

# 科技论文中直方图和条形图的错用案例\*

官鑫 韩宏志 姜瑾秋 李欣欣 王丽†

《吉林大学学报(医学版)》编辑部,130021,长春

**摘要** 为了提高科技论文中统计图使用正确率,研究了科技论文中直方图和条形图的定义和编辑加工要点。以 GB/T 3358.1—2009 为依据,结合科技论文中统计图的常见错用问题,就科技论文中直方图和条形图的错用情况进行分析并从编辑加工的角度提出相应的修改对策。

**关键词** 科技论文;直方图;条形图;错用;案例分析

**Misuse of histogram and bar graph in scientific papers**//GUAN Xin, HAN Hongzhi, JIANG Jinqiu, LI Xinxin, WANG Li

**Abstract** In order to improve the accuracy of using statistical graphs in scientific papers, we discussed the definition and the key points of editing and processing of histograms and bar graphs in scientific papers. Based on the national standard GB/T 3358.1-2009 and combined with the common misuse of statistical graphs in scientific papers, we analyzed the misuses of histograms and bar graphs in scientific papers and put forward the corresponding modification countermeasures from the angle of editing and processing.

**Keywords** scientific paper; histogram; bar graph; misuse; case analysis

**Authors' address** Editorial Board of Journal of Jilin University (Medicine Edition), 130021, Changchun, China

**DOI:**10.16811/j.cnki.1001-4314.2019.03.010

在高度信息化时代,研究者需要高效地传递各种信息,相对于文字而言,统计图可以科学、准确、直观和快捷地展现研究者的研究结果,发现事物间的联系和发展趋势,判定研究结果与结论的符合程度。一幅设计合理的统计图,有利于读者理解结果,寻找规律,引出结论,有时其作用远胜于文字。但目前科技论文中使用统计图的正确率不尽如人意。研究<sup>[1]</sup>显示:1998—2012 年中华医学期刊网收录的文献中统计图的使用率为 7.4%,统计图错误率为 14.3%。笔者在实际工作中发现:对比表达频数或频率分布的 2 种统计图,即直方图和条形图的误用现象特别突出。针对这一现象,依据 GB/T 3358.1—2009《统计学词汇及符号 第 1 部分:一般统计术语与用于概率的术语》<sup>[2]</sup>和《卫生统计学》<sup>[3]</sup>中常用统计图的相关内容,举例说明直方图和条形图的错用案例并提出了修改方案。

\* 吉林省卫生和计划生育委员会 2017 年科技能力提升项目(2017G014,2017G015);中国高校科技期刊研究会 2018 年青年基金课题(CUJS-QN-2018-013)

† 通信作者

## 1 直方图、条形图的定义及编辑加工要点

**1.1 定义** GB/T 3358.1—2009<sup>[2]</sup>对“直方图”的定义是:频数分布的一种图形表示,由一些相邻的长方形组成,每个长方形的底宽等于组距,面积与组的频数成比例。直方图的横轴表示数据类型,纵轴表示分布情况。纵轴用来描述连续变量的分布,如纵轴数据为频数,则叙述为频数分布直方图,各组样本的频数之和为样本总量;如纵轴数据为频率,则叙述为频率分布直方图,各组样本的频率之和为 100% 或者 1。

条形图(bar graph)也称直条图,由一组宽度相同,高度与频数成比例的长方形组成,表示名义特性频数分布的图形<sup>[2]</sup>。有时为了美观,将长方形绘制成三维图形,但这并不会增加任何信息,所以不推荐使用<sup>[2]</sup>。条形图可以反映非连续变量或者离散型变量的频数或频率分布。条形图可以横置或纵置。绘制纵置直条图时,横轴为各类别,纵轴为数据变量,即以长方形的高度表示各类别的频数或者频率。条形图分为单式和复式 2 种。

## 1.2 直方图编辑加工要点及绘制步骤

直方图的编辑加工要点:1)直方图的数值轴尺度要求从“0”开始,否则可能对数量间的大小关系形成误导;2)横轴一般取相等组距为宜,如果需取不等组距,应该按照“矩形高为组段频数与组距的比值”来调整矩形的高度;对于等距分组数据,一般用长方形的高直接表示频数分布,长方形的宽表示各组组距;3)由于分组数据具有连续性,直方图的各长方形通常呈连续排列,中间无空隙;4)直方图横轴坐标应能清楚地反映统计对象的分组情况及分组区间的开闭情况。

绘制直方图具体步骤:1)找出所有数据中的最小值和最大值;2)计算全距,全距为数据的最大值与最小值之差;3)确定组段数和组距:根据数据例数的多少,组段数一般可选 10~15 组,组段的左端点称为下限,右端点称为上限。组距为上限与下限之差。实际工作中常按照“组距为全距与预计的组段数的比值”的方法估计组距的大致长度。两端组段应包含最大值和最小值,尽量取较整齐数值作为组段的端点,组距以相等为宜。

### 1.3 条形图编辑加工要点

条形图编辑加工要点:1)条形图的数值轴尺度必须从“0”开始,否则可能对数量间的大小关系形成误导;2)条形图的绘制应遵循纵坐标:横坐标为5:7的比例(近似黄金分割比例);3)大部分情况下条形图的条宽为1.0 cm 或者0.8 cm,直条间间距必须相等,间距可等于条宽、1/2 条宽或2/3 条宽,但小于2 倍条宽;4)复式条形图是基于有联系的复合表的数据绘制而成,复式条形图中应给出具体的图例;5)条形图中体现组间比较情况时应给出各长方形上端的中点指示线条,并在组间比较差异有统计学意义的长方形上方给出相应的统计学符号。

## 2 错用案例辨析

**2.1 直方图的错用** 连续型变量的频数或者频率分布应绘制为直方图,部分作者在撰写论文时分不清直方图和条形图的区别,容易将直方图绘制成条形图。

**例1** 切削磨损统计数据见图1。图1是根据文献中数据绘制的频数分布直方图。可以看出,切削磨损的频数分布直方图为非负型的近似正态分布,因此可以近似运用平均值加2 倍标准偏差的形式来制定界限值<sup>[4]</sup>。

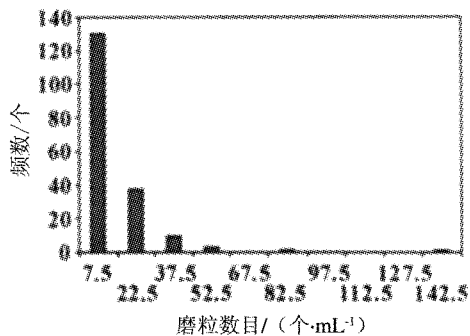


图1 切削磨损颗粒的频数分布

从文献[4]可见切削磨损颗粒的数据是一组连续性变量,该统计图的主要错误是:1)叙述为频数分布直方图,但呈现的却是条形图;2)切削磨损的频数分布直方图文字叙述为“非负型的近似正态分布”是错误的,应叙述为“正偏峰分布”(即峰向左侧偏移的分布);3)直方图的绘制未遵循5:7 的比例规律。根据文献的表中数据内容(切削磨损颗粒统计数据分组和颗粒的最大值、最小值及极差)并参考直方图的编辑加工要点,采用 excel 软件重新绘制调整为图2。

### 2.2 条形图的错用

对于离散型变量的分布状况可以通过频数分布表来概括,呈现为条形图。表示离散型变量的分布状况

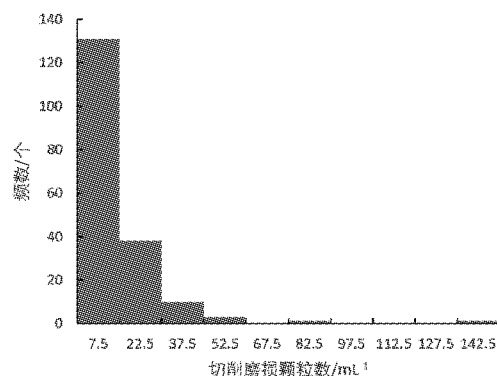


图2 切削磨损颗粒的频数分布

的条形图有时会被误绘制为直方图。

**例2** 2017 年调查某山区260 名孕妇产前检查次数分布情况并绘制统计图。检查0 次者5 人,检查1 次者18 人,检查2 次者22 人,检查3 次者96 人,检查4 次者73 人,检查5 次者30 人,检查多于5 次者16 人(见图3)。

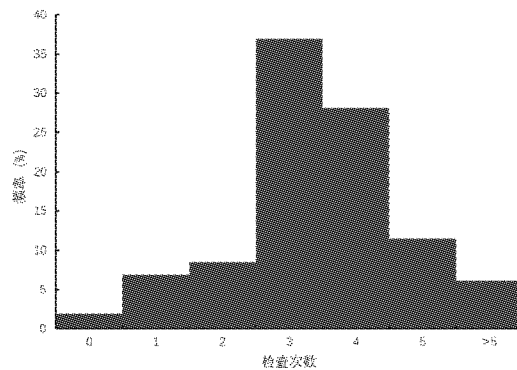


图3 2017 年某山区260 名孕妇产前检查次数

由例2 描述可知,该组数据描述的是一组简单的非连续、离散型变量的分布情况,产前检查次数是离散型变量,应绘制为条形图,而图3 误将产前检查次数理解为连续型变量绘制为直方图。经过修正,根据条形图的编辑加工要点采用 excel 软件重新绘制调整为图4。

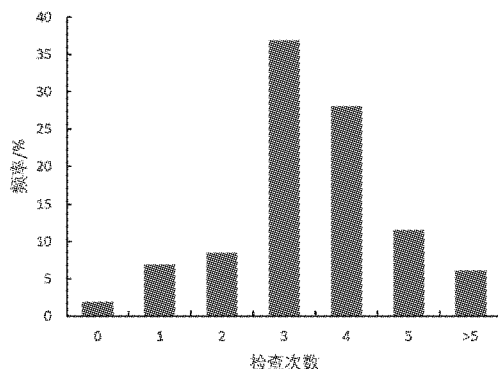


图4 2017 年某山区260 名孕妇产前检查次数

### 3 提高统计图质量的编辑对策

目前我国各阶段教育内容及专业培训中将科学研究结果形象化的理念还不够普及,科学、准确地使用统计图的意义未受到足够的重视<sup>[6]</sup>。针对目前科技论文中统计图使用不规范的现象,科技期刊编辑应该加强学习业务和稿件质量把关。科技期刊编辑应熟练掌握各种统计图的适用条件和使用注意事项,在编辑与作者关于统计图使用的互动中做到“在沟通中普及,在普及中提高”,使科技论文在“科学、准确”的前提下呈现出“灵活、多样”的精美版式,增加科技论文的可读性。

**3.1 加强统计图标准化知识学习** 随着计算机科学和信息技术的发展,各种绘图软件层出不穷,但由于论文作者对于软件使用的不熟练和各类统计图适用条件认知不够,常导致错用统计图<sup>[7]</sup>。科技期刊编辑应该加强对于统计图理论知识和实用操作流程的学习,积极向统计学方面的审稿专家请教,熟练掌握制图软件使用方法和专业的统计学科知识,帮助和指导论文撰写者更准确地绘制统计图,更好地以统计图的形式呈现实验结果。在“内容为王”的时代,应避免误用统计图并传播正确的使用方式,同时也要注意“眼球效应”,如在期刊网站、公众号或者 App 中推送关于统计图的正确使用和常见错用“案例”,方便投稿作者进行检索和学习<sup>[8]</sup>。

**3.2 积极向论文作者宣传相关要求** 对于编辑来讲,高质量的统计图须基于科学的编排原则并结合专业的图片调整方法,以保证统计图的规范性。建议科技期刊编辑出版单位根据相关国家标准(GB/T 3358.1—2009),制订标准化的统计图编辑加工流程,在投稿须知和征稿简则中明确提出对统计图的具体要求。

**3.3 制订统计图加工流程,做好编校工作** 科技期刊编辑应该在工作中制订规范的统计图加工流程,在规范原则下耐心、虚心地对每一位作者,并与作者保持及时沟通,告知作者按照流程做好统计图的编辑加工;由于统计图领域的标准规范也在随时更新,因此编辑应不断研究、发现和追踪新的标准规范,并在杂志的投稿须知和征稿简则中及时更新。在作者群中打造创新的学术内容,严谨的编辑流程,规范的文字加工,正确的统计图表,吸引高质量论文的投稿。

### 4 参考文献

- [1] 张娟,詹凤羽,杨珉. 医学文献中统计图使用比例及其应用错误的抽样调查[J]. 现代预防医学, 2015, 42(19): 3485
- [2] 统计学词汇及符号 第1部分:一般统计术语与用于概率的术语: GB/T 3358.1—2009[S]. 北京:中国标准出版社, 2009
- [3] 方积乾. 卫生统计学[M]. 5版. 北京:人民卫生出版社, 2005
- [4] 徐金龙,易如娟. 基于自动磨粒分析技术的直升机在用发动机油界限值的研究[J]. 润滑与密封, 2012, 37(2): 110
- [5] 王诗扬,卢天丕. 贵州省富锶(Sr)地下水赋存与分布规律探讨[J]. 贵州地质, 2018, 35(3): 225
- [6] 李晖. 第九讲 医学论文中如何正确使用统计图[J]. 中华预防医学杂志, 2002(3): 65
- [7] 马启,周庆泰. 医学论文中常见统计图、表错误分析[J]. 蚌埠医学院学报, 1991(3): 214
- [8] 吴向阳,张超群,张方林,等. 医疗设备数据库中的统计图制作[J]. 中国医疗设备, 2009, 24(9): 16

(2019-01-10 收稿;2019-03-14 修回)