

# 医学研究论文中构成比与率错误混用的典型案例分析

张 媛 刘雪松 霍永丰 朱 瑶 郭 瑞 宋国营 周 阳 陈新石

《中华医学杂志》编辑部,100710,北京

**摘 要** 患病率分析在医学研究论文中非常常见,但在一些已经发表的论文中,经常会看到将构成比和患病率错误混用的情况。文章通过一些实例来分析构成比和患病率的区别和联系,以及在实际应用中应该注意的问题。

**关键词** 构成比;患病率;医学研究论文

**Analysis of composition ratio and prevalence in medical research papers**//ZHANG Yuan, LIU Xuesong, HUO Yongfeng, ZHU Yao, GUO Rui, SONG Guoying, ZHOU Yang, CHEN Xinshi

**Abstract** Prevalence analysis is very common in medical research papers, but in some published papers, we often found that the composition ratio is mixed with the prevalence. Through some examples, this paper analyzes the difference and connection between composition ratio and prevalence, and the problems that should be paid attention to in practical application.

**Keywords** composition ratio; prevalence; medical research papers

**Authors' address** Editorial Department of National Medical Journal of China, 100710, Beijing, China

**DOI**:10.16811/j.cnki.1001-4314.2019.03.009

在医学研究论文中,常常要涉及统计表达及计算问题,正确的统计分析方法对于一篇临床研究类论文结论的得出至关重要。如果统计方法选用错误,将会导致错误的结论,甚至是完全相反的结论。统计分析资料包括计量资料和计数资料,而绝对数和相对数是研究计数资料的常见数据形式。在实际工作中,相对数的2种形式,即“比”和“率”非常容易混淆,其应用存在较多问题,编辑在遇到这一类情况时,往往不能准确辨析其正确与否。本文通过分析一些已经发表论文中存在的错误混用实例,旨在提醒作者和编者在处理相关论文时要注意审查相关表述及计算方法,避免得出错误的结论。

## 1 概念<sup>[1]</sup>

**1.1 绝对数** 在整理计数资料时,按研究对象的类别清点数目,直接得到的数字为“绝对数”。例如,甲乙2个学校的流感发病情况,甲校流感发病75例,乙校流感发病50例,这些都是绝对发生数。其只能说明甲校发病例数比乙校多,但不能说明甲校比乙校严重,因为2个学校的人口数不一定相同。

**1.2 相对数** 相对数是2个有关的绝对数之比。如果要比较2个学校流感发病的严重性,应知道2个学

校的人口数,并计算其“相对数”。设甲校有1万5000人,乙校有1万人,则甲校流感发病率 $R_{甲} = 75/15\ 000 = 5\%$ ,乙校流感发病率 $R_{乙} = 50/10\ 000 = 5\%$ ,可见,这2个学校的流感发病率是相同的。在医学研究论文中,常用的相对数包括构成指标(构成比)和频数指标(发病率、患病率等),经常见到将这2个指标混淆的情况,即“以比代率”,或“以率代比”。

**1.3 构成指标** 构成指标又称构成比,用来说明事物的内部各构成部分所占的比重。其公式为:构成比 $b = (n_1/N) \times 100\%$ 。其中 $n_1$ 代表事物内部某一构成部分的例数, $N$ 代表事物内部各构成部分例数的总和。需要注意的是,这2种构成比中的任意一个指标是不能拿出来单独比较的。在临床上经常使用的白细胞分类百分比,即属于构成比,其一个重要特征是各组成部分的百分比之和一定等于100%。

**1.4 频数指标** 频数指标是某现象或某事物发生的频率或强度。其公式为:频数指标 $k = n_2/N \times 100\%$ 。其中, $n_2$ 代表某种现象实际发生的例数, $N$ 代表该现象可能发生的例数。

## 2 实例分析

**2.1 以“比”代“率”** 构成比用来说明事物内部某种构成所占的比重,并不说明某现象发生的频率或强度。在很多论文中,经常会出现将构成比按照率的概念去解释和分析结论。

《西藏高原藏族和汉族妇女多囊卵巢综合征患病率及临床和环境特征的比较》<sup>[2]</sup>一文,题名为患病率的比较;但是通读全文,并没有进行藏族和汉族患病率的比较,而只是比较了构成比,犯了典型的“以比代率”的错误。

该文的具体信息如下:收集2015年12月1日至2016年11月30日期间西藏自治区人民医院妇产科内分泌门诊诊断为多囊卵巢综合征患者的临床资料,共调查了1520例门诊患者,其中藏族865例,占56.9%(865/1520),汉族617例,占40.6%(617/1520),其他民族38例,占2.5%(38/1520)。诊断为多囊卵巢综合征的患者165例,其中:藏族103例,占62.4%(103/165);汉族57例,占34.5%(57/165);其他民族5例,占3.1%(5/165)。最后得出的结论是:

藏族多囊卵巢综合征构成比大于汉族构成比。

此结论的错误在于构成比是不能单独拿出来比较的,而且并不符合文题的“患病率比较”,以构成比代替了患病率,犯了概念错误及计算错误。

正确的算法应该首先统计当年当地藏族与汉族各自的适龄妇女总例数,然后用实际发病例数除以总例数,才能得出2个民族各自不同的患病率。此外,该文章同时比较了藏族和汉族多囊卵巢综合征患者的各种临床特征的发生率,见表1。可见其发生率的计算方法也是错误的,藏族患者月经稀发发生率应为实际发生的例数除以藏族患者的例数,即 $35/103 \times 100\% = 24.3\%$ ;同理,汉族患者月经稀发发生率应为 $13/57 \times 100\% = 22.8\%$ ;其余项目以此类推。

《中国25392名儿童与青少年错颌畸形患病率的调查》<sup>[3]</sup>一文统计了全国7个地区共2万5392名儿童和青少年错颌畸形总的患病率,以及I类、II类、III类错颌畸形的患病率。结果见表2。该表表题为患病率,但是表头却都用的是百分比,进一步计算表内各个数据,发现其计算的确实是患病率,分母为参加调查的总例数,错就错在用错了名词,即表2中的“百分比”应改为“患病率”。接下来作者又分析了全体对象与各牙龄组的错颌构成比,如表3所示。该表表题与表

头均正确,但是发生了计算错误。第2行的乳牙期组II类、III类错颌畸形的构成比计算错误,因为构成比的一个明显特征是事物内部各组成部分所占的比例,其加和一定等于100%。只要细心观察,该表第2行的构成比末尾数字加和并不等于0,因此,通过重新计算,其正确的构成比应该分别为19.48%、28.81%。此外,以上2个表格还犯了一个明显的错误,即没有单位,应在表中“百分比”后注明单位“%”。

表1 2组患者临床和生化特征比较

特征	藏族(n=103)		汉族(n=57)		P值
	例数	比例/%	例数	比例/%	
月经稀发	35	21.2(35/165)	13	7.9	0.140
不孕	34	20.6(34/165)	17	10.3	0.679
闭经	27	16.4	15	9.1	0.989
痤疮	47	28.5	25	15.2	0.829
多毛	28	17.0	16	9.7	0.904
黑棘皮症	11	6.7	9	5.5	0.349
泌乳	6	3.6	6	3.6	0.280
超重	35	21.2	18	10.9	0.757
肥胖	4	2.4	0	0	0.298
腰围	60	36.4	32	9.4	0.161
LH/TSH 比值	63	38.2	24	14.5	0.083
睾酮>1.388 nmol/L	42	25.5	17	10.3	0.043

表2 全体调查对象与各牙龄组的错颌患病率

组别	例数	正常颌		错颌总计		I类错颌总计		II类错颌总计		III类错颌总计	
		人数	百分比	人数	百分比	人数	百分比	人数	百分比	人数	百分比
全体对象	25 932	8 172	32.18	17 220	67.82	8 877	34.96	5 090	20.05	3 253	12.81
乳牙期组	5 309	2 557	48.16	2 752	51.84	1 423	26.80	536	10.10	793	14.94
混牙期组	10 306	2 967	28.79	7 339	71.21	3 688	35.78	2 656	25.77	995	9.65
恒牙初期组	9 777	2 648	27.08	7 129	72.97	3 766	38.52	1 898	19.41	1 465	14.98

表3 全体调查对象与各牙龄组的错颌构成比

组别	错颌总人数	I类错颌总计		II类错颌总计		III类错颌总计	
		人数	百分比	人数	百分比	人数	百分比
全体对象	17 220	8 877	51.55	5 090	29.56	3 253	19.89
乳牙期组	2 752	1 423	51.71	536	19.84	793	28.82
混牙期组	7 339	3 688	50.25	2 656	36.1	995	13.56
恒牙初期组	7 129	3 766	52.83	1 898	26.62	1 465	20.55

**2.2 以“率”代“比”** 《中国孕妇、育龄妇女铁缺乏症患病率调查》<sup>[4]</sup>一文调查了我国孕妇及育龄妇女铁减少及缺铁性贫血的患病率,在缺铁性贫血患者中,又进一步分析了轻度贫血和中度贫血的患病率,如表4所示,并由此得出结论:“不同孕期孕妇轻度贫血及中度贫血患病率差异无显著性,孕妇与育龄妇女患病率差异亦无显著性。相同孕期孕妇及育龄妇女轻度贫血患病率分别为中度贫血患病率的7.40和6.00倍”。由表4可以看出,其中的患病率其实是“构成比”,将“构

成比”当成“患病率”是不妥的。欲求不同程度贫血的患病率,应先统计各个组别孕妇及育龄妇女的总例数,然后拿各个组别不同程度贫血例数除以该组别孕妇总例数,而不是除以贫血总例数,才是各个组别不同程度贫血的患病率。或者另一种修改方式是将表题及表头中的“患病率”均改为“构成比”,那么,相应的结果及结论也要随之做出修改。

**2.3 统计图表的表达错误** 《上海市2型糖尿病患病情况现状调查》<sup>[5]</sup>一文调查了上海市社区居民2型

糖尿病的患病率,并分别统计了城市、农村的不同年龄组的患病率,全文共5个表。选取其中1个表作为示例表5。该表表题为患病率,但是表头本应显示患病率的一列,却只给出了一个单位“%”,这是极其不规范的,缺少栏目词,正确的写法应该为“患病率/%”。

表4 孕妇及育龄妇女缺铁性贫血患者的轻、中度贫血的患病率

组别	例数	轻度贫血		中度贫血	
		例数	患病率/%	例数	患病率/%
孕妇	678	605	88.1	82	11.9
孕早期	146	133	91.1	13	8.9
孕中期	227	200	88.1	27	11.9
孕晚期	314	272	86.6	42	13.4
育龄妇女	562	478	85.1	84	14.9

表5 城市不同年龄组2型糖尿病、IGT患病率的性别分布

年龄组(岁)	例数	2型糖尿病		IGT	
		例数	%	例数	%
(表身具体内容省略)					

### 3 对策

**3.1 加强学习,提高对医学研究论文的辨识水平** 在刊出的医学研究论文中,有关这种统计学错误的发生率一直居高不下,编辑具有不可推卸的责任。医学期刊的编辑要认真学习统计学知识,只有自己认为存在错误的文章,或者有疑问的文章,才会去进一步追究其正确与否。若编辑自己看不出问题,就只能带着错误出版了。

**3.2 增强编辑责任心,认真审核稿件** 对该类论文,要认真审查统计表制作是否规范,计算方法是否正确,数据是否有误。对有疑问的文章,要与作者反复沟通,看作者自己是否能够讲清楚,是否已经请教了相关统计专家。若作者自己都说不清楚计算方法,那就要考虑文章的真实性的了。作者是源头,要杜绝虚假,端正学风。编辑是闸口,要起到督促、把关作用。

**3.3 合理利用同行评议,加强统计学审稿** 在三审、五定、编辑加工的实际操作流程中,未必每个编辑都能做到将文章送审统计学专家,所以,对有疑问的文章,务必请统计学专家审稿。

**3.4 提高编辑准入门槛,认真履行编辑职责** 医学期刊是医学领域中较为重要的传播媒体,具有很强的专业性,覆盖的知识范围较广,这就要求医学期刊编辑有

较高的综合素质,并经过系统的临床实验过程的培训和实践,具有一定的论文写作经验以及错误辨识水平。而目前医学期刊编辑从业人员众多,综合素质有一定差距,并非全部来自医学专业,或者即使是医学专业毕业,但并未做过实验,也未写过论文,对统计学知识知之甚少,在工作中专业知识与实际工作内容不能有效结合,从而导致医学论文中存在较多错误而编辑却未能辨识出来。因此,医学期刊的发展需要不断吸收高素质的医学专业毕业的编辑人员,从而更好地履行编辑职责。

### 4 结束语

在医学研究类论文中,临床工作者经常会用门诊或住院患者的资料来分析疾病与年龄、性别、职业等的关系,但是它的前提是务必采用正确的方法进行计算与评估。临床医生在医院系统上很容易得到计算构成比的资料,这种情况下计算出来的相对数绝大部分是构成比,不能当作发病率来分析;而要得到计算发病率的资料并不容易,这需要知道所统计年份当时的确切的受检人数。值得注意的是,在已发表的部分相关论文中,上述“以比代率”和“以率代比”的混用问题确实普遍存在,这有待作者在研究设计之初即严格按照相关规定执行,同时也要求编辑在加工稿件的过程中严格把关,对有疑问的文章一定不能轻易放过,要与作者和统计学专家反复沟通、确认,做到知其然,并知其所以然,以确保所发表文章具有真实性、科学性、准确性。

### 5 参考文献

- [1] 马斌荣. 医学科研中的统计方法[M]. 2版. 北京: 科学出版社, 2001: 112
- [2] 翟康路, 卓嘎, 迟洪滨, 等. 西藏高原藏族和汉族妇女多囊卵巢综合征患病率及临床和环境特征的比较[J]. 中华医学杂志, 2017, 97(37): 2928
- [3] 傅民魁, 张丁, 王邦康, 等. 中国25392名儿童与青少年错颌畸形患病率的调查[J]. 中华口腔医学杂志, 2002, 37(5): 371
- [4] 中国儿童、孕妇、育龄妇女铁缺乏症流行病学调查协作组. 中国孕妇、育龄妇女铁缺乏症患病率调查[J]. 中华血液学杂志, 2004, 25(11): 653
- [5] 李锐, 卢伟, 贾伟平, 等. 上海市2型糖尿病患病情况现状调查[J]. 中华医学杂志, 2006, 86(24): 1675

(2018-12-25 收稿; 2019-02-18 修回)