,Créquit等建议每6~12个月进行一次更新。无论NMA初始版本规模如何,其后续更新工作量一般不会超过初始工作量的15%,这对于一个成熟专业、结构合理的作者团队是可接受的,但具体的更新频率往往与研究问题的证据生产速度有关,最佳的更新频率意味着一个可接受的工作负荷,确保NMA结果保持最新,同时不会产生不必要的负担;此外, LNMA的具体更新时间需要根据证据使用者如指南制订团队、政策决策者需求来确定。如何更好地促进LNMA向动态指南转化,也是目前研究的热点与未来发展的方向^[26]。LNMA往往比动态MA具有更长的更新持续时间,综合现有已发表LNMA,建议持续更新时间应不少于2年,当持续更新的必要性经充分评估后不再满足时,可考虑停止动态更新。

一些计算机辅助自动化方法、策略和工具的出现,正在改变和简化LNMA的制作流程^[27,28]。人力和计算机辅助相辅相成,可以提高LNMA的可行性和持续性,使得研究人员更有效地编写、传播和更新LNMA。如Epistemonikos(epistemonikos.org)、PaperBot^[29]、GAPscreeener^[30]、RobotReviewer^[31]、ContentMine(contentmine.org)、GRADE pro GDT^[32]、CINeMA^[33]、StArt^[34]等工具的出现,可在文献检索与筛选、数据提取与分析、偏倚风险评估、证据分级、撰写报告、更新结论等步骤提供帮助。不难预见,随着计算机辅助系统的发展和智能化,未来在LNMA的更新流程中,人力所占的比重将会逐渐降低,届时当LNMA初始版本完成之后,理想情况下可由计算机系统完成自动化更新,达到真正意义上的实时LNMA。基于大数据和人工智能, LNMA生产、制作、更新等全阶段 workflows 的开发,将会使得人力和自动化以更有

效更高效的方式结合,为LNMA制作和传播效率的提高带来实质性的贡献和突破。

5 总结

本文梳理了LNMA的起源与发展,介绍了其统计学方法和更新步骤,并讨论了其现阶段制作的可行性以及机遇与挑战。现阶段LNMA并没有规范化报告发表模式,随着计算机辅助自动化技术的发展、统计学方法的成熟、评价工具的制定、发表传播平台的搭建,可以预见未来LNMA的报告质量将逐步得到提升,更新传播共享流程也将更加规范。将所有可用治疗方案及时纳入LNMA中,将进一步促进干预措施的疗效优先排序和决策制定,最终弥合证据和卫生保健实践之间的差距。

参考文献

- [1] Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses[J]. CMAJ, 2009, 181(8):488-93.
- [2] Sampson M, Shojania KG, Garrity C, et al. Systematic reviews can be produced and published faster[J]. Journal of clinical epidemiology, 2008, 61(6):531-6.
- [3] Shojania KG, Sampson M, Ansari MT, et al. How quickly do systematic reviews go out of date? A survival analysis[J]. Annals of internal medicine, 2007, 147(4):224-33.
- [4] Jadad AR, Cook DJ, Jones A, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals[J]. JAMA, 1998, 280(3):278-80.
- [5] Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up?[J]. PLoS medicine, 2010, 7(9):e1000326.
- [6] Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap[J]. PLoS medicine, 2014, 11(2):e1001603.
- [7] Vandvik PO, Brignardello-Petersen R, Guyatt GH. Living cumulative network meta-analysis to reduce waste in research: A paradigmatic shift for systematic reviews?[J]. BMC medicine, 2016, 14:59.
- [8] Elliott JH, Synnot A, Turner T, et al. Living systematic review: 1. Introduction—the why, what, when, and how[J]. Journal of clinical epidemiology, 2017, 91:23-30.
- [9] 田金徽, 李伦, 赵晔, 等. 网状Meta分析的撰写与报告[J]. 中国药物评价, 2013, 30(6):321-3, 333.
- [10] Créquit P, Trinquart L, Yavchitz A, et al. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer[J]. BMC medicine, 2016, 14:8.
- [11] Nikolakopoulou A, Mavridis D, Furukawa TA, et al. Living network meta-analysis compared with pairwise meta-analysis in comparative effectiveness research: empirical study[J]. BMJ, 2018, 360:k585.
- [12] Kanters S, Ford N, Druyts E, et al. Use of network meta-analysis in clinical guidelines[J]. Bulletin of the World Health Organization, 2016, 94(10):782-4.
- [13] 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.