Published in final edited form as:
J Epidemiol Community Health. 2014 December ; 68(12): 1207–1212. doi:10.1136/jech-2014-204206.

Do recruitment patterns of young men who have sex with men (YMSM) recruited through respondent-driven sampling (RDS) violate assumptions?

Gregory Phillips II[1], **Lisa M. Kuhns**[2,3], **Rob Garofalo**[2,3], and **Brian Mustanski**[1]

[1]Northwestern University, Feinberg School of Medicine, Department of Medical Social Sciences, Chicago, IL

[2]Division of Adolescent Medicine, Ann & Robert H. Lurie Children's Hospital, Chicago, IL

[3]Northwestern University, Feinberg School of Medicine, Department of Pediatrics, Chicago, IL

Abstract

Background—In order to generate unbiased estimates for data collected using respondent-driven sampling (RDS), a number of assumptions need to be met: individuals recruit randomly from their social network and people can accurately report their eligible network size. However, research has shown that these assumptions are often violated.

Methods—This study used baseline data from Crew 450, a longitudinal study of young men who have sex with men in Chicago who were recruited via a modified form of RDS, and its network substudy, in which a subset of 175 participants reported details on the composition and characteristics of their social network at either one or two years post-baseline.

Results—Nearly two-thirds of participants reported giving coupons to at least one alter (64.0%), and 56.3% believed their alter(s) used the coupons. Frequency of communication, closeness, and type of relationship played a major role in determining coupon distribution. Participants whose alters used coupons were significantly less likely to describe the strength of their relationship as "not at all close" (OR = 0.08; 95% CI: 0.02, 0.36) compared to "very close" and to communicate weekly (OR = 0.20; 95% CI: 0.08, 0.49) or 1–6 times in the last 6 months (OR = 0.18; 95% CI: 0.06, 0.59).

Conclusion—Contrary to RDS assumptions, we found that relationship characteristics played a significant role when individuals decided to whom they would give coupons.

Keywords

YMSM; RDS; recruitment methods

Please address correspondence and request for reprints to: Brian Mustanski, PhD, Northwestern University, Feinberg School of Medicine, Department of Medical Social Sciences, 625 N. Michigan Ave., Suite 2700, Chicago, IL 60611, brian@northwestern.edu.

Competing Interest
None declared

INTRODUCTION

Respondent-driven sampling (RDS) is a modified form of snowball sampling that was developed to recruit members of hard-to-reach or hidden populations. [1, 2] This method of sampling is considered to be superior to convenience sampling in that weighting of data collected using RDS leads to estimates that are generalizable to the networks from which they were drawn. [2, 3] However, there are a number of assumptions that need to be met in order to generate these unbiased estimates. One key assumption is that individuals recruit randomly from their own networks. [3–6] Although some prior research has shown that individuals recruit people who accurately reflect the demographics of their self-reported network (e.g., people with a large proportion of women in their network will be more likely to recruit women), [3, 5, 7] there have been conflicting findings [8–10] about whether this is an imprecise measure of random recruitment. These findings might more accurately reflect homophily – the tendency for an individual to recruit someone like themselves [1, 2] – than random selection of alters within a network. Violation of the random selection assumption can have a noticeable effect on results. Lu et al. conducted a simulation study on the performance of RDS when this assumption is violated and found that it biased the RDS estimator and increased the standard deviation and mean absolute error of estimates. [11]

In practice, the RDS assumption of random selection of recruits (alters) is rarely the case. A number of factors determine how an alter is selected: individuals tend to recruit people they have seen recently, [10] people who live near them or the interview site, [12] people with whom they have a close relationship, and people who they believe will participate in the study. [13] Studies that analyzed the reported relationship between an alter and his/her recruiter found that most were recruited by a friend or some other close associate. [14] However, it is difficult to accurately assess recruitment patterns, as there is rarely follow-up to see which alters were recruited from within the individual's network. Thus, the need for more data was highlighted by Gile and Handcock, who indicated that future studies are needed to ask "How do participants choose alters to whom to pass coupons?" and "To which alters do participants consider passing coupons?" [15]

Another assumption made by RDS is that participants can accurately report their degree, or network size. Participants in studies using RDS are usually required to identify the number of people they know, specifically those who could be eligible for the study. [16, 17] This self-reported network size is a major component in generating weights – people with small networks have a lower probability of being included in the study than people with large networks, so their data are weighted more heavily. [2, 3] Despite its importance, network size has been measured inconsistently across studies – some use a single question, while others use a series of questions; additionally, some fail to assess factors that comprise eligibility, such as geographic location. [18] Unfortunately, no standards have emerged to ensure response validity. Research has also shown that self-reported network size estimates are often inaccurate, [10, 19, 20] especially with larger networks. [9] Although some research has shown that errors in self-reported degree have few repercussions on RDS estimates with small sample sizes, [11] researchers remain concerned about the effect of inconsistent degree measures and reports. These potential issues in accurate assessment of

network size could have significant ramifications on the generalizability of weighted estimates.

To date, no thorough assessments of the validity of these factors inherent in RDS methodology have been conducted, mostly due to lack of data. In this manuscript, we review data collected in a longitudinal study of young men who have sex with men (YMSM) and its network substudy. At baseline, the parent study collected self-reported data on network size. Then, between one and two years later, a subset of participants was enrolled into the network substudy in which they reported the composition and characteristics of their social network, including elicitation of alters and alter characteristics. Hence, we had a unique opportunity to see how accurately self-reported network size matched self-reported network composition, since participants were required to provide names and details on each alter within their network. We also were able to assess the characteristics associated with selection of alters for participation in the parent study. Given that prior findings suggested inaccuracies in self-reported degree we hypothesized that the proportionate distribution of participants' self-reported eligible network size and the number of eligible social network alters would be highly correlated. We also hypothesized that participants would be more likely to recruit close friends into the study, rather than recruiting randomly from their social network.

METHODS

Participants

This study used baseline data from Crew 450, a longitudinal study of 450 YMSM in Chicago in which participants are followed over a period of two years, with study visits at 6-month intervals. Methodology for this study has previously been described. [21, 22] Briefly, participants were recruited into the study using a modified form of RDS. All eligible participants were given 3 coupons to distribute to their network members. Due to slow recruitment and unproductive seeds, reseeding continued throughout the study and resulted in a sample that consisted of 38.2% seeds. Additional description of the performance of RDS in this study has been presented elsewhere. [23] Individuals were considered to be eligible if they were between 16 and 20 years of age, male sex at birth, spoke English, had a previous sexual encounter with a man or identified as gay or bisexual, and were available for two years of follow-up.

Baseline data were matched with data collected in a social network substudy, in which selected Crew 450 participants were asked to provide additional information on their social and sexual networks (alters). All Crew 450 participants who had a one year follow-up visit (T3) or a two year follow-up visit (T5) during the network substudy recruitment period (2011–2012) were approached, resulting in 204 participants who were approached, 179 (87.7%) who agreed to participate, and 175 (85.8%) who completed the interview. More than two-thirds (n=135) were interviewed at T3, while 40 were interviewed at T5. Participants who were approached but did not participate in the substudy were similar to those who participated based on demographics and recruitment characteristics. For more information on the network substudy see Kuhns and colleagues. [23]

Name Generator

In the substudy, participants were asked to provide first and last names for all people in their social network. Five items were used to elicit names from the participants:

1. Name the people you are closest to, that is, people you see or talk to regularly and share your personal thoughts and feelings with.
2. Can you think of other people who would give time and energy to help you?
3. Can you think of other people who you could count on to lend or give you $25 or something of equal or greater value?
4. Can you think of other people who you could turn to for help or advice about gay-related issues or problems (for example, if you were being harassed)?
5. Can you think of other people you spend time with on a regular basis yet are not very close to?

Study staff capped the number of possible named alters at 40 for the substudy.

Measures - RDS

In Crew 450, participants were asked to report their network size – "*Approximately how many people do you know by name? These are people who you know and who also know you. You would know how to contact them and you have seen them in them in the past 6 months*." They were also asked to report the number of people in their network who might meet the study's eligibility criteria – "*Of those individuals that you know by name, how many people do you know who are young men between the ages of 16–20 who identify as gay, bisexual or queer or who have sex with other guys and who live in the Chicago area? These are people who you know and who also know you, who you know how to contact, and who you have seen in the past 6 months*." Additionally, they reported their relation to the person who recruited them into the study – "*How would you describe your relationship to the person who invited you to participate in this study. The person who gave you the coupon?*" For descriptive information on these responses see Kuhns et al. [24]

In the substudy, participants were asked about recruitment patterns in their network: (1) "*Who of the people on this list are people you gave coupons to for the Crew 450 study?*", (2) "*Who of the people on this list used the coupon you gave them to become a participant in the study (to the best of your knowledge)?*", (3) "*Who of the people on this list is the person who gave you a coupon to participate in the Crew 450 study?*"

Measures – Demographics

In Crew 450, participants provided their age, race/ethnicity, and sexual orientation. In the substudy, participants provided information on the age, race/ethnicity, gender, and sexual orientation for each of their alters. Self-reported information in the substudy (alter used coupon for Crew 450, participant was recruited for Crew 450 by alter, etc.) was matched with recruitment information in Crew 450 to determine if it corresponded with participants' reports.

In the substudy, participants were asked three questions to determine relationship type, frequency of communication, and strength of the relationship for each alter. Relationship type was assessed with the question: "*What is your relationship to each person?*" Frequency of communication was measured using the question "*How often have you talked with (or communicated by email or text) each of these people* ***over the last 6 months****?*" Response options were "*Not at all*," "*Once or twice*," "*Three to six times*," "*At least a couple of times a month*," "*Weekly*," and "*Daily*." Relationship strength was assessed by asking "*Rate the strength of your relationship with each person. Would you say your relationship is: very close, somewhat close, or not at all close?*"

Statistical Analysis

Reported network size in the Crew 450 questionnaire was compared to number of alters reported during the substudy interview, both for the larger social network and the eligible network (i.e., RDS network questions 1–2 above). A Shapiro-Wilk test was conducted on each variable to test for normal distribution of the data; if $p<0.05$ for at least one variable, then correlation between the two variables was assessed using Spearman ρ. These analyses were dichotomized by time of substudy survey administration to detect differences based on follow-up time.

Univariable analyses were conducted to investigate relationship characteristics of all alters, alters who were potentially eligible for participation in Crew 450 (participant said that alter was male, aged 17–21 years at T3 or aged 18–22 years at T5 to adjust for the time difference, and identified as gay/bisexual/queer), and alters who were given coupons for Crew 450. Bivariable and multivariable logistic regression models were developed to look at the characteristics of alters who were given coupons and those who were believed to have used coupons. All analyses included a clustering term to control for correlation between reported alters. Multivariable models controlled for any of the above demographic characteristic for which significant homophily was identified between the participant and his alters. SAS v9.3 (Cary, NC) was used for all analyses.

RESULTS

Network Size

Within the Crew 450 baseline questionnaire, participants reported a median of 50 people in their social network (range: 4 – 500). Comparatively, during the substudy interview these participants reported a median of 17 alters (range: 5 – 40), with 7 (4.0%) identifying the cap of 40 alters. There was a significant but moderate positive correlation between these reported network sizes (Spearman $\rho = 0.28$, $p < 0.001$). When participants were separated by follow-up time, a significant difference was seen: while the positive correlation remained for participants interviewed at T3 (Spearman $\rho = 0.28$, $p = 0.001$), it was no longer significant for those interviewed at T5 (Spearman $\rho = 0.15$, $p = 0.36$). After controlling for interview wave, there was a significant partial correlation (partial Spearman $\rho = 0.28$, $p < 0.001$) with the same magnitude as the unadjusted correlation.

Regarding potentially eligible network members (alters who were identified by their egos as male, aged 17–21 years at T3 or aged 18–22 years at T5, and a sexual identity of gay/bisexual/queer), participants reported a median of 11 people in the Crew 450 questionnaire (range: 0 – 500) and a median of 4 people in the substudy interview (range: 1 – 19). Again, there was a weak, significant positive correlation between these two measures (Spearman $\rho = 0.20$, $p = 0.01$). Similar to the overall reported network size, the significant correlation remained for participants interviewed at T3 (Spearman $\rho = 0.24$, $p = 0.01$), but not for those interviewed at T5 (Spearman $\rho = 0.06$, $p = 0.73$). Controlling for interview wave in the analysis resulted in a significant partial correlation (Spearman partial $\rho = 0.20$, $p = 0.01$) which was slightly larger than the unadjusted correlation.

Participant/Alter Relationships

In the substudy interview the majority of participants reported either daily (26.2%) or weekly (27.6%) communication with their alters (Table 1). Approximately three-quarters (74.9%) had a very close or somewhat close relationship with their alters. The vast majority of reported alters were friends (51.2%), with 17.3% reporting some other relationship with their alters. Predominantly, these other alters were ex-boyfriends (41.2%), with a smaller proportion identified as “booty calls,” “one night stands,” or “fuck buddies” (10.1%). Depiction of an ego-alter network is found in Figure 1.

Recruitment of Alters

Nearly two-thirds of participants reported giving coupons to at least one alter (64.0%), and 56.3% of these young men believed that at least one alter used the coupon. This was consistent with actual recruitment patterns in Crew 450; participants who thought one or more of their alters used a coupon in Crew 450 were significantly more likely to have had a person they recruited participate in the study (unadjusted odds ratio (OR) = 9.27; 95% confidence interval (CI): 4.16, 20.7).

The potentially eligible networks of participants demonstrated a high degree of homophily. Among Black YMSM, 87.4% of alters were also Black; among Latino YMSM, 72.7% of alters were also Latino; and among White YMSM, 75.4% of alters were also White. There was also a moderate, significant positive correlation between the participant’s age and the age of his identified alters (Spearman $\rho = 0.36$, $p < 0.001$).

Frequency of communication, closeness, and type of relationship played a major role in determining which alters were given coupons and which alters used the coupons. Alters who did not communicate with the participant, or who communicated less frequently than daily were significantly less likely to have been given coupons compared with alters who communicated at least on a daily basis, after controlling for homophilous characteristics (participant’s age and race/ethnicity) (Table 2). Compared with alters who were very close to the participant, alters who were somewhat close or not at all close were significantly less likely to have been given coupons, after controlling for participant’s age and race/ethnicity. With respect to relationship type, participants were less likely to give coupons to acquaintances/associates and others, compared to friends, after controlling for participant’s age and race/ethnicity.

In addition, alters who communicated weekly or less than monthly with the participant, those who were somewhat close or not at all close to the participant, and those who had another type of relationship with the participant were significantly less likely to have used the coupons given to them.

Recruitment by Alters

In addition to giving alters coupons, participants could also have received coupons from their alters. Nearly one-half of non-seed participants (45.1%) reported that they were recruited by one of their alters. Most participants reported communicating with their recruiter either daily (41.2%) or weekly (25.5%) and described their relationship as “very close” (62.8%). The majority of the recruiters were friends of the participant (72.6%); 7.8% were boyfriends and 5.9% were brothers.

DISCUSSION

Assessment of correlation demonstrated that the comparative magnitude of the eligible network sizes in Crew 450 and the substudy was approximately the same, at least for the network interviews conducted at the 12-month follow-up visit (T3). Although the weighting mechanism for data collected using RDS uses the inverse of the network size, the calculation relies on proportional distribution of eligible network sizes rather than absolute differences. [10] The moderate degree of correlation we found between the two eligible network size numbers demonstrates that the self-reported eligible network size from Crew 450 is a suitable relative surrogate for a person’s network size. Despite the fact that assessment of eligible network size through one or a series of questions in a survey has been identified as an imprecise measure, we found that, at least within this network of YMSM, new methods may not necessarily be needed to ensure accuracy for the purposes of RDS weighting. Still, more research is required to validate our findings, and we recommend that standardized degree questions are developed and implemented across all studies that utilize RDS. At least two questions should be assessed in the determination of degree – one for the individual’s entire network size and one for the individual’s eligible network size. The addition of each eligibility criterion increases the difficulty in identifying an accurate number, so precision of the network characteristics needs to be weighed against the precision of the estimated network size.

As was previously thought, YMSM in this study did not randomly select alters from their network to recruit into Crew 450. Instead they picked young men with specific characteristics – those with whom they had a close relationship and with whom they communicated frequently. The type of relationship also played a significant role – friends were chosen over acquaintances. Since we believed that previously published recruitment patterns might be due to homophily and not to random selection, we controlled for homophilous characteristics in our analyses by including these variables in the models. By removing its potential confounding effect, we can clearly see that random selection is not occurring. Since this key assumption was violated, weighted prevalence estimates might not be generalizable, even in the absence of other violations noted in this sample.

There are several possible solutions to this random selection assumption violation. One opportunity occurs when participants are given coupons - part of their recruiter training could stress that they distribute the coupons randomly to their eligible network. It is unlikely that this would have an effect on recruitment patterns, though; people preferentially give coupons to network members they know will participate in the study in order to earn the secondary recruiter incentive. One final potential intervention point comes at the end of data collection. Cognizant of inherent problems in RDS assumptions, many researchers have been working on developing new estimators that either take into account these potential violations or are independent of some of these assumptions. Therefore, in the near future, researchers who implement studies using RDS might be able to use novel estimators to abnegate any known affronts to assumptions.

Unexpectedly, less than one-half of non-seed participants reported being recruited by one of their alters. It is unlikely that they were recruited by a stranger, since no one in the Crew 450 study reported stranger recruitment. There are several more likely explanations for this phenomenon. One to two years had elapsed since they were initially recruited into Crew 450, so participants might no longer remember who gave them a coupon. Or, their recruiter might no longer be within their network due to ending a relationship or moving out of the area. Unfortunately, additional information about a participant's recruiter was not elicited for non-seeds whose network of alters did not contain their recruiter.

A potential limitation was the possibility that the composition of YMSMs networks could change significantly over time, since the substudy interview was conducted either one or two years after the initial baseline survey. We found that similarities remained in network composition at one year post baseline, but these disappeared by two years after the baseline survey. This reflects the dynamic nature of social networks and highlights the need for future research to measure network characteristics closer to the time that detailed alter information is collected. The decreasing association across the one year between interviews suggests we may have underestimated the associations compared to what we would have observed if the RDS and network interviews occurred at the same time. Of course, having both interviews occur concurrently would not be feasible as sufficient time is needed after coupons are provided in order to ascertain who redeems them. This also highlights the necessity for placing expiration dates on coupons for studies with more than one year of recruitment. Significant changes in networks over time mean that an individual recruited substantially long after survey administration might not be representative of the network characteristics reported in the questionnaire.

This study also had several other limitations. All data are reliant on self-report and are open to a number of biases, including recall and social desirability bias. Assessment of eligible network size differed between the two studies, and participants might not have been considering the same network members in each case. In the substudy interview, participants were asked to report on the age, race/ethnicity, gender, and sexual orientation of their alters; these are proxy reports and therefore might not be accurate. In addition, participants were not asked to provide more information on their alters (education, annual income, etc.) that could have been used to assess homophily and been added in adjusted analyses. Participants disclosed information on people who were within their social network (parents, friends,

boyfriends, etc.) and it is reasonable to assume that they are familiar enough with these individuals to provide basic demographic information. As part of the substudy interview, they were also asked to provide study staff with the names of their alters to facilitate matching with Crew 450 participants. Individuals might have been reticent to divulge the identities of people within their social/sexual network, and the number of alters might be an underestimate of the true value; however, participants had been involved with study staff for at least a year and most felt comfortable enough to disclose substantial information about their network alters. Additionally, the number of named alters was capped at 40; it is possible that without this upper limit, participants who reported 40 alters could have added many more. The number of named alters reported also involved a significant time commitment, and participants might have under-reported names in order to conserve time.

Conclusion

To our knowledge, this is one of the first studies to collect data from participants recruited via RDS before and after the distribution of coupons and collect detailed information on who they recruited into the study. Although asking participants to estimate their network size is imperfect, we found that it is an acceptable surrogate and may be useful in other studies. RDS assumes that people randomly select from their network when distributing their coupons; within this study, that was not the case. Participants were significantly more likely to recruit network members with whom they had closer relationships and who they saw on a frequent basis. Since this research was limited to a small group of YMSM, future study is needed on other hard-to-reach populations to assess if these findings are generalizable. Regardless, this aids in our understanding of RDS theory and implementation.

Licence for Publication

REFERENCES

1. Heckathorn DD. Respondent-driven sampling: A new approach to the study of hidden populations. Social Problems. 1997; 44:174–200.
2. Heckathorn DD. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. Social Problems. 2002; 49(1):11–34.
3. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology. 2004; 34:193–239.
4. Heckathorn DD. Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. Sociological Methodology. 2007; 37(1):151–207. doi:
5. Heckathorn DD, Semaan S, Broadhead RS, et al. Extensions of Respondent-Driven Sampling: A New Approach to the Study of Injection Drug Users Aged 18–25. AIDS Behav. 2002; 6(1):55–67.
6. Volz E, Heckathorn DD. Probability Based Estimation Theory for Respondent Driven Sampling. J Off Stat. 2008; 24(1):79–97.
7. Wang J, Carlson RG, Falck RS, et al. Respondent-driven sampling to recruit MDMA users: a methodological assessment. Drug Alcohol Depend. 2005; 78(2):147–157. [PubMed: 15845318]

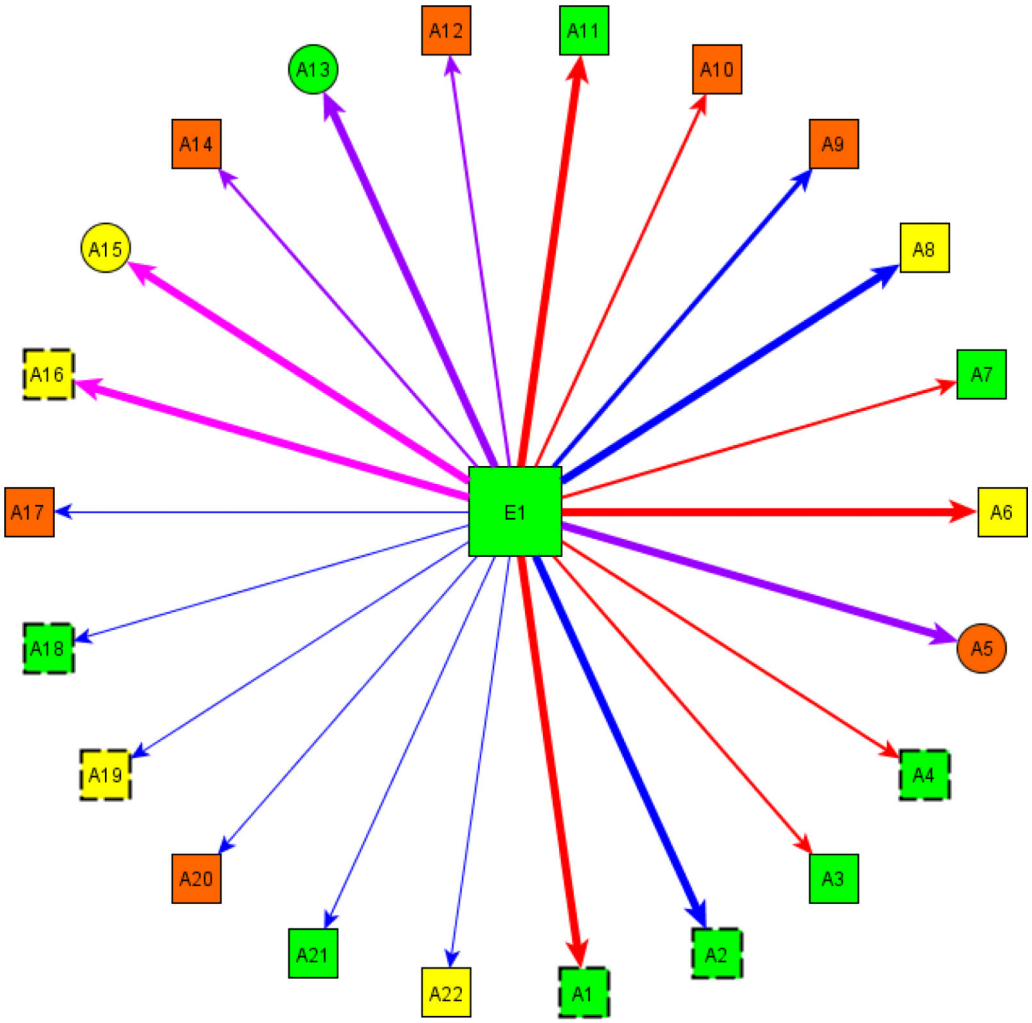
8. Wang J, Falck RS, Li L, et al. Respondent-driven sampling in the recruitment of illicit stimulant drug users in a rural setting: findings and technical issues. Addict Behav. 2007; 32(5):924–937. [PubMed: 16901654]
9. Iguchi MY, Ober AJ, Berry SH, et al. Simultaneous recruitment of drug users and men who have sex with men in the United States and Russia using respondent-driven sampling: sampling methods and implications. J Urban Health. 2009; 86(Suppl 1):5–31. [PubMed: 19472058]
10. Wejnert C. An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and out-of-Equilibrium Data. Sociol Methodol. 2009; 39(1):73–116. [PubMed: 20161130]
11. Lu X, Bengtsson L, Britton T, et al. The sensitivity of respondent-driven sampling. J R Stat Soc a Stat. 2012; 175:191–216. doi:
12. McCreesh N, Johnston LG, Copas A, et al. Evaluation of the role of location and distance in recruitment in respondent-driven sampling. International journal of health geographics. 2011; 10:56. [PubMed: 22008416]
13. Lu X, Malmros J, Liljeros F, et al. Respondent-driven sampling on directed networks. Electron J Stat. 2013; 7:292–322. doi:
14. Frost SD, Brouwer KC, Firestone Cruz MA, 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–i97. [PubMed: 17072761]
15. Gile KJ, Handcock MS. Respondent-Driven Sampling: An Assessment of Current Methodology. Sociol Methodol. 2010; 40(1):285–327. [PubMed: 22969167]
16. Magnani R, Sabin K, Saidel T, et al. Review of sampling hard-to-reach and hidden populations for HIV surveillance. Aids. 2005; 19(Suppl 2):S67–S72. [PubMed: 15930843]
17. Ramirez-Valles J, Heckathorn DD, Vazquez R, et al. 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]
18. Johnston LG, Malekinejad M, Kendall C, et al. 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–S141. [PubMed: 18535901]
19. Brewer DD, Garrett SB, Kulasingam S. Forgetting as a cause of incomplete reporting of sexual and drug injection partners. Sex Transm Dis. 1999; 26(3):166–176. [PubMed: 10100775]
20. Bell DC, Belli-McQueen B, Haider A. Partner Naming and Forgetting: Recall of Network Members. Social networks. 2007; 29(2):279–299. [PubMed: 17940583]
21. Newcomb ME, Ryan DT, Garofalo R, et al. The effects of sexual partnership and relationship characteristics on three sexual risk variables in young men who have sex with men. Arch Sex Behav. 2014; 43(1):61–72. [PubMed: 24217953]
22. Mustanski B, Johnson AK, Garofalo R, et al. Perceived Likelihood of Using HIV Pre-exposure Prophylaxis Medications Among Young Men Who Have Sex with Men. AIDS Behav. 2013; 17(6):2173–2179. [PubMed: 23128980]
23. Kuhns L, Birkett M, Muth SQ, et al. Modifying and adapting participant-aided sociograms for a study of social, sexual and substance-using networks among young men who have sex with men. Under Review. 2013
24. Kuhns LM, Kwon S, Ryan DT, et al. Evaluation of Respondent-driven Sampling in a Study of Urban Young Men Who Have Sex with Men. 2014 Submission.

What is already known on this subject?

Respondent-driven sampling has been shown to be effective at recruiting hard to reach populations. However, the theory behind this method is reliant on a number of assumptions that are often violated, specifically that individuals can accurately report their eligible network size and that they recruit randomly from their networks. Unfortunately, there is rarely the opportunity to follow-up with participants to see whom they recruited and accurately test these assumptions.

What does this study add?

We collected data at two time points (prior to coupon distribution and at least one year later) with two different surveying methods, which provided us with the unique opportunity to assess these assumptions. We found a significant positive correlation between numbers of eligible network members assessed two different ways, but that recruitment was not random – participants were more likely to recruit people with whom they communicated and saw most often.



Shape = gender: square = male; circle = female

Color = age: green = less than 25; yellow = 25-30; orange = more than 30

Line thickness = strength of relationship

Line color = type of relationship: red = friend; blue = sex partner; purple = relative; pink = co-worker

Shape outline: dashed = given coupon; solid = not given coupon

Figure 1.
Sample ego network with 22 alters (all Black members).

Table 1

Relationship characteristics of participants and alters.

	All alters (n=3390)	Potentially eligible alters* (n=846)	Potentially eligible alters given coupons (n=174)
	n (%)	n (%)	n (%)
Frequency of communication, last 6 months:			
Not at all	249 (7.3)	92 (10.9)	4 (2.3)
Once or twice	357 (10.5)	103 (12.2)	16 (9.2)
Three to six times	325 (9.6)	81 (9.6)	7 (4.0)
At least a couple of times a month	636 (18.8)	140 (16.6)	21 (12.1)
Weekly	934 (27.6)	203 (24.0)	48 (27.6)
Daily	888 (26.2)	227 (26.8)	78 (44.8)
Strength of relationship:			
Very close	1327 (39.2)	303 (35.8)	93 (53.5)
Somewhat close	1212 (35.8)	271 (32.0)	59 (33.9)
Not at all close	849 (25.1)	272 (32.2)	22 (12.6)
Type of relationship:			
Friend	1736 (51.2)	498 (58.9)	121 (69.5)
Acquaintance/Associate	319 (9.4)	94 (11.1)	6 (3.5)
Parent	151 (4.5)	0 (0.0)	0 (0.0)
Sibling	175 (5.2)	5 (0.6)	0 (0.0)
Other relative	287 (8.5)	8 (0.9)	2 (1.1)
Boyfriend	73 (2.2)	43 (5.1)	15 (8.6)
Coworker/Colleague	61 (1.8)	1 (0.1)	0 (0.0)
Other	587 (17.3)	197 (23.3)	30 (17.2)

*Potentially eligible alters were identified as male, gay/bisexual/queer, and 17–21 (T3) or 18–22 (T5) years of age by the participant.

Table 2

Association between coupon distribution/use and relationship.

	Alter given coupon		Alter used coupon	
	Unadjusted	**Adjusted***	**Unadjusted**	**Adjusted***
	OR (95% CI)	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**
Frequency of communication, last 6 months:				
Not at all	***0.08 (0.03, 0.19)***	***0.07 (0.03, 0.19)***	--**	--**
Once or twice	***0.31 (0.17, 0.58)***	***0.30 (0.16, 0.57)***	***0.25 (0.09, 0.66)*****	***0.18 (0.06, 0.59)*****
Three to six times	***0.16 (0.07, 0.36)***	***0.15 (0.07, 0.35)***		
At least a couple of times a month	***0.31 (0.17, 0.55)***	***0.31 (0.17, 0.55)***	0.59 (0.20, 1.80)	0.38 (0.10, 1.40)
Weekly	**0.54 (0.33, 0.89)**	**0.54 (0.33, 0.88)**	***0.22 (0.10, 0.52)***	***0.20 (0.08, 0.49)***
Daily	1.00 (--)	1.00 (--)	1.00 (--)	1.00 (--)
Strength of relationship:				
Very close	1.00 (--)	1.00 (--)	1.00 (--)	1.00 (--)
Somewhat close	***0.57 (0.39, 0.83)***	***0.54 (0.37, 0.78)***	0.60 (0.29, 1.23)	**0.41 (0.19, 0.91)**
Not at all close	***0.18 (0.10, 0.30)***	***0.17 (0.10, 0.28)***	***0.09 (0.02, 0.38)***	***0.08 (0.02, 0.36)***
Type of relationship:				
Friend	1.00 (--)	1.00 (--)	1.00 (--)	1.00 (--)
Acquaintance/Associate	***0.20 (0.08, 0.52)***	***0.20 (0.08, 0.54)***	1.05 (0.27, 4.16)	0.83 (0.10, 7.30)
Relative	0.52 (0.11, 2.39)	0.50 (0.11, 2.32)	--**	--**
Boyfriend	1.86 (0.94, 3.69)	1.87 (0.91, 3.84)	1.58 (0.53, 4.70)	1.69 (0.56, 5.09)
Other	***0.52 (0.33, 0.82)***	***0.52 (0.33, 0.83)***	**0.22 (0.07, 0.73)**	***0.16 (0.05, 0.49)***

* Adjusted for participant's race/ethnicity and age.

** Dropped due to zero cell size.

*** "Once or twice" and "three to six times" categories were merged due to small cell size.

Bold: p<0.5; ***Bold-italic***: p<0.1