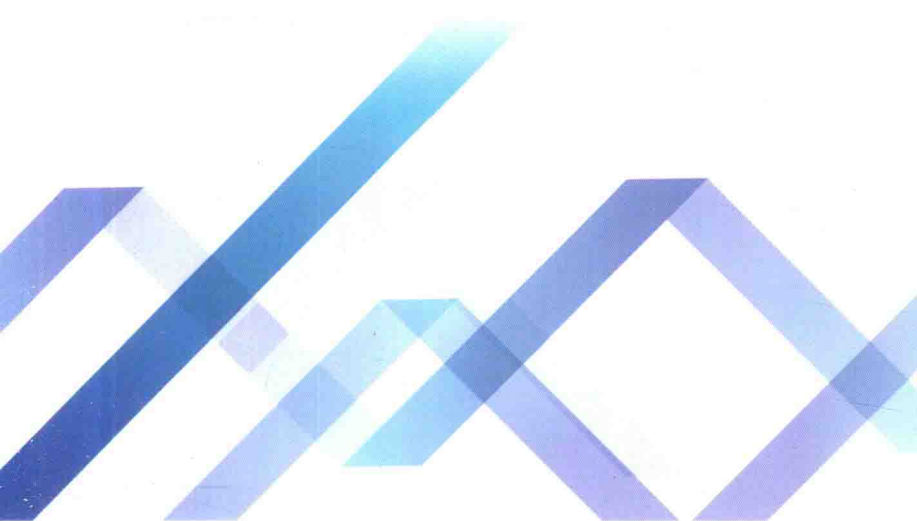


 北京名医世纪传媒

国家高技术研究发展计划课题资助

医学论文写作与编辑 统计学运用

胡良平 编著



中原出版传媒集团
中原传媒股份公司

 河南科学技术出版社

国家高技术研究发展计划课题资助

课题名称:心血管疾病大数据平台的构建和应用研究

课题编号:2015AA020102

医学论文写作与编辑统计学运用

YIXUE LUNWEN XIEZUO YU BIANJI TONGJIXUE YUNYONG

胡良平 编著



河南科学技术出版社

• 郑州 •

内容提要

本书以国际医学期刊中发表的高影响因子 SCI 论文为蓝本,总结了撰写和审读高水平医学论文所需要的基本功,从高水平学术论文的标志与写作要领、撰写高水平医学论文必须掌握的关键技术、如何从统计学角度审读医学论文三个角度展开论述,为希望撰写并发表高水平学术论文的生物医学和临床科研工作者,以及医学期刊和出版社的编辑和审稿专家提供有力工具和有益参考。本书适合临床医学、预防医学、基础医学、药学等领域的学者、编辑、学生、教师、研究人员和医护人员学习和使用。

图书在版编目 (CIP) 数据

医学论文写作与编辑统计学运用/胡良平编著. —郑州:河南科学技术出版社, 2019. 6

ISBN 978-7-5349-9510-1

I. ①医… II. ①胡… III. ①医学—论文—写作 ②医学—论文—编辑 IV. ①H152.3

中国版本图书馆 CIP 数据核字 (2019) 第 083686 号

出版发行: 河南科学技术出版社

北京名医世纪文化传媒有限公司

地址: 北京市丰台区万丰路 316 号万开基地 B 座 1-114 邮编: 100161

策划编辑: 欣逸

文字编辑: 宋建良

责任审读: 周晓洲

责任校对: 龚利霞

封面设计: 吴朝洪

版式设计: 崔刚工作室

责任印制: 陈震财

印刷: 郑州环发印务有限公司

经销: 全国新华书店、医学书店、网店

开本: 850 mm×1168 mm 1/32

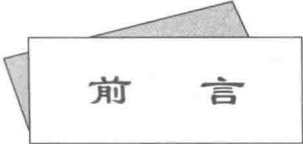
印张: 13.25 字数: 271 千字

版次: 2019 年 6 月第 1 版

2019 年 6 月第 1 次印刷

定 价: 59.00 元

如发现印、装质量问题, 影响阅读, 请与出版社联系并调换



前 言

从多年前开始,每当笔者应邀给全国各地的生物医学和临床科研工作者及期刊、出版社的编辑们讲授统计学时,他们热烈而又渴望的眼神总会给笔者带来一种冲动:一定要为他们写一本书,帮助科研工作者提高撰写学术论文的质量,提升编辑为医学稿件把好统计学质量关的技能。然而,当演讲结束后,由于种种原因,时间和精力又不得不花在其他一些“紧急任务”上。2015年初春的某一天,一位似乎读懂了笔者心声的医学编辑向笔者主动发出了召唤:请给广大医学科研工作者写一本关于医学论文写作的书吧!正是这一句看似很普通的邀请,极大地鼓足了笔者的勇气,努力去实现隐藏在自己内心多年的夙愿,并最终撰写了《医学论文写作与编辑统计学运用》一书。

本书共分三章:第一章“高水平学术论文的标志与写作要领”,从高水平学术论文的标志是什么、国际顶尖级高影响力因子SCI论文的统计学解读、怎样才能撰写出高水平学术论文和高水平学术论文的写作要领这四个方面诠释了什么是高水平学术论文和如何才能撰写出高水平学术论文;第二章“撰写高水平医学论文必须掌握的关键技术”,从制订科学完

善课题技术设计方案的基础知识、必须掌握运用统计软件实现常用统计分析的基础知识、必须掌握制订课题设计方案的关键技术和必须掌握合理选择统计分析方法的关键技术这四个方面提供了技术支撑；第三章“如何从统计学角度审读医学论文”，从已发表各类医学论文中统计学方面问题的辨析与释疑、医学论文统计学审读的要领与技巧这两个方面讲解了审读和编辑医学论文的统计学技巧。

本书以国际顶尖级医学期刊中发表的高影响因子 SCI 论文为蓝本，总结了撰写和审读高水平医学论文所需要的基本功和关键技术，希望能对广大读者有所帮助。

在本书即将出版之际，笔者真诚地感谢所有直接和间接提供帮助的同事、朋友和学生！还要感谢本书中所参考和引用的全部文献的原作者！正是因为他们的无私帮助，才使本书得以问世。

由于笔者水平有限，书中可能出现这样或那样的不妥，甚至错误之处，恳请广大读者不吝赐教，以便再版时修正。

胡良平



目 录

第一章 高水平学术论文的标志与写作要领	(1)
第一节 高水平学术论文的标志是什么	(1)
一、论文中提出的科学问题应具有哪些特性	(1)
二、论文中求证问题答案的过程应具有哪些特 性	(5)
三、论文中获得的问题答案应具有哪些特性	(10)
四、论文在学术和统计语言方面应具有哪些特 性	(12)
第二节 国际顶尖级高影响因子 SCI 论文的统计 学解读	(14)
一、《临床医师癌症杂志》统计学解读	(14)
二、《新英格兰医学期刊》统计学解读	(23)
三、《柳叶刀》统计学解读	(40)
第三节 怎样才能撰写出高水平学术论文	(49)
一、功夫应下在撰写论文之前	(49)
二、应制订出科学完善的课题设计方案	(50)
三、应实时严控研究过程中的质量	(52)
四、应确保研究数据的原始性和准确性	(54)
五、应正确运用统计学方法分析数据	(61)
六、应结合基本常识和专业知识和统计学知识	



解读分析结果	(66)
七、应充分了解目标期刊对稿件质量的要求	(83)
八、应在语言表达上精益求精	(84)
第四节 高水平学术论文的写作要领	(86)
一、生物医学学与临床研究学术论文的分类	(86)
二、国内外学术论文中统计分析部分的比较	(87)
三、学术论文中方法学部分撰写的要领	(100)
四、学术论文中统计分析部分撰写的要领	(131)
五、学术论文中从头至尾恰当运用统计学的要 领	(139)
第二章 撰写高水平医学论文必须掌握的关键技 术	(144)
第一节 制订科学完善课题技术设计方案的基础 知识	(144)
一、科学完善课题技术设计方案的标志	(144)
二、各类课题技术设计方案的核心内容	(149)
三、各类课题技术设计方案的要点与细节	(154)
第二节 必须掌握运用统计软件实现常用统计 分析的基础知识	(181)
一、SAS 软件及其相关内容简介	(181)
二、SAS 软件用法入门	(191)
三、SASPAL 软件用法简介	(207)
第三节 必须掌握制订课题设计方案的关键技 术	(213)
一、常见样本含量估计及其 SAS 实现	(213)
二、常见随机抽样方法及其 SAS 实现	(231)

三、常见随机分组方法及其 SAS 实现	(240)
四、常见多因素试验设计类型的构建	(254)
五、四种比较类型的选定	(269)
第四节 必须掌握合理选择统计分析方法的关键 技术	(272)
一、统计分析方法的分类	(272)
二、合理选择统计分析方法的要素	(288)
三、通过实例演示选择统计分析方法的关键技 术	(289)
四、常见问题与数据结构及统计分析方法的合理 选择	(293)
第三章 如何从统计学角度审读医学论文	(338)
第一节 已发表各类医学论文中统计学方面问题 的辨析与释疑	(338)
一、来自国内医学学术论文统计学方面问题的 辨析与释疑	(338)
二、来自国际医学学术论文统计学方面问题的 辨析与释疑	(374)
第二节 医学论文统计学审读的要领与技巧	(379)
一、医学论文中“题目和摘要”的审读要领与技 巧	(379)
二、医学论文中“方法”部分的审读要领与技巧 ...	(384)
参考文献	(401)
附录 A 胡良平统计学专著及配套软件简介	(405)

第一章

高水平学术论文的标志与写作要领

第一节 高水平学术论文的标志是什么

一、论文中提出的科学问题应具有哪些特性

(一)何为学术论文

首先,应明白何为“学术”。商务印书馆出版的《现代汉语词典》第6版第1479页上这样写道:学术是指有系统的、专门的学问。如学术界、学术思想;学术团体;钻研学术。上海辞书出版社出版的《辞海》缩印本1989版第1269页上这样写道:学术是指较为专门、有系统的学问。那么,何为“学问”?在前面述及的《现代汉语词典》第1479页上写道:学问是指①正确反映客观事物的系统知识:这是一门深奥的学问。②知识,学识:有学问,学问很大。

其次,应明白何为“论文”。在前面述及的《现代汉语词典》第854页上写道:论文是指讨论或研究某种问题的文章:学术论文;毕业论文。

由此可知,学术论文是指系统、专门讨论或研究某问题的文章。本书中所提及的医学学术论文则是指系统、专门讨



论或研究或论证与医学有关的某个或某些科学问题的学术文章,包括在国内、外期刊或学术会议上公开发表的学术论文和(或)未公开发表的硕士和博士学位论文。

(二)何为科学问题

人们提出的问题可以是各种各样的,有些属于一般性问题,如中国熟人之间见面时喜欢问:你吃饭了吗?你上哪儿去?你从哪儿来?有些则属于工作性问题,如有一段时间未见面的同事或朋友或同学之间见面时喜欢问:您最近身体好吗?您最近工作忙吗?您最近获得了科研课题吗?您的技术级别按时晋升了吗?然而,在学术论文中,学者们首先要提出科学问题,如人为什么要做梦?动物有没有思维?为什么有些人易患癌症?人类如何征服疾病?宇宙究竟有多大?最小的微生物究竟有多小?地球上气候的变化规律是什么样的?人类如何能准确地预知何时何地会发生多大震级的地震?人类为何会有战争?人类如何能避免战争?诸如此类的问题,永远也问不完,因此,人类要解决的问题将永无休止。人们不难总结出,何为“科学问题”。

事实上,要回答什么是科学问题,首先应回答什么是“问题”。“问题”有多个答案。第一,问题是要求回答或解释的题目:如我来释疑一下您提出的那几个问题;第二,问题是需要研究讨论并加以解决的矛盾、疑难:例如,这种药治疗胃溃疡很解决问题;第三,问题是指“关键、重要之点”:重要的问题在于先弄清事故的真相;第四,事故或麻烦:如我的自行车又出问题了。其次应明白什么是“科学”。“科学”是指反映自然、社会、思维等的客观规律的分科的知识体系。如科学发展观就是关于我国现阶段发展的符合科学的总体看法和



根本观点,其基本内容:坚持以人为本,通过统筹城乡发展、区域发展、经济社会发展、人与自然和谐发展、国内发展和对外开放等,实现全面、协调和可持续发展。由此可知,学术论文中通常提出的“科学问题”是指人类在认识自然、社会、历史、思维、政治、经济、军事、文化乃至文明等各个方面的客观规律中提出的问题,也包括与发现、创造、创新和发展有关的问题。不难想象:每篇学术论文至少要提出一个科学问题,有时,一个大的科学问题又会被分解成若干个彼此有一定联系但又有一定区分度的小科学问题。整篇学术论文就是围绕提出科学问题、寻找解决它的途径、报告科学问题被解决到了什么程度,最后还应交代此科学问题是彻底被解决了还是需要继续不断研究。

(三)科学问题应具有哪些特性

一般来说,一篇学术论文中提出的科学问题是十分有限的。这里所说的“十分有限”至少有两层含义:其一,数目有限。因为学术论文的篇幅非常有限,所以,在一篇学术论文中不可能提出许许多多的科学问题;其二,范围有限。因为科学问题所涉及的范围可大可小,研究者必须依据自己所掌握的人力、物力、财力、技术和时间,确定与其相称的科学问题。例如,若某项课题的核心内容是了解糖尿病的患病率,那么,对应的科学问题的提法就有如下多种:第一,北京市2010—2016年常住成年人糖尿病的年患病率的调查研究;第二,我国2010—2016年成年人糖尿病的年患病率的调查研究;第三,整个亚洲2010—2016年成年人糖尿病的年患病率的调查研究;第四,西方各国2010—2016年成年人糖尿病的年患病率的调查研究;第五,全球2010—2016年成年人糖

尿病的年患病率的调查研究。由此可知,一个科学问题是由“核心内容”+“范围或程度限制”组成的。不仅如此,一个科学问题至少应具有如下特性。

1. 科学性 一个科学问题本身必须是具有科学性的。所谓“科学性”,应与“反映自然、社会、思维等的客观规律”有关,而不是违背科学的。如若有人想不借助任何动力,让一架真实的飞机飞向高空,并可以长时间、长距离地飞行。显然,这个问题是不能被称之为“科学问题”的。

2. 创新性 一个科学问题应在一定程度上具有创新性。若一篇学术论文中所讨论的科学问题是别人已经提出并已被解决得很好的问题,即使这篇论文写得再好,也没有发表的价值,因为它不具有丁点的创新性。若有人已完成了“我国 2000—2004 年成年人糖尿病的年患病率的调查研究”课题,再有人准备开展“我国 2010—2014 年成年人糖尿病的年患病率的调查研究”课题,还是值得做的。虽然,别人已完成了此类课题,但随着时间的推移,我国的人口构成、人们的思维观念和生活方式、环境与生态状况等都已发生了较大变化,同样的观测指标(成年人糖尿病的年患病率)已经发生了很大的变化。即使原先课题研究者的调查研究方法原封不动地被借用或照搬,新开展的此项课题也具有一定的创新性,因为研究背景与规模变了,调查研究方法也需要做相应的调整。当然,科学问题的创新性更重要的是应体现在以下两个方面:①解决问题的理念、方法和技术上的创新;②研究成果的功能、质量和效益上的创新。

3. 前沿性 通常,前沿性是指当前或未来一段时间内世人特别关注的一类科学问题。这类问题的解决会对一个



学科、一个国家,乃至全人类产生巨大影响或改变。如美国在 20 世纪 60—70 年代就开展了载人登月工程;1990 年前后人类基因组计划就被全球有远见卓识的科学家们提出来了。这就是科学问题的前沿性。然而,一个非常有意义且解决起来却具有较高难度的科学问题,即使早就被人们提出来了,但迄今为止仍未得到很好解决,也可被称为“前沿问题”。如哥德巴赫猜想,早在 1742 年就被提出来了,直到 1966 年,我国数学家陈景润先生证明了“ $1+2$ ”成立。在当时,也应属于数学领域内的前沿科学问题。

4. 实用性 一般来说,人类提出和解决的科学问题都具有很强的实用性。爱因斯坦提出了相对论,为人类认识世界和宇宙找到了一把金钥匙;达尔文提出了生物进化论,为人类认识物种起源和变化提供了科学依据。虚无缥缈的问题常常对于大多数人来说没有吸引力,如有人常喜欢思考“人为什么活着”?也有人常苦恼:“我怎样才能一夜成名、一夜暴富?”

二、论文中求证答案的过程应具有哪些特性

(一)论文中求证答案的过程被研究者玷污了

相当多的论文中提出的科学问题本身是有意义的,其答案也是可以预知的,但求证答案的过程是违背科学的。换一句话说,求证答案的过程被研究者玷污了。请看下面的两个实例:实例一:某研究者提出的科学问题是经常锻炼身体的人会比不锻炼身体的人身体更健康。为了证明这个科学命题是否成立,研究者找来 5 名最近经常锻炼身体的朋友,又找来最近不锻炼身体的另外 5 名熟人。请有经验



的临床医师给他们仔细检查身体,结果却发现:最近经常锻炼身体的 5 名朋友的健康情况反倒比最近不锻炼身体的 5 名熟人差多了。这个结果似乎很不符合常理,但研究者不知道问题究竟出在哪里!实例二:某研究者提出的科学问题是人一旦患了骨质疏松症,其体内很多有关指标的取值就会变得很糟糕。为了证明这个科学命题是否成立,研究者找来 30 名年龄在 60 岁以上且患了骨质疏松症的女性,又找来 30 名年龄在 30 岁左右且未患骨质疏松症的女性。分别测定了这两组受试对象与骨质疏松症有关的多项评价指标的数值。经过正确的统计分析,得出患骨质疏松症的 30 名受试者的情况确实比未患骨质疏松症的 30 名受试者差。显然,这个结果非常符合常理,但求证答案的过程是充满了严重错误的!

以上两个实例都不是“个案”,类似的情况在现实学术杂志的论文中的确具有相当的普遍性。其根源就在于论文中求证答案的过程被研究者玷污了。用通俗的语言来表达,就是论文中求证答案的过程是违背基本常识的、是具有严重硬伤的“课题设计方案”的产物。在实例一中,5 名最近经常锻炼身体的朋友是已知自己身体出了问题(甚至是患了较严重的疾病),刚刚开始抓紧时间锻炼身体;而最近不锻炼身体的 5 名熟人深知自己年轻力壮,短时间内不锻炼身体,料也无妨。由基本常识可知:身体出了问题的人,经过短时间锻炼,对身体健康的改善是无济于事的;同理,身体健康的人,短期内不锻炼身体,对身体健康的损害也不会产生明显的作用。所以,得出上面违反常理的结果和结论也是能解释得通的。在实例二中,患骨质疏松症的 30 名



60 岁以上的女性都处在对身体健康不利的生理时期(绝经后期),她们体内的雌激素水平下降得很快,有些女性的更年期综合征表现得很重,因此,很多标志着女性身体是否健康的评价指标都处于“劣势”;而她们的对照组——30 名年龄在 30 岁左右且未患骨质疏松症的年青女性,正处在生命和生育的旺盛时期,她们体内的雌激素水平正处在高峰期,又不受更年期综合征的困扰,因此,很多标志着女性身体是否健康的评价指标都处于“强势”。所以,这两组受试对象之间在很多评价身体健康状况的指标上的差别不仅仅是由于“患”与“未患”骨质疏松症引起的,她们之间存在的巨大差别,实际上主要是由于她们所处的“生理时期”不同所致!

由此可知,即使论文中提出的问题是“科学问题”,但求证答案的过程(本质上就是课题研究设计方案是否科学完善、课题实施过程中是否始终遵守标准操作规程、质量控制是否实时和精准)被研究者玷污了,无论最终的答案正确与否,都是不值挂齿的!统计学的精华(制订出科学完善的课题设计方案、有严格的标准操作规程的制约、实时精准控制课题实施过程中的质量、确保研究数据的原始性和准确性)就在于此,所有认为统计学就是“数据处理或统计分析”工具的人,都是对统计学的误解或亵渎!因为任何盲目运用统计分析方法处理不值得分析的数据的人们,无异于是在“玩弄数字游戏”!

(二)论文中求证答案的过程应满足哪些条件

论文中求证答案的过程可以被分解成三个方面:第一个方面,求证过程中所依据的“课题设计方案是否科学完善”,



这是研究结论可能遭到一票否决的第一个重要方面。第二个方面,求证过程中是否实时精准地控制住了研究的质量,具体地说,就是:其一,制订并严格落实了标准操作规程;其二,影响评价指标取值的一切重要非试验都在掌控的范围之内;其三,对比组之间处于极好的均衡状态之下。这三点是研究结论可能遭到一票否决的第二个重要方面。第三个方面,求证过程中能否确保数据的原始性和准确性,也就是说,测定观测指标的方法是否准确、测定的时间点选择得是否恰当、测定的数据是否被人为“加工”过,这是研究结论可能遭到一票否决的第三个重要方面。由此可总结出论文中求证答案的过程应满足下列条件。

1. 具有科学完善的课题设计方案 一般来说,医学学术论文是建立在试验和(或)调查基础之上的。人们通过某些特定的试验或调查来收集证据(大多数证据是试验或调查获得的数据),从而来回答或论证事先提出的科学问题。如何高质量、高效率地去实施并完成相应的试验或调查?其态度或做法有以下几种:第一种,没有课题设计方案,直接糊里糊涂开展试验或调查。这类研究课题所得出的结果和结论肯定是不可信的!第二种,有漏洞百出的课题设计方案,在其指导下实施试验或调查。这类研究课题所得出的结果和结论基本上可断定也是不可信的!第三种,有比较完善但有严重硬伤的课题设计方案,在其指导下实施试验或调查。这类研究课题所得出的结果和结论是部分可信、部分不可信的!第四种,有科学完善的课题设计方案,在其指导下实施试验或调查。若后续四个关键部分(课题实施过程中始终按标准操作规程实施、有实时精准的质量控制体系、能确保



数据的原始性和准确性、数据分析合理且正确)都符合要求,则这类研究课题所得出的结果和结论是准确可信的。由常理可知,在课题正式启动之前,必须制订出科学完善的课题设计方案,这是保证基于科研课题而撰写出来的学术论文所得结果和结论准确可信的第一道坚强防线。

2. 具有完备的质量控制体系 通常而言,一个科研课题需要较长时间(平均 2~3 年,有些需要 5 年,还有些甚至需要 10 年)才能彻底完成。即使开始制订的课题设计方案非常“科学完善”,但在如此长的课题实施进程中,影响评价指标的重要非试验因素在不断地产生,同时,也在不断地侵蚀着正在实施的科研课题。如果没有完备的质量控制体系,就好比在风沙飞舞的露天里做所谓的“无菌试验”,岂不是在自欺欺人吗?据此试验条件下收集到的“科研数据”是不值得分析的,因为数据中已经出现了“嵌入式”的污染物。因此,在课题实施过程中,必须确保具有完备的质量控制体系并对其全过程实施实时且精准的质量控制,这是保证基于科研课题而撰写出来的学术论文所得结果和结论准确可信的第二道坚强防线。

3. 具有实时准确记录原始数据的措施 任何试验或调查的结果都是以“科研资料”的形式呈现出来的,常习惯上简称为“数据”。由基本常识可知,单纯的一串“数字”相当于“密码”。通常,这种“密码”是无法被破译的!数据,更确切地说,是科研资料,应该有一些辅助说明。比如,试验或调查的目的是什么?每个观测项目(通常用“变量”来泛指,包括一般变量、原因变量和结果变量)的名称和专业含义分别是什么?每个变量的取值形式、度量单位和范围分别是什么?