Assessing the Assumptions of Respondent-Driven Sampling in the National HIV Behavioral Surveillance System among Injecting Drug Users

Amy Lansky[*,1], Amy Drake[1], Cyprian Wejnert[1], Huong Pham[2], Melissa Cribbin[1] and Douglas D. Heckathorn[3]

[1]*Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, USA*

[2]*Northrop Grumman Corporation/BCA, Atlanta, Georgia, USA*

[3]*Department of Sociology, Cornell University, Ithaca, NY, USA*

Abstract: Several assumptions determine whether respondent-driven sampling (RDS) is an appropriate sampling method to use with a particular group, including the population being recruited must know one another as members of the group (i.e., injection drug users [IDUs] must know each other as IDUs) and be networked and that the sample size is small relative to the overall size of the group. To assess these three assumptions, we analyzed city-specific data collected using RDS through the US National HIV Behavioral Surveillance System among IDUs in 23 cities. Overall, 5% of non-seed participants reported that their recruiter was "a stranger." 20 cities with multiple field sites had ≥1 cross-recruitment, a proxy for linked networks. Sample sizes were small in relation to the IDU population size (median = 2.3%; range: 0.6%-8.0%). Researchers must evaluate whether these three assumptions were met to justify the basis for using RDS to sample specific populations.

INTRODUCTION

Behavioral surveillance of persons at risk of HIV infection is an important component of an overall HIV surveillance program [1,2]; these data are used to estimate prevalence, identify correlates of behaviors and determine prevention needs. Multiple methods have been used to sample populations at high risk of HIV infection including venue-based, time-space sampling; targeted sampling; snowball sampling; and respondent-driven sampling [3,4]. Respondent-driven sampling (RDS) [5,6] has been used successfully to reach injecting drug users (IDUs) in the United States [7,8] and elsewhere [9].

RDS has certain assumptions that must be met to determine if it is an appropriate sampling method to use with a particular group [10,11]. These assumptions require that the population being recruited must know one another as members of the target population (i.e., IDUs must know each other as IDUs). If members of the population cannot identify each other, then participants will not be able to produce eligible recruits and the method will fail to produce a sample. The population being recruited also must be adequately networked to accommodate a chain referral process; ideally, networks should form a single component (network of networks), rather than multiple, disconnected networks, so that referral chains can reach all subsets of the population in a defined area. Subsets of the population that are completely disconnected from the primary network cannot be reached by the peer recruitment process and thus the RDS findings will not be generalizable to these groups. A third assumption, that the sample size to be recruited using RDS is small relative to the overall size of the target population (i.e. a small sampling fraction), is required to ensure that each participant's ability to be recruited remains constant over time because the pool of potential recruiters is not noticeably diminished [10]. Given that respondents may only participate once, it is important to ensure that the sample size does not exhaust the pool of potential recruiters in the population as sampling progresses. Two other RDS assumptions, that participants can accurately report their personal network size and that recruitment is a random selection from the recruiter's network, are applicable to RDS analysis. Discussion of these assumptions is beyond the scope of this paper and has been reported elsewhere [12,13].

Few RDS studies have assessed these three assumptions. To build the literature on situations in which RDS does and does not work well as a recruitment and sampling strategy for reaching hard to reach groups, there is a need for quantitative indicators to assess the RDS assumptions. This paper defines quantitative measures to evaluate, post-hoc, the extent to which the three assumptions were met in the US National HIV Behavioral Surveillance System among injecting drug users for the first cycle of data collection from May 2005 to February 2006 (NHBS-IDU1). Based on this evaluation, we describe the lessons learned that were then applied to the second cycle (NHBS-IDU2).

MATERIALS AND METHODOLOGY

Methods for NHBS-IDU are reported in detail elsewhere [14] and briefly described here. NHBS-IDU1 was conducted by the Centers for Disease Control and Prevention (CDC) in collaboration with state and local health departments in 23

*Address correspondence to this author at the Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, USA; Tel: 404-639-5200; Fax: 404-639-0897; E-mail: alansky@cdc.gov

metropolitan statistical areas ("cities") within the United States. CDC determined that NHBS-IDU1 was not research; each local area obtained approval of human subjects in accordance with their institutions' determinations.

Local project staff in each city started the NHBS-IDU1 cycle with formative research to determine logistics of survey operations and to gather information on the local IDU population [15]. Each city set up at least one interview field site accessible to the various local drug-use networks and began RDS with a limited number (8-10) of initial recruiters or 'seeds' representing various drug networks and geographic or demographic characteristics.

NHBS-IDU1 procedures included eligibility screening, obtaining oral informed consent from participants, and an interviewer-administered survey. Eligibility for NHBS includes being of age 18 or older, being a resident of the city, not having already participated in the current NHBS data collection cycle, and being able to complete the survey in English or Spanish. An additional IDU cycle eligibility criterion was having injected drugs within 12 months preceding the interview date, measured by self-report and either evidence of recent injection or adequate description of injection practices [14]. The survey measured characteristics of participants' IDU networks (total number, gender and race/ethnicity), demographics, drug use and injection practices, sexual behaviors, HIV testing history, and use of HIV prevention services. Interviewers used handheld computers to administer the survey and record responses.

Participants could take the survey at any NHBS field site in their city. Participants who completed the survey were asked and trained to recruit others who also injected drugs by distributing number-coded coupons. Participants were compensated for their participation and for each eligible recruit who completed the survey; this dual-incentive structure is unique to RDS [5,6]. Compensation levels were determined in each city, but generally were about $25 for participation and $10 for recruitment.

NHBS-IDU1 was conducted from May 2005 through February 2006. Data collection duration varied across cities due to differences in timing for approval of human subjects, logistics, and speed of sample accrual.

Measures

Participants who agreed to be recruiters were told to give coupons to someone they knew as an IDU. Participants (excluding seeds) described their relationship to the person who gave them their coupon. Multiple responses were allowed, including: sex partner, drug partner, family, friend, colleague, acquaintance, and stranger ("you don't really know the person, just met him/her"). For analysis purposes, the participant's recruiter was categorized as a stranger if this response option was selected with no additional relationship reported.

Five variables which may affect participants' recruitment selections and introduce sampling bias were assessed: race/ethnicity, gender, age, preferred drug, and self-reported HIV status. Race and ethnicity were coded into one variable with mutually exclusive categories: white, black, Hispanic (regardless of race), and other (including Asians, Native Hawaiian and Pacific Islanders, multiracial persons, and

those with no recorded race). The variable "preferred drug" was derived from questions asking frequency of use of several drug types and then grouped into 5 categories: heroin only, heroin and cocaine (equal frequency or combined as speedball), cocaine or crack only, amphetamine (including methamphetamine), and other (all other drugs or combinations thereof). Self-reported HIV status was categorized as HIV-positive or not (which included those whose results were negative or indeterminate, those who never received the result or never tested, and those whose HIV status could not be determined).

Data Management and Analysis Methods

Coupon numbers and other information linking recruiters to their recruits were collected and maintained in RDS Coupon Manager (RDSCM) 2.0 software (Cornell University, Version 2.0, Ithaca, New York, USA). Survey data were transferred from the handheld to a computer and then uploaded to a secure server; some survey records were lost during collection or transfer and only the recruitment data from RDSCM 2.0 remained. Survey and RDSCM 2.0 data were merged using SAS software (SAS Institute Inc., Version 9.1, Cary, North Carolina, USA) and output to an electronic text file for analysis in RDSAT software (Cornell University, Version 5.6, Ithaca, New York, USA). The analyses for this paper included only eligible participants, except where otherwise noted. For some analyses, city-specific samples were aggregated to report on the whole NHBS-IDU1 sample.

Indicators for RDS Assumptions

Respondents know one another as members of the target population. Using SAS, we calculated the proportion of participants reporting that their recruiter was a stranger; a low proportion (2-4%) indicates that this assumption is met [16]. We also assessed the proportion of potential participants who were eligible as a way to determine the extent to which participants knew one another as IDUs; a high proportion of ineligible recruits would suggest that this assumption was not met.

Respondents' networks are linked and form a single network. We used RDSAT to create a matrix of cross-recruitments. To determine whether the IDU networks within each city were linked, cross-recruitment was assessed for field site, as networks often are defined by geography. An example of cross-recruitment is when a participant interviewed at Field Site B had received his/her coupon from a recruiter interviewed at Field Site A. We also assessed cross-recruitment for the 5 variables; we report data only for race/ethnicity as it had the most impact on sampling. To be considered linked at least one recruitment between any two field sites or any two racial/ethnic groups, respectively, was required. The presence of at least one cross-recruitment in the sample suggests the presence of a large number of connections across groups in the population; the higher the proportion of cross-recruitments, the greater the number of network connections among IDUs.

Sample size is small relative to size of the target population. The sampling fraction was defined as the number of persons screened for NHBS-IDU1 (regardless of

eligibility) divided by the total number of IDUs in each city [17].

RESULTS

Recruitment

From May 2005 to February 2006 a total of 13,519 persons were recruited, 384 of whom were seeds. A total of 1,563 (12%) persons were deemed ineligible and excluded from analysis: 196 did not meet NHBS general eligibility criteria (86 of whom were ineligible due to previous participation) and 1,367 did not meet current injection drug use criteria. Additionally, 46 persons had no recruitment information so their records could not be used. There were 334 persons with lost survey records. In addition, we did not include for analysis 38 persons with responses of highly questionable validity and 67 who were not classified as either male or female.

In the complete analysis dataset, there were 334 seeds and 11,137 peer-recruited participants recruited for a total of 11,471 participants. Table **1** displays characteristics of the overall sample; city-specific characteristics of NHBS-IDU1 participants are reported elsewhere [18]. Among the 11,471 participants, most (71%) were male and were of age 35 years and older (81%) (Table **1**). Nearly half (49%) were black, 25% white, and 21% Hispanic. Heroin was the preferred drug for 53% of the sample and 8% self-reported they were HIV-infected.

RDS Assumptions

Respondents know one another as members of the target population. Table **1** shows responses regarding the relationship to the recruiter (as reported by the participant). The most common (59%) relationship was "friend;" many reported relationships related to drug use such as someone they "buy drugs with" or "buy drugs from." Overall, 5% of non-seed participants reported that their recruiter was "a stranger" (with no other relationship; only 26 persons reported stranger and another relationship); this proportion varied by city (range 1.2%-20%), with 5 cities having >5% recruitment by strangers (Table **2**).

The proportion of potential participants who were eligible for NHBS-IDU1 was high overall (90%) and in each city (range 83%-98%, Table **2**). The majority of potential participants (61%, range 40%-86%) had physical signs of recent injection (data not shown). Although a higher proportion of ineligibles in cities with a high proportion of participants recruited by a stranger might be expected, we did not see this pattern (Table **2**).

Respondents' networks are linked and form a single network. Of the 23 NHBS-IDU1 cities, 3 used a single field site, so cross-recruitment was not assessed. All other cities had multiple field sites, ranging from 2 to 7 with an average of 4 field sites. In 3 cities with multiple field sites, each had 1 field site with no cross-recruitment to any other field site. In 1 of these cities, a new field site was opened after the existing ones were closed, making cross recruitment to this site impossible. We assume cross recruitment would have occurred from this field site had it been possible and therefore included all data in the analysis dataset. The other

Table 1.   **Characteristics of Participants--United States, National HIV Behavioral Surveillance System: Injecting Drug Users, May 2005-February 2006**

Characteristic	No.	%
Gender		
Male	8,158	71
Female	3,313	29
Age Category (Years)		
18-24	443	4
25-34	1,730	15
35-44	3,600	31
44-54	4,374	38
≥55	1,324	12
Race/Ethnicity		
White	2,841	25
Black	5,630	49
Hispanic	2,429	21
Other[a]	571	5
Preferred Drug		
Heroin	6,053	53
Heroin and Cocaine[b]	3,599	31
Cocaine or crack	788	7
Amphetamine[c]	626	6
Other[d]	405	4
HIV-Positive		
Yes	882	8
No[e]	10,589	92
Relationship to Recruiter[f]		
Main sex partner	355	3
Casual sex partner	178	2
Friend	6,543	59
Relative/family member	390	4
Person buy drugs from	332	3
Person buy drugs with	2,257	20
Person use drugs with	3,317	30
Person share needles with	609	6
Acquaintance	2,362	21
Stranger (only)[g]	519	5
Total	11,471	

Abbreviations: HIV, human immunodeficiency virus.
[a]Includes Asians, Native Hawaiian and Pacific Islanders, person who reported multiple races and those for whom race was not recorded.
[b]Heroin and cocaine use with equal frequency or combined as speedball.
[c]Includes methamphetamine.
[d]Includes all other drugs or combination of drugs.
[e]Includes those who tested HIV negative (n=9,048), those whose confirmatory test was indeterminate (n=41), those who never received a test result (n=532), those never tested (n=914) and those for whom HIV test status could not be ascertained (n=54).
[f]Relationships were reported by the participant; >1 response was allowed, therefore percentages do not add to 100. Seeds were not asked this question; percentages based on 11,137 participants
[g]Relationship was categorized as "stranger" if it was the only category chosen by the participant. If stranger was chosen as one of multiple categories, the responses appear in those categories but not in the "stranger" category.

Table 2. **Selected Characteristics of Samples, by city--United States, National HIV Behavioral Surveillance System: Injecting Drug Users, May 2005-February 2006**

Metropolitan Statistical Area ("City")	IDU Population Size[a]	NHBS-IDU Sample Size[b]	Sampling Fraction[c]	Proportion Eligible[d]	Proportion Recruited by a Stranger [e]	Cross-Recruitment by Field Site	Cross-Recruitment by Race/ Ethnicity [g]
	No.	No.	%	%	%	%	%
Atlanta, Georgia	14,602	616	4.2	91	12	18	17
Baltimore, Maryland	58,720	785	1.3	92	20	21	25
Boston, Massachusetts	67,044	540	0.8	88	2	30	35
Chicago, Illinois	32,206	653	2.0	83	4	46	18
Dallas, Texas	31,931	620	1.9	92	3	35	27
Denver, Colorado	20,689	612	3.0	87	4	74	43
Detroit, Michigan	27,166	568	2.1	96	3	n/a[f]	16
Fort Lauderdale, Florida	7,375	441	6.0	87	8	36	32
Houston, Texas	34,117	662	1.9	90	1	n/a[f]	32
Las Vegas, Nevada	13,708	341	2.5	98	17	8	40
Los Angeles, California	98,616	661	0.7	91	2	7	42
Miami, Florida	9,280	740	8.0	82	3	25	34
Nassau, New York	12,177	557	4.6	95	0.4	0.2	47
New Haven, Connecticut	13,629	593	4.4	90	2	11	34
New York City, New York	91,327	529	0.6	96	4	2	33
Newark, New Jersey	16,153	550	3.4	80	2	0.3	21
Norfolk, Virginia	10,259	580	5.7	86	4	14	9
Philadelphia, Pennsylvania	58,722	586	1.0	92	3	25	24
St Louis, Missouri	10,942	633	5.8	83	0.2	n/a[f]	8
San Diego, California	25,946	550	2.1	98	2	39	44
San Francisco, California	28,462	646	2.3	90	6	36	51
San Juan, Puerto Rico	15,031	585	3.9	98	4	2	--
Seattle, Washington	28,505	471	1.7	85	2	6	52
Total/Median[h]	726,607	13,519	2.3	90	5	N/A	NA

Abbreviations: HIV, human immunodeficiency virus; IDU, injecting drug user; NHBS-IDU, National HIV Behavioral Surveillance System: Injecting Drug Users; n/a, not applicable.
[a]Number of IDUs in the MSA was obtained from Brady *et al*. [17].
[b]"Sample size" includes all recruited persons regardless of eligibility.
[c]Sampling fraction was calculated as the NHBS-IDU sample size (column 2) divided by the IDU population size (column 1).
[d]Denominators do not include records without recruitment information, lost records, or persons excluded based on validity of response or gender.
[e]Percentage of non-seed participants who said the person who gave them the coupon was a stranger.
[f]Did not use multiple field sites.
[g]Race/ethnicity cross-recruitment not calculated for San Juan as 99% were Hispanic. The Norfolk sample was 87% Black and the St Louis sample was 91% Black.
[h]Total for population size, sample size, proportion eligible, and proportion recruited by a stranger; median value for sampling fraction.

2 cities had a field site located in an area that was geographically distant from the other locations, with limited hours of operation; there was no evidence suggesting that participants interviewed at these 2 field sites were part of the same networks as participants from other field sites. Therefore, data from these 2 field sites (n=90) were considered separate networks (i.e., not part of one component) and were excluded from the analysis dataset.

In all of the cities with multiple field sites there was at least 1 cross-recruitment by field site and by race/ethnicity. The proportion of cross-recruitments by field site ranged from 0.2% to 74% (Table **2**). The proportion of cross-recruitments by race/ethnicity ranged from 8% to 52% (Table **2**). In the two cities with the lowest proportion of cross-recruitments, nearly all the participants were Black (Table **2**).

Sample size is small relative to size of the target population. The sample sizes by city ranged from 341 to 785 (Table **2**). Overall, the sampling fraction was low, with less than 10% of the IDU population sampled in each city (median = 2.3%; range: 0.6%-8.0%).

DISCUSSION

In summary, NHBS-IDU1 met the three RDS assumptions we assessed based on the quantitative indicators we created. Results for each assumption varied by city. Related to the first assumption, that participants knew one another as members of the target population, we found that, for most cities, the proportion of recruitments by a stranger was low while the proportion of eligible recruits was high. In 5 cities the proportion recruited by a stranger was >5%, but these cities still had high eligibility rates suggesting that participants knew each

other well enough to recognize each other as IDUs. This assumption also has implications for analysis as RDS weighting is based on individuals with larger networks having greater likelihood of being recruited; if many participants recruit strangers (i.e., persons outside their network), then RDS weights based on network size would not be applicable. To examine the second RDS assumption, that the IDU networks within the NHBS cities were linked, we examined cross-recruitment by field site and by race/ethnicity. Cross-recruitment by field site ranged from 0.2% to 74%. Two cities had limited cross-recruitment by race/ethnicity, which may suggest that IDU networks in these cities are racially defined. When there is a low proportion of cross-recruitments, RDS analysis may still produce valid estimate; however the variance around these estimates will be noticeably high. For the third assumption, we found that in each city the sampling fraction was too small to noticeably diminish the recruiter pool, therefore allowing for robust recruitment.

This is the first paper to assess the extent to which the three RDS assumptions were met in samples from a standardized, multi-city behavioral surveillance system in the United States using quantitative indicators. The results from this paper can be used to guide other researchers to conduct similar evaluations of their own RDS studies. We created indicators for the assumptions that are easy to calculate; although we conducted our assessment post-hoc, the assumptions should be considered during formative research and the indicators can be used while planning an RDS study (e.g., considering sampling fraction by using existing population size estimates and planned sample size) or monitored as part of process evaluation during sample accrual (proportion recruited by a stranger and cross-recruitments) so that recruitment can be adjusted as needed. Rudolph *et al.* [19] also described ways they tested RDS assumptions in New York City among IDUs, using similar metrics reported here.

Two papers reviewing 123 RDS studies outside the US discussed challenges [20] and summarized characteristics of RDS studies [7]. Papers such as these have not reported data on whether these 3 assumptions were met empirically. Few other studies have reported on relationships between recruiters and recruits, including the proportion recruited by a stranger or cross-recruitments [19]. Other RDS studies have reported high proportions of eligible recruits, similar to the high proportion found in NHBS-IDU1 [21-23]. The hidden nature of most RDS target populations often precludes knowledge of population size and therefore makes calculation of the sampling fraction more challenging; we were able to use existing published estimates of the IDU population size in each NHBS city [17]. This is the first paper to report sampling fractions for 23 RDS samples collected using a standard protocol. Our data can contribute to refinement of theoretical work related to RDS estimation: in NHBS-IDU1, the overall sampling fraction was 2.3%, a figure well below the threshold of 50%, at which sampling-with-replacement can become a source of bias [24].

Our analyses had some limitations that suggest further development of quantitative indicators of the three RDS assumptions. Field site may not be the best variable to assess whether networks are sufficient to sustain a chain-referral process; other factors such as neighborhood of residence or zip code may be more relevant within each NHBS city to determine the extent to which networks are related. Our findings on cross-recruitment by race/ethnicity are similar to that reported in another IDU study in New York City [19]. Future research should consider what proportion of cross-recruitment is considered adequate to demonstrate linked networks; our standard of 1 cross-recruitment is a minimum level for lack of cross-recruitment to be ruled out, rather than a level of adequate cross-recruitment. Local NHBS project staff are encouraged to examine the assumptions considered here for their own data and staff from each NHBS city should consider their knowledge of the local IDU population to determine how well RDS sampled different groups of IDU within their city. The sample of IDUs reached by RDS can be compared to other methods of recruitment to determine if key sub-populations were missed [25].

Based on the analysis reported here, additional operational procedures were developed for NHBS-IDU2. A more refined definition of 'knowing' someone was added to the question assessing the relationship to the recruiter as well as to the recruiter training script (*By "know," I mean you know their name **OR** you see them around even if you don't know their name)*. Participants who reported that their recruiter was a stranger were probed using standardized questions; if participants reported never seeing the recruiter prior to being given a coupon or reported having first seen the recruiter in a situation related to NHBS-IDU, then the relationship classification of 'stranger' was considered validated. In addition, recruiters were trained not to give coupons to strangers. As part of their formative research, NHBS-IDU staff were required to analyze peer recruitment patterns in their NHBS-IDU1 data by race/ethnicity, gender, and other characteristics of potentially insular sub-populations of IDU (i.e., networks that are not linked to other networks). Based on this information, staff selected seeds from loosely networked sub-populations to ensure each group's representation, whereas closely networked sub-populations did not require the same extent of planning for selecting seeds. In addition, staff assessed potential field sites in part for the location's ability to serve as a "bridge" between major IDU sub-populations. Other formative research activities such as identifying studies of local IDU populations that describe networks and other characteristics of drug users can also help lay the foundation for the success of an RDS sample in reaching all groups of IDUs [15].

RDS is increasingly used to sample IDUs and other populations at high risk of HIV infection. As RDS is still a relatively new sampling and analysis method, it is important for investigators to share operational findings. As use of RDS increases, researchers must not only report on whether RDS assumptions were met to justify its use among specific populations, as we did here, but also plan formative research to ensure that assumptions can be met.

AUTHOR DISCLAIMER

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

REFERENCES

[1]   Lansky A, Sullivan PS, Gallagher K, *et al*. The HIV behavioral surveillance program in the US: A conceptual framework. Public Health Rep 2007; 122(Suppl1): 16-23.

[2]   Diaz T, DeCock K, Brown T, *et al*. New strategies for HIV surveillance in resource-constrained settings: an overview. AIDS 2005, 19 (Suppl 2): S1-S8.

[3]   Semaan S, Lauby J, Liebman J. Street and network sampling in evaluation studies of HIV risk-reduction interventions. AIDS Rev 2002; 4: 213-23.

[4]   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): S67-72.

[5]   Heckathorn, DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl 1997; 44: 174-99.

[6]   Heckathorn, DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. Soc Probl 2002; 49: 11-34.

[7]   Broadhead RS, Heckathorn DD, Weakliem DL, *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.

[8]   Abdul-Quader AS, Heckathorn DD, McKnight C, *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.

[9]   Malekinejad M, Johnston LG, Kendall C, *et al*. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. AIDS Behav 2008; 12(4 Suppl): S105-30.

[10]   Heckathorn DD. Extensions of respondent driven sampling: analyzing continuous variables and controlling for differential recruitment. Sociol Methodol 2007; 37(1): 151-207.

[11]   Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociol Methodol 2004; 34: 193-239.

[12]   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(1 Suppl): S5-S31.

[13]   Wejnert C. An empirical test of respondent-driven sampling: point estimates, variance, degree measures, and out-of-equilibrium data. Sociol Methodol 2009; 39: 73-116.

[14]   Lansky A, Abdul-Quader AS, Cribbin M, *et al*. Development of an HIV behavioral surveillance system for injecting drug users: the National HIV Behavioral Surveillance System. Public Health Rep 2007; 122(Suppl1): 48-55.

[15]   Allen DR, Finlayson T, Abdul-Quader A, *et al*. The role of formative research in the National HIV Behavioral Surveillance System. Public Health Rep 2009; 124: 26-33.

[16]   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: 387-402.

[17]   Brady JE, Friedman SR, Cooper HL, *et al*. Estimating the prevalence of injection drug users in the U.S. and in large U.S. metropolitan areas from 1992 to 2002. J Urban Health 2008; 85(3): 323-51.

[18]   Centers for Disease Control and Prevention. Risk, Prevention, and Testing Behaviors Related to HIV and Hepatitis Infections—National HIV Behavioral Surveillance System: Injecting Drug Users, May 2005-February 2006. HIV Special Surveillance Report 7. Available at: http://www.cdc.gov/hiv/topics/surveillance/resources/reports/. Published June 2011. [Accessed August 8, 2011].

[19]   Rudolph AE, Crawford ND, Latkin C, *et al*. Sub-populations 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: 280-9.

[20]   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): S131-41.

[21]   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(7): i60-i72.

[22]   Stormer A, Tun W, Guli L, *et al*. An analysis of respondent driven sampling with injection drug users (IDU) in Albania and the Russian Federation. J Urban Health 2006; 83(7): i73-i82.

[23]   Zhong F, Lin P, Xu H, *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: 1058-66.

[24]   Gile, K. Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. J Am Stat Assoc 2011; 106: 135-46.

[25]   Malekinejad M, McFarland W, Vaudrey J, Raymond HF. Accessing a diverse sample of injection drug users in San Francisco through respondent-driven sampling. Drug Alcohol Depend 2011; 118: 83-91.