

体育类核心期刊中运动医学论文的统计学 应用错误案例分析

张宁

(宁夏大学 体育学院, 宁夏 银川 750021)

摘 要:就近年来体育类核心期刊运动医学科研论文撰写中常见的统计学应用误区进行分析, 结果发现: ①实验分组不合理, 受试对象的选取违背同质性原则; ②统计方法不合理, 多个样本均数比较, 错误应用多个 t 检验代替, 误用单因素方差分析处理多因素方差分析; ③误判资料类型, 从而误用定量资料统计分析方法, 两样本率的检验错误用两样本的 t 检验来代替。

关 键 词: 统计学; 运动医学论文; 体育类核心期刊

中图分类号: G80-32 文献标识码: A 文章编号: 1006-7116(2010)11-0115-04

Analysis of statistical application errors in sports medical theses in core journals in the sport category

ZHANG Ning

(Institute of Physical Education, Ningxia University, Yinchuan 750021, China)

Abstract: The author analyzed statistical application mistakes commonly seen in the writing of sports medical research theses in core journals in the sport category in recent years: 1) the division of experiment groups was irrational; the selection of testees violated the principle of homogeneity; 2) the comparison of the average of multiple samples was wrongly replaced with multiple t tests; single factor variance analysis was wrongly used for multiple factor variance analysis; statistical methods were irrational; 3) data type was wrongly judged, thus a wrong quantitative data statistical analysis method was used; the test of two sample rates was wrongly replaced with the t test of two samples.

Key words: statistics; sports medicine theses; core journals in the sport category

统计学是科学研究的基础, 是研究结果科学性、可靠性的有力保障。而论文是研究成果的表达形式, 通过论文可展示研究者的科研成果, 也便于查阅者的了解、评价和引用。因此, 在论文中要完整、准确地进行统计学内容的表述, 以展示研究设计的合理性和研究结果的可靠性。但运动医学论文中存在的统计学错误仍相当严重, 其中实验设计和统计资料的应用上最为突出。本文介绍论文中出现错误的部分, 并对其进行了正确的统计学内容表述, 以便供广大运动医学科研工作者参考。

1 实验设计方面的问题

1) 实验分组不合理。

例 1: 《有氧健身操结合饮食控制改善代谢综合征患者胰岛素抵抗、血脂、血糖的实验研究》^[1]一文中研究对象为: 西安市碑林区某社区代谢综合征患者 77 人年龄 40~65(55.35 ± 7.26)岁。按照接诊顺序将入选 MS 患者随机分为以下 2 组, 综合干预组和对照组。

该文作者所选取代谢综合征患者年龄跨度较大(40~65 岁), 因为年龄的不同机能的代谢水平有很大的差异性, 不能认为他们来自于同一个总体。把他们随机分入综合干预组和对照组参加试验, 两组之间的可比性可能很差, 其结论的可信度很低; 如果, 按年龄段分层随机化, 组间的均衡性要高得多。

2) 受试对象的选取违背同质性原则。

例 2: 《不同类型休闲活动对老年女性身体机能影

响的差异性研究》^[2]一文中对成都市老年大学、锦江区老年活动中心,随机选取离退休 5 年以上,年龄 65 ~ 75 岁之间老年女性为研究对象。并依据她们的年龄段将活动类型分为:为棋牌队、曲艺队、门球队、舞蹈队;测量指标为:血压及脉率,肺活量,椎体骨密度。

从一般的常识可知,经常参加体育锻炼的人有利于其血压及脉率,肺活量、椎体骨密度指标向好的方向发展。原作者所选取的活动类型为:棋牌、曲艺队、门球队、舞蹈。其中棋牌是一种智力游戏和身体运动没有什么关系,而曲艺,门球,舞蹈主要以身体运动为主(曲艺有利于人体的肺活量的提高)。所以原作者在试验设计时违背同质性原则^[3]。

在实验设计时一定要遵守 3 大原则:对照、随机、重复。对照原则:要确定处理因素对实验指标的影响,如无对照是不能说明问题的。实验分组时有处理组和对照组。对照原则要求处理组和对照组除处理因素以外的其他可能影响实验的因素应力求一致(即齐同比较或有可比性)。随机原则:是使每个实验对象在接受分组处理时具有相等的机会,以减少偏性,使各种因素对各组的影响保持一致(均衡性好),通过随机化可减少分组人为误差。这是对资料分析时进行统计推断的前提。重复原则:是指可靠的实验应能在相同条件下重复出来(重现性),这就要求实验要有一定的例数(重复数)。因此,重复的含义是重现性与重复数。重现性可用统计学中显著性检验的值来衡量其是否满意^[4]。

2 定量资料统计分析存在的统计学错误

2.1 多个样本均数比较错误应用多个 t 检验

(原)表 1 不同有氧运动时间对大鼠血清 LDL 水平的影响(mmol/L, $n=10$)($\bar{x} \pm s$)

| 组别 | 0 周 | 2 周 | 4 周 | 6 周 |
|------|-----------|-----------|-------------------------|-------------------------|
| 对照组 | 1.01±0.18 | 1.00±0.16 | 1.07±0.19 | 1.03±0.19 |
| 小运动组 | 1.06±0.20 | 0.95±0.17 | 0.85±0.15 | 0.72±0.18 ¹⁾ |
| 大运动组 | 1.02±0.19 | 1.03±0.22 | 1.13±1.17 ²⁾ | 1.18±0.21 ³⁾ |

与对照组相比, 1) $P<0.05$; 与小运动组相比, 2) $P<0.05$, 3) $P<0.01$, “原表·”表示原作者的表格, “表·”表示取原作者表格的部分

在应用 t 检验和方差分析时要注意它们之间的共性与区别。 t 检验和方差分析都是对总体(样本)均数的检验。当对两个以上的总体(样本)均数是否存在显著性差异进行检验时, 如果用 t 检验, 4 个总体(样本)均数则需做 $c^2=6$ 次可能组合的检验。若在 $\alpha=0.05$ 的显著性水平上检验, 则将降低统计结论效度。所以, 两总体(样本)均数的检验用 t 检验, 两个以上总体(样本)均数的检验用方差分析。

例 3:《有氧运动对大鼠血糖、血脂和血液凝固功能的影响》^[5]一文中, 将 30 只大鼠随机分为 3 组: 对照组、小运动训练组和大运动训练组。对照组大鼠进行自由饮食和自由饮水; 运动训练组动物在此基础上给予为期 6 周的游泳训练。有氧运动(游泳)、运动后 2 周和运动后 4 周从眼眶后静脉丛取血, 以及实验结束时(运动后 6 周)断头取血, 分别做血糖、血脂和有关凝血指标的测定。用 SPSS 统计软件对原始数据进行常规数理统计分析, 计量资料用均数 \pm 标准差描述, 显著性检验用 t 检验, $P<0.05$ 表示有统计学意义。

这类错误是定量资料分析中最常见的错误, 而且是原则性错误, 会增加犯第一类错误的概率^[6]。案例(原)表 1 资料为对照组、小运动组、大运动组不同有氧运动时间对大鼠血清 LDL 水平的影响, 原作者用 t 检验分别对各组均数逐一进行比较。对照组大鼠在实验期间血清 LDL 水平没有明显变化 $P>0.05$ 。与对照组相比, 小运动组大鼠随着有氧运动时间的延长, 血清 LDL 浓度逐渐降低, 至第 6 周时, 差异具有统计学意义 $P<0.05$ 。大强度游泳运动训练组的大鼠血清 LDL 则呈逐渐升高趋势, 虽然与同期对照组大鼠相比, 差异没有统计学意义 $P>0.05$, 但与同期(第 4 周和第 6 周)小运动组大鼠血清 LDL 相比, 差异具有统计学意义 $P<0.05$ 和 $P<0.01$ 。这种检验方法有误, 因为该资料为多组均数间的比较, 正确的方法是用单因素方差分析, 只有在方差分析有显著性的基础上有必要再作均数间的两两比较, 用 q 检验, 而不是 t 检验, 同时在表 1 中应该列出方差分析的统计量 F 值及具体的 P 值。

2.2 处理多因素设计定量资料误用单因素设计定量资料的方差分析方法

例 4:《不同营养干预对运动性贫血大鼠红细胞及铁代谢指标的影响》^[7]一文中, 实验目的为:观察补充复方阿胶中药与补充铁剂营养对运动性贫血大鼠红细胞及铁代谢指标的影响。方法:以 32 只健康雄性 Wistar 大鼠为研究对象, 每只体重约 300 g, 随机分为 4 组, C 组为安静对照组($n=8$), E 组为递增负荷跑台运动组

($n=8$), EN I 组为递增负荷跑台运动+阿胶营养补充组 ($n=8$), EN II 组为递增负荷跑台运动+铁剂营养补充组 ($n=8$)。然后 EN I 组和 EN II 组营养补充 4 周。实验结束后 24h 内处死。结论: 9 周递增负荷跑台运动导致大鼠红细胞相关指标的显著性下降, 引起运动性贫血, 但血液铁代谢无显著变化; 补充 4 周复方阿胶中药制剂或铁制剂, 提高红细胞相关指标, 改善大鼠运动性贫血状况。数据统计方法实验数据用 SPSS13.0 软件中 one-way ANOVA 处理, 结果用均数 \pm 标准差表示, 显著性水平为 $P<0.05$, 非常显著性水平为 $P<0.01$ 。

原文采用单因素设计定量资料的方差分析不合适, 仔细考察各处理组之间的关系有: 是否递增负荷跑台运动, 是否补阿胶营养或补铁剂营养(两者都能促进 RBC 生成, 增加 Hb 含量的物质, 在本研究中可以

认为它们是同一因素), 原文在固定一个因素于某个水平下观测另一个因素的作用, 这种做法割裂了实验设计的整体性。正确的做法是, 先从专业角度上分析二个实验因素之间是否存在交互作用、是否存在二个因素有主次之分、是否存在交互作用或交互作用可以忽略不计等情况。也就说对事物的影响只考虑施加了一个条件(因素)为单因素方差分析; 如果施加了 2 个以上的条件(因素)为多因素方差分析。如果是多因素方差分析还要考虑施加的条件(因素)之间有没有联系、有没有主次之分。具体看(原)表 2 资料, 施加了两个条件(因素)即递增负荷跑台运动和补阿胶营养或补铁剂营养, 然后考虑两条条件(因素)之间的联系, 通过分析 2 个因素间存在交互作用的可能性比较大, 应采用交互作用多因素方差分析处理此定量资料为宜。

(原)表 2 9 周跑台运动后大鼠红细胞指标 ($\bar{x} \pm s$) 变化

| 组别 | N | ρ (N Hb)/(g·L ⁻¹) | $10^{12} \times \text{RBC}/\text{L}^{-1}$ | $10^2 \times \text{HCT}/\text{L}^{-1}$ |
|-------|---|------------------------------------|---|--|
| C | 8 | 157.13 \pm 6.98 | 9.16 \pm 0.49 | 48.15 \pm 2.43 |
| E | 8 | 141.88 \pm 9.14 ²⁾ | 7.83 \pm 1.45 ¹⁾ | 43.10 \pm 4.51 ²⁾ |
| EN I | 8 | 151.50 \pm 6.48 ³⁾ | 8.68 \pm 0.37 | 47.24 \pm 2.32 ³⁾ |
| EN II | 8 | 149.75 \pm 9.10 ³⁾ | 8.49 \pm 0.50 | 46.61 \pm 3.67 ³⁾ |

与对照组相比较, 1) $P<0.05$, 2) $P<0.01$; 营养补充组与运动组相比较, 3) $P<0.05$

2.3 统计方法不合理, 每一种统计方法都有其适用条件

例 5: 《高脂饮食对大鼠脂肪组织脂解调节因素的影响》⁸⁾, 原文中采用单因素方差分析的 LSD 方法对(原)表 3 各组 FFA 浓度进行分析, 得出结论具有统计学意义。

(原)表 3 10 周高脂饮食对大鼠血清 FFA 和甘油浓度 ($\bar{x} \pm s$) 的影响 $\mu\text{mol}/\text{L}$

| 组别 | c (FFA) | c (甘油) |
|------|------------------------------------|----------------------------------|
| C 组 | 234.09 \pm 14.64 | 76.81 \pm 7.65 |
| OR 组 | 273.30 \pm 30.3 ¹⁾ | 80.10 \pm 10.09 |
| OB 组 | 535.10 \pm 37.33 ²⁾⁴⁾ | 91.94 \pm 8.68 ²⁾³⁾ |

与 C 组相比较, 1) $P<0.05$, 2) $P<0.01$; 与 OR 组相比较, 3) $P<0.05$, 4) $P<0.01$

多个总体均数比较的方差分析, 要求方差齐性。方差齐性实际上是指要比较的各组数据分布是否一致, 通俗的说就是各组是否适合比较, 一般的经验是如果在比较的各组中, 如有一组的标准差是另一组的一倍时方差不齐性的概率较大。在做方差分析时, 做方差齐性检验主要是确认数据的合理性(不具相关性)。对(原)表 3 数据进行方差齐性检验, 可发现 C 组

与 OR 组及 OB 组 FFA 浓度指标不能满足方差齐性的要求, 故不能采用单因素方差分析的 LSD 方法。可以先进行对数、倒数或函数的转换, 选择适当的转换形式, 直到齐性检验变为不显著(不相关)。如果还不行就只能用非参数的单因素分析。

3 定性资料统计分析方面存在的统计学错误

3.1 误判资料类型, 而误用定量资料统计分析方法

例 6: 题目: 《传统体育对老年知识分子原发性高血压患者临床症状影响的观察研究》⁹⁾原作者对原发性高血压, 症状疗效评定标准是: 头痛、眩晕、心悸、耳鸣、失眠、烦躁。各症状按症状轻重不同采用“半定量”计分法, 按程度进行分级和计分, 共分为 4 级: (“无” 0 分、“轻” 1 分、“中” 2 分、“重” 3 分), EH(原发性高血压)患者经 6 个月的传统体育锻炼实验, 症状计分情况。治疗前后组内比较, 除 A 组心悸、耳鸣、失眠症状外, 3 组均能改善老年知识分子 EH(原发性高血压)患者临床症状($P<0.05$ 或 $P<0.01$); 实验后与 A 组比较, B 组眩晕、心悸, C 组头痛、眩晕、心悸、失眠、烦躁症状改善均优于对照组 $P<0.05$ 。

统计资料常常分为定量资料和定性资料两大类,

所谓定量资料是指对每个观察单位用计量方法测量某项指标所获得的数值;如身高(cm)、血压(mmHg)、脉搏(次/min)、红细胞($10^9/\text{mm}^3$)转氨酶(酶活性)等,都属于计量资料。而定性资料是指记录每个观察单位的某一方面的特征和性质^[10]:如血型(A、B、O、AB)、职业(工人、农民、军人、学生)、性别(男、女),等等。本资料观察的是原发性高血压症状疗效,这里的 0、1、2、3、4 代表的是原发性高血压患者症状的轻重程度,属于定性资料中有序资料,(如:无、轻、中、重;治愈、显效、好转、无效、死亡),但原作者却错误地将其判断为定量资料,表的标题后括号内写 $\bar{x} \pm s$ 的形式,表中给出也是平均数和标准差的数据。一般来说, t 检验仅适于分析来自“单组设计、配对设计和成组设计”的定量资料,用分析定量资料的方法分析定性资料,显然是错误的。正确判定统计资料的类型是选用统计分析方法的首要前提^[11]。本资料属定性资料,应根据分析目的,合理选用适合此类资料统计分析方法(如 Riditi 分析或者非参数检验秩和检验)。

3.2 两样本率的检验误用两样本的 t 检验来代替

例 7:《健身跑运动对中年人血脂异常者血脂、身体成分的影响》^[12]一文中将所有受试者随机分成低强度组(L组)、中等强度组(M组)和对照组(C组)。统计数据用平均值 \pm 标准差形式表示。相关指标数据用 t 检验和多因素方差分析进行分析, $P < 0.05$ 和 $P < 0.01$ 为显著性差异。脂肪百分比是脂肪体重除以体重,所以表 4 中实验前与实验后的数据是构成比。实验前后数据的统计处理方法,应采用两个样本率的检验,即 u 检验而不是 t 检验。如果处理的数据是样本率(或构成比)时,要对它们进行假设检验时,则采用率的假设检验,即:正态近似法。当两样本含量 n_1 及 n_2 足够大,此时两样本率的频数分布近似正态分布可用正态分布的原理作 u 检验。如果是两个率(或两个构成比)比较的假设检验或多个率(或多个构成比)比较的假设检验时,都可采用 χ^2 (卡方)检验^[13]。

(原)表 4 健身跑运动前后男性身高、体重指数和脂肪

| 指标 | 百分比的变化 ($\bar{x} \pm s$) | |
|--------|----------------------------|--------------------------------|
| | 实验前 | 实验后 |
| ∴ | ∴ | ∴ |
| 脂肪百分比% | | |
| L 组 | 20.27 \pm 3.50 | 19.09 \pm 3.31 ¹⁾ |
| M 组 | 23.86 \pm 3.38 | 22.37 \pm 3.43 ¹⁾ |
| C 组 | 20.31 \pm 3.79 | 20.49 \pm 3.59 |

与该组实验前比较, 1) $P < 0.01$

运动医学研究离不开统计设计和统计学方法,如果选择错误导致结果有偏倚或结果原则性的错误,对运动医学的危害性较大。正确的统计分析是获得正确的科研结论的要素之一。只有明确资料的性质、试验设计类型、分析目的,掌握相关统计方法应用的前提条件,才能进行正确的统计分析。

参考文献:

- [1] 郭汉,高晓华. 有氧健身操结合饮食控制改善代谢综合征患者胰岛素抵抗、血脂、血糖的实验研究[J]. 北京体育大学学报, 2009, 32(2): 77-82.
- [2] 王科峰. 不同类型休闲活动对老年女性身体机能影响的差异性研究[J]. 成都体育学院学报, 2009, 35(8): 70-72.
- [3] 潘发明,范引光,叶冬青. 卫生统计学教学形式和内容的拓展[J]. 疾病控制, 2004, 8(4): 1-2.
- [4] 赵书祥. 实用体育统计学[M]. 北京: 北京体育大学出版社, 2008.
- [5] 章罗庚. 有氧运动对大鼠血糖、血脂和血液凝固功能的影响[J]. 北京体育大学学报, 2009, 32(7): 66-68.
- [6] 胡良平,刘惠刚. 定量与定性资料统计分析错误辨析(2)[J]. 外科理论与实践, 2005, 10(6): 15-16.
- [7] 薛统,高颀. 不同营养干预对运动性贫血大鼠红细胞及铁代谢指标的影响[J]. 北京体育大学学报, 2009, 32(8): 62-64.
- [8] 田吉明,汪军,王瑞元. 高脂饮食对大鼠脂肪组织脂解调节因素的影响[J]. 成都体育学院学报, 2009, 35(1): 60-64.
- [9] 汤庆华. 传统体育对老年知识分子原发性高血压患者临床症状影响的观察研究[J]. 北京体育大学学报, 2009, 32(2): 67-69.
- [10] 安胜利. 统计学系列讲座(第 8 讲): 基于秩次的非参数检验[J]. 护理学报, 2006, 13(9): 94-95.
- [11] 周素华,陆云霞. t 检验和 χ^2 检验常见误用辨析[J]. 公共卫生与预防医学, 2007, 18(2): 110-111.
- [12] 郑健荣. 健身跑运动对中年人血脂异常者血脂、身体成分的影响[J]. 北京体育大学学报, 2009, 32(7): 56-59.
- [13] 丛湖平. 体育统计学[M]. 北京: 高等教育出版社, 1998: 122.