



国家卫生健康委员会“十四五”规划教材
全 国 高 等 学 校 教 材



供预防医学类专业用

卫生统计学

Health Statistics

第 9 版

主 编 郝元涛 刘美娜

副主编 赵 杨 王超龙 姚应水

数 字 主 编 王超龙

数字副主编 朱彩蓉 魏永越 秦国友

人民卫生出版社
· 北 京 ·

编委名单

编 委（以姓氏笔画为序）

王 玟	滨州医学院	张秋菊	哈尔滨医科大学
王学梅	内蒙古医科大学	武晓岩	桂林医科大学
王超龙	华中科技大学	尚 磊	空军军医大学
邓 丹	重庆医科大学	金丽娜	吉林大学
刘红波	中国医科大学	赵 杨	南京医科大学
刘丽亚	宁波大学	赵 星	四川大学
刘美娜	哈尔滨医科大学	郝元涛	北京大学
孙世权	西安交通大学	姚应水	皖南医学院
芦文丽	天津医科大学	秦国友	复旦大学
杨永利	郑州大学	秦家碧	中南大学
吴 莹	南方医科大学	顾 菁	中山大学
吴思英	福建医科大学	曹明芹	新疆医科大学
余红梅	山西医科大学	潘发明	安徽医科大学
张 涛	山东大学	魏永越	北京大学

编写秘书

武轶群 北京大学

数字编委



新形态教材使用说明



新形态教材是充分利用多种形式的数字资源及现代信息技术，通过二维码将纸书内容与数字资源进行深度融合的教材。本套教材全部以新形态教材形式出版，每本教材均配有特色的数字资源和电子教材，读者阅读纸书时可以扫描二维码，获取数字资源和电子教材。

电子教材是纸质教材的电子阅读版本，支持手机、平板及电脑等多终端浏览，具有目录导航、全文检索等功能，方便与纸质教材配合使用，随时随地进行阅读。

获取数字资源与电子教材的步骤

1 扫描封底红标二维码，获取图书“使用说明”。

2 揭开红标，扫描绿标激活码，注册/登录人卫账号获取数字资源与电子教材。

3 扫描书内二维码或封底绿标激活码随时查看数字资源和电子教材。

4 登录 zengzhi.ipmph.com 或下载应用体验更多功能和服务。

扫描下载应用

客户服务热线 400-111-8166

读者信息反馈方式



欢迎登录“人卫 e 教”平台官网“medu.pmph.com”，在首页注册登录后，即可通过输入书名、书号或主编姓名等关键字，查询我社已出版教材，并可对该教材进行读者反馈、图书纠错、撰写书评以及分享资源等。

修订说明

公共卫生与预防医学教育是现代医学教育的重要组成部分,在应对全球健康挑战、建设健康中国、提高国民健康素养、促进人群健康过程中,始终发挥着重要作用、承担着重大使命。在人类应对各种突发、新发传染病威胁过程中,公共卫生更是作用重大,不可或缺,都说明公共卫生学科专业的重要性与必要性。公共卫生不仅关系着公众的健康水平、公共安全和社会稳定,还影响着社会经济的发展和国际关系与世界格局的改变,是事关大国计、大民生的大学科、大专业。在我国公共卫生 40 余年的教学实践中也逐步形成了我国公共卫生与预防医学教育的一些特点。比如,我国的公共卫生教育是以强医学背景为主的公共卫生与预防医学教育,既体现了国家战略需求,也结合了本土化实践。现代公共卫生与预防医学教育强调“干中学”(learning by doing)这一主动学习、在实践中学习和终身学习的教育理念,因此公共卫生与预防医学教材建设和发展也必须始终坚持和围绕这一理念。

1978 年,在卫生部的指导下,人民卫生出版社启动了我国本科预防医学专业第一轮规划教材,组织了全国高等院校的知名专家和教师共同编写,于 1981 年全部出版。首轮教材共有 7 个品种,包括《卫生统计学》《流行病学》《分析化学》《劳动卫生与职业病学》《环境卫生学》《营养与食品卫生学》《儿童少年卫生学》,奠定了我国本科预防医学专业教育的规范化模式。此后,随着预防医学专业的发展和人才培养需求的变化,进行了多轮教材的修订、完善与出版工作,并于 1990 年成立了全国高等学校预防医学专业第一届教材评审委员会,至今已经是第五届。为了满足各院校教学的实际需求,规划教材的品种也在不断丰富。第二轮增加《卫生毒理学基础》《卫生微生物学》,第四轮增加《社会医学》,第五轮增加《卫生事业管理学》《卫生经济学》《卫生法规与监督学》《健康教育》《卫生信息管理学》《社会医疗保险学》,第八轮增加《公共卫生与预防医学导论》。由此,经过 40 余年的不断完善和补充,逐渐形成了一套具有中国本土特色的、完整的、科学的预防医学教材体系。

党的二十大报告提出“创新医防协同、医防融合机制,健全公共卫生体系”,我国新时代卫生健康工作方针明确坚持“预防为主”“将健康融入所有政策”,把公共卫生在国家建设发展中的基础性、全局性、战略性地位提到了空前高度。为贯彻落实党的二十大及二十届二中、三中全会精神,促进教育、科技、人才一体化发展,适应我国公共卫生体系重塑和高水平公共卫生学院建设的需要,经研究决定,于 2023 年启动了全国高等学校预防医学专业第九轮规划教材的修订工作。

预防医学专业第九轮规划教材的修订和编写特点如下:

1. 强化国家战略导向,坚持教材立德树人 教材修订编写工作认真贯彻落实教育部《高等学校课程思政建设指导纲要》,落实立德树人根本任务,以为党育人、为国育才为根本目标。在专业内容中融入思政元素,固本铸魂,阐释“人民至上、生命至上”的理念,引导学生热爱、专注、执着、奉献于公共卫生事业,打造政治过硬、心怀人民、专业能力强,既对国情有深刻理解,又对国际形势有充分认知,关键时刻能够靠得住、顶得上的公共卫生与预防医学专业人才队伍。

2. 培养公卫紧缺人才,坚持教材顶层设计 教材修订编写工作是在教育部、国家卫生健康委员会、国家疾病预防控制局的领导和支持下,由全国高等学校预防医学专业教材评审委员会审定,专家、教授把关,全国各医学院校知名专家教授和疾控专家共同编写,人民卫生出版社高质量出版。坚持顶层设计,按照教育部培养目标、国家公共卫生与疾控事业高质量发展的要求和社会用人需求,在全国进行科学调研的基础上,借鉴国内外公共卫生人才培养模式和教材建设经验,充分研究论证专业人才素质要求、学科体系构成、课程体系设置和教材体系规划。

3. 细化自强卓越目标,坚持教材编写原则 教材修订编写遵循教育模式的改革、教学方式的优化和教材体系的建设,立足中国本土,突出中国特色,夯实人才根基。在全国高等院校教材使用效果的调研、评价基础上,总结和汲取前八轮教材的编写经验和成果,对院校反馈意见比较集中的教材内容进行修改和完善。教材编写立足预防医学专业五年制本科教育,始终坚持教材“三基”(基础理论、基本知识、基本技能)、“五性”(思想性、科学性、先进性、启发性、适用性)和“三特定”(特定对象、特定要求、特定限制)的编写原则。

4. 深化数字科技赋能,坚持教材创新发展 为进一步满足预防医学专业教育数字化需求,更好地实现理论与实践结合,本轮教材采用纸质教材和数字资源融合的新形态教材出版形式。数字资源包括教学课件、拓展阅读、案例分析、实践操作、微课、视频、动画等,根据教学实际需求,突出公共卫生与预防医学学科特色资源建设,支持教学深度应用,有效服务线上教学、混合式教学等教学模式。

5. 全面服务教学育人,坚持教材立体建设 从第五轮教材修订开始,尝试编写和出版服务于教学与考核的配套教材,之后每轮教材修订时根据需要不断扩充和完善。本轮教材仍有10种理论教材配有学习指导与习题集、实习指导、实验指导类配套教材,供教师授课、学生学习和复习参考。

全国高等学校预防医学专业第九轮规划教材共17种,均为国家卫生健康委员会“十四五”规划教材。全套教材将于2025年出版发行,数字内容和电子教材也将同步上线。其他配套教材将于2026年陆续出版完成。另外,教育部公共卫生与预防医学“101计划”核心教材首轮共10种,也将同步出版,供全国广大院校师生选用参考。

希望全国广大院校在使用过程中能够多提宝贵意见,反馈使用信息,以便进一步修改和完善教材内容,提高教材质量,为第十轮教材的修订工作建言献策。

主编简介



郝元涛

1969年11月出生于贵州惠水。博士研究生导师，北京大学王宽诚讲席教授，国家“万人计划”教学名师。现任北京大学公众健康与重大疫情防控战略研究中心执行主任，担任教育部高等学校公共卫生与预防医学类专业教学指导委员会副主任委员、中华预防医学会生物统计分会主任委员、中华预防医学会公共卫生教育分会副主任委员。

从事教学工作近30年，主编或副主编教材与专著9部，主持国家级精品课程3门，建设教育部虚拟教研室1个，获得省级教学成果奖一等奖1项。主要研究领域包括传染病监测预警体系与模型、健康测量指标体系、基于现实世界数据的因果分析方法及其应用等。主持国家科技重大专项及国家自然科学基金项目7项，以通信作者身份发表高水平学术论文百余篇。



刘美娜

1962年5月出生于黑龙江哈尔滨。二级教授，博士研究生导师。现为中国卫生信息与健康医疗大数据学会卫生统计学教育专业委员会副主任委员、中国卫生信息与健康医疗大数据学会统计理论与方法专业委员会常务委员、中国康复医学会科技管理与评估委员会第二届委员会副主任委员。

自1984年开始，主要从事卫生统计学教学和科研工作。承担多轨道的医学统计学、卫生统计学、卫生管理统计与软件应用课程的教学任务；曾获全校教师讲课大赛第一名、黑龙江省优秀教师、学校“先进教师”、学院“先进个人”等称号；主编、副主编规划教材8部。科研中将统计理念和数理统计方法应用于疾病负担、发病预测、疾病治疗质量评价等社会关注点，理论联系实际，拓展了统计学应用的新领域；目前科研方向是疾病治疗质量综合评价模型构建及因果推断方法研究，获得国家自然科学基金项目4项，主持国家重大专项子课题3项。发表学术论文150余篇，其中SCI收录48篇。

副主编简介



赵 杨

1978 年 7 月出生于江苏如皋。教授, 博士研究生导师。现任南京医科大学公共卫生学院副院长、生物统计学系主任。兼任中国卫生信息与健康医疗大数据学会卫生统计学教育专业委员会主任委员, 中国临床肿瘤学会(CSCO)生物统计学专家委员会副主任委员, 中华预防医学会生物统计分会副主任委员。

研究方向包括复杂生物医学数据整合分析及因果推断方法研究、临床试验中的统计方法、真实世界研究中的统计理论与方法。曾在 *IJE*、*BMC Medicine*、*AJRCCM* 等期刊以第一作者或通信作者发表学术论文 80 余篇。先后主持国家自然科学基金项目 5 项, 国家重点研发计划课题 1 项。



王超龙

1985 年 6 月出生于福建南安。二级教授。现任华中科技大学同济医学院公共卫生学院副院长。国家杰出青年科学基金获得者, 入选国家级青年人才项目。

从事教学工作近 10 年, 曾任职哈佛大学生物统计系博士后, 新加坡基因组研究所研究员, 杜克-新加坡国立大学兼职助理教授。长期围绕重大疾病精准防控需求, 融合统计建模与生物信息学等技术开展群体遗传学与流行病学的方法与应用研究, 作为第一作者或通信作者在包括 *Cell*、*Nature*、*JAMA* 等学术期刊发表论文 50 余篇, 获钟南山青年科技创新奖、霍英东教育基金会高等院校青年科学奖等荣誉。



姚应水

1972 年 5 月出生于安徽铜陵。二级教授, 博士研究生导师。现任皖南医学院院长, 公共卫生与预防医学一级学科带头人。兼任首届全国中小学和高校健康教育教学指导委员会委员, 《中华疾病控制杂志》常务编委、《中华预防医学杂志》通信编委、《中国临床药理学与治疗学》常务编委和《皖南医学院学报》主编。

从事教学工作至今 30 年。为享受省政府特殊津贴专家、安徽省学术和技术带头人、安徽省“特支计划”领军人才、安徽省教学名师。主持国家自然科学基金项目 4 项, 省部级科研项目 16 项, 发表论文 200 余篇。主编教材 14 部、副主编 30 部。以第一完成人获安徽省教学成果奖特等奖 2 项。

前言

作为国家卫生健康委员会“十四五”规划教材、全国高等学校预防医学专业第九轮规划教材的《卫生统计学》(第9版),从1978年第1版问世至今,历时47年,发展到第9版,成为一本经久不衰的经典教材,为公共卫生领域培养了无数专业人才。回望历史,历版教材的主编、副主编、编委、教材秘书等都为教材发展和质量提升付出了艰辛努力,教材编写会上大家讨论交流的许多场景此刻浮现脑海,历历在目。原华西医科大学的杨树勤教授是第1、2、3版的主编,为本教材的建设奠定了坚实基础,今年恰逢杨树勤教授诞辰100周年,在此谨向杨教授致以崇高敬意!原华西医科大学的倪宗赞教授是第4版的主编,他在继承前三版优良传统的基础上进一步促进了教材发展。中山大学的方积乾教授是第5、6、7版的主编,适逢计算机和统计软件蓬勃发展之际,方教授与时俱进,引领教学改革,将本教材带向一个新的水平。四川大学的李晓松教授是第8版主编,他在教材编写中融入了大数据时代正确认识数据的重要性,对教材的内容和体系进行了有益的改革与创新。本版教材希望在前八版的基础上,力求实现传承与发展,促进卫生统计学教育在新时代知识创新中不断进步。

本书是全国高等学校预防医学专业本科规划教材,我们在修订和编写过程中严格遵循“三基”“五性”“三特定”的原则。其中,“三特定”中的特定对象,指本教材的主要读者是开设本科预防医学专业的院校所招收的五年制本科生;特定目标指培养“立地”人才,即本专业通用型高级人才;特定限制指字数与学时匹配,避免教材“越编越厚”。

根据上述修订与编写原则,本教材共包含19章,系统介绍了卫生统计学的基础理论、基本知识和基本技能,内容涵盖了国家执业医师资格考试的相关知识点,例如统计学的基本概念、定量和定性变量的统计描述方法、常见的概率分布、参数估计与假设检验基础、 t 检验、方差分析、 χ^2 检验、基于秩次的非参数检验、相关与回归、多重线性回归、logistic回归、Cox回归等常用分析方法,以及调查研究和实验研究的基本理论和方法。此外,本教材还介绍了当前卫生统计学领域的两个热门方向:一个是贝叶斯统计,另一个是人工智能技术与医学应用。第9版教材采用纸质教材与优质数字资源整合的新形态教材形式,我们将R语言介绍以及各章统计分析方法的R代码整合在数字资源中,供读者参考使用。

前言

全书的内容和编排体现了基础化、系统化、体系化的特点。这里,“基础化”是指本书所涵盖的均为经典的、基础的卫生统计学理论、知识和方法,具有“立地”的作用;“系统化”是指本书章节的顺序逻辑严密,由浅入深,层层递进,老师易教,学生易学;“体系化”则得益于历版编委的共同努力,使得卫生统计学的知识体系不断完善成熟,逐步形成了科学性与思想性兼具的学科架构。

本版教材于2024年5月28日在皖南医学院召开编写会,于2024年10月31日在北京大学召开定稿会。本教材能够在相对较短的时间里完成修订和编写,离不开全体编委的倾心付出,离不开皖南医学院和北京大学的大力支持,以及众多领导、老师和同学的关心与帮助。在此,向大家致以最诚挚的感谢!

尽管全体编委同心协力,期待本教材成为一本既有传承又有发展的精品教材,但囿于时间、能力等因素,恐错漏难免,恳请各位同行和读者不吝指正,以助教材进一步改进和完善。

郝元涛 刘美娜

2025年1月

目 录

第一章 绪论	1
第一节 卫生统计学的主要内容和作用	1
一、卫生统计学的定义与作用	1
二、卫生统计学工作的步骤	1
第二节 卫生统计学的若干基本概念	3
一、总体与样本	3
二、同质与变异	3
三、变量的类型	3
四、参数与统计量	4
五、设计与分析	4
六、误差	5
七、频率和概率	5
八、因果与关联	6
第三节 卫生统计学发展简史及未来	6
一、医学中统计思维的进化	6
二、统计学与公共卫生的相互促进	9
三、卫生统计学的未来	10
第四节 学习卫生统计学应注意的问题	11
一、学习目标与定位	11
二、基本概念、方法与技能	11
三、教与学的方法	12
四、本教材的内容安排	12
第二章 定量变量的统计描述	15
第一节 频数分布表与频数分布图	15
一、频数分布表(图)的编制	15
二、频率分布表(图)的用途	16
第二节 描述集中趋势的统计指标	18
一、算术均数	18
二、几何均数	19
三、中位数	21
第三节 描述离散趋势的统计指标	22
一、极差	23
二、四分位数间距	23

三、方差	24
四、标准差	25
五、变异系数	25
第四节 统计表和统计图	26
一、统计表	26
二、统计图	28
第三章 定性变量的统计描述	36
第一节 定性变量的频率分布	36
一、多分类变量的频率分布	36
二、二分类变量的频率分布	37
第二节 常用相对数指标	38
一、频率型指标	38
二、强度型指标	39
三、相对比型指标	40
四、应用相对数的注意事项	40
五、动态数列及其指标	41
第三节 率的标准化法	43
一、标准化法的意义和基本思想	43
二、标准化率的计算	44
三、应用标准化法的注意事项	46
第四节 医学人口健康统计常用指标	46
一、医学人口统计资料的来源	47
二、描述人口学特征的常用指标	47
三、生育和人口死亡的常用指标	48
四、疾病统计常用指标	51
第四章 常用概率分布	55
第一节 正态分布	55
一、正态分布的概念与特征	55
二、正态概率密度曲线下的面积	57
三、正态分布的应用	60
第二节 二项分布	62
一、二项分布的概念与特征	62
二、二项分布的正态近似	66
三、二项分布的应用	67
第三节 泊松分布	67
一、泊松分布的概念与特征	67
二、泊松分布的正态近似	69

三、泊松分布的应用	69
第四节 蒙特卡洛模拟	71
一、蒙特卡洛模拟的基本思想	71
二、常见分布的模拟抽样	71
三、蒙特卡洛模拟的步骤与应用	72
第五章 参数估计与假设检验基础	77
第一节 抽样分布与抽样误差	77
一、标准误	77
二、均数抽样误差的规律性	78
三、频率的抽样分布与抽样误差	80
第二节 t 分布	82
一、 t 分布的概念	82
二、 t 分布的图形与特征	82
第三节 参数估计	84
一、参数估计的基础理论	84
二、总体均数的区间估计	85
三、总体概率的区间估计	86
四、事件数的区间估计	86
第四节 假设检验的概念与原理	87
一、假设检验的思维逻辑	87
二、假设检验的基本步骤	88
三、假设检验中的两类错误	89
四、假设检验的正确应用	89
第五节 假设检验与区间估计的关系	90
第六章 t 检验	94
第一节 单样本 t 检验	94
第二节 配对 t 检验	95
第三节 两独立样本 t 检验	97
一、两样本均数之差的抽样分布	98
二、两样本所属总体方差相等(即具有方差齐性)	99
三、两样本所属总体方差不等(Satterthwaite 近似法)	100
四、两总体均数之差的置信区间	101
第四节 两独立样本资料的方差齐性检验	103
第五节 正态性检验	104
一、图示法	104
二、统计检验法	105
第六节 大样本资料的 Z 检验	105

一、单样本资料的 Z 检验	105
二、两独立样本资料的 Z 检验	106
第七章 方差分析基础	110
第一节 方差分析的基本思想与前提条件	110
一、方差分析的基本思想	110
二、方差分析的前提条件	114
第二节 完全随机设计资料的方差分析	114
一、离均差平方和与自由度的分解	115
二、完全随机设计资料方差分析的基本步骤	115
第三节 随机区组设计资料的方差分析	116
一、离均差平方和与自由度的分解	116
二、随机区组设计资料方差分析的基本步骤	116
第四节 均数的多重比较	118
一、SNK 法	119
二、Dunnett 法	120
第五节 方差齐性检验	121
一、Bartlett χ^2 检验	121
二、Levene 检验	121
三、残差图	122
四、数据变换	123
第八章 χ^2 检验	127
第一节 两个独立样本四格表资料的 χ^2 检验	127
一、 χ^2 检验的基本思想	127
二、四格表资料的 χ^2 检验	129
三、确切概率算法	131
第二节 多个独立样本列联表资料的 χ^2 检验	133
一、多个率比较的 χ^2 检验	134
二、多个频率分布比较的 χ^2 检验	134
三、多个率或频率分布的多重比较	135
第三节 配对设计资料的 χ^2 检验	136
一、配对设计四格表资料的 χ^2 检验	136
二、多分类配对列联表的 χ^2 检验	138
第四节 分类变量的独立性检验	139
一、独立性检验的基本思想	139
二、独立性检验的基本步骤	140
第五节 频数分布拟合优度的 χ^2 检验	142
一、频数分布的拟合优度	142

二、频数分布拟合优度的 χ^2 检验	142
第九章 基于秩次的非参数检验	147
第一节 非参数检验的应用	147
一、非参数检验的基本概念	147
二、非参数检验的优势和不足	147
第二节 单样本和配对设计资料的符号秩和检验	148
一、单样本数据的符号秩和检验	148
二、配对设计资料的符号秩和检验	149
第三节 两组独立样本比较的秩和检验	151
一、两组定量资料比较	151
二、两组等级资料比较	154
第四节 多组独立样本比较的秩和检验	155
一、多组定量资料比较	155
二、多组等级资料比较	157
三、多组独立样本资料的多重比较	158
第五节 随机区组设计资料的秩和检验	160
一、随机区组设计的定量资料比较	160
二、随机区组设计资料的多重比较	161
第十章 两变量的相关与回归分析	166
第一节 线性相关	166
一、线性相关的概念	166
二、线性相关系数	167
三、相关系数的统计推断	169
四、线性相关分析时应注意的问题	171
第二节 等级相关	171
一、秩相关的概念	171
二、秩相关系数的假设检验	172
三、分类变量的关联性分析	172
第三节 线性回归	173
一、线性回归模型	173
二、线性回归模型的参数估计	175
三、线性回归模型的假设检验	176
四、线性回归分析的应用	179
第四节 线性回归与相关的区别与联系	182
一、区别	182
二、联系	183

第十一章 多重线性回归	187
第一节 多重线性回归模型	187
一、数据与模型	187
二、偏回归系数的估计	188
第二节 多重线性回归的假设检验	189
一、整体回归效应的假设检验(方差分析)	189
二、偏回归系数的 t 检验	189
三、决定系数与复相关系数	190
四、偏相关系数	191
第三节 多重线性回归的自变量筛选	191
一、自变量筛选的标准与原则	191
二、自变量筛选的常用方法	193
第四节 多重线性回归的应用	194
一、多重线性回归应用的注意事项	195
二、案例讨论与结果报告	200
第十二章 logistic 回归分析	206
第一节 logistic 回归模型	206
一、logistic 回归模型的基本形式	206
二、logistic 回归模型参数的意义	207
第二节 logistic 回归模型的参数估计与假设检验	208
一、logistic 回归的参数估计	208
二、logistic 回归的参数假设检验	210
三、logistic 回归系数的区间估计	211
四、logistic 回归预测	211
第三节 其他常用的 logistic 回归模型	211
一、条件 logistic 回归模型	211
二、多分类 logistic 回归模型	214
第四节 logistic 回归的应用及注意的问题	215
一、logistic 回归的应用	215
二、logistic 回归应用中需注意的问题	216
第十三章 生存分析	220
第一节 生存分析的基本概念	220
一、生存时间与删失	220
二、生存概率与死亡概率	222
三、生存函数与风险函数	222
四、生存曲线与中位生存期	223

第二节 生存率的估计	223
一、Kaplan-Meier 法	223
二、寿命表法	225
三、生存率的区间估计	226
第三节 生存曲线的比较	227
一、log-rank 检验	227
二、log-rank 检验应用注意事项	228
第四节 Cox 回归模型及应用	230
一、Cox 回归模型	230
二、Cox 回归模型的实例应用	232
三、Cox 回归应用的注意事项	235
第十四章 调查研究设计	239
第一节 调查设计的基本内容	239
一、调查目的	239
二、调查对象与观察单位	240
三、调查方法与调查项目	240
四、数据整理分析计划	244
第二节 基本的概率抽样方法及其样本量估计	245
一、简单随机抽样	245
二、系统抽样	247
三、分层抽样	247
四、整群抽样	249
五、多阶段抽样	251
第三节 非概率抽样方法	253
一、非概率抽样方法的应用	253
二、常用的非概率抽样方法	253
第四节 调查问卷的质量评价	255
一、效度的概念及其评价指标	255
二、信度的概念及其评价指标	255
三、信度与效度的关系	256
第十五章 实验研究设计	259
第一节 实验设计的基本内容与原则	259
一、研究目的	259
二、实验设计的基本要素	260
三、实验研究的基本原则	262
四、质量控制	265
第二节 常用的实验设计类型及其样本量估计	266

一、完全随机设计	266
二、配对设计	268
三、随机区组设计	269
四、析因设计	270
五、重复测量设计	271
第三节 临床试验设计	272
一、基本概念	272
二、临床试验设计的基本类型	273
三、临床试验的三种试验类型	274
四、统计分析计划	275
第十六章 寿命表	280
第一节 寿命表的基本概念及用途	280
一、寿命表的基本概念	280
二、寿命表的用途	282
第二节 寿命表的编制方法	283
一、编制思想	283
二、数据来源	283
三、寿命表编制过程	283
第三节 去死因寿命表	287
第四节 健康预期寿命与伤残调整寿命年	289
一、健康预期寿命	289
二、伤残调整寿命年	291
第十七章 meta 分析	298
第一节 meta 分析概述	298
一、问题的提出	298
二、meta 分析的含义	298
三、meta 分析的用途	299
四、meta 分析的适用性	300
第二节 meta 分析的基本步骤与方法	300
一、meta 分析的基本步骤	300
二、meta 分析的统计方法	302
三、森林图的解读	311
第三节 meta 分析的偏倚及注意事项	311
第四节 meta 分析的进展及分析软件介绍	313
第十八章 贝叶斯统计简介	317
第一节 贝叶斯统计理论的基本观点与现状	317

一、贝叶斯学派与频率学派	317
二、贝叶斯统计的发展与现状	319
第二节 先验分布与后验分布	320
一、贝叶斯公式	320
二、先验分布	322
三、后验分布	327
第三节 贝叶斯统计推断	328
一、似然函数的概念与原理	328
二、贝叶斯参数估计	329
三、贝叶斯假设检验	330
第四节 贝叶斯统计应用	331
一、贝叶斯回归	331
二、贝叶斯决策	331
第十九章 人工智能技术与医学应用	335
第一节 人工智能简介	335
一、人工智能发展历程	335
二、人工智能与统计学	335
第二节 机器学习简介	337
一、机器学习的基本概念	337
二、机器学习的常见方法	337
三、机器学习的基本步骤	341
第三节 深度学习简介	345
一、深度学习的基本概念	345
二、常见的深度学习方法	346
三、大语言模型简介	351
第四节 人工智能与医学应用	353
一、人工智能与精准预防	353
二、人工智能与精准诊疗	353
三、人工智能与精准健康管理	354
四、人工智能与传染病精准防控	354
五、人工智能与精准药物研发	355
六、人工智能与精准全科医学服务	356
第五节 机遇与挑战	356
附录 统计用表	360
附表 1 标准正态分布界值表	360
附表 2 t 界值表	361
附表 3 二项分布概率 π 的置信区间	363

附表 4	F 界值表	373
附表 5	χ^2 分布界值表	385
附表 6	T 界值表(用于配对比较的 Wilcoxon 符号秩和检验)	386
附表 7	T 界值表(两样本比较的秩和检验用)	388
附表 8	H 界值表(用于三组比较的 Kruskal-Wallis H 检验)	390
附表 9	M 界值表(用于随机区组设计的 Friedman 秩和检验)	393
附表 10	r 界值表(用于 Pearson 相关系数检验)	394
附表 11	r_s 界值表(用于 Spearman 相关系数检验)	395
附表 12	随机数字表	397
附表 13	Ψ 界值表(用于多个样本均数比较时计算所需样本例数)	399
推荐阅读		402
中英文名词对照索引		403

统计软件数据包



第一章 绪论



统计学是一门既古老又现代的学科。在古代,人们清点并记录国家和地区的人口、财产、土地、军队等数量,对之进行分析,统计学应运而生。统计学的英文单词是 statistics,其前4个字母组成的词根 stat 就是国家和地区的意思,体现了该学科的起源。后来,统计学不断应用和发展,特别是吸纳了数学、概率论等理论和方法之后,统计学成为了一个具有完备知识体系的独立学科。现在,统计学得到广泛应用,与不同的应用领域相结合,形成了许多应用型的学科分支,例如:经济统计学、社会统计学、卫生统计学等。当今方兴未艾的人工智能,其核心就是统计学、计算机科学和大数据。

第一节 卫生统计学的主要内容和作用

一、卫生统计学的定义与作用

什么是统计学(statistics)?权威词典《韦氏国际词典》(Webster's International Dictionary)说,它是“a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data”。医学主题词表 MeSH(Medical Subject Headings)将统计学定义为“The science and art of collecting, summarizing, and analyzing data that are subject to random variation”。从上述权威定义中可以认识到,统计学是一门处理数据中变异性的科学与艺术,内容包括收集、分析、解释和呈现数据,目的是求得可靠的结果。

卫生统计学是统计学的一门重要分支学科,是统计学的原理和方法在公共卫生以及大健康领域的应用,它通过对医疗健康相关数据的收集、整理、分析与解释,处理公众健康中的不确定性问题,帮助人们回答诸如以下的问题:某病的发病率、患病率和死亡率分别是多少?某种干预措施或者治疗手段的效果究竟如何?不同药物的疗效是否存在差别?影响疾病发生发展的因素有哪些,其作用有多大?某病患者的总体生存状况如何?5年生存率是多少?某国某年的人群预期寿命是多少?特定的疾病对预期寿命的影响有多大?不同疾病或危险因素带来的疾病负担有多大?如何进行较优的资源配置、做好卫生决策?要解决上述问题,采用何种设计才是最科学高效的?凡此种种,均需要卫生统计学的理论、知识和方法。当前几乎所有医学专业学术期刊的文章都用统计图表呈现研究结果,反映出统计学已渗透到医学的方方面面,说明其在医学中的重要性。

二、卫生统计学工作的步骤

与其他科学研究类似,卫生统计学工作常常包括以下四个步骤。

(一) 统计设计

开展任何科学研究,设计是第一步的工作,也是研究成功的关键。如果没有好的设计,再严格的数据收集和再优秀的统计分析者都于事无补。科学的设计是研究结果有效可信的重要保证。统

计设计主要关注如何有效地收集、分析和解释数据,以确保研究结果的可靠性和有效性。一般而言,在开展研究之前,研究者必须通过文献综述了解研究的现状和前沿,明确本次研究的科学问题和目的,清晰定义研究对象和观察单位的纳排标准,选取恰当的抽样方法和样本量,如果是干预性研究,需要确定标准的干预措施,选择合适的对照。此外,在设计方案中,还要明确:需要收集哪些资料?如何获取这些资料?如何对资料进行整理汇总?需要计算哪些统计指标?采用什么统计分析方法?如何控制误差并做好质量控制?预计得到什么样的结果?需要多少人力、物力、财力?如何分工?上述问题都需要周密考虑、统筹安排。试想,如果要了解某年某地区成年人高血压患病情况,在设计方案中应该考虑哪些数据收集、分析和解释的内容?再设想,如果要在人群中观察某种新研制疫苗的效果,设计方案中应该包括哪些数据收集、分析和解释的主要内容?

本教材中调查研究设计和实验研究设计两章将系统介绍统计设计的内容。

(二) 资料收集

资料收集指采取各种方法收集准确可靠的数据。数据的收集方式有多种,在专项的科学研究中,根据设计方案的要求,可以通过访谈、问卷调查、体格检查、临床检查、生理生化检验、分子生物学检查等方式获得所需资料。另外,医疗卫生工作的统计资料还可以来源于统计报表,如法定的传染病报表、职业病报表等;经常性工作记录,最常见如健康体检记录、医院门诊病历以及住院病历等,这些是公共卫生执业医师与临床执业医师常用的研究数据来源之一;还有各种统计年鉴和统计数据专辑。以上途径都是医学研究数据的主要来源方式。此外,根据研究目的和内容,还可以收集气象、环境、人口社会经济发展等方面的数据。数据的来源不同,其收集数据的方法和过程也不尽相同,但是所有的数据收集过程均要力争准确、完整并且及时。

(三) 资料整理

资料整理是根据研究的目的,运用科学的方法,对所获得的资料进行审查、核对、分类、汇总等初步加工,使之净化、系统化,以便于下一步的计算分析。所谓净化,是指对原始数据的清理、检查、核对和纠错等。数据收集完成后,需要进行逻辑性检查,例如,当出现年龄超过100岁、身高超过2.2米的调查对象时,需要核实。所谓系统化,是指根据研究目的,将数据归类、整理或加工,以便分析时使用。例如,要比较不同地区、不同年龄段、不同性别的某种疾病的患病率时,需要汇总不同组别的数据,以便后期分析时直接利用。

(四) 资料分析

资料分析的主要目的在于通过运用合适的统计分析方法,表达数据特征并阐明事物的本质规律。统计分析包括两方面的工作:一是统计描述,运用恰当的统计指标、统计图、统计表,对资料的特征和分布规律进行描述。比如,分析每一组观察对象的年龄、性别分布、血压情况、既往病史等,以便了解不同组的观察对象之间的可比性。二是统计推断,指利用有限的样本信息,推断总体的特征并作出相应估计或决策的过程。例如,通过样本算得的高血压患病率推断该样本来源总体的患病率,通过组间样本均数的差别推断其总体均数是否相同。在抽样研究中,通过样本信息推断总体特征是必需的步骤,是回答研究假设的最终环节。

如何呈现资料统计分析的结果,以直观形象准确地表达给专业或非专业的读者?这也是资料分析步骤中需要考虑的重要内容,现在常称之为结果的可视化。

卫生统计工作主要分为以上四个步骤,实际工作中,它们紧密联系、环环相扣,任何一步出错,都会影响整个研究的结果。因此做好每一步的质量控制非常必要。统计设计的步骤至关重要,对整个研究过程起到统领的作用。

第二节 卫生统计学的若干基本概念

一、总体与样本

统计学家用总体(population)这个术语来表示根据研究目的确定的全体研究对象。例如,一个国家的所有成年人,某地的所有小学生。基于研究对时间、空间、人群范围的规定和研究对象的数量受限与否,可将总体分为有限总体和无限总体。另外,如果试图就某个总体下结论,这个总体便称为目标总体(target population);资料常来源于目标总体中的一个部分,称为研究总体(study population)。例如,关于吸烟与肺癌的研究以英国成年男子为目标总体;1951年英国全部注册医师只是成年男子中的一部分,是研究总体。需要谨慎的是,就研究总体所下的结论未必适用于目标总体。

研究总体中的个体(individual)往往很多,甚至无限多。一个不漏地观察其中的所有个体常常是不可能的;有时即使可能,也没有必要。科学的办法是从研究总体中抽取少量有代表性的个体,这个过程称为抽样(sampling),抽取的个体构成一份样本(sample);对这些个体组成的样本进行深入的观察与测量,获取数据(data);利用统计学方法,透过样本数据对研究总体的规律进行推断(inference)。

统计推断的基础是概率理论。如果某事件的结局具有多样性,事先不能肯定发生哪种结局,称之为随机事件。人们凭借样本数据来推断各种可能结局发生的概率。既然是推断,既然是由部分推断全体,统计学的结论从来就不是完全肯定或完全否定的。

能否成功达到从样本推断总体的目的,关键是抽样的方法、样本的代表性和推断的技术,这些是统计学的核心内容。

二、同质与变异

一个总体中有许多个体,它们之所以共同成为人们研究的对象,必定存在共性。说一些个体处于同一总体,就是指它们大同小异,具有同质性(homogeneity)。然而,同一总体内的个体间存在差异又是绝对的,这种现象就是变异(variation)。例如,同性别、同年龄的小学生具有同质性,属于同一个总体;但他(她)们的身高、体重又存在变异。没有同质性就不构成一个总体供人们研究,总体内没有变异性就无须统计学。

统计学的任务就是在变异的背景上描述同一总体的同质性,揭示不同总体的异质性(heterogeneity)。例如,为研究同性别、同年龄的中国小学生和日本小学生的平均身高是否不同,从两个总体中各抽取一份样本,各得一个平均数,数值不同,能不能就此推断两国同性别、同年龄小学生的平均身高不等?试想,如果再从中国抽取一份样本,再得一个平均数,数值也与前不同,也许你会认为这是同一总体的个体大同小异造成的。那么,日本的均数与中国的均数不相等,是因为小学生“大同小异”,还是因为两个总体本不相同?这就要靠统计学来回答。

三、变量的类型

总体中个体的特性总是通过一个或多个数量指标来描述,变异性的存在决定了要处理的是变量(variable)。识别变量的类型非常重要,因为不同类型的变量需用不同的统计学技术去分析。变量分类的方法很多,详细的讨论超出本课程的范围。本书把变量分成定性(qualitative)与定量(quantitative)两种类型。

定性变量中最常见的是分类变量(categorical variable)或名义变量(nominal variable)。例如,职业是一个分类变量,其可能的“取值”不是数字,而是工、农、商、学、兵等类别,这些被称为分类变量的水平(level);为便于输入计算机也可以采用代码(code)1、2、3、4、5等来表示各个水平。这些数仅仅是代码,不能进行计算。最简单也最常用的分类变量是二分类变量(binary variable),例如,性别(男女)、疾病(有无)和结局(生死)等。二分类变量常用0和1来编码,也可以称为0-1变量。与一般分类变量不同,0-1变量常称为哑变量(dummy variable)或假变量,其值可以直接参与计算。除了二分类变量之外,还有多分类变量。一个多分类变量可用几个哑变量来取代。例如,含工、农、商、学、兵5个水平的职业变量可转化为4个哑变量: $X_1=1$ 表示“工”,0表示“非工”; $X_2=1$ 表示“农”,0表示“非农”; $X_3=1$ 表示“商”,0表示“非商”; $X_4=1$ 表示“学”,0表示“非学”。这样, $\{X_1, X_2, X_3, X_4\} = \{1, 0, 0, 0\}$ 就代表“工”, $\{X_1, X_2, X_3, X_4\} = \{0, 1, 0, 0\}$ 代表“农”…… $\{X_1, X_2, X_3, X_4\} = \{0, 0, 0, 0\}$ 则代表“兵”。

另一类定性变量是有序变量(ordinal variable)或等级变量。不同于分类变量,它种种可能的“取值”中自然地存在着次序。例如,问卷调查常问对某件事情的满意程度,给出5项答案,请对象挑选:极不满意、不满意、满意、很满意、极满意。有些临床体检或实验室检验常用-、±、+、++和+++来表示测量结果,也属于有序变量。

定量变量可以分为两种类型:离散型变量(discrete variable)和连续型变量(continuous variable)。离散型变量只能取整数数值。例如,一个月中的手术患者数、一年里的新生儿数。连续型变量可以取实数轴上的任何数值。有些变量的数值由测量而得到,它们大多属于连续型变量,例如,血压、身高、体重等。“连续”是指该变量可以在实数轴上连续变动。有一些测量值,诸如红细胞计数,虽然以“个”为单位时只能取整数数值,但其数值很大,当以“千”或“万”为单位时,又可以取小数,所以通常把这些变量也视为连续型变量。

有时为了数据分析的方便,人们将一种类型的变量转化为另一种类型。但变量只能由“高级”(包含信息较多)向“低级”(包含信息较少)转化:定量→有序→分类→二分类/哑变量;不能作相反方向的转化。在定量变量中,离散型变量常常通过适当的变换或连续性校正后借用连续型变量的方法来分析;此外,为了使实际意义格外突出,有时也可以将连续型变量离散化。

四、参数与统计量

参数(parameter)是描述总体特征的指标。例如,某年某市14岁健康男孩的平均体重就是一个参数,某年某市成年女性中糖尿病的患病率也是一个参数。常常用希腊字母表示参数,例如用希腊字母 μ 表示总体均数,用 π 表示总体率。由于总体往往是无限的,或者是庞大的有限总体,不可能对总体中的每一位个体进行观察和测量,所以参数虽然是一个客观存在,但往往是未知的,需要通过样本来估计。通过样本计算出来的反映样本特征的指标称为统计量(statistic),如通过随机抽取1000名14岁的健康男孩,测量每一位男孩的体重,然后计算其平均体重,这个平均值就是一个样本统计量。样本均数常常用 $\bar{x}$ 表示。统计学关心的常常是总体参数,而参数又很难获得,实际工作中通常用样本统计量来估计未知的总体参数。

五、设计与分析

最简单的研究是收集一些有关卫生业务报表的常规统计资料,进行分析后写出报告。常规资料的好处是容易得到,并有一定的权威性。缺点是常常不适合特定研究的目的、危险因素记录

不全、不同地区和时间统计的“口径”不一致。更重要的是,常规资料往往是合计数据(aggreated data),诸如一个地区的百分数、平均数或总和。每当使用合计资料时,必须警惕,在合计资料中观察到的关系,在个体水平往往不存在。

和盖高楼、建大桥一样,任何科学研究必须精心设计在前。正如统计学家罗纳德·A·费希尔(Ronald·A·Fisher, 1890—1962)指出的,统计学方面的设计(design)是医药卫生科研设计不可或缺的部分。设计不仅要符合统计学原则,运用统计学方法和技术,而且在设计的时候要明确日后分析数据使用的统计方法。只有明确了设计的样式和分析的方法才得以进一步考虑数据应当如何收集、样本量应当多大。

特定的设计样式决定了特定的数据分析方法;不同设计下获得的资料常常要用不同的方法来分析。例如,以后将学到的方差分析(analysis of variance, ANOVA)是一大类统计方法,随机区组设计(randomized block design)资料的方差分析不同于析因设计(factorial design)资料的方差分析,更不同于回归分析(regression analysis)中的方差分析。又如,同是流行病学中的病例对照研究(case-control study),成组对照的资料和个体对照的资料形式上都是一个 2×2 表,但分析方法却不同。

由上可见,统计设计和统计分析是不可分割的两项内容。在学习和应用统计学时需要将两者结合起来。

六、误差

误差(error)是指实测值与真实值之差。按照其产生的原因可以分为随机误差和非随机误差两大类。

随机误差是各种偶然的、无法控制的因素引起的随机变化的误差。例如,在测量血压的时候,对同一个体进行连续多次测量,由于个体生理、心理的微小变化、测量人员技术不稳定、环境等原因,每次测量的结果可能不完全相同。由于影响因素太多、太复杂且不可控,所以随机误差不可避免,但是可以通过多次测量了解其分布规律,估计其大小;还可以通过增加测量次数求平均值,来降低测量结果随机误差的大小。通过了解随机误差的规律,利用统计学的方法可以估计随机误差的大小。

非随机误差又称为系统误差,也叫偏倚(bias),是由于测量仪器或人为因素等导致的实际测量值系统地偏离真实值的情况。不同于随机误差,系统误差需要事先了解其原因并通过周密的研究设计、规范的技术措施、严格的质量控制进行消除或控制。

七、频率和概率

医学研究中的许多现象都是随机现象,例如采用某种药物治疗患者,结局可能是治愈,也可能是没有治愈。统计学上用随机事件表示这种现象。在 n 次独立重复实验中,随机事件 A 出现 f 次,则称 f/n 为随机事件 A 出现的频率(frequency)。当 n 不断增大,频率 f/n 在一个常数附近作微小波动,则称该常数为随机事件 A 出现的概率(probability),可以记为 $P(A)$,简记为 P 。样本的频率总是围绕总体概率波动的,当样本量 n 较大的时候,频率波动较小,并且接近于总体概率,但是当样本量较小时,用频率估计概率是不可靠的。

概率用来度量随机事件发生的可能性大小,它的值介于0与1之间。 P 越接近0,表示事件发生的可能性越小; P 越接近1,表示事件发生的可能性越大。当 $P=0$ 时,称之为不可能事件;当 $P=1$ 时,称之为必然事件;不可能事件和必然事件可以视为随机事件的特例。另外,当随机事件发生的概率 $P \leq 0.05$ 或者 $P \leq 0.01$ 时,统计学上称之为小概率事件。对于小概率事件,人们认为在一次实

验、一次抽样中发生的可能性很小,可以认为不会发生;如果发生了,人们就会好奇,怎么这么巧?哪里出问题了?这是统计推断的理论依据。

八、因果与关联

公共卫生领域常常要探究疾病与危险因素(risk factor)之间的因果关系(causality)。例如,吸烟是不是肺癌的原因之一?探究因果关系的第一步应当确定是否存在关联(association)。然而,存在关联未必有因果关系。

如果两个变量之间的关系不受其他变量的影响,那么这两个变量之间的关联与因果关系相去不远。然而,事实上,每一个变量和其他无数变量都有关系。后者常常扮演混杂因素的角色。对混杂因素的存在视而不见,孤零零地考察两个变量之间的关系常常得出荒谬的结论。仅举一例:某君而立之年喜得贵子,栽一棵小树留作纪念,每月测子高和树高,发现两者密切联系,奉为神树……

正因为不可能完全控制所有的、重要的混杂因素,单靠人群调查和数据的统计学处理大多只能考察变量之间的关联。证明因果关系是十分困难的,目前因果分析已经成为统计学中的一个热点研究领域,在解释统计分析的结果以及下结论时务必对“因果”二字慎之又慎。

第三节 卫生统计学发展简史及未来

一、医学中统计思维的进化

医学研究者们普遍认同统计思维及其应用,这是统计学经过近200年进化并与医学科学不断结合的结果。

(一) 拉普拉斯的远见与路易斯的实践

法国数学家皮埃尔·西蒙·拉普拉斯(Pierre-Simon Laplace, 1749—1827)认为“概率论(probability theory)可以应用于整个人类知识系统,因为发现真理的主要方法是以概率为基础的”。这位被誉为“法国的牛顿”的数学家曾经预言,“医疗是概率论应用的一个重要领域”“随着观察数的增多,有效的治疗方法会充分地显示出来”。

法国医生菲利普·皮内尔(Philippe Pinel, 1745—1826)说,可以通过清点产生良好反应的次数来确定一种治疗的效果,若成功率较高,便认为是有效的。他甚至宣称,应用概率计算,医学才能成为一门真正的科学。

此后,另一位法国医生皮埃尔·查尔斯·亚历山大·路易斯(Pierre-Charles-Alexandre Louis, 1787—1872)审慎地评价了当时流行的放血疗法。他观察了52例重伤寒患者,39例放了血,平均生存时间是25.5天,而没有放血者的平均生存时间却是28天;88例恢复期伤寒患者中,62例放了血,平均带病期是32天,而没放血者的平均带病期是31天。路易斯还研究了放血治疗肺炎和扁桃腺炎、咽峡炎的疗效。他明确宣布放血疗法无效。在19世纪30年代后期关于碎石术的争论中,路易斯统计了传统手术和碎石术的死亡率,前者是21.6%(1 237/5 715),后者是2.3%(6/257)。他雄辩地指出,由于人类记忆的谬误,外科医生倾向于较多地记住他们成功的案例而不是失败的案例。路易斯强调用数据准确表达疗效与用“或多或少”“罕见”“频繁”之类用词的区别是“真理与谬误的区别,一方是明明白白、真正科学的,另一方则是含含糊糊、缺乏价值的”。现在,路易斯被认为是循证医学(evidence-based medicine)思想的早期奠基人之一。

(二) 统计学与医学统计学的开端

1834年,伦敦统计学会的创办者们将统计学的目标定为数据收集。学术界普遍认为概率论和实际数据是两码事,不可相混。由于弗朗西斯·高尔顿爵士(Francis Galton, 1822—1911)和卡尔·皮尔逊(Karl Pearson, 1857—1936)等英国生物计量学派学者的工作,人们改变了看法,统计学从一项简单的社会统计工作转变为一门应用数学的科学。

高尔顿是查尔斯·达尔文(Charles Darwin, 1809—1882)的表弟。1869年,他对优秀人才及其家系进行定量研究,认为和体力很像,智力也会遗传。他首创问卷(questionnaire)调查,借助英国皇家学会的长寿会员们,深入研究遗传效应和环境效应间的区别。他通过一代又一代豌豆种子的大小,以他自己发明的回归直线(regression line)为工具,描述和解释了“向均数回归”(regression toward the mean)的普遍现象。

高尔顿的门生卡尔·皮尔逊是现代统计学之父,他开创了统计方法学,并将其推广应用,让大家接受。皮尔逊将这个学科从描述性统计学发展到推断性统计学,相关与回归、 χ^2 检验、 P -value等假设检验和统计决策理论都是由他奠定的基础。1894年,皮尔逊开设了他的第一门统计学理论的高级课程,使得伦敦大学成为1920年之前现代统计学教育的唯一场所。

皮尔逊积极支持统计方法的普遍应用,并确信数学可以应用于生物学问题,而统计分析可以为植物、动物和人类生命的许多问题提供答案。在一篇文章被英国皇家学会拒稿后,他和高尔顿、沃尔特·韦尔顿(Walter Weldon)一起于1901年创办了一份学术期刊《生物统计》(*Biometrika*),发表他和他的生物统计学派的研究成果。但是,1903年皮尔逊给高尔顿的信中提到《生物统计》在剑桥只有两个订户,一个是皮尔逊的朋友,一个是韦尔顿的朋友。尽管皮尔逊的主要贡献是相关分析和 χ^2 拟合优度检验,1906年英国皇家学会的杂志却拒绝登载其论文,因为主编不了解相关系数的生物学意义。在皮尔逊的年代,医学专业人员按是否认同统计方法有用分成两类:临床医生强调医学“艺术”,认为统计学没有用,只能依靠个人经验;生理学家或细菌学家确信统计学能使观察结果更加客观,但并不认为统计结果能成为“科学”证据。

除皮尔逊之外,另一位现代统计学的奠基人是罗纳德·A·费希尔爵士,他在剑桥主修数学,研究误差理论、统计力学和量子理论。他22岁发表第一篇统计学论文,介绍极大似然方法(maximum likelihood method),3年后另一篇论文推导了皮尔逊相关系数的精确分布。费希尔的主要贡献之一是利用随机化方法做实验,从而在统计分析中可以考虑数据的变异性。他发展了实验设计和分析的统计方法,汇集在他的两本专著《研究工作者的统计方法》(*Statistical Methods for Research Workers*)和《实验设计》(*The Design of Experiments*)中。他提出了减小误差的三项原则:重复(replication)、随机化(randomization)和适当地组织实验。费希尔认为,为了由局部推断整体,统计学家在实验的设计阶段就必须参与。在一次演讲时,他说:“做完实验后才找统计学家,就好像要他做尸体解剖。他只会说,这实验‘死’于什么原因。”由于贡献巨大,费希尔被后人公认为几乎独立地奠定了现代统计学基础的天才。

(三) 从第一个随机化有对照的临床试验到法制化

梅杰·格林伍德(Major Greenwood, 1880—1949)是第一个响应皮尔逊关于医学专业“迫切需要”新统计方法观点的人。他在皇家伦敦医院当学生时写信给皮尔逊,并将统计分析应用于他的研究数据。在1904—1905年,取得行医执照并在《生物统计》上发表一篇文章之后,他决定师从皮尔逊。在与细菌学家阿尔姆罗斯·赖特爵士(Almroth Wright, 1861—1947)辩论疫苗疗法的有效性时,格林伍德指出赖特的结果包含了抽样带来的误差,得到医学界的注意。1903年利斯特(Lister)

预防医学研究所创建了第一个统计系,由格林伍德来领导,主要处理流行病学和病理学问题,而他的导师皮尔逊在伦敦大学则重点处理遗传学和纯数理统计学等问题。通过培养格林伍德,皮尔逊创造了医学统计学家(medical statistician)的角色,即既懂医学又懂统计的一种研究者。

雷蒙德·珀尔(Raymond Pearl, 1879—1940)在美国的地位与格林伍德在英国的地位相当。他在密歇根大学(University of Michigan)获得生物学博士学位后即去伦敦师从皮尔逊。1918年,他担任约翰斯·霍普金斯大学公共卫生学院生物统计和生命统计学的教授以及医院的统计学家。1921年,珀尔在《约翰斯·霍普金斯医院通报》(*Johns Hopkins Hospital Bulletin*)的一篇文章中说,现代医院产生的数据必须有统计学专家参与分析,以确保医学研究科学化。

格林伍德的学生之一奥斯汀·布拉德福德·希尔爵士(Austin Bradford Hill, 1897—1991)是现代临床试验(clinical trial)的主要推动者。1937年,《柳叶刀》(*The Lancet*)的编辑认为有必要向医生们解释统计学技术,便邀请希尔撰写一系列关于在医学中正确使用统计学的文章。这些文章后来汇总成册出版,名为《医学统计学原理》(*Principles of Medical Statistics*)。

英国临床医学研究理事会于1946年开始进行第一项具有适当随机化对照组的关于链霉素治疗肺结核的临床试验。从多个中心搜集患者,随机地分到两个处理组:链霉素加卧床休息或单纯卧床休息。患者的X线片由两位放射科医师和一位临床医师独立评价。结果发现链霉素组患者在生存和“X线片”两方面都有更好的改善。

希尔的工作为后来的临床试验确立了一种传统,即医生的观察和专业统计学家的统计设计结合起来。这两门独立学科的交叉合作是将概率论应用于临床试验的必备条件。拉普拉斯的先见,在概率的基础上评价医学治疗效果的主张终于实现了。

希尔的倡导得到大西洋两岸治疗改革者们的支持。支持者们雄辩,随机化、有对照的临床试验使得医生能够选择真正较好的治疗方法,可以防止他们过分热衷于一些较新的治疗方法。在1959年的一次学术会议上,牛津大学的医学教授乔治·皮克林爵士(George Pickering)赞扬随机化、有对照的临床试验,并宣称,相对而言,医生个人的临床经验是无计划、杂乱的,而且是非常不可靠的。

英国人率先将统计学应用于有对照的临床试验,美国人也进行了类似的实践。美国人1954年实施了人类历史上最大规模、花费最多的一项临床试验,旨在评价索尔克(Salk)疫苗预防脊髓灰质炎的效果。当时,脊髓灰质炎的年发病率约为1/2 000。该试验有180万名儿童参与,直接花费超过500万美元。开始时随机化分组遇到一些阻力,但最后约有1/4参与者得到了随机化。这项试验最终肯定了索尔克疫苗的效果。

20世纪60年代初,药物沙利度胺(thalidomide,又称反应停)曾引起胎儿畸形的暴发。美国食品药品监督管理局(Food and Drug Administration, FDA)发现有1 267名医生给19 822名患者(含3 760名育龄妇女)开处方,涉及250余万片药。这一事实提出一个问题:医学界的“专业判断”是否仍然可信?公众的呼声使得美国国会通过、总统批准了1962年药物修正案,即《基福弗-哈里斯修正案》。该修正案要求新药必须经过动物实验、加强监管,并规定使用未获批药物需患者知情同意,从简单的上市前通告制度变为复杂的上市前审批制度,大大加强了FDA的管理权限。这项法律把美国FDA转变为医药领域里药物有效性的最终仲裁者。美国FDA为临床试验规定了一整套制度,使之成为确定药物疗效的标准方法。

现在,统计学在医学中的应用被视为具有科学权威的“证据”,比个别专家意见更高级,更具“客观性”和“真实性”。这也和统计学这门学科定义的演变是一致的。1959年斯坦福大学教授赫尔曼·切尔诺夫(Herman Chernoff)和林肯·E·摩西(Lincoln E. Moses)合著的教材《基本决策理

论》(Elementary Decision Theory)中说,“多年以前统计学家也许会宣称,统计学是做数据列表的。如今的统计学家很可能说,统计学考虑的是面对不确定性如何作决策”。

二、统计学与公共卫生的相互促进

(一) 统计学是公共卫生专业人员的得力工具

与临床医师治疗个体患者不同,公共卫生考虑的是群体的健康。这就决定了公共卫生对统计学天然的亲和性,也决定了公共卫生专业人员视统计学为必备素养。人们采用各种抽样技术精心设计调查研究,掌握人群的健康状况和卫生服务需求;采用统计描述的手段反映疾病和卫生资源的分布特征;采用统计推断的技术在充满变异和偶然性的数据中识别危险因素、评价卫生措施、进行科学决策。

约翰·格朗特(John Graunt, 1620—1674)搜集死亡方面的数据,编制了寿命表,从而创造了人口统计这一学科。威廉·法尔(William Farr, 1807—1883)进一步改进寿命表方法,在英国建立了官方生命统计系统。美籍华裔生物统计学家蒋庆琅(Chin-Long Chiang, 1915—2014)在美国加州大学伯克利分校以现代统计学和随机过程的观点系统研究了寿命表的原理,使之成为生物统计学的重要组成部分,并出任世界卫生组织顾问,促进了寿命表在国际范围的应用,直至如今。

1848年,约翰·斯诺(John Snow, 1813—1858)首次详细研究了伦敦霍乱流行情况,发现传播途径,及时控制了疫情的蔓延。传染病的传播动力学模型应用了数学和统计学的方法。

第二次世界大战后不久,理查德·多尔(Richard Doll, 1912—2005)和奥斯汀·布拉德福德·希尔(Austin Bradford Hill, 1897—1991)两位爵士领导了吸烟与肺癌关系的创新性研究。1948年4月与1949年10月期间,他们向伦敦20所医院的709位肺癌住院患者询问有关他们吸烟的问题;与此同时,针对每一位肺癌患者,他们在同一所医院寻找一位同性别、年龄差别不超过5岁的非癌症患者作为对照,询问同样的问题。通过病例组与对照组吸烟量的比较,发现病例组吸烟量较大。1951年11月,多尔和希尔又开始了一项更大规模的纵向研究。他们向所有注册的英国医师约60 000人发出关于吸烟的问卷,收回约40 000份。然后,通过死亡登记获取他们的死亡和死亡原因等信息。1964年,他们报告了10年追踪的结果:对应于从未吸烟、每天1~4支、5~24支和25支或更多,英国男性医师的肺癌死亡率分别为0.07‰、0.57‰、1.39‰和2.27‰。后来,1974年和1994年又分别报告了20年和40年追踪的结果,除数字有所不同外,均存在类似的剂量-反应关系。

多尔堪称20世纪最具前瞻性的流行病学家。早在20世纪80年代,鉴于中国年轻人中吸烟者日益增多和社会控烟不力的情况,多尔警示肺癌等肿瘤和心血管疾病将成为中国社会的沉重疾病负担。

(二) 现代公共卫生领域对统计学的挑战

公共卫生领域不但充分应用现成的统计学理论与方法,还不断向统计学提出新要求和新问题。

20世纪后半叶,为了研究慢性非传染性疾病的流行,人们发展了一系列现代统计方法,用于危险因素的分析、生存时间的分析、疾病自然史等领域。这些方法不但大大提高了分析流行病学的能力和水平,而且也扩充了现代统计学和生物统计学的研究范围,出现了许多新的理论、方法和技术。

在评价化学毒物或药物时,动物毒理学和人群流行病学是两个主要的科学证据来源。长期致癌实验和生殖系统实验等动物实验数据的统计分析有定性检验和危险度定量评价两类。定性检验通过处理组和对照组之间的比较来确定化学物是否存在有害健康的效应;危险度定量评价更需借

助复杂的统计模型,例如广义线性混合效应模型、非线性混合效应模型、经验贝叶斯方法等。

横断面研究、病例对照研究和队列研究等传统的设计模式难以显示疾病的遗传规律。遗传流行病学注重家系资料的搜集,进而就父母向子女的传递、同胞对内部的异质性等进行统计分析。然而,同一家系内的成员并非相互独立的个体,且往往缺失甚多,这就大大推动了多水平模型和缺失数据填补方法等的研究。

艾滋病研究中,由于总体成员及其规模往往未知,个体感染和发病的起点也未知,原有的随机抽样调查技术难以施展,发病率、患病率以及潜伏期的常规估计方法不再适用。这一切促使人们不断发展逆向估计和捕获-再捕获等新的抽样与估计方法。

随着科学技术的快速发展,基因测序技术日新月异,通过基因测序技术锁定个人的疾病易感基因,进行提前预防和治疗成为公共卫生研究的热点。基因测序数据的一个特点是样本量小而变量多,一个测序数据库样本量通常为数百上千,最多上万,而变量数多达数百万个,表现为典型的超高维稀疏数据。因此,如何处理测序数据给统计学提出了新问题和新的挑战。在医学理论的指导下,寻找快速、高效的降维或投影策略和方法成为研究的重点领域之一。

传染病监测预警是预防和控制新发突发传染病的主要手段。运用动力学模型进行疾病的预警预测便成为此领域重要的研究分支之一。动力学模型主要包括微观动力学模型(例如病毒在体内的动力学)和宏观动力学模型(例如传染病在人群间的传播动力学),主要分析工具是微分方程。而疾病的发生发展以及传播与很多未知因素有关,充满了随机性,因此,从纯粹的动力学模型入手,难以描述疾病在发生发展和传播过程中的变异性。因此,随机微分方程是研究疾病动力学的一个新的发展方向。

总之,公共卫生不仅是统计学应用的一个重要领域,而且始终是现代统计学研究和发展的巨大动力。

三、卫生统计学的未来

21 世纪是信息爆炸的时代,信息的增长速度远远超过了人们处理信息和分辨信息的速度,被人们称为大数据时代。统计学是 20 世纪最伟大的学科之一,其自身的理论不断完善,方法层出不穷,应用领域越来越广,为各应用领域所作出的贡献不可估量。统计学丰富的历史积淀和当前的社会实践都促进和激励人们必须对统计学未来的发展进行思考。随着计算机技术的飞速发展,统计学的发展如虎添翼。21 世纪的统计学必将取得突破性发展,进入一个崭新的历史发展阶段。

卫生统计学未来的发展方向一定是整合多个学科的理论、方法和技术,团结不同专业背景的专家学者,利用大数据资源,为健康促进、疾病预防、临床治疗、康复等领域提供科学依据与决策支撑。目前的实践已经初现端倪,例如统计学的判别分析方法与现代计算技术结合,产生了有监督的机器学习方法;聚类分析与计算技术结合产生了无监督的机器学习方法;基于大数据,依靠强大的计算机算力,统计学的思想和方法结合现代计算技术,诞生了现在的人工智能。

卫生统计学与流行病学、数学、计算机科学、信息学、系统科学等相结合,研究处理健康医疗大数据的理论和方法,用于提升健康促进、医疗决策和生命科学研究的能力和水平,将形成一个新兴的学科和领域,目前人们称之为健康数据科学(health data science)。它的未来发展方向主要包括:①多模态数据整合。整合临床诊疗、基因组学、蛋白质组学、暴露组学、行为与环境等多源异构数据,实现多维度的健康数据分析,开展系统流行病学的研究。②人工智能和机器学习的应用。深度学习、强化学习等人工智能技术在医疗图像分析、疾病监测预测等方面的应用将会持续扩展,提高