

医学科技论文常见的统计学误用问题及其商榷

张振东^{1*}, 孙丽君¹, 孙竹平²

(1. 遵义医学院附属医院妇科, 遵义 563003; 2. 山东大学齐鲁医院《现代妇产科进展》编辑部)

医学统计学在现代医学研究中已得到广泛的应用,从课题的设计实施、数据采集、整理分析到结果表述、论文撰写等环节,均涉及到统计学的问题。一般来说,医学统计学方法分为基本统计方法与高级统计方法。基本统计方法中常见有 t 检验、卡方检验、常用方差分析、秩和检验、简单线性回归等;高级统计方法常用到的有析因分析、重复测量设计资料的方差分析、Logistic 回归分析、生存分析、判别分析、主成分分析、Meta 分析等。我们结合《现代妇产科进展》2010 年各期审校过程中发现的几个最为常见的统计学误用问题,探讨如何合理准确应用统计学,以期让统计学更好的为医学研究服务。

1 统计学应用中存在的常见问题

1.1 单因素方差分析(ANOVA)两两比较误用独立样本 t 检验 单因素方差分析设计 3 组以上的均数比较,如果总体比较有差异,需进行两两比较,一般用 SNK 法或 LSD 法。但部分研究者却将资料进行拆分,应用独立样本 t 检验进行两两比较,导致第 I 类统计学错误发生率(假阳性率)增加,从而掉进了一个常见的“统计陷阱”,使所得结论可信度大大降低甚至得出错误结论。SNK 法与 LSD 法虽然并非等价,实质是一致的。SNK 法一般用于经方差分析结果具有统计学意义时才决定进行的两两事后比较,而 LSD 法可用于方差分析不足以具有统计学意义时也能进行两两比较^[1]。比较两种方法在 SPSS 的输出结果形式,SNK 是“分堆”比较,一目了然,对于组别数较多的研究更为好用,但没有具体 P 值,而 LSD 是在进行“两两”比较时,能给出具体的 P 值。

1.2 两两比较时检验水准的重新调定 χ^2 检验或秩和检验 3 组以上整体比较有差异时,需应用分割法进行两两比较,这时检验水准应由原 0.05 调定为 0.0167,否则会增加第 I 类统计学错误的发生率。特别当 P 值处于 0.0167 ~ 0.05 时,按照 $P < 0.0167$

的标准,差异无统计学意义,而按照 $P < 0.05$ 的标准,却有意义,与事实相悖,出现假阳性,很容易得出错误结论。这种分割法有时很保守,当行列表资料分组多且为有序时可用 Mantel-Haenszel 卡方检验,也称线性趋势检验(test for linear trend)或定序检验(Linear-by-Linear test)^[2]。统计路径:用 SPSS 进行计数资料的趋势检验,在输出结果中读取线性关联检验统计量(Linear-by-Linear Association, LLA),如 $P < 0.05$ 可得出随着病种级别的升高,检测指标逐渐升高的趋势。

1.3 临床诊断试验中的统计学方法应用 在临床诊断试验研究中,经常选取单项计量指标或者联合计量指标以诊断某种疾病,若仅用初级统计学方法如 t 检验、单因素方差分析等往往不能有效挖掘信息,此时应采用受试者工作特征曲线(ROC)对检测结果进行分析评价。ROC 曲线分析基本原理是通过诊断界点的移动^[3],获得多对灵敏度和误诊率(1-特异度),以灵敏度为纵轴、误诊率为横轴,连接各点绘制曲线,然后计算曲线下的面积,面积越大诊断价值越高。ROC 曲线很直观,能根据敏感性与特异性之和最大化原则自动产生最有效的诊断临界点。具体路径可以参考相关统计专著^[3]。统计学处理一般描述为:采用 SPSS(版次)统计软件分析数据,对单项及联合检测结果作图绘成 ROC 曲线,计算曲线下面积(AUC)和标准误,其中联合检测结果变量即预测概率由 Logistic 回归产生(也可以用判别分析得出)。计量资料应用 $\bar{x} \pm s$ 表示,运用独立样本 t 检验及单因素方差分析,两两比较采用 SNK 及 LSD 法,计数资料采用 χ^2 检验。检验水准为 0.05。具体内容可据情而定。

1.4 重复测量资料的方差分析误用拆分文件的 t 检验或方差分析 如研究共设 3 组,每位患者在 3 个时间点均查某项血指标,部分作者在处理此类数据时,常误将纵向(同一时间点 3 组的比较)与横向

* 通讯作者 张振东(1976-),本刊特邀统计学编辑,遵义医学院附属医院妇科主治医师。Email:zhangzdx@126.com

(同组 3 个时间点的比较)数据均应用拆分文件的 t 检验或单因素方差分析来处理,结果导致统计学第 I 类错误发生。此组数据实质是重复测量资料,应采用重复测量资料的方差分析。SPSS 中的统计路径:数据-分析-一般线性模型-重复度量。研究者可以参考相关书籍进行处理^[3]。

1.5 2×2 析因设计及析因方差分析 实验是 2×2 析因设计时,分组有两个因素,A 与 B,故分组为 A、B、O、A+B,这个设计在析因设计研究中很常用,但常会出现分组设计正确,却没有用析因设计方差分析。析因设计与单因素方差分析不同^[4],它不但能分析治疗效果中处理因素的单独效应和主效应,还能分析因素间的交互效应,并能提高检验效能。非统计专业的研究者进行析因分析可能稍有难度,可参考相关统计学书籍提供的统计步骤进行此类分析^[3]。

1.6 Meta 分析 Meta 分析是循证医学系统评价常用的方法^[5],应用时需注意统计学处理中计数资料采用比值比(OR)作为效应变量。具体路径:先进行异质性检验,当 $P>0.05$ 时,认为同质,选择固定模型; $P\leq 0.05$ 时,不同质,此时可采用敏感性分析或分层分析等异质性处理,使之达到同质后再选择固定模型;若采用异质性处理仍未达到同质,则采用随机模型,以上统计路径均需交代清楚。Meta 分析的结果是以“森林树”体现的,审校中我们经常遇到作者绘制的“森林树”左上角“文献、对比、结果名称”等内容显示为“?”,这是由于部分版本的 Rev Man 软件不能输入中文,此时可以考虑省去,或用 Photoshop 软件添加相应中文。Meta 分析作为一种高级统计方法,专业性要求较高,作者可参考循证医学类权威杂志上的文章格式,如《中国循证医学杂志》中“论著·二次研究”栏目的循证文章。

2 科技论文中统计学处理的相关表述

2.1 资料与方法中具体统计路径的描述 “统计学处理”的内容常位于论文资料与方法的最后一段,一般来说包括统计软件名称及版次、统计描述、统计方法、检验标准等内容,亦可细致交待每个表格的具体统计方法。经典例子如下,“统计学处理:采用 SPSS(版次)统计软件分析数据。计量资料用均数±标准差表示,采用单因素方差分析,两两比较采用 SNK 法及 LSD 法。检验水准为 0.05”。上述内容包括了大致的统计方法,即具体的统计路径。此部分内容,没有绝对统一的规定^[6]。常见的问题有:统计学方法描述不全、内容过于简单、存在粘贴

抄写痕迹等。如部分论文的统计学处理中提及“以 $\alpha=0.05$ 为检验水准, $P<0.05$ 为差异有统计学意义”这句话,这在统计学上实质是一个重复句,保留其一即可。

2.2 结果中具体 P 值的标注 现在的统计学处理手工计算的较少,一般均应用统计软件,最常用的软件如 SPSS、SAS 均能给出具体 P 值。但部分论文的结果表述中却未标明具体 P 值,作为科技论文是不够严谨的,建议作者在表述研究结果时注明具体 P 值,增加论文可信度的同时,可用于再次分析。

2.3 表格制作 统计表设计需规范,应体现统计设计内容。部分表格存在内容割裂、组别名称违反表格简洁化原则、表下注释繁琐、横标目与纵标目颠倒等情况。建议作者写作论文时参阅相关统计学教程及杂志稿约。

医学统计学教学在中国的医学教育特别是高层次教育中举足轻重。一项课题从一个好的创意开始到实施、结题,统计学方法始终贯彻其中。不管是临床试验研究、实验性研究、观察性研究,还是临床测量误差与诊断试验,如果没有规范性的统计设计、精确的统计分析、科学的统计学解释与表述,很难获得学术上的认可^[1]。目前在统计学的具体应用中,研究课题开始设计时就有统计学专家参与已成为较流行的趋势。许多专业统计学专家或研究者在不停的进行“统计基础”的研究,以创造更多的数理统计方法。对于非统计专业的研究者来说,统计学的学习主要侧重于应用。在应用统计学教学中,一般分为 4 个档次或阶段来进行,首先通过学习医学统计学课本获取统计学思想、原理和方法,其次学习统计软件相关书籍掌握数据到统计结果的转化,再学习统计表达与描述书籍达到书写统计学报告能力,最后可以学习如何出具统计审阅报告。医学科技工作者亦可按照这个顺序学习以提高统计能力。

参 考 文 献

- [1] 颜虹. 医学统计学[M]. 北京:人民卫生出版社,2010:135-137
- [2] Norman GR, Streiner DL. 生物统计学基础[M]. 凌莉,译. 3 版. 北京:人民卫生出版社,2010:237
- [3] 陈平雁. SPSS13.0 统计软件应用教程[M]. 北京:人民卫生出版社,2006:143-147,160-169,322-326
- [4] 孙振球. 医学统计学[M]. 3 版. 北京:人民卫生出版社,2010:194-195
- [5] 王家良. 循证医学[M]. 北京:人民卫生出版社,2006:81-92
- [6] 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,2008:57-62

(收稿日期 2010-12-30)