

王 宇 杨功焕 主编

乌正赉 叶冬青 方法卷主编

Public  
Health  
in China

PUBLIC HEALTH IN CHINA

中国  
公共卫生

方 法 卷



中国协和医科大学出版社



Public Health in China  
中国公共卫生  
方法卷

---

主 编 王 宇 杨功焕  
方法卷主编 乌正赉 叶冬青

中国协和医科大学出版社

## 图书在版编目 (CIP) 数据

中国公共卫生·方法卷 / 王宇, 杨功焕主编. 乌正赓, 叶冬青方法卷主编. —北京: 中国协和医科大学出版社, 2010. 2

ISBN 978-7-81136-331-9

I. 中… II. ①王…②杨… III. 公共卫生-研究-中国 IV. R1

中国版本图书馆 CIP 数据核字 (2010) 第 003615 号

## 中国公共卫生·方法卷

---

主 编: 王 宇 杨功焕  
方法卷主编: 乌正赓 叶冬青  
责任编辑: 陈永生 左 谦 段江娟

---

出版发行: 中国协和医科大学出版社  
(北京东单三条九号 邮编 100730 电话 65260378)

网 址: [www.pumcp.com](http://www.pumcp.com)  
经 销: 新华书店总店北京发行所  
印 刷: 北京佳艺恒彩印刷有限公司

---

开 本: 889×1194 1/16 开  
印 张: 29.5  
字 数: 670 千字  
版 次: 2013 年 5 月第 1 版 2013 年 5 月第 1 次印刷  
印 数: 1—3000  
定 价: 95.00 元

---

ISBN 978-7-81136-331-9/R·331

---

(凡购本书, 如有缺页、倒页、脱页及其他质量问题, 由本社发行部调换)

## 《中国公共卫生》编撰委员会

顾问：陈春明 陈君石  
主编：王 宇 杨功焕

理论卷主编：曾 光 黄建始 张胜年

方法卷主编：乌正赓 叶冬青

实践卷主编：杨功焕 王若涛 李 辉 郑玉新

策划：袁 钟 胡永洁

学术秘书：段江娟 孟 媛

## 《中国公共卫生·方法卷》编委会

主编 编：乌正赓 叶冬青

编委：（按姓氏笔画为序）：

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| 丁国伟 | 于保荣 | 万 霞 | 王 健 | 王 璐 |
| 王亚东 | 王若涛 | 乌正赓 | 宁 艳 | 叶冬青 |
| 田向阳 | 刘 民 | 孙 强 | 许 群 | 吕书红 |
| 刘兴柱 | 严 非 | 杨 森 | 肖 璆 | 邹小农 |
| 张风雨 | 吴永宁 | 杨功焕 | 张学军 | 陈君石 |
| 陈园生 | 孟庆跃 | 高 敏 | 柴 燕 | 梁万年 |
| 龚幼龙 |     |     |     |     |

# 序

20 世纪,世界许多国家公共卫生状况空前改善。最近 30 多年来,全世界卫生事业取得显著成就,婴儿死亡率大幅度下降,营养不良获得普遍改善,医药卫生技术创新也势头迅猛。但是,经济发达国家和发展中国家在卫生投入和健康状况方面的差距依然如故。发展中国家每年有近 1100 万儿童死于可预防的传染性疾病。以疟疾这种可预防的疾病为例,每 30 秒就会夺去一名世界贫困地区儿童的生命。每年有超过 50 万妇女死于妊娠和分娩。结核病是可以治愈的疾病,但每年依然有 170 万人死于结核病。大多数低收入国家的艾滋病毒/艾滋(HIV/AIDS)疫情依然没有得到控制,全球大约已有 6000 万人感染艾滋病毒,2500 万人死于艾滋病相关疾病,而在中国估计目前存活艾滋病毒感染者和病人约有 78 万。

与此同时,慢性非传染性疾病,无论是发病数和死亡数都占总发病和死亡数的绝对多数,其中有 6 种重要慢性病(脑卒中、冠心病、糖尿病、肺癌、肝癌和乳腺癌)占总死亡的 35%,其标化死亡率呈上升趋势,这意味着危险因素在慢性病上升中起了关键作用。目前与慢性病相关的危险因素——烟草使用、酗酒、高盐高脂饮食以及静坐生活方式,要么处于高流行水平,要么呈进行性上升趋势。这些危险因素的流行趋势表明,在未来 20~30 年慢性病的发病和死亡率还会持续上升,其带来的疾病负担、劳动力的损失以及巨大的医疗费用,都将给社会、家庭和个人造成严重的影响。

儿童、青少年和劳动力人口中,伤害是第一位死因。大气和室内空气污染、不安全的饮用水和食品、工作环境,以及电离辐射等有害因素的流行水平增加,缺乏监管和控制,给健康带来了严重的危害。总之,在社会经济发展的进程中,新的健康问题不断增加。

过去 50 年,中国人群的健康状况得到了很大改善,在短短的几十年,人群期望寿命上升,婴儿死亡率、5 岁以下儿童死亡率和孕产妇死亡率呈明显下降趋势,营养不良疾病、主要的传染病,特别是疫苗可预防的传染病、肠道传染病,以及地方病,均呈明显下降趋势。但是各地发展不平衡,在贫困、偏远地区,这些应该得到良好控制的疾病和健康问题依然还很严重。传染病中,经性传播的传染病, HIV 感染仍呈上升趋势。

中国地域广阔,发展不平衡,许多应该得到有效控制的疾病和健康问题在偏远地区还未能有效控制,而偏远地区这些没有控制的传染病随着人口流动进入城市,使得这些问题更加严重。

中国人群中疾病模式发生了非常显著的变化,新出现的健康问题等,对公共卫生提出了新的要求需要利用新的理论、技术和方法应对这些新的挑战。然而我国目前尚没有较全面介绍公共卫生理论、方法,特别是总结中国公共卫生实践的书籍。为填补空白,2005 年中国疾病预防控制中心和中国协和医科大学出版社共同策划编写《中国公共卫生》一书,2006 年编写工作正式启动。

该书作为中国公共卫生领域的一部专著,由中国疾病预防控制中心牵头,联合全国公共卫生院校、临床、科研单位及社会学界专家共同编写。全书共分三卷,分别为公共卫生的理论卷、方法卷和实践卷。理论卷,阐述伴随公共卫生中新问题出现而产生的理论进展;方法卷,重点介绍公共卫生中常用的方法及技术;实践卷,主要反映过去半个多世纪中国公共卫生的实践,总结中国公共卫生实践的成功经验,同时反映随着快速的城市化、工业化,疾病模式快速转变,新的健康问题,中国公共卫生所面临的挑战和应对。实践卷是本书的特色,也是定名为《中国公共卫生》的依据。

本书作为学术专著,鼓励作者有独到、创新的观点。本书的读者对象定位于希望了解中国公共卫生实践,或致力于公共卫生事业的专家、学者,以及在公共卫生领域的工作人员和决策者;可作为疾病预防控制系统及公共卫生领域专业人员、决策者的工具书、参考书,也可作为其他行业了解公共卫生现状及相关知识的指导书。

参编专家们认同编著这样一部书的必要性和艰巨性,需要团结各领域的专家学者,共同努力、集中时间、全身心投入,完成这部高水平学术著作的编写;希望本书成为一部具有开拓性、科学性、客观性、权威性和全面性的高水平著作。但是在过去3年多的时间里,编写团队中的公共卫生专家们又经历了我国公共卫生中的诸多大事件,在投入公共卫生活动的同时仍在努力完成本书的撰写和编辑。当本书编撰成稿即将出版时,我们仍感到距理想状态相差甚大。我们怀着忐忑不安的心情将本书呈现给读者,希望能对广大读者有所裨益。对于书中存在的不足,望读者不吝指正,以便再版时更正。



## 前 言

在本书总主编的关怀和各位编者的努力下,《中国公共卫生》第二卷“方法卷”终于与读者见面了。方法学之于公共卫生,如同钥匙之于大门、阶梯之于高楼,它既是公共卫生的工具,又与公共卫生相生、相济。公共卫生工作者借用其他学科(如临床医学、流行病学、生物统计学、信息学、社会学、基础医学等)的方法,来观察、分析疾病和健康的人群现象,提出假设、验证假设、探索病因,提出干预措施并评价其效果,达到预防疾病、促进健康的目的。随着社会经济发展及医学模式的转变,公共卫生的格局发生了明显的变化,社会对公共卫生需求也相应改变,原有的公共卫生方法学已不能完全适应社会发展的新要求。

公共卫生的需求包括两方面,一是处理突发事件的应急需求,二是在较长一段时期内改善社会卫生状况的需求。突发事件应急要求公共卫生工作者能够迅速有效地控制传染病暴发及紧急环境暴露,最大限度地保障人群健康。公共卫生工作者应能运用流行病学、生物统计学、微生物学、毒理学等专业知识,判断传染病的暴发,紧急环境暴露的原因、波及范围、影响程度等;还应能运用医学、流行病学、社会学知识及行为干预,以在短时间内迅速做出决策、采取控制策略和措施。除应急外,从长远讲,公共卫生工作者还应通过加强疾病监测、预警、实验室检测、医院的传染病控制能力,控制传染病的流行,降低婴儿死亡率、孕产妇死亡率,减少环境污染、职业危害及慢性病,提高全人群健康水平。公共卫生工作者必须会针对不同的问题,根据公共卫生原则,采用不同方法,解决实际问题。

公共卫生涉及面广,包括自然科学、社会科学中的诸多学科。公共卫生方法学内容丰富,但迄今国内仍无一部能较全面、系统地介绍公共卫生方法的专著。在日常工作中,我们真切地感受到公共卫生工作者在方法问题上的迫切需求。考虑到这些实际情况,且根据我国公共卫生事业发展已有60年历史的事实,有必要利用多年的积累,对公共卫生方法学方面的一些理论和实际经验进行系统整理,以飨读者。这就是编纂本卷的初衷。

本卷涵盖流行病学与公共卫生、流行病学常用的测量指标,社会学研究方法在公共卫生中的应用,公共卫生监测和国家卫生信息系统,死因推断的方法学研究,卫生经济学在公共卫生中的应用,公共卫生项目管理,基因组学与公共卫生,共十章。书中较详细地介绍了不同领域方法学在公共卫生中的应用,其中流行病学与生物统计学是公共卫生之基础,其重要性毋庸置疑。疾病监测和信息系统是公共卫生的主链及载体,在现代疾病预防控制中,无论如何强调监测都不会过分,而疾病和健康信息是作出合理公共卫生决策的基础。卫生经济学、风险评估、公共卫生项目管理都是公共卫生方法学宏观整体的重要组成部分,公共卫生

的基因组学则是诠释如何从微观角度解决公共卫生的实际问题。公共卫生的发展很大程度需借鉴其他学科的知识及其先进的方法和技术,同时,我国公共卫生的发展还需学习国外的先进经验,提高我国公共卫生服务的能力。

本卷编写的特色,一是突出实用性,除强化了流行病学与生物统计学知识,还着重信息技术的利用,这些都是公共卫生工作者必备的基本素质和技能。二是注重可操作性,书中介绍的都是操作性强、切合实际的方法,适合现阶段我国社会经济发展水平及公共卫生实际情况。三是吸收了国内外相关研究的新知识和技术,特别是国外的先进经验,弥补我国在相应领域的不足。四是结合真实案例,介绍各种方法学的灵活、具体应用,以解决实际问题,便于读者理解。

在编写期间,编者不辞辛苦,从各地聚集一起确定编写大纲,召开编写会、定稿会达十多次。在沟通交流中,各位专家畅所欲言、各抒己见,达成很多共识,也不乏异议,体现了学者们的执著,做到了“君子和而不同”。作为本卷的主编,我们衷心感谢各位编者为本卷编写所作出的贡献。

在本书的编写中缺点和错误在所难免,祈望所有阅读本卷的读者不吝赐教,以待再版时订正。

方法卷主编 谨识

2012年7月



# 目 录

|                            |         |
|----------------------------|---------|
| 第一章 流行病学与公共卫生 .....        | ( 1 )   |
| 第一节 流行病学基本原理 .....         | ( 1 )   |
| 第二节 流行病学方法进展 .....         | ( 16 )  |
| 第三节 流行病学方法应用实例 .....       | ( 47 )  |
| 第二章 流行病学常用的测量指标 .....      | ( 85 )  |
| 第一节 测量指标的分类 .....          | ( 85 )  |
| 第二节 测量疾病频率的指标 .....        | ( 86 )  |
| 第三节 测量暴露与疾病之间联系的指标 .....   | ( 97 )  |
| 第四节 效用指标 .....             | ( 112 ) |
| 第三章 社会学研究方法在公共卫生中的应用 ..... | ( 132 ) |
| 第一节 社会学研究方法概述 .....        | ( 132 ) |
| 第二节 社会调查研究的基本步骤 .....      | ( 137 ) |
| 第三节 定量研究 .....             | ( 142 ) |
| 第四节 定性研究 .....             | ( 155 ) |
| 第五节 实验研究 .....             | ( 161 ) |
| 第六节 评价 .....               | ( 164 ) |
| 第七节 社会学研究方法比较 .....        | ( 167 ) |
| 第四章 公共卫生监测和国家卫生信息系统 .....  | ( 172 ) |
| 第一节 公共卫生监测的起源和发展 .....     | ( 172 ) |
| 第二节 公共卫生监测的基本环节 .....      | ( 176 ) |
| 第三节 公共卫生监测的新进展 .....       | ( 183 ) |
| 第四节 几种重要的监测系统 .....        | ( 187 ) |
| 第五节 建立、运行和评价监测系统 .....     | ( 199 ) |
| 第六节 强化国家卫生信息系统的建设 .....    | ( 212 ) |
| 第五章 死因推断 .....             | ( 232 ) |
| 第一节 背景 .....               | ( 232 ) |
| 第二节 死因推断量表 .....           | ( 237 ) |
| 第三节 根据死因推断量表推断死因 .....     | ( 240 ) |
| 第四节 死因推断量表的效度研究 .....      | ( 241 ) |

|                                     |              |
|-------------------------------------|--------------|
| 第五节 死因推断量表效度与死因构成 .....             | (247)        |
| 第六节 死因推断量表的运用 .....                 | (249)        |
| <b>第六章 风险评估 .....</b>               | <b>(255)</b> |
| 第一节 风险评估在风险分析框架中的地位 .....           | (256)        |
| 第二节 风险评估的基本内容 .....                 | (258)        |
| 第三节 风险评估的应用 .....                   | (274)        |
| <b>第七章 健康教育与健康促进方法与应用 .....</b>     | <b>(283)</b> |
| 第一节 健康传播策略与方法 .....                 | (283)        |
| 第二节 健康心理学方法 .....                   | (292)        |
| 第三节 健康教育与健康促进的计划、实施与评价 .....        | (296)        |
| 第四节 社会动员 .....                      | (300)        |
| 第五节 健康教育与健康促进方法在公共卫生中的应用 .....      | (303)        |
| 第六节 健康教育和健康促进方法在不同人群与场所中的应用 .....   | (309)        |
| 第七节 健康教育与健康促进方法在行为与生活方式干预中的应用 ..... | (317)        |
| 第八节 国外健康教育与健康促进案例 .....             | (325)        |
| 第九节 中国健康教育与健康促进的发展 .....            | (336)        |
| <b>第八章 卫生经济学在公共卫生中的应用 .....</b>     | <b>(344)</b> |
| 第一节 卫生经济学发展简史 .....                 | (344)        |
| 第二节 卫生经济学基本理论和方法 .....              | (348)        |
| 第三节 卫生经济研究分析框架 .....                | (360)        |
| 第四节 卫生经济学在公共卫生中的应用 .....            | (369)        |
| <b>第九章 公共卫生项目管理 .....</b>           | <b>(379)</b> |
| 第一节 项目管理概述 .....                    | (379)        |
| 第二节 项目需求论证 .....                    | (383)        |
| 第三节 卫生项目的准备与设计 .....                | (386)        |
| 第四节 卫生项目的实施与监督 .....                | (388)        |
| 第五节 公共卫生项目的评估 .....                 | (392)        |
| <b>第十章 基因组学与公共卫生 .....</b>          | <b>(403)</b> |
| 第一节 基因组学概述 .....                    | (403)        |
| 第二节 人类疾病的基因组学研究 .....               | (406)        |
| 第三节 后基因组时代的系统生物学 .....              | (416)        |
| 第四节 基因组学的关联分析与人类复杂疾病 .....          | (421)        |
| 第五节 表观遗传学与肿瘤 .....                  | (438)        |
| 第六节 基因组在临床实践及疾病预防中的应用评估 .....       | (442)        |
| 第七节 公共卫生基因检测中的问题 .....              | (446)        |
| <b>索引 .....</b>                     | <b>(456)</b> |

发达国家,人群慢性非传染性疾病发病率超过传染病。随之西方国家进行了一系列大规模的流行病学研究,包括20世纪50年代中期英国医师Doll和Hill对吸烟与肺癌关系的研究,1948年迄今长达60多年的美国弗雷明汉心脏研究(Framingham heart study),以及1954年由Jonas Edward Salk组织实施的涉及美国、加拿大和芬兰3国150余万人参加的脊髓灰质炎疫苗现场试验等在流行病学发展中具有里程碑性质的事件,确定了吸烟与肺癌,高血压、高胆固醇、肥胖与冠心病的关系,肯定了脊髓灰质炎疫苗预防脊髓灰质炎的效果,并促使流行病学多病因理论的逐步形成和发展,流行病学研究设计(病例对照研究、队列研究、巢式病例对照研究等)不断完善。统计学分析技术(效应测量指标,分层分析,偏倚的识别、控制、校正,多变量分析方法等)的进步,又推进流行病学研究的发展。

自20世纪90年代开始,伴随着生物学特别是分子生物学、基因组学等的发展,流行病学与其交叉融合并形成了分子流行病学、分子遗传流行病学等诸多流行病学分支学科。流行病学似乎摆脱了长期以来囿于危险因素研究的尴尬局面,并有望揭开危险因素与疾病之间联系“黑箱”的内在机制,为提出更有效的公共卫生干预措施提供了可能。然而,过去20年间被人们寄予厚望的流行病学却遭遇了前所未有的困惑与批评。主要表现为:首先,流行病学关于疾病与危险因素之间弱联系的研究结果常不可重复;其次,观察性研究结论与干预性研究结论常有不一致;第三,对流行病学观察性研究内在效度(由于混杂不可避免及难以控制而导致)的怀疑;第四,对流行病学研究的反思,例如,现代流行病学对公共卫生影响研究多限于个体水平,忽略了对群体水平的影响,尤其是群体、个体影响因素之间的相互作用等等。

当然,怀疑与反思并非意味着流行病学学科发展的终结,相反,是流行病学学科发展的重要动力,促使流行学家进一步研究观察性研究的理论、设计及分析方法,扩展流行病学研究的内容。近年来,在传统流行病学研究设计基础上衍生出许多新的研究设计类型,群体水平的公共卫生影响因素也开始受到重视,特别是多水平模型、结构方程模型等的提出,为同时研究公共卫生群体、个体以及两者套叠的相互作用,并从历史、现代与未来多个维度研究疾病与健康问题提供了方法基础。随着信息化时代的到来,循证医学的提出、发展和壮大,为系统总结单项流行病学研究结果、循证公共卫生决策构建了框架,同时也为流行病学研究的公共卫生导向提供方法和思维基础。

流行病学的根本性特点是,其研究对象是以多个个体(每一个个体为基本单位)组成的群体,此群体生存的宏观环境、研究范围涉及各种健康(卫生)问题及现象,以及其与环境的相互联系,这有别于研究每一个个体内的系统、器官、组织、细胞、分子的基础医学诸学科(如组织胚胎学、生理学、病理学、药理学等),也有别于研究微观环境科学中的诸学科(如微生物学、环境毒理学等)。流行病学的另一个根本性特点是其实用性,它是为公共卫生决策需要服务的。流行病学工作者不可能像实验室工作者那样自己去“安排”卫生事件或环境,而只能采用适合社会实际情况、切实可行的方法,发现并解决当地的公共卫生问题。所以,流行病学又被称为公共卫生中“运用可行方法的艺术”。流行病学工作者的经常性工作是从人群疾病及健康问题,当地各种公共卫生问题在不同时间、不同空间及不同特征人群间的分布(常称为“三间”分布)着手,发现人群疾病及健康问题的原因、人群中主要的公共卫生问题,并找到可以干预、影响疾病及健康问题及其他各种公共卫生问题分布的相关因素,提出并实施相应的可行决策,达到预防、控制疾病及其他公共卫生问题的目的。

此外,流行病学的发展始终与生物(卫生)统计学的进展紧密相伴。流行病学发展中的

需求,不断地给统计学提出挑战,而统计学的进展又为流行病学提供了定量描绘、分析群体疾病、公共卫生事件及其影响因素的支持性工具技术。例如,相对危险度和比值比的概念及其计算方法。虽然近年来流行病学越来越多地涉及人类行为的研究,加之全球定位系统(GPS)的进展,社会学和人类学现场的“定性研究”与传统的统计学定量方法结合在一起应用于流行病学及公共卫生中,但是,与生物统计学方法的紧密结合,运用数学定量方法来处理疾病、健康及其他公共卫生问题仍然是流行病学一个很突出的特征,合起来称为公共卫生中的“流行病与统计学方法”。

回顾公共卫生实践中流行病学的发展,可以将其基本原理概括为“以群体生态思维的方式,运用统计学定量工具,从比较疾病、健康事件在不同人群中的分布入手,探索疾病、健康问题及其他公共卫生问题可干预性因素,组织实施干预,达到为公共卫生决策服务,维护和促进人群健康目的”的方法学。它是集“思维”、“工具”和“服务”三位一体的学科,既是“科学”又是“艺术”。公共卫生工作者应深刻理解流行病学思维,掌握其各种方法、技术,因地制宜地为公共卫生需要服务。

## 二、流行病学指标的有效测量

流行病学方法的第一步,就是对疾病、死亡、健康问题或其他卫生事件,在明确定义的基础上,相对于其发生的时间、空间和人群,通过观察测量而赋予数字或数值。可以赋值的范围称为变量的定义域。对变量所赋予的数字,只不过是一个数学符号,同一个数学符号在不同变量中代表的意义可能完全不一样。按变量可赋予数字值的性质,变量可分为属性(二分、多分)变量及连续变量,或计数、等级、计量变量,或列名、等级、间隔及比变量(按有无绝对零点划分),其中二分变量只可赋值为0或1。通过改变赋值方法,计量数据可转换成等级数据,等级变量也可转换成计数变量,但是反方向的转换却是不可能的。

对疾病、死亡、健康问题或其他卫生事件的赋值,就可得到有关变量的数值。公共卫生中最常用的绝对数是疾病(死亡)发生的例数(case number)和发病(死亡)的时间(incidence time)。两个绝对数通过一次除法运算得到的数,称为相对数。人群中的疾病、死亡、健康事件或其他卫生事件,除少数罕见疾病可用绝对数表述之外,多用相对数指标来测量。

常用的相对数指标有4类:①构成比(例),例如,患病比例,无量纲,取值在0到1闭区间内,反映构成概率。患病比例习惯上又称患病率(prevalence)。②发生(频)率,例如,累积发病频率,无量纲,取值在0到1闭区间内,反映发生概率。累积发病频率习惯上又称发病率(incidence)。③速率指标,例如,发病密度(incidence density),量纲为时间的倒数,取值在0到无穷大,反映疾病发生的时点密度,不是概率。④比(ratio),例如,比数(odds),量纲可有/可无,取值由具体的情况而异,反映分子分母对比的指数关系。

在此4类基本相对数指标基础上可以派生出多种指标。例如,发病率差( $R_1 - R_0$ )、相对危险度(relative risk,  $RR$ )和比数比(odds ratio,  $OR$ ),反映了暴露与疾病之间相关的强度,等等。判断某个指标的性质一定要依据其如何测量赋值,而不能由名称猜测。

现以某病为例说明流行病学中几类最基本的相对数指标的相互关系。第一步给予明确的定义,这是测量病例(相对数的分子部分)的基础,然后进一步明确(存在或发生)这些病例的相关人群(相对数的分母部分),从而得到两个基础性相对数指标,患病率和发病率。



在此基础上,把发病率作为一个连续性变量而求对其时间的导数就可得到时点发病密度。对某个稳定人群而言,当观察期短并且发病人数远远小于全体观察人数时(大多数疾病满足此条件),在此观察期发病密度恒定,则上述四类基本相对数指标有以下关系:

(平均)时点患病构成比数=发病密度×平均患病时间

累积发病率=发病密度在观察期的积分= $1-e^{-(\text{发病密度} \times \text{观察时间})}$

由此还可以近似地估计:时点患病率=发病密度×患病时间

累积发病率=发病密度×观察时间

收集数据的方法可以间接引用常规报告或现存资料,也可通过直接观察、检测、调查、公共卫生监测或筛查。原始数据的真实、可靠及精确,决定了结果的真实、可靠及精确。因此,整个资料的收集过程包括记录、计算机录入、校对、整理、归纳,都必须实行全面质量保障(total quality assurance)。最后,还应对测量的结果(赋值)开展有效性(validity)和可靠性(reliability)评价。

测量的有效性,有时也称为真实性或效度,指的是测量得到的数值反映真值的程度。测量的可靠性,有时也称为可重复性或信度,指的是反复测量同一真值得到的数值之间的一致性程度。真值常用所谓的“金标准”来表示。可靠性不是有效性的充分条件,但却是有有效性的必要(前提)条件。在理化测量中,常称测量的有效性为准确度(accuracy),称可靠性为精密度或精度(precision)。有些流行病学工作者将“准确度”定义为可靠性与有效性的合称。

如真值是确定的常数时,反复多次测量同一真值得到的许多测量值的总方差等于系统误差方差与随机误差方差两者之和。随机误差,是指测量无法避免的、均数为0、呈正态分布的偶然性误差。测量的总方差是测量可靠性的指标,总方差越大可靠性越差。测量值的均数与真值的差称为绝对测量误差,绝对测量误差与真值的比为相对测量误差,这都是反映有效性的指标,测量误差越小其测量的有效性越好。如绝对测量误差为0时,有效性最好,反映样本测量为无系统误差,称为无偏测量。反之,如绝对误差不为0,称此测量为有偏测量,绝对误差为正值时称为高估偏性测量,绝对误差为负值时称为低估偏性测量。分析有效性时,要判断有无系统误差、系统误差的方向及大小(即偏倚是否存在,若存在,其方向及大小如何)。

如被测量的变量是二分变量,判断其可靠性的指标常用kappa值(统计量),判断其有效性常用敏感度(sensitivity)和特异度(specificity)来评价。如测量的变量是计量变量,可靠性常用两次独立测量值的相关系数来评价,有效性常用测量值的回归系数来评价。如真值是随机变量,测量值的总方差等于真值方差加系统误差方差加随机误差方差三者之和,其可靠性的指标可定义为真值方差与系统误差方差之和占测量总方差的比例,该比例越大越可靠;其有效性定义为真值方差占测量总方差的比例,该比例越大越有效。所以,有效性高的测量其可靠性一定高,而可靠性高的测量其有效性不一定高。如没有“金标准”,就无法得到真值,可靠性只能用测量本身的可重复性来评价,有效性也常用表面效度(face validity)、内容效度(content validity)和结构效度(construct validity)来估计。常用的区分效度(discrimination validity)本质上是一种结构效度。

总之,流行病学方法源于卫生事件的观察与测量。因此,原始数据的质量控制是公共卫生决策的基础。

### 三、选择适宜统计学技术估计随机变异及相关性

医学中公共卫生的许多事件可以有多种结局，个体会呈现什么结局并无规律，但随着观察由个体组成的群体数量增加，各种事件结局发生可能性却呈现出确定性分布规律。概率是可能性的数学表达，它是一个可以取值（定义域）为0到1的闭区间无量纲的数，如某种结局发生概率等于1时，该结局必然发生，如概率为0时，该结局绝不发生，其他情况下结局发生的可能性介于0到1之间。发生某结局概率为0或1的事件，称为确定性事件，而在个体中呈现无规律、在群体中呈现确定性概率分布的事件，称为随机或偶然事件。若不但在个体中呈现无规律、在群体中也不呈现确定性概率分布的事件，称为混沌（chaos）事件。群体中多种结局呈现的概率分布，就是对个体发生该事件的各种可能性大小的估计，或者是对随机事件偶然性的估计。例如，观察群体身高得到的均数，是该人群各个体身高最可能的数值。人群患病率、发病率及死亡率所反映的是该群体每个个体患某病、发生某病及死亡的可能性，这些都是相应的随机公共卫生事件的统计参数。统计学提供了可以定量地估计、比较这类随机事件统计指标的许多技术和方法，从而使在个体中呈现无规律的公共卫生及医学的随机现象，可以用概率语言对偶然性大小进行估计，也就可以利用统计学参数，对在群体中呈现的这些现象进行定量估计、描述、比较和分析。

随机事件的概率本质，一直有两种不同的理解。一种看法认为，随机事件的多种可能结局受到（总体）系统内部大量无法控制的因素影响，这些影响因素及其作用的本质是不可确定性测量的，尽管在个体下其呈现无规律性，但随着抽取样本量的增加，各种可能结局的相对频数却存在极限，也就是说有确定性的概率分布。各种可能结局发生的概率是随机事件（总体）系统的内在本质特征，是随机事件（总体）系统的一个物理常数，与测量方法、系统外部无关。另一种看法认为，随机事件各种可能结局的概率是在已知证据基础上对其相信程度的度量，运用某种测量获得证据之后的验后概率可由 Bayes 定理对验前概率用似然函数（likelihood function）估计（修正）而得。如对随机事件各种可能结局完全不知道时，Laplace 提出了用等概率表示验前概率（Bayes 假设），这时验后概率与测量方法高度相关（与测量似然函数成正比），它不是事件内在本质性特征，而是现有证据支持该随机事件各种可能结局的可相信程度，这是一种从完全无知探索完全知道的指标。

根据第一种看法，统计学就是通过样本信息对总体本质特性进行推断的过程，由此诞生了“推断统计学”，即由样本信息对总体参数作统计推断（statistical inference）。按第二种看法，统计学是依据现有信息进行决策的假设检验过程，由此产生了“Bayes 统计学”，依据测量的似然函数对某假设的可相信程度进行统计学决策（statistical decision-making）。这两种观点都推动了统计学的发展，正如物理学中既可以把光看作“粒子”，又可以将其看作“波”一样。采取互补原理，可以兼收并容，既可将统计学技术看作是对随机事件总体特性的推断，又可看作是决策假设检验的工具。

在公共卫生中常用的统计学技术大致包括：计算各种公共卫生事件发生频率的指标及表达其分布的图表的统计学描述；由样本信息估计总体参数的统计推断；比较样本统计学指标的差异显著性的假说检验；科学研究的统计学设计；应用各种统计学模型进行模拟及分析。近年来，由于计算机和各种应用统计学软件的成功，公共卫生工作者完成复杂的统计学计算已无困难，关键问题在于如何正确理解各种统计学方法的适用条件、基本的统计学思想和原

理,以及选择合适的统计学模型进行分析。例如,在由样本信息估计总体参数的统计学推断中,应注意只有比例这种指标可以按二项分布进行统计学推断,而速率、比这类指标一般不能用二项分布进行推断。样本应是按随机原则抽于同质总体,这包含两个基本假设,即随机抽样和同质总体。因此必须清楚什么是随机抽样,为什么必须随机抽样;什么是同质总体,为什么样本必须来自同质总体。又例如,在应用各种统计学模型并进行参数估计时,尤其是在应用多变量统计模型时,应该掌握应用这些模型的前提(假设)条件、检验原理及方法、所获得参数的意义,这样才不会发生错误。

多种统计学技术对公共卫生中流行病学方法的贡献是工具性贡献,到目前为止主要是在两个方面提供了定量工具,即定量估计随机变异的大小、控制混杂作用,以及定量估计控制混杂后的残存相关性大小。前者用得最多的是各种统计学显著性的假说检验,后者用得最多的是分层分析技术或建立在统计学模型基础上的各种多变量分析技术。流行病学大量数据的分析现都要用计算机处理,并依赖各种应用统计学分析软件。但更重要的是应首先对原始数据进行核对、改错、清理,继而对数据进行描述性分析,适当处理各种丢失值(missing value),最后才是按研究的要求进行统计学检验和估计。近年来发展的等级回归(hierarchical regression),又称多层次回归(multilevel regression),被用于概括各种回归技术,提高统计学模型及统计学估计的适应性及精度。

统计学假说检验的根本目的只是在于评估随机变异的作用有多大。因此,必须正确对待所谓的统计显著性或非常显著性的标准( $P$ 值是否 $<5\%$ 或 $1\%$ ),这是历史习惯形成的,只具有相对意义,不是客观的绝对标准,更不能反映生物学、公共卫生的实际差异大小。此外,经验信息只能“否定(falsify)”某种假说,而不能“证实”某种假说。所以,统计学假说检验只能在一定可信度下去否定无效假设(null hypothesis),并在这个意义上接受备择假设(alternative hypothesis)。具体地说,因为无效假设通常是特异性高的假设,所以它才可能被现有给定的信息所否定。而备择假设却有多种可能,用现有给定的信息不但无法否定,而且只能由于无效假设被否证了,才使得某些可能的备择假设相对为真实的可能性增加了,仅此而已。

运用分层分析技术控制混杂作用,常要用“加权(weight)”方法整合多个分层指标为一个加权平均估计值,这里加权方法的选择是必须关注的。用多元分析技术估计同时校正多个混杂因素后的残存相关性大小时,统计学提供了选择统计数学模型的指南,而选择什么因素、各因素以什么编码方式进入模型,主要应根据流行病学、公共卫生实践的考虑。例如,为了预测目的,可采用逐步统计技术(stepwise statistical procedures),任由计算机“自动”筛选自变量。但如果目的是为了控制混杂作用,就不能单纯依赖逐步统计技术筛选自变量,必须根据流行病学考虑预设必须选入模型的自变量,而忽略逐步统计技术对其的选择。

#### 四、因果推理及偏倚分析

因果概念涉及人类的全部生活,是哲学和各种科学学科有限的几个基本概念之一。它的哲学含义和在各种学科中的定义不尽相同,而且争议从未终止。自然科学有其时空连续性基础上的因果定域(locality)假设,即某事件(结果)只能接受在空间上与其定域相联的另一事件(原因)的影响,不存在超越(空间)作用,也不存在超光速(时间)传递。某事件(原因)的改变必然有(时空)路径(pathway)接触另一事件(结果)而“引起”其变化。

这个从原因到结果的(时空)路径,称为因果路径或因果机制。如因果路径中没有中间事件(即原因直接与结果定域相联)时,该原因称为直接原因(第一亲代事件),该结果是直接结果(第一子代事件)。基础医学学科中的生物学因果关系,通常是指这种机制性因果关系(mechanism causality)。

本节仅根据流行病学的基本原理,简述所谓的“流行病学因果推理”(epidemiological cause-effect inference),“那些能使人群发病概率增加的因素,就可被认为是疾病病因,如其中一个或多个因素不存在,疾病的频率就下降。”如某个被认为是疾病病因的公共卫生可干预性因素,经干预控制后疾病发生频率就下降了,在流行病学因果推理上就认为该因素与疾病的因果关系得到检验“证实”。它不同于直接实验性科学(如物理学和化学)的验证,也有别于基础医学各学科(如病理学和药理学、致病微生物学、毒理学等)中的病因和因果推理。流行病学因果推理,要探索引起由个体组成的群体健康事件中各种原因及可能的因果路径,在该健康事件发生前直接与个体定域相关联的因素都可能是直接原因,通过这些直接原因而影响健康事件发生的其他因素都只可能为间接原因。流行病学因果推理目的,在于找到公共卫生可干预性因素及估计如果控制该因素会产生多大的效果,而不在于过分严格地区分直接原因、还是间接原因。

Rubin 在 1974 年提出了流行病学中直接因果关系的定义。假设存在一个与某个体 A 完全相同的某个体 B,在时间  $t_1$  分别对 A 和 B 施加干预 C 和不施加干预 C,至时间  $t_2$ ,个体 A 与 B 反应值之差即定义为“在时间区间( $t_1$ 、 $t_2$ )内施加干预 C 相对于不施加干预 C 对个体 A 的因果作用。”同理,假设存在一个与某总体 A 完全相同的某总体 B,在时间  $t_1$  分别对 A 和 B 施加干预 C 和不施加干预 C,至时间  $t_2$ ,总体 A 与总体 B 的平均反应值之差即定义为“在时间区间( $t_1$ 、 $t_2$ )内施加干预 C 相对于不施加干预 C 对总体 A 的平均因果作用。”由于现实中不可能存在完全相同的两个个体或总体,因此 B 是虚拟的。Rubin 的因果模型是一种虚拟真实的模型。近年来,一些学者也提出了用与 Rubin 等价的“非真实”(counterfactual)因果定义,又称因果的“潜在结局模型”(potential-outcome model)。这些虚拟真实的因果模型都揭示出流行病学因果关系的相对性,即依赖于(相对于)其所选择的特定干预及对照而存在,没有绝对、抽象的流行病学因果关系。

在探索吸烟与肺癌之间因果关系的实践中,Hill (1965 年)提出了 9 条区分流行病学因果联系(causal association)、还是非因果联系(non-causal association)的标准:即联系强度(strength)、一致性(consistency)、特异性(specificity)、时间顺序(temporality)、生物学梯度(biological gradient)、生物学机制的可解释性(plausibility)、(证据间)无矛盾(coherence)、实验性证据(experimental evidence),以及与先例的类似性(analogy)。这是 Hill 对美国“吸烟与健康(1964 年)报告”中提出的 5 条标准的一个扩展,已经成为当前流行病学因果推理的指南。其中,“时间顺序”,即因果发生的时间顺序,显然是作出因果推理必须条件(绝对标准)。但是,对观察性流行病学研究来说,许多因素的发生时间有时很难确定。其余 8 条都只是使因果关系可能性增加而已,并非必须条件,不能以这 8 条是否存在作为判断因果关系的唯一依据。例如,联系的强度越大,留有未控制的偏倚或误差的可能性就小,越可能为因果性联系。但是,因果联系弱常是因为样本量不足,联系未达到统计学显著性水平而被忽略(假阴性),扩大样本量或者运用 Meta-分析,整合类似数据才有可能检出强度弱的因果联系。此外,严格控制对比各组的同质性,可大大提高流行病学因果联系推理的灵敏性。“血清流行病学”、“遗传流行病学”和“分子流行病学”实验室技术的发展,为建立



“同质”对比各组提供了新的基础。Kenneth J. Rothman 等在《Modern Epidemiology (现代流行病学)》第3版(2008)中介绍了“充分病因及组分病因模型”(model of sufficient and component causes)和“因果图模型”(model of causal graphs)。其中“充分病因及组分病因模型”较好地说明了流行病学多种充分病因及组合病因的因果推理及其偏倚分析,指出传统的单因、单果联系只是特例,并给出在多因、多果假说下估计某单一病因在其中的贡献的方法,即“病因分数(etiological fraction)”计算公式。“因果图模型”属于“结构方程”类模型,取决于建立模拟结构的因果机制,可以较好地说明(时间)纵向的因果理论,是近年来用直接非圈图(directed acyclic graph, DAG)方法探索流行病学复杂因果关系的一种尝试。

通过疾病分布及其相关因素差异的比较和分析进行流行病学因果推理,有两个要素:①样本的内部有效性(internal validity)及从样本推论总体的外部有效性(external validity);②按一定抽样误差(随机误差)估计的统计学相关性、经过偏倚(系统误差)分析到流行病学因果推理。两者结合,首先是必须保证样本的内部有效性,统计学(显著性)假设检验及参数的区间估计方法提供了适宜的抽样误差估计技术;其次在有统计显著性相关的前提下,根据Hill的9条标准,采用流行病学测量的信息偏倚、混杂偏倚的分析技术,完成因果推理的内部有效性评估;第三步才是在因果联系内部有效性的前提下,进行从样本推广到总体的外部有效性评估,要求从可能存在样本选择偏倚、总体混杂偏倚的角度,对总体的因果联系可能性进行评估。这里再次强调,用5%或1%假阳性率作为统计学显著性标准进行决策,弱的因果联系常会被忽略掉。

把“充分病因及组分病因模型”和Rubin的因果定义结合起来,可以理解处理(暴露)组与对照(非暴露)组共同(除暴露外)包括的充分病因组分,以决定暴露-疾病因果性联系的真实强度,此即流行病学意义上的效应修饰(effect modification),这些共同的病因组分又称为暴露的效应修饰因素,是因果联系的本质特性;而处理(暴露)组与对照(非暴露)之间的背景因素差异,是产生暴露对疾病的特异因果作用混杂的必要条件(注意:不是充分条件),可能产生对暴露-疾病因果性联系强度的歪曲(即混杂作用)。因此,从暴露-疾病的相关性向因果性推理过程中,必须发现暴露的效应修饰因素,消除暴露的混杂。两者都可以“影响”暴露-疾病特异性因果联系强度,差别在于前者是对因果联系本质特性的修饰,后者是对因果本质特性的混杂。流行病学上的效应修饰,在统计学中又称交互作用,但其不同于生物学意义上(结构机制)的交互作用,也不同于公共卫生意义上的交互作用。

如处理(暴露)组与对照(非暴露)组按背景因素分层,可认为近似满足了Rubin的因果模型的条件,这时每层内不存在混杂,评价层间因果联系强度有无显著性差异,尤其是评价联系的方向(正、负)是否相同,对判断是否存在效应修饰十分重要。若有显著性差异,且层间联系的方向不同,表示存在效应修饰。只有层间不存在效应修饰时,才可将各层效应合并估计总的联系强度,近似地看作是推广到总体的、校正混杂后的因果联系强度的有效性评估。注意:用于帮助确定层间是否存在效应修饰的技术是“统计学交互作用分析”,如层间联系强度无统计学显著性差异时,说明不存在效应修饰,才可把各层效应合并求出总的联系强度。这样做有两个风险,一是统计学交互作用与其所选择的交互作用模型有关,而公共卫生意义上的交互作用通常只在相加性模型中才有意义,因此,如用相乘性模型无统计交互作用,反而说明存在公共卫生意义上的交互作用。另一是如用假阴性率作为决策的标准,而目前统计学决策却是采用控制假阳率的推断技术。混杂偏倚可以通过改善研究设计来控制,也可以用适宜的统计学方法进行校正,值得注意的是如在因果通路(causal pathway)上