Published in final edited form as:
Ann Epidemiol. 2011 April ; 21(4): 280–289. doi:10.1016/j.annepidem.2010.11.007.

Sub-populations of illicit drug users reached by targeted street outreach and respondent driven sampling strategies: Implications for research and public health practice

Abby E Rudolph, MPH,
Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland

Natalie D Crawford, MPH,
Department of Epidemiology, Columbia University Mailman School of Public Health, New York

Carl Latkin, PhD,
Department of Health, Behavior, and Society, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland

Robert Heimer, PhD,
Center for Interdisciplinary Research on AIDS, Yale School of Public Health, New Haven, Connecticut

Ebele O Benjamin, MPH,
The Center for Urban Epidemiologic Studies, The New York Academy of Medicine, New York City, New York

Kandice C Jones, MPH, and
The Center for Urban Epidemiologic Studies, The New York Academy of Medicine, New York City, New York

Crystal M Fuller, PhD
Department of Epidemiology, Columbia University Mailman School of Public Health, New York; The Center for Urban Epidemiologic Studies, The New York Academy of Medicine, New York City, New York

Abstract

Purpose—To determine whether illicit drug users recruited through Respondent-Driven Sampling (RDS) and targeted street outreach (TSO) differ by comparing two samples recruited concurrently with respect to sample selection and potential recruitment biases.

Methods—217 heroin, crack, and cocaine users aged 18-40 were recruited through TSO in New York City (2006-2009). 46 RDS seeds were recruited similarly and concurrently yielding a

Correspondence: Abby Rudolph, Mailing Address: 3422 Keswick Rd, Baltimore MD 21211, Abby.Rudolph@gmail.com, Phone: 267-261-6395.

Conflict of interest: none declared.

This work originated from the Department of Epidemiology at Columbia University Mailman School of Public Health and the Center for Urban Epidemiologic Studies at The New York Academy of Medicine

maximum of 14 recruitment waves and 357 peer recruits. Baseline questionnaires ascertained sociodemographic, drug use, and drug network characteristics. Descriptive statistics and log-binomial regression were used to compare RDS and TSO samples.

Results—RDS recruits were more likely to be male (prevalence ratio [PR]:1.28), Hispanic (PR: 1.45), black (PR:1.58), older (PR:1.02), homeless (PR:1.19), and crack users (PR:1.37). RDS recruited fewer injectors (PR:0.35) and heroin users (PR:0.74). Among injectors, RDS recruits injected less frequently (PR:0.77) and were less likely to use Needle Exchange Programs (PR: 0.35).

Conclusion—These data suggest that RDS and TSO strategies reach different sub-groups of drug users. Understanding the differing capabilities of each recruitment strategy will enable researchers and public health practitioners to select an appropriate recruitment tool for future research and public health practice.

Keywords

Respondent Driven Sampling; Targeted Street Outreach; HIV; illicit drug users; sampling bias

---

Generating a representative sample of illicit drug users for substance abuse and HIV research can be challenging because there is no sampling frame. While convenience, targeted, snowball and time-location sampling methods are used to recruit this population, volunteer and masking biases may result from the inability to sample randomly from the target population. Respondent-Driven Sampling (RDS) aims to reduce these biases with a modified form of chain-referral sampling that regulates peer recruitment and uses probability weights to offset nonrandom recruitment.

RDS has recruited geographically(1-3) and demographically diverse samples of injection drug users (IDUs).(4-9) Like samples recruited with other strategies, respondent-driven samples include IDUs from both institutional settings and public venues. RDS also recruits those missed by traditional strategies: those only identified through social networking approaches or through their participation in activities that define the hidden population.(10)

Studies comparing respondent-driven samples with chain referral(11), targeted(12,13), traditional outreach(1,2), snowball (14) and time-location samples(3,14,15) have reported differences in demographic characteristics(1,11,12,14), but not in HIV risk behaviors(11). Results from comparison studies should be interpreted with caution, because comparison samples were often recruited over different enrollment periods(14-16) and from different geographic locations.(17) For example, in a study comparing samples recruited with time-location sampling (2002), snowball sampling (2002), and RDS (2005) (14), the respondent-driven sample was considered to be the most representative because it's social class structure most closely resembled that of the AIDS cases among men who had sex with men in Brazil.(14) Sample differences were attributed to differing recruitment strategies, but changes in the target population over the three years separating the studies were not considered. Another study attributed differences in HIV seroprevalence estimates from respondent-driven (2006) and time-location (2003) samples(15) to the recruitment strategy, but did not discuss the role that temporal trends may have played in the observed discrepancy.(15) In a third study, the authors concluded that respondent-driven samples were more generalizable than earlier samples recruited using outreach workers because of a higher female-to-male ratio of drug users in the respondent-driven sample (2006), but again the study did not account for temporal or geographic differences between samples.(17)

Other studies compared respondent-driven samples with simulated/theoretical time-location samples to prove greater representation, but it is unknown how well simulated/theoretical

samples approximate actual samples.(3) In one study, estimates from RDS and simulated time-location samples converged but the respondent-driven sample was more diverse(10) Another study that constructed a time-location sample from an existing respondent-driven sample reported that RDS recruited a more diverse population, but recommended that future studies compare samples recruited contemporaneously but with different strategies.(18) Other studies asserted that RDS recruited hidden populations that were inaccessible through traditional sampling techniques and claim to have recruited a more representative sample because the sample distribution reached a point of equilibrium, or the stable composition that is attained when the sample composition ceases to change between subsequent waves of respondent-driven recruitment.(7,19,20) However, the validity of RDS population-based estimates is dependent on several assumptions which are typically not tested.(18,21,22) Understanding which assumptions are not met can provide information on sample biases.

While these studies suggest that RDS may generate a more representative sample than other population-based sampling methods (e.g., targeted-street outreach), which are subject to biases that cannot be mitigated through statistics further research is needed to 1) compare RDS with other recruitment strategies that take time and geographic location into account and do not rely on comparisons with theoretical samples, 2) validate RDS in a variety of settings(16), and 3) identify sub-populations that can be reached with one recruitment tool but not another. This study compared participants concurrently recruited using RDS and TSO in order to determine whether illicit drug users recruited through these approaches differed with respect to demographics, drug use behaviors, HIV status, and drug network size. While TSO cannot mitigate the biases associated with non-random sampling, the resulting sample is not merely a convenience sample, because it employs a systematic approach to sampling when true random sampling is not feasible.(23) RDS has the potential to account for these biases through a post-stratification process involving weights that account for differences in recruitment and homophily and variations in network size, however this correction process is dependent on a series of assumptions which are often not met.' Therefore, we also evaluated the RDS assumptions and enumerated the potential biases of each recruitment approach.

METHODS

The data were collected as part of Social Ties Associated with Risk of Transition into injection drug use (START), a longitudinal study aiming to identify risk factors for transitioning into injection drug use among young adult injection and non-injection drug users (heroin, crack, and cocaine) in New York City (NYC). Non-injection drug users (NIDUs) were followed prospectively for 18 months and newly initiated IDUs were evaluated cross-sectionally. NIDUs and IDUs were recruited concurrently through TSO and RDS between July 2006 and June 2009.

Targeted Street Outreach

As previously described, economically disadvantaged and racially diverse NYC communities with high rates of HIV infection and overdose mortality were ethnographically mapped and targeted.(24) Outreach recruitment followed a targeted sampling plan, which was developed for HIV prevention studies and has been used to recruit those at increased risk for HIV.(23,25) Of 812 illicit drug users screened, 217 were eligible. Of the 217, 17 were originally approached as RDS seeds but were enrolled as TSO-recruits because they did not agree to recruit peers (an eligibility requirement for RDS seeds but not peer recruits). Because bias could have been introduced by including these 17 in the TSO sample due to differing eligibility criteria (willingness to recruit peers), we conducted a sub-analysis that compared these 17 individuals with the 200 TSO-recruits; significant differences by race existed (5 were Hispanic and 12 were black; P=0.02), but not by gender, education, income,

homelessness, injection status, HIV status, drugs used, the number of drug using networks, or number of sex partners (data not shown). Thus, including these 17 in the TSO sample will make the two recruitment approaches more similar on race/ethnicity and will consequently dilute the association between recruitment approach and race/ethnicity.

RDS

48 seeds were recruited through TSO in Brooklyn, Bronx, Manhattan and Queens (12 from each borough). Each seed was asked to recruit up to three peers, each of whom were asked to recruit three additional peers, and so on until recruitment was administratively ended in June 2009. Of 621 participants screened to participate in the respondent-driven sample, 439 were eligible. Participants were initially interviewed at a Harlem community-based site and on a mobile van, however those enrolled on the van (and their peer recruits) were removed from the analysis because the van's constant relocation made RDS peer referral and follow-up using the van infeasible. An exception was made for participants recruited on the van who had access to the community-based site (removing 32 individuals from the analysis). Of the 407 remaining, two seeds and two peer recruits were dropped due to inconsistent self-report of drug use (N=403; 46 seeds and 357 peer recruits). Removal of study participants resulted in the removal of subsequent peer recruits for this analysis.

Eligibility

Eligible START participants were 18-40 years of age (verified with a photo ID) and active drug users. Eligible IDUs reported injecting heroin, crack or cocaine for ≤4 years and injecting ≥once in the past 6 months; injection drug use was verified by visible track marks. NIDUs reported non-injection use of heroin, crack or cocaine for ≥1 year and used heroin, crack or cocaine 2-3 times per week in the last three months. Self-reported drug use was verified with rapid drug tests which screened for opiate and cocaine metabolites in urine. Those with a negative drug test were not eligible and were compensated for round-trip travel to the research site.

Baseline

Participants provided informed consent and completed a 90-minute interviewer-administered questionnaire approved by both Columbia University and the New York Academy of Medicine institutional review boards. This instrument ascertained demographic and social contextual characteristics, social and behavior characteristics, parental drug use, depression, conduct disorder, suicidal attempt/ideation and victimization. Participants received $30 and a round-trip Metrocard for completing the questionnaire.

RDS Procedures

RDS participants received 1) three RDS coupons to recruit drug-using peers to participate in START, 2) an individual recruitment training (IRT) to emphasize the importance of peer recruitment and provide recruiting tips, and 3) an invitation to attend up to two group-facilitated peer recruitment training sessions (RDSTs). Participants speaking only Spanish received an extended IRT because there were too few Spanish-speaking participants to conduct RDSTs in Spanish. Those attending RDSTs received $20 and a round-trip Metrocard after completing a post-session survey that asked about their experiences with peer recruitment and for feedback regarding the session.

RDS participants received $10 for each eligible peer-recruit (up to 3) and a $10 bonus if 3 eligible peers were recruited. RDS participants recruiting ineligible peers who met the criteria specified on the recruitment coupon received $5 and a round-trip Metrocard. For each ineligible recruit, seeds received a new coupon. While seeds were given unlimited

coupons to recruit 3 eligible peer networks, peer networks were given five attempts to recruit 3 eligible peers.

RDS coupons had unique 9-digit numbers linking each participant to 1) the seed initiating the recruitment chain, 2) the individual recruiting him/her, and 3) his/her peer recruits. It also tracked the number of recruitment waves. RDS participants were encouraged to give coupons to drug users they knew and not to strangers. Reverse identifiers (e.g., mother's maiden name, last 4 digits of social security number, birth date) were also collected to prevent participants from enrolling multiple times or with false identities. Peer-recruits presenting without a coupon were asked to provide the study ID or full name of the study participant who referred him/her before they were screened for eligibility.

Data analysis

Recruitment matrices, homophily indexes, and equilibrium statistics for a set of pre-specified sociodemographic characteristics were calculated using RDS Analysis Tool (RDSAT) Version 5.6.(26) Recruitment homophily, or the propensity for people to recruit others with similar characteristics (e.g. age, gender, race/ethnicity) is measured as an index ranging from negative one to one and it influences the number of waves that are required to attain equilibrium.(4,5) It is positive when people are more likely to recruit others with similar characteristics, negative when individuals are more likely to recruit peers with different characteristics than themselves, and zero when individuals recruit peers randomly. (6) Equilibrium distributions were set to fall within 2% of the sample distribution.

Estimated population proportions were obtained by weighting the data based on participants' recruitment patterns (who recruited whom) and the size of their drug-using network (total number of people in his/her network who use drugs, including those who he/she does and does not use drugs with). As several individuals who were recruited by eligible peers and/or who recruited eligible peers reported zero drug-using network members, self-reported drug-using network sizes were post-fit to account for misreporting. Thus, the number of self-reported drug-using network members for those who were recruited by an eligible peer was corrected to be at least one. Similarly, the corrected number of drug-using network members for those who were recruited by a peer and who had recruited one, two, or three eligible peers was corrected so that at least 2, 3, or 4 drug-using network members were reported, respectively. While the main objective of this procedure was to correct the values for those reporting zero drug-using network members, it is possible that the corrected values for others may have been over-estimated by this procedure. Unpublished data from Heimer and colleagues suggest that the square root of this corrected value provides the least biased estimate; those with corrected drug-using network sizes of one will remain the same and those most likely to be inflated will be slightly attenuated.

T-tests and chi-square statistics were used to assess non-random sampling from within networks and the independence of seeds and recruits (SAS v9.2)(27)(Table 4). RDS- and TSO-recruited participants were compared using descriptive statistics and log-binomial regression using SAS v9.2 (27). As this analysis aims to compare the groups of individuals recruited via RDS and TSO and not the representativeness of the samples recruited, RDS weights were not applied. Because RDS seeds who did not recruit eligible peers did not contribute to the final RDS sample, they were excluded from the analysis comparing RDS and TSO.

RESULTS

Respondent-driven recruits

Forty-six seeds (28 of whom recruited eligible peers) and a maximum of 14 recruitment waves produced 357 peer-recruits. Two seeds, each extending ≥13 waves recruited over half the peer-recruits (n=203). Five seeds (extending ≥6 waves each) recruited 255 individuals and 311 individuals were recruited by 10 seeds with recruitment waves extending ≥4 waves. 18 seeds did not recruit any eligible peers. The majority of the sample was NIDU (91%), male 75%, and Black (54%) or Hispanic (37%) with a median age of 34 years (Table 1). The average number of drug-using network members was 1.52 (post-fit=1.57), 36% had ≥ a high school degree or GED, 85% had an annual income ≤$10,000, and 10% were HIV positive. In the past 6 months, 70% were homeless, 47% used heroin, 78% used cocaine, and 84% used crack (Table 1).

All variables examined reached equilibrium and there were no significant differences between the RDS-weighted population estimates and sample estimates for any of these variables (Table 2). There were no major differences in homophily or drug-using network size by any variables considered and the weights corresponding with each of these characteristics were low, which explains why the RDS-adjusted population estimates were not significantly different from the sample compositions.

As seen in Table 3, several RDS assumptions were not met: 1) reciprocal recruitment ties, 2) random recruitment from within peer networks, 3) accurately self-reported degree weights, 4) independence of seeds and peer recruits (Table 4), random recruitment from within peer networks, and (Table 4) and 5) one recruit per respondent.

TSO recruits

Among TSO recruits, the median age was 32, the average number of drug-using network members was 1.28, 62% were male, 36% Hispanic, 43% Black, 34% had ≥ a high school degree or GED, 81% had an annual income ≤$10,000, 57% were NIDUs, 43% were IDUs, and 8% were HIV positive. In the past 6 months, 58% were homeless, 68% used heroin, 74% used cocaine, and 72% used crack. Among IDUs, 23% never used needle exchange programs (NEPs) (Table 5).

Differences by recruitment strategy

Table 5 depicts differences in the RDS and TSO samples. There were significant differences with respect to gender, race/ethnicity, age, homelessness, injection status, heroin use and crack use. Compared with TSO-recruits, RDS-recruits were significantly more likely to be male [Prevalence Ratio ((PR):1.28; [95% Confidence Interval (95%CI):1.10, 1.49)]], Hispanic (PR:1.45; 95%CI:1.11, 1.88), black (PR:1.58; 95%CI:1.23, 2.03), older (PR:1.02; 95%CI:1.01, 1.03), homeless (PR:1.19; 95%CI:1.04, 1.36), and to use crack (PR:1.37; 95%CI:1.13, 1.67). RDS-recruits were less likely to inject (PR:0.35; 95%CI:0.26, 0.47) and to use heroin (PR:0.74; 95%CI:0.66, 0.83).

Among IDUs, RDS recruits injected les frequently (PR:0.77; 95%CI:0.72, 0.83) and were less likely to have ever used NEPs (PR:0.35; 95%CI:0.19, 0.62) than TSO-recruits.

DISCUSSION

RDS and TSO recruited slightly different populations and both had limitations. In this study, RDS captured more males, racial/ethnic minorities, and individuals who were older, homeless, and crack users. TSO recruited more IDUs and heroin users. TSO-recruited IDUs

injected more frequently and were more likely to have used NEPs. Although including the 17 individuals who were originally recruited as RDS seeds in the TSO sample attenuated the association between race/ethnicity and recruitment approach slightly, the two recruitment approaches were still significantly different by race/ethnicity. Understanding the differences in the sub-populations of drug users recruited with these tools has important implications for both research and practice.

For example, studies that target a higher risk injecting population might benefit more from a targeted approach, which was better able to capture IDUs who injected frequently and used NEPs. However, RDS may be more appropriate for studies targeting more mobile or minority populations of drug users. RDS may also be better for studies planning to enroll both IDUs and NIDUs and studies that are interested in examining both the injection and sexual risk behaviors associated with drug use.

Both the RDS and TSO samples were limited in their ability to recruit drug using individuals who perceived the risks of being involved in a study about drug use to outweigh the potential benefits (non-response bias). As demonstrated in Table 5, both samples represent a predominately lower income population. Thus it is likely that neither of these approaches will generate a sample that accurately represents the entire scope of drug use within a population. Future studies might consider offering coupons specifically for participants to recruit more isolated or less connected network members, purposefully selecting a few seeds from these under-represented groups, or allowing chains to terminate naturally to increase the probability of reaching these individuals.

Since sensitive information was asked of all participants, it is possible that both samples are biased by social desirability. However, we did observe a number of high risk behaviors and the proportions of individuals reporting these behaviors are consistent with previous studies in this population, which suggests that this bias may be minimal.

RDS also has unique biases that result when model assumptions do not hold. For example, not all RDS recruits reported knowing their recruiter, which is a violation of the reciprocity assumption. High homophily, with respect to HIV status suggests that peer recruitment techniques may work well for recruiting other HIV positives, but that to generate an unbiased estimate of HIV prevalence in the target population, a large sample size that allows for many waves of recruitment is required. It is unknown whether the homophily observed with respect to HIV status can be attributed to knowing other HIV positives through HIV support groups, or whether it reflects a homophily with respect to high risk sex and injecting practices. There were also inconsistencies between the characteristics of the population recruited and those of the theoretical population based on the characteristics of drug-using network members. However, it is unknown how well the theoretical population approximates the target population, especially given the large number of ineligible respondents. Biases may have also been introduced by inaccurate degree weights, as is suggested by the relatively low estimates compared to other studies in the field. Finally, even though equilibrium was reached for all of the characteristics examined, there were some significant differences between seeds and peer recruits.

Future RDS studies should carefully select seeds that represent the diversity of the target population and should include some members who are often not reached (e.g., wealthier drug users). Additionally, those using RDS to generate valid HIV prevalence estimates should consider collecting information on potential biases in order to correct estimates to more accurately reflect the truth. For example, comparing network attributes with those of peer recruits can provide information about how the sample misrepresents the source population and weights can be applied to account for non-random selection. In addition,

quantifying degree weight biases could be used to adjust the prevalence estimates or to create a more accurate measure of uncertainty. In our sample, degree estimates were lower than the number of drug-using network members reported in comparable studies. Some of these differences may reflect the fact that this analysis was conducted among IDUs and NIDUs, while the studies conducted by Weeks and colleages (2002) and Latkin and colleagues (1995) were among IDUs, who may have larger drug-using networks. However, a study conducted among NIDUs by Pilowsky and colleagues (2007) reported a median total network size of 5 and a median drug-using network size of 2(14), which is also higher than what we reported. It is possible that the interviewers did not sufficiently probe for additional networks or that the duration of the survey led individuals to under-report their social network size, as a series of questions about each social network member reported followed. Because RDS assumes that recruitment is proportional to degree size the relative degree size is more important than the absolute degree size, so it is unknown how much this error in reporting would bias population estimates. Additional information on how degree estimates are biased could be used to weight RDS statistics so that more accurate HIV prevalence estimates can be attained.

Since TSO does not sample randomly from the target populations, it likely results in a biased sample which potentially lacks external validity. It is commonly understood in substance abuse and HIV prevention research that TSO is used to target a group of at risk individuals and that those sampled may not be representative of the target population, but rather a select group of individuals (typically high risk) who the researchers wish to study

With limitations acknowledged, this is one of few studies that compare recruitment strategies in a controlled setting and carefully examine the limitations of each. Future studies are needed to compare RDS with other recruitment tools and in a variety of populations to better understand when RDS is a more appropriate recruitment approach than others and when RDS is less likely to produce biased estimates. Since most drug using populations are connected by social ties created through buying or sharing drugs, RDS has the potential to reach members of the target population not reached with other sampling strategies. However, RDS estimators are not unbiased when model assumptions are not met. Thus, there are some situations where the breakdown of assumptions in RDS could result in a sample with more biases than other more traditional sampling approaches that assume random sampling. On the contrary, in terms of recruiting a particular population to target with public health messages or an intervention, there may be situations where RDS (even if the assumptions are not entirely met), is more appropriate. More formative work is needed to understand the population structure of those being targeted for recruitment and researchers must carefully examine their desired target population and the potential biases of each recruitment strategy before selecting a recruitment tool.

Acknowledgments

This research was supported by the National Institute on Drug Abuse (R01 DA019964-01, T32 DA007282).

The authors thank Dr. Lisa Jacobson, Dr. Frank Sifakis, Dr. David Celentano, and the START research team for their valuable contributions.

Abbreviations

CI	Confidence Interval
HIV	Human Immunodeficiency Virus
IRT	Individual Recruitment Training

IDU	Injection Drug User
IQR	Interquartile Range
NEP	Needle exchange program
NIDU	Non-injection Drug User
PR	Prevalence Ratio
RDS	Respondent Driven Sampling
RDST	Respondent Driven Sampling Training
START	Social Ties Associated with Risk of Transition into injection drug use
TSO	Targeted Street Outreach

References

1. Broadhead RS, Heckathorn DD, Weakliem DL, Anthony DL, Madray H, Mills RJ, et al. Harnessing peer networks as an instrument for AIDS prevention: results from a peer-driven intervention. Public Health Rep. 1998; 113(Suppl 1):42–57. [PubMed: 9722809]
2. Heckathorn D, Broadhead R, Anthony D, Weakliem D. AIDS and Social Networks: HIV Prevention Through Network Mobilization. Sociological Focus. 1999; 32(2):159–179.
3. Yeka W, Maibani-Michie G, Prybylski D, Colby D. Application of respondent driven sampling to collect baseline data on FSWs and MSM for HIV risk reduction interventions in two urban centres in Papua New Guinea. J Urban Health. 2006; 83(6 Suppl):i60–72. [PubMed: 17066328]
4. Heckathorn D. Respondent-Driven Sampling: A new approach to the study of hidden populations. Social Problems. 1997; 44(2):174–199.
5. Heckathorn D. Respondent-Driven Sampling II. Deriving valid population estimates from chain-referral samples of hidden populations. Social Problems. 2002; 49(1):11–34.
6. Abdul-Quader AS, Heckathorn DD, McKnight C, Bramson H, Nemeth C, Sabin K, et al. Effectiveness of respondent-driven sampling for recruiting drug users in New York City: findings from a pilot study. J Urban Health. 2006; 83(3):459–76. [PubMed: 16739048]
7. Heckathorn D, Semann S, Broadhead R, Hughes J. Extensions of Respondent-Driven Sampling: A new approach to the study of injection drug users aged 18-25. AIDS and Behavior. 2002; 6(1):55–67.
8. 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(4 Suppl):S131–41. [PubMed: 18535901]
9. 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]
10. Ramirez-Valles J, Heckathorn DD, Vazquez R, Diaz RM, Campbell RT. From networks to populations: the development and application of respondent-driven sampling among IDUs and Latino gay men. AIDS Behav. 2005; 9(4):387–402. [PubMed: 16235135]
11. Platt L, Wall M, Rhodes T, Judd A, Hickman M, Johnston LG, et al. Methods to recruit hard-to-reach groups: comparing two chain referral sampling methods of recruiting injecting drug users across nine studies in Russia and Estonia. J Urban Health. 2006; 83(6 Suppl):i39–53. [PubMed: 17096189]
12. 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):i29–38. [PubMed: 16933101]
13. Lansky A, Abdul-Quader AS, Cribbin M, Hall T, Finlayson TJ, Garfein RS, et al. Developing an HIV behavioral surveillance system for injecting drug users: the National HIV Behavioral Surveillance System. Public Health Rep. 2007; 122(Suppl 1):48–55. [PubMed: 17354527]

14. Kendall C, Kerr LR, Gondim RC, Werneck GL, Macena RH, Pontes MK, et al. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. AIDS Behav. 2008; 12(4 Suppl):S97–104. [PubMed: 18389357]
15. Frost SD, Brouwer KC, Firestone Cruz MA, Ramos R, Ramos ME, Lozada RM, et al. Respondent-driven sampling of injection drug users in two U.S.-Mexico border cities: recruitment dynamics and impact on estimates of HIV and syphilis prevalence. J Urban Health. 2006; 83(6 Suppl):i83–97. [PubMed: 17072761]
16. 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. 20(2):159–67. [PubMed: 20123167]
17. Stormer A, Tun W, Guli L, Harxhi A, Bodanovskaia Z, Yakovleva A, et al. An analysis of respondent driven sampling with Injection Drug Users (IDU) in Albania and the Russian Federation. J Urban Health. 2006; 83(6 Suppl):i73–82. [PubMed: 17075727]
18. Johnston LG, Sabin K, Mai TH, Pham TH. Assessment of respondent driven sampling for recruiting female sex workers in two Vietnamese cities: reaching the unseen sex worker. J Urban Health. 2006; 83(6 Suppl):i16–28. [PubMed: 17031567]
19. 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):S67–72. [PubMed: 15930843]
20. Johnston LG, Khanam R, Reza M, Khan SI, Banu S, Alam MS, et al. The effectiveness of respondent driven sampling for recruiting males who have sex with males in Dhaka, Bangladesh. AIDS Behav. 2008; 12(2):294–304. [PubMed: 17712620]
21. Wang J, Carlson RG, Falck RS, Siegal HA, Rahman A, Li L. Respondent-driven sampling to recruit MDMA users: a methodological assessment. Drug Alcohol Depend. 2005; 78(2):147–57. [PubMed: 15845318]
22. Gile KJ, Handcock MS. Respondent-driven sampling: An assessment of current methodology. Sociological Methodology. 2010 in press.
23. Watters JK, Biernacki P. Targeted Sampling - Options for the Study of Hidden Populations. Social Problems. 1989; 36(4):416–430.
24. Ompad DC, Galea S, Marshall G, Fuller CM, Weiss L, Beard JR, et al. Sampling and recruitment in multilevel studies among marginalized urban populations: the IMPACT studies. J Urban Health. 2008; 85(2):268–80. [PubMed: 18214686]
25. Weeks MR, Clair S, Borgatti SP, Radda K, Schensul JJ. Social Networks of Drug Users in High-Risk Sites: Finding the Connections. AIDS Behav. 2002; 6(2):193–206.
26. Volz E, Heckathorn DD. Probability Based Estimation Theory for Respondent-Driven Sampling. Journal of Official Statistics. 2008; 24(1):79–97.
27. SAS. SAS Institute Inc.; Cary, NC, USA: 2008. Copyright ©In
28. Pilowsky DJ, Hoover D, Hadden B, Fuller C, Ompad DC, Andrews HF, et al. Impact of social network characteristics on high-risk sexual behaviors among non-injection drug users. Subst Use Misuse. 2007; 42(11):1629–49. [PubMed: 17934988]
29. Heckathorn, D. Extensions of Respondent-Driven Sampling: Analyzing continuous variables and controlling for differential recruitment. Cornell University; 2007. p. 151-208.
30. Liu H, Liu H, Cai Y, Rhodes AG, Hong F. Money boys, HIV risks, and the associations between norms and safer sex: a respondent-driven sampling study in Shenzhen, China. AIDS Behav. 2009; 13(4):652–62. [PubMed: 18841459]
31. Weeks M, Clair S, Borgatti SP, Radda K, Schensul JJ. Social Networks of Drug Users in High-Risk Sites: Finding the Connections. AIDS Behav. 2002; 6(2):193–206. R.
32. Latkin C, Mandell W, Oziemkowska M, Celentano D, Vlahov D, Ensminger M, et al. Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. Drug and Alcohol Dependence. 1995; 38:1–9. [PubMed: 7648991]
33. Weeks M, Clair S, Borgatti SP, Radda K, Schensul JJ. Social Networks of Drug Users in High-Risk Sites: Finding the Connections. AIDS Behav. 2002; 6(2):193–206. R.

Table 1

Socio-demographic and Drug Use Characteristics of Illicit Drug Users Recruited Through RDS in New York City, 2006-2009

	Study Sample* (N=357)	
	N	**%**
Gender		
Male	271	75.1
Female	85	24.9
Race/Ethnicity		
Hispanic	122	36.5
Black	199	54.4
White/Other	36	9.1
Education		
< High School	231	64.3
≥ High School	126	35.7
Income		
≤ 10,000	288	85.2
> 10,000	50	14.8
Age		
18-27	63	18.2
28-30	49	14.3
31-34	70	19.2
35-36	51	14.5
37-38	55	17.5
39-40	69	19.2
Homeless in last 6 months		
No	108	30.2
Yes	249	69.8
Injection status		
NIDU	331	90.6
IDU	26	9.4
HIV status		
Negative	302	90.1
Positive	33	9.9
Heroin use in the past 6 months (with or without crack/cocaine)		
No	188	53.3
Yes	165	46.7
Cocaine use in the past 6 months (with or without heroin)		
No	81	22.5
Yes	274	77.5
Crack use in the past 6 months (with or without heroin)		

	Study Sample* (N=357)	
	N	**%**
No	52	15,8
Yes	304	84.2
Number of drug-using network members (mean, post-fit mean)	1.52	1.57

Abbreviations: HIV, Human Immunodeficiency Virus; RDS, Respondent Driven Sampling

*This does not include seeds

Table 2

Homophily, Cross-group Recruitment, Mean Number of Drug-Using Network Members, and Equilibrium Statistics in a Sample of Illicit Drug Users in New York City

	Homophily	Cross-group Recruitment (%)	Mean Number of Drug-Using Network Members (adjusted)	# waves required to reach equilibrium	Equilibrium sample %	RDS-weighted population estimate	Sample estimate %	2-sided *P*-value (population vs. sample estimate)
Gender								0.60
Female	0.277	55.7	1.333	4	22.6	23.1	24.9	
Male	0.297	16.2	1.371	5	77.4	76.9	75.1	
Race/Ethnicity								0.44
Hispanic	0.384	38.7	1.383	5	37.2	36.6	36.5	
Black	0.327	31.3	1.341	5	52.6	53.4	54.4	
White/Other	0.080	82.9	1.389	4	10.1	9.9	9.1	
Education								0.44
<High School	0.065	32.6	1.353	2	64.8	65.1	64.3	
≥High School	0.079	60.0	1.373	2	35.2	34.9	35.7	
Income								0.48
≤ 10,000	0.037	14.2	1.360	2	85.1	85.3	85.2	
> 10,000	0.049	81.1	1.377	2	14.9	14.7	14.8	
Age								0.11
18-27	0.056		1.506	3	17.8	17.6	18.2	
28-30	0.036		2.285	3	13.7	11.9	14.3	
31-34	0.086		1.468	3	19.6	19.9	19.2	
35-36	0.060		1.423	3	13.9	14.3	14.5	
37-38	0.014		1.472	3	15.5	15.8	17.5	
39-40	0.041		1.426	3	19.4	20.5	19.2	
Homeless in last 6 months								0.72
No	0.170	59.6	1.412	3	29.7	28.6	30.2	
Yes	0.125	25.0	1.339	3	70.3	71.4	69.8	
Injection status								0.35

	Homophily	Cross-group Recruitment (%)	Mean Number of Drug-Using Network Members (adjusted)	# waves required to reach equilibrium	Equilibrium sample %	RDS-weighted population estimate	Sample estimate %	2-sided *P*-value (population vs. sample estimate)
NIDU	0.016	6.5	1.349	2	92.6	93.3	90.6	
IDU	0.123	81.8	1.514	3	7.4	6.7	9.4	
HIV status								0.13
Negative	0.552	3.0	1.368	7	93.2	93.3	90.1	
Positive	0.560	42.5	1.393	12	6.8	6.7	9.9	
Heroin use (past 6 months)								0.59
No	0.348	31.1	1.349	4	51.9	52.3	53.3	
Yes	0.359	33.3	1.373	4	48.1	47.7	46.7	
Cocaine use (past 6 months)								0.43
No	0.178	63.1	1.321	4	22.6	23.3	22.5	
Yes	0.208	18.4	1.372	3	77.4	76.6	77.5	
Crack use (past 6 months)								0.50
No	0.191	69.8	1.415	4	15.2	14.6	15,8	
Yes	0.151	12.4	1.349	3	84.8	85.4	84.2	

Abbreviations: HIV, Human Immunodeficiency Virus; IDU, Injection Drug User; NIDU, Non-injection Drug User; RDS, Respondent Driven Sampling

Table 3

Testing RDS assumptions for population estimation, New York City (2006-2009)

Assumption	Test of Assumption	Conclusions
1) Respondents know one another as members of the target population and recruitment ties are reciprocal(29).	Participants were asked to describe their relationship with the person who recruited him/her.	Most participants described their recruiter as a friend or acquaintance (83%); 6% as a relative, 7% as a stranger, and 4% as other.
2) There is sufficient cross-over between subgroups and networks are dense enough to sustain a chain referral process(30).	This assumption could not be tested because peer recruitment did not terminate naturally, however we did evaluate cross-over recruitment.	There was a tendency to recruit within-group for all variables considered, but cross-group recruitment was substantial for gender (36%), race (39%), income (22%), education (43%), and homelessness (36%), so chains did not become trapped within a single group. Cross-group recruitment was low for injection status (11%) and HIV status (8%). HIV positives comprised 10% of the sample and recruited other HIV positives 59% of the time. HIV negatives comprised 90% of the population and recruited other HIV negatives 97% of the time.
3) Sampling occurs with replacement(29).		This is not a reasonable assumption for well-connected networks because many people share the same networks. Thus as recruitment waves extend, the proportion of networks available for recruitment (not enrolled or previously screened for eligibility) decreases, and this assumption is less likely to be valid.
4) Respondents are recruited from one's network at random(29).	T-tests and Chi-square tests were used to compare peer-recruit characteristics with the characteristics that individuals reported that their networks possessed.	Networks and recruits were not significantly different with respect to race and injection status. There were significant differences with respect to gender, age, education, crack smoking, and heroin snorting. Relative to what would be expected if recruitment reflected the composition of the self-reported personal networks, males were over-recruited by 20%, high school educated individuals were underrepresented by 56%, crack smokers were overrepresented by 11%, and heroin snorters were overrepresented by 101%. The sample was also younger than what would have been expected had the sample's age distribution reflected the composition of self-reported drug-using network members.
5) Respondents can accurately report their personal drug-using network size, defined as the number of relatives, friends, and acquaintances who can be considered members of the target population(29).	Compare degree estimates to the number of drug-using network members reported in comparable studies.	Degree estimates, which were based on the number of drug users that each respondent reported in his/her network are also likely to be inaccurate. As only 59% of those recruited by their peers were eligible, the number of self-reported drug-using network members may be an overestimate. However fewer drug-using network members were reported in this study than in other studies among NIDUs(28) and IDUs(31,32). For example, Weeks and colleagues (2002) reported an average of 4.5 drug using peers(33) and Latkin and colleagues (1995) reported an average of 5.22 drug-using network members(32), both of which are larger than the 1.52 drug-using network members reported here. Each of these studies also reported more total network members (including non-drug network members; means: 5.6(33) and 10.3(32)) than what is reported here (mean=3.33).
6) Each respondent recruits a single peer(29).	To prevent recruitment chains from terminating early, most studies break this assumption. Instead of allowing each participant to recruit only one peer, most RDS studies allow 3 peer recruits.	Of the 390 RDS participants (including seeds) who were given peer recruitment coupons (13 participants were not given coupons), 87 recruited 1 eligible peer (22%), 47 (12%) recruited 2 eligible peers, 61 (16%) recruited 3 eligible peers, and the remaining did not refer any eligible peers.
7) The composition of those in the final sample is independent of those selected as seeds at equilibrium.	While an ideal test of this assumption would evaluate the independence of each seed and the respondents comprising its referral chain, this was not possible. Instead, t-tests and chi-square tests were used to assess the independence of seeds and peer recruits.	Seeds and peer recruits were independent with respect to gender, age, education, crack smoking and heroin snorting. Compared with recruits, seeds were more likely to be Hispanic and Black and to inject drugs

Table 4

Evaluating Sampling Biases and the Independence of Seeds and Peer Recruits in an RDS Sample of Illicit Drug Users in New York City, 2006-2009

	Seeds (n=46)	Recruits (n=357)	Networks (mean number of drug-using network members=1.52) N=534	*P*-value for difference between networks and recruits (2-sided)	*P*-value for independence of seeds and recruits (2-sided)
Gender				<0.01	0.11
Female (%)	34.8	23.9	36.9		
Male (%)	65.2	76.1	63.1		
Race/Ethnicity				0.2214	0.01
Hispanic (%)	54.4	34.2	29.4		
Black (%)	43.5	55.7	61.6		
White/Other (%)	2.2	10.1	9.1		
Median age	33.5	34	35	<0.01	0.62
≥ High School (%)	37.0	35.3	62.6	<0.01	0.82
Inject (%)	23.9	7.3	7.7	0.81	<0.01
Smoke crack (%)	76.1	85.4	77.1	<0.01	0.10
Snort heroin (%)	50.0	43.8	21.8	<0.01	0.43

Abbreviations: Respondent Driven Sampling (RDS)

Table 5

Sampling Differences by Recruitment Strategy., New York City (2006-2009)

	Targeted Street Outreach N=217		RDS (productive seeds and peer recruits) N=385			RDS (productive seeds and peer recruits) vs. Targeted Street Outreach	
	n	%	n	%	*P*-value (2-sided)	PR	95%CI
Gender					<0.01		
Female	82	38.3	95	24.7		Ref	Ref
Male	132	61.7	289	75.3		1.28	1.10, 1.49
Race/Ethnicity					<0.01		
Hispanic	77	35.5	135	35.1		1.45	1.11, 1.88
Black	93	42.9	213	55.3		1.58	1.23, 2.03
White/Other	47	21.7	37	9.6		Ref	Ref
Education					0.66		
< High School	143	66.2	248	64.4		Ref	Ref
≥ High School	73	33.8	137	35.6		1.03	0.91, 1.16
Income					0.20		
≤ 10,000	167	81.1	310	85.2		Ref	Ref
> 10,000	39	18.9	54	14.8		0.89	0.74, 1.07
Age (median, IQR)	32	27-37	34	29-38	<0.01	1.02	1.01-1.03
Homeless in last 6 months					<0.01		
No	92	42.4	122	31.7		Ref	Ref
Yes	125	57.6	263	68.3		1.19	1.04, 1.36
Injection status					<0.01		
NIDU	124	57.1	353	91.7		Ref	Ref
IDU	93	42.9	32	8.3		0.35	0.26, 0.47
HIV status					0.27		
Negative	185	92.5	321	89.7		Ref	Ref
Positive	15	7.5	37	10.3		1.12	0.93, 1.35
Heroin use in the past 6 months					<0.01		

	Targeted Street Outreach N=217		RDS (productive seeds and peer recruits) N=385			RDS (productive seeds and peer recruits) vs. Targeted Street Outreach	
	n	%	n	%	*P*-value (2-sided)	PR	95%CI
No	69	31.8	200	52.5		Ref	Ref
Yes	148	68.2	181	47.5		0.74	0.66, 0.83
Cocaine use in the past 6 months					0.47		
No	55	25.6	88	23.0		Ref	Ref
Yes	160	74.4	295	77.0		1.05	0.91, 1.22
Crack use in the past 6 months					<0.01		
No	61	28.1	59	15,4		Ref	Ref
Yes	156	71.9	325	84.6		1.37	1.13, 1.67
Number of people use drugs with					0.14		
None	87	40.9	131	34.0		Ref	Ref
≥ 1	130	59.2	254	66.0		1.10	0.97, 1.25
Number of people in network who use drugs							
None	74	34.3	103	26.8	0.05	Ref	Ref
≥ 1	142	65.7	282	73.3		1.14	0.99, 1.32
IDUs only							
Ever used needle exchange program					<0.01		
No	21	22.8	19	57.6		Ref	Ref
Yes	71	77.2	14	42.4		0.35	0.19, 0.62
Injection frequency (Median, IQR)[1]	6	4-6	4	1-6	<0.01	0.77	0.72-0.83

Abbreviations: CI, Confidence Interval; IDU, Injection Drug User; NIDU, Non-injection Drug User; RDS, Respondent Driven Sampling; IQR, Interquartile Range

[1] 0=never, 1=once a month or less, 2=2-3 days/month, 3=once/month, 4=2-3 days/week, 5=4-6 days/week, 6=everyday