

· 述评 ·

医学研究论文中应避免的统计学问题

董时富

医学论文是概括性描述研究过程、客观报告研究结果、体现研究水平与科学价值的论证性文章。论文质量取决于该研究的科学性、创新性、实用性与论文的表现形式,而医学研究设计的内容是体现论文科学价值与研究结论可靠性的基础。统计学处理在医学期刊论文中占有重要的地位,绝大多数医学期刊论文都涉及到统计学处理。然而,在应用统计学处理过程中所出现的问题也屡见不鲜。如何正确地应用统计学方法来解决医学研究中的实际问题,保证论文的科学性,这是值得论文作者和编辑探讨的。现将医学研究论文中常见的统计学问题作如下评述,期望能对提高医学研究论文质量有所帮助。

一、医学研究设计的主要内容应该在论文中有所交代

研究设计是对一项科学试验全过程的科学化思考与周密安排,是试验研究的科学价值灵魂及试验成功与失败的关键,是体现论文科学价值与研究结论可靠性的基础。因此,有关研究设计的重要内容如研究对象的选择过程、处理因素的分配方法、对照组的设置方式、效应指标的观测方法与标准、随机化方法等都应该在论文(材料和方法)中加以阐述,以增加读者对该研究全过程的了解及该研究结果科学性和可信度的认同。在医学论文中常常出现的问题,是对研究设计的内容表述不全,最多见于研究对象的选入标准模糊、对照组设置过程与随机化方法在论文中交代不清,从而降低了论文的科学性,同时也反映了论文作者缺乏对科学研究的严谨态度。

二、缺乏对照或不正确的对照会导致不科学的结论

有比较与才有鉴别,对照是试验研究中普遍采用的一种科学方法。研究中设置对照组的目的在于控制非试验因素对试验因素效应的干扰,缺乏对照或不正确的对照都会导致不科学的结论。

某作者应用 15% 硝酸银局部涂药治疗变应性鼻炎(每天 1 次,5 次为 1 疗程),其疗效观察结果如表 1 所示。作者由表 1 的结果得出“该法对各症状的有效率均在 75.5% 以上,值得推广应用”的结论。

表 1 15% 硝酸银治疗变应性鼻炎的结果

症 状	观察例数	显效	有效	无效	总有效率(%)
打喷嚏	138	67	42	29	79.0
流鼻涕	138	63	43	32	76.8
鼻痒	106	51	30	25	76.4
鼻塞	98	43	31	24	75.5

本例中的问题很明显是没有设置对照组,其结论是不科学的。因为对于患感冒所出现的这些症状,即使不采用任何处理,经过 5 d 后也会有相当高的自愈率。只有以对照组作为本底,方能通过比较的方法对纯净处理的效应作出评价。

在研究中设置对照组的方式不正确也同样导致不科学的结论。例如某研究者探讨 X 线照射脑垂体是否有降血压效应,选择 30 例原发性高血压患者进行试验,对每位患者实施 X 线脑垂体照射的试验前、后分别测量患者的血压值,由 30 对试验前、后的差值数据,经配对 t 检验($P < 0.05$),从而得出“X 线照射脑垂体有显著性降血压效果”的结论,这也是不科学的。虽然该研究以患者试验前的血压值为对照(称为自身对照、内对照),但因为缺乏无 X 线脑垂体照射的对照组(称为外对照)数据,故无法证实患者试验后的血压值比试验前的血压值的降低是否就一定归于 X 线照射脑垂体的效应。很显然,我们有理由怀疑:该试验的血压值下降是由于患者的心理作用而不是 X 线照射脑垂体的效应。

在医学期刊论文中设置对照时容易出现的另一个问题,是没有考虑试验组与对照组的可比性。如果本次研究中借用外地的同类资料作为对照组,应考虑外地的试验条件与本地的试验条件是否相同;若采用本地的同类历史资料作为对照组,就应该考虑是否存在因为时间的推移使本次试验的条件与取得历史资料时的试验条件发生了变化。也就是说,除了被研究的处理因素外,应该使试验组和对照组中的所有其他试验条件因素都尽可能地保持一致,即所谓的齐同对照。非齐同对照下的研究结果是不可靠的。

三、就诊或住院患者作为信息源的资料应避免得出不正确的结论

以医院就诊或住院患者为对象在医学期刊论文中频繁可见,由此信息源获取的资料具有相对准确、详

实、容易获取的优点。但是要特别注意,不应由医院内的患者资料信息来推断反映自然居住人群总体特征的统计指标。因为来院就医的患者人群不是自然居住人群总体的随机样本,所以无法计算发病率、罹患率等指标。而在医学期刊论文中时有这类错误发生。例如某作者为了研究伤害的原因及其流行病学特征,以某市城区一所医院在 1998~2000 年间所有因伤害收治的 1352 例患者为对象,通过回顾性调查方式获取相关信息。经数据分析,认为受伤害者中男性发生率高于女性(2.23:1);机动车交通事故发生率(31.07%)最高,是伤害的头号杀手,他杀(28.70%)、意外跌落(19.08%)的发生率相对次之;伤害以 15~30 岁最高(32.91%)、30~40 岁次之(30.55%);职业以农民发生率最高(29.88%)、工人次之(19.16%),科教人员最低(1.48%)、学龄前儿童较低(3.99%),如此等等。殊不知该研究的对象为因伤害来院就医的患者人群,因为受到医院、患者的双向选择以及其他各种原因,尽管该样本达千例以上,但却是一个偏性样本,不能成为自然居住人群的流行病学总体中的随机样本。换言之,这 1352 名患者以不成比例的方式来自不同特征的人群,无法找到计算事件发生率的分母人群数。该文中所谓的“率”只不过是一些“比例”而已。这是将比例指标误作频率指标的典型案例。切记:来自医院的数据为偏性数据,称为伯克森偏倚(Bekson's Bias),不要轻易地用此类数据计算流行率。

四、足够多的重复试验数量是保证统计推断稳定性的基础

根据统计学原理,为了用样本的信息推断相应总体的统计学特征,必须保证从该总体中随机化抽取的研究单位有足够多的数量,即样本含量应足够大。理论上认为,当样本含量趋近于总体的单位数量时,其样本统计量值将趋近于总体参数真值。实际上,我们无法也无须花费巨大的研究成本来获得总体参数的真值,但是我们可以通过有一定数量的样本信息来推断相应总体参数的最大可能估计值,即用样本统计量作为总体参数的无偏估计量。然而,样本统计量的这种无偏性的统计学特征只有在样本含量足够大的条件下才具有,当样本含量较小时就变得极不稳定。所以在小样本含量条件下计算出的统计量并不具有统计学推断的价值。

五、医学研究中的随机化问题

随机化是指在研究对象的选择和处理因素的分配时,使所有被选择对象都能以同等的机会(概率)进入试验的过程。随机化的目的是使非试验因素在试验组和对照组之间达到均衡,从而保证试验结果的科学性以及分析结论的可靠性。在医学期刊论文(特别是临

床观察研究)中,最常见到的问题是遵循随机化原则不严格。表现在临床观察的病例组和对照组的病人分配上,并不是在严格的统计随机化意义下执行病例分组,而是按照研究者的意愿随便地实施分组,这种“随意”和“随便”不等于统计学上的随机化。有少数期刊论文看上去只是一份“回顾性”临床总结报告,即在撰写论文之前并无严格的试验设计,而是搜集手头已有的可得病例资料,能有多少就多少,没有对照随便找。依据这种随意性资料不可能获得有科学价值的结果。

六、关于统计学假设检验问题

假设检验是应用统计学原理,通过观察样本的统计信息对某种关于相应总体特征假设的可信性所作推理性判断的方法。绝大多数医学期刊论文都涉及到统计假设检验,出现的问题也不少见。在此仅对医学期刊论文中出现的一些比较基本的常见问题作如下简述。

1. 忽视了统计检验方法的应用条件与适用范围:例如四格表专用卡方检验要求四个格内的理论频数均大于 5,否则应采用校正卡方检验公式;若其中某一格的理论频数等于或小于 1,则应选用精确概率检验法。 t 检验仅适合两组均数差异的比较,而方差分析可用于两组或两组以上均数间差异的比较;而且 t 和 F 检验都要求变量服从正态分布及各被比较组的方差齐同。

例如某医学学报期刊报道冠心病患者凝血酶-抗凝血酶 III 复合物(TAT)检测与 t 检验分析结果如表 2 所示。

表 2 冠心病患者与正常对照组 TAT 测定结果($\bar{x} \pm s$)

组别	样本含量(n)	TAT($\mu\text{g/L}$)
正常组	45	7.32 ± 4.37
AMI	44	$16.00 \pm 13.20^{**}$
UA	21	$13.80 \pm 8.95^{**}$
OMI	16	$8.76 \pm 6.74^{*}$
SA	15	7.30 ± 4.67

注:与正常组比较,* $P < 0.05$,** $P < 0.01$

从表 2 的结果可发现两个应用统计检验的错误。首先,该资料有四个冠心病组与一个正常对照组间进行比较,已经超出了 t 检验的能力范围,应该采用方差分析的检验方法。其次,由表 2 中的第 3 列数据可见该指标(TAT)测量值在各组中的方差并非齐性,最大方差与最小方差的比值 $= (13.20^2 / 4.37^2) = 9.12$,而且各组的变异系数都在 59.70% ~ 82.50% 之间,明显显示该指标缺乏正态分布特征。

2. 应在论文中给出统计检验方法的名称、检验统计量值与概率值;医学期刊论文中经常出现的另一个问题是在论文中未写出所采用的统计检验方法的名

称,只给出概率值;有的则既无统计检验的名称、统计量值,又无概率 P 值得出某组显著性地高于某组的结论。

3. 统计结果的解释应与统计检验的能力范围一致:如多个率、比的行列表经过卡方检验呈现显著性统计结果时,只表示这多个率、比之中至少有两个率、比之间的差异呈显著性。而有些作者在此情况下将结果表述为各个率、比之间均有显著性。

七、正确应用统计图与统计表

统计表是采用表格的形式归纳试验结果数据的方式,具有逻辑关系清晰、一目了然的优点。所以在应用上首先应注意避免在一个表中容纳过多的主题,使统计表变得臃肿而失去一目了然的优点。其次是要规范化制作统计表格,表格中用数据整齐对位填充而不应留有空格。

统计图是将具体数值模糊化,用点的位置、线段的长短以及几何图形的面积大小等来直观表达事物的相对关系的方法。应用统计图的重点在于选择合适的图形来表达论文中要表达的内容;有些期刊中很少见到统计图,可能是难于制版的缘故。但是,在必要的时候

采用统计图是胜于用表和文字表达作者意愿的好方法。

八、统计学方法只能帮助研究者发现规律,而不能创造规律

科学假设是根据已知的科学事实和科学理论,对所研究的自然现象及其变化的规律提出假定性推测和说明。统计方法论采用的这种科学思维方式有时被误解了,以致产生了许多关于统计学的怪论。如认为“统计学家是一个从无根据的假设到预料中的必然结局之间划一条精确界限的人”。这就产生了对统计方法的滥用,这种滥用现象在杂志、报纸以及电视广告中屡见不鲜。其中也不乏有利用统计学说谎者;有些人对待统计学方法采用实用主义的态度,他们拿着数据来选用统计学方法,目的仅在于要用统计学上的一个 P 值来支持他的结论。要知道统计学是帮助研究者发现规律的科学方法,而不是用以“创造规律”的工具。他们应用统计学就如同醉汉应用街旁的路灯,不是用路灯为他照明,而是用路灯柱来支撑自己。

(收稿日期:2002-03-08)

(本文编辑:郭铁成)

《The Chinese – German Journal of Clinical Oncology》 新刊出版声明

《The Chinese – German Journal of Clinical Oncology》(《中德临床肿瘤学杂志》)的前身《德国医学》杂志创办于 1984 年,是中德合作出版的优秀学术性刊物。

随着我国医学科学发展的需要和中国入世后的国际需求,中德双方一致决定,在 2002 年《德国医学》杂志全面改刊为《The Chinese – German Journal of Clinical Oncology》(《中德临床肿瘤学杂志》)。

新刊将继承原有《德国医学》杂志高水平、高质量、国际化的发展优势,主要刊登世界各国作者,特别是中国作者在肿瘤领域领先的科研成果和临床诊疗经验,以及与肿瘤临床密切相关的基础理论研究等内容,并全文以英语发表,国内外公开发行。它仍保留与德国合作出版的性质,仍属中德 – 德中医学协会会刊。编委会仍由中德双方知名专家组成,中方名誉主编为中国科学院资深院士裘法祖教授;中方主编为中国科学院院士吴孟超教授、中德医学协会主席龚非力教授。

新刊将一如既往地继续以广大读者为中心,不断为读者奉献学术价值高、临床实用性强、信息量大、内容新颖、刊登及时、印刷精美的优秀文章;新刊也将一如既往地继续以广大作者为依托,积极为作者服务,严格做到对所有来稿处理及时、审稿认真、退修详细、发稿迅速;新刊还将全面实现与国际出版惯例接轨,求实求精,为国际权威检索机构,如 Medline、SCI 等收录本刊做好准备。

改版后新刊拟于 2002 年以季刊正式出版,逢季末 25 日出版,大 16 开版,64 ~ 80 内页,激光照排,胶印,每册订价:人民币 ¥28.00、美元 \$30.00,全年订价:人民币 ¥112.00、美元 \$120.00。统一刊号:ISSN 0178 – 3351/CN/42 – 1654/R,邮政代号:38 – 121。全国各地邮局均可直接订阅,如订期已过,请直接向编辑部订购。

在此本刊衷心感谢广大读者多年来对原《德国医学》杂志的支持与合作!希望改刊后的《The Chinese – German Journal of Clinical Oncology》(《中德临床肿瘤学杂志》)能得到您一如既往的欢迎。

编辑部地址:430030 武汉市解放大道 1095 号同济医院《中德临床肿瘤学杂志》编辑部

电话:(027)83662630;传真:(027)83662645

E-mail: dmedizin@tjh.tjmu.edu.cn

《中德临床肿瘤学杂志》编辑部