

# 医学论文统计表的“不一致”现象

徐苗蓁<sup>1)</sup> 周晴<sup>2)</sup> 宋艳<sup>1)</sup> 周英智<sup>1)†</sup>

1) 山东大学自然科学学报编辑部, 250061; 2) 山东大学基础医学院, 250012; 济南

**摘要** 列举医学论文统计表的“不一致”现象: 表题与栏目不一致, 表中数字之和与合计不一致, 表注与表身不一致, 摘要与表不一致, 表与“统计学分析”不一致, “讨论”与表不一致。分析出现这些问题的原因, 并给出了建议。

**关键词** 医学论文; 统计表; 不一致; 原因分析; 建议

**Inconsistencies related to statistical tables in medical papers//**

XU Miaozhen, ZHOU Qing, SONG Yan, ZHOU Yingzhi

**Abstract** This paper lists the inconsistencies related to statistical tables in medical papers between table title and column heading, between the sum of the numbers and the total, between footnote and table, between abstract and table, between table and the content described in “statistical analysis”, and between discussion and table. The reasons are analyzed and some suggestions given on how to avoid these problems.

**Keywords** medical paper; statistical table; inconsistency; reason analysis; suggestion

**First-author's address** Editorial Department of Journal of Shandong University (Science and Technology), 250061, Jinan, China

DOI: 10.16811/j.cnki.1001-4314.2017.04.014

统计表是医学论文中最常见的数据表达形式, 其质量在一定程度上反映医学论文的水平<sup>[1]</sup>。“不一致”现象是医学论文中常见的问题之一<sup>[2]</sup>, 在统计表中尤为突出, 如表本身的相关内容不一致, 或统计表的内容与文中其他内容相矛盾, 这些问题严重影响论文的学术质量。众多学者对统计表的编制及其常见问题做了相关研究<sup>[3-4]</sup>, 但尚未见专门针对统计表“不一致”现象的报道。现结合典型案例做一系统总结。所有案例均来自 2016 年 CNKI 收录的论文。除案例中存在的统计表“不一致”现象外, 对其他问题不做分析讨论。

## 1 统计表常见的“不一致”现象

**1.1 表题与栏目不一致** 例 1<sup>[5]</sup> 为观察 DNA 甲基化调控对预处理致敏百草枯 V79 细胞的影响, 实验设 4 个组, A 组为 5-A za-2'-dc 组, B 组为百草枯组, C 组为 5-A za-2'-dc 与百草枯联合作用组, D 组为空白对照组。采用台盼蓝染色检测细胞死亡情况, 结果见表 1。

表 1 台盼蓝染色结果检测细胞死亡率结果( $\bar{x} \pm s$ )

组别	细胞存活率/%
A	7.44 ± 2.31 *
B	12.90 ± 1.21 *
C	53.58 ± 1.57
D	4.58 ± 2.11 *

注: 与 C 组比较, \*  $P < 0.05$ 。

表 1 题名中的“细胞死亡率”与栏目中的“细胞存活率”不一致, 栏目中的“细胞存活率”应为“细胞死亡率”。

**1.2 表中数字之和与合计不一致** 例 2<sup>[6]</sup> 为观察排石汤联合速效救心丸治疗输尿管结石的效果, 将 136 例输尿管结石患者分为治疗组和对照组, 分别给予不同的治疗, 2 个疗程(20 d)后分析腰腹疼痛缓解时间及结石排出时间, 结果见表 2。

表 2 两组腰腹疼痛缓解时间及结石排出时间

组别	n	腰腹疼痛缓解时间				结石排出体外时间		
		10 min	20 min	30 min	>30 min	10 d	20 d	>20 d
治疗组	68	12	23	24	22	62	5	1
对照组	68	12	15	16	5	54	7	7
组间比较		Z = -4.408				Z = -4.716		
P 值		<0.01				<0.01		

将腰腹疼痛缓解时间分为 4 段, 治疗组各时间例数之和为 81, 对照组为 48, 与每组 68 例均不符合。根据表中数据, 我们测算腰腹疼痛缓解时间  $Z = -2.111$ ,  $P = 0.035$ ; 结石排出体外时间  $Z = -2.023$ ,  $P = 0.043$ , 与表中结果不同。

**1.3 表注与表身不一致** 例 3<sup>[7]</sup> 在体外利用  $H_2O_2$  处理外源性骨髓间充质干细胞(BMMSCs), 模拟去卵巢(OVX)大鼠内源性 BMMSCs 的状态, 比较外源性 BMMSCs 未处理、用  $H_2O_2$  处理、用  $H_2O_2$  和杨梅素(My)处理 3 种状态下 SOD、ALP、Osterix 基因的表达。结果见表 3。

表 3 各组大鼠外源性 BMMSCs 基因的表达( $n = 4, \bar{x} \pm s$ )

组别	SOD	ALP	Osterix
BMMSCs	6.73 ± 0.53	9.74 ± 0.54	12.18 ± 0.79
BMMSCs + $H_2O_2$	2.29 ± 0.33 <sup>a</sup>	3.42 ± 0.93 <sup>a</sup>	4.79 ± 0.81 <sup>a</sup>
BMMSCs + $H_2O_2$ + Myr	6.40 ± 0.36 <sup>b</sup>	8.43 ± 1.01 <sup>b</sup>	7.68 ± 0.69 <sup>ab</sup>

注: a:  $P < 0.05$ , 与假手术组比较; b:  $P < 0.05$ , 与 OVX 组比较。

† 通信作者

表3表身未出现假手术组、OVX组,而表注出现与这2个组比较的内容,令人不知所云。仔细查看原文,发现该表注与前面统计表的完全相同,很可能是在文字复制过程中出现了问题。此外,表注中a、b应为上标,以便与表身格式一致。

**1.4 摘要与表不一致 例4<sup>[8]</sup>** 研究非瓣膜性心房颤动患者并发缺血性脑卒中的影响因素,对入选的901例非瓣膜性心房颤动患者随访3年,收集一般状况、相关疾病病史和服用药物的临床资料,按照随访期间是否发生缺血性脑卒中分为卒中组和非卒中组,应用Logistic回归分析发生卒中的影响因素。

作者按患者一般情况、疾病史、服用药物做了3次Logistic回归,其中一般情况的Logistic分析见表4。

表4 患者一般情况的Logistic回归分析

变量	$\beta$	标准误	Wald $\chi^2$ 值	P 值	OR(95% CI) 值
性别	-0.577	0.434	1.765	0.184	0.562(0.240 ~ 1.315)
年龄	0.083	0.021	15.511	0.000	1.087(1.043 ~ 1.133)
身高	-0.038	0.024	2.456	0.117	0.962(0.917 ~ 1.010)
体质量	-0.010	0.017	0.357	0.550	0.990(0.957 ~ 1.023)
房颤病程	0.107	0.158	0.460	0.498	1.113(0.816 ~ 1.519)

摘要中有关年龄的描述为:“发生缺血性脑卒中的危险因素有年龄(OR = 1.087, 95% CI: 0.240 ~ 1.315)……”。95% CI 包含1,说明年龄在Logistic回归分析中无统计学意义,认为年龄是危险因素不对。再看表4数据,年龄的95% CI 为1.043 ~ 1.133,性别的95% CI 才是0.240 ~ 1.315,摘要中误把性别的95% CI 作为年龄的95% CI。

**1.5 表与统计学分析不一致 例5<sup>[9]</sup>** 研究心脏康复治疗对缺血性心肌病室性心律失常的影响,将99例患者分为对照组和治疗组。对照组给予常规治疗,治疗组在常规治疗的基础上给予心脏康复治疗。疗效分为治愈、显效、有效、无效,有效率指有效、显效和治愈病例所占的比例。

在“统计学分析”中交代采用秩和检验比较等级资料,相关结果见表5。由此认为,治疗组有效率、显效率均高于对照组。

表5 2组临床疗效比较 例

组别	例数	治愈	显效	有效	无效	有效率/%
对照组	54	6	15	19	14	74.1
治疗组	45	5	19 <sup>①</sup>	13	8	82.2 <sup>②</sup>

注:与对照组比较,① $P < 0.01$ ,② $P < 0.05$ 。

从表5可以看出,作者对2组显效、有效率做了比较,但未给出统计量及具体P值,推测原文作者采用了 $\chi^2$ 检验。我们采用SPSS 22.0对2组“显效”一例做比较, $\chi^2 = 2.271$ , $P = 0.132$ ,不是表中所交代的 $P < 0.01$ ;对“有效率”一例做比较, $\chi^2 = 0.943$ , $P = 0.332$ ,也不是表中所交代的 $P < 0.05$ 。表中资料为等级资料,作者并没有按照“统计学分析”中所说的那样做秩和检验。我们做的秩和检验结果为 $Z = -1.257$ , $P = 0.209$ ,2组疗效差异并无统计学意义。

**1.6 讨论与表不一致 例6<sup>[10]</sup>** 调查上海市静安区社区老年人缺乏身体活动暴露情况,结果见表6。

表6 上海市静安区老年人缺乏身体活动暴露情况

年份	男/%	女/%	总人群/%
2002	48.49	50.79	49.81
2007	39.37	42.46	41.02
2011	37.33	41.05	39.39
Pearson $\chi^2$ 值	27.28	27.29	55.17
P 值	<0.01	<0.01	<0.01
线性趋势 $\chi^2$ 值	23.49	23.06	46.75
P 值	<0.01	<0.01	<0.01

在“讨论”中,作者提及2002年上海市静安区老年人群中缺乏身体活动的暴露率,为49.84%,与表中的49.81%不符。

## 2 原因分析

对于统计表的不一致问题,作者、审稿专家、编辑都有责任。下面仅从编辑的角度分析出现问题的原因:

- 1)对统计学的基本知识掌握不够,不能发现一些不一致甚至前后矛盾的错误。
- 2)编辑加工不仔细,不注意全文内容的一致性。
- 3)复制出错,特别是1篇论文中有多个结构相似的统计表时,复制后有的内容没有做相应改动。

## 3 建议

1)系统掌握统计学的基本知识,注意发现统计学内容的不一致现象。如例4摘要中年龄95% CI 包含1,说明年龄在Logistic回归分析中无统计学意义,而作者认为年龄的影响有统计学意义,两者相矛盾,再去核实正文中的内容,就会发现摘要与统计表的不一致问题。

2)编辑应学会常用统计软件,如用SPSS进行统计学处理,根据统计表提供的数据,特别是计数资料、等级资料,核算统计学结果是否正确。必要时请作者提供原始数据进行核算。