

网络规模迭加法在艾滋病人群 规模估计中的应用

王 俊¹, 杨 瑛², 张铁军¹

(1. 复旦大学公共卫生学院, 上海 230032;
2. 上海市闵行区疾病预防控制中心艾滋病性病防治科, 上海 200237)

摘要: 目的: 估计上海市闵行区艾滋病人群(HIV感染者和艾滋病患者)的规模, 了解闵行区艾滋病疫情情况, 制定并实施有针对性的防治干预措施。方法: 采用多阶段随机抽样的方法在上海市闵行区社区人群中招募调查对象, 并进行问卷调查, 利用网络规模迭加法估计艾滋病人群规模。结果: 共调查合格对象3 907人, 其中男性1 912人(48.9%), 女性1 995人(51.1%); 调查对象平均年龄为(40.54±11.69)岁。初步估计上海市闵行区居民社交网络规模均值为206人, 经过反向预测法调整校正后为240人; 进一步依据网络规模迭加法并结合态度法校正得该区艾滋病人群规模为368人(95%CI: 291.19~444.17), 约占上海市闵行区常住人数的0.03%。结论: 采用网络规模迭加法对艾滋病人群的规模进行估计, 其操作简单易行, 但仍需对结果进行适当校正。

关键词: 流行病学; 网络规模迭加法; 艾滋病; 人群规模估计

中图分类号: R181 文献标识码: A 文章编号: 1674-2850(2017)05-0574-06

Application of network scale-up method on estimation of the size of people living with HIV/AIDS

WANG Jun¹, YANG Ying², ZHANG Tiejun¹

(1. School of Public Health, Fudan University, Shanghai 230032, China;
2. Department of AIDS, Shanghai Minhang Center for Disease Control and Prevention, Shanghai 200237, China)

Abstract: Objective: To estimate the population size of people living with HIV/AIDS (PLWHA) in Minhang district, Shanghai, and understand the epidemic of AIDS as well as develop specific prevention measures. Methods: A multiple stages random sampling method was used to recruit eligible participants among general population in Minhang district, Shanghai, and the questionnaire interviews were followed. The population size of PLWHA was further estimated by network scale-up method. Results: A total of 3 907 subjects, including 1 912 (48.9%) males and 1 995 (51.1%) females, were finally interviewed, and the mean age of the subjects was (40.54±11.69). The average social network size of studied subjects from Minhang district was 206. After the adjustment of backward method, it turned out to be 240. Based on the estimation of network scale-up method and attitude adjustment, the population size of PLWHA was 368 (95%CI: 291.19-444.17), which accounted for 0.03% of the total residents of Minhang district. Conclusion: The process of estimation of PLWHA population based on network scale-up method was relatively simple, but the results need adjustment appropriately.

Key words: epidemiology; network scale-up method; AIDS; population size estimation

基金项目: 高等学校博士学科点专项科研基金(20120071120050); 中央高校基本科研业务费专项资金(20520133104)

作者简介: 王俊(1989—), 女, 硕士研究生, 主要研究方向: 传染病流行病学

通信联系人: 张铁军, 副教授, 主要研究方向: 传染病流行病学. E-mail: tjzhang@shmu.edu.cn

0 引言

艾滋病是 HIV 感染所导致的一种慢性传染性疾病, 已在全球肆虐, 截至 2014 年 2 月, 全国报告艾滋病患者 176 602 例^[1]。由于各种原因, 可能会造成艾滋病病例少报、迟报等现象, 这不可避免地使病例报告数与实际数之间存在一定的差异, 而疫情的估计和预测对艾滋病流行现状、防治、干预措施的制定, 以及对卫生资源的分配和优化有极其重要的价值^[2]。因此, HIV 感染者和艾滋病病人的规模估计对于了解人群的实际感染情况具有重要作用。目前, 在中国用于艾滋病疫情估计和预测最广泛的方法为工作簿法 (workbook method), 亚洲疫情模型 (Asia epidemic model, AEM) 和估计与预测软件包 (estimation and projection package, EPP) 模型也逐渐受到关注。本研究采用另一种较新颖的方法——网络规模迭加法^[3]来估计上海市闵行区的艾滋病人群规模。

1 对象与方法

1.1 调查对象

多阶段随机抽样: 首先在上海市闵行区, 按照街道、居委会、小区、单元依次进行随机抽样, 对最终选取的单元进行整群抽样, 同时每户家庭中只抽取一人作为调查对象。调查对象纳入标准为: 1) 在上海市闵行区居住时间不少于 6 个月, 年龄为 18~60 岁; 2) 无任何精神疾病, 能顺利完成调查; 3) 取得知情同意, 愿意参加本研究的上海市常住居民。

1.2 调查方法与内容

由经过专门培训的公共卫生人员担任调查员, 采用面对面的询问式匿名调查方法, 调查内容主要包括: 一般人口学特征、调查对象对一些亚人群规模的了解情况和对艾滋病人群的尊重程度等, 所有调查内容均不涉及调查对象的任何隐私。在参考国外研究^[4-5]的基础上, 本研究将所认识的人定义为: 在上海市居住 6 个月及以上; 见到对方的名字或者看到对方能够互相认识; 在过去 2 年内有过联系, 但要排除在网上交谈而在现实生活中未曾谋面的人。

1.3 网络规模迭加法的计算步骤

网络规模迭加法的理论基础^[4]在于: 假定其他方面均相同, 调查对象所认识的属于目标人群的人数是其所认识所有人数的函数。结合式 (1)^[5]和问卷收集到的数据初步得到调查对象的社交网络规模均值 (c)。

$$\frac{m}{c} = \frac{e}{t}, \quad c_i = \frac{\sum_{j=1}^L M_{ij}}{\sum_{j=1}^L e_j} \cdot t, \quad (1)$$

其中, c_i 为调查对象的社交网络规模数; M_{ij} 为调查对象 i 认识亚人群 j 的人数 (亚人群共有 L 个); e_j 为亚人群 j 的规模; t 为一般人群的总数。同时, 对 c 值进行调整和校正: 首先运用式 (1) 初步计算出调查对象的社交网络规模值; 其次采用反向预测法对其进行校正以提高可信度; 再根据式 (2) (已有研究证明此公式为无偏公式)^[5]

$$e = t \cdot \frac{\sum_{i=1}^l M_{ij}}{\sum_{i=1}^l c_i} \quad (2)$$

和调整校正后的 c 值可得出未知的人群规模; 最后进一步采用尊重程度权重校正系数对艾滋病人群规模进行校正。权重校正系数公式为

$$W_i = \frac{M_i}{M_3}, \quad (3)$$

其中, W_i 为调整的权重; M_i 为调查对象对艾滋病人群某一尊重程度时所认识的艾滋病人数均值; M_3 为调查对象对艾滋病人群一般尊重程度时所认识的艾滋病人数均值。

1.4 统计学分析

采用 EpiData 3.1 建立数据库, 利用 SPSS 18.0 进行数据的整理分析。统计描述主要采用均数±标准差; 利用网络规模迭加法相关公式估计当地艾滋病人群的规模, 同时采用反向预测法和尊重程度对其进行校正。

2 结果

2.1 一般人口学特征

此次调查共收集 4 017 份问卷, 剔除不符合要求 (年龄小于 18 岁或大于 60 岁) 的 110 份问卷, 有效问卷 3 907 份。调查对象的平均年龄为 (40.54±11.69) 岁, 其中男性占 48.9%, 女性占 51.1%。已婚者 3 069 人, 占 78.6%, 文化程度为高中和大学及以上者所占比例分别为 30.5% (1 190/3 907) 和 43.6% (1 705/3 907), 如表 1 所示。

表 1 调查对象的一般人口学特征
Tab. 1 Sociodemographic characteristics of participants

人口学特征		样本量/人	所占比例/%
性别	男	1 912	48.9
	女	1 995	51.1
年龄/岁	18~25	440	11.3
	26~35	1 160	29.7
	36~45	835	21.4
	46~55	977	25.0
	56~60	495	12.7
文化程度	文盲	34	0.9
	小学	242	6.2
	初中	736	18.8
	高中	1 190	30.5
	大学及以上	1 705	43.6
婚姻状况	单身	571	14.6
	结婚	3 069	78.6
	同居	82	2.1
	离婚	146	3.7
	丧偶	36	0.9

注: 由于部分调查者未填写婚姻状况, 婚姻状况样本量加和不为 3 907

2.2 调查对象社交网络规模的估计与校正

利用式 (1) 初步计算出调查对象的社交网络规模均值 (c) 为 206, 标准差为 188.32, 中位数为 152。

根据反向预测法对其进行校正, c 值为 240, 标准差为 212.02, 中位数为 175, 如表 2 所示。

表 2 利用反向预测法对已知人群进行校正

已知人群	官方数据/人	估计值/人	估计值/官方数值	排除或纳入
20~30 岁男性	61 261	105 818	1.73	纳入
15~17 岁男性	5 576	38 164	6.84	排除
>70 岁男性	51 839	37 512	0.72	纳入
20~30 岁女性	54 839	104 053	1.90	纳入
15~17 岁女性	5 462	37 277	6.82	排除
>70 岁女性	62 057	42 292	0.68	纳入
10~13 岁人数	17 467	25 869	1.48	纳入
在读初中生	47 434	17 474	0.37	排除
离婚男性	4 668	3 596	0.77	纳入
结婚男性	12 198	13 602	1.12	纳入
外国人	23 152	3 470	0.15	排除
医生	2 813	13 623	4.84	排除
专职律师	527	2 686	5.10	排除
去世人数	7 257	4 366	0.60	纳入
出生人数	8 598	10 812	1.26	纳入
出国人数	6 500	2 825	0.43	排除
从事零售业	65 662	9 368	0.14	排除
出车祸	226	1 400	6.19	排除
因交通事故死亡	82	609	7.42	排除
无偿献血	19 857	6 931	0.35	排除

2.3 艾滋病人群的估计与校正

调查结果表明共有 307 人 (7.86%) 表示认识艾滋病人群, 根据上述调整校正后的 c 值和式 (2) 可以初步计算出上海市闵行区艾滋病人群规模为 334 人。考虑到不同的对象对待艾滋病这一特殊人群持有不同的尊重程度 (如表 3 所示), 根据问卷的调查结果以及式 (3) 得出对艾滋病人群规模调整校正的权重系数为 1.10, 进一步算得艾滋病人群规模为 368 人 (95%CI: 291.19~444.17), 约占上海市闵行区常住人数的 0.03%。

表 3 不同尊重程度时认识艾滋病人数的均值及权重

尊重程度	样本量/人	认识艾滋病人数的均值	不同程度权重系数
非常尊重	131	0.13	1.63
尊重	651	0.08	1.00
一般尊重	1 489	0.08	1.00
不尊重	842	0.09	1.13
非常不尊重	792	0.06	0.75
总的权重系数			1.10

3 结论

艾滋病人群具有一定的特殊性和隐蔽性^[6-7], 加上社会上的歧视等因素, 导致 HIV 检测率比较低, 感染者具体数量不明, 因此艾滋病人群规模的估计对艾滋病疫情以及卫生资源的分配和优化有极其重要的价值。网络规模迭加法是一类较新的特殊人群规模估计法, 该方法源于 KILLWORTH 等^[3]用于估计一些难以接触到的人群, 国内也有不少学者在艾滋病高危人群中开展了一些有意义的研究^[8-9]。因该方法不需要针对目标人群进行调查, 只需要对全人群进行代表性的抽样问卷调查, 不涉及个人隐私, 相比于传统的针对特殊人群的调查方法如乘法法^[10-11]、捕获-再捕获法^[12]更易获得数据。同时, 该方法可以在同

一调查中对多个人群进行人群规模的估计, 节省人力物力^[13]。但该方法仍处于探索阶段, 可能存在一些偏倚^[14]如传播效应和回忆偏移, 使得该方法会对人群规模的估计产生不同程度的低估和高估^[5], 因此需要通过合适的方法进行校正, 以提高估计的准确性。

本研究通过网络规模迭加法估计上海市闵行区艾滋病人群的规模, 结果显示上海市闵行区艾滋病人群的规模为 368 人 (95%CI: 291.19~444.17), 约占上海市闵行区常住人数的 0.03%, 低于 2011 年全人群感染率的低值 (0.04%), 同时也低于闵行区利用工作簿法估计出的结果 (0.04%)。本次研究结果稍低于其他研究结果, 其原因一方面可能是其他研究的估计结果常包括 HIV 感染者和艾滋病患者, 而其中 HIV 感染者不易被普通人群识别和接触, 在本研究中对调查对象的理解可能更加倾向于艾滋病患者, 因而会有所低估; 另一方面可能是近年来人口流动更加频繁, 当地导入人口增加迅速, 从而引起分母增加, 造成估计偏低。同时值得注意的是, HIV 感染者和艾滋病患者的估计常受到一些社会因素的影响, 虽然研究根据人群认知态度进行了校正, 但仍然可能存在一定程度的低估。

网络规模迭加法中已知人群的选择是基础, 不同的人群选择可能会影响结果的估计, 所以在采用此方法时要慎重选择已知人群。本研究运用反向预测法对已知人群进行筛选, 以尽可能地保证数据的准确性, 从而使结果更具有说服力。同时, 本研究是首次利用网络规模迭加法对上海地区艾滋病人群规模进行估计, 估计结果有待进一步验证。该方法能够对多个人群进行一次性调查估计, 以后可与其他方法配合, 提高方案的科学性, 为高危人群艾滋病防治工作提供参考依据。

[参考文献] (References)

- [1] 中国疾病预防控制中心, 性病艾滋病预防控制中心, 性病控制中心. 2014 年 2 月全国艾滋病性病疫情及主要防治工作进展[J]. 中国艾滋病性病, 2014, 20 (4): 227.
NCAIDS, NCSTD, China CDC. Update on the AIDS/STD epidemic in China and main response in control and prevention in February, 2014[J]. Chinese Journal of AIDS & STD, 2014, 20(4): 227. (in Chinese)
- [2] 吕繁, 张大鹏, 贺雄, 等. 艾滋病高危人群基数估计及其方法[J]. 中华流行病学杂志, 2003, 24 (11): 987-990.
LÜ F, ZHANG D P, HE X, et al. The size estimation and method in high risk population of AIDS[J]. Chinese Journal of Epidemiology, 2003, 24(11): 987-990. (in Chinese)
- [3] KILLWORTH P D, JOHNSEN E C, McCARTY C, et al. A social network approach to estimating seroprevalence in the United States[J]. Social Networks, 1998, 20(1): 23-50.
- [4] KADUSHIN C, KILLWORTH P D, BERNARD H R, et al. Scale-up methods as applied to estimates of heroin use[J]. Journal of Drug Issues, 2006, 36(2): 417-440.
- [5] KILLWORTH P D, McCARTY C, BERNARD H R, et al. Estimation of seroprevalence, rape and homelessness in the US using a social network approach[J]. Evaluation Review, 1998, 22(2): 289-308.
- [6] 郑锡文, 张家鹏, 王小善, 等. 云南省瑞丽市吸毒人群艾滋病病毒感染自然史研究[J]. 中华流行病学杂志, 2000, 21 (1): 17-18.
ZHENG X W, ZHANG J P, WANG X S, et al. The natural history of HIV infection among IDUs in Ruili, Yunnan province, China[J]. Chinese Journal of Epidemiology, 2000, 21(1): 17-18. (in Chinese)
- [7] RANGSIN R, CHIU J, KHAMBOONRUANG C, et al. The natural history of HIV-1 infection in young Thai men after seroconversion[J]. J. Acquir. Immune Defic. Syndr., 2004, 36(1): 622-629.
- [8] 惠珊, 王璐, 郭巍. 网络规模迭加法及其在男男性接触者人群规模估计中的应用[J]. 国际流行病学传染病学杂志, 2010, 37 (6): 405-408.
HUI S, WANG L, GUO W. Estimation on the scale of men who have sex with men using the network scale-up method[J]. International Journal of Epidemiology and Infectious Disease, 2010, 37(6): 405-408. (in Chinese)

- [9] 郭静, 黄晓娟, 王秀彬, 等. 利用网络规模迭加法估计北京市大学生男男性行为人群规模[J]. 中华流行病学杂志, 2013, 34 (11): 1080-1082.
GUO J, HUANG X J, WANG X B, et al. Estimation on the size of men who have sex with men among college students in Beijing through the network scale-up method (NSUM)[J]. Chinese Journal of Epidemiology, 2013, 34(11): 1080-1082. (in Chinese)
- [10] 马焯, 吕繁, 卢培能, 等. 应用乘法法估计暗娼规模的方法学研究[J]. 疾病控制杂志, 2005, 9 (3): 205-208.
MA Y, LÜ F, LU P N, et al. A methodological study on commercial sex workers population size estimation by multiplier method[J]. Chinese Journal of Disease Control & Prevention, 2005, 9(3): 205-208. (in Chinese)
- [11] 杨春霞, 栾荣生, 毛宇嵘, 等. 乘法法对绵阳市暗娼基数的研究[J]. 中国艾滋病性病, 2004, 10 (1): 45-47.
YANG C X, LUAN R S, MAO Y R, et al. The estimation of baseline number of female commercial sex workers (CSWs) in Mianyang city[J]. Chinese Journal of AIDS & STD, 2004, 10(1): 45-47. (in Chinese)
- [12] 郑军, 覃碧云, 刘军, 等. 衡阳市三县(区市)暗娼、吸毒人群规模估计与艾滋病知识、行为调查研究[J]. 实用预防医学, 2011, 18 (2): 211-213.
ZHENG J, QIN B Y, LIU J, et al. Study on the estimated number, AIDS knowledge level and risk behavior of female sex workers and drug users in 3 counties of Hengyang[J]. Practical Preventive Medicine, 2011, 18(2): 211-213. (in Chinese)
- [13] BERNARD H R, HALLETT T, IOVITA A, et al. Counting hard-to-count populations: the network scale-up method for public health[J]. Sexually Transmitted Infections, 2010, 86(Suppl. 2): i11-i15.
- [14] MALTIEL R. Estimating population size using the network scale up method[D]. Seattle: University of Washington, 2013.