

# 临床医学论文中常见的统计学问题及对策\*

罗承丽<sup>1)</sup> 冷怀明<sup>2)</sup> 李海鸥<sup>3)</sup> 赵彬<sup>1)</sup> 姜军<sup>1)</sup>

(1)第三军医大学西南医院《中华乳腺病杂志(电子版)》编辑部;2)第三军医大学科研部学报编辑部;  
3)第三军医大学《免疫学杂志》编辑部:400038,重庆)

**摘要** 分析临床医学论文中常见的统计学问题,如组间可比性差,误用卡方检验,统计数据错误,临床诊断试验类论文统计表不规范,生存分析随访率低等。认为:编辑在初审和加工稿件时应按照实验设计的随机、对照、均衡、重复等4原则审查论文;对于特殊类型的研究论文,还应根据其特点进行审查。提出了审查临床医学论文的注意事项,以期能有助于编辑作好临床医学论文的质量控制。

**关键词** 临床医学论文;统计学;质量控制

**Common problems and countermeasures of statistics in clinical papers**// LUO Chengli, LENG Huaiming, LI Haiou, ZHAO Bin, JIANG Jun

**Abstract** The authors analyzed common problems of statistics in clinical papers, such as poor comparability between groups, misuse of chi-square test, statistical data errors, non-standardization/mistakes in tables of papers on clinical diagnoses, low follow-up rate of survival analysis etc, and proposed that editors should do the initial evaluation and editing of manuscripts based on the four principles of trial design: randomization, control, balance and replication. For special research papers the evaluation or review should be done according to their characteristics, and the authors have given concrete suggestions in order to help editors achieve the quality control of clinical papers.

**Key words** clinical paper; statistics; quality control

**First-author's address** Editorial Department of Chinese Journal of Breast Disease (Electronic Edition), Southwest Hospital, Third Military Medical University, 400038, Chongqing, China

临床医学论文是临床医学工作者的科研成果及临床实践经验的总结,对疾病的预防、诊疗和康复有一定的参考价值,并且有利于临床新技术、新方法、新成果的推广;但是,如果论文中存在某些学术方面的问题或科学性方面的差错,则会误导读者,影响临床诊疗技术的传播、交流与进步。目前,由于同行专家的审稿质量参差不齐<sup>[1]</sup>,临床医学期刊发表的论文仍存在诸多问题<sup>[2-5]</sup>;中华医学会系列杂志2008年质量审读报告显示,临床医学论文的统计学问题仍比较普遍<sup>[6]</sup>;因此,分析临床医学论文存在的常见统计学问题并探讨其对策,对编辑提高统计学审查能力,从而提高临床医学期刊论文质量具有重要的意义。

笔者结合自己的编辑工作实践,分析临床医学论文中常见的统计学问题以及识别统计学差错的注意事项,以期能有助于编辑同人作好临床医学论文的质量控制。

## 1 临床医学论文的常见问题

**1.1 组间可比性差** 组间可比性是指对比组之间除对比因素(药物或处理)外,其他特征基本相同,即所谓组间均衡。如组间均衡、组间的差异有统计学意义,则可认为结果(如疗效等)是由于对比因素的作用;如组间均衡性差,任何有统计学意义的组间差异均不能说明是对比因素的作用,因为其他因素也可能造成组间差异。

临床医学论文中常见的问题有如下2种。

1)组间均衡性差。处理组和对照组间年龄、性别、病程、病情、术式等非处理因素的分布不均匀。

2)将自己的试验资料与文献或其他资料进行对比。由于比较资料之间的基本条件,如患者的年龄、性别、入院时间的早晚、病情、病程、营养状况及合并症,诊断标准、疗效判断标准、观察指标及方法,使用的仪器、试剂、条件等均可能影响比较结果,此种对比必然会造成选择性偏倚。

**1.2 误用卡方检验** 卡方检验是一种常用于定性资料的假设检验,主要用于2个或多个样本率或构成比的比较。普通卡方检验的应用有一定的前提条件,要求定性资料的样本数 $n > 40$ 例且所有理论频数 $T > 5$ 。如果 $n \geq 40$ 例且有 $1 \leq T < 5$ 时应使用校正的卡方检验,如果 $n < 40$ 或有 $T < 1$ 时则应使用Fisher确切概率检验<sup>[7]</sup><sup>162</sup>。

临床医学论文中常见的问题有如下2种:

1)该用校正的卡方检验和Fisher确切概率检验时,不考虑卡方检验的适用条件,一律盲目套用卡方检验。

2)误用卡方检验代替Spearman秩相关分析。

**1.3 统计数据差错** 正确的统计数据是得出可靠结论的前提,然而,在临床医学论文数据核查的过程中,笔者发现论文中常存在数据差错。差错类型有如下几种。

\* 中国高校自然科学学报研究会基金项目(GBJXC0932);重庆市高校学报研究会科研项目(XB200916)

1) 数据抄写差错。例如,在一篇题为《Survivin 的表达与 TE 方案新辅助化疗疗效相关性研究》的稿件中,作者误将 SPSS 输出结果中的统计量 0.030 看成  $P$  值,因而得出  $P < 0.050$ ,Survivin 的表达与疗效有关的错误结论。笔者复核数据时发现有问题,便与作者联系重新进行统计学分析,然而,重新计算后  $P > 0.05$ ,研究结果与国内外相关研究不一致,最后作者要求撤回稿件,重新整理数据。

2) 摘要和正文中描述的数据与表或图中列出的数据不一致。

3) 同一数据在摘要和正文前后不一致。

4) 同一数据的小数位在摘要、正文前后及图表内均不一致。

5) 各分项病例数之和不等于总病例数。

6) 表中横向栏目的数据之和正确,而纵向栏目的数据却错误。

7) 文中列出的统计表与实际输入软件进行分析的列联表不同,导致编辑复核的数据与文中列出的不符。统计数据差错在临床回顾性研究论文中尤为常见。

**1.4 临床诊断试验问题** 临床上对某种新的诊断试验进行评价时,必须以“金标准”作为新诊断是否正确的标准。

所谓金标准,是指当前被医学界同行公认的标准诊断方法,是区分“有病”与“无病”的依据,包括病理学检查、外科手术发现、特殊影像诊断、临床综合诊断标准及长期临床随访所获得的肯定诊断<sup>[8]</sup>。临床诊断试验实用价值的评价指标包括敏感度、特异度等。这些指标的计算需要严格按照特定的统计表<sup>[8]</sup>进行,但在临床诊断试验论文原稿中常发现统计表制作不规范,诊断试验检测的阳性和阴性结果位置互换,金标准诊断结果与诊断试验结果位置互换等问题。这既不利于阅读理解,又易导致临床诊断试验评价指标如敏感度、特异度等的计算错误。

**1.5 生存分析存在的问题** 在临床医学研究中,有时需要对患者的生存情况作出评价,例如肿瘤患者接受某种治疗后的生存情况等。此类研究需要界定随访的开始时间和研究终止时间。由于失访、意外死亡、试验到期等原因,部分研究对象不能随访到确切的存活时间,所以无法以生存时间直接进行统计分析,而需要应用一些特殊的统计方法。

目前,常用 Kaplan-Meier 法计算生存函数,用 log-rank 检验比较几组生存时间,用 Cox 比例风险模型探讨协变量对生存时间的影响<sup>[9]</sup>。生存分析类论文常见的问题有如下几种:

1) 随访率较低<sup>[10]</sup>,降低了患者生存时间在统计学上的真值代表性;

2) 只描述生存时间计算的起点和研究的终点,未说明随访率和中位生存时间;

3) 未说明截尾数据(失访、死于其他原因、随访截止时仍存活病例)的情况;

4) 用 log-rank 检验比较几组生存时间时,未在生存曲线图的注解中描述卡方值和  $P$  值。

## 2 对策

针对临床医学论文的常见错误,编辑在初审和加工稿件时应注意按照实验设计的随机、对照、均衡、重复 4 原则审查论文。对于特殊类型的研究论文,还应根据其特点进行审查。

**2.1 组间可比性审查** 在组间可比性方面,应审查对照组是否符合以下 3 个条件:

1) 非处理因素对等。除处理因素外,对照组的非处理因素应与实验组相同。

2) 时间和空间同步。在整个研究进程中对照组和实验组始终处于同一时间和空间。

3) 专设对照组。任何一个对照组都是为相应的实验组专门设立的,不能借用文献或其他资料与本研究作对照。对于根据临床资料作回顾性总结的论文,如描述“全部病例被随机分配到治疗组和对照组”,则应删除“随机”2 字,并且要求作者给出组间各种重要非处理因素的详细资料和数据、非处理因素的统计学检验结果( $P$  值均应  $> 0.05$ ),以便说明组间具有可比性。对于前瞻性临床医学研究论文,应要求作者详细说明如何随机抽取研究对象,以及应用何种方法将全部研究对象真正随机分配到试验组和对照组。

**2.2 卡方检验的审查** 应注意检查如下内容:

1) 样本数和理论频数的大小,是否该用校正卡方检验或 Fisher 确切概率检验。

2) 对统计学检验结果的解释是否前后一致,与图表内列出的结果一致。如果统计表中列出的是  $\chi^2$  统计量,而在结果部分却描述为“相关”,或结果中描述“差异有统计学意义”却在讨论中描述为“具有相关性”,则应请作者明确研究目的是想了解组间结果变量的“差异性”还是变量间的“相关性”,并选用合适的检验方法。

**2.3 核对数据** 注意核查论文中同一数据是否前后一致,与图表内一致;各分项病例数之和是否等于总例数;同一数据的小数位在摘要、正文和图表内是否一致。

对于定性资料的统计表,尤其要注意核对横向栏目数据之和与纵向栏目数据之和是否正确。对于根据

公式计算的指标,应注意将统计表中的数据代入公式核实。

对于定量资料,由于无原始数据,编辑不便核实;但如为定性资料,统计表中一般列出了原始数据,编辑可以根据作者在“材料与方法”中介绍的统计学方法,用SPSS统计软件复核统计量和 $P$ 值。如发现问题可及时与作者联系。另外,编辑还可在退修稿件时要求作者将SPSS的原始统计结果连同修改稿一并寄回,以便复核统计数据。

**2.4 临床诊断试验类论文** 对这一类论文,应注意审查如下内容:

1)新的诊断试验与金标准比较时,应审查统计表制作是否标准,数据是否有误,是否计算敏感性、特异性、准确性、阳性似然比、阴性似然比、Youden指数、阳性预测值、阴性预测值。需注意,阳性预测值和阴性预测值受患病率的影响,不能用于评价肿瘤标志物的诊断价值<sup>[12]</sup>。

2)如新的诊断试验是与非金标准方法比较,则应作一致性(Kappa)比较评价。此时需注意:Kappa  $\geq 0.75$ 表明2种检验方法的一致性较好,0.75 > Kappa  $\geq 0.4$ 表明一致性一般,Kappa < 0.4则表明一致性较差<sup>[13]</sup>。

**2.5 生存分析类论文** 对这一类论文,应注意审查如下内容:

1)是否描述生存时间计算的起点、生存的结局(疾病的复发或死亡等)和研究的终止日期;

2)随访率如何(一般要求随访率在95%以上,不能低于90%<sup>[10]</sup>);

3)是否描述患者的中位生存期;

4)截尾(删失)数据的情况及其处理(需注意,不论删失数据的产生原因为何,删失生存时间的计算均为规定的起点至删失点<sup>[7]400</sup>);

5)几组生存时间比较时,生存曲线图的注解中是否描述卡方值和 $P$ 值;

6)进行log-rank检验时,是否满足除比较因素外,影响生存率的各混杂因素组间均衡可比<sup>[7]408</sup>;

7)用Cox比例风险模型分析时,统计表是否包括变量、回归系数( $\beta$ 值)、标准误、统计量(如 $\chi^2$ 值)、统计学意义( $P$ 值)、相对危险度( $R_R$ )以及95%的可信区间;

8)生存分析方法的使用是否正确。

### 3 结语

提高临床医学论文的统计学质量任重而道远。长期以来,编辑都过分依赖审稿专家,而具备审稿资格的专家常常身兼数职,很难有更多的精力详细审阅稿件,致使有些质量不高的稿件得以发表,影响了临床医学期刊的声誉<sup>[2-4]</sup>;因此,除了请统计学专家作医学统计学讲座,以及不断发掘年富力强、对期刊工作有热情的审稿专家外,临床医学编辑自身也要加强统计学知识的学习,努力提高对统计学问题的识别能力,并制订出临床医学论文的质量控制标准,以确保临床医学期刊发表的论文更加科学而严谨。

### 4 参考文献

- [1] 苏振伦.从《中华耳科学杂志》82份外审意见量化分析探讨同行评议的优势与不足[J].中国科技期刊研究,2009,20(1):149-151
- [2] 李廷谦,王刚,王蕾,等.我国中医药临床研究的现状和评价[J].中国循证医学杂志,2005,5(6):431-437
- [3] 胡良平.我国临床科研的现状与思考[N].中国医学论坛报,2006-02-16(19)
- [4] 顾琴龙.我国临床科研中存在的问题与思考[J].外科理论与实践,2007,12(5):407-408
- [5] 杨莉丽,马擘,黄新文,等.论文摘要的数据准确性问题[J].编辑学报,2008,20(1):25-26
- [6] 中华医学会系列杂志社总编室.中华医学会系列杂志2008年质量审读报告[J].杂志工作通信,2008(11):3-13
- [7] 方积乾,孙振球.卫生统计学[M].6版.北京:人民卫生出版社,2008
- [8] 熊鸿燕,易东.医学科研方法:设计、测量与评价[M].重庆:西南师范大学出版社,2005:42
- [9] 方积乾,宇传华.第十三讲 如何处理随访资料[J].中华预防医学杂志,2003,37(1):63-65
- [10] 董秀玥.医学期刊中的生存分析及存在的问题[J].数理医学杂志,2003,16(5):421-422
- [11] 胡良平.统计学三型理论在统计表达与描述中的应用[M].北京:人民军医出版社,2008:290
- [12] 高云朝.国内肿瘤标志物诊断性试验论文存在的问题[J].放射免疫学杂志,2006,19(6):512-514
- [13] 张文彤.SPSS11统计学分析教程:基础篇[M].北京:北京希望电子出版社,2002:263

(2009-05-07 收稿;2009-07-02 修回)