American Journal of Epidemiology

Vol. 174, No. 10
DOI: 10.1093/aje/kwr246
Advance Access publication:
October 14, 2011

Practice of Epidemiology

Assessing Network Scale-up Estimates for Groups Most at Risk of HIV/AIDS: Evidence From a Multiple-Method Study of Heavy Drug Users in Curitiba, Brazil

Matthew J. Salganik*, Dimitri Fazito, Neilane Bertoni, Alexandre H. Abdo, Maeve B. Mello, and Francisco I. Bastos

* Correspondence to Dr. Matthew J. Salganik, Department of Sociology and Office of Population Research, Princeton University, 145 Wallace Hall, Princeton, NJ 08544 (e-mail: mjs3@princeton.edu).

Initially submitted April 23, 2011; accepted for publication June 22, 2011.

One of the many challenges hindering the global response to the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic is the difficulty of collecting reliable information about the populations most at risk for the disease. Thus, the authors empirically assessed a promising new method for estimating the sizes of most at-risk populations: the network scale-up method. Using 4 different data sources, 2 of which were from other researchers, the authors produced 5 estimates of the number of heavy drug users in Curitiba, Brazil. The authors found that the network scale-up and generalized network scale-up estimators produced estimates 5–10 times higher than estimates made using standard methods (the multiplier method and the direct estimation method using data from 2004 and 2010). Given that equally plausible methods produced such a wide range of results, the authors recommend that additional studies be undertaken to compare estimates based on the scale-up method with those made using other methods. If scale-up-based methods routinely produce higher estimates, this would suggest that scale-up-based methods are inappropriate for populations most at risk of HIV/AIDS or that standard methods may tend to underestimate the sizes of these populations.

acquired immunodeficiency syndrome; epidemiologic methods; HIV; network sampling; population size estimation; social networks

Abbreviations: AIDS, acquired immunodeficiency syndrome; CAPS, Centro de Atenção Psicossocial; CI, confidence interval; HIV, human immunodeficiency virus; PCAP, Pesquisa de Conhecimento, Atitudes e Práticas Relacionadas ao HIV/AIDS na População Brasileira de 15 a 54 Anos de Idade.

One of the challenges hindering the global response to the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic is the difficulty of collecting reliable information about the populations most at risk. This information is difficult to collect because in many countries, HIV/AIDS risk is concentrated in populations—illicit drug users, female sex workers, and men who have sex with men—that are difficult to sample using standard statistical methods. The resulting lack of accurate, timely, and comprehensive information makes evidence-based approaches to targeting prevention programs and monitoring effectiveness difficult. Consider one of the most basic questions one might ask: How large are the most at-risk populations around the world, and how are the sizes of these populations changing over time? Despite enormous amounts of work carried out using a variety of methods, much uncertainty remains (1). For example, in many countries where injecting drug use has been reported, no reliable estimate of the number of drug injectors exists (2). Even the estimates that do exist are difficult to interpret because of methodological differences between countries and over time within countries (3). Similar uncertainties exist about the numbers of female sex workers and men who have sex with men (4–7).

One promising approach for estimating the sizes of groups most at risk of HIV infection is the network scale-up method, a technique that is new to epidemiology but has established roots in anthropology and social network analysis (8–10). The method uses information about the personal networks of a random sample of the general population to make size estimates, and it has a number of attractive features for global public health (10): 1) it can easily be standardized across countries and time because it requires a random sample of the general population, perhaps the most widely used sampling design in the world; 2) it can produce estimates of the sizes of many target populations in the same data collection, whereas many alternative methods require distinct data collections for each population of interest; 3) it can be partially self-validating because it can easily be applied to populations of known size; 4) depending on the sampling frame, it can produce estimates at either the city level or the national level, whereas many alternative methods can only be applied on 1 geographic scale; 5) it does not require respondents to report that they are members of a stigmatized group; and 6) it is relatively inexpensive and does not require extensive administrative records, which makes it feasible to use at frequent time intervals, even in middle- and low-income countries.

Despite these appealing characteristics, the applicability of the network scale-up method for global HIV/AIDS research remains unclear. Therefore, we empirically assessed the utility of the network scale-up method and the newer generalized network scale-up method in this context. Ideally, we would assess the accuracy of these scale-up-based estimates, but that is difficult, because for most at-risk populations we lack a ‘‘gold standard’’ size estimate. Therefore, we conducted our study in a most-at-risk population whose size had been estimated previously: heavy drug users in Curitiba, Brazil. Curitiba is an optimal location for this study because Brazil is a middle-income country with a concentrated HIV/AIDS epidemic and a strong governmental response to HIV/AIDS (11), and Curitiba, a city of 1.8 million people in southern Brazil, was the site of a 2004 Brazilian Ministry of Health study that yielded an estimate of the number of heavy drug users in the city. In addition to this previous estimate, we also estimated the number of heavy drug users in Curitiba using 2 standard methods: the multiplier method and the direct estimation method (12). These 3 estimates provided a background that we could use to assess the scale-up and generalized scale-up estimates. Thus, while most studies of hard-to-count populations produce only a single estimate, our study produced 5 different estimates based on 4 distinct data sources, 2 of which were from other researchers.

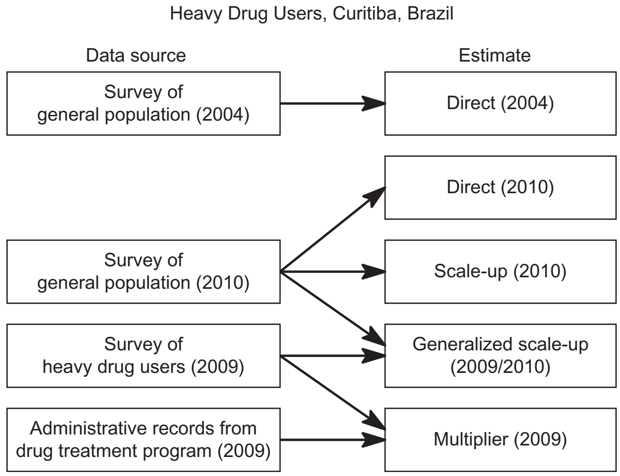


Figure 1. Design of a study for estimating the number of heavy drug users in Curitiba, Brazil. Four distinct data sources, 2 of which were from other researchers, were used to produce 5 estimates.

MATERIALS AND METHODS

The target population in our study was heavy drug users, defined as people who had used illegal drugs other than marijuana more than 25 times in the past 6 months. This target population is appropriate to the current state of the HIV/AIDS epidemic in Brazil, where injecting drugs is unusual and heavy drug users show high rates of HIV infection relative to the general population (11).

Data sources

Our study used 4 data sources to produce 5 estimates, as summarized in Figure 1. One source of data, which were collected by our research team, was a face-to-face survey administered to a household-based random sample of 500 adult (i.e., aged ≥18 years) residents of Curitiba in 2010. The second source of data, also collected by our research team, was a respondent-driven sample (13–17) of 303 heavy drug users in Curitiba selected in 2009 (18, 19). The third source of data, collected by an independent group of researchers, was the 2004 Brazilian Ministry of Health PCAP survey (Pesquisa de Conhecimento, Atitudes e Práticas Relacionadas ao HIV/AIDS na População Brasileira), which measured the knowledge, attitudes, and practices of the Brazilian population with respect to HIV/AIDS (20). The final data source we used was administrative records from the Centro de Atenção Psicossocial (CAPS) drug treatment program in Curitiba. For more on the definitional consistency across these data sources, see the Web Appendix (http://aje.oxfordjournals.org/).

Network scale-up method and generalized network scale-up method

The network scale-up method estimates population sizes using information about the personal networks of survey respondents under the assumption that personal networks are, on average, representative of the general population. For example, if a respondent reports knowing 2 illicit drug users and knows 200 people overall, we can estimate that 2/200, or 1%, of the population are illicit drug users. This estimate can be improved by averaging data over many respondents (9). The data needed for the network scale-up method come from interviews with a random sample of the general population. In addition to basic demographic questions, respondents are asked how many people they ‘‘know’’ in the target population. Following standard practice (10), in our study ‘‘know’’ was defined as follows: ‘‘You know them and they know you, and you have been in contact with them in the last 2 years.’’ Respondents are then asked a battery of questions to estimate the number of people they know (i.e., the size of their personal network).

From these survey data, one can estimate the size of the target population as

$$\hat{p} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} \hat{d}_i}, \qquad (1)$$

where y_i is the number of people known in the target population and $\hat{d}_i$ is the estimated personal network size (9). Thus, one can view the network scale-up estimator as a generalization of the familiar sample proportion, which is the number of sample members in the target population divided by the sample size. The network scale-up estimator is instead the total number of target population members known by the respondents divided by the total number of people known by respondents.

In this context, the 2 methods most appropriate for estimating the total number of people known by each respondent are the known population method and the summation method (10). Because it was not clear a priori which method would produce more accurate estimates in this context, we used both methods in our study. Tests (described in the Web Appendix) showed that in this study, the data from the known population method were preferable, and therefore those data will be presented throughout.

The network scale-up method makes some strong implicit assumptions, and for that reason, we also collected the data needed for the generalized scale-up estimator. These data come from a sample of the target population—in this case, heavy drug users—and are then combined with the data from the general population to produce 2 correction factors: one for the lack of information flow and one for the differential network size between the target population and the general population. These correction factors and the procedures needed to estimate them are described in detail in the Web Appendix.

Direct estimation and multiplier method

For comparison with the scale-up and generalized scale-up estimates, we estimated the number of heavy drug users using 2 common methods: direct estimation and the multiplier method (12). Direct estimation involves asking a sample of the general population whether they are heavy drug users. The multiplier method estimates the size of the target population based on 2 pieces of information: 1) the number of people in the target population with some specific characteristic (e.g., the number of people in a specific drug treatment program) and 2) the estimated prevalence of that characteristic in the target population. This information is combined as follows:

$$\hat{N}_T = \frac{N_c}{\hat{p}_c}, \qquad (2)$$

where N_c is the number of people in the target population with that characteristic and $\hat{p}_c$ is the estimated proportion of the target population with that characteristic. In our study, N_c was the number of heavy drug users in a specific treatment program and $\hat{p}_c$ was the estimated proportion of heavy drug users who were in that program.

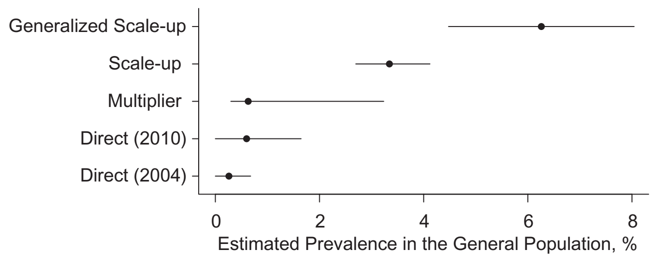


Figure 2. Five estimates of the prevalence of heavy drug use in Curitiba, Brazil, 2004 and 2009–2010. Scale-up and generalized scale-up estimates were substantially higher than those obtained from standard methods (direct estimation and the multiplier method). Estimates of the number of heavy drug users in Curitiba ranged from 4,700 to 114,000. Bars, estimated 95% confidence interval.

RESULTS

The results from all 5 estimates are presented in Figure 2 and described in detail below.

Commonly used methods: direct estimation and multiplier method

We had 2 different sources of data for direct estimates. First, in 2004, the Brazilian Ministry of Health conducted the PCAP survey, which included approximately 1,000 people in Curitiba, and asked directly about the use of powder cocaine and injected cocaine. From these data, we estimated a prevalence of heavy drug use within the Curitiba population of 0.3% (95% confidence interval (CI): 0, 0.7) (see Web Appendix). In our 2010 survey of the general population, we produced a direct estimate of the prevalence of heavy drug use in the general population of 0.6% (95% CI: 0, 1.6) (see Web Appendix).

To produce our multiplier estimate, we learned from administrative records that 423 heavy drug users were enrolled in the CAPS drug treatment program in August 2009. We also estimated from our sample of heavy drug users (data collected in 2009) that 3.7% of heavy drug users were in the CAPS program (95% CI: 0.7, 7.8). Therefore, our multiplier estimate for the number of heavy drug users was 423/0.037 (see equation 2) or 11,459 people, which corresponds to 0.6% of the population (95% CI: 0.3, 3.2) (see Web Appendix). Thus, we see that these 2 commonly used methods produced similar estimates.

While this is somewhat reassuring, there are reasons to suspect that direct estimation and the multiplier method both produce underestimates. Direct estimates of the prevalence of drug use can be plagued by nonsampling error (21) and are suspected to be underestimates for 2 reasons (22). First, several studies that compared self-reported drug-use data with drug-testing data found that respondents underreport their drug use, in some cases substantially (23–25). Second, heavy drug users appear to be more difficult to reach in standard household surveys, which creates differential nonresponse (26, 27). For these reasons, some researchers place more confidence in ‘‘indirect’’ estimation methods, of which the multiplier method is an example (2). However, the multiplier-based estimates are only as good as the data used to create them. Multiplier methods will tend to produce underestimates if the members of the target population that appear in administrative data are overrepresented in the sample of the target population—akin to problems with capture-recapture when capture probabilities are correlated (28). For example, we suspect that participants in the CAPS treatment programs were overrepresented in our sample of heavy drug users, because middle- and upper-class heavy drug users were less likely to participate in CAPS (because it is a free government program) and less likely to participate in our respondent-driven sampling study (because the financial incentives for participation were less attractive for middle- and upper-class drug users). If this pattern did occur, these middle- and upper-class heavy drug users would be essentially invisible to the multiplier method.

Given these sources of concern about the commonly used methods, we now turn to another indirect method, the network scale-up method. By using information embedded in respondents' personal networks, this method allows researchers to make indirect estimates that include people whose drug use is not recorded in any administrative records.

Network scale-up method and generalized network scale-up method

Respondents in our general population survey reported knowing a total of 3,075 heavy drug users in Curitiba. Further, we estimated that our respondents knew a total of 92,003 people in Curitiba. Therefore, the scale-up estimator produced an estimated proportion of heavy drug users of 3.3% (95% CI: 2.7, 4.1) (see Web Appendix).

The generalized network scale-up estimator relaxes 2 assumptions of the network scale-up estimator. It relaxes both the assumption that people are aware of everything about the people they are connected to and the assumption that the target population has the same average personal network size as the population as a whole. As we describe in more detail in the Web Appendix, we estimated the 2 necessary correction factors using data from our sample of heavy drug users and produced a generalized scale-up estimate of the proportion of heavy drug users of 6.3% (95% CI: 4.5, 8.0).

DISCUSSION

The estimates derived from the network scale-up method and the generalized network scale-up method were substantially higher than those from standard methods (Figure 2). However, our scale-up-based estimates of drug use are roughly comparable to those of previous national-level studies in Brazil and international benchmarks (see Web Appendix). Further, because the scale-up method allows researchers to estimate the sizes of multiple target populations, our study also estimated the number of female sex workers and men who have sex with men in Curitiba. We find that these estimates too are roughly comparable with those of other studies from Brazil and international meta-analysis (see Web Appendix). We caution, however, that all of these comparisons have a large degree of uncertainty because of differences between the studies and ambiguities in the definitions of the target populations.

Although these consistency checks are somewhat encouraging, they cannot assess the accuracy of the estimates. Therefore, as a final check, we note that we also asked respondents how many people they knew in 20 populations of known size—for example, women who have given birth in the last 12 months, students enrolled in public universities, and employees of the city of Curitiba (see the Web Appendix for a list of the 20 populations). Therefore, to assess the network scale-up method, we estimated the size of each of these populations using our sample and the scale-up estimator (equation 1). Figure 3 reveals that for most of the 20 populations, the size estimates, while not perfect, were quite reasonable. However, Figure 3 also reveals a tendency to overestimate the sizes of smaller populations and underestimate the sizes of larger populations, a finding that is consistent with previous studies (29, 30). The fact that this exact estimator in this exact sample can produce reasonable estimates for quantities we can check gives us some additional confidence about the estimates for quantities we cannot check.

Estimates made in populations of known size also reveal that the estimated 95% confidence intervals, which in this case were generated using a bootstrap procedure accounting for our 2-stage cluster sample design (see Web Appendix), are not wide enough: The purported 95% confidence intervals

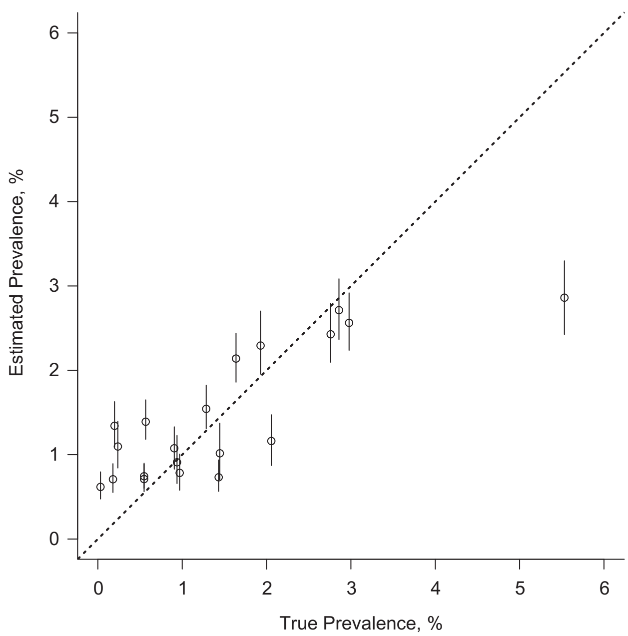


Figure 3. Validation of network scale-up estimates for 20 populations of known size in Curitiba, Brazil, 2010. (A list of the 20 populations is presented in the Web Appendix.) The estimates were generally similar to the true values, but there was a tendency to overestimate the sizes of small groups and underestimate the sizes of large groups, a pattern that has been observed in other scale-up studies as well (29, 30). The purported 95% confidence intervals also had poor coverage properties. The 1 outlier in the plot represents the group "middle school students in a public school." Bars, estimated 95% confidence interval.

have an empirical coverage rate of 25%. While this is somewhat discouraging, it is also exciting that we can actually detect this problem (we suspect that the confidence intervals for many comparable methods are also too small, but this problem is largely invisible). Therefore, we suggest that in future research, investigators also address nonsampling sources of uncertainty, such as those introduced by response bias or recall errors.

Because the scale-up-based estimates were so much higher than those obtained with existing methods, we considered many possible sources of error that might have inflated our estimates (of course, as explained above, there are reasons to suspect that the standard estimates are too low). One possible source of overestimation could be the order of the questions in our survey: Heavy drug users were the first group asked about, and this might have led to inflated responses. However, no effects of question order were found in a previous telephone-based network scale-up survey in Italy (31). Unfortunately, we were unable to randomize the order of our questions for logistical reasons, so we cannot address this possibility directly with our data. We recommend that future researchers randomize the order of the questions if possible.

An alternative explanation for these apparently high estimates is that some interviewers may not have followed the study protocol. More specifically, rather than asking, "How many people do you know who live in Curitiba and have used illegal drugs other than marijuana more than 25 times in the last 6 months (i.e., average of once a week)?," some interviewers could have shortened the question to, "How many people do you know that use drugs?". This shorter question could have produced much higher responses that would have led to a higher estimated population size. Although we had no reason to believe that this occurred, we assessed the robustness of our estimates to data from a single interviewer by systematically dropping the data collected by each of our 9 interviewers. This analysis showed that no particular interviewer had a large effect on the estimate (see Web Appendix).

An additional source of upward bias is "drug use inflation," that is, respondents' including in their answers people who used drugs but did not use them heavily enough to match our study criteria. For example, if a respondent knew someone who drank alcohol every day, the respondent might have included this person in his or her count even though that person did not match our study criteria. In fact, a previous survey in Brazil suggests that this "drug use inflation" might occur: Approximately 20% of respondents reported that someone who drank alcohol twice per week was at severe risk due to substance misuse, and almost half of respondents reported that someone who had used marijuana once or twice in his/her lifetime was at severe risk due to substance misuse (32). The differences between this previous survey and our study prevent any firm conclusions, but these previous results at least suggest that "drug use inflation" might have occurred in our survey; this will have to be a topic for future research.

A further possible source of error in the scale-up estimates is problems with the sampling frame. If residents of Curitiba differ in their propensity to know heavy drug users (30) and if the sampling frame was less likely to include persons with higher propensities to know heavy drug users (possibly the homeless), then our scale-up estimates could be too low. Conversely, if the sampling frame systematically excluded persons who have a lower propensity to know heavy drug users (possibly those living in gated communities), then our scale-up estimates could be too high. The relative magnitude of these problems is difficult to assess empirically, and the sensitivity of the network scale-up method to sampling frame problems is an important question for future research.

Prior to data collection, we expected that the scale-up-based estimates might be higher than those made with other methods, but we did not expect them to be so much higher. As was described above, we suspect that direct estimates and multiplier estimates will be too low in our setting (and possibly in many other settings). However, the generalized scale-up method, which we believe is more statistically appropriate than the scale-up method, produced estimates that were much higher than expected. Since these equally plausible methods produced such different results, we recommend that in additional studies investigators compare scale-up-based estimates with those made using other methods; conducting additional scale-up studies without having results from other methods for comparison will not address this challenge.

Fortunately, our research design (Figure 1) can be easily replicated in other settings. In many cities, there are routine behavioral surveillance surveys of populations most at risk of HIV/AIDS, and there are routine studies involving samples taken from the general population. In this case, by adding a few additional questions to each data collection effort, investigators can replicate our study at virtually no cost. Further, our research design could be enriched with additional sources of data and additional estimation methods. For example, distributing a unique object to members of the target population before sampling from the target population could produce a capture-recapture estimate (33), although some features of the target population sampling—in this case, respondent-driven sampling—may complicate this approach (17, 33–36). Other sources of administrative data, such as HIV registry data, could be used to produce additional multiplier method estimates (37), but the accuracy of these estimates will depend on the availability of administrative data and possible statistical dependencies between data sources. An additional variation in design would be to use alternative sampling methods to reach the target population (e.g., time-location sampling).

If additional studies are undertaken and it is found that scale-up-based methods routinely produce higher estimates than existing methods, scale-up-based methods may not be appropriate for estimating the sizes of populations most at risk of HIV/AIDS. Alternatively, it may be the case that existing methods have been systematically underestimating the sizes of these populations. At this point, we do not have enough evidence to definitively address this important possibility.

Procedures used for prevention, management, and treatment of HIV/AIDS in Brazil are recognized worldwide as a benchmark for other middle-income countries. However, Brazil is an emergent country with a growing population and competing health priorities (38). Human and material resources in Brazil and in other countries should be mobilized in the most equitable way possible, on the basis of sound empirical evidence. Therefore, estimates of the sizes of the groups most at risk of HIV/AIDS should be as accurate as possible. The present study shows that scale-up-based estimators are a promising alternative to commonly used approaches, but more research is needed. General, standardizable, and cost-effective approaches to collecting data about the most-at-risk groups are necessary for the formulation of evidence-based policies designed to curb the spread of HIV/AIDS.

ACKNOWLEDGMENTS

Author affiliations: Department of Sociology and Office of Population Research, Princeton University, Princeton, New Jersey (Matthew J. Salganik); CEDEPLAR, Federal University of Minas Gerais, Belo Horizonte, Brazil (Dimitri Fazito); and FIOCRUZ, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil (Neilane Bertoni, Alexandre H. Abdo, Maeve B. Mello, Francisco I. Bastos).

This research was supported by the Joint United Nations Programme on HIV/AIDS, the Brazilian Ministry of Health, a CAPES/Fulbright fellowship (F. I. B.), the US National Science Foundation (grant CNS-0905086 to M. J. S.), and the US Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant R01HD062366 to M. J. S.).

The authors thank Dr. M. Mahy, Dr. R. Lyerla, Dr. H. R. Bernard, Dr. C. McCarty, Dr. T. Zheng, Dr. T. McCormick, K. Levy, and D. Feehan for helpful comments. They also thank Dr. A. R. P. Pascom, M. L. Vettorazzi, Dr. S. T. Moyses, D. Blitzkow, C. H. Venetikides, and M. Thomaz for assistance. Finally, the authors thank the Curitiba Municipal Health Secretariat/CAPS for sharing data about the drug treatment program.

The opinions expressed here represent the views of the authors and not the funding agencies.

Conflict of interest: none declared.

REFERENCES

1. Brookmeyer R. Measuring the HIV/AIDS epidemic: approaches and challenges. *Epidemiol Rev*. 2010;32(1):26–37.
2. Mathers BM, Degenhardt L, Phillips B, et al. Global epidemiology of injecting drug use and HIV among people who inject drugs: a systematic review. *Lancet*. 2008;372(9651):1733–1745.
3. Mathers B, Cook C, Degenhardt L. Improving the data to strengthen the global response to HIV among people who inject drugs. *Int J Drug Policy*. 2010;21(2):100–102.
4. Baral S, Sifakis F, Cleghorn F, et al. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000–2006: a systematic review. *PLoS Med*. 2007;4(12):e339. (doi:10.1371/journal.pmed.0040339).
5. Cáceres C, Konda K, Pecheny M, et al. Estimating the number of men who have sex with men in low and middle income countries. *Sex Transm Infect*. 2006;82(suppl 3):iii3–iii9.
6. Cáceres CF, Konda K, Segura ER, et al. Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003–2007 estimates. *Sex Transm Infect*. 2008;84(suppl 1):i49–i56.
7. Vandepitte J, Lyerla R, Dallabetta G, et al. Estimates of the number of female sex workers in different regions of the world. *Sex Transm Infect*. 2006;82(suppl 3):iii18–iii25.
8. Bernard HR, Johnsen EC, Killworth PD, et al. Estimating the size of an average personal network and of an event subpopulation: some empirical results. *Soc Sci Res*. 1991;20(2):109–121.
9. Killworth PD, McCarty C, Bernard HR, et al. Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. *Eval Rev*. 1998;22(2):289–308.
10. Bernard HR, Hallett T, Iovita A, et al. Counting hard-to-count populations: the network scale-up method for public health. *Sex Transm Infect*. 2010;86(suppl 2):ii11–ii15.
11. Malta M, Magnanini MM, Mello MB, et al. HIV prevalence among female sex workers, drug users and men who have sex with men in Brazil: a systematic review and meta-analysis. *BMC Public Health*. 2010;10(1):317. (doi:10.1186/1471-2458-10-317).
12. Joint United Nations Programme on HIV/AIDS. *Estimating the Size of Populations at Risk for HIV*. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2003.
13. Heckathorn DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc Probl*. 2002;49(1):11–34.
14. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl*. 1997;44(2):174–199.
15. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol Methodol*. 2004;34(1):193–239.
16. Volz E, Heckathorn DD. Probability-based estimation theory for respondent-driven sampling. *J Off Stat*. 2008;24(1):79–97.
17. Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Stat Med*. 2009;28(17):2202–2229.
18. Bastos FI, Malta M, Albuquerque EM, et al. *Taxas de Infecção de HIV e Sífilis e Inventário de Conhecimento, Atitudes e Práticas de Risco Relacionadas À̀s Infecções Sexualmente Transmissíveis Entre Usuários de Drogas em 10 Municípios Brasileiros*. Rio de Janeiro, Brazil: Oswaldo Cruz Foundation; 2009.
19. Salganik MJ, Mello MB, Abdo AH, et al. The game of contacts: estimating the social visibility of groups. *Soc Networks*. 2011;33(1):70–78.
20. Szwarcwald CL, Barbosa-Júnior A, Pascom AR, et al. Knowledge, practices and behaviours related to HIV transmission among the Brazilian population in the 15–54 years age group, 2004. *AIDS*. 2005;19(suppl 4):S51–S58.
21. Turner CF, Lessler JT, Gfroerer JG, eds. *Survey Measurement of Drug Use: Methodological Studies*. Rockville, MD: National Institute on Drug Abuse; 1992.
22. Manksi CF, Pepper JV, Petrie CV. *Informing America's Policy on Illegal Drugs: What We Don't Know Keeps Hurting Us*. Washington, DC: The National Academies Press; 2001.
23. Fendrich M, Johnson TP, Sudman S, et al. Validity of drug use reporting in a high-risk community sample: a comparison of cocaine and heroin survey reports with hair tests. *Am J Epidemiol*. 1999;149(10):955–962.
24. Colón HM, Robles RR, Sahai H. The validity of drug use responses in a household survey in Puerto Rico: comparison of survey responses of cocaine and heroin use with hair tests. *Int J Epidemiol*. 2001;30(5):1042–1049.

25. Delaney-Black V, Chiodo LM, Hannigan JH, et al. Just say "I don't": lack of concordance between teen report and biological measures of drug use. *Pediatrics*. 2010;126(5):887–893.
26. Caspar RA. Follow-up of nonrespondents in 1990. In: *Survey Measurement of Drug Use: Methodological Studies*. Rockville, MD: National Institute on Drug Abuse; 1992:155–173.
27. Zhao J, Stockwell T, Macdonald S. Non-response bias in alcohol and drug population surveys. *Drug Alcohol Rev*. 2009; 28(6):648–657.
28. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev*. 1995;17(2): 243–264.
29. Killworth PD, McCarty C, Bernard HR, et al. Two interpretations of reports of knowledge of subpopulation sizes. *Soc Networks*. 2003;25(2):141–160.
30. Zheng T, Salganik MJ, Gelman A. How many people do you know in prison?: using overdispersion in count data to estimate social structure in networks. *J Am Stat Assoc*. 2006;101(474): 409–423.
31. Snidero S, Zobec F, Berchialla P, et al. Question order and interviewer effects in CATI scale-up surveys. *Sociol Method Res*. 2009;38(2):287–305.
32. Secretaria Nacional de Políticas sobre Drogas. *Relatório Brasileiro sobre Drogas*. Brasília, Brazil: Secretaria Nacional de Políticas sobre Drogas; 2009.
33. Paz-Bailey G, Jacobson JO, Guardado ME, et al. How many men who have sex with men and female sex workers live in El Salvador? Using respondent-driven sampling and capture-recapture to estimate population sizes. *Sex Transm Infect*. 2011;87(4):279–282.
34. Berchenko Y, Frost SD. Capture-recapture methods and respondent-driven sampling: their potential and limitations. *Sex Transm Infect*. 2011;87(4):267–268.
35. Gile K, Handcock M. Respondent-driven sampling: an assessment of current methodology. *Sociol Methodol*. 2010; 40(1):285–327.
36. Goel S, Salganik MJ. Assessing respondent-driven sampling. *Proc Natl Acad Sci U S A*. 2010;107(15):6743–6747.
37. Heimer R, White E. Estimation of the number of injection drug users in St. Petersburg, Russia. *Drug Alcohol Depend*. 2010; 109(1–3):79–83.
38. Victora CG, Barreto ML, do Carmo Leal M, et al. Health conditions and health-policy innovations in Brazil: the way forward. *Lancet*. 2011;377(9782):2042–2053.

Web Appendix

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

Matthew J. Salganik Dimitri Fazito Neilane Bertoni Alexandre H. Abdo
Maeve B. Mello Francisco I. Bastos

These supporting online materials have 9 sections. Section 1 describes the known population and summation methods of personal network size estimation and how we used the data from our study to decide between them. These personal network size estimates are used in the network scale-up and generalized network scale-up estimates described in Sections 2 and 3. Next, Sections 4 and 5 describe the multiplier and direct estimates, two commonly used alternatives to the network scale-up method. Section 6 considers the possibility of interviewer effects in the network scale-up estimates. Section 7 compares the scale-up based estimates to Brazilian and international benchmarks. Section 8 discusses the definitions used in each of our data sources. Finally, Section 9 provides the exact question text in Portuguese (along with English translations) that were used for the direct estimates.

1 Personal network size estimation

As described previously, the network scale-up and generalized network scale-up methods depend on the estimated personal network size of respondents. The two methods most appropriate for estimating this from a survey are the known population method (Killworth et al., 1998) and the summation method (McCarty et al., 2001), each with its own strengths and weaknesses (Bernard et al., 2010; McCormick et al., 2010). Because neither of these methods had been used in Brazil before, we were unsure which would perform better in this context. Therefore, we included both methods in our general population survey. To summarize our findings, the two estimation methods produced similar results, but the known population method was empirically demonstrated to produce more accurate scale-up based estimates. For this reason, we only presented results from the known population method in the main paper.

To estimate personal network size using the known population method, each respondent is asked the number of people he or she knows in various groups of known size. For example, if a respondent reports knowing one woman living in Curitiba who was married in the last 12 months, one could combine that with the fact that there were 9,960 women in Curitiba that were married in the last 12 months to estimate that the respondent knows about one-ten-thousandth (1/9,960) of all Curitibans. As there are about 1.8 million Curitibans, we would estimate that the respondent has a personal network size of $\frac{1}{9{,}960} \cdot 1.8$ million $\approx$ 180 people. To improve the accuracy of this estimate, we can ask about many groups of known size which leads to the following estimator (Killworth et al., 1998)

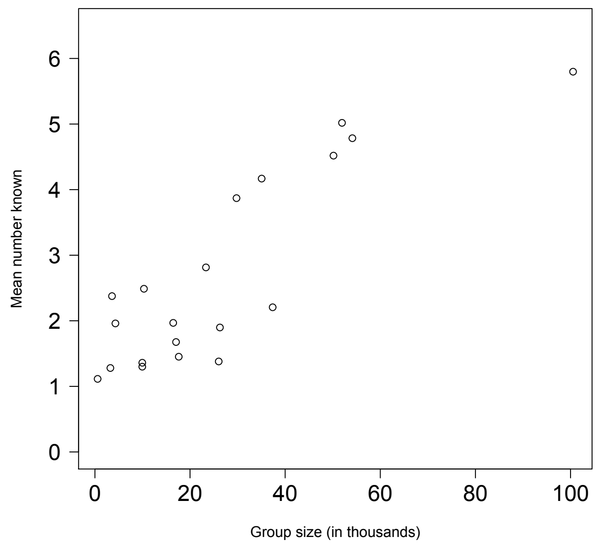
$$\hat{d}_i = \frac{\sum_k y_{ik}}{\sum_k N_k} \times N \tag{1}$$

where $\hat{d}_i$ is the estimated personal network size (i.e., degree) of person i, y_{ik} is the number of people in group k known by person i, N_k is the number of people in group k, and N is number of people in the general population. In our study, all of these quantities were restricted to people in Curitiba ($N = 1,817,434$). Our survey used 20 populations of known size such that $\sum_k N_k = 527,710$, and, therefore, $\frac{\sum_k N_k}{N} \approx 0.3$. The 20 groups and their sizes are presented in Web Table 1.

Group	Size (absolute)	Size (percent)	Source
Middle school student in a public school	100,527	5.5%	MEC[a]
Boy younger than 5	54,129	3.0%	DATASUS[b]
Girl younger than 5	51,948	2.9%	DATASUS[b]
Woman over 70	50,159	2.8%	DATASUS[b]
Employee of City of Curitiba	37,372	2.1%	IBGE/MUNIC[c]
Construction worker	35,056	1.9%	RAIS/CAGED[d]
Man over 70	29,768	1.6%	DATASUS[b]
Student in a public university	26,282	1.4%	MEC[a]
Retired because of disability	26,029	1.4%	MPS[e]
Woman 20 and older who had a baby in the last 12 months	23,344	1.3%	IBGE/SIDRA[f]
High school student in a private school	17,627	1.0%	MEC[a]
Bank teller	17,056	0.9%	RAIS/CAGED[d]
Middle school student in a private school	16,461	0.9%	MEC[a]
Died in the last 12 months	10,310	0.6%	DATASUS[b]
Woman married in the last 12 months	9,960	0.5%	IBGE/MUNIC[c]
Man married in the last 12 months	9,960	0.5%	IBGE/MUNIC[c]
Bus driver	4,309	0.2%	RAIS/CAGED[d]
Woman under 20 who had a baby in the last 12 months	3,593	0.2%	IBGE/SIDRA[f]
Taxi driver	3,252	0.2%	RAIS/CAGED[d]
Hospitalized for a traffic accident in the last 12 months	568	0.03%	DATASUS[b]
Total	527,710	29.0%	

[a] Ministry of Education, Educational Census (2009)
[b] Ministry of Health Public Database (2009)
[c] Brazilian National Institute of Geography and Statistics, Municipalities Demographic Information Database (2009)
[d] Ministry of Labor and Business, Administrative Record Center (2009)
[e] Ministry of Public Insurance (2009)
[f] Brazilian National Institute of Geography and Statistics, Aggregated Database of Demographic Information and Vital Statistics (2009)

Web Table 1: The 20 groups of known size that were used to estimate respondents' personal network sizes.



Web Figure 1: Mean number known in each of the 20 groups of known size compared to the size of that group. In general, respondents report knowing more people in larger groups ($r = 0.86$), which suggests that the responses are reasonable.

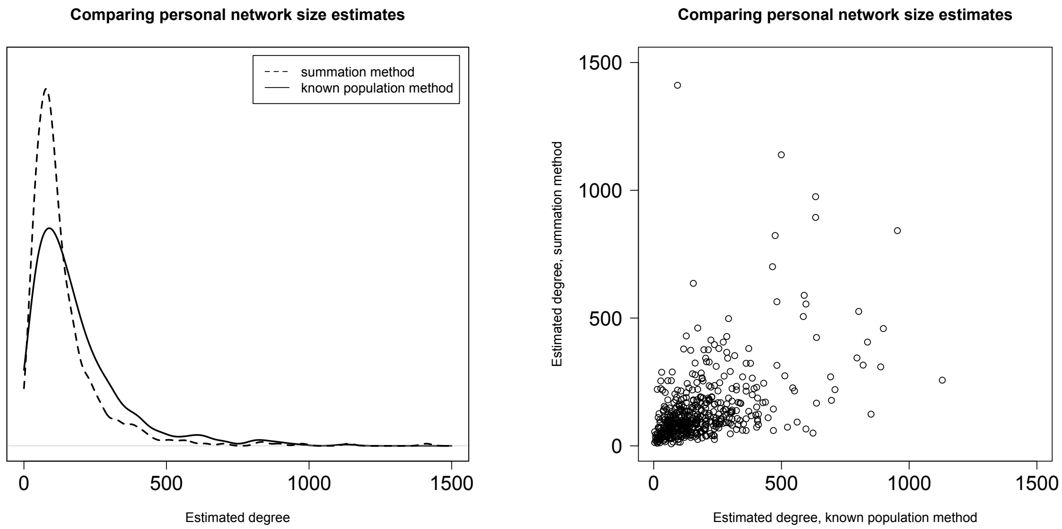
When analyzing data from the known population method, we follow the standard practice of top coding all responses at 30 (Zheng et al., 2006; McCormick et al., 2010); this affected less than 1% of survey responses. As a first check of the responses, we note that the mean number of people known in each group ($\frac{\sum_{i=1}^{n} y_{ik}}{n}$) is strongly correlated with the size of that group ($r = 0.86$) (Fig. 1). This provides us some confidence that, as has been found in other studies (Bernard et al., 2010), respondents in our study were able to answer these questions in a reasonable manner. Using the responses to questions about these 20 groups and the estimator described in Equation 1, the mean estimated personal network size in Curitiba is 184 (median: 138). Web Figure 2(a) plots the distribution of estimated personal network sizes (i.e., degrees) and we see that qualitatively this is similar to degree distributions estimated from previous studies: the data are right skewed and have a non-zero mode (McCarty et al., 2001; Zheng et al., 2006; McCormick et al., 2010).

To estimate personal network size using the summation method, one attempts to creates a set of exhaustive and mutually exclusive relationship types (McCarty et al., 2001). For example, through a series of interviews and focus groups, our study developed 22 relationship types appropriate to the Brazilian context including immediate family members living in your home, immediate family members living in other homes, friends from work, acquaintances from work, etc.; a full list in presented in Web Figure 3. From this data we can estimate personal network size by simply summing the responses of each respondent,

$$\hat{d}_{i,sum} = \sum_j y_{ij} \tag{2}$$

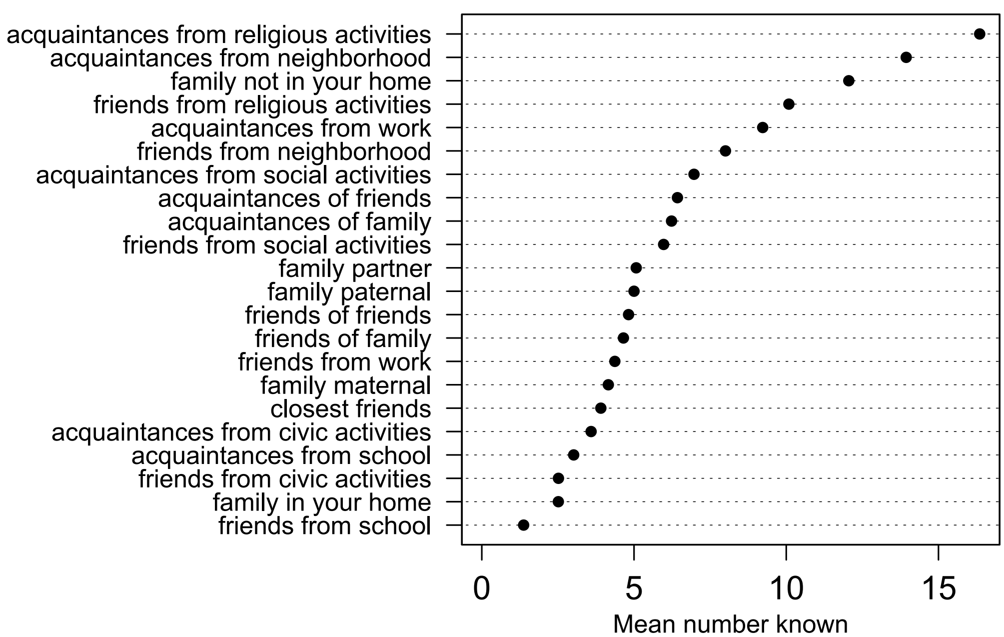
where y_{ij} is the number of people that respondent i knows in category j.

The mean estimated personal network size in Curitiba using the summation method is 140 (median: 99). Web Figure 2(a) plots a histogram of the estimated degrees and shows that at the population-level, both methods produce similar estimates for the distribution of personal network size in Curitiba. Further, at the individual-level, these two estimates are also highly correlated ($r = 0.49$) (Fig. 2(b)), a finding similar to a previous study in the United States (McCarty et al., 2001). While it is reassuring that the two personal network size estimation methods produce similar results, quantitatively the estimates are different (mean of 184 vs. mean of 140) and this difference would result in target population size estimates that are about 30% larger ($\sim \frac{184}{140}$) if the summation method estimated personal network size was used.



(a) (b)

Web Figure 2: Estimated degree distribution using the known population method and summation method. Both methods show qualitatively similar results at the population level. At the individual level, the correlation between the estimates is $r = 0.49$.



Web Figure 3: Mean number of people known in each of the 22 categories for used in the summation estimate.

Fortunately, we can make the choice of which personal network size method to use a data-driven decision by leveraging the fact that we asked about 20 populations of known size (Web Table 1). The intuition behind our approach is to see which measurement of personal network size can better predict the size of the 20 groups of known size. We consider two criteria for evaluating these estimates: mean absolute error (MAE) and root mean square error ($RMSE$) with the difference being that $RMSE$ penalizes larger errors more heavily because errors are squared. That is,

$$MAE_m = \frac{\sum_k \mid (\widehat{p_{k,m}} - p_k) \mid}{k} \tag{3}$$

and

$$RMSE_m = \sqrt{\frac{\sum_k (\widehat{p_{k,m}} - p_k)^2}{k}} \tag{4}$$

where m is the method of personal network size estimation (e.g., summation method or known population method), k is a population of known size (e.g., man more than 70 years old), p_k is the size of group k, and $\widehat{p_{k,m}}$ is the size of group k estimated using m as the method of personal network size estimation. Further, when estimating the size of group k using the known population method, we do not use group k in the personal network size estimation process in order to provide the fairest test. Web Figure 4 shows the estimated sizes of the 20 known groups using both network size estimation techniques. Broadly speaking, the estimates seem reasonable using either personal network size method, but by both criteria—$RMSE$ and MAE—the know population method personal network size estimates are preferred (known population method: $MAE = 0.57$, $RMSE = 0.80$; summation method: $MAE = 0.70$, $RMSE = 0.81$). For this reason, the results in the main paper are presented using the known population method.

Web Figure 4 also clearly shows one outlier, the largest group of known size: "middle school student in a public school." This pattern of underestimating the size of larger groups has been noted in earlier scale-up studies as well (Killworth et al., 2003; Zheng et al., 2006). Further, by noting that the group "middle school students in a private school" is estimated reasonably well ($\hat{p} = 1.0\%$, $p = 0.9\%$), we concluded that this outlier was the result of the size of the group, not something specific about middle school students.

We conclude by noting that the preceding approach for choosing between different design options, which we believe has not been used previously, is quite general and can be used in many other ways. For example, it could be used to measure which definition of "to know" leads to more accurate estimates, which interviewing protocol leads to more accurate estimates, etc. The fact that these research design decisions can be made based on empirical data rather than educated guesses is a huge advantage of the scale-up method and suggests the possibility of a cumulation of knowledge in this area. If replicated in other studies as well, we believe that the result that the known population method should be preferred to the summation method is one step in this process of gradual and cumulative improvement.

2 Network scale-up estimate

As defined in the main text, the network scale-up estimator is:

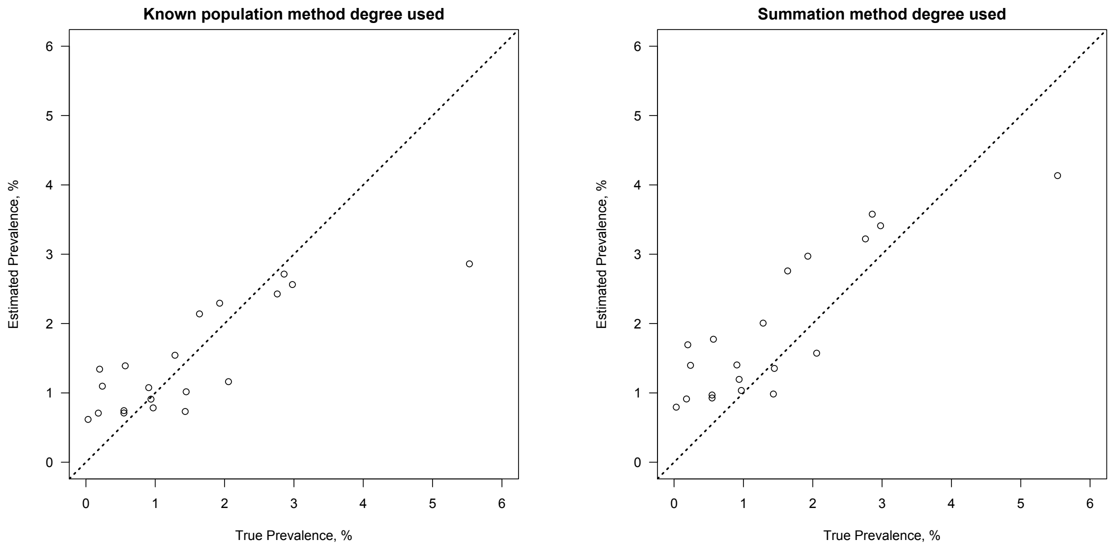
$$\hat{p} = \frac{\sum_i y_i}{\sum_i \hat{d}_i} \tag{5}$$

Using the $\hat{d}_i$ estimated via the known population method (see Section 1) we estimate

$$\hat{p} = \frac{3075}{92003} = 3.3\% \tag{6}$$

Thus, we can see that although our sample size was only 500 people, our survey collected information about an estimated 92,000 Curitibans. We caution that these are not necessarily 92,000 distinct people because of the possibility that the personal networks of our respondents overlap. To construct confidence intervals around this estimates, we begin by noting that the standard scale-up confidence interval procedure developed in Killworth et al. (1998) results in an estimated standard error of

$$\hat{se}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{\sum_i \hat{d}_i}}. \tag{7}$$



(a) Using known population method degree estimates (b) Using summation method degree estimates

Web Figure 4: Validation plots compared the actual size of 20 populations to their estimated size. Estimates made with the known population personal network size are closer to the true values as measured in terms of mean absolute error (MAE) and root mean squared error ($RMSE$).

However, this approach will underestimate the variance in our case for two reasons. First, the previous result assumes that we have independently sampled $\sum_i \hat{d}_i$ people (see Killworth et al. (1998) for a derivation), when in fact these $\sum_i \hat{d}_i$ people are nested within our 500 respondents and are therefore *not* independent. Second, the previous result assumes simple random sampling, but in our study we had a two-stage cluster sample design. To resolve both of these problems, we used a bootstrap approach (Efron and Tibshirani, 1993) where we generated replicate samples by first randomly re-sampling clusters and then re-sampling respondents within clusters. More specifically, let $C_1, C_2, \ldots C_{54}$ be the set of 54 census tracts in our sample and let $r_1, r_2, \ldots r_{n_C}$ be the n_C respondents from census tract C. We first resampled 54 census tracts with replacement from the set of 54 census tracts, and then for each census tract, C, in our replicate sample, we resample n_C respondents from the set of respondents from census tract C. We generated 10,000 replicate samples using this procedures and used the percentile method to produce an 95% confidence interval of [2.7%, 4.1%].

3 Generalized network scale-up method

The network scale-up estimator described in the previous section makes some implicit assumptions that are widely believed to be incorrect. For that reason Salganik and Feehan (2011) generalize the network scale-up estimator to relax these strong assumptions by introducing two additional components in the estimator and provide a method for estimating these additional components. We briefly review that work here.

One major implicit assumption of the network scale-up method is that people are aware of everything about those that they are connected to. However, this is unlikely to be the case for traits that are stigmatized or illegal (e.g., heavy drug use). In other words, respondents in our survey might be connected to a heavy drug user, but not aware of the heavy drug use. This problem is called "transmission error" in the scale-up literature because information about network ties is not always "transmitted." Generally, transmission error results in the number of people known in the target population (y_i in Eq. 5) being too low, yielding underestimates of target population size (Shelley et al., 1995, 2006). A second implicit assumption of the network scale-up method is that the target population, in our case heavy drug users, has the same average personal network size as the general population. Intuitively, we can imagine that if heavy drug users know fewer people on average they will be underrepresented in the set of people that we learn about using the scale-up method. Or, if heavy drug users tend to know more people than average, they would be overrepresented. Salganik and Feehan (2011) formalize these ideas and show that a new maximum likelihood estimator, which they call the generalized scale-up estimator, becomes:

$$\widehat{p} = \frac{\sum_i y_i}{\sum_i \hat{d}_i} \cdot \frac{1}{\hat{\delta}} \cdot \frac{1}{\hat{\tau}} \tag{8}$$

where $\hat{\tau}$ is the estimated information transmission rate and $\hat{\delta}$ is the estimated popularity ratio. One can see that if the estimated transmission rate ($\hat{\tau}$) is 1 and the estimated popularity ratio ($\hat{\delta}$) is 1, the generalized scale-up estimator reduces to the original scale-up estimator presented in Equation 1 in the main paper.

In our study, we collected the information needed to estimate τ and δ using a combination of data from the survey of the general population and a survey of the heavy drug users. The procedure to estimate τ involves only the survey of heavy drug users and was done using the *game of contacts* procedure, as described extensively elsewhere (Salganik et al., 2011). When describing this procedure, we follow standard network terminology and refer the respondents as *egos* and the people connected to them as *alters*. For a given ego, the game of contacts uses a series of questions about people with specific first names to sample from the set of alters of that ego (McCarty et al., 1997). Then, information is collected about each alter. For example, the ego can be asked "How many people do you know named Osvaldo?" and then for each Osvaldo known, the ego can be asked whether the alter is aware that the ego is in the target population. This procedure can be repeated for many names to generate a large sample of alters. The estimated transmission rate is then:

$$\hat{\tau} = \frac{\sum_i \frac{w_i}{\pi_i}}{\sum_i \frac{x_i}{\pi_i}} \tag{9}$$

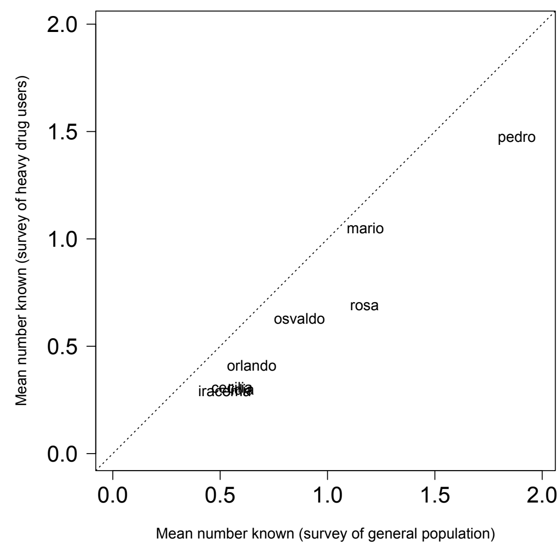
where w_i is the number of alters of respondent i that are aware that she is in the target population, x_i is the number of alters generated by the game for respondent i, and π_i is respondent i's probability of selection. Based on our sample of heavy drug users, we estimated $\hat{\tau} = 0.77$, with estimated 95% confidence interval of [0.73, 0.83]. In other words, we estimated that there is about a 75% chance that an alter connected to a heavy drug user in Curitiba will be aware that the given ego is a heavy drug user.

There are a number of possible sources of error in the estimated information transmission rate. Our procedure measures heavy drug users' perceptions of information transmission rates, not objective-truth. While these perceptions may be more important than actual behavior in some cases (Kitts, 2003), we speculate that there will be both systematic and random differences between the two. Random differences are less of a concern because our estimates are about heavy drug users in Curitiba as a group, not any particular individual, and therefore, involve a large amount of averaging. However, we suspect that there may be systematic error as well which could cause more serious problems. For example, ego's reports about alter's knowledge about ego may be inflated because of the "illusion of transparency," which leads people to believe that knowledge about their behavior is widely known (Gilovich et al., 1998). This and other possible sources of error are described in detail elsewhere (Salganik et al., 2011).

To estimate popularity ratio (δ), we combine information from the sample of heavy drug users with data from the survey of the general population. More specifically, respondents in both surveys were asked how many people they know with eight different names—Pedro, Rosa, Orlando, Lídia, Mário, Cecília, Osvaldo, and Iracema. By comparing these responses, we can estimate the differential popularity using the following intuition: if we find that members of the target population know, on average, one person named Pedro and members of the general population know, on average, two people named Pedro, then assuming that members of the target population are not especially likely or unlikely to know people named Pedro, we can estimate that members of the target population have 50% smaller networks or that $\hat{\delta} = 0.5$. Averaging over the eight names, we come up with the following estimator:

$$\hat{\delta} = \frac{\bar{y}_T}{\bar{y}_G} \tag{10}$$

where $\bar{y}_T$ if the average number of people known with those eight names by the people in the target population and $\bar{y}_G$ is the average number of people known with those names in the general population (Salganik and



Web Figure 5: Comparing the mean number people known with each of eight names in the survey of the general population and the survey of heavy drug users; dotted line represents line of equality. From these data we estimated a popularity ratio ($\hat{\delta}$) of 0.69. The figure shows that heavy drug users tend to give lower responses, a pattern that is consistent across names. The three partially overlapping names in the lower left portion of the figure are Lídia, Cecília, and Iracema.

Feehan, 2011). In our data, $\bar{y}_T = 0.644$ and $\bar{y}_G = 0.928$ resulting in $\hat{\delta} = 0.69$.[1]

There are a number of possible sources of error in the estimated popularity ratio. First, a portion of the difference in responses between the two populations could be due to survey mode effects. In the survey of heavy drug users, respondents were required to provide information about each person they knew with a specific name (see the description of the game of contacts above). However, in the general population survey, no additional information was collected about each alter. This difference in data collection procedures could lead respondents in the game of contacts to report fewer alters which would lead to an underestimate of the popularity ratio. Another potential source of error could arise if heavy drug users are more likely to know people by nickname, rather than proper name. This pattern has been reported anecdotally and would also lead to an underestimate of the popularity ratio. A third possible source of error is that our sample of heavy drug users was collected via respondent-driven sampling (Heckathorn, 1997, 2002; Salganik and Heckathorn, 2004) and this sampling method over-samples respondents who are known by many members of the target population (Salganik and Heckathorn, 2004). Therefore, the nature of our sample of heavy drug users could lead to an overestimate of the popularity ratio. A final potential source of error could come from the set of names used (for additional discussion on how the names were selected see Salganik et al. (2011)). If people with the chosen names are more likely to be known by heavy drug users this will lead to an overestimate of the popularity ratio, or likewise, if people with these names are less likely to be known by heavy drug users this would lead to an underestimate of the popularity ratio. However, Web Figure 5 shows that the popularity ratio for each specific name was similar and suggests that our choice of names did not have a strong effect on our findings.

Using the estimated transmission rate ($\hat{\tau}$) and the estimated popularity ratio ($\hat{\delta}$) we can construct the generalized scale-up estimate,

[1] Note that this calculation excludes one outlier from the general population survery who reported knowing 400 people named Pedro, 400 named Rosa, 100 named Orlando, 50 named Lidia, 200 named Mario, 50 named Cecilia, 50 named Osvaldo, and 50 named Iracema. This respondent was deemed to be an outlier for two reasons. First, of the 500 respondents in the scale-up survey, this was the only one who reported knowing more than 30 people with any name and second, the respondent had a summation method estimated degree of 232. If one were to include this outlier, $\bar{y}_G = 0.9857$ and therefore $\hat{\delta} = 0.65$.

$$\widehat{p}_g = \frac{\sum_i y_i}{\sum_i \widehat{d}_i} \cdot \frac{1}{0.69} \cdot \frac{1}{0.77} = \frac{3075}{92003} \cdot 1.88 = 6.3\% \tag{11}$$

To estimate the variance of $\widehat{p}_g$ we could linearize Equation (8) using a Taylor series expansion (see Sarndal et al. (1992)), but these results are likely to be unwieldy and their accuracy would be unclear. Therefore, we used a bootstrap approach (Efron and Tibshirani, 1993) where we generated replicates of the heavy drug users sample (using the procedure described in Salganik et al. (2011)) and the general population sample (using the procedure described in Section 2). Then for each pair of replicate samples, we recomputed $\sum_i y_i$, $\sum_i \widehat{d}_i$, $\hat{\delta}$, and $\hat{\tau}$. This approach preserves any covariance in the estimates that make up $\widehat{p}_g$. Performing the procedure for 10,000 replicates results in a bootstrap estimated 95% confidence interval of [4.5%, 8.0%]. As should be expected, the estimated confidence interval of the generalized scale-up estimate is larger than the standard scale-up estimate because of the additional estimates involved ($\hat{\delta}$ and $\hat{\tau}$). In this case, $\frac{\widehat{se}(\widehat{p}_g)}{\widehat{se}(\widehat{p})} \approx 2.5$ which seems reasonable.

Because this is the first time that the generalized network scale-up estimator has been used we cannot compare its performance to that in previous studies. Further, since we did not estimate the transmission rate, τ, or popularity ratio, δ, for other groups we cannot apply the generalized scale-up method to any other groups of unknown size or known size.

4 Multiplier method

Another method for estimating the sizes of hard-to-count populations is the multiplier method (UNAIDS, 2003). This method estimates the size of the target population based on two pieces of information: 1) a count of people in the target population with some specific characteristic and 2) an estimated prevalence of that characteristic in the target population. In our case, we learned that 423 heavy drug users were enrolled in the CAPS drug treatment program in Curitiba in August 2009. We attempted to ensure that the count from CAPS involved *unique* people (e.g., did not double-count individuals) and matched our study criteria. For example, there are a large number of people in CAPS who are addicted to alcohol. Since these people did not match our study criteria they were not included in the count of 423.

In a study of heavy drug users in Curitiba using respondent-driven sampling (RDS), we estimated that 3.7% of heavy drug users were in a CAPS treatment program (Bastos, 2009). Therefore, our multiplier estimate for the number of heavy drug users is

$$\frac{423}{\widehat{N}_{hdu}} = 3.7\% \tag{12}$$

which yields

$$\widehat{N}_{hdu} = \frac{423}{0.037} = 11459 \text{ people} \tag{13}$$

When converted to be expressed as a proportion of the general population this becomes:

$$\widehat{p}_m = \frac{11459}{1817434} = 0.6\% \tag{14}$$

To construct a confidence interval around this estimate, we start by constructing a confidence interval around our estimate that 3.7% of the heavy drug users in Curitiba are in CAPS treatment. The standard respondent-driven sampling bootstrap procedure (Salganik, 2006) results in a 95% confidence interval of [0.7%, 7.8%]. Using the endpoints of this interval and Eqs. 13 and 14, we create a 95% confidence interval of [0.3%, 3.2%] for the prevalence of heavy drug users in Curitiba. This interval is so wide because the estimated prevalence of treatment in a CAPS program is so low (less than 5%) and that estimate appears in the denominator of Equation (13). This leads to an unstable estimator where a small difference in estimated prevalence of CAPS treatment among heavy drug users leads to a large difference in the estimated number of heavy drug users. Note that the confidence interval for $\widehat{p}_m$ is asymmetric because the confidence interval for the RDS-estimated proportion of heavy drug users in CAPS treatment is asymmetric and because that estimate appears in the denominator of Eq. 13.

As wide as the confidence interval is around the multiplier estimate, it probably underestimates the true uncertainty because the procedure used to generate the RDS confidence intervals has been shown to produce intervals that are too small (Goel and Salganik, 2010). Unfortunately, a more accurate RDS confidence interval procedure is not yet known. Another source of uncertainty is the count of heavy drug users in CAPS treatment, but this uncertainty is difficult to assess statistically because it is not caused by sampling. However, we note that errors of, say, 10% in either direction (i.e., true count of 380 or 465) would increase or decrease our multiplier estimate by 10%, which is equivalent to a change of 0.3 percentage points. Finally, as noted in the main text, a potential source of bias in our estimate comes from our estimate of the proportion of heavy drug users in CAPS: 3.7%. If those who are in CAPS treatment are overrepresented in the RDS estimate, then the estimated prevalence of heavy drug users in CAPS treatment will be too high. The consequence of this overestimate, is that we will have an underestimate of the number of heavy drug users in Curitiba.

5 Direct estimation method

The final method used for estimating the number of heavy drug users was direct estimation (UNAIDS, 2003). This method involves asking a large sample of the general population if they are heavy drug users. While statistically well-grounded, direct estimation is likely to produce an underestimate, because as described in the main text, heavy drug users are less likely to be included in a standard household-based survey (Caspar, 1992; Zhao et al., 2009) and respondents will probably under-report their drug use in a survey (Fendrich et al., 1999; Colón et al., 2001; Delaney-Black et al., 2010).

We have two difference sources of data for direct estimates. First, in 2004 the Brazilian Ministry of Health conducted the PCAP survey which involved a sample of approximately 1,000 people in Curitiba and asked directly about use of powder cocaine and injected cocaine (Szwarcwald et al., 2005). Three people reported using these drugs frequently; Section 9 has the exactly question text in English and Portuguese and the response frequencies. From these data, a reasonable estimate for the prevalence of heavy drug users in the general population is

$$\hat{p}_{dir,2004} = \frac{1+2}{\frac{1188+1117}{2}} = 0.3\%. \tag{15}$$

Without the raw data files from the PCAP, to which we do not have access, we cannot fully account for the complex sample design when constructing confidence intervals around our estimates. However, we can approximate the design effect for our estimate by comparing the estimated standard errors in Szwarcwald et al. (2005), which were estimated in a way to account for the complex sample design, to the estimated standard errors that would have resulted from simple random sampling. Using four estimates presented in Table 4 of Szwarcwald et al. (2005)—injected cocaine at least once, injected cocaine currently, snorted cocaine at least once, snorted cocaine currently—we estimated approximate design effects ranging from 1.27 to 2. Therefore, being maximally conservative, we will assume a design effect of 2 for our estimate, which means that we need to inflate the standard error our estimate by $\sqrt{2}$:

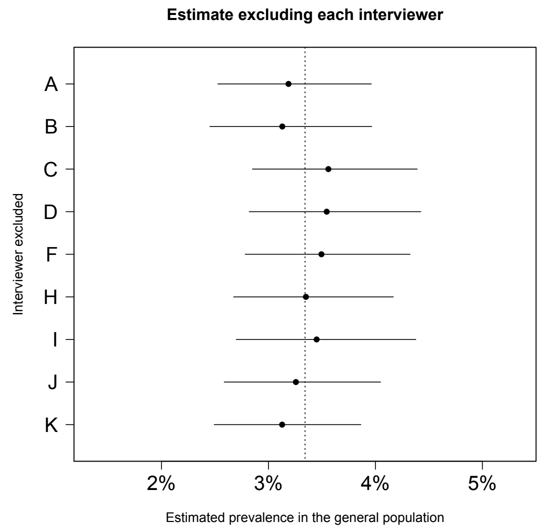
$$\hat{se}(\hat{p}_{dir,2004}) = \sqrt{\frac{\hat{p}_{dir,2004}(1-\hat{p}_{dir,2004})}{\frac{1188+1117}{2}}} \times \sqrt{2} = 0.002124 \tag{16}$$

which yields a 95% confidence interval of [0, 0.7%]. Note that the 2004 PCAP may underestimate the rate of heavy drug use because it only asked about cocaine. However, other drugs that were common in 2010 (e.g., crack and ecstasy) were likely uncommon in 2004 (Fonseca et al., 2010).

In 2010 on our general population survey, we asked respondents directly about the use of all illegal drugs other than marijuana; Section 9 has the exact question text in English and Portuguese and the response frequencies. From that data we estimate:

$$\hat{p}_{dir,2010} = \frac{3}{500} = 0.6\% \tag{17}$$

and using the bootstrapping procedure described in Section 2 we estimate a 95% confidence interval of [0%, 1.6%].



Web Figure 6: Estimated proportion of heavy drug users in Curitiba, Brazil excluding each interviewer one at a time. The vertical line is the estimate based on the entire sample. These results suggest that no one interviewer had an unusually large impact on our estimates.

6 Interviewer effects

Our network scale-up and generalized network scale-up estimates were substantially higher than those using other methods. One possible explanation for this pattern is that respondents reported knowing more people who were heavy drug users than they actually knew. As mentioned in the main text, we suspected that one source of such a problem could have been interviewers who did not properly carry out our protocol (although we had no evidence that this might have occurred). Specifically we were worried that some interviewers might have changed the somewhat complex question "How many people do you know who live in Curitiba and who have used illegal drugs other than marijuana more than 25 times in the last six months (i.e., average of once a week)?" to the simpler "How many people do you know that use drugs?" Such a change could obviously lead to responses that are too high which would in turn lead us to overestimate the target population size. Therefore, in this section we investigate whether there were interviewer effects on the responses.

Twelve different interviewers took part in the study, but 99% of the interviews were conducted by nine people. Here we explore the effects of these nine interviewers on estimates. Previous work on the network scale-up method has found relatively large interviewer effects on outcomes related to survey logistics—item-level missing data, interrupted surveys, and complete refusals—but did not explore the seemingly more important issue of interviewer effects on actual survey responses (Snidero et al., 2009). Web Figure 6 plots the scale-up estimates excluding data from each interviewer. That is, instead of one estimate based on data from nine interviewers, we produce nine estimates, each based on the data from eight interviewers. The resulting estimates range from 3.1% to 3.6%, and since the estimate using all interviewers was 3.3%, we conclude that no one interviewer had a substantial impact on our estimate.

7 Comparison to Brazilian and international benchmarks

7.1 Illicit drug users

In order to better understand the scale-up based estimates, we compared them to Brazilian and international benchmarks. Readers familiar with international estimates of the prevalence of drug users may suspect that our estimate are quite high, but a critical difference is that most studies are of *injecting drug users*, whereas our study was of *heavy drug users*, a potentially much larger group that includes non-injectors. As mentioned previously, we choose to study heavy drug users, rather than injecting drug users, because that is the most appropriate group to study given the current state of the HIV/AIDS epidemic in Brazil where injecting drugs is unusual and heavy drug users show high rates of HIV relative to the general population (Malta et al., 2010).

A 2008 study by the Brazilian Ministry of Health estimated that 0.5% of Brazilians nationally between the ages of 15 and 49 *injected* drugs in 2008 (Ministério da Saúde, 2011). We can express our network scale-up estimated number of heavy drug users to be with respect to Curitibans between the ages of 15 and 49, and we get an estimate of 5.8% (95% confidence interval: 4.7%, 7.1%).[2] We also note than an earlier study of heavy drug users in Curitiba estimated that only about 10% of them injected (Bastos, 2009), with others reporting use of ecstasy, cocaine, and crack, a finding that is consistent with the rapid spread of crack in Brazil (Duailibi et al., 2008). Therefore, after the conversions needed for comparison, we would estimate that about 0.58% of Curitibians between 15 and 49 inject drugs currently, an estimate roughly consistent with the previous national-level estimate by the Brazilian Ministry of Health. Our generalized network scale-up estimate, when converted to be an estimate of injecting drug user prevalence of Curitibians between 15 and 49 would be 1.1%, above the previous national-level estimate. Thus, our scale-up based estimates are at least consistent with national-level estimates in Brazil although it is clearly difficult to compare them precisely.

Our scale-up based estimates are also roughly consistent with international estimates at the country level summarized in the meta-analyses of Aceijas et al. (2004) and Mathers et al. (2008), but again, precise comparison is difficult. Because the previous meta-analyses present estimates in terms of the population between 15 and 64 years old, we convert our scale-up estimate to that scale and get an estimate of heavy drug use prevalence of 4.7% (95% confidence interval 3.8%, 5.6%)[3], which translates to an 0.47% estimated prevalence of injecting drug use among 15- to 64-year olds. This estimate is generally in line with country-level estimates found elsewhere in the world, and is similar to the estimated injecting drug use prevalence in Portugal (0.47%) and Slovakia (0.49%) (Mathers et al., 2008). Our generalized scale-up estimate (after suitable conversion) would be about 0.89% which is similar to the estimated prevalence in Italy (0.89%) and Ukraine (1.16%) (Mathers et al., 2008).

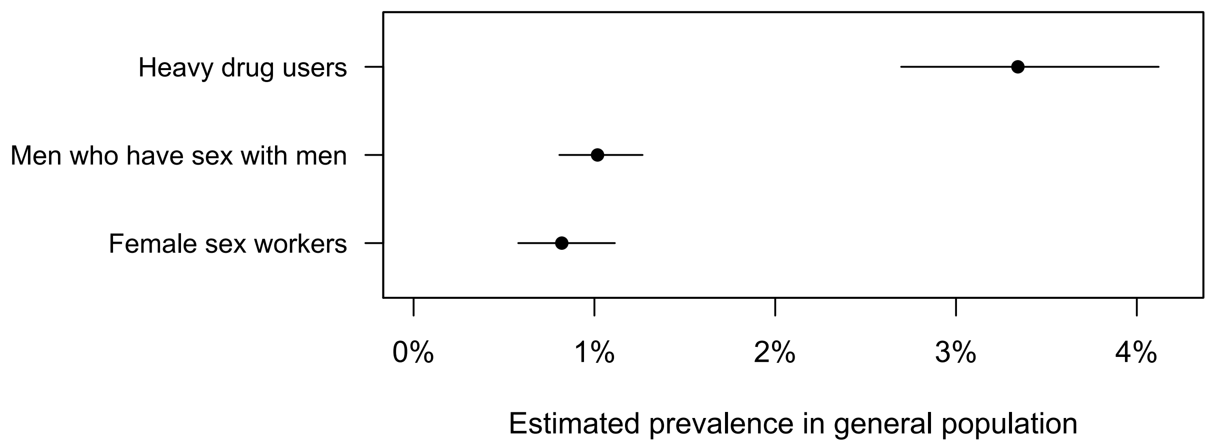
Finally, our scale-up based estimates are also roughly consistent with the estimated prevalence of injecting drug use (of 15-64 year olds) in 96 metropolitan areas in the United States (Brady et al., 2008). Our scale-up estimate of 0.47% for Curitiba and our generalized scale-up estimate of 0.89%, are at the lower-to-middle end of the estimates from US metropolitan areas which ranged from 0.37% (of population aged 15-64) in Ann Arbor, MI to 3.36% (of population aged 15-64) for Baltimore, MD. Thus, provided that one adjusts for the difference between heavy drug users and injecting drug users, our scale-up based estimates appear plausible when compared to national-level estimates from Brazil, national-level estimates from other countries, and city-level estimates from other countries. However, we again caution that these comparisons are difficult to do with any precision.

7.2 Female sex workers and men who have sex with men

As mentioned in the main text, the scale-up method allows researchers to estimate the sizes of several target populations in a single data collection. Therefore, our study also estimated the prevalence of female sex workers and men who have sex with men, two other stigmatized groups at risk for HIV/AIDS. Using the scale-up estimator (Eq. 5), we estimate that prevalence of female sex workers in Curitiba is 0.8% (95% confidence interval: 0.6%, 1.1%) (Web Figure 7); note that because we did not have a sample of female sex workers we could not calculate the correction factors needed for the generalized scale-up. A 2008 Brazilian Ministry of Health study produced a national-level estimate of 1.2% prevalence among women between 15 and 49 (Ministério da Saúde, 2011). If we convert our network scale-up estimate for the prevalence of female sex workers to be with respect to women between 15 and 49, we get an estimate of 2.8% (95% confidence interval: 2.0%, 3.8%), which is quite a bit higher than this previous Brazilian estimate. Estimates from other countries in Latin America range from 0.2% to 2.0% (with the exception of Belize which had an estimate of 7.4%) (Vandepitte et al., 2006), again suggesting that our estimate is somewhat higher than

[2] To convert our estimated prevalence for all of Curitiba to be an estimated prevalence for Curitibans between 15 and 49 we multiply our estimate by the number of people in Curitiba (1,817,434) and divide by the number of people between 15 and 49 (1,045,604).

[3] To convert our estimated prevalence for all of Curitiba to be an estimated prevalence for Curitibans between 15 and 64 we multiply our estimate by the number of people in Curitiba (1,817,434) and divide by the number of people between 15 and 64 (1,304,987).



Web Figure 7: Estimated prevalence of heavy drug users, men who have sex with men, and female sex workers in Curitiba, Brazil.

previous estimates. Turning to our estimate of men who have sex with men, using the scale-up estimator (Eq. 5), we estimate that prevalence of men who have sex with men is 1.0% (95% confidence interval: 0.8%, 1.3%) (Web Figure 7); note that because we did not have a sample of men who have sex with men we could not calculate the correction factors needed for the generalized scale-up. A 2008 Brazilian Ministry of Health study estimated a prevalence of 3.1% for men between 15 and 49 (Ministério da Saúde, 2011) which is slightly lower than our scale-up estimate converted to this scale: 3.6% (95% confidence interval: 2.9%, 4.5%). International estimates are often presented as a prevalence among men, not men aged 15 to 49, and if we convert to that scale, our estimate becomes 2.1% (95% confidence interval: 1.7%, 2.6%) which is within the ranges found in international meta-analyses although precise comparison is difficult because of definitional ambiguity (Caceres et al., 2006; Baral et al., 2007; Caceres et al., 2008). Thus, our scale-up estimates for these two other target groups are not radically different from national-level estimates in Brazil and are within internationally recognized ranges.

Overall, these comparisons of our estimates to previous Brazilian and international benchmarks do not in any way ensure that the network scale-up estimates are accurate. However, they do provide some sense that the estimates at least plausible.

8 Consistency of data sources

Because we used four different data sources, two of which were collected by other researchers, there were challenges in ensuring the consistency of the target population definition. Our target population was heavy drug users, defined to be people who have used illegal drugs other than marijuana more than 25 times in the past six months.

One data source was a face-to-face survey that was administered to a household-based random sample of 500 adult (i.e., 18 years and older) residents of Curitiba in 2010. This sample was used for the direct estimates in 2010 and the network scale-up and generalized network scale-up estimates. When asking about the respondent's drug use we asked two questions: 1) "Have you used illegal drugs other than marijuana in the past six months?" 2) If yes, we asked: "Have you used illegal drugs, other than marijuana, more than 25 times in the last 6 months (average of once a week)?" The exact question text in Portuguese and response frequencies are presented in Section 9. When asking about how many heavy drug users the respondent knows, we asked, "How many people do you know that live in Curitiba and used illegal drugs other than marijuana more than 25 times in the last 6 months (average of once a week)?"[4] Thus the scale-up, generalized scale-up,

[4]The question text in Portuguese was: "Quantas pessoas você conhece que vivem em Curitiba que usaram drogas ilícitas, que não a maconha, mais de 25 vezes nos últimos 6 meses (média de uma vez por semana)?"

and direct estimate from 2010 all use the same question wording, but the direct estimate is restricted to people 18 years and older while the scale-up based estimates have no age restriction.

A second data source was the respondent-driven sampling study of heavy drug users that was collected between July 28, 2009 and October 18, 2009 (Bastos, 2009). Because it was funded through a different process, that study used a slightly different criteria for screening participants: people 18 and older who have used cocaine, crack, methamphetamines, heroin or hallucinogens at least 25 days in the past 6 month and/or injected drugs at least once in the past six months.[5] There were only two respondents out of 303 in this sample who reported injecting drug in the past six months, but did not report using illegal drugs other than marijuana more than 25 times in the past six months. We have removed both of these respondents when estimating transmission rate (see Section 3) and the multiplier estimate (see Section 4). Including these cases does not change the estimates noticeably.

A third source of data were administrative records from the CAPS (Centro de Atenção Psicossocial) drug treatment program in Curitiba. From these administrative records, we believe that 423 people were enrolled in CAPS in August 2009 who met our definition of heavy drug users. This count is supposed to represent *unique* people (e.g., did not double-count individuals) and excluded people who were not heavy drug users by our definition (e.g., people who were addicted to alcohol).

Our final source of data was the 2004 Brazilian Ministry of Health PCAP survey, which measured the knowledge, attitudes, and practices of the Brazilian population with respect to HIV/AIDS (Szwarcwald et al., 2005). This study, which was conducted six years before our study, was not designed with our exact target population in mind. However, the survey had two drug use questions that we used for our estimates: "In relation to cocaine powder, you . . ." and "In relation to injecting cocaine, you . . ." with the answers choices being: never used, experimented but don't use any more, use infrequently, and use frequently. The exact question text in Portuguese and response frequencies are presented in Section 9. We considered the respondent a heavy drug user if he or she reported using either drug frequently. Note that the 2004 PCAP may underestimate the rate of heavy drug use because it only asked about cocaine. However, other drugs that were common in 2010 (e.g., crack and ecstasy) were likely uncommon in 2004 (Fonseca et al., 2010). Thus, despite having four different data sources, two of which were collected by other researchers for different purposes, we believe that definitional inconsistency, while it might have created minor differences, is not a major source of problems when comparing our five estimates.

9 Questions used for direct estimates

The questions and responses from the 2004 PCAP in both English and Portuguese are below (Ministério da Saúde, 2005).

Q: In relation to cocaine powder, you ["Em relação à cocaina em pó, você"] $n = 1188$

- never used ["nunca cheirei"], $n = 1115$
- experimented, but don't use any more ["já experimentei, mas não uso mais"], $n = 67$
- use infrequently ["uso de vez em quando"], $n = 5$
- use frequently ["uso frequentemente"], $n = 1$

Q: In relation to injecting cocaine, you ["Em relação à cocaina injetada na veia, você"], $n = 1117$

- never used ["nunca tomei"], $n = 1107$
- experimented, but don't use any more ["já experimentei, mas não uso mais"], $n = 8$
- use infrequently ["uso de vez em quando"], $n = 0$
- use frequently ["uso frequentemente"], $n = 2$

[5] The question texts in Portuguese were: "Você usou drogas injetáveis, pelo menos uma vez, nos últimos 6 meses?" and "Você usou cocaína, crack, metanfetaminas, heroína ou alucinógenos em pelo menos 25 dias, nos últimos 6 meses?

In 2010 we directly replicated the questions in 2004 and received similar responses. We also asked respondents directly in a way that matched our definition of heavy drug user.

Q: In relation to cocaine powder, you [“Em relação à cocaina em pó, você”] $n = 500$

- never used [“nunca cheirei”], $n = 453$
- experimented, but don’t use any more [“já experimentei, mas não uso mais”], $n = 42$
- use infrequently [“uso de vez em quando”], $n = 5$
- use frequently [“uso frequentemente”], $n = 0$

Q: In relation to injecting cocaine, you [“Em relação à cocaina injetada na veia, você”], $n = 500$

- nunca tomei [“never used”], $n = 491$
- já experimentei, mas não uso mais [“experimented, but don’t use any more”], $n = 8$
- uso de vez em quando [“use infrequently”], $n = 1$
- uso frequentemente [“use frequently”], $n = 0$

Q: Have you used illegal drugs other than marijuana in the past six months? [“Você usou qualquer droga ilícita, que não maconha, nos últimos 6 meses?”]

- Yes [“Sim”], $n = 11$
- No [“Não”], $n = 489$

If respondents answered yes to the previous question they were asked:
Q: Have you used illegal drugs, other than marijuana, more than 25 times in the last 6 months (average of once a week)? [“Você usou qualquer droga ilícita, que não maconha, mais de 25 vezes nesse período dos últimos 6 meses (média de uma vez por semana)?”]

- Sim [“Yes”], $n = 3$
- Não [“No”], $n = 8$

References

Aceijas, C., Stimson, G. V., Hickman, M., and Rhodes, T. (2004). Global overview of injecting drug use and HIV infection among injecting drug users. *AIDS*, 18(17):2295–2303.

Baral, S., Sifakis, F., Cleghorn, F., and Beyrer, C. (2007). Elevated risk for HIV infection among men who have sex with men in low- and middle- income countries 2000-2006: A systematic review. *PLoS Medicine*, 4(12):e339.

Bastos, F. I. (2009). Taxas de infecção de HIV e sífilis e inventário de conhecimento, atitudes e práticas de risco relacionadas às infecções sexualmente transmissíveis entre usuários de drogas em 10 municípios brasileiros. Technical report, Oswaldo Cruz Foundation, (Submitted on behalf of the Brazilian Multicity Study Group on Drug Misuse).

Bernard, H. R., Hallett, T., Iovita, A., Johnsen, E. C., Lyerla, R., McCarty, C., Mahy, M., Salganik, M. J., Saliuk, T., Scutelniciuc, O., Shelley, G. A., Sirinirund, P., Weir, S., and Stroup, D. F. (2010). Counting hard-to-count populations: the network scale-up method for public health. *Sexually Transmitted Infections*, 86(Suppl 2):ii11–ii15.

Brady, J., Friedman, S., Cooper, H., Flom, P., Tempalski, B., and Gostnell, K. (2008). Estimating the prevalence of injection drug users in the U.S. and in large U.S. metropolitan areas from 1992 to 2002. *Journal of Urban Health*, 85(3):323–351.

Caceres, C., Konda, K., Pecheny, M., Chatterjee, A., and Lyerla, R. (2006). Estimating the number of men who have sex with men in low and middle income countries. *Sexually Transmitted Infections*, 82 Suppl 3(82):3–9.

Caceres, C. F., Konda, K., Segura, E. R., and Lyerla, R. (2008). Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003-2007 estimates. *Sexually Transmitted Infections*, 84(Suppl 1):i49–i56.

Caspar, R. A. (1992). Follow-up of nonrespondents in 1990. In *Survey Measurement of Drug Use: Methodological Studies*, pages 155–173. National Institute on Drug Abuse, Rockville, MD.

Colón, H. M., Robles, R. R., and Sahai, H. (2001). The validity of drug use responses in a household survey in puerto rico: comparison of survey responses of cocaine and heroin use with hair tests. *International Journal of Epidemiology*, 30(5):1042 –1049.

Delaney-Black, V., Chiodo, L. M., Hannigan, J. H., Greenwald, M. K., Janisse, J., Patterson, G., Huestis, M. A., Ager, J., and Sokol, R. J. (2010). Just say "I don't": Lack of concordance between teen report and biological measures of drug use. *Pediatrics*, 126(5):887–893.

Duailibi, L. B., Riberio, M., and Laranjeira, R. (2008). Profile of cocaine and crack users in Brazil. *Cadernos de Saude Publica*, 24(S4):545–557.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall/CRC.

Fendrich, M., Johnson, T. P., Sudman, S., Wislar, J. S., and Spiehler, V. (1999). Validity of drug use reporting in a High-Risk community sample: A comparison of cocaine and heroin survey reports with hair tests. *American Journal of Epidemiology*, 149(10):955 –962.

Fonseca, A. M., Galduróz, J. C., Noto, A. R., and Carlini, E. L. (2010). Comparison between two household surveys on psychotropic drug use in Brazil: 2001 and 2004. *Ciência & Saúde Coletiva*, 15(3):663–670.

Gilovich, T., Savitsky, K., and Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, 75(2):332–346.

Goel, S. and Salganik, M. J. (2010). Assessing Respondent-Driven sampling. *Proceedings of the National Academy of Science, USA*, 107(15):6743–6747.

Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, 44(2):174–199.

Heckathorn, D. D. (2002). Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems*, 49(1):11–34.

Killworth, P. D., McCarty, C., Bernard, H. R., Johnsen, E. C., Domini, J., and Shelly, G. A. (2003). Two interpretations of reports of knowledge of subpopulation sizes. *Social Networks*, 25:141–160.

Killworth, P. D., McCarty, C., Bernard, H. R., Shelly, G. A., and Johnsen, E. C. (1998). Estimation of seroprevalence, rape, and homelessness in the U.S. using a social network approach. *Evaluation Review*, 22:289–308.

Kitts, J. A. (2003). Egocentric bias or information management? selective disclosure and the social roots of norm misperception. *Social Psychology Quarterly*, 66(3):222–237.

Malta, M., Magnanini, M. M., Mello, M. B., Pascom, A. R., Linhares, Y., and Bastos, F. I. (2010). HIV prevalence among female sex workers, drug users and men who have sex with men in Brazil: A systematic review and meta-analysis. *BMC Public Health*, 10(1):317.

Mathers, B., Degenhardt, L., Phillips, B., Wiessing, L., Hickman, M., Strathdee, S., Wodak, A., Panda, S., Tyndall, M., and Toufik, A. (2008). Global epidemiology of injecting drug use and HIV among people who inject drugs: A systematic review. *The Lancet*, 372(9651):1733–1745.

McCarty, C., Bernard, H. R., Killworth, P. D., Shelley, G. A., and Johnsen, E. C. (1997). Eliciting representative samples of personal networks. *Social Networks*, 19(4):303–323.

McCarty, C., Killworth, P. D., Bernard, H. R., Johnsen, E., and Shelley, G. A. (2001). Comparing two methods for estimating network size. *Human Organization*, 60:28–39.

McCormick, T., Salganik, M. J., and Zheng, T. (2010). How many people do you know?: Efficiently estimating personal network size. *Journal of the American Statistical Association*, 105(489):59–70.

Ministério da Saúde (2005). *Pesquisa de Conhecimento, Atitudes e Práticas na População Brasileira de 15 a 54 anos, 2004*. Ministério da Saúde, Secretaria de Vigilância Saúde, Programa Nacional de DST e Aids. - Brasília: Ministério da Saúde.

Ministério da Saúde (2011). MonitorAIDS. http://sistemas.aids.gov.br/monitoraids/, (Acessed: 2011-04-15).

Salganik, M. J. (2006). Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *Journal of Urban Health*, 83(7):98–112.

Salganik, M. J. and Feehan, D. (2011). Generalized network scale-up method for estimating the size of hard-to-count populations. Technical report, Office of Population Research, Princeton University.

Salganik, M. J. and Heckathorn, D. D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, 34:193–239.

Salganik, M. J., Mello, M. B., Abdo, A. H., Bertoni, N., Fazito, D., and Bastos, F. I. (2011). The game of contacts: Estimating the social visibility of groups. *Social Networks*, 33(1):70–78.

Sarndal, C., Swensson, B., and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer, New York.

Shelley, G. A., Bernard, H. R., Killworth, P., Johnsen, E., and McCarty, C. (1995). Who knows your HIV status? what HIV+ patients and their network members know about each other. *Social Networks*, 17(3-4):189–217.

Shelley, G. A., Killworth, P. D., Bernard, H. R., McCarty, C., Johnsen, E. C., and Rice, R. E. (2006). Who knows your HIV status II?: information propagation within social networks of seropositive people. *Human Organization*, 65(4):430–444.

Snidero, S., Zobec, F., Berchialla, P., Corradetti, R., and Gregori, D. (2009). Question order and interviewer effects in CATI scale-up surveys. *Sociological Methods & Research*, 38(2):287–305.

Szwarcwald, C. L., Barbosa-Júnior, A., Pascom, A. R., and de Souza-Júnior, P. R. (2005). Knowledge, practices and behaviours related to HIV transmission among the Brazilian population in the 15-54 years age group, 2004. *AIDS*, 19(Suppl 4):S51–58.

UNAIDS (2003). *Estimating the Size of Populations at Risk for HIV*. Number 03.36E. UNAIDS, Geneva.

Vandepitte, J., Lyerla, R., Dallabetta, G., Crabbe, F., Alary, M., and Buve, A. (2006). Estimates of the number of female sex workers in different regions of the world. *Sexually Transmitted Infections*, 82 Suppl 3:iii18–iii25.

Zhao, J., Stockwell, T., and Macdonald, S. (2009). Non-response bias in alcohol and drug population surveys. *Drug and Alcohol Review*, 28(6):648–657.

Zheng, T., Salganik, M. J., and Gelman, A. (2006). How many people do you know in prison?: Using overdispersion in count data to estimate social structure in networks. *Journal of the American Statistical Association*, 101:409–423.