Evaluating Consistency in Repeat Surveys of Injection Drug Users Recruited by Respondent-Driven Sampling in the Seattle Area: Results from the NHBS-IDU1 and NHBS-IDU2 Surveys

RICHARD D. BURT, PhD, AND HANNE THIEDE, DVM, MPH

PURPOSE: We compared data from two respondent-driven sampling (RDS) surveys of Seattle-area injection drug users (IDU) to evaluate consistency in repeat RDS surveys.

METHODS: The RDS-adjusted estimates for 16 key sociodemographic, drug-related, sexual behavior, and HIV- and hepatitis C virus-related variables were compared in the 2005 and the 2009 National HIV Behavioral Surveillance system surveys (NHBS-IDU1 and NHBS-IDU2). Time trends that might influence the comparisons were assessed by the use of data from reported HIV cases in IDU, surveys of needle exchange users, and two previous IDU studies.

RESULTS: NHBS-IDU2 participants were more likely than NHBS-IDU1 participants to report older age, heroin as their primary injection drug, male-to-male sex, unprotected sex with a partner of nonconcordant HIV status, and to self-report HIV-positive status. NHBS-IDU2 participants were less likely to report residence in downtown Seattle, amphetamine injection, and a recent HIV test. Time trends among Seattle-area IDU in age, male-to-male sex, and HIV testing could have influenced these differences.

CONCLUSIONS: The number and magnitude of the estimated differences between the two RDS surveys appeared to describe materially different populations. This could be a result of changes in the characteristics of Seattle-area IDU over time, of accessing differing subpopulations of Seattle IDU, or of high variability in RDS measurements.

INTRODUCTION

Surveys of injection drug users (IDU) can provide information on the status of the HIV epidemic in this population and on behaviors that affect the likelihood of HIV transmission that can be useful in guiding prevention efforts. However, it is difficult to recruit a survey population that can be said with confidence to represent the IDU in any particular community (1, 2). In the late 1990s respondent-driven sampling (RDS) was introduced as a variation on traditional snowball sampling as a means of obtaining unbiased estimates of characteristics of hidden study populations (3, 4).

In RDS, participants are issued coupons with which to recruit their peers and are paid for each person recruited. The new recruits are in turn given coupons with which to recruit further waves of participants. Via the use of data obtained during interview, RDS incorporates adjustments for its estimates of population characteristics on the basis of the size of participants' networks and by differential recruitment patterns within and across groups of interest, determined by information of who recruited whom. Mathematical theory and results from modeling studies have asserted that statistically unbiased estimates of population characteristics can be calculated in an RDS-recruited study population (5, 6).

RDS estimates necessarily incorporate certain assumptions, notably that each member of participants' relevant social networks is recruited into the study with equal probability and that participants describe the size of their social networks accurately and consistently. However, the assumptions underlying both modeling and theoretical results may not be fulfilled in practice. Modeling studies have suggested the potential for bias in RDS-adjusted estimates (7) and found high variance, and thus low power, in RDS results (8).
Appendix: 3.7.3

Selected Abbreviations and Acronyms

IDU = injection drug user
RDS = respondent-driven sampling
NHBS = National HIV/AIDS Behavior Surveillance
HARS = HIV/AIDS Reporting System
MSM = men who have sex with men
HCV = hepatitis C virus
NEX = needle exchange

a study modeling the effects of deviations from RDS assumptions, investigators found that RDS estimators are subject to error insofar as participants are more likely to recruit persons in their networks, who they are closer to or whom they more closely resemble, both likely propositions in practice (9).

Empirical results assessing RDS in real-world settings can offer a useful complement to the evaluation of RDS methodology. One means of evaluation is to investigate the consistency of RDS measurements in repeated surveys. In 2005 and 2009, the National HIV/AIDS Behavior Surveillance (NHBS) system used RDS to recruit IDU in several U.S. cities, including Seattle (10). We report here a comparison of results from the two surveys in the Seattle area for sociodemographic and HIV transmission-associated variables commonly used to describe IDU populations. To help assess the potential influence of time trends on our findings, we include data from the HIV/AIDS Reporting System (HARS) on newly diagnosed HIV cases in IDU, from repeat surveys of Seattle-area needle exchange users, and from two previous local studies of IDU: the RAVEN and Kiwi studies.

METHODS

Recruitment in NHBS-IDU2

Recruitment methods in NHBS-IDU2 have been described in print (11). Initial seeds were chosen to broadly represent the racial, gender, men who have sex with men (MSM) status, drug preference, and residential diversity of Seattle-area IDU. Eligibility requirements mandated that participants be 18 years of age or older, demonstrate evidence of injection by physical signs or detailed knowledge of injection practices, be able to complete the survey in English, and reside in King or Snohomish Counties.

A 30- to 40-minute face-to-face interview was conducted in study offices in downtown Seattle with the use of a handheld computer. HIV testing was performed with a rapid test on a fingerstick sample (OraSure Technologies, Bethlehem PA) or a standard test on a venous blood sample (Bio-Rad, Hercules, CA). Confirmatory testing by Western Blot was performed on a blood sample (Bio-Rad) for those with reactive initial tests. Participants were paid $25 for the interview, $15 for the HIV test, and $10 for each coupon returned by an eligible participant. Study procedures were reviewed and approved by the Washington State Institutional Review Board.

Differences in Recruitment Methods Between NBHS-IDU1 and NHBS-IDU2

NHBS-IDU1 followed a protocol similar to that of NHBS-IDU2 (12). The primary interview site in NHBS-IDU1 was also in downtown Seattle, although in a different location. NHBS-IDU1 also recruited one or two days a week in offices in south King County, which produced only 10% of participants. All participants in NHBS-IDU1 were issued three coupons. NHBS-IDU2 used schemes of offering from one to five coupons to participants on the basis of their age and residence to enhance participation of younger IDU and IDU living outside of downtown, a tactic with precedent in the RDS literature (13). The NHBS-IDU1 study questionnaire was quite similar (and in most variables identical) to that of NHBS-IDU2 for the variables we are reporting. Neither HIV nor hepatitis C (HCV) testing was offered in NHBS-IDU1, whereas all NHBS-IDU2 participants were offered an HIV test, and those recruited after July 22, 2009 were offered HCV testing.

Data from HIV Surveillance and Local Studies

We assessed consistency across data sources and time trends in variables of interest using data from HARS on persons reported with HIV in King County 1990-2009, the RAVEN study (recruited from a collection of institutional settings 1994-1997) (14), the Kiwi study (recruited from King County jails 1998-2002) (15), and from five Public Health surveys (in 2003, 2004, 2006, 2009, and 2011) of persons exchanging needles at King County needle exchanges (16). Because there were no participants in NHBS-IDU1 from Snohomish County and only two in NHBS-IDU2, analysis was restricted to King County residents in all data sources, and to persons 18 years of age or older. All data sources include MSM/IDU.

Statistical Analysis

We used RDSAT to produce adjusted estimates of the population proportions of characteristics of interest (17). For statistical evaluation of our results, we compared overlap in RDSAT-derived 95% confidence intervals (18). We also evaluated significance in univariate logistic regression analyses incorporating individual weights for study participants. These weights reflect adjustments for differential recruitment characteristics and network size. We combined the two study populations, with RDSAT-generated individual weights, and conducted logistic regression analyses using study (NHBS-IDU1 or NHBS-IDU2) as the dependent variable and the variable of interest as the independent variable. Multivariate weighted logistic regression analyses used a dichotomized
variable of interest as the dependent variable with study and potentially confounding variables as independent variables.

Where the graphical representation suggested the existence of long terms across data sources, we used a linear-by-linear $\chi^2$ statistic to evaluate the trend across all years covered by available data. Because our statistical measures may well overestimate significance (8), we present exact $p$-values, even when very small, as a means of judging the likelihood that the differences observed would truly attain significance. A $p$-value less than $10^{-3}$ was used as a threshold to identify results to highlight. With two exceptions (noted in the text) the RDSAT 95% confidence intervals of these highlighted results overlapped in the two surveys. Analyses were conducted in SPSS (19).

RESULTS

NHBS-IDU2 Recruitment

In NHBS-IDU2, 497 participants were recruited who met the eligibility requirements for the current analysis between June 17 and November 25, 2009. These derived from six initial seeds (Fig. 1). The majority of participants (70%) derived from one seed. There were 16 waves of recruitment. Six participants reported being recruited by a stranger. Of the 1333 coupons distributed, 652 (49%) were returned by a potential participant.

The difference between the sample population proportions and the estimated equilibrium proportions (an indication of the degree to which the sample is independent of the characteristics of the seeds) was $\leq 0.02$ for each category of the variables listed in Tables 1 and 2 (Tables S1 and S2). The most pronounced levels of homophily (which measures the extent to which participants with a certain characteristic preferentially recruited persons sharing that characteristic) were seen for amphetamine injection (0.72), having an HIV test in the previous 3 months ($-0.56$), male-to-male sex (0.53), heroin injection (0.43), and black race (0.41).

Sociodemographics

There were differences in the age distribution of the NHBS-IDU1 and NHBS-IDU2 populations ($p = 4 \times 10^{-4}$), with NHBS-IDU2 participants estimated to be more likely to be 50 years of age or older compared with NHBS-IDU1 participants (42% vs. 28%) (Table 1). After the introduction of differential coupon distribution in NHBS-IDU2, the estimated proportion of participants older than 50 decreased from 58% to 37%. Thus, if anything, the difference in the proportion of older participants between...
TABLE 1. Comparing sociodemographic characteristics of Seattle-area IDU participating in the NHBS-IDU1 and NHBS-IDU2 surveys

<table>
<thead>
<tr>
<th></th>
<th>NHBS-IDU1 (n=368)</th>
<th>NHBS-IDU2 (n=497)</th>
<th>p-value (with RDS weighting)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–29</td>
<td>12 (6–21)</td>
<td>9 (5–13)</td>
<td>4 x 10^{-4}</td>
</tr>
<tr>
<td>30–39</td>
<td>21 (13–31)</td>
<td>21 (15–27)</td>
<td></td>
</tr>
<tr>
<td>40–49</td>
<td>38 (29–49)</td>
<td>28 (21–36)</td>
<td></td>
</tr>
<tr>
<td>&gt;=50</td>
<td>28 (19–37)</td>
<td>42 (33–52)</td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>53 (41–65)</td>
<td>58 (48–66)</td>
<td>.04</td>
</tr>
<tr>
<td>Black</td>
<td>21 (10–30)</td>
<td>21 (13–30)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>12 (4–22)</td>
<td>7 (3–13)</td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td>2 (0.2–1)</td>
<td>4 (2–17)</td>
<td></td>
</tr>
<tr>
<td>Asian or other race</td>
<td>1 (0.2–1.4)</td>
<td>0* – -</td>
<td></td>
</tr>
<tr>
<td>Multiple races</td>
<td>12 (7–19)</td>
<td>10 (6–15)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>76 (66–84)</td>
<td>66 (59–73)</td>
<td>.002</td>
</tr>
<tr>
<td>Female</td>
<td>24 (16–34)</td>
<td>34 (27–41)</td>
<td></td>
</tr>
<tr>
<td><strong>Area of residence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Seattle</td>
<td>7 (3–11)</td>
<td>14 (8–21)</td>
<td>1 x 10^{-4}</td>
</tr>
<tr>
<td>Downtown Seattle</td>
<td>52 (41–65)</td>
<td>39 (30–48)</td>
<td></td>
</tr>
<tr>
<td>South Seattle</td>
<td>25 (15–35)</td>
<td>27 (21–37)</td>
<td></td>
</tr>
<tr>
<td>South King County</td>
<td>11 (3–22)</td>
<td>10 (6–15)</td>
<td></td>
</tr>
<tr>
<td>East King County</td>
<td>6 (1–12)</td>
<td>9 (3–14)</td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;High school grad.</td>
<td>23 (16–33)</td>
<td>25 (18–32)</td>
<td>.07</td>
</tr>
<tr>
<td>High school grad.</td>
<td>40 (31–50)</td>
<td>43 (35–50)</td>
<td></td>
</tr>
<tr>
<td>Post high school</td>
<td>37 (26–47)</td>
<td>32 (25–40)</td>
<td></td>
</tr>
<tr>
<td>n (total)</td>
<td>368</td>
<td>497</td>
<td></td>
</tr>
</tbody>
</table>

IDU = injection drug user; RDS = respondent-driven sampling.

*Based on univariate logistic regression incorporating RDSAT generated individual weights.

A single Asian in NHBS-IDU2 was excluded from the racial analysis (only) as RDSAT failed because of a lack of cross-group recruitment for the Asians/other race category; other race was not a response offered in NHBS-IDU2.

NHBS-IDU1 and NHBS-IDU2 was reduced by the differential coupon distribution scheme.

There was evidence of a long-term trend towards older age among Seattle-area IDU; a linear regression of age against year in all non-NHBS data sources was significant (β = 0.09; p < 10^{-3}; Fig. 2). The figure demonstrates an apparent concordance in the time trend in the HARS, RAVEN, Kiwi, and the needle-exchange survey data. The proportions from the NBHS surveys appear greater than what would be expected from the trend in the other data sources and the difference between the two NHBS surveys more pronounced than the differences within the other data sources.

Differences in residential distribution were seen between NHBS-IDU1 and NHBS-IDU2 (p = 1 x 10^{-4}), motivated in large part by NHBS-IDU2 participants’ lower likelihood of living in downtown Seattle (39% vs. 52% in NHBS-IDU1). After introduction of the differential coupon distribution in NHBS-IDU2, the estimated proportion of downtown participants decreased from 65% to 26%. This could have influenced the lower proportion of downtown residents in NHBS-IDU2. The proportions of downtown residents in both NHBS surveys were greater than seen in the other data sources (Fig. 3). Because the overwhelming majority of participants in both NHBS surveys were interviewed in offices downtown, this raises the question whether RDS might preferentially recruit participants residing near interview sites.

Drug-Related Behavior

NHBS-IDU2 participants reported a different pattern of drug preference compared with NHBS-IDU1 participants (p = 1 x 10^{-5}) (Table 2). NHBS-IDU2 participants were more likely than NHBS-IDU1 participants to report heroin as the drug they injected most often (81% vs. 68%) and less likely to report amphetamines (7% vs. 17%). The tendency to report less amphetamine injection in NHBS-IDU2 runs counter to the greater proportion of participants in NHBS-MSM2 reporting male-to-male sex (see below). Amphetamine injection was more commonly reported by MSM/IDU than non-MSM/IDU in both NHBS surveys (46% vs. 5% overall).

Sexual Behavior

NHBS-IDU2 participants were more likely to report male-to-male sex within the previous 12 months than NHBS-IDU1 participants (15% vs. 6%; p = 3 x 10^{-5}). The proportion of MSM/IDU was much greater among HARS cases than in the other data sources (Fig. 4). The figure suggests an increase over time in the proportion of MSM/IDU in combined data from the RAVEN, Kiwi, and needle exchange surveys, and this tests as statistically significant (p = 3 x 10^{-13}). Although the proportions of MSM/IDU in the two NHBS surveys appeared generally consistent with this trend, the variation between the two NHBS surveys appeared substantially more pronounced than seen in the other data sources.

There were notable differences in the representation of MSM/IDU in the different NHBS-IDU2 recruitment chains (Fig. 1). The chain deriving from seed 4 consisted predominantly of MSM/IDU, whereas MSM/IDU representation in the most prolific chain was sparse. Excluding participants from chain 4, the estimated proportion of MSM in NHBS-IDU2 decreased from 15% to 9%. There was evidence of isolation between MSM/IDU and non-MSM/IDU in NHBS-IDU2: the homophily was 0.61 for MSM/IDU and 0.53 for non-MSM/IDU (Table S2). It appears that RDS-adjusted estimates can be influenced by
seed selection, perhaps as a consequence of substantial network isolation.

We constructed a composite variable for sexual HIV transmission risk: vaginal or anal sex without a condom with a partner of unknown or opposite HIV status (unprotected nonconcordant sex) at last sexual contact. This was more likely to be reported by NHBS-IDU2 participants than those in NHBS-IDU1 (32% vs. 19%; \( p = 5 \times 10^{-5} \)).

### HIV- and HCV-Related Variables

NHBS-IDU2 participants were less likely than NHBS-IDU1 participants to report an HIV test both within the previous 12 months (48% vs. 65%; \( p = 2 \times 10^{-6} \)) and within the previous 3 months (14% vs. 29%; \( p = 4 \times 10^{-7} \)). The RDSAT 95% confidence intervals did not overlap for the 3-month time frame. The variability in time trends of HIV testing in the past 12 months appears greater between the two NHBS surveys than within the other studies (Fig. 5). Both the needle exchange and NHBS surveys indicate higher testing levels than seen in older data and both show a decrease in the interval from 2005 to 2011.

Self-reported HIV status differed strikingly between the two NHBS surveys (\( p = 9 \times 10^{-14} \)), with NHBS-IDU2 participants strikingly more likely than NHBS-IDU1 participants to self-report being HIV positive (7% vs. 0.4%). The 95% confidence intervals around the two estimates did not overlap. Greater levels of HIV infection in NHBS-IDU2 than in NHBS-IDU1 were reported both among MSM/
IDU (37% vs. 3%) and among non-MSM/IDU (2.5% vs. 0.6%). In contrast, similar proportions of participants in both surveys were estimated to self-report HCV positive status (58% vs. 59%). The 58% figure in NHBS-IDU2 was unchanged if consideration was restricted to NHBS-IDU2 participants who (like NHBS-IDU1 participants) were not offered HCV testing.

Multivariate Control for Confounding in Differences Between NHBS-IDU1 and NHBS-IDU2

We evaluated the extent to which the difference between NHBS-IDU1 and NHBS-IDU2 in HIV testing in the past 12 months could be a product of confounding by sociodemographic characteristics, drug preference, and MSM status. In logistic regression analyses there was a significant, independent association between HIV testing and race ($p = 1 \times 10^{-9}$) and area of residence ($p = .01$). There was no significant association with age ($p = .19$), sex ($p = .59$), education ($p = .69$), drug most frequently injected ($p = .16$) or male-to-male sex in the previous 12 months ($p = .15$). After control for race and area of residence, NHBS-IDU2 participants remained less likely than NHBS-IDU1 participants to report an HIV test in the previous 12 months (odds ratio = 0.49; $p = 4 \times 10^{-6}$). In analogous models investigating the effects of the same collection of potential confounders, significant differences between NBHS-IDU1 and NHBS-IDU2 persisted for: heroin as the drug most frequently injected, amphetamines as the drug most frequently injected, male-to-male sex in the previous 12 months, unprotected non-concordant sex at last sexual contact, and self-reported HIV positive status ($p \leq 10^{-5}$ in each case).

**DISCUSSION**

Among the 16 variables we compared, we found significant differences between the NHBS-IDU1 and NHBS-IDU2 populations in two variables based on a criterion of non-overlapping RDSAT 95% confidence intervals: HIV testing in the past 3 months and self-reported HIV status. In a further six variables differences between the studies attained a significance of $<10^{-3}$ in weighted logistic regression analyses: age, area of residence, drug most frequently injected, male-to-male sex in the previous 12 months, unprotected non-concordant sex, and HIV testing in the past 12 months. In our judgment the magnitude of the
differences seen in these variables was sufficient to materially affect the characterization of the study population.

We found strikingly different levels of statistical significance in our data depending on whether evaluation was based on non-overlapping confidence intervals or weighted logistic regression. Recent publications are reporting findings based on both these methods (20–24). The discrepancy between methods highlights an urgent need for the development of reliable methods to allow the use of the conventional statistical tools in RDS-recruited study populations. We suggest, however, that statistical measures, structured as they are to reject a difference between groups being compared in the absence of conclusive evidence to the contrary, do not provide a definitive guide to evaluating the comparisons of this study; rather, we use them to direct attention to the most divergent results.

Temporal trends could account for some of the differences we observed between NHBS surveys. The difference between NHBS-IDU1 and NHBS-IDU2 in the proportion of participants older than 50 appears to be only partially accounted for by an overall increase in age among Seattle-area IDU. Differences in the proportion of MSM/IDU seem to be more a product of variability in the two RDS studies than of their accessing distinct populations. It is difficult to determine with confidence the role played by temporal trends and variation in study populations in the differences seen in HIV testing, given the more complicated temporal pattern among the different studies. Investigation of truly replicate RDS samples would require contemporaneous recruitment of parallel RDS-recruited study populations.

Differences in self-reported HIV-positive status between NHBS-IDU1 and NHBS-IDU2 were particularly pronounced. The number of IDU presumed to be living with HIV increased by 6% from 2005 to 2009 (25), so the difference does not reflect a drastic rise in diagnosed HIV infections. The greater levels of HIV testing in NHBS-IDU1 is inconsistent with a hypothesis that their lower rate of self-reported HIV-positive status is a product of having a greater proportion of HIV-positives unaware of their status. NHBS-IDU2 participants were offered an HIV test, whereas those in NHBS-IDU1 were not, so the difference in self-reported status could be a product of inaccurate reporting among participants not subject to HIV testing. Arguing against this is the observation that participants in both surveys self-reported considerable levels of other stigmatized behavior (sharing injection equipment, unprotected sex).
and similar levels of HCV infection. In national NHBS data the proportion of IDU self-reporting HIV-positive status was actually somewhat greater in NHBS-IDU1 (8%) than in NHBS-IDU2 (5%) (26, 27). On the other hand, if the difference in self-reported status were a product of accessing distinctly different subpopulations of IDU in the two surveys, it is surprising that the difference is seen among both MSM/IDU and non-MSM/IDU, groups which appear to be substantially isolated from one another.

In two published reports, one surveying heterosexuals in Cape Town (23), the other among MSM in Beijing (28), authors found material differences between the populations surveyed at different time points, which they interpreted as evidence for changes over time. The possibility of variation attributable to RDS methodology is also worth consideration. The question has been raised by us and others whether RDS may access only a subset of a target population (12, 29). Alternatively, RDS-derived estimates of population characteristics may be statistically unbiased but still have such high variance that attaining estimates with adequate precision is difficult (8). It has been suggested that sample size calculations be multiplied by the magnitude of the design effect (18). Given the design effect of approximately three in our data, this would require that 1500 participants be surveyed rather than the 500, which had been calculated by the Centers for Disease Control and Prevention to have adequate power. Insofar as they do not reflect real change over time, we cannot distinguish in our data whether differences we observed between NHBS-IDU1 and -IDU2 are a product of intrinsically high variance or recruitment of only a limited subset of the target population. Both possibilities would substantially impact the utility of RDS methodology.

Alternatives to RDS recruitment have well-recognized disadvantages. Venue-based and institutionally-based recruitment will not access members of the target population who do not frequent the points of recruitment. These biases are easy to understand, if not to satisfactorily adjust for. Inaccuracies in an RDS-derived population are less overt and may escape the attention of researchers. Nonetheless, RDS offers the opportunity to access hidden populations. It has clear advantages in the simplicity and standardization of its methods, and is particularly advantageous in environments where researchers have limited prior knowledge of how to access a target population. We do not by any means conclude that our data suggest that RDS methodology is fatally flawed, but rather urge that its advantages and
disadvantages continue to be realistically evaluated and weighed against those of alternative recruitment methods.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.annepidem.2012.02.012.

REFERENCES


FIGURE 5. Time trends in the percent reporting an HIV test in the previous 12 months: among Seattle-area IDUs, 1994–2011.


