

· 现场调查 ·

应用捕获-再捕获方法估计湖南省 洪江区吸毒人数的调查

陈曦 吕繁 郑军 蒲洪波 尹逊强 张艳辉 覃碧云 潘虎

【摘要】 目的 估计湖南省洪江区吸毒人数。方法 应用捕获-再捕获法对洪江区吸毒者人数进行估计,同时,应用最优分配随机分层抽样的方法验证。一种捕获-再捕获法(CR1法)以抽样调查调查人数为第一次捕获,以公安部门登记在册 687 人为再捕获。应用 Seber 调整公式,计算估计吸毒人员总人数。另一种(CR2法)采用戒毒所登记的记录,以 4 个月为间隔,前后两次为捕获-再捕获。最优分配随机分层抽样方法则采用访谈,获得数据。结果 抽样调查共调查 1388 人,发现吸毒人员 24 人,占 1.73%;洪江区总人口为 72 709 人,估计总吸毒人数 1258 人。CR1 法估计总吸毒人数 904 人;CR2 法估计总吸毒人数 1069 人。三种估计数差别无统计学意义,但高于公安部门在册人数 1.3~1.6 倍。结论 捕获-再捕获法可以满足对吸毒人数的估计。

【关键词】 吸毒;捕获-再捕获方法;随机分层抽样

Estimation on the number of drug users in Hongjiang, Hunan: a capture-recapture study CHEN Xi*, LV Fan, ZHENG Jun, PU Hong-bo, YIN Xun-qiang, ZHANG Yan-hui, QIN Bi-yun, PAN Hu. *Hunan Provincial Control of Disease and Prevention, Changsha 410005, China

【Abstract】 Objective To estimate the number of drug users in Hongjiang, Hunan and to develop strategy for drug reduction in the future. **Methods** Two capture-recapture methods were used to estimate the numbers of drug addicts. Random stratified sampling survey was used to verify the optimum allocation. The first capture-recapture method (CR1) referred to the number from optimum allocation random stratified sampling survey conducted in the communities and the number from local Public Security Bureau list being the second capture. The second capture-recapture method (CR2) referred to the collection of records in the detoxification unit with an interval of 4 months. The estimated number was calculated under Seber's adjustment formula. Face to face interview was carried out during the optimum allocation random stratified sampling survey process. **Results** Of 1388 interviewed in the communities, 24 (1.73%) were identified as drug addicts under the optimum allocation random stratified sampling survey. When the figure 1.73% was applied to the total population (72 709) in Hongjiang, the result yielded an estimation of 1258 drug addicts. The estimated numbers of CR1 and CR2 were 904 and 1069 respectively. However, the number was 1.3 to 1.6 fold higher than the reported number (687) by local Public Security Bureau.

Conclusion The capture-recapture method seemed a better method in estimating the number of drug addicts.

【Key words】 Drug users; Capture-recapture method; Optimum allocation random stratified sampling survey

目前常用的 HIV/AIDS 实际例数估计及预测方法大都需要 HIV 感染高危人群基数作为参数,而现有的高危人群基数主要来源于相关部门的资料或依

据经验推测,如吸毒人数来源于公安部门登记在册的吸毒者的资料,与实际吸毒人数有较大的出入,造成对 HIV 流行形势分析不确切,不利于干预对策的制定、干预效果的评价和卫生资源的合理分配。采用科学而有效的统计学方法估计高危人群基数十分重要。为此我们尝试应用捕获-再捕获法对湖南省洪江区吸毒者人数进行估计,同时应用最优分配随机分层抽样的方法印证,结果报告如下。

基金项目:卫生部科技司 2002 年度资助课题项目(WA2002-03-01)

作者单位:410005 长沙,湖南省疾病预防控制中心(陈曦、郑军、张艳辉、覃碧云);中国疾病预防控制中心性病艾滋病预防控制中心(吕繁);湖南省怀化市洪江区卫生防疫站(蒲洪波、潘虎);中南大学公共卫生学院(尹逊强)

资料与方法

1. 背景资料: 洪江区辖 4 个街道办事处、3 个乡镇, 总人口 72 709 人。近年来, 在市场经济冲击下, 工业发展相对滞后, 失业率近 10%, 造成许多年轻人精神和生活上没有依托。加之该区交通便利, 又毗邻桂黔滇三省区, 吸毒现象日趋严重, 1999 年当地公安部门开始建立强制戒毒所, 现有在册吸毒人员 687 人。吸毒人员 90% 以上有共用注射器行为, 共用注射器静脉吸毒及不良的生活行为方式给艾滋病的发生与流行提供了机会。1998 年 4 月, 该区在吸毒人员中发现首例输入性 HIV 感染者后, 艾滋病在该区开始传播, 所有艾滋病感染者及患者均为吸毒。由于性乱和以“性换毒”现象的存在, HIV 有向暗娼人群和各类人群播散的迹象。

2. 方法:

(1) 抽样调查: 根据公安部门的资料, 洪江区吸毒人数占总人口的 0.94%, 参照 Poisson 分布期望值可信限表, 确定样本大小为 1200 人。采用最优分配随机分层抽样的方法, 先将洪江区居民分为城、乡两层, 根据每层的人口确定城、乡的样本量, 城镇和乡村分别为 900 人和 300 人。以居委会(村)为单位进行随机抽样, 每个居委会人口约 300~400 人(不足 300 人的进行合并), 共抽取新街煤炭坝-冒天井、新街二凉亭和高坡街木栗冲 3 个居委会、常青乡岩门村 1 个村共 1597 人, 进行问卷调查。询问调查对象的个人识别信息以及是否吸毒、是否进过当地的戒毒所、认识的人中是否有人吸毒及认识的吸毒人员是否是问卷调查的对象。调查员对吸毒人员再通过当地派出所、居委会、吸毒人员本人及其街坊邻居进行核实, 将核实的吸毒人员(包括进过和未进过戒毒所, 公安部门登记和未登记的)作为调查所获得的吸毒人员。

(2) 以抽样调查人数和公安登记人数进行的捕获-再捕获(CR1): 以抽样调查获得的吸毒人员及公安部门登记在册的吸毒人员, 分别建立包括姓名、性别、年龄、住址、身份证号码等资料的名单, 两份名单录入计算机, 以抽样调查获得的吸毒人员为第一次捕获(标记)的人群(M); 以公安部门登记在册的吸毒人员为第二次捕获的人群(n)。统计出第二次捕获人群中已被标记了的人群(m), 计算洪江区吸毒总人数。

(3) 以强制戒毒所出入所记录进行的捕获-再捕

获(CR2): 根据 1999 年强制戒毒所建所以来出入所的记录, 建立包括有姓名、性别、年龄、住址、身份证号码、出入所时间等资料的名单并录入计算机, 利用 SPSS 10.0 软件进行分析, 确定捕获期及两次捕获的间隔期, 将资料按时间段分组, 捕获期($L1$)里所有登记的吸毒人员即是第一次捕获(标记)的人群(M); 在捕获期($L2$)里所有登记的吸毒人员即为第二次捕获的人群(n)。统计出第二次捕获(n)的人群中已被标记了的人群(m), 估计出洪江区吸毒人数。

3. 统计学分析: 采用 Seber^[1] 的调整公式, 计算总人数 $(N) = [(M + 1)(n + 1)/(m + 1)] - 1$ 。95% $CI = N \pm 1.96[\text{Var}(N)]^{1/2}$ 。方差计算公式为: $\text{Var}(N) = [(M + 1)(n + 1)(M - m)(n - m)] / [(m + 1)^2(m + 2)]$ 。采用 χ^2 检验不同方法间的差异。

结 果

1. 抽样调查: 2003 年 6 月 20 - 30 日共调查 1388 人, 发现吸毒人员 24 人, 占 1.73%; 洪江区总人口为 72 709 人, 估计总吸毒人数 1258 人。

2. CR1: 以抽样调查获得吸毒人员 24 人为第一次捕获(M), 2003 年 8 月洪江区公安部门最新登记在册吸毒人员 687 人为第二次捕获(n), 比较两次捕获的名单, 有 18 人同时出现在两次捕获的名单中, 即 $m = 18$ 。那么, 吸毒总人数为 904 人(95% $CI: 712 \sim 1096$)。

3. CR2:

(1) 两次捕获间隔期的确定: 1999 年 5 月至 2003 年 8 月共有 280 人次 2 次以上进入强制戒毒所, 其中每次间隔时间 < 3 个月有 63 人, 6 个月内 62 人, 9 个月内 40 人, 13 个月内 31 人, 15 个月内 24 人, 18 个月内 19 人, ≥ 18 个月为 41 人。由此可见, 两次或两次以上入所的时间间隔主要集中在 < 3 个月及 3 - 6 个月两个区间, 故两次捕获的间隔期应选择 < 3 个月或 3 - 6 个月。若选择 < 3 个月的区间作为两次捕获的间隔期, 则第一次捕获的吸毒人员在第二次捕获期内仍在戒毒所内未被放出去的人数较多, 为了满足捕获-再捕获法目标人群的所有成员都有同样被捕获的机会这个假设, 故选择 3 - 6 个月为两次捕获间隔期的区间, 以 4 个月作为两次捕获的间隔期。

(2) 捕获期的确定: 根据戒毒所资料, 除极少数

提前解除戒毒的人员外,自 2002 年以来进入该所的戒毒人员的戒毒时间都为 6 个月,故将捕获期定为 6 个月。为了使估计数更接近当前实际,以最近的 6 个月即 2003 年 2 月 28 日至 8 月 30 日为第二次捕获时间。根据两次捕获的间隔 4 个月反推出第一次捕获时间为 2002 年 10 月 31 日至 2003 年 4 月 30 日。

(3) 估计吸毒人员总数: $M = 106$, $n = 109$, 两次捕获的吸毒者中 $m = 10$, 估计洪江区吸毒人员(包括进过和未进过戒毒所的)总数为 1069 人(95% CI: 460~1614)。

4. CR1、CR2 与抽样调查法的比较: CR1、CR2 估算的吸毒人数与抽样调查估算的吸毒人数差异无统计学意义(表 1)。

表1 三种方法估计的吸毒人数比较

方法	调查人数	吸毒人数	吸毒率(%)	χ^2 值
抽样调查	1 388	24	1.73	-
CR1	72 709	904	1.24	2.60
CR2	72 709	1 069	1.47	0.63

5. 估计的吸毒人数与公安部门在册的吸毒人数的比较: 估计的吸毒人数与公安部门在册的吸毒人数比较差异有统计学意义($\chi^2 = 86.57$, $P < 0.005$), 见表 2。

表2 估计的吸毒人数与公安部门在册的吸毒人数的比较

方法	调查人数	吸毒人数	吸毒率(%)	χ^2 值	P 值
公安部门	72 709	687	0.94	-	-
抽样调查	1 388	24	1.73	8.8	<0.005
CR1	72 709	904	1.24	29.9	<0.005
CR2	72 709	1 069	1.47	84.1	<0.005

讨 论

近年来国外曾用捕获-再捕获法估算吸毒人数、麻风患者致畸人数和评估有关先天性风疹综合征、破伤风以及百日咳的死亡率全国性报告的完整性等。以往的工作普遍认为此种方法在用于接触比较困难的对象数量估计有效,因此,世界卫生组织/联合国 AIDS 规划署推荐为用于吸毒人群基数估计的方法之一^[2,4]。

我们研究结果显示,采用捕获-再捕获法估计的吸毒人数是公安部门在册人数的 1.3~1.6 倍,说明当地吸毒的实际形势远比目前政府部门掌握的情况严重。但也没有出现过去经验推断吸毒人数是报告人数的 3~6 倍现象,分析原因不能排除我们选择的

地点范围比较小、对象容易识别和公安打击力度较大等因素。然而,就方法学而言,洪江区两种捕获-再捕获法估计的吸毒人数分别为 904 人和 1069 人,与抽样调查得出的吸毒人数 1258 人差异无统计学意义,则说明此方法是完全可以满足研究的需要。

为了得到可靠的数据,减少偏性,我们认为在应用两次捕获的办法时,必须要求在时间、空间和对象满足三个假设:第一个假设是目标人群应该是封闭的。事实上,没有哪座城市是封闭的,因此这个假设只能在某种程度上得到满足。如果吸毒人群出入该座城市的频率很高,再捕获时,就不太可能包括第一次捕获到的人群,导致重复的人数减少,数量高估的情况就会出现。我们的研究中,利用对调查时间间隔的控制满足相对封闭的需要。CR1 第一次捕获时间为 6 月底,第二次捕获时间为 8 月底,间隔 2 个月,CR2 采用两次捕获的间隔为 4 个月。尽管不能确定出入当地的吸毒者的人数,但是在这段时间内不可能有相当比例的人群出入该地,因而可以认为在研究期间目标人群是相对封闭的。第二个假设是目标人群的所有成员都有同样被捕获的机会。CR1 第一次捕获吸毒人员采用随机的原则,第二次捕获是公安部门登记在册的吸毒人员名单;CR2 吸毒人员两次捕获都是不同时期进过强制戒毒所。第三个假设是被捕获的对象是独立的。我们在进行调查时,CR1 吸毒人员的资料严格遵守保密的原则,不会因为接受了调查而被公安部门抓获;CR2 吸毒人员都是不同时期进过强制戒毒所,这样就满足了调查对象的独立性。

在采用调查访谈进行捕获时(CR1),由于该工作是在社区内进行,询问吸毒是一个敏感的话题,故应注意公安部门未登记的那些吸毒人员可能不承认自己有吸毒行为,造成第二次捕获(公安部门登记在册的名单)中标记数(m)过高,低估吸毒人群的基数;而调查对象提供虚假的个人资料,如姓名、住址、年龄、身份证号码、绰号等,造成第二次捕获(公安部门登记在册的名单)中标记数(n)过低,吸毒人群的基数将被高估。解决由此带来的偏性的方法有三:一是调查前对调查人员进行培训,提高调查人员的访谈技巧,使得他们在调查中更好地与调查对象进行沟通,打消调查对象的疑虑,尽可能获得真实的信息;其次,吸毒人员一般有一定的特征,其街坊、亲友中也有知情者,调查中发现可疑的吸毒人员而其本人不承认的,可以通过其街坊、亲友进一步核实;三

是在调查中捕获的吸毒人员,应对其提供的个人资料从多个渠道进行核实。由于吸毒是一个敏感问题,调查对象可能不愿承认自己吸毒,虽然我们在问卷中通过其他的渠道,如询问街坊邻居及朋友来获得吸毒人员名单,但是抽样问卷调查获得的吸毒人员估计数可能低于实际吸毒人数。而用 CR 估计的吸毒人数与抽样调查获得的吸毒人数间的差异无统计学意义,故在吸毒人员基数估计时应考虑估计的吸毒人员数量低于实际吸毒人员数量。

抽样调查、CR1 和 CR2 估计的都是当地吸毒人员总数,包括进过戒毒所和没有进过戒毒所的吸毒人员。比较 CR1 和 CR2,CR2 工作量小,简便易行,个人资料准确齐全,能准确确定第二次捕获中被标记的人群(n)。但也有观点认为,吸毒人员被公安部门登记在册(抓获)及进过强制戒毒所后,因为公安部门掌握了其情况,再次被抓获而进入强制戒毒所的几率要更大,利用强制戒毒所出入所记录进行捕获-再捕获估计吸毒人员基数时会造成低估。另外,部分吸毒人群的亚人群,如经济条件富裕的吸毒者由于没有被公安部门抓获强制戒毒,在 CR2 进行捕获时有可能会被漏掉。

对吸毒人员数量进行估计是一项十分困难的工作,受到诸多因素的影响。此次应用捕获-再捕获方法进行了尝试,根据我们的体会,认为尽管不容易得到可靠的数据,各地在实践过程中,可以借鉴 CR1 方法用来估算吸毒人群的规模。

参 考 文 献

- 1 Seber GAF. The estimation of animal abundance and related parameters. London: Charles Griffin, 1982. ISBN 0-85264-262-8.
- 2 Mastro TD, Kitayaporn D, Weniger BG, et al. Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method. Am J Public Health, 1994, 84: 1094-1099.
- 3 Broek JD, Jaarsveld T, Rijk AD, et al. Capture-recapture method to assess the prevalence of disabled leprosy patients. Lepr Rev, 2001, 72: 292-301.
- 4 Taylor R. A review of methods for estimating the size of subgroups particularly at risk of infection with HIV and development of proposals which could be used to enumerate these populations in the field with particular reference to the use of capture release methods for estimating populations of IV drug users and prostitutes. Geneva, Switzerland: WHO, 1989.

(收稿日期:2004-04-09)

(本文编辑:张林东)

· 疾病控制 ·

安徽省 36 所幼儿园儿童蛲虫感染流行病学调查

许礼发 赵丹 李朝品

为了解安徽省幼儿园儿童蛲虫感染情况,于 2004 年 10-12 月对全省 36 所市区及郊区幼儿园的儿童进行了蛲虫感染抽样调查。抽样选取安徽省不同卫生状况和地理环境下的 18 所市区幼儿园及 18 所郊区幼儿园 2~6 岁的儿童,共计 737 人,采用透明胶纸肛拭子法查获蛲虫卵(注意事项:检查前一天晚上不清洗肛门,当天清晨尽量不解大便,每位儿童连续 3 天取样)。

结果与分析:城市和郊区共调查儿童 737 人,感染人数为 149 人,感染率为 20.2% (149/737)。其中,市区调查 410 人,感染 73 人,感染率为 17.8% (73/410);郊区调查 327 人,感染 76 人,感染率为 23.2% (76/327)。市区和郊区感染率有差异。各年龄组中 2 岁儿童 161 人,感染 20 人,感染率为 12.4% (20/161);3 岁儿童 190 人,感染 38 人,感染率为 20% (38/190);4 岁儿童 210 人,感染 54 人,感染率为 25.7% (54/210);5~6 岁儿童 176 人,感染 37 人,感染率为 20.2% (37/

176)。各年龄组之间感染率有差异,随着年龄的增长感染率增高。调查男性 379 人,感染 80 人,感染率为 21.1% (80/379);女性为 358 人,感染 69 人,感染率为 19.3% (69/358)。不同性别之间感染率无明显差异。蛲虫的生活史及传播途径很简单,以肛门-手-口直接传播为主。儿童在幼儿园集体生活、学习,而幼儿本身存在的各种不良卫生习惯,使得蛲虫在幼儿园感染率高,传播速度快,且容易交叉反复感染。蛲虫病的主要症状是雌虫的产卵活动引起肛门及会阴部皮肤瘙痒及继发性炎症。患者还常有烦躁不安、失眠、食欲减退、夜惊等表现,长期反复感染,会影响儿童的健康生长。因此,提示我们要改善幼儿园的卫生设施,加强幼儿园的管理和对幼儿的教育,被褥及玩具应定期消毒;同时应提高家长对幼儿养成良好卫生习惯的意识,做到饭前便后洗手,勤洗澡,勤换衣裤,经常修剪指甲,还应定期进行药物驱虫治疗,以降低感染率。

(收稿日期:2005-04-05)

(本文编辑:尹廉)

作者单位:232001 淮南,安徽理工大学医学院病原学与免疫学教研室