Assessing differences in groups randomized by recruitment chain in a respondent-driven sample of Seattle-area injection drug users

Richard D. Burt, Ph.D.[1,*] and **Hanne Thiede, DVM, MPH**[1]

[1]HIV/STD Program, Public Health-Seattle & King County, 401 Fifth Ave., Suite 1152, Seattle WA 98104

Abstract

Purpose—Respondent-driven sampling (RDS) is a form of peer-based study recruitment and analysis that incorporates features designed to limit and adjust for biases in traditional snowball sampling. It is being widely used in studies of hidden populations. We report an empirical evaluation of RDS's consistency and variability, comparing groups recruited contemporaneously, by identical methods and using identical survey instruments.

Methods—We randomized recruitment chains from the RDS-based 2012 National HIV Behavioral Surveillance survey of Seattle-area injection drug users into two groups and compared them in terms of sociodemographic characteristics, drug-associated risk behaviors, sexual risk behaviors, HIV status and HIV testing frequency.

Results—The two groups differed in 5 of the 18 variables examined (p .001): race (for example, 60% white vs. 47%), gender (52% male vs. 67%), area of residence (32% downtown Seattle vs. 44%), an HIV test in the previous 12 months (51% vs. 38%). The difference in serologic HIV status was particularly pronounced (4% positive vs. 18%). In 4 further randomizations, differences in one to five variables attained this level of significance, though the specific variables involved differed.

Conclusions—We found some material differences between the randomized groups. While the variability of the present study was less than has been reported in serial RDS surveys, these findings indicate caution in the interpretation of RDS results.

Purpose

Surveys of populations at risk for HIV can provide important information on HIV prevalence, risk behavior, testing practices and access to medical care which can help guide public health response to HIV. However, accessing populations at elevated risk for HIV, such as injection drug users (IDU) and men who have sex with men (MSM), can be

challenging, as these populations are to greater or lesser extent covert due to stigma and legal jeopardy associated with homosexuality and drug injection.

Respondent-driven sampling (RDS) is an approach which has been proposed to be advantageous for surveying such hidden populations (1). In RDS, participants are provided with coupons with which to recruit their peers and are compensated when the coupons are redeemed by new participants. Methods have been developed to analyze RDS-recruited study populations which provide adjustment for differences among participants in their social network size and for differential recruitment among participants with differing characteristics (2–6). Mathematical theory and modeling studies have asserted that the resulting estimates of population characteristics are asymptotically unbiased and independent of the characteristics of the initial participants (4;5;7). In recent years RDS has become a widely used methodology for surveying populations at risk for HIV throughout the world (8–10).

The mathematics of RDS adjustments, however, are based on a number of assumptions (such as random recruitment within a participant's social network and consistent reporting of network sizes among participants) which may not reflect actual conditions (11–13). The accuracy and variability of RDS have been assessed by several approaches: The characteristics of the same target population recruited by RDS and by other methods have been compared (14–23). Variability in RDS measurements has been evaluated in computer modeling based on populations with a known network structure (24;25), and with computer-generated network structures (2;11). Sequential RDS-derived study populations have been compared (21;26–30). While useful, the interpretation of each of these approaches has limitations: comparisons with other methods begs the question of which method more accurately reflects the target population; computer modeling methods are dependent on the extent to which the models reflect reality; sequential comparisons are affected by temporal changes in study populations and potential differences in survey methods and administration.

In 2005, 2009 and 2012 the National HIV Behavioral Surveillance system (NHBS) conducted surveys of IDU using RDS in some 20 U.S. cities, including Seattle, as part of a program of serial surveys of IDU, MSM and persons at elevated risk of heterosexual HIV transmission (31). In this report, we use the 2012 Seattle-area NHBS survey of IDU to evaluate consistency and variability in an RDS-recruited study population. We divided the study population into two groups based on allocating recruitment chains by a randomization algorithm. We then compared the groups in terms of sociodemographic characteristics, drug-associated risk behaviors, sexual risk behaviors, and HIV status and testing behavior. This study design allows a comparison of two groups recruited simultaneously, by identical methods, and evaluated by the same survey instruments. It thus avoids the effects of changes over time and differences in study design and implementation which could have affected previous evaluations of RDS methodology.

Methods

Recruitment

Following standardized NHBS protocols, seeds were recruited to provide representation of the diversity of Seattle-area IDU in race, sex, age, area of residence, drug of choice and sexual orientation. Seeds were given 5 coupons. Subsequent participants were originally issued three coupons, which was decreased to two, then one coupon in order to balance the number of interviews with the study appointment slots available. The study protocol closely matched that of the 2009 NHBS IDU survey (29), though study offices were located in a different area of downtown Seattle. Eligibility criteria required participants to be 18 years or older, residents of King or Snohomish Counties, able to complete the interview in English and to display either physical evidence of recent drug injection or demonstrate convincing knowledge of injection practices. Participants were screened, interviewed and gave informed consent in face-to-face interviews conducted with hand-held computers. HIV testing was by a rapid test on a finger stick blood sample (OraSure technologies, Bethlehem PA), followed by a blood-based Western Blot on those with reactive rapid test results (Bio-Rad, Hercules, CA). Study procedures were approved by the Washington State Institutional Review Board.

Randomization procedure

We used recruitment chain as the basis to assign participants to one of two groups. This ensured availability of complete data on who recruited whom, from which the adjustments of RDS-based estimates could be calculated. A random number between 0 and 1 was assigned to each recruitment chain. In many RDS study populations there are substantial differences in the size of the different recruitment chains. Preliminary investigations using groups simply defined by random number (>.50 vs. .50) produced one group which exceeded 70% of the total survey study population more frequently than thought desirable; for instance, this occurs in the randomization depicted in Table 1.

To ensure more comparable group sizes, we used a two-step randomization procedure (illustrated in Table 1). First, the chains were ordered by random number. Participants in recruitment chains below a certain breakpoint would be assigned to group 1, those above to group 2. The breakpoint was defined in the following manner: For each potential breakpoint in the randomization the number of participants in recruitment chains above and below the breakpoint was calculated. The breakpoint which produced the smallest difference in number of participants between groups was chosen to define the two analysis groups for this randomization. A priori, we chose the first randomization performed to present more detailed findings, and then summarize results from all five randomizations that were conducted using this procedure to further assess variability across randomizations.

Variable definitions

We compared the randomized groups in terms of a collection of 16 variables with a total of 46 variable categories. There were constructed to be comparable with a previous comparison of participants in the 2005 and 2009 NHBS surveys of Seattle-area IDU (29), and used a questionnaire that was similar to, and in most cases identical to, the 2009 survey. One difference is that in the 2012 questionnaire unprotected, HIV non-concordant, male-to-male

anal sex was evaluated by a series of questions on the number of male-to-male main and casual anal sex partners, the number with whom a condom was not used, the number for whom HIV status was known, and the number HIV-positive and HIV-negative. For heterosexual contacts, the same more general question was asked as in previous surveys: "Did you have vaginal or anal sex without a condom with a woman (or for women, a man) who was HIV-negative?" followed by analogous questions for HIV-positive partners and partners of unknown status. As serologic testing for HIV and hepatitis C (HCV) was performed in 2012, we present serologic status for these viruses rather than the self-reported status of the earlier study.

Statistical evaluation

We used statistical testing (using a criterion of p   .001) as a means to identify differences between groups that merit attention, to provide an objective measure of the extent of such differences across randomizations, and to compare the number of such differences with those found in previous comparisons of serial RDS study populations (21;26–30). Several means of adjusting RDS-generated data have been proposed (2–6). We used the Salganik-Heckathorn estimator-based RDSAT software package, which is freely available and widely employed (32). Statistical testing in RDS-generated data remains problematic and no method has gained general acceptance. The p-values we present incorporate RDSAT-derived weights for individual participants in logistic regression analyses (27;30). This allows a summary measure across multiple categories of a variable incorporating adjustments for network size and differential recruitment patterns.

We used RDSAT to generate individual weights with respect to the outcome variable of interest. The weights were then applied in univariate logistic regression models using randomized group as the dependent variable and the variable being evaluated as the independent variable. P-values are based on a likelihood ratio test. We present exact p-values, even when very small, to help guide interpretation of the significance of differences observed. In doing so we are not claiming high precision in the likelihood ratio p-values, but rather providing an indication of the probability that the differences measured would retain significance even after large (but unknown) correction for the higher variability of RDS methods compared to simple random sampling, and after correction for multiple testing considerations.

Possible confounding of HIV status by MSM status was assessed by including a term for MSM status in the logistic regression model for HIV status. In order to compare our findings with previous literature, we also indicated where the 95% confidence intervals of the two randomized groups do not overlap; this analysis is based on category-by-category comparisons, rather than a comparison across all categories of a variable. Analyses were conducted in SPSS (33).

Results

Recruitment

Nine seeds were interviewed, 8 of whom recruited additional participants. Of the 1,274 recruitment coupons distributed, 744 (58%) were returned by potential participants between 7/9/2012 and 11/29/2012. Altogether, 750 persons (including the 9 seeds) were screened for study eligibility (3 interviews were lost due to computer malfunction): 45 of these were excluded because they lacked evidence of injection, 7 were previous participants, 6 were judged to have provided invalid data, 2 resided outside of the study area, and 2 were incapacitated. This left 688 participants available for the present analysis.

The recruitment chain of seed 6 was the largest, with 229 eligible participants (33% of the total study population), followed by seed 2 (23%) and seed 8 (22%) (Table 1; Figure 1). The maximum number of recruitment waves was 15; 81% of participants were recruited at wave 4 or higher and 50% at wave 7 or higher. Ten participants (1.5%) reported being recruited by a stranger.

First Randomization

In the first randomization, among the 46 categories of the variables investigated, three in group 1 (all were categories of age) and none in group 2 had an absolute difference of greater than 2% (but less than 3%) between the sample population proportion and the calculated equilibrium estimate (Supplemental on-line material, Tables S1–S4). This comparison has been suggested as a measure of the independence of an RDS-recruited study population from seed characteristics. Homophily is a measure of the tendency of participants to recruit others similar to themselves in a given characteristic, based on a scale from +1 (all recruitment occurred between persons sharing a given characteristic) to –1 (no such recruitment). Combining group 1 and group 2 together, homophily over an absolute value of 0.50 was found for: serologic HIV status (both positive and negative), reporting no male-to-male sex, and heroin as the drug most frequently injected. The design effect evaluates the proportional difference between the RDS bootstrap-derived variance and variance that would be expected from simple random sampling. The design effect had a median value of 2.91 (range 0.52 – 10.32).

In the first randomization, 5 of the 16 variables examined differed between the randomized groups with a p-value $\leq$ .001 (race, sex, area of residence, HIV testing in the previous 12 months, and serologic HIV status) (Tables 2 and 3). The largest absolute differences was for female sex (48% vs. 33%), followed by HIV seropositivity (4% positive vs. 18%), white race (60% vs. 47%), 12-month HIV testing (51% vs. 38%), and residence in downtown Seattle (32% vs. 44%). The logistic regression results incorporating RDSAT-derived individual weight differed substantially from analogous unweighted models (data not shown). There was no variable category among the 46 evaluated in which 95% confidence intervals for the two groups did not overlap.

Results from all 5 randomizations

As three recruitment chains (2, 5 and 8) dominated the study population in terms of number of participants, it was possible that the randomization procedure would tend to favor a specific assignment of these chains with respect to one another (for instance chains 5 and 2 vs. chain 8), so that the randomizations would differ mostly in the assignment of the less populated chains. However, among the five randomizations there was at least one in which a group containing each one of these three chains was compared to a group containing the other two chains (Tables 4, S5, S8, S11, S14; Figures S1–S4).

Among the 16 variables investigated, the number for which differences between the randomized groups had a p-value .001 varied from 1 to 5 (Table 4; details in tables S6, S7, S9, S10, S12, S13, S15, S16). Among the 46 variable categories, the number for which the 95% confidence intervals in the two randomized groups did not overlap varied from 0 to 3. The number of significant differences by each measure in the first randomization was within the range seen among the other randomizations.

Serologic HIV status showed particularly marked differences between randomized groups in 4 of the 5 randomizations: 4% vs. 18% (1st randomization), 15% vs. 8% (2nd), 3% vs. 25% (3rd) and 6% vs. 17% (5th). The difference was less pronounced in the in the 4th randomization, 10% vs. 13% (Tables 3, S7, S10, S13, S16). It is also of interest that none of the randomizations with p-values <.01 for differences in HIV testing between groups when testing was evaluated within the past 12 months showed a comparably significant difference when testing was evaluated within a three month time frame (Tables 3, 4, S10).

MSM/IDU in the Seattle area, and especially amphetamine-injecting MSM/IDU, have dramatically higher HIV prevalence than non-MSM IDU (34–36). The differences in HIV status between randomized groups could thus reflect an uneven distribution of such MSM/IDU across recruitment chains. When an additional term for MSM status was included in the logistic regression analysis for HIV status, HIV status still differed between randomized groups in these 4 randomizations noted above with a p-value 10^{-4}.

Conclusions

We have attempted to evaluate the inherent variability of RDS by randomizing participants in an RDS-recruited survey of IDU into two groups and comparing the characteristics of those groups. This study design evades problems of changes in characteristics over time and of differences in study design and implementation. We found differences in between 1 and 5 of 16 variables examined in 5 separate randomizations, based on a p-value of .001. Identifying differences in terms of non-overlapping 95% confidence intervals, we found differences in from 0 to 3 of 46 variable categories. The two methods of evaluating statistical significance did not always identify the same variable as differing between groups.

Using statistical testing as the arbiter of such comparisons, however, is not entirely satisfactory. Such statistical testing is structured to reject the hypothesis of differences between groups unless there is clear and convincing evidence of a difference. For present purposes, this would seem to unduly prejudice the evaluation towards a finding of no

difference. Rather, we suggest using the statistical testing as a guide to identify variables to inspect more closely and to form a necessarily subjective judgment as to what difference would materially affect the characterization of the study population. In the first randomization, for instance, our personal judgment is that the differences between two groups in HIV status and sex describe materially different populations while the differences in HIV testing and area of residence are of less moment.

In a comparison of the 2005 and 2009 Seattle-area NHBS IDU RDS-recruited surveys, 8 of 16 variables differed on the basis of a p-value .001, more than in the present study (29). This suggests that some, but not all, of the variability between the 2005 and 2009 surveys could be a product of changes over time in the study population, differences in study office location or in survey administration. An even higher order of difference was seen in a comparison between 2005 NHBS survey and two previous institutionally-based surveys (22).

Other studies have reported differences between RDS-recruited study populations from the same city surveyed at different times (Table 5). Based on a criterion of non-overlapping 95% confidence intervals, 9 of 26 variable categories differed between sequential surveys in Cape Town (30), 12 of 36 in Jinan (28), 3 of 8 in Guangzhou (21;27), and 4 of 30 in Beijing (26). As already noted, these findings are qualified by potential changes over time and differences in study implementation. That being said, all of these comparisons had a higher proportion of study variables differing between surveys than the present study.

In a comparison of Seattle NHBS surveys of MSM in 2008 and 2011 recruited by venue-day-time sampling (VDTS), only one of 21 variables differed between surveys with a p-value .001, and for that variable there was independent evidence of a true change over time (37). The higher proportion of variables differing at this level or higher in the present study could reflect differences between MSM and IDU, but also raises the possibility of higher variability in RDS than in VDTS study populations.

Several limitations should be recognized in the interpretation of our data. Our data derive from one survey of an IDU population at one point in time in one city, so that our findings would require independent replication elsewhere and in other target populations to assess their generalizability. The logistic regression p-values we present incorporate adjustments of RDS in their point estimates but are not adjusted for the higher variance associated with the RDS procedure compared to simple random sampling (24;38;39), and so are likely to overestimate the true significance of the differences observed. Different levels of social desirability bias may pertain to differing subpopulations of IDU, which could produce inaccuracy in comparisons of self-reported data. The groups defined by the randomization scheme of the present study cannot be strictly claimed to represent simultaneous independent samples of the same underlying population. For instance, participants recruited in one chain may have been accessible through shared social networks for recruitment into another chain but would have been ineligible because they previously participated. Finally, the number of participants in the randomized groups varied from 291 to 412. While these figures are not wholly out of line with the numbers being reported in RDS studies, higher numbers of participants would have increased the resolution of the study.

In summary, we assessed variability among participants in a RDS-recruited study population randomized by recruitment chain. While precise statistical evaluation remains difficult, we judge that in each randomization there were differences between the randomized groups that materially affected the characterization of the study population. The variability found in the present analysis was less than what had been seen between RDS-recruited serial surveys in Seattle and elsewhere. This difference could be a product of genuine changes over time in the target population or could possibly reflect sensitivity of RDS methods to difference in the details of RDS implementation. Serial VDTS-recruited surveys of Seattle-area MSM had less variability than seen in the present study. Modeling studies (2;11;24), and the comparison of RDS-recruited populations with known population characteristics (19;40), have raised questions about the precision of RDS results. Combined, these findings raise a note of caution with regard to the accuracy of RDS estimates of the characteristics of specific study populations. Alternative methods of recruiting IDU also have well-recognized problems (41;42) and on the basis of current knowledge it is difficult to find common terms upon which to compare the relative inherent variability of RDS and its alternatives. Serial VTDS surveys of a population with a history of RDS surveys could offer the potential of such a head-to-head comparison.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations Used

HIV	human immunodeficiency virus
HCV	Hepatitis C virus
IDU	injection drug users
NHBS	National HIV Behavioral Survey
RDS	respondent-driven sampling

Reference List

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

2. Gile KJ. Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. J Am Stat Soc. 2012; 103(493):135–146.

3. Heckathorn DD. Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. Sociological Methodology. 2007; 37:151–207.

4. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology. 2004; 34(1):193–240.

5. Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. J of Official Statistics. 2011; 24:79–97.

6. Tomas A, Gile KJ. The effect of differential recruitment, non-response and non-recruitment on estimators for respondent-driven sampling. Elec J Statistics. 2011; 5:899–934.

7. Heckathorn DD. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. Social Problems. 2002; 29:11–34.

8. Malekinejad M, Johnston LG, Kendall C, Kerr LR, Rifkin MR, Rutherford GW. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. AIDS Behav. 2008; 12(4 Suppl):S105–S130. [PubMed: 18561018]

9. Johnston LG, Malekinejad M, Kendall C, Iuppa IM, Rutherford GW. Implementation challenges to using respondent-driven sampling methodology for HIV biological and behavioral surveillance: Field experiences in international settings. AIDS Behav. 2008; 12(Suppl):S131–S141. [PubMed: 18535901]

10. Montealegre JR, Johnston LG, Murrill C, Monterroso E. Respondent driven sampling for HIV biological and behavioral surveillance in Latin America and the Caribbean. AIDS Behav. 2013; 17:2312–2340.

11. Gile KJ, Handcock MS. Respondent-driven sampling: An assessment of current methodology. Sociological Methodology. 2010; 40(1):285–327. [PubMed: 22969167]

12. Paquette D, Bryant J, de Wit J. Respondent-driven sampling and the recruitment of people with small injecting networks. AIDS B. 2012; 16:890–899.

13. Rudolph AE, Fuller CM, Latkin C. The importance of measuring and accounting for potential biases in respondent-driven sampling. AIDS Behav. 2014; 17:2244–2252. [PubMed: 23515641]

14. Guo Y, Li X, Fang S, Lin X, Song Y, Jiang S. A comparison of four sampling methods among men having sex with men in China: implications for HIV/STD surveillance and prevention. AIDS Care. 2011; 23:1400–1409. [PubMed: 21711162]

15. Weir SS, Merli MG, Li J, Gandha AD, Neeley WW, Edwards JK. A comparison of respondent-driven and venue-based sampling of female sex workers in Liuzhou, China. Sex Transm Infect. 2013; 88(Suppl 2):i95–101. [PubMed: 23172350]

16. Rudolph AE, Crawford ND, Latkin C, Heimer R, Benjamin EO, Jones KC, et al. Subpopulations of illicit drug users reached by targeted street outreach and respondent-driven sampling strategies: implications for research and public health practice. Ann Epidemiol. 2011; 21(4):280–289. [PubMed: 21376275]

17. Kral AH, Malekinejad M, Vaudrey J, Martinez AN, Lorvick J, McFarland W, et al. Comparing respondent-driven sampling and targeted sampling methods of recruiting injection drug users in San Francisco. J Urban Health. 2010; 87(5):839–850. [PubMed: 20582573]

18. Robinson WT, Risser JM, McGoy S, Becker AB, Rehman H, Jefferson M, et al. Recruiting injection drug users: a three-site comparison of results and experiences with respondent-driven and targeted sampling procedures. J Urban Health. 2006; 83(6 Suppl):i29–i38. [PubMed: 16933101]

19. McCreesh N, Frost S, Seeley J, Katongole J, Tarsh MN, Ndunguse R, et al. Evaluation of Respondent-driven sampling. Epidemiology. 2012; 23(1):138–147. [PubMed: 22157309]

20. Wei C, McFarland W, Colfax GN, Fuqua V, Raymond HF. Reaching black men who have sex with men: a comparison between respondent-driven sampling and time-location sampling. Sex Transm Infect. 2012; 88:622–626. [PubMed: 22750886]

21. He Q, Wang Y, Li Y, Zhang Y, Lin P, Yang F, et al. Accessing men who have sex with men through long-chain referral recruitment, Guangzhou, China. AIDS Behav. 2008; 12(4 Suppl):S93–S96. [PubMed: 18389358]

22. Burt RD, Hagan H, Sabin K, Thiede H. Evaluating respondent-driven sampling in a major metropolitan area: Comparing injection drug users in the 2005 Seattle area national HIV behavioral surveillance system survey with participants in the RAVEN and Kiwi studies. Ann Epidemiol. 2010; 20(2):159–167. [PubMed: 20123167]

23. Paz-Bailey G, Jacobson JO, Hernandez FM, Nieto AI, Estrada M, Creswell J. 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; 87:279–282. [PubMed: 21385892]

24. Goel S, Salganik MJ. Assessing respondent-driven sampling. Proc Natl Acad Sci U S A. 2010; 107(15):6743–6747. [PubMed: 20351258]

25. Lu X, Bengtsson L, Britton T, Camitz M, Kim VJ, Thorson A, et al. The sensitivity of respondent-driven sampling method. J Royal Stat Soc: Series A. 2011; 175(1):191–216.

26. Ma X, Zhang Q, He X, Sun W, Yue H, Chen S, et al. Trends in prevalence of HIV, syphilis, hepatitis C, hepatitis B, and sexual risk behavior among men who have sex with men. Results of 3 consecutive respondent-driven sampling surveys in Beijing, 2004 through 2006. J Acquir Immune Defic Syndr. 2007; 45(5):581–587. [PubMed: 17577125]

27. Zhong F, Lin P, Xu H, Wang Y, Wang M, He Q, et al. Possible Increase in HIV and Syphilis Prevalence Among Men Who Have Sex with Men in Guangzhou, China: Results from a Respondent-Driven Sampling Survey. AIDS Behav. 2011; 15(5):1058–1066. [PubMed: 19826942]

28. Ruan R, Yang H, Zhu Y, Wang M, Ma Y, Zhao J, et al. Rising HIV prevalence among married and unmarried among men who have sex with men: Jinan, China. AIDS Behav. 2019; 13:671–676. [PubMed: 19440833]

29. Burt RD, Thiede H. Evaluating consistency in repeat surveys of injection drug users recruited by respondent-driven sampling in the Seattle area: Results from the NHBS-IDU1 and NHBS-IDU2 surveys. Ann Epidemiol. 2012; 22(5):354–363. [PubMed: 22420929]

30. Townsend L, Johnston LG, Flisher AJ, Mathews C, Zembe Y. Effectiveness of respondent-driven sampling to recruit high risk heterosexual men who have multiple female sexual partners: differences in HIV prevalence and sexual risk behaviors measured at two time points. AIDS Behav. 2010; 14(6):1330–1339. [PubMed: 20625926]

31. Gallagher K, Sullivan PS, Lansky A, Onorato I. Behavioral surveillance among people at risk for HIV infection in the U.S: The National HIV Behavioral Surveillance System. Public Health Rep. 2007; 122:32–38. [PubMed: 17354525]

32. Respondentdrivensampling.org. 2013. Available at: http://www.respondentdrivensampling.org

33. SPSS. version 20. SPSS; Chicago, IL: 2011.

34. Harris NV, Thiede H, McGough JP, Gordon D. Risk factors for HIV infection among injection drug users: results of blinded surveys in drug treatment centers, King County, Washington 1988–1991. J Acquir Immune Defic Syndr. 1993; 6(11):1275–1282. [PubMed: 8229659]

35. Burt RD, Thiede H. Evidence for risk reduction among amphetamine-injecting men who have sex with men: Results from the National HIV Behavioral Surveillance surveys in the Seattle area 2008–2012. AIDS Behav. published on-line April 12, 2014 2014.

36. Thiede H, Burt RD. Results from the Kiwi Study: HIV and Hepatitis C prevalence and risk behaviors in recently arrested injection drug users in King County. Washington State/Seattle-King County HIV/AIDS Epidemiology Report 1st Half. 2004; 2003:25–35. Available at: http://www.kingcounty.gov/healthservices/health/communicable/hiv/epi/reports.aspx.

37. Burt RD, Oster AM, Golden MR, Thiede H. Comparing study populations of men who have sex with men: Evaluating consistency within repeat studies and across studies in the Seattle area using different recruitment methodologies. AIDS Behav. 2014; 18:S370–S381.

38. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. J Urban Health. 2006; 83(6 Suppl):i98–112. [PubMed: 16937083]

39. Wejnert C, Pham H, Krishna N, Le B, DiNenno E. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. AIDS Behav. 2012; 16(4):797–806. [PubMed: 22350828]

40. Wejnert C, Heckathorn DD. Web-based network sampling efficiency and efficacy of respondent-driven sampling for online research. Sociological Methods & Research. 2008; 37(134)

41. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS. 2005; 19 (Suppl 2):S67–S72. [PubMed: 15930843]

42. Semaan S, Lauby J, Liebman J. Street and network sampling in evaluation studies of HIV risk-reduction interventions. AIDS Rev. 2002; 4(4):213–223. [PubMed: 12555695]
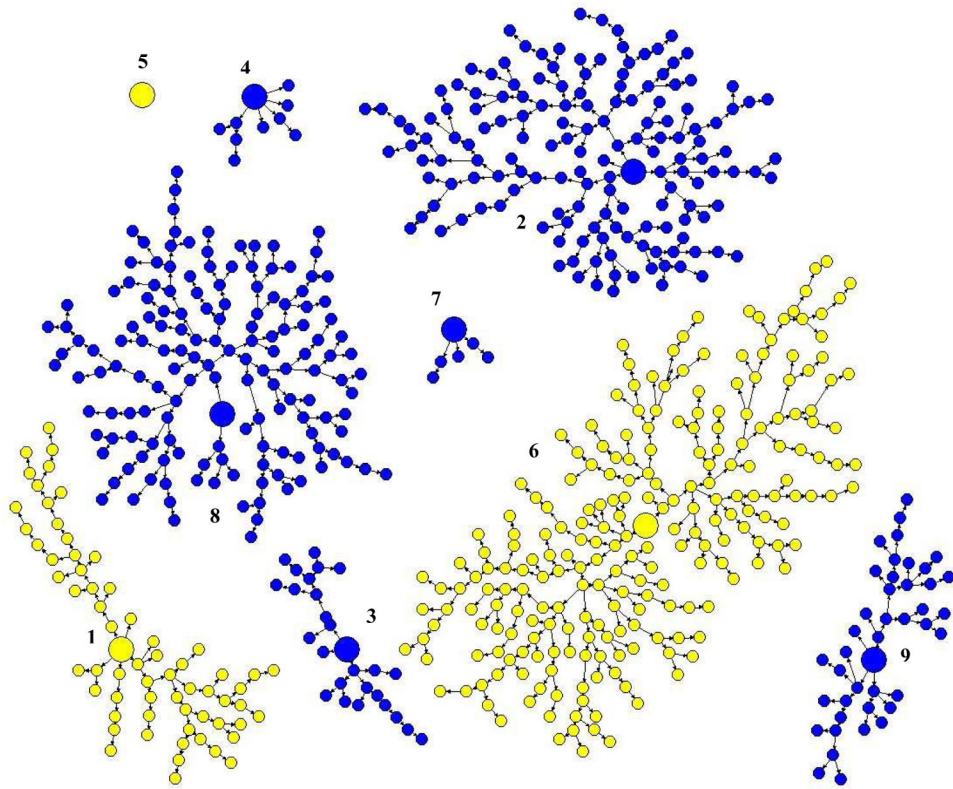
Figure 1.
Random groups defined by recruitment chains among Seattle-area participants in the 2012 NBHS IDU survey, First randomization: Yellow = Group 1; Blue = Group 2: Enlarged circles = seeds; Numbers = Recruitment chain

Table 1

Details of the procedure for randomization by recruitment chain, Among Seattle-area participants in the 2012 NBHS IDU survey: First randomization

	Recruitment chain	Random number	N participants	Cumulative number: group 1	Cumulative number: group 2	Difference group1:group2
Group 1	6	.05	229	229	459	230
	1	.18	65	294	394	100
	5	.19	1	295	393	98
Group 2	2	.32	156	451	237	214
	8	.39	152	603	85	518
	4	.54	10	613	75	538
	3	.58	27	640	48	592
	7	.66	6	646	42	604
	9	.79	42	688	0	688

Table 2

Comparing sociodemographics and drug-related variables between groups randomized by recruitment chain among Seattle-area participants in the 2012 NHBS-IDU survey: First randomization

	Group 1 (Chains 1,5,6)		Group 2 (Chains 2,3,4,7,8,9)		p-value[a,b]
	RDS-adjusted estimate	95% C.I.	RDS-adjusted estimate	95% C.I.	
	(Percentages)				
Age					
18 – 29	17	(8 – 28)	10	(6 – 15)	.01
30 – 39	25	(17 – 33)	21	(15 – 27)	
40 – 49	22	(15 – 30)	29	(21 – 38)	
50	36	(25 – 49)	41	(32 – 50)	
Race					
White	60	(47 – 70)	47	(36 – 57)	3×10^{-4}
Black	22	(10 – 34)	28	(19 – 38)	
Hispanic	5	(3 – 8)	10	(5 – 15)	
Other race	2	(1 – 4)	6	(2 – 10)	
Multiple races	12	(8 – 18)	10	(6 – 15)	
Sex					
Male	52	(41 – 63)	67	(59 – 75)	6×10^{-5}
Female	48	(37 – 59)	33	(26 – 41)	
Area of residence					
North Seattle	20	(12 – 28)	17	(11 – 25)	.001
Downtown Seattle	32	(22 – 43)	44	(33 – 54)	
South Seattle	35	(23 – 47)	22	(16 – 30)	
South King County	11	(3 – 21)	14	(8 – 21)	
East King County	3	(1 – 7)	3	(1 – 6)	
Education					
< High school grad.	26	(17 – 37)	23	(17 – 29)	.59
High school grad.	40	(30 – 49)	41	(32 – 49)	
Post high school	34	(25 – 44)	37	(29 – 45)	
Primary injection drug					
Heroin	86	(78 – 92)	78	(67 – 87)	.007

	Group 1 (Chains 1,5,6)		Group 2 (Chains 2,3,4,7,8,9)		
	RDS-adjusted estimate	95% C.I.	RDS-adjusted estimate	95% C.I.	p-value[a,b]
Speedballs	1	(1 – 3)	5	(3 – 8)	
Cocaine	3	(0.3 – 6)	2	(0.1 – 5)	
Amphetamines	10	(4 – 18)	15	(6 – 27)	
Shared Needle[c]					
No	71	(62 – 78)	76	(70 – 82)	.11
Yes	29	(22 – 38)	24	(18 – 30)	
Shared Cooker[c]					
No	54	(43 – 63)	49	(40 – 58)	.20
Yes	46	(37 – 57)	51	(42 – 60)	
Backloaded[c]					
No	77	(71 – 84)	77	(71 – 83)	.66
Yes	23	(16 – 30)	23	(17 – 29)	
N (total)	295		393		

[a] p-values derive from univariate logistic regression models incorporating RDSAT-generated individual weights.

[b] There was no variable category in the first randomization for which the 95% confidence intervals of the two groups did not overlap.

[c] Previous 12 months.

Table 3

Comparing sexual, HIV and HCV - related variables between groups randomized by recruitment chain, among Seattle-area participants in the 2012 NHBS-IDU survey: First randomization

	Group 1 (Chains 1,5,6)		Group 2 (Chains 2,3,4,7,8,9)		p-value
	RDS- adjusted estimate	95% C.I.	RDS- adjusted estimate	95% C.I.	
	(Percentages)				
Male-to-male sex[a]					
No	91	(87 – 96)	86	(75 – 95)	.03
Yes	9	(4 – 13)	14	(6 – 25)	
Number sex partners[a]					.70
0	20	(10 – 30)	17	(11 – 22)	
1	39	(30 – 50)	41	(33 – 50)	
2 – 4	28	(20 – 35)	30	(23 – 37)	
5 – 9	8	(5 – 13)	6	(3 – 10)	
10 +	6	(3 – 9)	6	(3 – 9)	
Unprotected, non-concordant sex[a,b]					.25
No	73	(65 – 81)	77	(71 – 83)	
Yes	27	(20 – 35)	23	(17 – 29)	
HIV test, 12 months					.001
No	49	(38 – 60)	62	(55 – 71)	
Yes	51	(41 – 62)	38	(29 – 45)	
HIV test, 3 months					.68
No	83	(75 – 88)	84	(78 – 90)	
Yes	17	(12 – 25)	16	(10 – 22)	
Serologic HIV status					2×10^{-10}
Negative	96	(92 – 99.6)	82	(69 – 94)	
Positive	4	(0.4 – 8)	18	(7 – 31)	
Serologic HCV status					.47
Negative	29	(20 – 38)	32	(23 – 40)	
Positive	71	(62 – 80)	69	(60 – 77)	
N (total)	295		393		

[a] Previous 12 months

[b] Vaginal or anal sex without a condom with a partner of unknown HIV status or a status opposite to that of the participant.

Table 4

P-values comparing differences between groups randomized by recruitment chain, among Seattle-area participants in the 2012 NHBS-IDU survey: comparison across five randomizations

Variable	Randomization				
	1	**2**	**3**	**4**	**5**
Age	.01	.004	.04	.81	.94
Race	3×10^{-4}	5×10^{-9}	7×10^{-9}	.36	.34
Sex	6×10^{-5}	2×10^{-4}	.002	.85	.61
Area of residence	.001	.08	8×10^{-4}	.004	.004
Education	.59	.38	.19	.76	.63
Drug most frequently injected	.007	.003	.04	.01	.51
Shared needle	.11	.60	.03	.004	.07
Shared cooker	.20	.006	.90	.001	.06
Backloaded	.66	.30	.23	.06	.08
Male-to-male sex	.03	.26	3×10^{-5}	5×10^{-4}	.80
Number sex partners	.70	.43	.03	.48	.12
Unprotected, non-concordant sex	.25	.82	.009	.76	.19
HIV test, 12 months	8×10^{-4}	.008	5×10^{-4}	.01	.005
HIV test, 3 months	.68	.64	.69	.10	.09
HIV status	2×10^{-10}	.007	$<10^{-16}$	.35	3×10^{-6}
HCV status	.47	.77	.52	.45	.83
Recruitment chains: Group 1	1,5,6	2,8,9	2,3,6	4,6,7,8	6,8
Recruitment chains: Group 2	2,3,4,5,8,9	1,3,4,5,6,7	1,4,5,7,8,9	1,2,3,5,9	1,2,3,4,5,7,9
N (group1):N (group2)	295:393	350:338	412:276	397:291	381:307
# p .001[a]	5	2	5	2	1
# Non-overlapping 95% confidence intervals[b]	0	2	3	0	0

[a] Among 16 variables.

[b] Number of categories (among 46 total) in which the group 1 95% confidence intervals do not overlap the group 2 95% confidence intervals.

Table 5

Variable categories with non-overlapping confidence intervals in studies comparing RDS study populations at two points in time

Study	First survey 95% C.I[a]	Second survey 95% C.I	Non- overlapping 95% C.I. [b]
	(Percentages)		
Townsend (Cape Town) [26]	**2006**	**2008**	9/26
Employed, no	26.5 (20.8–34.1)	49.3 (40.3–56.5)	
Employed, yes	73.5 (65.9–79.2)	50.7 (43.5–59.7)	
Sexual partners, 2–6	66.6 (59.2–71.9)	89.2 (84.7–93.2)	
Sexual partners, >6	33.4 (28.1–40.8)	10.8 (6.8–15.3)	
Condom with main partner, always	11.0 (7.1–16.0)	23.0 (16.8–30.8)	
Condom with non-regular partner, sometimes	18.4 (13.9–25.3)	34.7 (27.7–42.2)	
Condom with non-regular partner, never	32.2 (26.4–38.7)	13.4 (8.3–18.9)	
Alcohol, 0–4 drinks	23.2 (17.3–29.3)	42.8 (35.0–50.2)	
Alcohol, 5+ drinks	76.8 (70.7–82.7)	57.2 (49.8–65.0)	
Ruan (Jinan) [24]	**2007**	**2008**	12/36
Married	17.5 (12.5–22.1)	28.5 (22.7–34.5)	
Lives with spouse	8.5 (5.0–11.9)	20.2 (15.6–25.0)	
Sexual identity, homosexual	65.4 (59.4–71.5)	46.4 (41.4–51.6)	
Sexual identity, bisexual	26.5 (21.0–32.3)	41.5 (36.3–47.0)	
Sexual identity, questioning	5.5 (3.4–7.6)	12.1 (8.4–15.5)	
Education, jr. high	51.7 (45.2–58.4)	28.4 (26.3–33.8)	
Education, technical school	17.5 (12.5–23.1)	47.8 (41.6–53.8)	
HIV+, among unmarried	0.5 (0–1.1)	3.4 (1.3–5.8)	
HIV+, among married	0	2.7 (0.2–5.7)	
UAI, with male, among unmarried	66.1 (60.2–71.5)	36.8 (32.0–41.8)	
Sex with female, among married	97.7 (93.9–100.0)	78.1 (71.3–83.6)	
Ever illicit drugs, among unmarried	2.4 (0.5–4.6)	0.1 (0–0.1)	
He, [16] **Zhong,** [23] **(Guangzhou)**	**He, 2006**	**Zhong 2008**	3/8
Syphilis	3.8 (2.2–6.7)	17.5 (13.6–21.5)	

Study	First survey 95% C.I[a]	Second survey 95% C.I	Non- overlapping 95% C.I. [b]
Official Guangzhou resident	13.9 (9.6–17.7)	21.3 (17.9–26.8)	
Tested for HIV	7.1 (4.4–9.7)	14.3 (10.7–18.5)	
Ma (Beijing) [22,c]	2005	2006	
Single	75.3 (69.4–80.4)	64.3 (58.3–68.8)	4/30
Married	17.9 (13.3–22.8)	29.2 (24.8–34.4)	
Receptive UAI	55.1 (48.7–61.3)	42.1 (36.7–47.5)	
HCV-positive	1.3 (0.5–2.1)	5.2 (2.3–8.2)	

[a]C.I. = Confidence Interval.

[b]Number of categories with non-overlapping confidence intervals/ total number of variable categories.

[c]A further comparison in the data of Ma et al. between surveys in 2004 and 2006 showed difference in 14 of 26 variable categories but interpretation of this data is complicated by somewhat differing survey methods in the 2004 survey: there was only one seed and for a third of the study population network size data was not directly collected but rather imputed.