

# 医学论文统计学报告水平评价量表简表的研制\*

刘清海<sup>1)</sup> 方积乾<sup>2)</sup>

收稿日期:2009-01-20  
修回日期:2009-03-23

1) 中山大学学报编辑部//编辑学与出版研究中心, 510080 广州中山二路74号, E-mail: liuqh@mail.sysu.edu.cn

2) 中山大学公共卫生学院流行病学与统计学系

**摘要** 在以前的工作中,我们研制了一份对医学论文统计学报告水平的评价量表,可作为实时、内部评价医学论文的科学水平。为了适应于不同背景的需要,我们进一步研制了一份评价量表的简表,该简表共含10个项目,其中7项选自于能够最大效力地区分不同级别杂志论文报告水平的统计学项目,补充3项为对应于Jadad量表的评价论文科学水平的3项。该量表简表适用于快速评价国内医学杂志或论文的科学报告水平。

**关键词** 医学论文 统计学报告水平 评价量表简表 论文评价 文献质量评价

基于已发表医学论文仍存在较多的统计学问题,且这些统计学问题突出地体现在报告水平方面<sup>[1]</sup>,于是,在我们以前的系列研究中,我们研制了一份医学论文统计学报告自查清单<sup>[2]</sup>、一份医学论文统计学报告水平评价量表<sup>[3]</sup>,同时,对国内外的几种典型杂志进行了统计学报告水平的统计和比较,发现国内医学论文的统计学报告水平与国外相差较大<sup>[1]</sup>。国内几年前已有医学编辑认为应将统计学评审内容纳入医学期刊的评价内容,加强医学编辑对论文统计学问题的鉴审工作<sup>[4]</sup>,也有医学统计学者研制了临床试验文献质量评价量表<sup>[5]</sup>,含19项共100分,与我们前期的工作成果基本一致。然而,若用于快速评价,20项还是偏多,我们拟进一步压缩项目,希望以少于10项的简表既达到能区分不同级别统计学报告水平的医学论文,又不失核心项目的目的。

## 1 量表简表项目的选取方法

前期的研究工作中,我们得到了一份医学论文统计学报告水平的评价量表全表,共27项,经5位评价者应用该量表对国外2种著名医学杂志、国内2种中华医学系列杂志和2种一般医学杂志的有对照的临床研究论文进行统计学报告水平的评价,结果,不同评价者之间的一致性较好,对不同级别杂志的统计学报告水平的区分度也很好<sup>[3]</sup>。然而,我们希望能进一步简化量表,于是需要对量表进行进一步的分析。我们采取了两个原则:一是最少项目最大区分原则,二是核心项目保留原则。

### 1.1 最少项目最大区分原则

完成信度和效度的分析后,我们希望能以最少的项目区分出上述三个级别杂志的统计学报告水平,于是再对27项

的量表全表应用的结果计算各项项目的方差及仅取中文文献计算方差,结果如表1(按中文文献方差排序)。将评价中文文献时各项目方差最大的逐项保留下来时,发现至少要保留7项时(表1中方差等于0.26及0.25者),才能区分出三种级别的杂志(表2);若保留方差最大的6项,则只能区分出国内一般医学杂志,而中华医学系列则和外文杂志重叠。据此,依最少项目最大区分原则可认为量表简表以保留7项为宜,此时总分为7分,可分为3分以内,4~5分,6分及以上三级。

表1 应用量表全表评价中外医学文献时各项项目的方差(n=54)

项目代号	中文方差	全部方差	项目代号	中文方差	全部方差
M4	0.00	0.19	R4	0.23	0.18
M6	0.08	0.19	R6	0.23	0.23
R3	0.10	0.22	R7	0.24	0.21
R9	0.12	0.10	D4	0.24	0.24
R5	0.12	0.24	M5	0.24	0.19
D3	0.12	0.09	R13	0.24	0.19
R8	0.14	0.13	M3	0.25	0.21
M2	0.14	0.14	R10	0.26	0.25
D2	0.16	0.24	M8	0.26	0.25
R12	0.16	0.25	R14	0.26	0.25
R11	0.19	0.21	D5	0.26	0.23
R2	0.19	0.25	R1	0.26	0.23
M7	0.21	0.25	M1	0.26	0.25
D1	0.21	0.23			

表中项目编号含义同文献[3]

\* 基金项目:广东省软科学课题资助项目,编号:2005B70101121

表2 取最大方差7项时各级杂志得分情况

	篇数	均数	标准差	95%可信区间	最小值	最大值
外文文献	18	5.44	1.097	4.90, 5.99	3	7
中华医学系列杂志	18	4.56	1.723	3.70, 5.41	1	7
国内一般医学杂志	18	2.61	0.916	2.16, 3.07	1	4
合计	54	4.2	1.742	3.73, 4.68	1	7

$F=22.633, P<0.001$ , 经 SNK-q 检验, 在 0.05 的水平能完全检验出三组的不同水平。

### 1.2 核心项目保留原则

依照最少项目最大区分原则定下来的7项指标, 它的意义在于区分, 反过来说, 即各级杂志都会报告的核心项目和各级杂志都不太会报告的边缘项目因为都不具有区分意义而不会被保留, 而保留下来的则是那些有些杂志报告了有些杂志则未报告的中等重要项目。而作为一份评价量表, 因其具有导向性, 因此应该保留核心项目。核心项目比较难以量化界定, 因此我们参考了有关文献。参照周登远综述的国外在文

献质量评价中的多个量表<sup>[6]</sup>, 国外有不少于25种评价临床试验文献质量的量表, 其中 CONSORT 声明(22 项或细分成 57 小项)、Jadad 量表(3 项)、Delphi 清单(8 项)和 Chalmers 量表(约 20 项, 各项权重不同)应用较多, 其项目数以 Jadad 量表最少最核心, 为 3 项。Jadad 量表于 1996 年确定<sup>[7]</sup>, 现在还大量应用于国内医学文献质量的评价<sup>[8]</sup>, 尤其是 Meta 分析时作为文献入选的标准。其他几个量表影响也很大, 尤其是 CONSORT 声明, 不过由于项目较多, 都属于统计学报告的重要项目, 但没有界定为核心项目。于是我们采用了 Jadad 的最简 3 项指标作为核心项目, 在我们评价量表全表中保留下来, 即: 随机的描述、盲法的描述、随访失访数的描述。

## 2 量表简表的最终项目

按照上述的最小项目最大区分原则和核心项目保留原则, 我们从评价量表的全表中保留下来 10 项作为有对照的临床研究统计学报告质量评分简表, 两种原则保留的项目无交叉, 见表 3, 其中带 \* 号者为核心项目, 项目编号已经重排。

表3 有对照的临床研究统计学报告质量评分简表

编号	项目描述	评分说明	评分
M1	目标人群描述是否清楚	如年龄、地理、转诊	1, 0
M2	是否包括入选标准与排除标准	有一项即得分	1, 0
M3	有无定义观察或研究终点	个体或整体皆可	1, 0
M4*	有无说明分组的具体方法	仅“随机”二字可计 1 分	1, 0
M5*	有无说明准“盲”对什么因素	未采用盲法亦应如是说	1, 0
R1	研究与随访的起止时间是否清楚	有一项不明则不得分	1, 0
R2*	是否说明随访失访数	若无失访应如是说	1, 0
R3	主要指标统计检验的实际方法是否清楚	若与方法中的描述同, 亦应如是说	1, 0
R4	有无分组报告负性反应或事件的人次与程度	未分组不得分, 讨论中不计分	1, 0
D1	有无结合其他文献加强或平衡本文结论	纯解释性不计分	1, 0

答案肯定者得 1 分, 否则得 0 分。M: 材料与方法部分; R: 结果部分; D: 讨论部分。R1 也可在方法部分。\* 代表核心项目

## 3 各种量表的比较与应用

鉴于已经发表的医学文献仍然存在着不少统计学问题, 因此国内外的专家自上世纪中叶开始一直在致力于提高医学论文统计学方法的使用与报告水平, 最初为对已发表文献统计学问题的发现与总结, 到了上世纪 80 年代才有针对医学论文的统计学指南, 后来发现将统计学方法的正误和报告的全面与否合在一起仍无成效。于是上世纪 90 年代开始出现了以指导医学论文作者如何报告统计学项目的 CONSORT 声明等<sup>[9]</sup>。同时, 随着循征医学的发展及 Meta 分析的需要, 评价医学文献质量的各种量表也逐渐多了起来。不过, 国外在 1993 年以前的量表基本没有按照正规的量表制作过程进行项目的筛选与评价, 多数为参照已发表的临床试验报告和作者理解筛选项目, 因此应用范围受到影响。按周登远的总结<sup>[6]</sup>, 2004 年以前国外有不少于 25 种比较正规的医学文献质量评价量表, 应用较多的 CONSORT 声明、Jadad 量表、

Delphi 清单和 Chalmers 量表等。

与国外相比, 我国的相关量表评价工作和研究一直落后, 有些是沿用、翻译或修改国外量表, 有些是自制评价量表。国内的评价量表不论是翻译、修改或是自制, 基本都没有遵照严格的量表制作过程进行, 也没有进行信度和效度的评价, 有些类似国外 1993 年以前的工作。汪培山的量表采用了 Delphi 法和小组讨论相结合的方法, 并对量表进行了信度和效度的考核<sup>[5]</sup>。我们自查清单的项目筛选也采用了 Delphi 法<sup>[2]</sup>, 量表项目的筛选来自文献评价的应用实践, 并进行了不同评价者间评价一致性的考核与区分效度的考核<sup>[3]</sup>。从所得结果来看, 汪培山的 19 项量表与我们的 27 项量表大体是一致的, 我们的量表包含了汪培山的量表。汪培山等还同时应用 CONSORT 声明进行了评价, 把自身量表结果与 CONSORT 评价结果进行了相关性分析, 结果 Kendall 相关系数为 0.723。由于我们量表包含了汪培山的量表, 相信我们的 27 项量表与 CONSORT 的相关性也是很强的。所不

同的是,我们初步确定的条目较多(达125项,汪培山的最初条目仅为31项),联系到的专家数较多(全国范围56人,汪的为10人),重测信度的评价人数较多(5人,汪的为2人),应用的文献较多(预调查9篇、扩大应用45篇、信度效度评价54篇共90篇,汪的文献为30篇)。总的来说,我们的量表应该更加全面、可靠。

同时,为了提高评价的效率,也为了适用于不同背景的需要,我们还制作了本评价简表,共10项。本简表的遵循了最少项目最大区分原则和核心项目保留原则,同时满足了评价区分的需要和量表导向的功用需要。与评价量表全表一样,本简表适用的文献对象为有对照的临床研究,适用评价者为期刊管理者对医学期刊科学质量的评价及期刊编辑对医学论文统计学报告水平的评价。由于本简表仅含10个项目,因此比较适用于快速和初步评价,不宜用于全面评价。

到目前为止,国内有关医学文献质量评价的量表仅见汪培山的量表<sup>[5]</sup>和我们的量表<sup>[3]</sup>,且都适用于临床研究,说明我国在这方面的研究仍然是非常初步的。我们认为,进一步的工作可以往更全面、更权威、更规范的方向深入,也可以往专项评价量表方向进行<sup>[10]</sup>。

#### 参考文献

1 刘清海,方积乾. 国内外医学论文统计学报告质量的比较研究.

- 中国科技期刊研究,2008,19(2):236-239
- 2 刘清海,方积乾. 应重视医学临床试验论文统计学问题——统计学报告项目自查清单的研制. 中华医学杂志,2007,87(34):2446-2448
- 3 刘清海,方积乾. 医学论文统计学报告水平评价量表的研制及其意义. 编辑学报,2008,20(3):278-280
- 4 杨扬,沈志超. 循证医学和医学论文中统计学问题编辑鉴审的必要性. 编辑学报,2001,13(2):77-78
- 5 汪培山,周登远,贾真琳等. 临床试验文献质量评价量表的制作和评价. 药物流行病学杂志,2007,16(1):46-50
- 6 周登远,汪培山. 临床对照试验文献质量评价的发展与现状. 药物流行病学杂志,2004,13(4):211-214
- 7 Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reporters of randomized clinical trials: is blinding necessary? *Controll Clin Trials*, 1996,17:1-12
- 8 张琰,刘慧林,王麟鹏. 针刺预防性治疗偏头痛的文献质量评价. 中国循证医学杂志,2008,8(6):461-465
- 9 Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*, 1996,276:637-639
- 10 刘清海. 医学论文科学水平的评价探讨. 中华医学科管理杂志,2008,21(2):82-85

## 温州医学院 1998 ~ 2007 年中文论文产出统计与分析

毛文明 徐晓泉 郑俊海\*

收稿日期:2008-10-26

修回日期:2009-03-16

温州医学院期刊研究所,325035 浙江温州市茶山高教园区, E-mail: mwm@wzmc.net

**摘要** 参考文献计量学方法,在《中国期刊全文数据库》、《中国引文数据库》内检索温州医学院著者在国内期刊发表的中文论文及被引用情况。结果表明,1998~2007年温州医学院发表中文论文总数为15383篇,人均4.31篇,其中,在核心期刊论文占发文总数的24.2%,发文数呈逐年增长;论文分布在1156种期刊,平均每刊13.3篇;医学类期刊占73.8%,社会科学类占26.2%;发文最多的著者为78篇,发文7篇以上的著者占著者总数的17.5%,其发文量接近温州医学院总发文量的50%。被引论文数为6451篇,总被引频次为21725。我们认为,近十年温州医学院产文数增长较快,反映学院科技水平不断发展。

**关键词** 中文论文 被引频次 文献计量学 温州医学院

科技论文是科技成果的重要表现形式,是科技人员交流科技成果和心得、获取科技领域最新动态的最佳手段,也是衡量一个单位科技实力的重要指标。为了解温州医学院(以下简称温医)近十年来科技论文发表情况及其在国内学术界

的影响力,本研究以中国知网为平台,利用其《中国期刊全文数据库》和《中国引文数据库》对温医1998~2007年在国内期刊发表中文论文及被引用情况进行统计(英文论文的发表情况已另文统计<sup>[1]</sup>),结果报告如下。

\* 通讯作者:郑俊海,男,编审, E-mail: zjh@wzmc.net。