

系列讲座

现代实用流行病学方法

第一讲 探索性研究方法

章扬熙

编者按 近年来, 流行病学进展很快, 研究方法在不断丰富与更新, 研究范围无论在深度方面还是广度方面均有开拓。本刊特邀中国医学科学院皮肤病研究所流行病学名誉教授章扬熙主任医师撰写这一系列讲座, 内容包括探索性、分析性研究的设计与分析, 疾病预测预报方法, 对策措施的研究与效果评价, 生存分析方法, 目标规划方法等。这是本刊“高级现代流行病学培训班”教材的部分内容, 如果读者想深入、全面学习这些新方法、新内容, 欢迎大家参加本刊的培训班的学习。

流行病学是研究人群的疾病与健康状态的分布及其有关因素, 进而探索其发生与分布的原因以及消长规律, 研究制订预防保健的对策和措施并评价其效果的学科。按照认识论的观点, 流行病学方法的类型可概括为探索性研究、分析性研究、实验性研究与理论性研究四个基本类型。

探索性研究 (exploratory study), 又称描述性研究 (descriptive study) (确切地说, 描述性研究属于探索性研究的范畴), 是在于了解疾病与健康状态的分布及其有关因素, 揭示其动态分布的特征, 结合有关专业知识进行分析与推理, 从感性认识提高到理性认识, 形成有关因素 (危险因素或保护因素) 的假设, 而这个假设是否正确还有待于通过分析性研究与实验性研究来加以证实。应当认为, 探索性研究的提法似比描述性研究更准确, 因为描述总有一个目的, 不是为描述而描述, 而这个目的即为通过探索来形成假设。在进行探索性研究时, 首先遇到的问题是健康谱与疾病谱以及疾病模式。

一、健康谱 (spectrum of health): 现代医学的发展, 已改变了“无病即是健康”的概念。按照世界卫生组织的定义, “健康是指机体的、精神的及社会的康乐”。为了研究人群的健康状态, 就要建立机体健康、精神健康及社会健康的测量方法。Breslow等对此做了有益的尝试。

例1: Breslow等建立了七个等级的人群机体健康标准 (表1)。

该氏并应用这个标准于1965年对美国阿拉美达县20岁以上成人及不满20岁的已婚居民6 938人调查了机体健康情况, 所得的人群机体健康谱列于表1的第

表1 人群的机体健康标准及阿拉美达县人群的机体健康谱

健康等级	等级标准	调查人数	构成比 (%)
1	严重残废; 生活不能自理或已不能工作达6个月以上	485	7
2	轻度残废; 工作时间减少或工作性质改变或减少其它活动达6个月以上	554	8
3	无残废, 但在既往12个月内有两种或两种以上的机体损伤或慢性情况	624	9
4	无残废, 但在既往12个月内有一种慢性情况或损伤	1 316	19
5	无上列任何问题, 但报告有一种或一种以上的症状	1 940	28
6	没有任何主诉, 但根据对调查表中问题的回答显示精力不足或一般	1 593	23
7	没有任何主诉, 而且精力旺盛	416	6

末栏。然后, 对此等级资料进行了Ridit分析。Breslow等计算了平均Ridit值(\bar{R})及不同生活习惯人群的 \bar{R} 值, 以反映该人群的健康水平。根据Ridit分析原理, 总人群的 \bar{R} 值为0.5, 某一群组机体健康状况越好, \bar{R} 值越低; 机体健康状况越差, \bar{R} 值越高。该氏对有不同生活习惯者的 \bar{R} 值做了比较, 筛选出与机体健康呈正相关的生活习惯有: (1) 不吸烟; (2) 适量的业余体力活动; (3) 少饮酒; (4) 每日7~8小时睡眠; (5) 保持与身长相适应的正常体重; (6) 吃早餐; (7) 不吃零食。并且上述七项好

的生活习惯具备越多的人群，其健康水平越高（表2）。Ridit分析的计算过程，详见表3。其第（4）栏为第（3）栏调查人数的一半，第（5）栏为第（3）栏的累计值并下移一行所得的数值，第（6）栏为第（4）栏与第（5）栏数值的合计，第7栏为第（6）栏除以总人数，即R值。

表2 具有不同良好生活习惯数人群的R值

	良好生活习惯数							
	0	1	2	3	4	5	6	7
R值	0.82	0.63	0.64	0.58	0.54	0.50	0.46	0.41

表3 Ridit 分析

健康等级 (1)	具备七种好生活习惯人数(f ₇) (2)	调查人数(f) (3)	(3)/2 (4)	累计(3) 下移一行 (5)	(4)+(5) (6)	R值 (6)/总人数 (7)
7	18	416	208		208	0.030
6	21	1593	796.5	416	1212.5	0.175
5	32	1940	970	2009	2979	0.430
4	15	1316	658	3949	4607	0.665
3	11	624	312	5265	5577	0.805
2	4	554	277	5889	6166	0.890
1	2	485	242.5	6443	6685.5	0.965
合计	103	6928	—	—	—	—

R值用下式计算：

$$\bar{R} = \frac{\sum fR}{\sum f} \quad (1)$$

总人群的R为：

$$\bar{R}_{总} = \frac{416 \times 0.03 + 1593 \times 0.175 + \dots + 485 \times 0.965}{6928} = 0.5$$

具备七种好生活习惯人群的R为：

$$\bar{R}_7 = \frac{18 \times 0.03 + 21 \times 0.175 + \dots + 2 \times 0.965}{103} = 0.41$$

余类推。

以上研究说明，良好的生活习惯数与R值呈负相关，与人群的健康呈正相关。但是，由于这个分析基于横断面的现况研究，不能判断两个事件的时间顺序，何为因何为果，遂进行了前瞻性观察。1974年，即9年后对该群组再次做了调查。这次调查对象是原群组中70岁以下存活者5108人，实际应答者为3892人，应答率达76.2%。对应答者与非应答者于首次调查时（1965年）的机体健康状况及其与生活习惯的关系进行了对比分析，结果二者未见显著性差异，说明由无应答而产生的偏性不大。为了更合理地比较不同生活习惯的人群在9年中健康水平的变化，原作者首先把用七个等级表

示的健康谱改为用各等级人群9年期间年龄标化生存率来表示的定量测量指标，然后按性别、年龄及1965年的健康状况对1974年健康状况测量值进行调整。具体的调整方法分两步，第一步是对不同性别和年龄组人群健康状况测量值分别建立回归模型，第二步是计算调整健康状况测量值。

$$\hat{H}(74) = a_i + b_i H(65)$$

$$\text{调整测量值} = H(74) - \hat{H}(74)$$

式中H(65)及H(74)分别为1965年及1974年实际健康状况的定量测量值， $\hat{H}(74)$ 为用上述回归模型计算出来的期望测量值， a_i 为回归方程的截距， b_i 为回归系数。调整测量值为1974年的实测值与期望值之差，此值若为正值说明1974年的实际健康水平比期望值高，健康水平有相对的提高，此值若为负值则反之。研究表明，除吃早餐和不吃零食两个习惯与健康水平之间未见显著正相关外，其余五种生活习惯均与机体健康水平呈正相关，而且好的生活习惯愈多，健康水平愈高，延长平均寿命越长。

二、疾病谱与疾病模式：

1. 疾病谱 (spectrum of disease)：要研究疾病的分布特征，首先要认识疾病，疾病在人群中的表现是多种多样的，构成各种疾病固有的疾病谱。各种传染

病有其固有的感染谱 (spectrum of infection)。当一种传染病流行时,若查出全部暴露者和全部病例,可以有六种情况:①未发生感染;②隐性感染;③轻型疾病;④中型疾病;⑤重型疾病;⑥病死。这六种情况的构成比则构成感染谱。只有认识了这六种情况发生的频率及其相对比,才是认识了该传染病的全貌。有些传染病,如脊髓灰质炎、流行性脑脊髓膜炎、埃尔特霍乱、结核病等以隐性感染为主,目力所及的病例有如只见露出水面的冰山尖顶部分,而隐性感染者有如冰山的大部分隐于水中而无法窥见,流行病学上称之为传染病的冰山概念 ("iceberg" concept)。为了弄清人群的隐性感染情况,往往要进行人群免疫水平的检测。感染谱是宿主与病原体斗争的表现,它受病原体的属、种、型和数量的影响,也受宿主的免疫力和抵抗力的影响,也受环境的影响,不是不变化的,对其规律的研究亦属重要课题。

2. 疾病模式 (disease pattern): 从宏观角度,研究各种人类疾病的总情况,就要研究疾病模式。疾病模式是指在特定时间内,某一国家或地区人群中各类疾病的发病频率、相对比和分布特征。对疾病模式的研究,不仅要摸清疾病发生的频率和相对比,还要探索这种疾病模式形成的原因和影响其变化的有关因素,为制定卫生事业发展规划、疾病防治对策、考核卫生工作效果、确定医学研究方向提供科学依据。自人类诞生以来,就同疾病进行着不懈的斗争,在整个人类的历史过程中,疾病模式的演变大体分为三个阶段:①饥荒、营养不良和传染性、感染性疾病阶段;②降低传染性、感染性疾病阶段;③慢性疾病阶段。

3. 疾病死亡模式 (mortality pattern of disease): 疾病死亡模式是指在特定时间内,某一国家或地区人群中总的和各类疾病死亡的频率及其分布特征。疾病模式的调查工作量很大,由于疾病谱的存在,使疾病发病率的确立,尤其是发病的确立较为困难,对其研究开展较少。疾病最严重的危害是对人群生命的威胁,而死亡与死亡原因的概念是较为明确,容易确立的,所以,许多国家尤其是发达国家把生命统计 (vital statistics) 作为卫生立法的重要组成部分,定期进行疾病死亡模式的研究。目前,世界上已有97个国家开展了生命统计工作,每年向世界卫生组织 (WHO) 报送死因统计年报,WHO每年出版生命统计年鉴并系统分析世界各国疾病死亡模式及其演变特征。

例2: 我国疾病死亡模式的演变特征及有关对策的

探讨: 解放前,我国死亡率为25%~32%,婴儿死亡率为200%,其中城市为120%,平均期望寿命仅为35岁。解放后,随着社会经济与卫生事业的发展,人民生活水平的提高,总死亡率在不断下降,1949年为20%,1957年为10.8%,1965年为9.5%,1990年为6.22%,婴儿死亡率也在逐步下降,1958年为80.8%,1989年为44%。平均期望寿命在逐年上升,1957年为57岁,1975年为66.82岁,1982年为67.9岁,1985年为68.9岁。据中国卫生统计年鉴记载,按卫生部疾病分类法统计,1957年我国部分城市前10位死亡顺位依次为呼吸系病(120.3/10万)、急性传染病(56.6/10万)、肺结核(54.6/10万)、消化系病(52.1/10万)、心脏病(47.2/10万)、脑血管病(39.0/10万)、恶性肿瘤(36.9/10万)、神经系病(29.1/10万)、外伤及中毒(19.0/10万)、其它结核(14.1/10万)。至1986年前10位死亡顺位已发生很大变化,城市依次为心脏病(124.88/10万)、恶性肿瘤(114.69/10万)、脑血管病(114.47/10万)、呼吸系病(48.14/10万)、消化系病(22.85/10万)、外伤(22.30/10万)、肺结核(9.30/10万)、泌尿系病(8.85/10万)、中毒(8.77/10万)、新生儿病(622.5/10万,以出生人数为分母计算的);农村依次为心脏病(154.52/10万)、脑血管病(98.75/10万)、恶性肿瘤(95.74/10万)、呼吸系病(79.34/10万)、消化系病(34.44/10万)、中毒(24.33/10万)、外伤(22.81/10万)、肺结核(21.79/10万)、新生儿病(905.13/10万)、泌尿系病(7.98/10万)。以上是我国部分市、县的资料,这个资料是否符合全国的情况呢?据1990年中国疾病监测年报报道,全国分层分区整群抽样获得的145个疾病监测点近1000万人口的资料,依ICD-9来分类填报死因,用WHO推荐的疾病、外伤和死因的国际统计分类手册,把死因归为56类疾病统计,得前8位死因顺位,城市为心血管疾病(包括脑血管病)、肿瘤、呼吸系病(包括肺心病)、意外死亡、消化系病、感染性疾病、泌尿系病、新生儿病,农村为呼吸系病、心血管疾病、肿瘤、意外死亡、感染性疾病、消化系病、新生儿病、泌尿系病。虽然两者的分类不尽相同,但总的变化趋势是一致的,说明上述两个资料基本上能代表全国情况。

纵观世界疾病死亡模式演变的趋势是由传染病、寄生虫病、消化道和呼吸道感染以及营养不良为主要死因转变为心脑血管病、癌症为主要死因的过程。欧美的发达国家已基本完成了这种转变。非洲、南亚的不发达国家基本上未见显著转变。亚、非、拉的发展中国家则正

处于这种转变过程之中。

我国疾病死亡模式正在演变之中，其特征为：

(1) 总死亡率与婴儿死亡率下降速度较快，平均期望寿命上升速度也较快，已接近发达国家水平，这几项在欧美用了百余年才达到的指标，我国仅用了30多年时间就已基本达到了。(2) 心血管疾病(包括脑血管疾病)、恶性肿瘤的死亡率在急速增高，1990年约为1957年的2~3倍，跃居死因顺位的前三位。若不采取对策，还会继续上升。(3) 传染病与寄生虫病虽已退居死因顺位的第五位(农村)或第六位(城市)，但由于我国处于社会主义初级阶段，工农业生产水平低，经济文化不发达，造成肠道传染病、虫媒传染病等发生和流行的环境和条件还广泛存在，而水改粪管等卫生治本措施还需一个较长时间才能解决。由于这个特点，我国传染病的发病率和死亡率仍比发达国家为高，且有很大波动性，有些病弄不好还会出现反复。这种传染病防治问题与非传染病防治问题并存的局面，势必会持续一个相当长的时期。

显见，防治疾病工作任重而道远，为此提出以下基本的对策建议：

(1) 进一步贯彻预防为主方针，在卫生资源的投入上进一步向预防战线倾斜。

(2) 进一步加强传染病的防治工作，认真做好监测和监督工作，加强对策、措施的评价和研究工作，不断改进和优化对策、措施，加强科学规划和管理，使防病更有效，且省人力财力。

(3) 对非传染病的防治工作需从战略决策的高度来认真研究，加以统筹规划。心血管病和恶性肿瘤多发生于中老年人，培养多年成才的中年人夭折则损失巨大。对这两类病的治疗办法有限，但都是可以预防的。欧美有些国家在这方面作了大量预防工作，取得了控制心血管病和肿瘤发病率下降的显著效果，这些经验值得借鉴。近年来，我国有些地区如天津、辽宁等地开展了恶性肿瘤、心血管疾病的发病报告和不同程度的防治工作，但是，这是一项全国性的大事，要从战略的高度认真研究，予以决策。希望能做好这件事，造福于人民。

(未完待续)

公路车祸发生因素流行病学调查分析

何 荣¹ 高德爱¹ 许汉生¹ 陈汉江¹ 王凡武² 王永如²

我们对汉中市1988~1991年发生的470起(一般、重大、特大)车祸事故进行了流行病学描述，就发生的原因进行了初步探讨。

一、车祸发生情况：在4年中发生车祸470起，平均发生率为134.18/万，1988~1991年分别为186.29/万、150.07/万、104.92/万、110.38/万。470起车祸中伤亡率为95.35/万。其中死亡率为35.11/万，年龄分布以26~36岁者为高，其次51岁以上的老年人，15岁以下儿童最低。车祸伤亡以农民为主，其次为工人和职员、儿童。

二、车祸发生原因分析：470起车祸中发生事故的主要原因系驾驶员责任的占82.76%，骑自行车人过失的占7.44%，非驾驶员开车占5.52%，驾驶员责任主要是违章操作占71.70%，其次为超速行驶占13.19%，无

证开车者占5.53%。车辆以大货车为多，占总事故数的46.60%。车属以企事业单位和个体运输户车祸事故为最高，分别占事故总数的52.34%和28.72%。车祸发生以国、省、县道为主，占事故总数的89.15%，这与道路宽、直、平，视线良好，车流量大，驾驶员及行人易产生麻痹思想有关。车祸事故主要发生在晴天，占事故总数的74.04%。其次为阴天，雨雪雾天最低。470起车祸中直接经济损失515 340元，平均每起经济损失为1 096元以上，特大事故损失超万元也有发生。

(本次调查承蒙汉中市公安交警大队事故处理科韩文轩、王世儒同志的大力协助，谨致谢意)

(收稿：1992-11-20 修回：1993-02-15)

1. 陕西省汉中市卫生防疫站 723000

2. 陕西省汉中地区卫生防疫站