

## Implementation of respondent-driven sampling among female sex workers in Brazil, 2009

Implementação do método de amostragem *respondent-driven sampling* entre mulheres profissionais do sexo no Brasil, 2009

Giseli Nogueira Damacena <sup>1</sup>  
Célia Landmann Szwarcwald <sup>1</sup>  
Aristides Barbosa Júnior <sup>2</sup>

### Abstract

*Female sex workers are known in Brazil and elsewhere in the world as one of the most-at-risk populations for risk of HIV infection, due to their social vulnerability and factors related to their work. However, the use of conventional sampling strategies in studies on most-at-risk subgroups for HIV is generally problematic, since such subgroups are small in size and are associated with stigmatized behaviors and/or illegal activities. In 1997, a probabilistic sampling technique was proposed for hard-to-reach populations, called Respondent-Driven Sampling (RDS). The method is considered a variant of chain sampling and allows the statistical estimation of target variables. This article describes some assumptions of RDS and all the implementation stages in a study of 2,523 female sex workers in 10 Brazilian cities. RDS proved appropriate for recruiting sex workers, allowing the selection of a probabilistic sample and the collection of previously missing information on this group in Brazil.*

*HIV; Prostitution; Sexual Behavior*

<sup>1</sup> Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

<sup>2</sup> Departamento de DST, AIDS e Hepatites Virais, Ministério da Saúde, Brasília, Brasil.

#### Correspondence

G. N. Damacena

Laboratório de Informações em Saúde, Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz.

Av. Brasil 4365, Pavilhão Haity Moussatché, Rio de Janeiro, RJ 21040-900, Brasil.  
giselind@fiocruz.br

### Introduction

At the national and international levels, HIV prevention work has been based on the natural history of infection, experiences in specific AIDS programs, and mathematical and simulation models <sup>1,2</sup>. Increasing importance has been assigned to most-at-risk population groups, who, although relatively small in size, play a preponderant role in the spread of the disease, especially in countries with concentrated epidemics <sup>3</sup>.

In relation to sexual transmission, female sex workers as a group are recognized as a most-at-risk population for HIV infection, both in Brazil <sup>4,5,6</sup> and in various other countries <sup>7,8,9,10,11,12</sup>, due to their social vulnerability and factors related to their work, such as multiple sex partners. In addition to sexually transmitted infections (STI), which act as co-factors <sup>13</sup>, studies of sex workers have shown that time in the profession and the consumption of illegal drugs are also associated with greater risk of HIV infection <sup>4,5,14,15,16</sup>.

In Brazil, the size of the subgroup of female sex workers has been estimated at 1% of the Brazilian female population 15-49 years of age <sup>17</sup>, that is, more than half a million women. Since 2002, prostitution has been officially recognized as a profession in the country (Ministry of Labor and Employment, Brazilian Classification of Occupations. <http://www.mtecbo.gov.br/cbosite/pages/home.jsf>, accessed on Oct/2009), but is still

surrounded by stigma and discrimination, which hinders access to the health system and other social interventions<sup>18</sup>.

The situation of increased risk of HIV infection, syphilis, and other STI led to several studies on female sex workers in Brazil beginning in the 1990s<sup>5,19</sup>. However, these studies generally used convenience samples, thus hindering the estimation of variables for nationwide monitoring of the HIV epidemic in this population group<sup>20</sup>.

Studies in most-at-risk subgroups for STI through conventional sampling strategies are generally problematic. Since these subgroups are numerically small and linked to stigmatized or illegal activities, they are considered hard-to-reach populations<sup>21</sup>. Since the mid-1990s, the development of probabilistic sampling methods has been encouraged, not only to ease the difficulties in access, but also to allow statistical estimation of the target variables<sup>21,22,23</sup>. In this context, time-space sampling (TSS)<sup>24</sup> and respondent-driven sampling (RDS) were proposed<sup>25</sup>.

Since it was first proposed, RDS has been widely used in various countries in studies of most-at-risk population subgroups for HIV<sup>26,27,28,29,30,31,32,33</sup>. In a recent review of studies in the literature that used RDS in hard-to-reach populations, Johnston et al.<sup>34</sup> identified 128 HIV surveillance studies performed outside the United States from 2003 and 2007. Likewise, Malekinejad et al.<sup>35</sup> identified 123 studies, 59 of which in Europe, 40 in Asia/Pacific, 14 in Latin America, 7 in Africa, and 3 in Oceania.

In Brazil, the commitment to HIV surveillance in most-at-risk groups required nationally representative studies. In 2006, the Department of STD, AIDS, and Viral Hepatitis, under the Ministry of Health, promoted the transfer of sampling methodology for hard-to-reach populations<sup>36</sup>. In the years 2008 and 2009, research projects were conducted in three population subgroups (men who have sex with men – MSM, female sex workers, and drug users), using RDS as the sampling method. The current article aims to describe some assumptions of the RDS sampling method and all the stages in the method's implementation in a study of female sex workers, disclosure of a recently proposed sampling methodology for hard-to-reach populations in the Brazilian literature.

### Assumptions of the RDS method

The RDS method was proposed in 1997 by Heckathorn<sup>25</sup>. It is considered a variant of chain sampling, since members of the target population group recruit their own peers to participate in

the study, assuming that capturing individuals in hard-to-reach populations is facilitated when the recruitment is done by individuals from the same population, who are part of a social network.

However, unlike other, non-probabilistic chain sampling methods like the so-called “snowball” technique<sup>20</sup>, RDS is implemented under statistical assumptions that allow calculating the selection probabilities, thus making it a probabilistic sampling method.

RDS assumes that persons with a given characteristic or activity are connected in a social network and have links to other persons with similar characteristics. The data collection is done through successive recruitment cycles called “waves”. First, individuals from the target population called “seeds” are selected non-randomly to participate in the study<sup>25</sup>. The seeds are asked to recruit a fixed number of peers from among their friends and acquaintances in the same population subgroup. The peers recruited by the seed also recruit other peers, and so on.

The participants' tendency to recruit peers with similar characteristics to their own has been called the “homophilia effect”<sup>37</sup>. To account for the non-random selection of individuals and the possible over-representation of individuals with given characteristics in the study population, the recruitment is modeled by a Markov process<sup>22,38</sup>.

A Markov process is a given phenomenon that can be classified in finite and discrete states, and in which the probability of transition between such states, also in a discrete time interval, only depends on the current and previous state<sup>39</sup>. And furthermore, according to the “law of large numbers”<sup>40</sup>, the probability of the process in any state over the course of a large number of steps will be independent of the state in which it began. In the case of recruitment in the RDS method, this means that memory of the recruitment occurs wave-to-wave, that is, the characteristics in the recruited individual only depend on the characteristics of his/her direct recruiter, and not on those of the recruiter's recruiter or of any participant in previous waves. After a sufficient number of waves, the characteristics of the individuals in the final sample are independent of the seed's characteristics.

Theoretically, the suggested number of waves for reaching equilibrium is six, since according to Watts<sup>41</sup>, all the individuals are connected to each other through only a half dozen links. However, if the homophilia effect is large, the sampling process needs to consist of a sufficiently large number of waves in order to reach Markov equilibrium (for given variables) and the minimum estimated sample size.

In the RDS method, recruitment is done on the basis of a set number of coupons (usually 2 to 4) distributed to participants in order to invite their peers. To participate in the study, the invited individuals must present their coupons as a guarantee of the network recruitment. Each coupon has a serial number which is used later for the social network design and recruitment patterns. Distribution of a set number of coupons to all the participants decreases the possibility that recruiters with large social networks will be over-represented in the sample as compared to those with small networks. It also reduces the initial influence of the seeds on the final sample and stimulates long recruitment chains, increasing the sample's power to capture "hidden" individuals in the study population<sup>25,38</sup>.

The RDS method also draws on the strategy of giving incentives to the participants. The first or "primary" incentive is given to participants when they complete their participation in the study. The second, called the secondary incentive, is given to participants for each peer successfully recruited into the study. Incentives have been used to promote the recruitment and thus the growth of the networks. They can be monetary or not, as for example non-perishable foods or food stamps, or tickets to shows or nightclubs. Choice of the type of incentive depends on the target group's characteristics.

In relation to the study venue, it is important that the study be carried out in an easy-to-reach place for the participants and that it is free of prejudices. Health services are often chosen, in order to promote their use by the study population.

### Analysis of data collected with RDS

In 2004, Salganick & Heckathorn<sup>22</sup> proposed that data collected with RDS be weighted according to an expansion factor based on the inverse of the selection probability. Under the hypothesis that the seed is selected with probability proportional to her/his network size, and that participants randomly recruit their peers from their network of friends and acquaintances, the authors demonstrated that the probability of selecting each participant is also proportional to the network size, i.e., to the number of acquaintances known in the target population.

In order to measure the effect of the sampling plan on the variance of the mean estimates, one uses the so-called "design effect", which is calculated as the ratio between the estimate of the variance determined by the sampling plan and the estimate of the variance obtained by a simple random sample of the same size. Design effect is

also used to calculate the sample size of future studies that may use the same complex sampling design<sup>42</sup>.

In the case of samples collected with RDS, Salganick<sup>43</sup> initially estimated design effects of around 2, through "bootstrap" simulation methods, whose general principle is to use the observed sample to generate a set of identical samples and produce a set of replicas of the estimates, thereby allowing calculation of the variance and confidence intervals. Later, simulation models that take into account the homophilia effect and dependence of observations showed much larger design effects, greater than 4<sup>44</sup>. In 2010, the same authors criticized the previously proposed estimators to calculate the variance and confidence intervals<sup>45</sup>.

For statistical analysis of data collected with RDS, an application was developed called "Respondent-Driven Sampling Analysis Tool" (RDSAT – [http://www.respondentdrivensampling.org/reports/RDSAT\\_56\\_Manual.pdf](http://www.respondentdrivensampling.org/reports/RDSAT_56_Manual.pdf), accessed on 14/Aug/2009), which weighs and adjusts the data according to the recruitment patterns observed during the study, considering the homophilia effect and the size of each individual's social network. However, the application presents important limitations, like the impossibility of performing multivariate analyses.

Since RDS is a complex and relatively new sampling method, the appropriate statistical techniques for analyzing the collected data are still being developed. For example, over time, two distinct estimators have been proposed for estimating prevalence rates, which have been discussed in the recent literature on RDS<sup>46</sup>. While the point estimators are quite similar, some questions remain about the best way to estimate variance<sup>47</sup>.

### Implementation of RDS among female sex workers in Brazil

The study, called the "Health Chain", was carried out in Brazil from August 2008 to July 2009, aimed at estimating the prevalence rates for HIV and syphilis and identify knowledge, attitudes, and practices related to HIV infection and other sexually transmitted diseases in female sex workers. The research project was approved by the Institutional Review Board of the Oswaldo Cruz Foundation (case no. 395/07).

To conduct the research, the Department of STD, AIDS, and Viral Hepatitis chose 10 Brazilian municipalities (counties) based on the size and importance of the local AIDS epidemic. The sample size (2,500 women) was calculated by es-

timating a 6% HIV prevalence rate, with a 95% confidence interval, and two-tailed error of 1.5%, considering a design effect of 1.5. In each of the municipalities, the attempt was made to distribute the sample proportionally to the municipality's population, while setting a minimum sample of 100 women. Table 1 shows the distribution of the planned and collected sample in the 10 municipalities.

Table 1

Planned and actual (collected) sample size in the ten municipalities in the study on respondent-driven sampling in female sex workers. Brazil, 2009.

Municipalities	Sample size	
	Planned	Collected
Manaus	200	199
Recife	200	237
Salvador	300	260
Campo Grande	150	147
Brasília	300	308
Belo Horizonte	300	289
Santos	150	191
Rio de Janeiro	600	601
Curitiba	200	201
Itajaí	100	90
<b>Total</b>	<b>2,500</b>	<b>2,523</b>

The eligibility criteria were the following: (1) age 18 years or older; (2) being a woman; (3) working as a sex worker in the municipality where the study was being conducted; (4) having had sex in exchange for money at least once in the previous four months; (5) accept participating under the study conditions and agree to sign the free and informed consent form; (6) present a valid invitation to participate in the study; (7) not having participated in the study previously; and (8) not being under the influence of illegal drugs or alcohol at the time of the interview.

The study consisted of a preliminary interview to verify the eligibility criteria and characterize the network and work venue, a self-completed computer questionnaire, and rapid tests for syphilis and HIV. In all ten municipalities, the fieldwork was done in health services, with the teams consisting of: a supervisor, in charge of organizing the field work in the health service, administering the invitations, and supervising

the teamwork in the service; an interviewer, in charge of applying the behavioral questionnaire; a counselor, in charge of pre- and post-test counseling; and the testing professional, a health worker with university training in a biomedical field, responsible for applying and interpreting the HIV and syphilis tests. The research teams were also responsible for monitoring the networks and encouraging participation by the sex workers. For this purpose, promotional materials for the study were prepared and distributed in the main prostitution venues in the target cities, with the following slogan: "If you receive an invitation from the Health Chain, be sure to participate. Help form the chain of health".

Before the data collection phase, in each municipality a preparatory survey was done to facilitate the study's implementation at each site according to the characteristics of prostitution in each city. The following issues were addressed: identification of the main prostitution venues, diversity of types of female sex workers present in the city, choice of the health service for conducting the study, and identification of possible seeds. In the preparatory survey, focus groups and/or in-depth interviews were held with the sex workers and their local representatives, using a previously prepared script and qualitative methodology<sup>48</sup>.

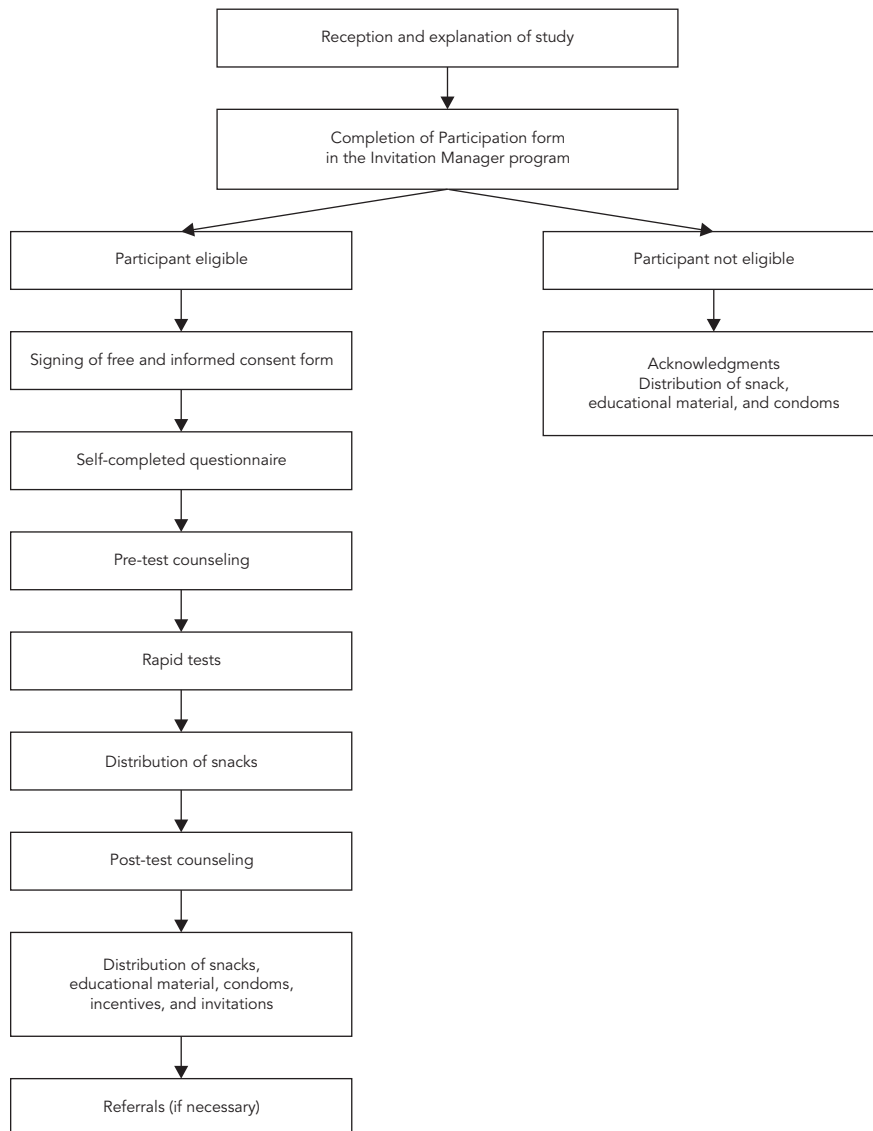
In each municipality, five to ten initial participants were chosen, called "seeds", selected non-randomly. Seeds were chosen with differences in age group, color/race, socioeconomic class, schooling, and types of work venues, like the street, nightclubs, saunas, hotels, call services, and others. In addition, this initial focus was on sex workers that were well-connected in the community and that had reported extensive social networks during the preparatory survey.

Each seed received three coupons to give to her friends or acquaintances. The women invited by the seeds that participated in the study comprised the first "wave". After participating in the interview, they received three coupons themselves to invite their friends and acquaintances. This process was repeated until the sample was reached in each site. Figure 1 shows the flow of participation by each sex worker, beginning with her arrival at the health service.

The coupons were numbered in order to prevent counterfeiting and to allow identifying the recruitment patterns in the population. The participants were only identified by the identification code on the coupon, consisting of two letters (the municipality's initials) and a set of numbers indicating the seed and the wave. The number of coupons handed out for each participant to recruit her peers was predetermined (three). In

Figure 1

Participation stages in the Chain of Health project.



order to manage the invitations, a program was generated on ACCESS (Microsoft Corp., USA), called "Invitations Manager".

The study questionnaire was self-completed and included questions on the following themes: (1) socio-demographic characteristics (age, schooling, marital status, color/race, religion, economic class, etc.); (2) knowledge, attitudes, perceptions, and risk practices for HIV and other STI; (3) HIV transmission routes and forms of

prevention; (4) history of HIV/STI testing, knowledge of serological status; (5) history of STI and presentation of symptoms; (6) sexual initiation; (7) number and types of sex partners (men, women; occasional, stable, commercial); (8) condom use (last sexual relations, regularity, unprotected relations with different types of partners and different types of sexual practices); (9) characteristic of sex trade; (10) drug and alcohol use; and (11) access to preventive activities.

The questionnaire was filled out directly by the interviewee on a computer using a program called Audio Computer-Assisted Self-Interview (ACASI – Tufts University, Boston, USA. <http://www.tufts.edu/med/nutrition-infection/tnc-cdaar/acasi/acasi.html>). This program allowed the interviewee to listen to all the questions and possible answers, thus facilitating understanding of the questionnaire and simultaneously guaranteeing confidentiality.

To allow calculating the prevalence rates for HIV and syphilis among female sex workers, rapid tests were performed (using capillary blood samples) in the participants, complying with the guidelines and protocols of the Department of STD, AIDS, and Viral Hepatitis, with pre- and post-test counseling by the team counselor. The health workers that performed the rapid tests had university training in a health field and had been specifically trained by the Department of STD, AIDS, and Viral Hepatitis before the study began.

The participants that tested positive received counseling in order to minimize the psychological risks and notify the steady partners, having been referred to health services under the Unified National Health System (SUS) for the appropriate follow-up according to the guidelines of the Department of STD, AIDS, and Viral Hepatitis. The services for referral of positive cases in each participating municipality were identified before launching the fieldwork.

During the research project's implementation, each participant received the primary incentive when she completed the interview, plus an additional reward (secondary incentive) in the form of cash (varying from BR\$5.00 to BR\$10.00) for recruiting each peer that successfully completed the study. The primary incentive consisted of condoms, lubricant, educational material, food stamps, and transportation vouchers, a snack along with the interview, and a little prize picked among a variety of beauty products, like a hairbrush, make-up, or manicure kit.

As for the statistical analysis, the data were weighted using the following question to measure the network size: "*How many female sex workers do you know personally?*". In each municipality, the expansion factors were inversely proportional to the network size. In addition, since the study was done in 10 municipalities, in order to produce results for the entire sample, calibration was performed according to the relative number of women 15-49 years of age in each municipality, assuming a constant proportion of female sex workers in all the sites studied. Interestingly, unlike other studies that used the RDS methodology, the statistical analysis was done for entire sample, with the data collected in the ten

municipalities. The methods proposed for data analysis, including estimation of the variance and design effect in the prevalence rates, as well as the multivariate logistic regression models, are beyond the scope of the current article and will be presented in subsequent publications.

Concerning the possible selection bias described by Heckathorn in 2002<sup>38</sup>, in the case of the Chain of Health project, the homophilia effect occurred according to HIV serological status, since participants with HIV infection tended to recruit other HIV-positive participants. These findings indicate a great dependency in the observations, which needs to be considered in the data analysis. By way of illustration, Figure 2 shows the network of female sex workers participating in the study in the city of Rio de Janeiro, where this tendency can be observed.

## Discussion

RDS is a method based on various statistical assumptions in order to reach a representative sample of the study population. Despite its potential to produce probabilistic samples, since it is a relatively recent method, it has been analyzed and rethought with each new application of the procedure<sup>20,21,22,25,37,43,49</sup>.

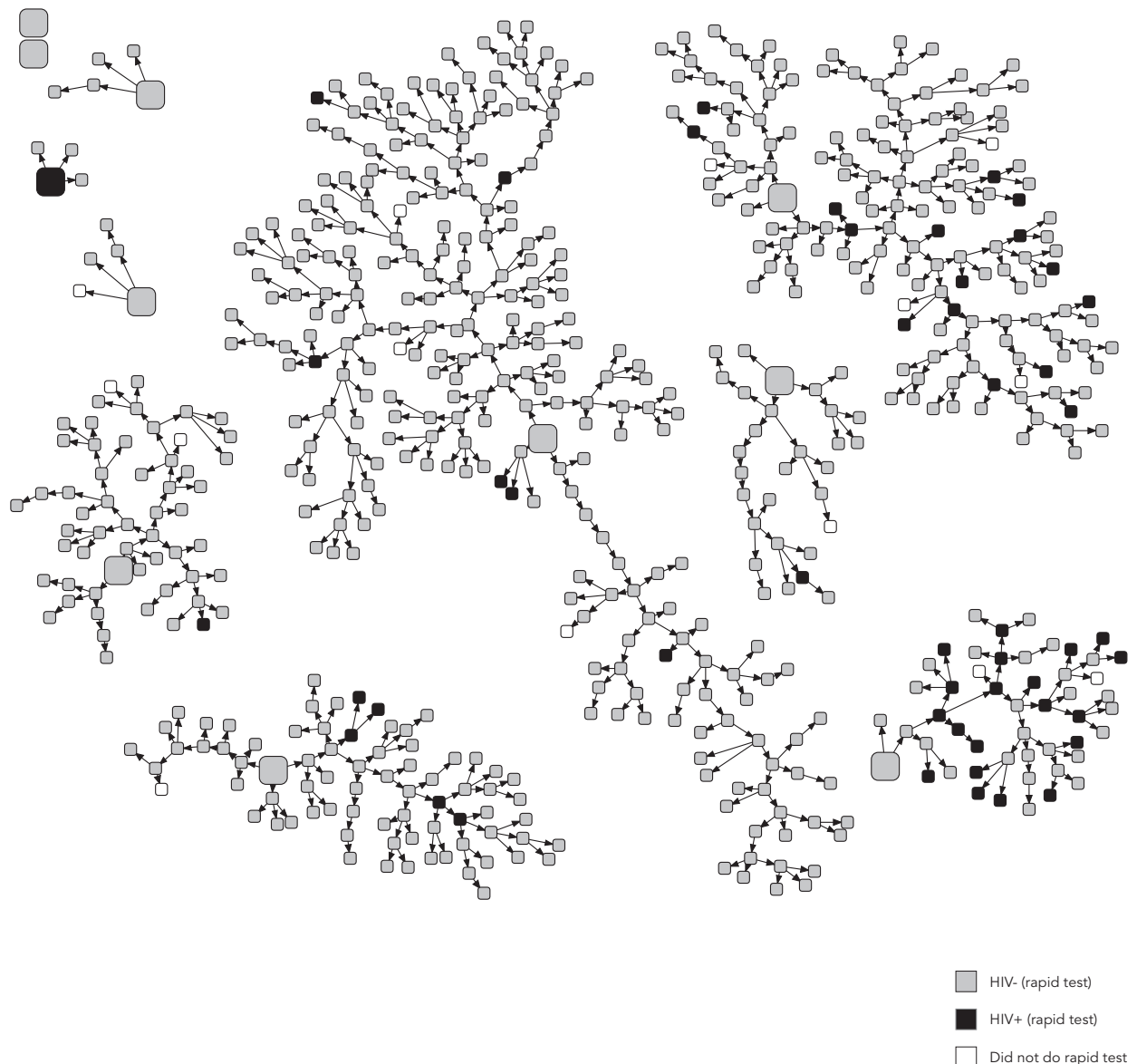
In Brazil, besides making it possible to fill in existing information gaps on the population of female sex workers, the study enabled evaluating the experience in the application of the RDS sampling method in this population subgroup in ten Brazilian municipalities.

In relation to the incentives, the primary incentive was used since the beginning of the study in all the sites. The secondary incentive was not used at the beginning of the study, on the assumption that there would be spontaneous interest in participating, based on the concepts of solidarity and the social networks among the female sex workers. However, in the majority of the participating municipalities, it proved necessary to use the secondary incentive in order to bolster participation in the study, since the networks were developing too slowly and there were constraints on time and resources to conduct the research. According to Heckathorn<sup>25</sup>, while the primary incentive motivates individual participation based on autonomous decisions in the concept of cooperation, the secondary incentive brings the social influence to the surface, which is taken advantage of in the sampling process.

However, provision of the secondary incentive has generated controversy in the international literature<sup>50</sup>. On the one hand, if the dual incentive stimulates both recruitment and

Figure 2

Distribution of female sex workers' network in the study in Rio de Janeiro, Brazil, according to HIV sero-prevalence.



\* Large symbols represent the seeds.

participation in the study, on the other it becomes a kind of parallel currency in groups with low socioeconomic status, generating duplicities in the sample, participation of non-eligible individuals, and even clashes among individuals in the population segment.

In the case of the Health Chain study, the use of the secondary incentive functioned to stimu-

late and speed up the inclusion of participants, and had the merit of obtaining the participation of "hidden", socially excluded sex workers, without the slightest awareness of the importance of their participation, but only interested in the incentives they were going to receive. However, although the process fostered the use of health services (even though artificially), the desire to

participate without meeting the eligibility criteria and the improper receipt of the incentive led to cases of physical assault against the research team.

In the application of RDS among female sex workers in Brazil, a link was also made with the social movement. In each of the municipalities in the study, focal points were chosen, that is, persons with influence in the community of sex workers, who helped to publicize the study and accompany the participants with invitations to the health service. The use of focal points helped overcome the distrust and fear among sex workers, who are routinely exposed to violence and social discrimination<sup>51</sup>.

As for the network size, the study showed that some seeds take a long time to “germinate”, while some types germinate faster than others, interfering in the sample’s diversity. In addition, due to the homophilia effect, a long network does not always guarantee capturing the diversity of types of individuals in the population segment, an essential characteristic for the sample to be representative of the target population. In this sense, the data weighting is essential for the sample to be representative of the population group.

As for the data weighting in the current study, the size of the network among sex workers varies according to the work venue, and is smaller among women that work the streets as compared to other places (nightclubs, saunas, hotels). Thus, the original weighting method will have to be improved, considering the possibility of stratification by work area.

## Resumo

*O grupo das mulheres profissionais do sexo é reconhecido, nacional e internacionalmente, como uma população de maior risco à infecção pelo HIV pela sua vulnerabilidade social e por fatores relacionados à atividade profissional. Porém, a utilização de estratégias convencionais de amostragem em estudos nos subgrupos de maior risco de contraírem HIV é, em geral, problemática, já que estes têm pequena magnitude em termos populacionais e estão vinculados a comportamentos estigmatizados ou atividades ilegais. Em 1997, foi proposta uma técnica de amostragem probabilística para populações de difícil acesso, o Respondent-*

*Driven Sampling (RDS). O método é considerado como uma variante da amostragem em cadeia e possibilita a estimação estatística dos parâmetros de interesse. Este artigo descreve alguns pressupostos do RDS e todas as etapas de implementação em um estudo com 2.523 profissionais do sexo em dez municípios brasileiros. A utilização do RDS mostrou-se apropriada para recrutamento de profissionais do sexo, possibilitando a seleção de uma amostra probabilística e a coleta de informações inexistentes sobre este grupo no Brasil.*

In relation to the homophilia effect, a tendency was observed among participants with HIV infection to also recruit HIV-positive peers. Considering that the great majority (75%) of HIV-positive participants were unaware of their serological status, this tendency is probably explained by common characteristics in their peer network, like age, prostitution site, type of client, social condition, price charged for sex, drug use, and other factors, and should be investigated in future studies. From the epidemiological point of view, while dependence of observations is interesting because it allows understanding the risk factors associated with the social networks, from the statistical point of view the structure of dependence should be considered when estimating variables and requires specific statistical techniques for the data analysis.

In sum, international literature, many authors have used traditional multivariate statistical techniques considering independent observations<sup>52,53,54,55,56,57,58</sup>, while few studies have taken into account the structure of dependence of observations in the data analysis<sup>29,59</sup>.

In short, the use of RDS proved appropriate for studying female sex workers in Brazil, since it facilitated recruitment and allowed the selection of a probabilistic sample in this population subgroup. On the other hand, the experience with the implementation of RDS leaves some challenges to be explored, like improvement of the methodology for statistical analysis of data collected with RDS, adequate data calibration, and interpretation of the social networks.

*HIV; Prostituição; Comportamento Sexual*



## Contributors

G. N. Damacena conceived the article's content and was responsible for elaborating the text, table, figures, and references. C. L. Szwarcwald conceived the article's content and participated in elaborating the text. A. Barbosa Júnior participated in elaborating the text.

## References

1. d'Cruz-Grote D. Prevention of HIV infection in developing countries. *Lancet* 1996; 348:1071-4.
2. van Vliet C, Meester EI, Korenromp EL, Singer B, Bakker R, Habbema JD. Focusing strategies of condom use against HIV in different behavioural settings: an evaluation based on a simulation model. *Bull World Health Organ* 2001; 79:442-54.
3. Joint United Nations Programme on HIV/AIDS. Monitoring the declaration of commitment on HIV/AIDS: guidelines on constructions of core indicators. Geneva: Joint United Nations Programme on HIV/AIDS; 2005.
4. Ministério da Saúde. Avaliação da efetividade das ações de prevenção dirigidas às profissionais do sexo, em três regiões brasileiras. Brasília: Ministério da Saúde; 2004. (Série Estudos, Pesquisas e Avaliação, 7).
5. Szwarcwald CL, Bastos FI, Gravato N, Lacerda R, Chequer PN, Castilho EA. The relationship of illicit drug use to HIV infection among commercial sex workers in the city of Santos, São Paulo, Brazil. *Int J Drug Policy* 1998; 9:427-36.
6. Szwarcwald CL, Barbosa Júnior A, Souza Jr. PRB, Lemos KR, Frias PG, Luhm KR, et al. HIV testing during pregnancy: use of secondary data to estimate 2006 test coverage and prevalence in Brazil. *Braz J Infect Dis* 2008; 12:167-72.
7. Carter KH, Harry BP, Jeune M, Nicholson D. HIV risk perception, risk behavior, and seroprevalence among female commercial sex workers in Georgetown, Guyana. *Rev Panam Salud Pública* 1997; 1:451-9.
8. Jamal N, Khushk IA, Naeem Z. Knowledge and attitudes regarding AIDS among female commercial sex workers at Hyderabad City, Pakistan. *J Coll Physicians Surg Pak* 2006; 16:91-3.
9. Wei SB, Chen ZD, Zhou W, Wu FB, Li SP, Shan JG. A study of commercial sex and HIV/STI-related risk factors among hospitality girls in entertainment establishments in Wuhan, China. *Sex Health* 2004; 1:141-4.
10. Singh TN, Kananbala S, Thongam W, Devi KhS, Singh NB. Increasing trend of HIV seropositivity among commercial sex workers attending the Voluntary and Confidential Counseling and Testing Centre in Manipur, India. *Int J STD AIDS* 2005; 16:166-9.

11. Shakarishvili A, Dubovskaya LK, Zohrabyan LS, St. Lawrence JS, Aral SO, Dugasheva LG, et al. Sex work, drug use, HIV infection, and spread of sexually transmitted infections in Moscow, Russian Federation. *Lancet* 2005; 366:57-60.
12. Hamers FF, Downs AM. HIV in Central and Eastern Europe. *Lancet* 2003; 361:1035-44.
13. Wasseheit JN. Epidemiological synergy: interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis* 1992; 19:61-77.
14. Bautista CT, Sanchez JL, Montano SM, Laguna-Torres A, Suarez L, Sanchez J, et al. Seroprevalence of and risk factors for HIV-1 infection among female commercial sex workers in South America. *Sex Transm Infect* 2006; 82:311-6.
15. Lau JT, Tsui HY, Siah PC, Zhang KL. A study on female sex workers in southern China (Shenzhen): HIV-related knowledge, condom use and STD history. *AIDS Care* 2002; 14:219-33.
16. Miguez-Burbano MJ, Angarita I, Shultz JM, Shor-Posner G, Klaskala W, Duque JL, et al. HIV-related high risk sexual behaviors and practices among women in Bogota, Colombia. *Women & Health* 2000; 30:109-19.
17. Programa Nacional de DST e AIDS, Secretaria de Vigilância em Saúde, Ministério da Saúde. Pesquisa de conhecimento atitudes e práticas na população brasileira de 15 a 54 anos, 2004. Brasília: Ministério da Saúde; 2006.
18. Ministério da Saúde. Profissionais do sexo: documento referencial para ações de prevenção das DST e da AIDS. Brasília: Ministério da Saúde; 2002. (Série Manuais, 47).
19. Silva NG, Peres CA, Morell MGGP, Areco KN. Fatores associados à infecção pelo HIV em trabalhadoras do sexo (TS) em Santos – SP. *Saúde Coletiva* 2007; 4:168-72.
20. Semaan S, Lauby J, Liebman J. Street and network sampling in evaluation studies of HIV risk-reduction interventions. *AIDS* 2002; 4:213-23.
21. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* 2005; 19 Suppl 2:S67-72.
22. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol Methodol* 2004; 34:193-240.
23. Family Health International. Evaluating programs for HIV/AIDS prevention and care in developing countries: a handbook for program managers and decision makers. Arlington: Family Health International; 2001.
24. Stueve A, O'Donnell LN, Duran R, Doval AS, Blome J. Time-space sampling in minority communities: results with young Latino men who have sex with men. *Am J Public Health* 2001; 91:922-6.
25. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl* 1997; 44:174-99.
26. Jarlais DCD, Arasteh K, Semaan S, Wood E. HIV among injection drug users: current epidemiology, biologic markers, respondent-driven sampling, and supervised-injection facilities. *Curr Opin HIV AIDS* 2009; 4:308-13.
27. Johnston L, Sabin K, Hien MT, Huong PT. Assessment of respondent-driven sampling for recruiting female sex workers in two Vietnamese cities: reaching the unseen sex worker. *J Urban Health* 2006; 83:16-28.
28. Liu H, Liu H, Cai Y, Rhodes AG, Hong F. Money boys, HIV risks, and the associations between norms and safer sex: a respondent-driven sampling study in Shenzhen, China. *AIDS Behav* 2008; 13:652-62.
29. Pollini RA, Alvelais J, Gallardo M, Vera A, Lozada R, Magis-Rodríguez C, et al. The harm inside: injection during incarceration among male injection drug users in Tijuana, Mexico. *Drug Alcohol Depend* 2009; 103:52-8.
30. Reisner S, Mimiaga MJ, Case P, Johnson CV, Safren AS, Mayer KH. Predictors of identifying as a barebacker among high-risk New England HIV seronegative men who have sex with men. *J Urban Health* 2008; 86:250-62.
31. Wattana W, van Griensven F, Rhucharoenpornpanich O, Manopaiboon C, Thienkrua W, Bannatham R, et al. Respondent-driven sampling to assess characteristics and estimate the number of injection drug users in Bangkok, Thailand. *Drug Alcohol Depend* 2007; 90:228-33.
32. Simic M, Johnston LG, Platt L, Baros S, Andjelkovic V, Novotny T, et al. Exploring Barriers to “respondent driven sampling” in sex worker and drug-injecting sex worker populations in Eastern Europe. *J Urban Health* 2006; 83:i6-15.
33. 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:i60-72.
34. 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.
35. Malekinejad M, Johnston LG, Kendall C, Kerr LRFS, 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.
36. Barbosa Júnior A, Pascom ARP, Szwarcwald CL, Kendall C, McFarland W. Transfer of sampling methods for studies on most-at-risk populations (MARPs) in Brazil. *Cad Saúde Pública* 2011; 27 Suppl 1:S36-44.
37. Heckathorn DD, Semaan S, Broadhead RS, Hughes JJ. Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18-25. *AIDS Behav* 2002; 6:55-67.
38. Heckathorn DD. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc Probl* 2002; 49:11-34.
39. Shamblin JE. Pesquisa operacional: uma abordagem básica. São Paulo: Editora Atlas; 1979.
40. Kemeny JG, Snell JL. Finite Markov Chains. 6th Ed. London: D. van Nostrand Company; 1960.

41. Watts DJ. Six degrees: the science of a connected age. New York: W.W. Norton & Company; 2004.
42. Szwarcwald CL, Damascena GN. Amostras complexas em inquéritos populacionais: planejamento e implicações na análise estatística dos dados. *Rev Bras Epidemiol* 2008; 11 Suppl 1:S38-45.
43. Salganik M. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J Urban Health* 2006; 83:i98-112.
44. Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Stat Med* 2009; 28:2202-29.
45. Goel S, Salganik MJ. Assessing respondent-driven sampling. *Proc Natl Acad Sci USA* 2010; 107:6743-7.
46. Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. *J Off Stat* 2008; 24:79-97.
47. 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.
48. Minayo MCS. O desafio do conhecimento. Rio de Janeiro: ABRASCO/São Paulo: Editora Hucitec; 1994.
49. Abdul-Quader AS, Heckathorn DD, Sabin K, Saidel T. Implementation and analysis of respondent driven sampling: lessons learned from the field. *J Urban Health* 2006; 83:i1-5.
50. Semaan S, Santibanez S, Garfein RS, Heckathorn DD, Jarlais DCD. Ethical and regulatory considerations in HIV prevention studies employing respondent-driven sampling. *Int J Drug Policy* 2009; 20:14-27.
51. Mazzeiro JB. Sexualidade criminalizada: prostituição, lenocínio e outros delitos – São Paulo 1870/1920. *Rev Bras Hist* 1998; 18:247-85.
52. Shahmanesh M, Wayal S, Copas A, Patel V, Mabey D, Cowan F. A study comparing sexually transmitted infections and HIV among ex-red-light district and non-red-light district sex workers after the demolition of Baina red-light district. *J Acquir Immune Defic Syndr* 2009; 52:253-7.
53. Bobashev GV, Zule WA, Osilla KC, Kline TL, Wechsberg WM. Transactional sex among men and women in the South at high risk for HIV and other STIs. *J Urban Health* 2009; 86 Suppl 1:S32-47.
54. Clark MA, Rogers ML, Armstrong GF, Rakowski W, Bowen DJ, Hughes T, et al. Comprehensive cancer screening among unmarried women aged 40-75 years: results from the cancer screening project for women. *J Women's Health (Larchmt)* 2009; 18:451-9.
55. Lauby JL, Millett GA, LaPollo AB, Bond L, Murrill CS, Marks G. Sexual risk behaviors of HIV-positive, HIV-negative, and serostatus-unknown Black men who have sex with men and women. *Arch Sex Behav* 2008; 37:708-19.
56. Wheeler DP, Lauby JL, Liu KL, Van Sluytman LG, Murrill C. A comparative analysis of sexual risk characteristics of black men who have sex with men or with men and women. *Arch Sex Behav* 2008; 37:697-707.
57. Deiss RG, Brouwer KC, Loza O, Lozada RM, Ramos R, Cruz MA, et al. High-risk sexual and drug using behaviors among male injection drug users who have sex with men in 2 Mexico-US border cities. *Sex Transm Dis* 2008; 35:243-9.
58. Platt L, Wall M, Rhodes T, Judd A, Hickman M, Johnston LG, et al. Methods to recruit hard-to-reach groups: comparing two chain referral sampling methods of recruiting injecting drug users across nine studies in Russia and Estonia. *J Urban Health* 2006; 83 Suppl 6:S39-53.
59. Strathdee SA, Lozada R, Pollini RA, Brouwer KC, Mantsios A, Abramovitz DA, et al. Individual, social, and environmental influences associated with HIV infection among injection drug users in Tijuana, Mexico. *J Acquir Immune Defic Syndr* 2008; 47:369-76.

---

Submitted on 21/Dec/2009

Final version resubmitted on 08/Jun/2010

Approved on 10/Aug/2010