Meta 分析的综合评价

沈 宁¹, 胡良平^{1,2*}

(1. 军事医学科学院生物医学统计学咨询中心, 北京 100850;

2. 世界中医药学会联合会临床科研统计学专业委员会, 北京 100029

* 通信作者: 胡良平, E-mail: lphu812@sina.com)

【摘要】 本文旨在提出一种新的评价方式对 meta 分析方法及相关文章进行综合评价, 即对大量的单个 meta 分析研究进行汇总评价。古人云: 万变不离其宗。国内外众多不同领域的学者应用 meta 分析这种方法发表了大量的学术论文, 这些文章在研究设计、实施和写作思路等方面都有相通之处。然而, “循证而行”的众多 meta 分析类文章, 呈现出在“设计”“实施”“分析”和“结论”等环节方面良莠不齐, 其中有些还存在严重的瑕疵。笔者从研究设计、文献来源和提取信息三个角度对 meta 分析类文章进行分析, 指出其存在的要害问题, 以期提升单个 meta 分析类研究课题质量提供对策与建议。

【关键词】 系统评价; Meta 分析; 综合评价; 研究设计; 数据库; 文献检索

中图分类号: R195.1

文献标识码: A

doi: 10.11886/j.issn.1007-3256.2017.03.004

Comprehensive evaluation on the meta - analysis

Shen Ning¹, Hu Liangping^{1,2*}

(1. Consulting Center of Biomedical Statistics, Academy of Military Medical Sciences, Beijing 100850, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China

* Corresponding author: Hu Liangping, E-mail: lphu812@sina.com)

【Abstract】 This paper aims to provide a new angle to evaluate the method of meta - analysis and the articles using this method, that is assessing large amount of individual articles in a comprehensive way. As the old saying goes while the appearances may vary the essence remains unchanged. Many scholars worldwide published a large number of academic papers in different fields by using the same method, meta - analysis. Therefore, we found similarities in the aspects of research planning, implementation and writing through these papers. However, we also found that, although, a lot of meta - analysis articles use this evidence - based approach, the standardization of design, implementation, analysis and conclusion presented in these papers seems uneven. Especially, some of the articles have serious flaws. We analyze the meta - analysis articles from the following three aspects: research design, literature source and information extraction, point out the critical problems and provide methods and suggestions in order to improve the quality of individual meta - analysis research.

【Keywords】 Systemic review; Meta - analysis; Comprehensive evaluation; Study design; Database; Literature searching

1 概 述

1.1 五大类研究课题

当今世界, 人类所开展的科学研究非常之多, 所涉猎的领域和范围非常之广, 学者们很难用简短的语言将其全部囊括进去。但是, 按研究的手段或课题的类型划分, 不外乎有以下五大类。

第一类, 试验性研究课题。即研究者通过设定某些特定的环境或条件, 按一定的操作规程和质量控制手段, 去实施某些特定的处理, 从而发现已经呈现出来的现象或揭示出事物内在的固有的变化和规律; 或者制造或加工出具有某种功能或作用的物件, 如药物、仪器设备、武器、粮食、蔬菜等。

第二类, 观察性研究课题。研究者通过谋划某些希望了解或探究的内容、事物或现象, 按一定的操

作规程和质量控制手段, 去调查、打听、采访、观察或测量, 以期获得必要的信息或资料, 再通过对其整理、归纳、分析与提炼, 得出对所关心的事物或现象的认识和评价。

第三类, 理论性研究课题。纯方法学上的思考和推理, 从而揭示某些事物或现象内在的固有的变化和规律。例如, 我国数学家陈景润研究哥德巴赫猜想问题以及王元研究数论问题^[1]。

第四类, 文献性研究课题。基于文献资料, 对已发表的多项同类研究课题进行归纳、整合、总结、分析和推理, 期望得出关于所探究问题的一个基于“更大样本量、集成更多信息”且在更大范围内具有“高度概括性、重现性和可信性结论”的一类研究课题。“系统评价”^[2]是此类课题的“广义通用名称”, 而“meta 分析”是此类课题的“狭义通用名称”。

第五类, 真实世界研究课题。采用类似于“观察性研究”的手法, 但没有系统完善的“调查设计方

案”指导,仅对与研究者关心的某些事物或现象有关且自然产生或发生的过程或结果进行记录、描述、呈现、总结、分析和推理,得出近乎关于“总体”的结果和结论。近年来,关于“大数据”^[3]的研究热潮或许是此类研究课题的“代名词”。

1.2 何为 meta 分析的综合评价

文献[2]第3页上指出,meta 分析最早于 1861 年由英国皇家天文学家 George Biddell Airy 提出。自此至今,全世界的学者已经发表的有关 meta 分析的学术论文不计其数。

人们不禁要问:系统评价/meta 分析提供的“证据”一定是真实、重要、绝对可信的吗?答案是否定的!理由很简单:不同的研究者所做的 meta 分析研究课题,由于受其学问深浅、知识结构、设计方案的科学和严谨程度、制定的标准操作规程的质量、质量控制的实时和严格程度、数据搜集和管理的质量、资料统计分析的精准和全面程度等的影响,所得结果和结论的精准程度与可信度可能参差不齐。于是,证据质量分级和推荐强度系统(简称 GRADE 系统)于 2004 年应运而生^[2]。

笔者脑海中迸发出一个新的问题:谁会采用前述提及的“GRADE 系统”对每一篇拟发表的“meta 分析研究论文”进行评价呢?事实上,“GRADE 系统”只是一面业已存在的“镜子”或“标准”,为每一位拟发表“meta 分析论文”的作者提供了一定的参考。但是,对全世界拟做这方面课题研究者的科学研究素养和技能的提高,其影响力微乎其微。因为很多已发表的 meta 分析方面的论文,在许多环节和细节上存在较为严重、甚至非常严重的瑕疵。

正因如此,笔者认为:有必要对迄今为止的所有“meta 分析”类学术论文做一下全面“扫描”和“透析”,从头至尾,从其“表现型”找出其“原型”,并进行系统全面地梳理、分析与论证,呈现其“标准型”^[4-5],从而为人们开展此类课题研究提供一个可资借鉴或模仿的标准模板,这就是对“meta 分析”的综合评价,取名为“meta 分析的 meta 分析”。显然,它属于“广义综合评价”之一^[6]。

2 实施单个 meta 分析的关键要素

2.1 选题

阅读每篇 meta 分析的论文,首先映入读者眼帘的是关于“XXXX 的 meta 分析”。这里的“XXXX”就是研究者希望回答的科学问题,通常都是生物医学或临床上的热点问题或尚无定论的专业问题。事实上,它就是一个“meta 分析研究课题”。选题非常重

要,因为它涉及到该项研究是否有价值,应具有新颖性、重要性、实用性。

2.2 设计

有人认为:开展 meta 分析课题研究,不同于试验或调查研究,就是把别人所获得的试验研究资料拿来,运用统计分析软件将它们进行合并计算即可。因此,不存在“设计的问题”。这是被“meta 分析”的“表现型”蒙蔽了眼睛。事实上,即便是常规的统计分析方法,如差异性检验、相关与回归分析等,在使用之前也应当考察“资料是否值得分析”,而不是盲目地选择某种统计分析方法去直接处理。

验证所收集的资料是否值得进行 meta 分析的关键在于考察从定下“选题”到“准备开始进行 meta 分析”的全过程是否存在“瑕疵”,严格地说,只有在前述的“全过程”无懈可击的前提下,所收集到的资料才是值得进行 meta 分析的。然而,做到这一点的依据或保证就是 meta 分析之前的“设计”。

“设计”一词听起来很“玄乎”,其实,“设计”就是普通人最容易理解的“安排”。研究者怎样“安排”,才能把与“题目”有关的文献“找准找全”;才能把收集到的文献中所包含的全部与完成好“课题”有关的信息都呈现出来。

2.3 实施

所谓实施,就是按照“设计”去具体落实。涉及到检索数据库的种类、检索策略(包括时间范围、语种、关键词等)、纳入与排除文献的标准、从文献中提取信息的人员素质、所使用统计分析软件的功能和权威性、实际实施 meta 分析人员的统计分析水平及解读统计分析结果人员的统计与专业水平等。

3 从多种不同角度考察单个 meta 分析的质量

3.1 从研究设计角度对 meta 分析原始文献的分析

可以从八个方面来展开:①对原始文献中处置伦理学问题的分析;②对原始文献中处置研究类型问题的分析;③对原始文献中处置设计类型问题的分析;④对原始文献中处置比较类型问题的分析;⑤对原始文献中处置三要素问题的分析;⑥对原始文献中处置四原则问题的分析;⑦对原始文献中处置标准操作规程问题的分析;⑧对原始文献中处置质量控制问题的分析。

在上述八个方面是否做得到位、高质量,是体现单个 meta 分析研究课题在实施 meta 分析之前的“整体质量”的“命脉”。

3.2 从文献来源角度对 meta 分析的分析

可以从三个方面来展开:①对 meta 分析文献检索涉及数据库的分析;②对 meta 分析中原始文献所用语种的分析;③对检索文献策略的分析。

在以上三个方面是否考虑得尽可能全面、细致,是体现单个 meta 分析研究课题在实施 meta 分析之前的“文献质量”的“命脉”。

3.3 从原始文献提取信息角度对 meta 分析的分析

可以从三个方面来展开:①对 meta 分析文章中数据提取者资质和提取方式的呈现分析;②对 meta 分析文章中信息提取方案的内容分析;③对 meta 分析文章中原始研究质量分析的呈现内容分析。

以上三个方面是否考虑得尽可能全面、细致,是体现单个 meta 分析研究课题在实施 meta 分析之前的“资料质量”的“命脉”。

综上所述,从以上三个大的视角来考察或审视单个 meta 分析的质量,笔者可以谨慎地得出如下基本判断:在国内外已发表的 meta 分析类论文中,存在的问题还是相当普遍和严重的,值得未来 meta 分析类研究课题的作者们引以为戒。

例如,在文献[7-8]中,对于伦理学问题的考虑就不够到位(具体地说,漏掉了伦理委员会名称以及文件批号);在文献[9]中,缺乏对伦理问题的声明;在文献[10]中,未明确交代对照组是如何设置的。

在文献[11]中,涉及了五种试验研究类型,其中在纳入与排除标准中提到纳入的试验类型不限,包括随机对照试验、非随机的同期对照试验、病例对照研究、队列研究和病例系列报告。这样的 meta 分析是把“荟萃”当作“大杂烩”了,其意义和价值也就可想而知了。

由于篇幅有限,不能将存在的问题逐一举例并辅以全面、细致剖析。笔者将选择合适的时机或场合予以阐释,此处从略。

4 从多种不同角度提升单个 meta 分析的质量

4.1 注重提升实施 meta 分析之前的设计质量

如同开展试验研究(包括临床试验研究)和调查研究课题一样,开展 meta 分析研究课题时,也必须高度重视“设计”。因为“具有高质量的设计”是课题成功的“根基与保证”,其内容十分丰富,但主要应抓好八个方面,参见本文“3.1 节”;还可参阅本刊自 2016 年第 3 期至 2017 年第 2 期“科研方法专

题”中的所有文章(约 24 篇)。

4.2 注重提升实施 meta 分析之中的操作质量

在具有高质量“meta 分析设计方案”的指导下,实施“文献检索”“信息提取”“信息整合”“数据分析”“结果呈现与解读”和“结论陈述”,这是实施 meta 分析之中不可避免的一系列序贯操作。在前述的每一道“工序”中,都必须严格按照事先制订好的系统、科学、全面且具有可操作性的“标准操作规程(SOP)”^[12]具体实施,同时,从头至尾都必须进行严格、实时、精准的“质量控制”。确保各环节或工序之间“无缝衔接”,确保每个环节或工序上的“低质率”降至为零(这就是曾经风靡一时的“六西格玛管理理念”^[13-14])。

参考文献

- [1] 日本数学会. 数学百科辞典[M]. 北京: 科学出版社, 1984: 321-403.
- [2] 罗杰, 冷卫东. 系统评价/Meta 分析理论与实践[M]. 北京: 军事医学科学出版社, 2013: 3-178.
- [3] 赵勇. 架构大数据: 大数据技术及算法解析[M]. 北京: 电子工业出版社, 2015: 1-19.
- [4] 胡良平. 统计学三型理论在实验设计中的应用[M]. 北京: 人民军医出版社, 2006: 1-10.
- [5] 胡良平, 张天明. 新概念统计学: 统计思维与三型理论在医药领域中的应用[M]. 北京: 中国中医药出版社, 2015: 1-274.
- [6] 郭春雪, 沈宁, 胡良平. 整合科研设计与统计分析的方法: 广义综合评价[J]. 四川精神卫生, 2017, 30(1): 16-21.
- [7] 重组人脑利钠肽多中心研究协作组. 重组人脑利钠肽治疗心力衰竭安全性和疗效的开放性随机对照多中心临床研究[J]. 中华心血管病杂志, 2011, 39(4): 305-308.
- [8] 武剑, 赵璟妍, 黄疆, 等. 中国缺血性脑血管病患者脑动脉狭窄分布与危险因素的茶萃分析[J]. 中华医学杂志, 2015, 95(29): 2396-2402.
- [9] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement [J]. Int J Surg, 2010, 8(5): 336-341.
- [10] 周蔚然, 沈颖, 刘小梅. 唤醒疗法与去氨加压素治疗儿童原发性遗尿症的非随机对照试验[J]. 中国循证儿科杂志, 2015, 10(2): 155-159.
- [11] 黄紫涵, 王小虎, 田金徽, 等. 早期乳腺癌保乳术后单导管球囊近距离加速部分乳腺放疗单组率的 Meta 分析[J]. 中国循证医学杂志, 2016, 16(11): 1298-1305.
- [12] 胡良平. 课题设计与数据分析: 关键技术与标准模板[M]. 北京: 军事医学科学出版社, 2014: 104-136.
- [13] 张驰. 六西格玛试验设计[M]. 广州: 广东经济出版社, 2003: 1-30.
- [14] 马逢时, 周晔, 刘传冰. 六西格玛管理统计指南[M]. 北京: 中国人民大学出版社, 2007: 1-9, 535-650.

(收稿日期: 2017-06-04)

(本文编辑: 吴俊林)