

常见药理学数据处理错误的分析

袁秉祥 (西安交通大学医学院药理学教研室, 西安 710061)

中图分类号: R965 文献标识码: A 文章编号: 1005-9903(2002)01-0059-03

药理学研究数据统计是建立在设计合理, 客观观察, 资料完整, 记录准确的基础上, 是分析样本, 推论总体, 透过偶然, 找出规律的过程。统计分析的主要任务是进行显著性检验, 而显著性检验是分析无效假设的概率(P), 因此, 经统计所得出的结论是概率性的, 不是绝对的肯定和否定^[1]。熟悉统计学的基本概念, 正确分析数据性质, 就能恰当地选择统计方法, 进行正确分析, 得出正确结论。统计软件也无法帮助你进行数据性质的分析和选择正确的统计方法。很多数据处理错误是由实验设计不当引起的, 而实验设计不当又是因为缺乏数据处理技能造成的, 因此正确掌握数据处理方法可帮助进行合理的实验设计。现就药理学论文常见数据处理失误进行了分类, 并对各类失误提出如下建议。

1 统计结论的描述性错误非常普遍

1.1 $P < 0.05$ 或 $P < 0.01$ 的统计结论描述 当显著性检验所得出的概率是 $P < 0.05$ 或 $P < 0.01$ 时, 结论不应是差别显著或非常显著, 而应是差别有显著意义或有非常显著意义。然后根据专业知识判断差别是否显著或非常显著。例如: 某降压药使原发性高血压患者的平均血压降低 $0.507 \pm 0.293 \text{ kPa}$ ($3.8 \pm 2.2 \text{ mmHg}$, $P < 0.01$) 时, 只能说使血压降低有非常显著的意义, 根据临床专业知识, 不能说降压作用非常显著, 可见统计结论不能替代专业结论。

1.2 $P > 0.05$ 的统计结论描述 结论应是差异无显著性意义, 不应得出无差别或相等的结论, 实际上是得到一个没有结论的结论。因为 $P > 0.05$ 只说明无效假设的可能性较大, 不能说明两组数据等效。为了判断数据间是否无差异或相等, 应该进行生物等效性分析, 不能说明两组数据等效。生物等效性分析可用等效界值判断, 规定采用双向单侧 t 值法, 判断标准可定为 $0.90 \sim 1.10$ (药效学) 或 $0.80 \sim 1.20$ (药动学) 范围内。由于 AUC 和 C_{max} 呈偏态分布, 应进行对数转换, 生物利用度的等效性分析的判断标准为 $0.8 \sim 1.25$ ^[2,3]。我国新药研究已经要求在临床药效学和药动学研究方面进行生物等效性分析。

2 随机分组不均衡 在对实验动物进行随机分组时, 应按动物的体重、性别、状态等影响观察指标的因素进行均衡随机分组。如果分组时未注意对照组和实验组的均衡关系, 那么, 观察指标在给药前, 对照组与给药组就会出现差异, 使给

药后的实验数据可比性减小, 统计学意义下降。如果实验确实难以均衡分组, 不应以原始数据进行统计处理, 而应求出给药前后变化差值, 或最好求出给药前后变化率以后, 再进行统计处理^[4], 这样可有效地消除不均衡分组因素, 但在表格中必需给出给药前的数据, 以表明数据的均衡性。

变化率 = (给药后数据 - 给药前数据) / 给药前

3 非参数不能用参数统计方法 对计量资料的参数数据经常用 t 检验或多因素方差分析 (F 检验), 但也经常出现把非参数计量资料的数据用参数统计方法进行处理。时反应数据经常出现正向偏态分布; 微量元素 (如血铅、血锌) 等数据经常出现负向偏态分布。当数据呈偏态分布时, 不能用参数统计方法 (t 检验或 F 检验), 而应该用非参数统计方法 (参比差值法或秩和检验) 或进行数据转换 (倒数或对数) 后用参数统计方法。建议在出现时反应数据和微量元素数据时, 应在文章中注明常态分布判断和数据转化过程。

4 配对关系 在进行配对 t 检验和有配对关系的两率对比时, 要特别注意其适用范围。配对检验只适应于同体对比 (同一观察对象的两种观察方法) 和前后对比 (用药前后和两疗程对比), 并有充分理由说明其可比性。配对对比 (一一配对) 由于其必须有严格的配对关系, 所以, 不易轻易使用。有配对关系的两均数对比用配对 t 检验; 有配对关系的两率的对比用关联 χ^2 法和优势 χ^2 法处理。时序性资料存在给药前后多组数据, 需要一一计算出其差值或变化率^[4], 由于不属于配对资料, 不能用配对 t 检验, 应将其差值或变化率进行非配对 t 检验。

5 应判断计量资料方差是否整齐和计数资料中样本是否足够大 在进行两均数对比时, 要事先判断其方差是否整齐。如果方差不整齐, 不能简单直接用 t 检验, 可用非参数统计方法或校正其自由度后用 t 检验 (t' 检验)。在进行两率对比时要事先判断资料中样本数是否足够大, 或是否含有“0”、“1”或“2”, 如 50 例病人, 48 例有效, 2 例无效。此时, 则不能用 χ^2 法进行检验, 而应该用较繁琐的 Fisher 氏直接机率法, 或用简化机率法进行检验^[1]。

6 有顺序关系的计数资料 过去经常出现用 $\chi^2 (2 \times 2)$ 法处理的错误, 现在此类错误已经少见。某些临床药效学研究经常出现无效、有效、显效、治愈等指标, 基础研究也经常出现此类有顺序关系的半定量指标, 如 -、+、#、## 等。有顺序关系的计数资料应该采用 Ridit 法或序值法检验, 不宜用 χ^2

(2×2)法进行单项指标比较。临床研究最后计算疗效(有效率)时,只应把显效和治愈计算在内,不必把有效项计算在内^[5]。

7 有效数字的取舍 生物学实验的误差大小,取决于个体差异、抽样误差、条件误差、系统误差和过失误差,误差又决定着有效数字的取舍。一般说来,生物学实验的均数(\bar{x})的有效数字保留 3 位,标准差与均数的精确度应保持一致。有效数字的多少应由实验精确度和标准差的大小来决定。精确度越小,标准差越大,有效数字应该越少。有些刊物规定用 s 的三分之一来确定 $\bar{x} \pm s$ 的有效数字,然后再增加一位估计有效数字,例如 28.37 ± 1.56 的有效数字为 28.4 ± 1.6 ,增加一位估计有效数字,可写成 28.37 ± 1.56 ,如果标准差增大到 8.78,则应写为 28.4 ± 8.8 。

8 量效关系的判断 剂量依赖性 or 量效关系的判断不宜根据均数大小直接判断,或用一般的显著性检验进行判断,而应把数据对打成散点,用线性回归的 t_r 检验方法^[6]。不能根据实验数据就直接得出药物或疗法的最大效应(E_{max}), E_{max} 是用直线回归方法推导出来的。

9 选择合适的对比组 药效学实验中常设正常组、模型组、阳性药组、实验组大、中、小剂量组,一般常以模型组为对比组,这样只用一种标志就可标记出各组的显著性。不要以正常组为对比组,更不应以此与各给药组进行比较。建议以实验组小剂量为对比组。这样还可以同时比较剂量间的差异和与阳性药的差异。

10 剂量呼应关系 人与动物间或不同动物间由于体表面积不同,用药时存在着一定的剂量呼应关系。以 g/kg 或 mg/kg 表示剂量,不同动物间的剂量应不相同。不同动物用同一剂量是错误的。剂量呼应关系还表现在毒理学实验和药效学实验之间,例如药效学实验的大剂量不能小于长期毒性实验的小剂量。人和动物及不同动物间的剂量量呼应关系可粗略估计如下(/kg):

人: 犬 : 猴猫兔: 豚鼠: 大鼠: 小鼠
1 : 2~ 3: 3~ 5 : 5~ 7: 7~ 9: 9~ 11

11 $\chi^2(R \times C)$ 检验后和 F 检验后的两两对比 在进行药效学研究时,多行多列的 $\chi^2(R \times C)$ 检验后和 F 检验后,很多作者忘记继续进行两两对比,使信息量缺失。 $\chi^2(R \times C)$ 法和 F 检验可分析出因素间趋势性的变异,不继续进行两两对比就无法特意分析某两组数据间的差异。卫生统计学认为多因素方差分析后,可用 q, q' 检验进行两两对比,但这个规则不适合于药理学研究资料。 q, q' 检验在进行两两对比时,不但要考虑参与对比的两因素,还要考虑其它因素,这对大样本的流行病学研究是正确的,对小样本的药理学研究是不合适的,因为,此时不存在与对比两因素有关的其它因素,所以,只要进行两两对比的 t 检验即可。

12 其他类型错误 多批资料简单用单批资料的方法处理:不同时间(季节)和地点(地区或实验室)协同完成的同一项实验属多批资料,应首先分析批间的变异,然后用析因 t 检验进行分析,不可将多批资料简单相加,一并处理。

案例报道(单率)缺少泊松分布或二项分布的可信限范围说明。

数据或资料不全 这类错误包括:配对资料中缺少原始数据;未说明统计方法;时序性资料缺少给药前原始数据;例数不够等失误。

注:数据性质和数据处理方法小结

12.1 计数资料^[4]

单率描述:泊松分布或二项分布法(可信限区间)
两率对比:

非配对 | 构成比同 | 样本数大(例数 ≥ 3): $\chi^2(2 \times 2)$ 法
 | | 样本数小(例数 < 3): 机率法
 | 构成比不同: χ^2 权重法

有配对关系: 关联 χ^2 法+ 优势 χ^2 法

多率对比 | 有等级顺序关系: 等级序值法 Ridit 法
 | | 两行多列: $\chi^2(2 \times K)$ 法
 | 无等级顺序关系 | 多行多列: $\chi^2(R \times C)$ 法
 | 两两对比: $\chi^2(2 \times 2)$ 法

12.2 计量资料^[5]

两均数对比:

常态分布 | 无配对关系 | 方差整齐 | 例数不同: t 值法
 | | | 例数相同: t 值简法
 | | 方差不齐: t' 检验
 | 有配对关系: 配对 t 值法

偏态分布: 秩和法或序值法

多均数对比:

多因素分析 | 二水平一级交互: $2 \times 2 F$ 检验
(F 检验) | 二水平二级交互: $2 \times 2 \times 2 F$ 检验
 | 拉丁方设计 F 检验
 | 两两对比: t, d', q, q' 检验
多批资料: 析因 t 值法

12.3 生物等效性检验 双向单侧 t 检验^[7]。

两均数对比 | 随机分组: 等效界值 = $(D \times Ms - T \times Se)$
 | 交叉配对 | $T_L = [(0.09691 + M_{\log(t/r)}) / Se$
 | $T_H = [(0.09691 + M_{\log(r/t)}) / Se$

两率对比:

$$u = \frac{\Delta - |p_T - p_C|}{\sqrt{p(1-p)(1/n_T + 1/n_C)}}$$

其中 $p = (r_T + r_C) / (n_T + n_C)$

参考文献:

[1] 袁秉祥,王清莲. 医师实用统计手册[M]. 西安: 陕西科技出版社, 1998. 21, 44.
[2] 桂常青, 郑青山, 孙卫民. 等. 生物利用度等效性检验的对数值直接算法[J]. 中国临床药理学与治疗学杂志, 1998, 3(3): 203-209.
[3] 袁秉祥, 朱芳. 生物等效性的计算和计算程序[J]. 西安医科大学学报, 2001, 22(2): 164-166.

[4] 程能能, 孙瑞元. 关于新药生物统计指导原则(草案)的说明[J]. 中国临床药理学与治疗学杂志, 1996, 1(2): 136-145.

[5] 刘玉秀, 洪立基. 新药临床研究与统计分析[M]. 南京: 南京大学出版社, 1999. 98.

[6] 王丽荣, 孙卫民, 孙瑞元. 药物量效相关性的显著性检验[J]. 中国临床药理学与治疗学杂志, 1996, 1(2): 146-148.

[7] 刘玉秀, 杨友春. 临床试验疗效的等效评价[J]. 中国临床药理学与治疗学杂志, 1999, 4(3): 220-226.