



HHS Public Access

Author manuscript

Surv Pract. Author manuscript; available in PMC 2015 October 06.

Published in final edited form as:

Surv Pract. 2014 ; 7(1): 1–7.

Respondent-Driven Sampling of Heterosexuals at Increased Risk of HIV Infection

Neale Batra

Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention

Simone C. Gray

Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention

Nevin Krishna

Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention

Nikhil Prachand

Chicago Department of Public Health-STI/HIV Division

William T. Robinson

Louisiana Office of Public Health – STD/HIV Program, Louisiana State University Health Sciences Center in New Orleans School of Public Health

Cyprian Wejnert

Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD and TB Prevention

Sampling hard-to-reach populations can be difficult with traditional survey methods. Challenges arise because sampling frames are typically unknown, and individuals can be wary of authority or afraid of being identified by the stigmatized or illegal nature of their behaviors. A contemporary solution is respondent-driven sampling (RDS), a survey method spearheaded by Heckathorn (1997) that utilizes links in underlying social networks to create branching referral chains of respondents. Obtaining samples for the target population is heavily impacted by RDS parameters, such as field site and the location of the initial respondents (“seeds”). However, the relationship between the RDS parameters and the eventual sample is not yet thoroughly understood. This paper reviews spatial patterns in two RDS samples from Chicago and New Orleans to determine the relative importance of seed and field site location, as well as assess the impact of physical barriers.

RDS is a sampling technique based on the principle that individuals are better able to locate and recruit persons with similar characteristics to themselves through their own social networks. In RDS, seeds receive incentives to recruit eligible peers to the study. Each “wave” of respondents visits an established field site for eligibility screening and interviews. Respondents are then compensated and given incentives to continue recruiting the next wave of eligible peers to the study (Abdul-Quader and Heckathorn 2006).

Although RDS has been shown to be widely applicable, social structure within a population has been found to significantly bias RDS results (Goel and Salganik 2009). The effect of nonrandom selection in personal networks has also been cited as a possible study design flaw (Toledo et al. 2011; Wang et al. 2007). Similarly, relationships between geographic distribution and sample characteristics have not been adequately explored and even less is known about how varying parameters of RDS study design such as interview field site and seed location affect the respondent sample distribution (Doreian and Conti 2012; Rudolph et al. 2010; Toledo et al. 2011).

In this paper, we explore some of the aforementioned limitations and present findings from an RDS study used to recruit individuals considered at increased risk for HIV in Chicago and New Orleans. We concentrate on the respondent patterns in relationship to seed residence and the placement of the field site.

The data for this study were obtained from the second cycle of the National HIV Behavioral Surveillance which recruited heterosexuals at highest risk for HIV infection (Gallagher et al. 2007). The study was completed in 2010 and our analyses focused on the metropolitan statistical areas of Chicago and New Orleans. Study participants had to meet eligibility criteria based on residency, age, income, education, drug use history, and recent heterosexual sexual activity. Participants were interviewed about their behaviors, offered an HIV test, and received monetary compensation and incentives for further recruitment of their peers (average of \$25 for the survey and test, and \$10 per eligible referral recruited to the project).

Seeds were selected to be diverse with regard to locally relevant characteristics, typically: race/ethnicity, gender, and/or age. Additionally, seeds were also required to be residents of high-risk areas (HRAs; poverty areas created by selecting census tracts where household poverty is >20 percent), be knowledgeable about the community, and to have large social networks. Seeds were asked to recruit 1–5 eligible persons from their social network. The process was repeated with eligible respondents until the final sample size of mapped respondents was 547 in Chicago and 582 in New Orleans.

Interview field site locations were chosen based on proximity to an HRA, there being no real or perceived barriers to visiting the field site (confidentiality of participants should not be compromised), close proximity to public transportation, and not compromising staff safety. Figure 1 shows maps of the two study regions with seed and field site locations as well as the distribution of median household income by census tract. The New Orleans study had one centralized field site in the relatively low-income downtown area. In contrast, the Chicago study had five field sites covering areas with varying levels of income distribution. There were five seeds in New Orleans located within 5 miles of the field site, while the 14 seeds in Chicago had wide coverage of target HRAs. Participant residence was determined by having them point to their residence on a map overlaid with census tracts. The data were then geocoded and analyzed using ArcGIS (ESRI, Version 9.3.1), and residence was mapped based on the geocentroid of their census tract.

Maps of the resulting geographic distribution of respondents for both New Orleans and Chicago revealed a high level of clustering around field sites (Figure 2). In New Orleans, the highest densities of respondents were immediately surrounding the only field site. In Chicago, respondents were tightly clustered around three main field sites: the two west side sites of South Austin and Breaking Ground, and the Liberation Christian Center site towards the southern side. Less clustering was apparent around the two easternmost sites. This trend, however, masks the fact that many respondents in Chicago did not visit the site closest to their residence. Figure 3 shows residences of the Chicago respondents by the field sites they visited. The spatial distribution of respondents shows that respondents often lived in a different area of the city than the field site they visited (Figure 3). This could be explained by Chicago's schedule of five rotating field sites and 24-h delay on coupon activation. Respondents living in the immediate area of a field site appear to have traveled to another site open the following day, instead of waiting an entire week to redeem their coupon. This observation could direct future RDS study research and design.

It is common for only a subset of the seeds chosen for RDS studies to actually produce long recruitment chains; such chains are termed "productive chains." Productive chains were defined as those containing over 15 individuals. When mapped by productive chain, respondents revealed the broad reach of a single seed. New Orleans had three productive chains representing 99 percent of the sample; two of which represented 93 percent of the sample. Chicago had two productive chains representing 95 percent of the sample. The geographic coverage of productive chains in both cities differed little by chain size, but still appeared subject to the social network accessed. In New Orleans, the smallest productive chain (n=37) reached the same targeted census tracts as the largest chain (n=289). Similarly, in Chicago, a small chain of 82 participants covered all the target HRAs and classically low-income areas that were covered by a larger chain of 437 participants. The larger chain only added more dense coverage around certain field sites, likely related to that particular seed. This suggests that while productive chains generally reach all accessible HRAs, the extent to which an area is sampled can depend greatly on the social networks being accessed.

An examination of the Chicago field sites emphasizes the important role field site location plays in reaching a target population. The sites were all located in or near HRAs, with the exception of the catch-all Lakeview site in downtown Chicago. The Breaking Ground site, located directly in an HRA, produced a sample that included hundreds of respondents from within that HRA (Figures 1 and 3). The Matthew House and Liberation Christian Center sites in the southern region of Chicago were not able to be located in the center of an HRA and instead operated in gentrifying neighborhoods in the HRA's periphery. Although many of the seeds were placed within the targeted HRA (Figure 1), the distribution of respondents in Figure 2 shows that few respondents actually resided in the southern HRA between these two sites. Instead, respondent residences tend to be clustered around the location of the field sites, suggesting that proximity to a site is an important predictor of study participation (Figure 2).

Our observations suggest that geographical barriers can constrain sample distribution. In New Orleans, there is a marked drop-off in respondents south of the Mississippi River. This river in New Orleans has two bridges and one ferry, presenting a significant barrier for low-

income respondents traveling north to the single field site. This is contrasted with the situation in Chicago, where the Chicago River is spanned by 38 bridges and the vast majority of the city is crisscrossed by the extensive public transit system. Without a geographical barrier, chains originating on the west side of Chicago sampled extensively from HRAs across the city.

The purpose of this study was to start a conversation about spatial patterns in relation to field site and seed location for a population of RDS respondents considered at increased risk for HIV. While our results are ultimately descriptive, the spatial relationships suggested by the maps lead to conclusions important for future RDS implementation of this population. Our evidence suggests that in the absence of geographical barriers, field site location is an important determinant of respondent distribution, much more so than seed location. With due consideration to constraints such as staff safety, field sites should be located directly in target areas. Respondents are generally willing to travel to open field sites that are located 8–15 miles away from their residence in order to participate in the study, even if there was a field site within 5 miles of their residence that would be open on an alternative date. Further research is needed to extrapolate these results to other H2R populations and regions.

References

- Abdul-Quader A, Heckathorn D. Effectiveness of respondent-driven sampling for recruiting drug users in New York City: findings from a pilot study. *Journal of Urban Health*. 2006; 83(3):459–476. [PubMed: 16739048]
- Doreian P, Conti N. Social context, spatial structure and social network structure. *Social Networks*. 2012; 34(1):32–46.
- Gallagher KM, Sullivan PS, Lansky A, Onorato IM. Behavioral surveillance among people at risk for HIV infection in the U.S.: The National HIV Behavioral Surveillance System. *Public Health Reports*. 2007; 122(Suppl 1):32–38. [PubMed: 17354525]
- Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Statistics in Medicine*. 2009; 28:2202–2229. [PubMed: 19572381]
- Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems*. 1997; 44(2):174–199.
- Rudolph AE, Crawford ND, Latkin C, Heimer R, Benjamin EO, Jones KC, Fuller CM. Subpopulations of illicit drug users reached by targeted street outreach and respondent-driven sampling strategies: implications for research and public health practice. *Annals of Epidemiology*. 2010; 21(4):280–289. [PubMed: 21376275]
- Toledo L, Codeco CT, Bertoni N, Albuquerque E, Malta M, Bastos FI, Brazilian Multicity Study Group on Drug Misuse. Putting respondent-driven sampling on the map: insights from Rio de Janeiro, Brazil. *Journal of Acquired Immune Deficiency Syndrome*. 2011; 57(Suppl 3):S136–S143.
- Wang J, Falck RS, Li L, Rahman A, Carlson RG. Respondent-driven sampling in the recruitment of illicit stimulant drug users in a rural setting: findings and technical issues. *Addictive Behaviors*. 2007; 32(5):924–937. [PubMed: 16901654]

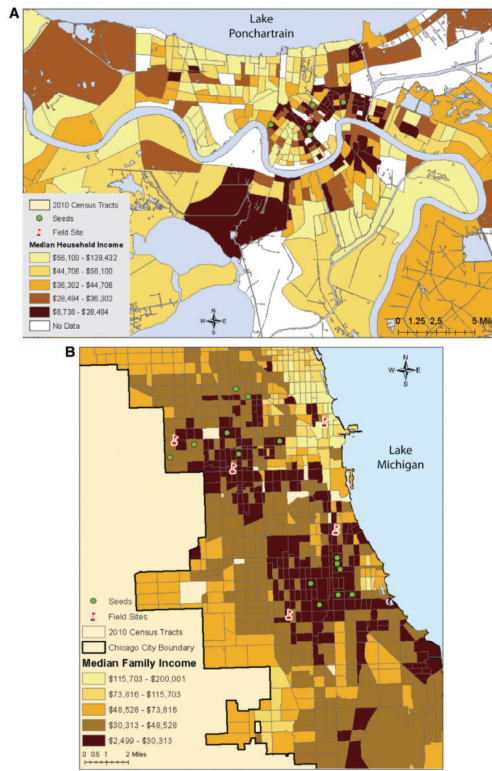


Figure 1.
Field site, income, and seeds in (A) New Orleans and (B) Chicago

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

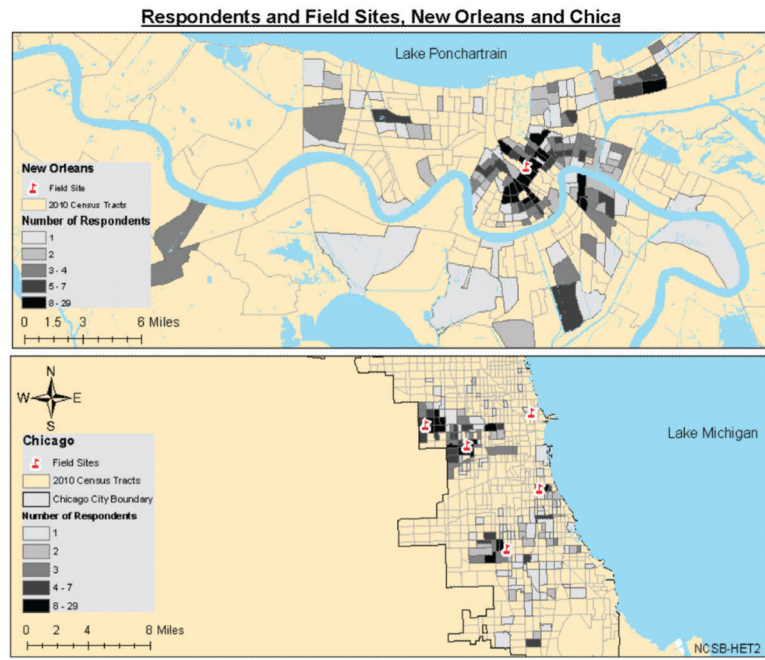


Figure 2. Distribution of respondents in New Orleans and Chicago

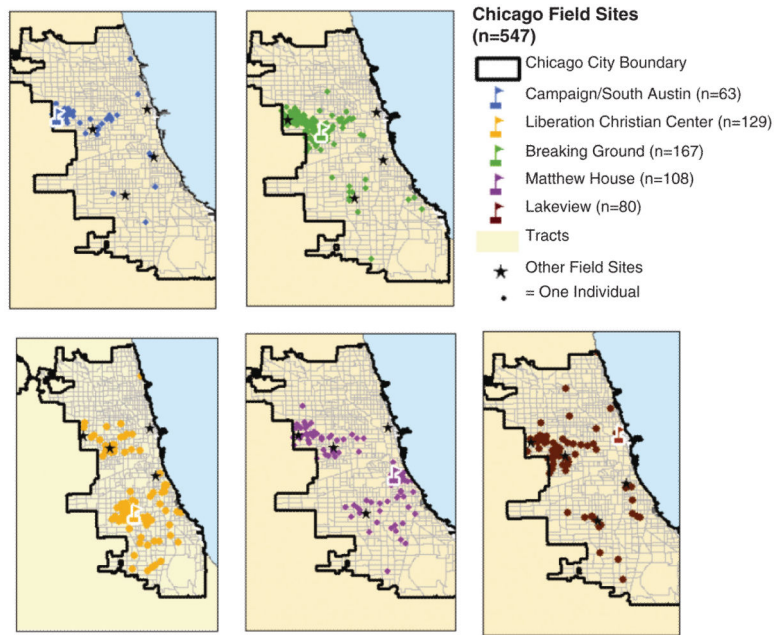


Figure 3.
Chicago respondents by the field site they visited

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript