

利用网络规模迭加法估计重庆市男男性行为人群规模的研究

包水莲 吴国辉 张维 周超 郭巍 周芯蕾 王璐

【摘要】 目的 利用网络规模迭加法估计重庆市男男性行为者(MSM)人群规模。方法 采用多阶段随机抽样的方法在重庆市开展社区人群的问卷调查,利用网络规模迭加法估计重庆市一般人群的社交网络规模(c),进而根据社交网络规模和调查对象认识MSM的人数及歧视程度,估计和调整重庆市MSM人群的规模。结果 重庆市居民的社交网络规模平均为330人,不同区县($P<0.001$)、性别($P<0.0001$)、年龄($P<0.0001$)及文化程度($P<0.0001$)的社交网络规模差异明显;按照性别、年龄、文化程度校正,稳定性分析及异常值剔除后,居民社交网络规模的均值为310人;MSM人群规模为16 767人(95%CI:14 602~18 932),约占重庆市15~49岁男性人口数的0.21%。结论 网络规模迭加法是一种全新的人群规模估计方法,具有简便和快捷的特性及一定的推广价值。

【关键词】 网络规模迭加法;男男性行为者;人群规模估计

Application on size estimation through the network scale-up method on men who have sex with men in Chongqing municipality BAO Shui-lian¹, WU Guo-hui², ZHANG Wei², ZHOU Chao², GUO Wei¹, ZHOU Xin-lei¹, WANG Lu¹. 1 National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2 Chongqing Center for Disease Control and Prevention

Corresponding author: WANG Lu, Email: wanglu64@chinaaids.cn

This work was supported by grants from the US. Center for Disease Control and Prevention Global AIDS Program (No. 1U2GPS001188) and the Chinese Center for Disease Control and Prevention Foundation for Young Science 2011 (No. 2011A104).

【Abstract】 Objective To estimate the size of men who have sex with men (MSM) in Chongqing municipality. **Methods** A community based investigation was conducted in Chongqing to estimate the number of personal social network (c value) in the general population through the network scale-up method (NSUM). All the participants were selected by multistage sampling randomly. Factors as c value, the number of acquaintance of MSM and its respective coefficient were used to estimate and adjust the size of MSM in Chongqing. **Results** The average size of personal social network for Chongqing residents was 330, with differences among place ($P<0.001$), sex ($P<0.0001$), age ($P<0.0001$), education level ($P<0.0001$). The size of social network changed into 310 after the adjustment on demography, back estimation and elimination of the outliers. The estimated population size for MSM was 16 767 after the adjustment according to the respective levels, with 95% CI from 14 602 to 18 932, which accounted for 0.21% of the total number of males, aged from 15 to 49 in Chongqing municipality. **Conclusion** NSUM was a new method for estimation of population size, which seemed worthy in promoting the practice of NSUM for its efficiency.

【Key words】 Network scale-up method; Men who have sex with men; Size estimation

男男性行为者(MSM)是艾滋病的高危人群之一,2011年估计在中国现存活的HIV/AIDS中,

MSM占17.4%^[1]。由于担心受到歧视,MSM人群较为隐蔽、难以直接接触,应用以往的估计方法如捕获-再捕获法、乘法法存在一定的限制性且成本较高;本研究应用一种较新的估计方法网络规模迭加法(network scale-up method, NSUM),估计重庆市MSM人群规模。

对象与方法

1. 原理: NSUM是一种基于社会结构理论形成

DOI: 10.3760/cma.j.issn.0254-6450.2012.10.009

基金项目: 美国疾病预防控制中心全球艾滋病项目(1U2GPS001188); 中国疾病预防控制中心2011年青年科研基金(2011A104)

作者单位: 102206 北京, 中国疾病预防控制中心性病艾滋病预防控制中心(包水莲、郭巍、周芯蕾、王璐); 重庆市疾病预防控制中心(吴国辉、张维、周超)

包水莲、吴国辉同为第一作者

通信作者: 王璐, Email: wanglu64@chinaaids.cn

的用于估计难以计数人群的方法^[2],其理论基础是:假定其他方面等同,调查对象认识其社交网络中属于某个亚人群的人的概率等于该亚人群占调查地区一般人群总数的概率^[3]。用公式表示为^[2]:

$$\frac{m}{c} = \frac{e}{t} \quad (1)$$

其中 m 为调查对象所认识的属于某个亚人群的人数, c 值为调查对象的社交网络规模, e 值为亚人群的规模, t 代表一般人群总数。

c 值的估计是 NSUM 方法的关键,目前 c 值的估计包括已知人群法和总和法两种^[4],本研究采用已知人群法进行 c 值的估计,利用最大似然法估计个人社交网络规模的公式^[2]:

$$c_i = \frac{\sum_j m_j}{\sum_j e_j} \cdot t \quad (2)$$

其中 c_i 为调查对象的社交网络规模, m_j 为调查对象所认识的亚人群的人数, e_j 为亚人群的规模, t 为一般人群的总数。

2. 主要定义:本研究“认识的人”的定义以国外研究为基础并根据重庆市具体情况进行部分修正^[2-4]:在重庆市居住时间 ≥ 6 个月;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。

3. 已知人群数据收集:本研究拟采用 20~24 岁男性人数、70 岁以上女性人数、公务员、残疾人等 19 个有明确规模数据的人群作为已知人群,对 c 值进行估计。人群规模数据来自于重庆市 2011 年统计年鉴^[5]、重庆市公安局、人事局、妇幼保健院及残联。

4. 现场调查:①研究对象入选标准:在重庆市居住时间不少于 6 个月,年龄 18~60 岁,无任何精神疾病能顺利完成调查,知情同意愿意参加本研究。②抽样方法:采用多阶段随机抽样的方法,抽取重庆市 40 个区县中 2 区 1 县;在城区按照街道、居委会、小区、单元进行抽样,在农村按照镇、村的方式进行抽样,对最终选取的单元和村进行整群抽样,每户家庭中只抽取一人作为调查对象,现场调查时间为 2011 年 9—10 月。③调查方法:采用自填问卷的方式。

5. 统计学分析:采用 EpiData 3.1 软件建立数据库,进行数据双录入,用 SPSS 18.0 软件对数据进行整理和分析。

(1) c 值的调整(按以下步骤逐步进行):①按照人口学特征校正:以 2010 年重庆市 18~60 岁常住人口的性别、年龄、文化程度构成为标准,对调查对象

认识的已知人群的人数进行校正,进而调整 c 值。②反向预测:先从已知人群中剔除一个人群,而后利用其他已知人群估计得到的 c 值反推被剔除人群的规模^[3];保留反向预测值与官方数据的比值在 0.5~2.0 之间的人群进行 c 值的估计。③异常值剔除:对 c 值进行对数转换并利用箱式图剔除异常值。

(2) MSM 规模估计和校正:在借鉴乌克兰研究的基础上^[6],通过收集调查对象对 MSM 的尊重程度对 MSM 的估计结果进行校正;尊重程度划分为 5 层:1=非常尊重,2=尊重,3=一般尊重,4=不尊重,5=非常不尊重;假设当调查对象给予 MSM 中等程度的尊重时,应答结果较为准确,利用公式(3)为不同尊重程度的调查对象分配权重:

$$W_i = M_i / M_3 \quad (3)$$

其中 W_i 为调整权重, M_i 为给予 MSM 某一尊重程度时调查对象所认识 MSM 人数的均值, M_3 为给予 MSM 一般尊重的调查对象认识 MSM 人数的均值。

结 果

1. 人口学特征:共收集问卷 3026 份,其中有效问卷 2957 份,问卷有效率为 97.72%。在巫山县调查 998 人(33.75%),北碚区 1313 人(44.40%),沙坪坝区 646 人(21.85%);男性 1173 人(39.67%),女性 1784 人(60.33%);调查对象的平均年龄为(39.21±11.51)岁,年龄分布较为均匀,已婚比例为 82.36%。

2. 社交网络规模(c)的估计和调整:

(1)利用原始数据估计:利用公式(2)估计出个人社交网络规模的均值为 330 人,标准差为 367 人,中位数为 208 人;利用多组独立样本比较的秩和检验对不同人口学特征的社交网络规模进行比较,结果显示不同区县、性别、年龄组和文化程度居民的社交网络规模差异有统计学意义(表 1)。

(2)人口学特征校正:按照人口学特征对 c 值进行校正,校正后个人的社交网络规模的均值为 311 人,标准差为 360 人,中位数为 195 人。

(3) c 值的稳定性:本研究利用反向预测的方法评价 c 值估计结果的稳定性,结果显示不同的已知人群估计 c 值的结果差异明显,其中律师人群高估最为明显,残疾人数低估最为明显;以反向预测值和官方数据的比值在 0.5~2.0 之间作为标准,共保留 8 个已知人群(20~24 岁男性、20~24 岁女性、70 岁以上女性、在读高中生、2010 年结婚、2010 年死亡、生育女性、拘留或判刑者)用于 c 值的估计,均值为 311 人,标准差为 370 人,中位数为 195 人。

表 1 重庆市 MSM 社交网络规模人口学特征

人口学特征	样本人数	个人社交网络规模(均值)	P 值
调查地区			<0.0001
巫山县	998	397	
北碚区	1313	280	
沙坪坝区	646	326	
性别			<0.0001
男	1173	383	
女	1784	295	
年龄组(岁)			<0.0001
18~	71	552	
20~	630	487	
30~	833	331	
40~	769	293	
50~60	654	195	
文化程度			<0.0001
文盲	116	159	
小学	516	204	
初中	919	251	
高中或中专	756	300	
大专及以上	650	605	
婚姻状况			<0.0001
同居	36	610	
在婚	2437	299	
未婚	391	512	
离异或丧偶	93	249	
居住时间(年)			0.008
0.5~1	140	298	
1~3	387	320	
>3	2430	333	

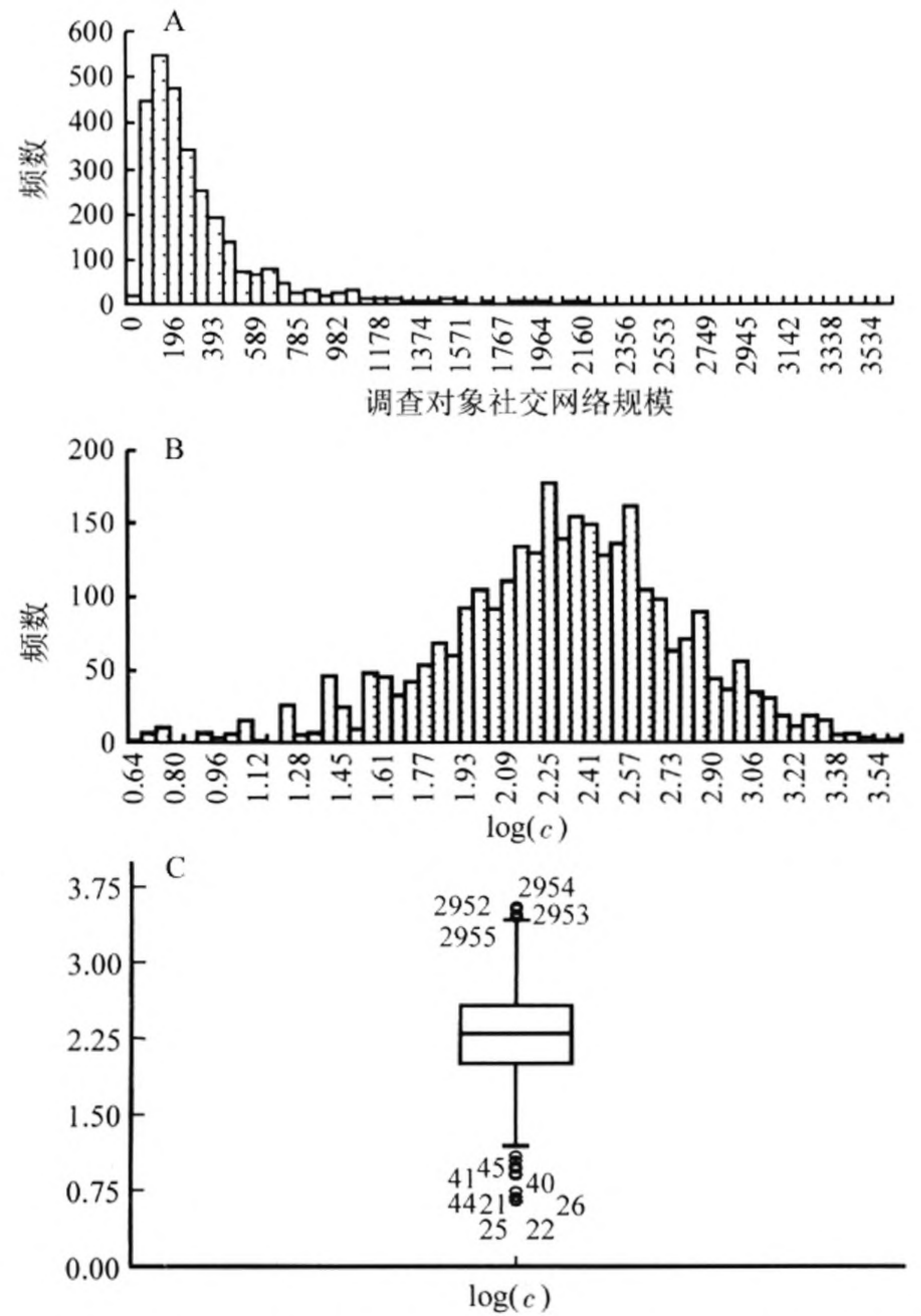
(4) *c* 值分布和异常值处理: 研究结果显示 *c* 值的频数分布呈左偏态分布(图 1-A), 对 *c* 值进行对数转换并利用箱式图(图 1-B、C)剔除异常值后, 结果显示 $\log(c)$ 符合正态分布 ($P > 0.05$), 此时相对应 *c* 值的均值为 310 人, 标准差为 341 人, 中位数为 199 人。利用上述方法对 *c* 值进行逐步调整的结果见表 2。

表 2 重庆市社交网络人口数据 *c* 值逐步校正情况

<i>c</i> 值校正过程(逐步进行)	均值	标准差	中位数
原始数据	330	367	208
按照人口学特征校正	311	360	195
反向预测	311	370	195
异常值剔除	310	341	-

3. MSM 规模估计与校正: 共有 115 人 (3.89%) 表示“认识的人”中有 MSM (问卷中体现为“男同性恋”), 调查对象认识 MSM 的平均人数为 0.08; 结合上述调整后的 *c* 值, 利用公式 (2) 初步估计出重庆市 MSM 人群规模为 7452 人。

考虑到 MSM 人群的隐匿和社会对该人群的歧视, 按照尊重程度对调查对象认识 MSM 人群的人数



注: A: 调查对象社会网络规模分布; B: 调查对象社交网络规模对数转换后分布; C: 调查对象社交网络规模异常值分布

图 1 重庆市 MSM 社交网络规模 (*c* 值) 的分布与转换

进行分层并按照公式 (3) 计算调整权重, 得出调整权重 $1/W_i$ 为 2.25, 由此估计出重庆市 MSM 人群规模约为 16 767 人, 约占重庆市 15~49 岁男性人口数的 0.21%。

讨 论

艾滋病相关高危人群规模是艾滋病防治工作的一项重要数据, 本研究利用 NSUM 对重庆市居民的社交网络规模进行估计, 进而估计该市 MSM 人群的规模。研究结果显示重庆市居民社交网络规模 *c* 值的分布呈左偏态分布, 与国外研究中 *c* 值的分布情况一致^[2,7]。不同城区和人群社交网络规模存在显著差异, 经过调整后, 重庆市居民社交网络规模约为 310 人, 该结果与美国利用两种调查方法获得的 290 人较为接近^[4]。

MSM 人群的规模估计是艾滋病研究中公认的难题, 目前国内较为常用的方法包括乘数法与捕获再捕获法, 但国内的研究结果差异非常大, 张北川等^[8]的研究显示, MSM 约占男性人口的 10%~15%, 在北京和哈尔滨市进行的相关研究结果显示^[9], 北京市 MSM 规模数占当地 15~49 岁男性人口数的 1.0%, 哈尔滨市则为 0.3%; 在上海市两个网站进行

的研究显示^[10],估计的 MSM 人数占 15~49 岁男性人口的 7.1% 和 6.6%。本次调查通过人口学特征校正、稳定性分析、异常值剔除和尊重程度分层加权的方法对 c 值和目标人群规模估计结果进行调整,结果显示,重庆市 MSM 人群的规模估计值为 16 767 人,按照 NSUM 中单纯随机抽样方法计算 95%CI 为 16 767(14 602~18 932) 人,估计 MSM 人数占重庆市 15~49 岁男性人口数的 0.21%,略低于哈尔滨市利用乘数法的研究结果。可能原因为:①MSM 人群较为隐匿,本研究的调查对象为社区居民,其认识的 MSM 人群可能是最近两年内表现最为活跃的 MSM 人群。②MSM 对社区居民是一个较生僻的概念,包括男同性恋者、男双性恋者和部分男异性恋者^[11];在实际调查中,考虑到调查对象对 MSM 的理解程度及“性”的敏感性,应用“男男同性恋”进行了替代,可能会低估 MSM 的规模。③MSM 人群流动性较大,而本研究的估计结果为重庆市常住人口中的 MSM 人群,因而与其他研究相比,估计结果可能偏低。

国外最近的研究显示,应用 NSUM 进行估计时,容易高估规模较小的人群和低估规模较大人群^[2],造成这种情况的原因主要有三种偏倚:屏障效应(barrier effect)、传播效应(transmission effect)和估计效应(estimation effect)^[4]。本研究在应用 NSUM 对 MSM 人群进行规模估计时,虽然对“认识的人”有严格的定义,但由于社会对 MSM 人群的普遍歧视^[12],MSM 人群在表露自己的身份时有一定的选择性。有研究表明,MSM 人群仅向同志圈内及少数要好的朋友坦白自己的身份^[11,13],且更多的 MSM 倾向在网络上寻找性伴侣^[10]。这些现实情况,使得本研究仍然会受到传播效应和屏障效应偏倚的影响;此外由于调查对象很难在短时间内准确回忆出认识属于某个人群的人数,由此会造成信息偏倚而使估计结果不准确。

NSUM 方法是一种全新的高危人群规模估计方法,在国际上正处于探索阶段,与传统的估计结果相比,NSUM 估计结果存在不同程度的低估或高估^[6,14,15]。本研究是在我国开展的第一个现场研究,处于方法的试探阶段,估计结果仍有待进一步验证。与传统人群规模估计方法相比,该方法以社区人群为调查对象,有明确的抽样框架,可以一次对多个人群进行估计^[16]。但是,该方法对抽样方法、样本来源和现场质量控制均有较高的要求,另外,该方法几种偏倚的评价方法也处于探索之中。考虑到该方法能够对多个人群进行快捷估计,如果能够对调查

现场进行严格控制,或能够与人口学调查或营养调查等工作相结合,仍不失为一种有效的方法。

参 考 文 献

- [1] Chinese Ministry of Health, UNAIDS, WHO. 2011 estimates for the HIV/AIDS epidemic in China. Beijing: 2011. (in Chinese) 中华人民共和国卫生部,联合国艾滋病规划署,世界卫生组织. 2011 年中国艾滋病疫情估计. 北京: 2011.
- [2] Killworth PD, McCarty C, Bernard HR, et al. Estimation of seroprevalence, rape and homelessness in the US using a social network approach. *Evaluation Review*, 1998, 22: 289-308.
- [3] Kadushin C, Killworth PD, Bernard HR, et al. Scale-up methods as applied to estimates of heroin use. *J Drug Issues*, 2006 (2): 417-440.
- [4] McCarty C, Killworth PD, Bernard HR, et al. Comparing two methods for estimating network size. *Human Organization*, 2001, 60(1): 28-39.
- [5] Chongqing Statistic Bureau. Chongqing statistical yearbook of 2011. Chongqing: China Statistics Press, 2011. (in Chinese) 重庆市统计局. 重庆市 2011 年统计年鉴. 重庆: 中国统计出版社, 2011.
- [6] Paniotto V, Petrenko T, Kupriyanov V, et al. Estimating the size of populations with high risk for HIV using the network scale-up method [EB/OL]. http://nersp.osg.ufl.edu/~ufruss/scale-up/Ukraine_Final%20Report_scale-up_eng_July24.pdf.
- [7] McCormick TH, Salganik MJ, Zheng T. How many people do you know? Efficiently estimating personal network size. *J Am Stat Assoc*, 2010, 105: 59-70.
- [8] Zhang BC, Li XF, Shi TX, et al. The preliminary estimates of the number and HIV infection rate for gay men/ bisexual men in China. *Chin J STD/AIDS Prev Cont*, 2002, 8(4): 197-199. (in Chinese) 张北川,李秀芳,史同新,等. 对中国男同/双性爱者人口数量与艾滋病病毒感染率的初步估测. *中国艾滋病性病防治*, 2002, 8(4): 197-199.
- [9] Wang LY, Xia DY, Wu YH, et al. Application of a multiplier method to estimate the population size of men who have sex with men. *South China J Prev Med*, 2006, 32(3): 9-11, 15. (in Chinese) 王丽艳,夏冬艳,吴玉华,等. 乘数法估计北京、哈尔滨两市男性同性恋人群规模的研究. *华南预防医学*, 2006, 32(3): 9-11, 15.
- [10] Ning Z, Pan QC, Zheng XH, et al. Study on using the multiplier method in estimating the size of men who have sex with men population in Shanghai. *Chin J Epidemiol*, 2007, 28(9): 848-850. (in Chinese) 宁镇,潘启超,郑晓虹,等. 乘数法估计上海市男男性接触者人群规模. *中华流行病学杂志*, 2007, 28(9): 848-850.
- [11] Wang QY, Suo LD, Li XY. Study on the background and status of homosexual behavior. *J Pub Health Prev Med*, 2006, 17(5): 44-45. (in Chinese) 王全意,索罗丹,黎新宇. 男男性行为的历史背景及现状. *公共卫生与预防医学*, 2006, 17(5): 44-45.
- [12] Wang Y, Zhang HB, Xu J, et al. Analysis on current situation of discrimination experience and feeling of MSM. *Pract Prev Med*, 2010, 17(4): 636-638. (in Chinese) 王毅,张洪波,徐杰,等. 男男性行为者歧视经历及感受现状分析. *实用预防医学*, 2010, 17(4): 636-638.
- [13] Yang L, Xie YJ. A quality research on gay self-identity development. *Chin J Human Sexuality*, 2011, 20(5): 42-45, 58. (in Chinese) 杨玲,谢远俊. 男同性恋者自我认同发展的质性研究. *中国性科学*, 2011, 20(5): 42-45, 58.
- [14] Salganik MJ, Fazito D, Bertoni N, et al. Assessing network scale-up estimates for Moldova UNAIDS Country Office. Size estimation for high risk groups Moldova. <http://aids.md/aids/files/166/report-estimations-sizes-vulnerable-populations-mission-2010-en.pdf>.
- [15] Salganik MJ, Fazito D, Bertoni N, et al. Assessing network scale-up estimates for groups most at risk for HIV/AIDS: Evidence from a multiple method study of heavy drug users in Curitiba, Brazil. *Am J Epidemiol*, 2011, 174(10): 1190-1196.
- [16] Bernard HR, Hallett T, Lovita A, et al. Counting hard-to-count populations: the network scale-up method for public health. *Sex Transm Infect*, 2010, 86 Suppl 2: S11-15.

(收稿日期: 2012-05-02)

(本文编辑: 尹廉)