ORIGINAL ARTICLE

Estimating the size of key populations for HIV in Singapore using the network scale-up method

Alvin Kuo Jing Teo ⓘ ,[1] Kiesha Prem,[1] Mark I C Chen,[1,2] Adrian Roellin,[1,3] Mee Lian Wong,[1] Hanh Hao La,[4,5] Alex R Cook[1,3]

[1]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[2]Infectious Disease Research and Training Office, National Centre for Infectious Diseases, Singapore
[3]Department of Statistics and Applied Probability, National University of Singapore, Singapore
[4]Center for High Impact Philanthropy, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[5]Center for Public Health Initiatives, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence to
Dr Alex R Cook, National University Singapore Saw Swee Hock School of Public Health, Singapore 117549; alex.richard.cook@gmail.com

AKJT and KP contributed equally.

ABSTRACT

Objectives To develop a localised instrument and Bayesian statistical method to generate size estimates—adjusted for transmission error and barrier effects—of at-risk populations in Singapore.

Methods We conducted indepth interviews and focus group to guide the development of the survey questionnaire. The questionnaire was administered between July and August 2017 in Singapore. Using the network scale-up method (NSUM), we developed a Bayesian hierarchical model to estimate the number of individuals in four hidden populations at risk of HIV. The method accounted for both transmission error and barrier effects using social acceptance measures and demographics.

Results The adjusted size estimate of the population of male clients of female sex workers was 72 000 (95% CI 51 000 to 100 000), of female sex workers 4200 (95% CI 1600 to 10 000), of men who have sex with men 210 000 (95% CI 140 000 to 300 000) and of intravenous drug users 11 000 (95% CI 6500 to 17 000).

Conclusions The NSUM with adjustment for attitudes and demographics allows national-level estimates of multiple priority populations to be determined from simple surveys of the general population, even in relatively conservative societies.

INTRODUCTION

The HIV epidemic in Singapore is classified as low level but is likely to be a concentrated epidemic among key populations such as men who have sex with men (MSM).[1 2] Most new HIV infections occur among men, and 95% of cases were acquired through sexual transmission in 2016.[3] Like many Asian countries, Singapore does not currently have a systematic approach to collect data on the size of key populations that are most likely to acquire and transmit HIV. Without these populations size estimates, it is challenging to generate reliable projections of the number of people with HIV, effectively plan or evaluate prevention and treatment responses, or allocate resources.

The network scale-up method (NSUM) is a relatively new and promising approach to estimate the size of hard-to-reach populations at risk of HIV/AIDS.[4] It involves estimation of the personal network size of a representative sample of the general population and, with this information, estimation of the number of members of a hidden subpopulation. This method has been applied in various countries to estimate sizes of at-risk populations. It was simple to implement and less expensive,[5] and yielded important information for individual country-level HIV prevention programme planning, monitoring and evaluation.[6–8] Given these successes, the NSUM may be suitable for behavioural surveillance in Singapore, and it is therefore of interest that we evaluate the feasibility of using a similar strategy locally. Findings from this study will inform the development of the tools and methodologies to estimate the size of populations at risk of HIV and provide pertinent information to complement the national-level HIV surveillance system.

The objectives of this study are to generate estimates of the size of four key populations at risk of HIV—male clients of female sex workers (MCFSW), MSM, female sex workers (FSW) and intravenous drug users (IVDU)—and develop methods to incorporate parameters that may account for both transmission error and barrier effects that are inherent in the NSUM.

METHODS

Overview

We developed and fielded the NSUM survey[4] and developed a Bayesian hierarchical NSUM model to determine the national-level size estimates of key populations at risk of HIV in Singapore. The approach is outlined below and detailed in the online supplementary material.

Questionnaire design

First, we conducted formative research between March and May 2017 to adapt the NSUM to the Singapore context. Four focus group discussions using homogeneous participants and nine indepth interviews were conducted with key stakeholders working in the field of HIV/AIDS—clinicians, academics, programme officers, social workers, counsellors and policymakers—to guide the development of the size estimation component of the HIV surveillance system in Singapore. Participants were purposively sampled by the research team to maximise efficiency and data validity. Information collected from the formative assessment was used to develop the survey questionnaire, as detailed in online supplementary table 1 and online supplementary table 2. The questionnaire was pilot-tested prior to the main survey.

Data source

The Singapore Population Health Studies (SPHS) is a consortium of several population health cohorts in Singapore, namely the Multi-Ethnic Cohort, Diabetic Cohort and Singapore Health Study. The sampling method for the cohort studies has been previously described.[9][10] We recruited SPHS participants between 21 and 70 years old. Those who were illiterate or did not fulfil the age criteria were excluded.

Data collection

Out of the 269 individuals approached between July and August 2017, 17 did not fulfil the age criteria, while 53 were either illiterate or declined to participate. In total, 199 individuals were recruited. The SPHS and research staff explained the study and answered any questions that the participant had before the questionnaire was completed onsite. The self-administered questionnaire was made available in the three main languages in Singapore—Chinese, English and Malay. The questionnaire was anonymous and verbal consent was obtained from all participants. Participants were reimbursed with S$10 (~US$7.30) on completion.

Measures

The questionnaire sought participants' sociodemographic information, their opinions about certain behaviours, the social standing of different groups of people and their opinions on penal code section 377A.[11] Participants were requested to quantify the number of people whom they knew from a list of 24 specified populations (the 19 known populations are listed in online supplementary table 4) and how have they been in contact with one person from each category within the last year (the one person being the most recently contacted).

In this study, we defined knowing a person as follows: A knows B if (1) A knows B by name and sight, and *vice versa*; (2) B is currently living in Singapore; and (3) A had contact with B at least once in the last 12 months.

Statistical analyses

We analysed participants' perceptions of behaviours and social standing of subpopulations in the society associated with the hidden populations. Poisson regression on the number of people in the four at-risk populations was performed to determine factors—participants' demographics and social acceptability rating of selected behaviours—that were associated with knowing more people in the specified populations. Subsequently, we developed a Bayesian hierarchical network scale-up model to estimate the size of four at-risk populations of HIV/AIDS, together with the average personal network size.

Bayesian NSUM model

This model, an adaptation of the NSUM,[4][12] uses Bayesian hierarchical modelling to estimate the number of individuals in hidden populations. This method provides a flexible framework to estimate both individual-level and population-level parameters and is a natural way to handle data with a complex structure such as repeated measurements.[13]

In this model, as depicted in online supplementary figure 2, $N_{i,j}^K$ and $N_{i,l}^H$, the number of contacts reported by individual i with someone in known subpopulation j and hidden subpopulation l, are assumed to be Poisson with means $\lambda\alpha_i S_j^K$ and $\lambda\alpha_i S_l^H$, respectively. The parameter λ is the scaling parameter mapping from the total subpopulation S_j^K or S_l^H to a typical individual's number of contacts; it is a key estimand to derive the hidden population

sizes. A mean 1 random effect α_i for each participant was used to characterise variability in network size. Non-informative prior distributions were assigned to the subpopulation sizes and the scaling parameter, and a non-informative hyperprior for the precision of the random effect. The fitted model was also used to estimate the average social network size, λS, where S is the size of the population living in Singapore.

Because the basic model does not account for two structural features that may bias estimates of the four key populations at risk of HIV/AIDS, two variants of the basic model were considered—*transmission error model* and *transmission error and barrier effect model*—both described below. The former accounts for transmission error by incorporating respondents' perception of that population. The latter builds on the former by incorporating demographics, which may be associated with the chance of knowing people in the hidden populations.

Transmission error model

We attempted to adjust for transmission error—the possibility that members of the hidden population might not divulge membership to some of their contacts—by introducing a correction factor to estimate the size of the at-risk populations. In this model, the mean of the Poisson variate of the number of at-risk populations is modified to account for the individual's perception of that population measured through the variable $x_{i,j}$:

$$N_{i,l}^H \sim Po\left(\lambda\alpha_i \exp\left\{\beta_l\left[x_{i,l} - U_l\right]\right\} S_l^H\right)$$

where U_l is the upper bound of the Likert scale for the question to which $x_{i,l}$ corresponds, and β (if >0) lowers the mean number of people belonging to a hidden population who are known to the individual and whose membership of the hidden population is known to the individual, if that individual expresses an unfavourable attitude towards that population. For instance, someone who has an unfavourable attitude towards MSM may know fewer people who have confided their sexuality to that person, even if he or she knows them. This parameter is also assigned a non-informative prior.

Transmission error and barrier effect model

Building on the *transmission error model*, we accounted for differential types of contacts in different parts of the population (barrier effect), by modifying the mean number of people known in each subpopulation, as follows:

$$N_{i,l}^H \sim Po\left(\begin{array}{c} \lambda\alpha_i \exp\left\{\beta_l\left[x_{i,l} - U_l\right]\right\} \exp\left\{\gamma_{age,l}\left[z_i^{age} - \bar{z}^{age}\right] \right. \\ + \gamma_{sex,l}\left[z_i^{sex} - \bar{z}^{sex}\right] + \gamma_{Malay,l}\left[z_i^{Malay} - \bar{z}^{Malay}\right] \\ \left. + \gamma_{Indian,l}\left[z_i^{Indian} - \bar{z}^{Indian}\right]\right\} S_l^H \end{array}\right)$$

and analogously for the known populations, where z_i^{age} is the age of individual i, z_i^{sex}, z_i^{Malay} and z_i^{Indian} are indicator variables with values of 1 if individual i is male, Malay or Indian, respectively. Here, $\bar{z}^{age}$ is the median age and $\bar{z}^{sex}$, $\bar{z}^{Malay}$ and $\bar{z}^{Indian}$ are the mean values of these binary variables across the Singapore resident population. The parameters γ_{age}, γ_{sex}, γ_{Malay} and γ_{Indian} account for heterogeneity in social structure and were given non-informative priors.

To identify the known populations that provided the most stable comparators for the hidden populations, we undertook leave-one-out validation. For each known population, in turn, we assumed the size was unknown and attempted to re-estimate it, using the other populations. The subpopulations that were most consistently back-estimated were identified by ranking the discrepancy (log ratio) between the scaled-up and the actual

Table 1 Demographics of study participants

Demographic variables		
Age, in years	Mean (SD)*	Median (range)
	49.7 (12.1)	52 (23–70)
	n	%
Sex†		
Female	104	53.3
Male	91	46.7
Education attained		
Primary	22	11.1
Secondary	80	40.2
Pretertiary‡	57	28.6
University	40	20.1
Ethnicity		
Chinese	92	46.2
Malay	62	31.2
Indian	40	20.1
Others	5	2.5
Housing§		
1–3 rooms	36	18.1
4+ rooms	146	73.4
Private	17	8.5

*SD denotes the sample standard deviation.
†Four participants did not indicate their sex.
‡Pretertiary education represents education attained at junior/preuniversity/vocational colleges and polytechnics.
§Housing was used as a proxy for socioeconomic status. The majority of Singaporeans live in public housing (1–3 rooms and 4+ rooms). Private housing refers to condominiums and landed properties.

population sizes, and those with larger discrepancies were removed from subsequent analysis.

The models were fit using a Markov chain Monte Carlo algorithm with 50 000 iterations with a burn-in of 5000, storing 1 out of 10 iterations. Convergence was assessed visually with trace plots. The data analyses were performed in R,[14] and the model building was done in Just Another Gibbs Sampler (JAGS).[15 16] Deviance information criterion (DIC)[17] was used to compare the models. JAGS and R codes are presented in online supplementary materials appendix 1. Data are available as online supplementary files.

We validated the Bayesian method against the classical estimation procedure,[4 18 19] for the simpler model without barrier effects or transmission error for which a classical estimator is available. We also performed sensitivity analyses to assess the robustness of the population size estimates to the transmission error adjustment parameter, β, and ran bootstrap simulations to reweight the survey participants to the general population (presented in the online supplementary materials).

RESULTS

In total, 199 adults aged 23–70 years old completed the questionnaires (table 1). There were more women and minority ethnic groups in the sample; the latter is attributed to deliberate oversampling in the SPHS cohorts to improve the precision of estimates within these groups.

Demographics were significantly associated with the number of contacts in the four hidden populations (table 2 and online supplementary figure 3). Younger participants knew more MCFSW, MSM and IVDU than did older participants, while Malays knew fewer MCFSW but more IVDU. None of the female participants reported knowing any FSW.

The basic model was compared against the classical estimation procedure and similar results were obtained (online supplementary figure 4), demonstrating that our method provides a close analogue to the classical approximation in the simpler case where both are applicable, despite the different formulation.

The ten most reliably estimated known populations to be used to estimate hidden population sizes are presented in online supplementary figure 5. The mean number of contacts of each of the selected known and hidden populations in an individual's network is presented in figure 1 and online supplementary figure 6: typical participants reported knowing few FSW (0.1, 95%CI 0.0–0.2) and IVDU (0.2, 95%CI 0.1–0.3), but an average of one MSM (0.8, 95%CI 0.5–1.3) and one MCFSW (0.7, 95%CI 0.4–1.1).

Perceptions of how socially acceptable behaviours associated or not with the hidden populations are illustrated in online supplementary figure 7. Injecting drugs was about as socially acceptable as drink driving, while sexual behaviours identified with the hidden populations were comparable with a woman smoking or children putting their elderly parents in a nursing home.

Table 2 Results from Poisson regression

	Key populations at risk of HIV/AIDS							
	MCFSW		FSW		MSM		IVDU	
	Ratio of mean (95% CI)*	P value	Ratio of mean (95% CI)*	P value	Ratio of mean (95% CI)*	P value	Ratio of mean (95% CI)*	P value
Age, in decades	0.84 (0.74 to 0.95)	0.006	1.13 (0.84 to 1.51)	0.431	0.58 (0.52 to 0.65)	<0.001	0.68 (0.56 to 0.83)	<0.001
Sex†								
Female	Ref				Ref		Ref	
Male	5.46 (3.58 to 8.33)	<0.001	–		0.72 (0.54 to 0.96)	0.025	1.26 (0.76 to 2.08)	0.375
Ethnicity								
Chinese	Ref		Ref		Ref		Ref	
Indian	1.43 (1 to 2.06)	0.052	1.79 (0.8 to 3.99)	0.154	1.17 (0.85 to 1.63)	0.333	2.23 (1.04 to 4.81)	0.04
Malay	0.19 (0.1 to 0.36)	<0.001	0.73 (0.29 to 1.81)	0.492	0.42 (0.28 to 0.61)	<0.001	3.09 (1.57 to 6.08)	0.001
Other‡	2.01 (1.14 to 3.55)	0.016	–		1.02 (0.4 to 2.56)	0.974	–	

*CI refers to confidence interval of the ratio of the means.
†No women reported knowing any FSWs. Hence, we could not estimate the sex effect of knowing more FSW.
‡No one who was from other ethnic groups knew any FSW or IVDU.
FSW, female sex worker; IVDU, intravenous drug users; MCFSW, male clients of female sex workers; MSM, men who have sex with men; Ref, reference group.
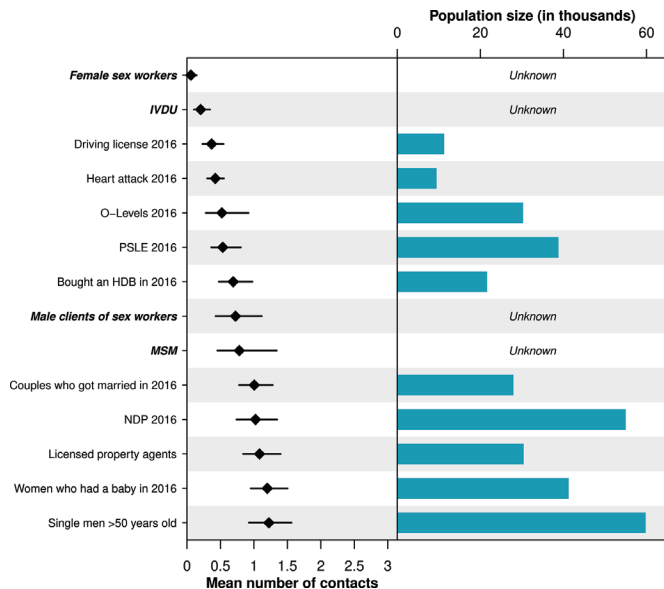
Figure 1 Mean number of contacts of selected populations in an individual's network and subpopulation size. The bootstrapped mean number of contact is represented by the point, and its 95% CI of the bootstrapped mean is indicated by the line. *O-Levels 2016* refers to students who sat for the General Certificate of Education Ordinary Level examinations in 2016, typically at the end of secondary school education. *PSLE 2016* refers to students who sat for the Primary School Leaving Examination in 2016, typically at the end of primary school education. *NDP 2016* refers to individuals who attended the Singapore National Day Parade in 2016. *Bought an HDB in 2016* refers to all individuals who bought a flat by the Housing and Development Board in Singapore in 2016. *Heart attack 2016* refers to individuals who had suffered a heart attack in 2016. IVDU, intravenous drug user; MSM, men who have sex with men.

We estimated the size of the four hidden populations using the basic and extended models (figure 2). The adjusted size estimate of the population of MCFSW was 72 000 (95% CI 51 000 to 100 000), of FSW 4200 (95% CI 1600 to 10 000), of MSM 210 000 (95% CI 140 000 to 300 000) and of IVDU

11 000 (95% CI 6500 to 17 000). There was strong support in favour of the model which considered both transmission error and barrier effect compared with the basic NSUM model (ΔDIC=1617) or the model that accounted only for transmission error (ΔDIC=1592) (details in online supplementary table 5), although the transmission error model was also a substantial improvement over the basic NSUM model (ΔDIC=25). The sizes of the MCFSW, MSM and IVDU populations increased on incorporating the correction factors.

DISCUSSION

This study provides the first estimates of the number of FSWs and their clients in Singapore, and the first empirical estimates of MSM and IVDU.[20][21] A previous estimate of the size of the MSM population was 90 000[22] (our estimate was 210 000), using community organisations' guesstimates. In 2004, the United Nations Reference Group on HIV/AIDS Prevention and Care among IDU estimated the IVDU population in Singapore to be 15 000, with estimation bounds of 10 000–20 000[23] (our estimate was 11 000).

We assessed several fundamental concepts in this survey— definition of contacts, personal and community perceptions of selected behaviours, and populations—which was instrumental in building the final NSUM Bayesian model. To determine the best way to define a contact in the local setting, we quantified the proportion of the means of the communications via text messages, phone call, sharing a meal or a drink and interacting inperson undertaken by study respondents with their recent contacts. While the majority reported contacting their recent contacts in-person, we found some level of agreements between the different means of communications (online supplementary material, Definition of a contact). Therefore, we recommend the utilisation of face-to-face communication, text messaging and phone call in the final working definition of contact. In this survey, we elicited study respondents' self-perception and their perception of the attitudes of others in their community towards selected behaviours and subpopulations in the society. We found the attitudes of respondents and their perceptions of attitudes of others in the community to be highly correlated (online supplementary figure 8). Furthermore, we observed only
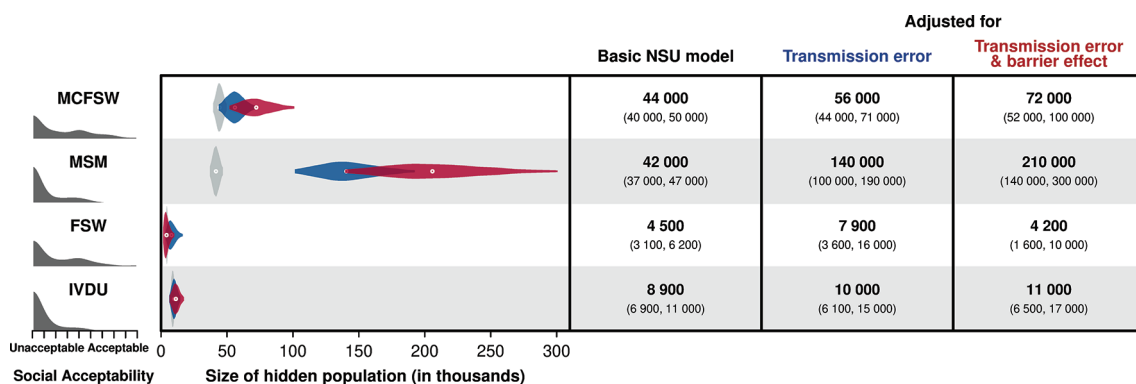

Figure 2 Size of the four key populations at risk: (1) male clients of female sex workers (MCFSW), (2) men who have sex with men (MSM), (3) female sex worker (FSW) and (4) injection drug users (IVDU) estimated by the basic NSUM model (in grey), the NSUM model adjusting for transmission error (in blue), and the NSUM model adjusting for both transmission error and barrier effect (in red). The estimated size of the populations at risk is presented with its 95% credible interval. The interpretation of violin plots is similar to box plots; they display the probability density of the prevalence estimates at different values. Points are posterior median prevalence estimates, and curves are posterior distributions of the parameters truncated to within 95% CIs (all tabulated in the table on the right). Distribution of individual perceptions of how socially acceptable the four at-risk populations are is presented on the left. This correction factor was introduced into the adjusted models to account for transmission error. NSUM, network scale-up method.

weak correlations between the community's perceptions and parameters involving at-risk populations. Hence, we considered respondents' self-perception in the final analysis, and we recommend future NSUM studies to focus on that.

The act of gay sex between men is illegal in Singapore.[24] While the statute (section 377A) is not actively enforced, its retention amplifies discrimination against MSM.[25] Therefore, membership to this group may not always be known to their social contacts, resulting in *transmission error*.[26] Other high-risk behaviours—using illicit drugs, selling and buying sex—may be stigmatised or illegal, making research on these hidden populations challenging. We identified that social acceptability of same-sex behaviour and men who pay for sex could potentially explain the transmission errors in the MSM and MCFSW populations, and that perceived social standing of FSWs was more suitable for addressing the transmission errors in the corresponding group. No real adjustment was possible for the IVDU populations because the correlation between the attitude towards illicit drug users and the number of contacts they knew was low.

In principle, with a sample that penetrates all parts of the population, the effect of assortative mixing (ie, that like mix with like) is mitigated, but samples that are unable to represent high-risk individuals among survey participants adequately may suffer barrier effects, leading to estimates of the corresponding subpopulation sizes that are biased downwards. Because the prevalence of the high-risk groups may differ by age or gender, barrier effects can be partially mitigated by accounting for heterogeneities in the make-up of social networks, which we accounted for through demographic covariates. This may not fully address barrier effects, however, if the high-risk group is particularly isolated from the general population. We suspect that of the high-risk groups considered, the barrier effect of IVDU may be more prominent due to many members of that community being incarcerated.

The findings of this study should be considered together with several limitations. The sample did not perfectly resemble the Singapore population. In particular, younger participants were more likely to be liberal and to have favourable attitudes towards the four hidden populations, so their under-representation in this sample may bias down our estimates. We address this via a model that accounts for demographics. The self-reported number of contacts in this study might be inherently affected by recall bias.[27] It has been shown that 20% of the critical details of personal events were irretrievable after one year.[28] We sought to address this in the study design by shortening the time window that defines a contact from two years, which is the norm in the literature, to one. As a result, we anticipated that study respondents would have less difficulty enumerating their contacts, and the mean number of contacts per group was correspondingly low (<3 contacts) (online supplementary figure 6), which as Feehan and colleagues[29] argue ought to result in more precise size estimates. We assumed that different numbers of contacts reported in the high-risk groups are due solely to differences in disclosure, and not due to systematic differences in the actual number of the high-risk groups known. This assumption cannot be supported empirically under the current study design. Others have argued[30] convincingly that additional data on disclosure rates from the high-risk groups themselves may overcome this limitation, but in Singapore, as in some other settings, obtaining a sample of marginalised or criminalised groups may be intractable.

CONCLUSION

In this study, we attempted to adjust for both barrier effects and transmission error using sociodemographic and perceptions of the marginalised groups to estimate the size of these hidden populations at risk of HIV, extending previous work that did not adequately accommodate such information. The national size estimates of at-risk populations generated will help determine the magnitude and resources required for national HIV prevention and intervention efforts. This approach could be considered by countries seeking to augment their national HIV surveillance systems or to estimate other hard-to-reach populations, such as victims of slave trade and domestic violence.

Key messages

► The network scale-up method (NSUM) is relatively simple, feasible and inexpensive, and it is capable of estimating multiple hard-to-reach populations in a single survey, even in relatively conservative societies.

► After adjusting for transmission error and barrier effect, the Bayesian NSUM estimated 72 000 male clients of female sex workers, 4200 female sex workers, 210 000 men who have sex with men and 11 000 intravenous drug users in Singapore.

► The NSUM could be considered by countries seeking to estimate the size of key populations to enhance HIV surveillance system, prevention and intervention efforts.

Present affiliations The present affiliation of Hanh Hao La is: Center for High Impact Philanthropy, Center for Public Health Initiatives, University of Pennsylvania, Pennsylvania, Philadelphia, USA.

ORCID iD
Alvin Kuo Jing Teo http://orcid.org/0000-0002-0569-3518

REFERENCES
1 UNAIDS. *Global AIDS response progress reporting (GARPR) 2014—country progress report Singapore*. Geneva: UNAIDS, 2014.

2   Action for AIDS Singapore. *PreP consultation with community and health Service providers*. Singapore: Action for AIDS, 2016.

3   Ministry of Health. Update on HIV/AIDS situation in Singapore 2016. Singapore: : Ministry of Health, 2017. Available: https://www.moh.gov.sg/content/moh_web/home/statistics/infectiousDiseasesStatistics/HIV_Stats/update-on-the-hiv-aids-situation-in-singapore-2016--june-2017-.html[Accessed 28 Aug 2017].

4   Bernard HR, Hallett T, Iovita A, *et al*. Counting hard-to-count populations: the network scale-up method for public health. *Sexually Transmitted Infections* 2010;86(Suppl 2):ii11–15.

5   Ezoe S, Morooka T, Noda T, *et al*. Population size estimation of men who have sex with men through the network scale-up method in Japan. *PLoS ONE* 2012;7.

6   Nikfarjam A, Shokoohi M, Shahesmaeili A, *et al*. National population size estimation of illicit drug users through the network scale-up method in 2013 in Iran. *Int J Drug Policy* 2016;31:147–52.

7   Sulaberidze L, Mirzazadeh A, Chikovani I, *et al*. Population size estimation of men who have sex with men in Tbilisi, Georgia; multiple methods and triangulation of findings. *PLoS ONE* 2016;11.

8   Quaye S, Fisher RH, Atuahene K, *et al*. Critique and lessons learned from using multiple methods to estimate population size of men who have sex with men in Ghana. *AIDS Behav* 2015;19:16–23.

9   Chen MIC, Lee VJM, Lim W-Y, *et al*. Influenza A(H1N1) seroconversion rates and risk factors among distinct adult cohorts in Singapore. *JAMA* 2009;2010:1383–91.

10 Tan NC, Barbier S, Lim WY, *et al*. 5-year longitudinal study of determinants of glycemic control for multi-ethnic Asian patients with type 2 diabetes mellitus managed in primary care. *Diabetes Res Clin Pract* 2015;110:218–23.

11 The statutes of the Republic of Singapore. Penal code: chapter 224. *Original Enactment: Ordinance 4 of 1871)* 2017.

12 Bernard HR, Johnsen EC, Killworth PD, *et al*. Estimating the size of an average personal network and of an event subpopulation: some empirical results. *Social Science Research* 1991;20:109–21.

13 Gelman A, Carlin J, Stern H, *et al*. *Bayesian data analysis*. 3rd edition. United States: Taylor Francis Ltd, 2013.

14 R Core Team. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2017.

15 Plummer M, Stukalov A, Denwood M. Bayesian graphical models using MCMC. R package version 4-6., 2016. Available: https://CRAN.R-project.org/package=rjags [Accessed 10 Nov 2017].

16 Plummer M. Just another Gibbs sampler (JAGS), 2017. Available: http://mcmc-jags.sourceforge.net/ [Accessed 10 Nov 2017].

17 Spiegelhalter DJ, Best NG, Carlin BP, *et al*. Bayesian measures of model complexity and fit. *J Royal Statistical Soc B* 2002;64:583–639.

18 McCarty C, Killworth PD, Bernard HR, *et al*. Comparing two methods for estimating network size. *Human Organization* 2001;60:28–39.

19 Habecker P, Dombrowski K, Khan B. Improving the network scale-up estimator: incorporating means of sums, recursive back estimation, and sampling weights. *PLoS ONE* 2015;10:e0143406.

20 Mathers BM, Degenhardt L, Phillips B, *et al*. Global epidemiology of injecting drug use and HIV among people who inject drugs: a systematic review. *The Lancet* 2008;372:1733–45.

21 Degenhardt L, Peacock A, Colledge S, *et al*. Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: a multistage systematic review. *The Lancet Global Health* 2017;5:e1192–207.

22 International Planned Parenthood Federation (IPPF). *Sex Between Men In Your City*. London: IPPF, 2011.

23 Aceijas C, Stimson GV, Hickman M, *et al*. Global overview of injecting drug use and HIV infection among injecting drug users. *AIDS* 2004;18:2295–303.

24 Chen J. Singapore's culture war over section 377A: through the lens of public choice and Multilingual research. *Law Soc. Inq.* 2013;38:106–37.

25 Chua LJ, Su D, Tan RKJ, *et al*. Decriminalisation of same-sex relations and social attitudes in Singapore: an empirical study. *Hong Kong Law Journal* 2017;47:793–824.

26 Salganik MJ, Mello MB, Abdo AH, *et al*. The game of contacts: estimating the social visibility of groups. *Social Networks* 2011;33:70–8.

27 Killworth PD, McCarty C, Bernard HR, *et al*. Two interpretations of reports of knowledge of subpopulation sizes. *Social Networks* 2003;25:141–60.

28 Bradburn NM, Rips LJ, Shevell SK. Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 1987;236:157–61.

29 Feehan DM, Umubyeyi A, Mahy M, *et al*. Quantity versus quality: a survey experiment to improve the network scale-up method. *Am J Epidemiol* 2016;183:747–57.

30 Feehan DM, Salganik MJ. Generalizing the network scale-up method: a new estimator for the size of hidden populations. *Sociol Methodol* 2016;46:153–86.