

Factors Associated with Productive Recruiting in a Respondent-Driven Sample of Men who Have Sex with Men in Vancouver, Canada

Jamie I. Forrest, Nathan J. Lachowsky, Allan Lal, Zishan Cui, Paul Sereda, Henry F. Raymond, Gina Ogilvie, Eric A. Roth, David Moore, and Robert S. Hogg

ABSTRACT *Respondent-driven sampling (RDS) has become a preferred sampling strategy for HIV research and surveillance in many global settings. Methodological investigation into the validity of RDS-generated samples has helped improve theoretical components of design. However, the operational challenges of implementing RDS remain underreported. We sought to identify factors independently associated with productive recruiting in an urban RDS-generated sample of gay, bisexual, and other men who have sex with men (MSM). Data were collected from the Momentum Health Study, a cohort of MSM recruited by RDS in Vancouver, Canada. Eligible men were given up to six RDS coupons to recruit their peers. The primary outcome was a count variable of each participant's number of eligible recruits. Multivariable Poisson regression identified independent predictors of productive recruitment. In total, 719 individuals comprised this analysis, of which 119 were seeds. The distribution of eligible recruits was right skewed, with 391 (54.4 %) having never recruited another participant and only eight participants (1.1 %) having recruited five. Significant, independent predictors of recruiting one additional participant included network size per ten unit increase (adjusted risk ratio [aRR] 1.03), being of Aboriginal race/ethnicity compared with White (aRR 1.51), being HIV-positive (aRR 1.31), being sexually active with only males (aRR 2.48), being single compared with common law/married (aRR 1.37), having recently read gay newspapers (aRR 1.58), having recently sought sex partners online (aRR 1.33) and being out to a male parent (aRR 1.30). This analysis demonstrates the importance of social network size in RDS adjustment, but also identifies other socio-demographic and behavioral variables that increased RDS coupon return, which may help researchers better operationalize the implementation of RDS.*

KEYWORDS *HIV, Gay and bisexual men, MSM, Respondent-driven sampling, RDS, Recruitment*

Forrest, Lachowsky, Ogilvie, and Moore are with the Faculty of Medicine, University of British Columbia, Vancouver, Canada; Lachowsky, Lal, Cui, Sereda, Moore, and Hogg are with the British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada; Raymond is with the San Francisco Department of Public Health, San Francisco, USA; Raymond is with the University of California San Francisco, San Francisco, USA; Roth is with the Department of Anthropology, University of Victoria, Victoria, Canada; Hogg is with the Faculty of Health Sciences, Simon Fraser University, Burnaby, Canada.

Correspondence: Jamie I. Forrest, Faculty of Medicine, University of British Columbia, Vancouver, Canada. (E-mail: jamie.forrest@alumni.ubc.ca)

INTRODUCTION

Sampling strategies for harder-to-reach populations continue to pose challenges for population health researchers. In the context of the HIV epidemic, groups most at risk for HIV, notably sex workers, people who inject drugs, and gay, bisexual, and transgender individuals, may also experience marginalization that limits their engagement with health services and research.¹ Traditional forms of probabilistic sampling, that rely on a sampling frame, are prohibitive given the lack of information of this kind among marginalized at-risk populations. Conventional forms of sampling harder-to-reach populations for HIV epidemiological research include convenience sampling and venue-based or time-location sampling.¹ These strategies, however, come with their own inherent sampling biases.

Respondent-driven sampling (RDS) has become a popular alternative, adaptive sampling strategy² in many global settings.³ RDS is a method of long-chain peer referral recruitment with a strict set of assumptions for statistical adjustments that aim to account for the biases inherent in social network sampling.⁴ RDS adjustments for network size and homophily can be applied to descriptive statistics as well as multivariable model building for statistical inference.⁵

The theoretical constructs of RDS,⁶ improving the validity of sample estimates⁷⁻¹⁰ and comparisons of RDS with other sampling strategies,¹¹ are extensively debated in the literature. In comparison, very little is published on the practical and operational realities of collecting these samples, which has led to a renewed call for more robust reporting practices on these issues.¹² Where extensively reviewed, the focus is on implementation challenges of conducting RDS in low-income settings.¹³ Among RDS samples of MSM in high-income settings, the limited reporting has focused on identifying factors associated with the productivity of seeds, the purposively sampled individuals who begin the long-chain referral sampling process. For example, Reisner and colleagues, who sampled MSM in Massachusetts, USA, identified a number of sociodemographic factors that increased the odds of a seed recruiting two or more subsequent participants.¹⁴ This important work has identified some of the challenges associated with operationalizing RDS, but limited in its ability to infer factors associated with productive recruiting in an entire RDS-generated sample. For example, a seed may be productive (having recruited his maximum allowable number of recruits), but if these participants are not, the network will not grow. Thus, the purpose of this study is to identify independent predictors of more productive recruiters in a RDS urban, high-income setting sample of men who have sex with men by using a “count” outcome.

METHODS

Data Source and Study Procedures

Data were taken from the Momentum Health Study, a longitudinal cohort of MSM, recruited by RDS in Metro Vancouver, Canada. Eligibility criteria for the study included being at least 16 years of age, gender identifying as a man (regardless of sex at birth), having reported sex with a man in the previous 6 months, being able to complete a questionnaire in English, currently residing in Metro Vancouver, and being and recruited by a peer (or be a purposively selected seed). Study staff administered a screening tool to identify potentially ineligible participants. After a free and informed consent was given by the participant, men were asked to complete a computer-assisted self-interview (CASI) behavioral questionnaire that asked about

their demographics, HIV testing (and treatment), sexual behaviors, substance use patterns, community participation and connectedness, RDS-specific social network data, and knowledge and attitudes towards HIV and prevention. Participants then completed a nurse-administered clinical questionnaire and rapid point-of-care HIV test and serological tests for syphilis and hepatitis C virus.

Following completion of the study procedures, men were asked to help the study by recruiting their peers. Following a short training session by study staff, who explained the importance of recruiting other eligible participants within their social or sexual networks, participants were given up to six RDS coupons to pass on to their potential recruits. Compensation for participation included a \$50 honourarium and \$10 for each eligible recruit who successfully completed a study visit. The University of British Columbia -Providence Healthcare Research Institute, Simon Fraser University, and University of Victoria granted ethical approval for this study.

Data Measures

The primary outcome measure for this analysis was a count of each participant's total number of recruited eligible participants. The maximum number of RDS coupons given out by study staff to each participant was six, although no participants recruited more than five eligible peers. Thus, the range of responses of the primary outcome variable was zero to six. Independent variables of interest included age (continuous), self-reported participant network size (continuous), sexual identity (gay, bisexual, or other, including two-spirit and queer), race/ethnicity (White, Asian, Aboriginal, Latin American, or other), annual income (<\$30,000 versus \geq \$30,000), HIV status (negative, positive), ever tested for HIV versus not, partnership status (common law/married, not common law/married, or not partnered/single), recent types of sexual partners (only males, mostly males, or mostly/only females), whether the participant reported being out to their paternal parent or guardian versus not, whether the participant read gay newspapers in the past 6 months or not, whether the participant sought sex partners online in the past 6 months or not, and how much of the participant's social time was spent with other gay guys (≤ 25 , 26–50, 50–75, or >75 %).

Data Analysis

Descriptive statistics were used to characterize the data using SAS version 9.3 (SAS Institute, Cary, North Carolina). Recruitment weights were generated to account for differences in recruitment and homophily (characteristics determined a priori to be strongly associated with the formation of social ties) and degree weights were generated to account for variations in network size (reported number of members of the population known by each participant). These were calculated using RDSAT version 7.1.46 developed by Heckathorn.⁴ Poisson regression was preferred over negative binomial regression as the data dispersion was not deemed significant. Poisson regression was determined most appropriate to identify participant characteristics associated with an increased count in the number of each participant's eligible recruits. Application of the Vuong test¹⁵ found that a zero-inflated Poisson model was significantly preferred over a standard Poisson model. The final multivariable model was selected using backward selection criteria that minimized the Akaike Information Criterion (AIC). Results are presented as risk ratios (RRs), interpreted as the increased risk associated with recruiting one additional recruit, and 95 % confidence intervals (CI) are presented. All significance

tests were two-sided, and p values of less than 0.05 were considered statistically significant.

RESULTS

In total, 719 individuals contributed to this analysis, of whom 119 were seeds. Of these, 587 (81.6 %) were given six RDS coupons, 78 (10.8 %) were given one to five RDS coupons, and 54 (7.5 %) were not given any coupons, either because our sample size was nearly achieved or because study staff questioned the eligibility of participants and their recent recruits. Across all participants, a total of 3705 RDS coupons were issued. In terms of successful recruitment of eligible MSM, 391 (54.4 %) recruited none, 161 (22.4 %) recruited one, 94 (13.1 %) recruited two, 49 (6.8 %) recruited three, 16 (2.2 %) recruited four, and eight (1.1 %) recruited five. Figure 1 shows the distribution of the primary outcome variable, number of RDS coupons redeemed by eligible participants.

Table 1 shows the overall characteristics of the sample. For variables characterized by proportions, an RDS adjusted proportion and associated 95 % CI are presented. A number of crude proportions for independent variables did not fall within 95 % CI around RDS estimates: bisexual identity, self-identifying as White, income, reporting sex with only males, being out to a paternal parent/guardian, reading gay newspapers, and social time spent with other gay men. The median age of was 33 years (Q1, Q3: 26, 47) and a majority identified as gay (80.7 %, 95 % CI 76.2–85.5). With respect to race/ethnicity, 68.1 % (95 % CI 60.9–74.1) self-identified as White, 9.9 % (95 % CI 6.5–14.9) as Asian, 10.4 % (95 % CI 5.6–16.1) as Aboriginal, 6.9 % (95 % CI 3.0–11.4) as Latin American and 4.6 % (95 % CI 2.3–7.8) with another identity group. Nearly three quarters of participants (74.1 %; 95 % CI 68.9–79.6) reported an annual income less than \$30,000 and 23.3 % (95 % CI 16.1–31.1) were HIV-positive.

Table 2 presents the results of the final multivariable Poisson model. Significant independent predictors of recruiting one additional participant included network

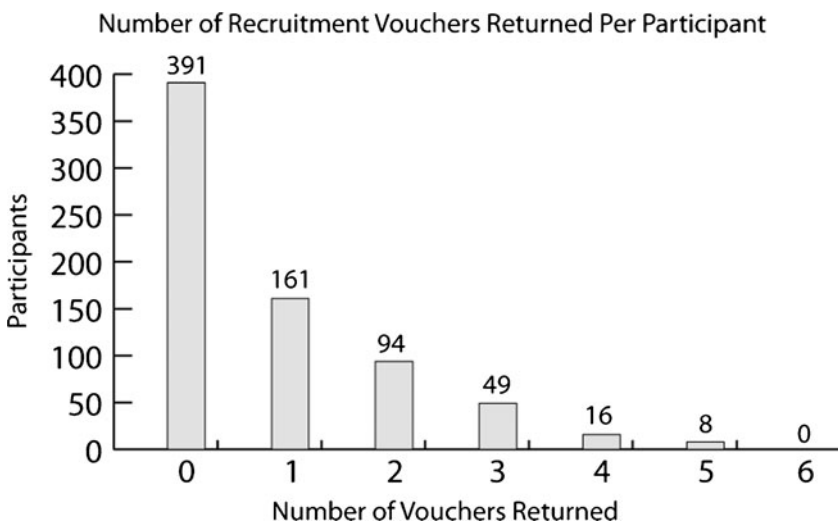


FIG. 1 Distribution of recruitment productivity: number of RDS coupons returned per participant.

TABLE 1 Crude and RDS-adjusted descriptive statistics of study sample

	Crude	RDS-adjusted
	<i>n</i> (%)	% (95 % CI)
Number of recruits		
0	337 (50.7)	—
1	161 (24.2)	—
2	94 (14.1)	—
3	49 (7.4)	—
4	16 (2.4)	—
5	8 (1.2)	—
Age (continuous)		
Median [Q1, Q3]	33 [26, 47]	—
Participant network size (continuous)		
Median [Q1, Q3]	10 [5, 20]	—
Sexual identity		
Gay	612 (85.1)	80.7 (76.2–85.5)
Bisexual	66 (9.2)	15.3 (10.4–19.5)
Other	41 (5.7)	4.0 (2.4–6.1)
Race/ethnicity		
White	539 (75.0)	68.1 (60.1–74.1)
Asian	72 (10.0)	9.9 (6.5–14.9)
Aboriginal	50 (7.0)	10.4 (5.6–16.1)
Latin American	31 (4.3)	6.9 (3.0–11.4)
Other	27 (3.8)	4.6 (2.3–7.8)
Annual income		
< \$30,000/year	457 (63.6)	74.1 (68.9–79.6)
≥ \$30,000/year	262 (36.3)	25.9 (20.4–31.1)
Self-reported HIV status		
Negative	520 (72.3)	76.7 (68.9–83.9)
Positive	199 (27.7)	23.3 (16.1–31.1)
Ever tested for HIV		
No	39 (5.4)	6.4 (4.0–9.0)
Yes	680 (94.6)	93.6 (91.0–96.0)
Common law relationship status		
Not common law/married	148 (20.6)	19.3 (14.5–24.0)
Common law/married	125 (17.4)	18.1 (13.7–24.7)
No regular partner	446 (62.0)	62.6 (55.9–68.4)
Sexual activity in P2Y		
With only males	39 (85.8)	78.6 (74.3–84.7)
With mostly males	70 (9.7)	12.9 (7.9–16.2)
With mostly/only females	32 (4.5)	8.6 (4.9–12.3)
Out to paternal parent/guardian		
No	193 (28.8)	41.7 (35.5–49.4)
Yes	478 (71.2)	58.3 (50.6–64.5)
Reads gay newspapers in P6M		
No	121 (16.8)	23.2 (18.4–29.2)
Yes	598 (83.2)	76.8 (70.8–81.6)
Cruises online in P6M		
No	267 (37.1)	41.6 (35.4–48.3)
Yes	452 (62.9)	58.4 (51.7–64.6)
Social time with other gay men		
25 % or less	162 (22.8)	31.5 (26.0–37.0)

TABLE 1 *Continued*

	Crude	RDS-adjusted
	<i>n</i> (%)	% (95 % CI)
26–50 % of social time	183 (25.7)	26.8 (21.9–33.1)
50–75 % of social time	227 (31.9)	28.8 (23.0–34.0)
More than 75 % of social time	139 (19.6)	12.9 (9.2–17.1)

RDS respondent driven sample, *95 % CI* confidence interval, *Q1*, *Q3* first and third quartiles, *P2Y* past two years, *P6M* past six months

size per ten unit increase (adjusted risk ratio [aRR] 1.03; 95 % [95 %] CI 1.00–1.06, being of Aboriginal race/ethnicity compared with White (aRR: 1.51; 95 % CI: 1.09–2.09), being HIV-positive (aRR 1.31; 95 % CI 1.08–1.59), being sexually active with only males compared with mostly or only females (aRR 2.48; 95 % CI 1.10–5.55), being single compared with common law/married (aRR 1.37; 95 % CI 1.01–1.87), having recently read gay newspapers (aRR 1.58; 95 % CI 1.19–2.10), having

TABLE 2 Factors independently associated with productive recruitment in a multivariable Poisson model

	aRR	95 % CI
Participant network size (per 10 unit increase)	1.03	1.00–1.06
Race/ethnicity		
White	1.00	–
Aboriginal	1.51	1.09–2.09
Latin American	1.41	0.94–2.11
Asian	0.94	0.67–1.33
Other	0.91	0.53–1.56
Self-reported HIV status		
Negative	1.00	–
Positive	1.31	1.08–1.59
Sexual activity in P2Y		
Mostly or only females	1.00	–
Mostly males	2.41	1.02–5.67
Only males	2.48	1.10–5.55
Common Law relationship status		
Not common law/married	1.00	–
Common law/married	0.73	0.53–0.99
No regular partner	0.94	0.75–1.19
Reads gay newspapers in P6M		
No	1.00	–
Yes	1.58	1.19–2.10
Cruise online in P6M		
No	1.00	–
Yes	1.33	1.09–1.63
Out to male parent		
No	1.00	–
Yes	1.30	1.03–1.65

aRR adjusted risk ratio, *95 % CI* 95 % confidence interval, *P2Y* past 2 years, *P6M* past 6 months

recently sought sex partners online (aRR 1.33; 95 % CI 1.09–1.63) and reporting being out to their male parent (aRR 1.30; 95 % CI 1.03–1.65).

DISCUSSION

Our study found that having a larger network size, being Aboriginal, HIV-positive, exclusively sexual with males, relationally single, reading gay newspapers, and cruising online as independent predictors of recruiting additional participants to our sample. This is the first analysis, to our knowledge, to build a Poisson regression multivariable model to identify independent predictors of having recruited additional participants within an RDS study.

A major strength of our work is its use of all recruitment data, not restricted to the limited data provided by RDS seeds, with the aim of identify factors associated with productive recruiting in the entire sample. This approach adds additional value to the study conducted by Reisner and colleagues¹⁴ that applied logistic regression to identify factors associated with seeds recruiting two or more recruits. By applying Poisson regression methods, this study was able to use a more nuanced outcome. However, readers should be cautious in interpreting our findings as they may be specific to an urban-based MSM population and not generalizable to other settings.

There are three important dividends for RDS implementation that can be ascertained from our work. Most notably, our finding demonstrating social network size to be a significant independent predictor of productive recruitment underlines the importance of making statistical adjustments to RDS samples to account for a network size bias. Network size bias is a key component of RDS statistical adjustment⁶ (along with homophily) and this study's finding that network size independently predicts productive recruitment, validates the importance of this bias adjustment in generating RDS-weighted proportions.

Second, our study finding indicates that gay community connectedness is an important factor in determining productive recruitment. As noted here, men who only have male sex partners were 2.5 times more likely to recruit an additional participant, compared with men who reported having mostly or only female partners, after controlling for all other variables in the model, which included network size. Men who reported reading gay newspapers were also significantly more likely to recruit an additional participant. These findings suggest that men participating in the social and sexual networks of Vancouver's gay communities and demonstrating greater community connectedness are most likely to recruit subsequent participants. This finding speaks to the importance of targeted seed selection and real-time recruitment monitoring. A further exploration of homophily would help understand whether the factors associated with more productive recruitment are similar to those with high homophily (more likely to recruit similar individuals).

Finally, our work demonstrates that RDS is an important tool for the recruitment of marginalized MSM in settings often considered very vulnerable for participants.³ When this vulnerability is evenly distributed across multiple sub-identities of gay communities, it can be assumed that RDS would generate samples more representative of their populations. However, when vulnerability across the target population is more heterogeneous, the representativeness of the sample may come into question. For example, this study found that being HIV-positive and being Aboriginal were both independent predictors of productive recruitment. These groups have been shown to have a lower socioeconomic status in Vancouver,^{16,17}

and the fact that these groups were productive recruiters may speak to the advantage of RDS in recruiting more vulnerable populations that may exhibit closer social connectedness as a result of their marginalization. Conversely, the previous sample of gay and bisexual men collected in Vancouver by time-location sampling, recruited a proportionately smaller number of low-income and HIV-positive MSM.¹⁸ However, it is worth noting that income and age, two other predictors of socioeconomic status, were not significant in our final multivariable Poisson model. This is supported by Kendall et al.,¹⁹ but in contrast to a finding reported by McCreesh and colleagues that younger men and lower SES less likely to be recruited in their sample.²⁰ It also may explain the finding reported by Kuhns and colleagues²¹ of short recruitment chains and a segmentation in patterns of recruitment by race/ethnicity among young MSM in Chicago, where community connectedness and the vulnerabilities experienced by this population are not homogenous.

This analysis presents an opportunity to explore further strategies to balance RDS theory and practice in an effort to promote recruitment and generate the most representative samples.²² This includes innovative strategies for recruitment, such as electronic RDS coupons, and operational factors like study site location and coupon design, that may also facilitate more productive recruiting.²³ Consideration of these factors when implementing RDS means researchers must think beyond the theoretical underpinnings of the methodology and consider operational factors that will help optimize data collection in each context-dependent setting.

In conclusion, our study identified a number of sociodemographic, network, and social variables that independently predicted an increased number of eligible recruits by participants in a RDS sample of MSM in Metro Vancouver, Canada. More effort should be made to characterize recruitment dynamics in RDS generated samples and share strategies that may help researchers improve the operational realities of implementing this widely applied sampling strategy. As we prepare to recruit a second serial cross section of MSM in Vancouver and these findings will be considered when implementing the next RDS study.

ACKNOWLEDGMENTS

The authors would like to thank the Momentum Study participants, office staff and community advisory board, as well as our community partner agencies, the Health Initiative for Men, YouthCO HIV, and HepC Society of BC, and the Positive Living Society of BC. Momentum is funded through the National Institute on Drug Abuse (Grant #R01DA031055-01A1) and the Canadian Institutes for Health Research (Grant # MOP-107544). NJL is supported by a CANFAR/CTN Postdoctoral Fellowship. DMM is supported by a Scholar Award from the Michael Smith Foundation for Health Research.

REFERENCES

1. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS*. 2005; 19(2): S67–72.
2. Thompson SK, Collins LM. Adaptive sampling in research on risk-related behaviors. *Drug Alcohol Depend*. 2002; 68(Suppl 1): S57–67.

3. Malekinejad M, Johnston LG, Kendall C, Kerr LR, Rifkin MR, Rutherford GW. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. *AIDS Behav.* 2008; 12(4 Suppl): S105–30.
4. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl.* 1997; 44(2): 174–99.
5. Heckathorn DD. Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. *Sociol Methodol.* 2007; 37(1): 151–207.
6. Gile KJ, Handcock MS. Respondent-driven sampling: an assessment of current methodology. *Sociol Methodol.* 2010; 40(1): 285–327.
7. Liu H, Li J, Ha T, Li J. Assessment of random recruitment assumption in respondent-driven sampling in egocentric network data. *Soc Netw.* 2012; 1(2): 13–21.
8. Mills HL, Johnson S, Hickman M, Jones NS, Colijn C. Errors in reported degrees and respondent driven sampling: implications for bias. *Drug Alcohol Depend.* 2014; 142: 120–6.
9. Rudolph AE, Fuller CM, Latkin C. The importance of measuring and accounting for potential biases in respondent-driven samples. *AIDS Behav.* 2013; 17(6): 2244–52.
10. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J Urban Health.* 2006; 83(6 Suppl): i98–112.
11. Sabin KM, Johnston LG. Epidemiological challenges to the assessment of HIV burdens among key populations: respondent-driven sampling, time-location sampling and demographic and health surveys. *Curr Opin HIV AIDS.* 2014; 9(2): 101–6.
12. White RG, Hakim AJ, Salganik MJ, et al. Strengthening the reporting of observational studies in epidemiology for respondent-driven sampling studies: “STROBE-RDS” statement. *J Clin Epidemiol.* 2015; 68(12): 1463–71.
13. 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.
14. Reisner SL, Mimiaga MJ, Johnson CV, et al. What makes a respondent-driven sampling “seed” productive? Example of finding at-risk Massachusetts men who have sex with men. *J Urban Health.* 2010; 87(3): 467–79.
15. Vuong QH. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: J Econ Soc.* 1989; 57(2): 307–33.
16. Duncan KC, Reading C, Borwein AM, et al. HIV incidence and prevalence among aboriginal peoples in Canada. *AIDS Behav.* 2011; 15(1): 214–27.
17. Joy R, Druyts EF, Brandson EK, et al. Impact of neighborhood-level socioeconomic status on HIV disease progression in a universal health care setting. *J Acquir Immune Defic Syndr.* 2008; 47(4): 500–5.
18. Moore DM, Kanters S, Michelow W, et al. Implications for HIV prevention programs from a serobehavioural survey of men who have sex with men in Vancouver, British Columbia: the ManCount study. *Can J Public Health.* 2012; 103(2): 142–6.
19. Kendall C, Kerr LR, Gondim RC, 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): S97–104.
20. McCreesh N, Copas A, Seeley J, et al. Respondent driven sampling: determinants of recruitment and a method to improve point estimation. *PLoS One.* 2013; 8(10), e78402.
21. Kuhns LM, Kwon S, Ryan DT, Garofalo R, Phillips G 2nd, Mustanski BS. Evaluation of respondent-driven sampling in a study of urban young men who have sex with men. *J Urban Health.* 2015; 92(1): 151–67.
22. Truong HM, Grasso M, Chen YH, et al. Balancing theory and practice in respondent-driven sampling: a case study of innovations developed to overcome recruitment challenges. *PLoS One.* 2013; 8(8), e70344.
23. Forrest JI, Stevenson B, Rich A, et al. Community mapping and respondent-driven sampling of gay and bisexual men’s communities in Vancouver, Canada. *Cult Health Sex.* 2014; 16(3): 288–301.