

第二阶段疾病监测点的选取及其代表性

中国预防医科院流行病学微生物学研究所*

杨功焕 郑锡文 曾光 王骊山 陈艳麟 陈爱平 黄正京 葛为民

提要 本文探讨了第二阶段疾病监测点的选取原则, 由于全国各地社会经济的发展水平和自然条件极不平衡, 卫生事业发展水平和人群健康状况存在着明显的差异。为了保证所获得的样本能够推论全国情况, 故采用了分层原则: 1. 保证样本中在不同卫生状况的地区的人口比例与全国类似; 2. 保证地理分布的均衡性, 进行分层整群随机抽样。同时考虑到可行性, 允许抽到的监测点, 不能胜任工作者在同类地区进行调换。并对这些监测点的代表性作了验证。选用人均国民生产总值等指标, 检验证明各层样本均数与该层总体没有显著性区别, 总样本均数与总体没有显著性区别。同时, 1990年疾病监测点报告甲、乙类的传染病发病率虽然是全国疫情系统的2~3倍。但是疾病的构成比是类似的, 月发病率趋势也是相同的。进一步佐证了样本的代表性。由此可以认为, 目前的疾病监测系统基本上能够反映总体人群健康水平。

关键词 监测 随机抽样

中国是一个有11亿人口的发展中国家, 目前在全国各地均进行准确而及时的信息报告还难以做到。因而自1980年起, 在卫生部及中国预防医学科学院的领导及支持下, 由中国预防医学科学院流行病学微生物学研究所牵头, 全国29个省、自治区及直辖市卫生防疫站参加, 逐步建立了全国疾病监测网, 包括71个城乡监测点, 监测人口达1千万, 占全国人口的百分之一。监测系统提供了丰富的信息。但是这些信息还很难代表全国的情况。这是由于这些监测点是自愿参加, 多集中在大城市和富裕农村地区; 而全国各地社会经济的发展水平和自然条件极不平衡, 卫生事业发展水平和人群健康状况存在着明显的差异, 因而监测人群与实际人群比较, 存在以下问题: 农村与城市人口的比例不合理; 各类地区人口的比例不合理。为了使监测系统提供的信息发挥更大的作用, 因而第二阶段全国疾病监测点的确定, 采用了分层抽样, 以求获得一个大致与全国基本情况类似的样本。

抽样原则

一、分层原则:

1. 保证样本中在不同卫生状况的地区的人口比例与全国类似。

根据一定的指标, 把城市农村均再细分为不同类型的城市和农村。

2. 保证地理分布的均衡性。

二、可行性的考虑: 虽然一方面采用分层随机抽样原则, 以求得一个能代表全国情况的样本。另一方面也考虑到现实情况要能长期坚持进行工作。允许抽到的监测点, 不能胜任工作者在同类地区进行调换(实际上大约有15%的监测点在同层的情况下进行了调换)。

抽样方法及结果

一、抽样方法: 采用多阶段分层系统整群抽样, 各层按层的大小(人口数)抽取观察单位数。第一阶段: 以地理区域作为分层标准, 分为13个地理区域, 然后按不同卫生状况的区域进行分层, 有7个层次。共计有71个层次。分层后, 按初级单位含量比例的概率抽样法(PPS), 每一个县或每一个城市为一初级抽样单位, 列出初级单位名单, 人口数及其累计秩序范围, 用系统抽样法确定初级抽样单位所

*北京, 邮政编码102206

在的县、市。在已确定的县、市中，再抽到1~2个乡或几条街道，作为监测区。每一个监测区为2~10万人。

二、选点结果：第二阶段疾病监测点的选取，经过了反复论证，得到了第一批样本点，

经长沙会议、九江会议与各省疾病监测负责人协商，为了保证数据质量，对部分不合适的县在同层进行了调换，从而得到了目前的监测点。

总体和样本中各层人数的构成比列在图1。

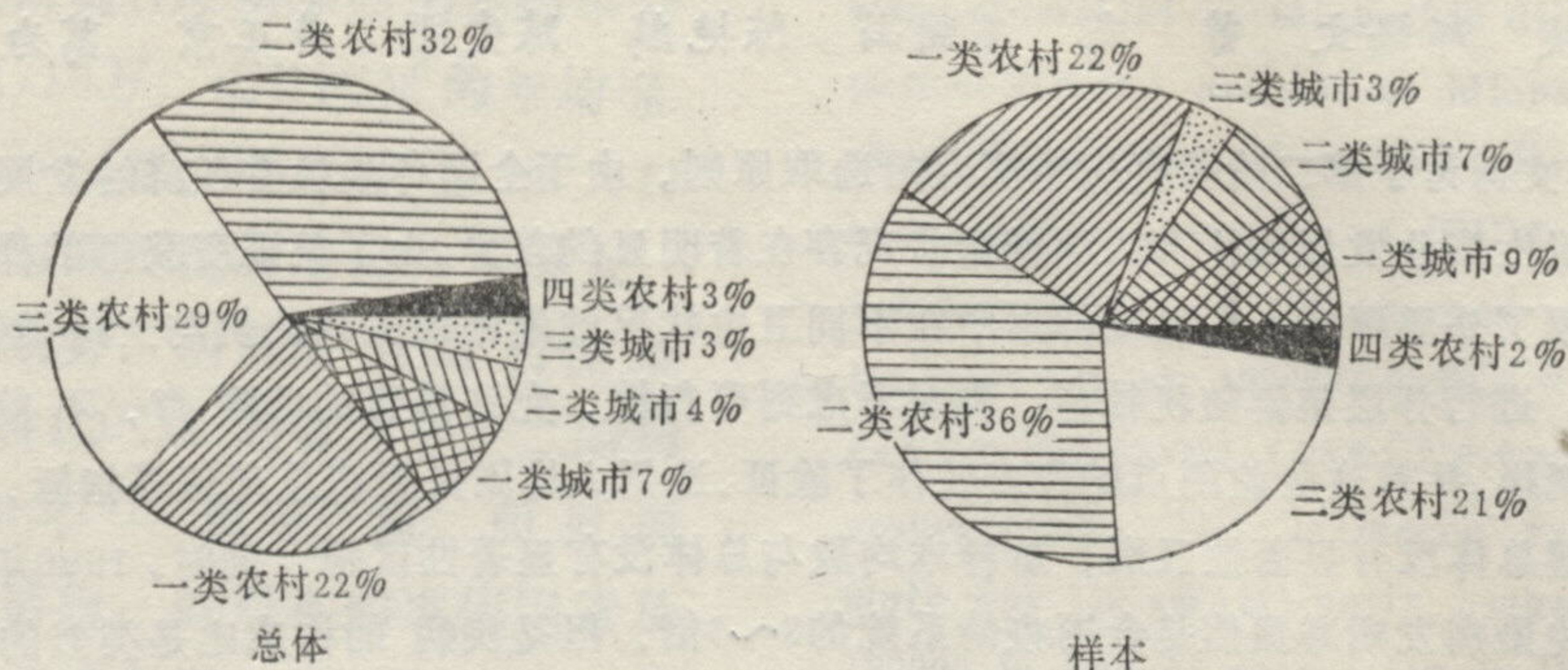


图1 样本和总体人口构成
(根据1982年人口普查资料和监测点资料)

样本均衡性的判断

由于该样本是作为代表全国进行综合卫生信息监测点之长期观察点，因此对于能否代表全国人群的健康水平是必须慎重考虑的。前面已经提到，在抽样过程中，约有15%的监测点进行了调换，因此我们对目前由1000万人口组成的145个监测点，能否代表全国的情况，进行了论证。所谓能否代表全国的情况，是指在同等条件下，这145个监测点所得到的各项数据和2341个县、市所得到的数据是类似的。由于当时我们还不可能有145个监测点人群的数据，暂以145个监测点所在的县、市作为样本。以2341个县、市作为总体。

那么，怎样来判断样本和总体是否类似呢？我们从三个方面进行判断：第一个方面，社会经济背景情况，选用人均国民生产总值，12岁及以上人群的文盲率2个指标；第二个方面，人口结构情况，选用出生率、14岁及以下人口所在百分比、65岁及以上人口所在百分比3个指标；第三方面，居民健康状况，选用死亡率、婴儿死亡率2个指标。这些数据均来自1982年人口普查数据。

经分层后，这几个指标均近似正态分布，我们检验各层这7个指标的均数 X_i 是否来自均数为 X_i 的层总体。经加权后，检验总样本这7个指标的均数 X 是否来自均数为 X 的总体。分别列出按不同卫生状况分层和地理区域分层以及总的样本与总体的均数、标准差及 u 值（由于篇幅所限，本刊略）。从附表可见，在 $\alpha=0.05$ 的水平， u 值均小于1.96， $P>0.05$ 。即是说，该层样本是来自同一层总体。换句话说，该层样本是能代表该层总体。而总样本能代表总体。

同时，1990年疾病监测点开始报告数据，虽然疾病监测系统报告甲、乙类的传染病发病率是全国疫情系统的2~3倍。但是几类疾病的构成比是类似的（图2）；月发病率趋势也相同（图3）。这也进一步佐证了样本的代表性。

由此可以认为，目前的疾病监测系统基本上能够反映总体人群健康水平。

讨 论

在样本均衡性的判断中，我们暗含了一个假定，即如果样本的7个指标的均数与总体没有显著性区别，则表明样本来自于同一总体，

附表

总体和样本N1...N7均数的比较

		总体均值	总体标准差	样本均值	样本标准差	u 值
大城市	n ₁	3337.37	4 148.00	2 652.20	1 403.00	-0.15647
	n ₅	24.59	4.01	24.67	4.28	0.01364
	n ₆	4.61	1.25	4.65	1.32	0.02200
	n ₈	16.84	3.23	15.83	2.60	-0.24358
	n ₉	5.14	1.00	5.10	0.94	-0.02915
	n ₁₀	16.63	6.16	17.23	6.55	0.06673
	n ₁₃	14.50	5.85	13.71	3.73	-0.11387
中等城市	n ₁	2 553.30	1 772.00	3 146.54	2 359.00	0.20107
	n ₅	28.70	4.86	27.58	3.41	-0.18865
	n ₆	3.90	1.13	3.99	1.05	0.05835
	n ₈	18.36	3.99	18.88	3.97	0.09239
	n ₉	4.94	1.16	5.12	1.06	0.11455
	n ₁₀	20.96	11.23	18.92	8.23	-0.14652
	n ₁₃	17.67	9.41	17.32	4.29	-0.03384
小城市	n ₁	1 631.70	1 095.00	1 735.78	1 022.00	0.06949
	n ₅	32.07	4.39	28.99	4.34	-0.49894
	n ₆	5.08	1.34	4.76	1.30	-0.18934
	n ₈	19.63	3.69	17.23	1.99	-0.57246
	n ₉	6.15	1.03	5.39	1.08	-0.50924
	n ₁₀	26.36	11.92	20.00	13.48	-0.35344
	n ₁₃	27.57	8.32	19.37	5.37	-0.82
最富裕农村	n ₁	852.72	488.66	803.35	412.99	-0.07716
	n ₅	32.07	4.39	32.56	4.69	-0.07628
	n ₆	5.08	1.34	5.03	1.49	0.02495
	n ₈	19.63	3.69	19.69	4.38	-0.01048
	n ₉	6.15	1.03	6.19	0.95	0.02855
	n ₁₀	26.36	11.92	27.87	13.11	0.08522
	n ₁₃	27.57	8.32	28.84	9.01	0.10356
富裕农村	n ₁	462.37	168.44	457.67	55.89	-0.02648
	n ₅	35.49	3.35	35.19	3.99	-0.05758
	n ₆	4.75	1.04	4.80	1.09	0.03319
	n ₈	21.79	4.39	22.07	4.02	0.04704
	n ₉	6.41	1.22	6.18	0.96	-0.14816
	n ₁₀	34.41	18.75	30.71	14.71	-0.15526
	n ₁₃	31.97	9.65	29.53	8.76	-0.18722
贫穷农村	n ₁	305.29	83.44	299.65	60.33	-0.05478
	n ₅	37.23	3.25	37.09	2.69	-0.03318
	n ₆	4.62	0.92	4.78	1.07	0.11338
	n ₈	24.22	5.82	22.89	4.21	-0.18516
	n ₉	7.10	1.74	6.32	1.14	-0.37497
	n ₁₀	41.58	24.20	39.35	19.33	-0.07200
	n ₁₃	41.85	12.30	37.17	9.57	-0.30030
最贫穷农村	n ₁	305.24	317.54	199.29	57.44	-0.32833
	n ₅	39.65	3.58	39.76	3.05	0.02339
	n ₆	4.36	1.13	4.06	0.78	-0.21849
	n ₈	32.26	6.76	31.69	6.84	-0.05927
	n ₉	10.89	3.70	10.11	2.19	-0.18141
	n ₁₀	72.42	64.28	67.86	34.12	-0.06266
	n ₁₃	65.69	14.49	59.77	11.99	-0.31477

注: n₁人均国民生产总值; n₅0~14岁人口占总人口的百分比; n₆≥65岁人口占总人口的百分比; n₈出生率(‰); n₉粗死亡率(‰); n₁₀婴儿死亡率(‰); n₁₃≥12岁人群占总人口的百分比

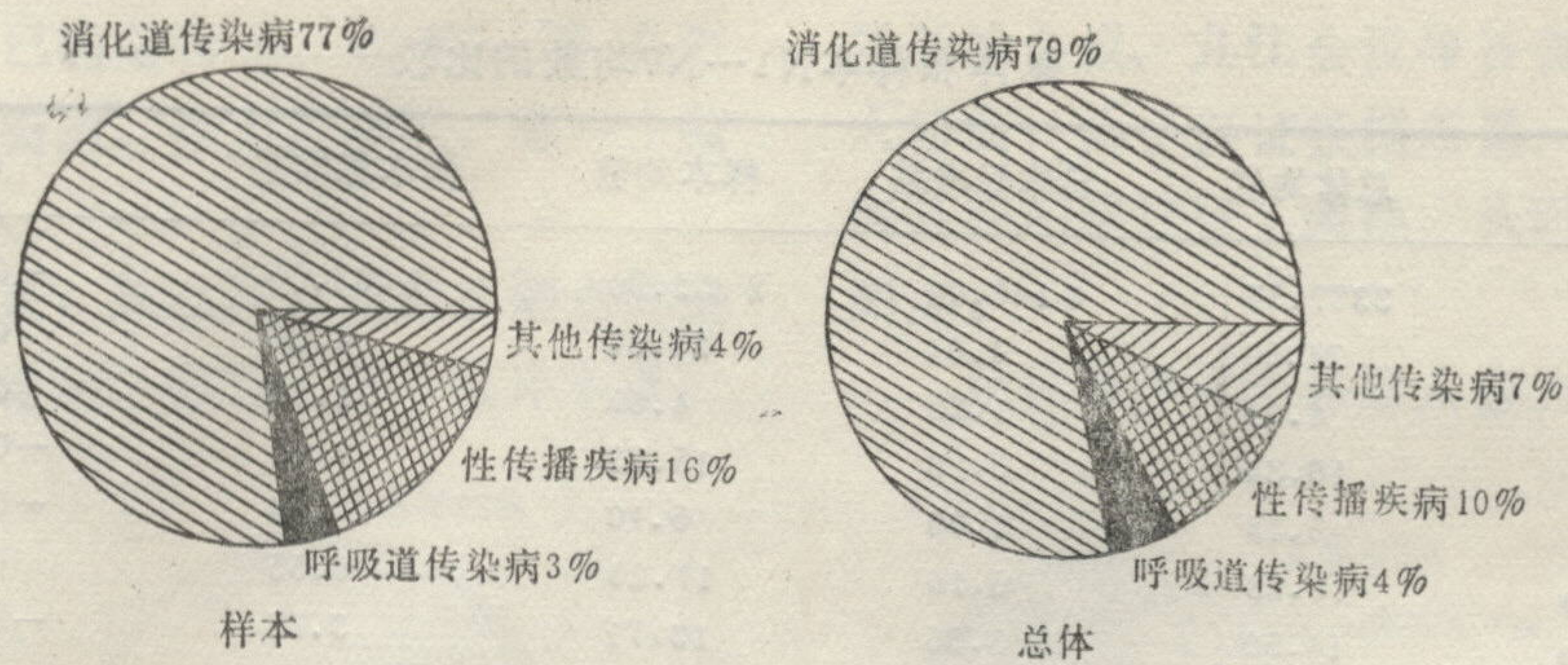


图2 甲、乙类传染病构成
(根据1990年127个监测点和全国资料)

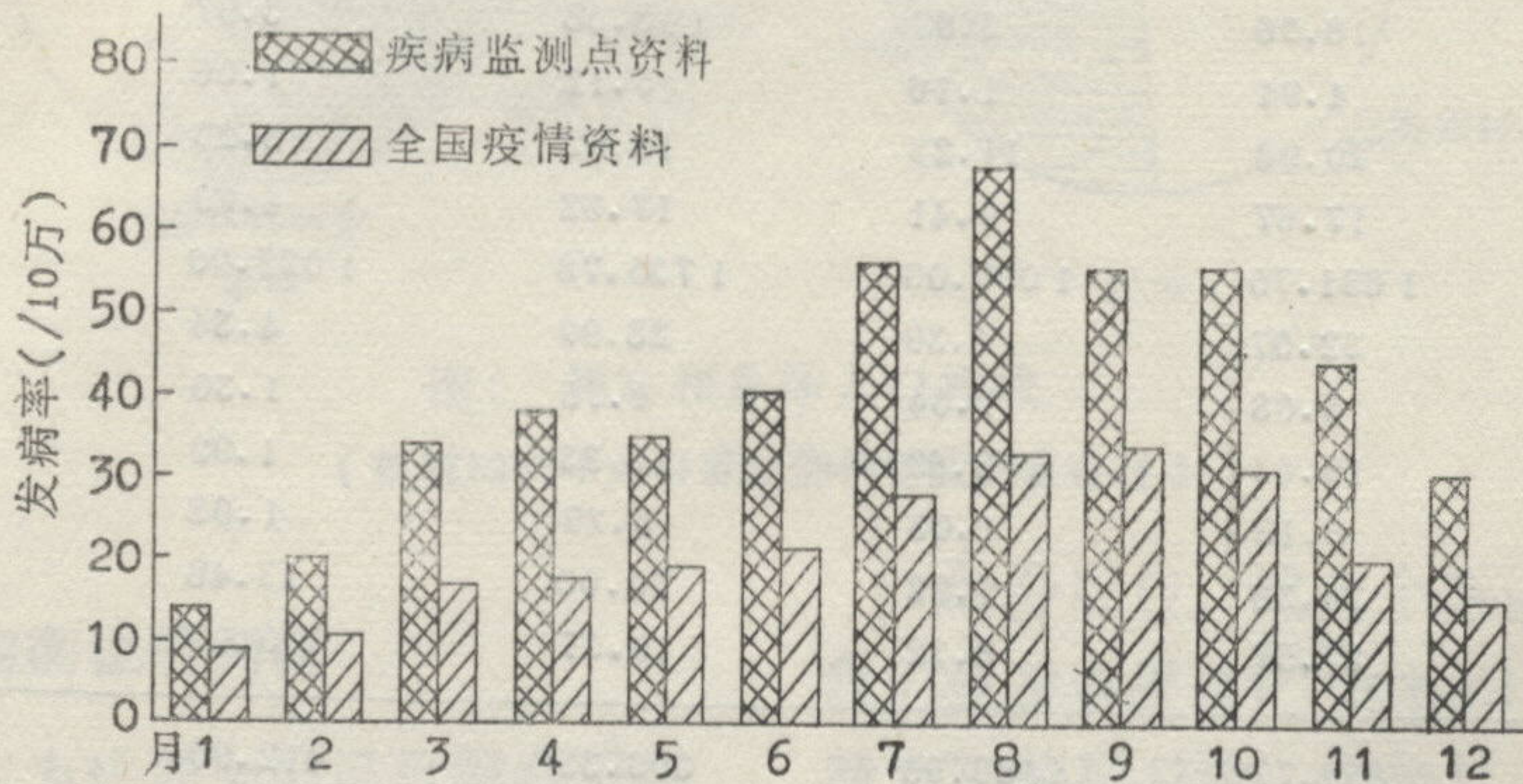


图3 疾病监测点发病率资料和全国资料比较
(根据1990年监测资料)

则样本是可以代表总体。

但具体分析来看，本文的 u 值共有 147 个。其中人均工农业总产值，有 11 个为正值，10 个为负值。文盲率有 2 个为正值，19 个为负值。这表明，从社会经济状况来看，样本与总体是没有倾斜的。15 岁及以下人口比重有 3 个为正值，18 个为负值。65 岁及以上人口比重有 12 个为正值，9 个为负值。出生率有 3 个为正值，18 个为负值。这表明，从人口结构来看，虽然没有显著性区别，但样本中人群的出生率偏低，人口相对老化。死亡率有 4 个为正值，17 个为负值。婴儿死亡率有 3 个为正值，18 个为负值。文盲率有 3 个为正值，18 个为负值。这表明，人群健康水平仍高于总体。因此必须认识到，虽然没有显著性区别，样本仍较总体偏好。

本文利用的是第三次人口普查的资料，主要基于这样一个考虑，各地的数据有可比性，但是由于 10 年来各地的变化不一，不少地区与

1982 年人口普查的资料反映的情况已有很大的不同，因此对结果有一定的影响。从某种意义上说，这种分类并不完全符合实际情况，有必要利用 1990 年人口普查数据进一步验证和调整。

(本研究在选点和本文撰写过程中，得到了田凤调教授、汤旦林教授、章扬熙主任医师的指导，特此致谢)

(1991 年 12 月 10 日收稿，1992 年 3 月 7 日修回)

Selection of DSP Points in Second Stage and Their Presentation Yang Gonghuan, et al., The Institute of Epidemiology & Microbiology, Chinese Academy of Preventive Medicine, Beijing

The principal of selecting DSP points is discussed. Because social and economic development is not balance, developing level of health business and health situation of population are obvious varies in different areas of China. In order to ensure that health information from surveillance population can be used to infer

health situation of national population, the feature of geographic, administrative areas and various health situation in different areas are considered. The principals of selecting DSP are:

1. To assure well-balance of geographic distribution of DSP.
2. To assure well-balance of distribution of DSP in different kind of areas, according to above 9 index: GNP, illiterate rate, birth rate, infant death rate, rough death rate, ratio of 0~14 years people to total people, ratio of about 65 years people to total people,

ratio of labor in industry to total people, and ratio of labor in agriculture to total people. The way of select DSP is: Multistage stratified random sampling (PPS). Meanwhile in order to avoid systematic error, about 15% points which are no qualified were changed. Therefore whether data from DSP can be represented health level of population of whole nation has been tested. The conclusion is that the data from DSP can indeed do so.

Key words Surveillance

《医学微生物电子显微镜图谱》征订启事

中国药品生物制品检定所蔡保健所著《医学微生物电子显微镜图谱》一书, 1991年由华夏出版社出版。我国著名学者高尚萌、刘秉阳教授为本书作序; 严玉辰、杨正时、丁绍卿、李雪东等写了评审意见。

著者从事电镜超微结构研究近30年, 积累了丰富的实际工作经验, 并取得一定科研成果; 书中刊用的450张电镜照片, 是从2万余张照片中精选出来的, 其中部分照片是国内外未见报道的新材料。

全书共4章30节, 计有概论、细菌各论、真菌和病毒各论。

在第一章概论中, 全面介绍了细菌的形态和超微结构。其中包括球菌、杆菌、弯曲菌, 并对细菌的基本构造、细胞壁、细胞质膜、细胞浆以及核区, 细菌的特殊构造, 如荚膜、鞭毛、中介体、纤毛、质粒、微管、微泡以及细菌素等; 细菌的接合、噬菌体、噬菌过程、细菌的分裂繁殖以及DNA病毒和RNA病毒等, 都分别作了介绍, 均附电镜照片及说明。

第二章为细菌各论, 对葡萄球菌、链球菌、脑膜炎球菌、大肠杆菌、沙门氏菌、志贺氏菌、变形杆菌、耶尔森氏菌、蛭弧菌、拟杆菌、霍乱弧菌、弯曲菌、布鲁氏杆菌、鼠疫杆菌、军团菌、芽胞杆菌、结核杆菌、绿脓杆菌等分别作了论述, 并附电镜照片及说明。

第三章为真菌和“三体”各论, 对真菌、黄曲霉菌、桔青霉菌、镰刀菌、白色念珠菌以及钩端螺旋体、立克次体、支原体、衣原体等分别作了论述, 并附电镜照片。

第四章为病毒各论, 对痘病毒、疱疹病毒、腺病毒、多瘤病毒、呼肠孤病毒、反录病毒、脊髓灰质炎病毒、鼻病毒、盖他病毒、马雅罗病毒、西门利克森林病毒、流行性乙型脑炎病毒、登革热病毒、新疆出血热病毒、流行性出血热病毒、乙型脑炎病毒等分别进行论述, 并附电镜照片。

全书以图片及图片解说为主, 内容丰富, 文字简洁, 比较全面地包括了各类型医学微生物的形态和超微结构, 反映了我国目前医学科学发展的水平, 是医学科研、教学单位中、高级专业人员的有益参考资料, 为大专院校微生物教学提供直观教材; 对卫生防疫部门、医学临床单位从事细菌学、病毒学、生物医学、医学教学人员亦具有一定参考价值。

本书为精装16开本, 总印数仅1000册。现已开始办理订购手续, 每册订价55元(含邮费)。汇款收到后, 书及发票即寄。

邮局汇款请寄: 102206北京昌平流字5号

《中国媒介生物学及控制杂志》编辑部(联系人: 陈秀丽)

银行汇款

单位名称: 中国预防医学科学院流行病学微生物学研究所

开户银行: 北京昌平沙河镇办事处; 帐号: 880110。

无论邮局或银行汇款, 请均注明“购电镜图谱”字样, 以防与其他书刊混淆。