Respondent-driven sampling for identification of HIV- and HCV-infected people who inject drugs and men who have sex with men in India: A cross-sectional, community-based analysis

Sunil S. Solomon¹,²,³*, Allison M. McFall², Gregory M. Lucas¹, Aylur K. Srikrishnan³, Muniratnam S. Kumar³, Santhanam Anand³, Thomas C. Quinn¹,⁴, David D. Celentano², Shruti H. Mehta²

¹ Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, ² Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, ³ Y.R. Gaitonde Centre for AIDS Research and Education, Chennai, India, ⁴ Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, United States of America

* sss@jhmi.edu

Abstract

Background
A major barrier to achieving ambitious targets for global control of HIV and hepatitis C virus (HCV) is low levels of awareness of infection among key populations such as men who have sex with men (MSM) and people who inject drugs (PWID). We explored the potential of a strategy routinely used for surveillance in these groups, respondent-driven sampling (RDS), to be used as an intervention to identify HIV- and HCV-infected PWID and MSM who are unaware of their status and those who are viremic across 26 Indian cities at various epidemic stages.

Methods and findings
Data were collected as part of the baseline assessment of an ongoing cluster-randomized trial. RDS was used to accrue participants at 27 sites (15 PWID sites and 12 MSM sites) selected to reflect varying stages of the HIV epidemic among MSM and PWID in India. A total of 56 seeds recruited a sample of 26,447 persons (approximately 1,000 participants per site) between October 1, 2012, and December 19, 2013. Across MSM sites (n = 11,997), the median age was 25 years and the median number of lifetime male partners was 8. Across PWID sites (n = 14,450), 92.4% were male, the median age was 30 years, and 87.5% reported injection in the prior 6 months. RDS identified 4,051 HIV-infected persons, of whom 2,325 (57.4%) were unaware of their HIV infection and 2,816 (69.5%) were HIV viremic. It also identified 5,777 HCV-infected persons, of whom 5,337 (92.4%) were unaware that they were infected with HCV and 4,728 (81.8%) were viremic. In the overall
sample (both MSM and PWID), the prevalence of HIV-infected persons who were unaware of their status increased with sampling depth, from 7.9% in participants recruited in waves 1 through 5 to 12.8% among those recruited in waves 26 and above (p-value for trend < 0.001). The overall detection rate of people unaware of their HIV infection was 0.5 persons per day, and the detection rate of HIV-infected persons with viremia (regardless of their awareness status) was 0.7 per day. The detection rate of HIV viremic individuals was positively associated with underlying HIV prevalence and the prevalence of HIV viremia (linear regression coefficient per 1-percentage-point increase in prevalence: 0.05 and 0.07, respectively). The median detection rate of PWID who were unaware of their HCV infection was 2.5 per day. The cost of identifying 1 unaware HIV-infected individual ranged from US$51 to US$2,072 across PWID sites and from US$189 to US$5,367 across MSM sites. The mean additional cost of identifying 1 unaware HCV-infected PWID was US$13 (site range: US$7–US$140). Limitations of the study include the exclusivity of study sites to India, lack of prior HIV/HCV diagnosis confirmation with clinic records, and lack of cost data from other case-finding approaches commonly used in India.

Conclusions

In this study, RDS was able to rapidly identify at nominal cost a substantial number of unaware and viremic HIV-infected and HCV-infected individuals who were currently not being reached by existing programs and who were at high risk for transmission. Combining RDS (or other network-driven recruitment approaches) with strategies focused on linkage to care, particularly in high-burden settings, may be a viable option for achieving the 90-90-90 targets in key populations in resource-limited settings.

Author summary

Why was this study done?

- Men who have sex with men (MSM) and people who inject drugs (PWID) bear a high burden of HIV and HCV, and fall behind other at-risk groups in achieving Joint United Nations Programme on HIV/AIDS (UNAIDS) 90-90-90 targets.
- The behaviors of PWID and MSM (injection drug use and engagement in same-sex relationships) are highly stigmatized in many low- and middle-income countries, resulting in large segments of these populations being hidden.
- Finding PWID and MSM who are HIV-infected but unaware of their status and not in treatment is critical to curbing the HIV epidemic.

What did the researchers do and find?

- We used data collected as part of an ongoing cluster-randomized trial to explore the potential of respondent-driven sampling (RDS), a chain-referral recruitment strategy, to be used as an intervention to identify unaware and viremic HIV-infected MSM and
PWID across 26 Indian cities. We further evaluated the potential of incorporating hepatitis C virus (HCV) testing for PWID.

- RDS was effective in identifying unaware and viremic HIV-infected MSM and PWID, as well as unaware and viremic HCV-infected PWID.
- The longer RDS ran (more waves), the more likely it was to identify unaware and viremic HIV-infected MSM and PWID.
- RDS was most effective in settings with high HIV and HCV burden and low service availability.
- Incorporating HCV testing into HIV programs for PWID is a cost-effective way of identifying PWID unaware of their HCV infection.

What do these findings mean?

- As the HIV community embraces the UNAIDS 90-90-90 targets, respondent-driven recruitment could be a critical tool in achieving the first 90 target among MSM and PWID in low-and-middle-income countries.
- Programs targeting drug-using populations should consider combining HIV and HCV screening for maximal benefit to the population.

Introduction

Effective treatment for HIV and hepatitis C virus (HCV) infection has generated optimism that epidemic control and even elimination may be possible with high treatment coverage rates [1–3]. For HIV, the advent of highly active antiretroviral therapy (HAART) and subsequent therapeutics that are more potent, less cumbersome, and less toxic have transformed HIV from an invariably fatal disease into a chronic, manageable condition [4]. For HCV, while therapeutic developments have come more recently, they have been equally remarkable. New direct antiviral agents have ushered in an era where 95% of those treated can be cured with 8–12 weeks of therapy [5–7]. With these developments for both HIV and HCV, the scientific community has recognized the treatment benefits, both for the individual and the population, of “treatment as prevention.” Accordingly, the Joint United Nations Programme on HIV/AIDS (UNAIDS) has set targets such as “zero new HIV infections” and, more recently, the 90-90-90 targets, which set 90% benchmarks for diagnosis of infected persons, consistent HAART use in diagnosed persons, and viral suppression in treated persons [3]. For HCV, even though therapeutic developments have been more recent, conversations have already turned towards eradication, given the availability of curative therapy [2,8].

However, both HIV and HCV, which together affect 100–120 million people globally and account for more than 1.5 million deaths each year, face similar challenges in meeting these ambitious targets [9,10]. Moreover, programs for both will need to reach vulnerable populations such as people who inject drugs (PWID) and men who have sex with men (MSM), who bear a disproportionately high burden of these infections in low- and middle-income countries (LMICs). In such settings, the greatest barrier to achieving these ambitious targets is low levels of awareness of infection—more than 50% of MSM and PWID in LMICs are unaware of their
HIV status [11], and as many as 90% of PWID in LMICs are unaware of their HCV status or aware but not linked to care [12]. It is thus critical to identify interventions to close this gap.

MSM and PWID have several characteristics in common including, in some settings, criminalization. Injection drug use is universally illegal, and there are laws that criminalize homosexuality (anti-sodomy laws) in more than 70 countries, including 33 countries in Africa and 25 in Asia. As a consequence, both MSM and PWID experience high levels of stigma in most LMICs, which often drives them further underground. Yet both MSM and PWID have strong social ties, leading to high levels of interconnectedness that have previously been leveraged for public health research and surveillance. For example, a key sampling strategy for measuring disease burden in MSM and PWID, called respondent-driven sampling (RDS), relies on individuals to recruit their peers through a reciprocal incentive system [13,14]. While RDS was in some ways initially conceived as a means for achieving HIV prevention [15,16], it has primarily been utilized as a tool for surveillance, and its potential to identify infected individuals and engage them in care has been insufficiently evaluated.

Accordingly, we explored the potential for RDS to be used as an intervention to identify HIV- and HCV-infected PWID and MSM who are unaware of their infection status or who are aware but not engaged in optimal care (detectable viremia) across 26 Indian cities at various epidemic stages. Specifically, using RDS, we examined the rate of, and the factors associated with, identification of people unaware of their infection status or aware of their infection status but with incomplete viral suppression, as well as associated costs.

**Methods**

**Ethical clearances**

All participants were required to provide verbal informed consent to be included in this study. This study was approved by the institutional review boards of the Johns Hopkins University Schools of Medicine and Public Health and the Y.R. Gaitonde Centre for AIDS Research and Education.

**Study setting**

Data used in these analyses were collected as part of the baseline assessment of an ongoing cluster-randomized trial among MSM and PWID in India, Integrated Care Centers to Improve HIV Outcomes in Vulnerable Indian Populations: National Collaboration on AIDS (NCA study) (ClinicalTrials.gov identifier: NCT01686750) [17]. RDS was used to accrue participants for the baseline assessment in order to characterize HIV and HCV (PWID sites only) epidemiology among these key populations in candidate cities for the trial (n = 27 sites [15 PWID and 12 MSM] across 26 Indian cities; New Delhi had both an MSM and a PWID site). Recruitment took place between October 1, 2012, and December 19, 2013. Sites were selected to reflect varying stages of the HIV epidemic among MSM and PWID in India. One RDS field site was established in each city. The sites were on average about 1,000 square feet and usually comprised 6 soundproof rooms (1 for the site coordinator, 1 for the counselor, and 4 for the interviewers). Additionally, there was a laboratory and a phlebotomist at each site, and a logistics assistant to guide participants within the sites.

**Study population**

The study population was accrued using RDS, a chain-referral recruitment strategy that is effective in recruiting hard-to-reach populations [13,14]. At each site, prior to the RDS, ethnography was conducted to identify seeds—individuals considered influential in the local (i.e.,
city) MSM/PWID community who served as the starting point for RDS. Each seed was given 2 hologram-labeled referral coupons to recruit up to 2 network members (i.e., others living in their community that they know inject drugs or have sex with men). Unique identification numbers on the coupons established the recruiter-recruit relationships. Though RDS studies have previously provided up to 5 coupons, we limited to 2 coupons since RDS theory assumes that as the RDS penetrates deeper into a community (i.e., more waves), the bias introduced by ad hoc seed selection is diminished, resulting potentially in a more representative sample. The network members returned to the site with the coupon and, if eligible, were enrolled and asked to complete study assessments. These recruited individuals were considered wave 1 of recruitment and were each given 2 referral coupons to recruit up to 2 members of their network. The next round of individuals recruited and enrolled were considered wave 2, and so on. Recruitment continued at each site until 1,000 eligible participants (prespecified sample size) had been sampled. Two seeds (started simultaneously) were used at all but 1 MSM site (New Delhi) and 1 PWID site (Gangtok), where a third seed was included. The third seed was initiated since recruitment in these 2 cities was relatively slower than in other cities included in this sample. Participants were reimbursed both for participation (INR 250 [US$3.8]) and for each eligible study participant they referred to the study (INR 50 [US$0.8]). Participants provided a fingerprint, which generated a unique hexadecimal code to prevent duplicate enrollment. Additional details on the recruitment process are included in S1 Text.

Eligibility criteria for both MSM and PWID included (1) age ≥ 18 years, (2) provision of informed consent, and (3) possession of a valid RDS referral coupon. In addition to these criteria, MSM participants also had to self-identify as male (transgender/hijras were excluded) and report sex with a man in the prior 12 months. Similarly, PWID participants had to report drug injection in the prior 24 months. There were no exclusions based on sex at PWID sites. Transgender individuals who visited a PWID site with a valid referral coupon were included in the study.

Study procedures

Participants completed an interviewer-administered electronic survey that captured information on demographics, risk behaviors, and access to HIV and HCV diagnostic and treatment services. Upon completion of the survey, participants provided a venous sample following HIV pre-test counseling. Rapid HIV testing was performed using 3 rapid tests: Alere Determine HIV-1/2 (Alere Medical, Chiba, Japan), First Response HIV Card Test 1–2.0 (Premier Medical Corporation, Daman, India), and Signal Flow Through HIV 1+2 Spot/Immunodot Test Kit (Span Diagnostics, Surat, India). HIV results were delivered to participants on the same day with appropriate post-test counseling. All samples were shipped to the central laboratory in Chennai, India, for further testing and storage. Absolute CD4+ count results were made available to HIV-infected study participants within 14 days of their study visit. HIV-1 RNA was quantified in HIV-positive participants using a RealTime HIV-1 assay (Abbott Laboratories, Abbott Park, Illinois, US) with a lower limit of quantification of 150 copies/ml. Among participants from PWID sites, HCV antibody testing was performed on stored specimens in 2014 using the Genedia HCV ELISA 3.0 (Green Cross Medical Science, Chungbuk, Korea), and HCV RNA was quantified using the Abbott RealTime HCV assay (Abbott Laboratories, Abbott Park, Illinois, US).

Statistical methods

The analyses presented in this report were not prespecified as part of the study protocol for the cluster-randomized trial from which these data derive, and a prospective analytical plan was
The primary objective of this analysis was to assess the utility of RDS as an intervention to identify out-of-care (viremic) individuals in key populations across India. As treatment targets for both HIV and HCV are evolving, we considered several definitions of target groups to enhance applicability across settings—Group 1: HIV-infected persons unaware of their status (based on self-report); Group 2: HIV-infected persons with HIV viremia regardless of their awareness or care status; Group 3: HCV-infected PWID unaware of their status (based on self-report); and Group 4: HCV-infected PWID with HCV viremia regardless of their awareness status. Group 1 addresses the first UNAIDS 90-90-90 target, while Group 2 addresses the third target and is relevant to the evolving consensus that viral suppression in a high proportion of infected persons can slow or stop the epidemic [3]. HIV viremia was defined as HIV RNA greater than 150 copies/ml in an HIV-infected person regardless of awareness or care status. HCV viremia was defined as HCV RNA greater than 30 IU/ml in an HCV-infected person regardless of awareness, representing persons with active HCV infection.

We explored several outcomes that we hypothesized could impact the utility of an RDS recruitment approach to help achieve critical HIV and HCV program targets. First, we examined the association between depth of RDS (i.e., recruitment wave) and prevalence of viremic/unaware individuals using line graphs and logistic regression. Second, we estimated the rate at which RDS identified all 4 target groups; this rate was calculated by dividing the number of unaware/viremic individuals identified in a city by the number of days it took to complete the RDS in the city, resulting in a rate of detection per day. Linear regression was used to explore community-level characteristics associated with detection rates across the 27 sites. Community-level characteristics such as HIV/HCV prevalence were calculated using the RDS-II estimator, which weights values using the inverse of network size (i.e., number of MSM/PWID seen in the prior 30 days). Third, we estimated the cost per unaware/viremic individual identified. The cost per individual was calculated by dividing the total cost to complete RDS in a city (e.g., site staffing, rent, utilities, incentives and compensation for RDS, site communication costs), after subtracting research-related expenses (e.g., sample shipping for storage, assays for diseases not related to the outcomes of interest such as HSV-2 and syphilis, investigator costs), by the number of unaware/viremic individuals identified in a city. For estimation of costs, we assumed US$1 was equal to 65 Indian rupees (INR) and utilized market value for tests: HIV antibody, INR 200 (US$3.1); HIV RNA, INR 4,400 (US$67.7); HCV antibody, INR 300 (US$4.6); and HCV RNA, INR 6,500 (US$100.0). Seeds were excluded from all analyses. There were minimal missing data; therefore, we used a complete case analysis approach (observations with missing data were dropped); 10 HIV-infected individuals did not have HIV RNA measurement, and 4 HCV-infected PWID did not have HCV RNA measurement. Pin (postal) code of residence by wave was mapped using ArcGIS version 10.2 (Redlands, CA, US). All statistical analyses were performed using Stata version 13 (StataCorp, College Station, TX, US).

**Results**

**Respondent-driven sampling and population summary metrics**

A total of 56 seeds led to the recruitment of 26,447 ($n = 11,997$ MSM; $n = 14,450$ PWID) persons at 27 sites distributed across 26 Indian cities in 15 states (Table 1) (S1 Fig depicts the flow diagram of the study population). Recruitment of MSM was initiated on October 1, 2012, and ended June 24, 2013; recruitment of PWID was initiated on January 1, 2013, and ended December 19, 2013. Approximately 1,000 participants were recruited per site, with the exception of Moreh (in Manipur), where RDS had to be terminated due to civil unrest. In each city, only 1 RDS field site was established (except New Delhi, which had both an MSM and a PWID...
Despite this, participants were recruited across large distances, some of which require travel times above 2 hours (Fig 1). The site median time to recruit MSM and PWID samples was 99 and 135 days, respectively, ranging from 52 days in Churachandpur (in Manipur; PWID sample) to 200 days in Chandigarh (Union Territory; PWID sample). The median depth of the MSM and PWID samples was 21 and 22 waves, respectively. At MSM sites, the median age was 25 years and the median number of lifetime male partners was 8. Across PWID sites, 92.4% were male, the median age was 30 years, and 87.5% reported injection in the prior 6 months. The median site-level HIV prevalence was 13.0% (range: 2.0%–43.3%) and the median site-level HCV antibody prevalence across PWID sites was 45.5% (range: 4.1%–68.4%). Other characteristics of these samples and factors associated with HIV incidence and prevalence have been described elsewhere [18,19].

Of the 4,051 HIV-infected persons recruited, 2,325 (57.4%) reported that they were unaware of their HIV infection and 2,816 (69.5%) had HIV viremia. At PWID sites, of the 5,777 PWID positive for HCV antibodies, 5,337 (92.4%) were unaware that they were infected with HCV and 4,728 were HCV viremic (81.8%).

### Table 1. Summary respondent-driven sampling recruitment metrics and HIV/HCV characteristics.

<table>
<thead>
<tr>
<th>Metric/characteristic</th>
<th>Overall (27 sites)</th>
<th>MSM sites only (12 sites)</th>
<th>PWID sites only (15 sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number recruited (seeds)</td>
<td>26,447 (56)</td>
<td>11,997 (25)</td>
<td>14,450 (31)</td>
</tr>
<tr>
<td>Median recruitment time in days (range)</td>
<td>112 (52–200)</td>
<td>99 (70–157)</td>
<td>135 (52–200)</td>
</tr>
<tr>
<td>Median number of waves (range)</td>
<td>21 (11–50)</td>
<td>21 (11–28)</td>
<td>22 (12–50)</td>
</tr>
<tr>
<td>Median site-level HIV prevalence (%) (range)</td>
<td>13.0 (2.0–43.3)</td>
<td>8.6 (2.0–18.8)</td>
<td>19.8 (6.1–43.3)</td>
</tr>
<tr>
<td>Weighted median site-level HIV prevalence (%) (range)*</td>
<td>11.2 (1.7–44.9)</td>
<td>6.4 (1.7–13.1)</td>
<td>18.1 (6.0–44.9)</td>
</tr>
<tr>
<td>Median site-level proportion of HIV-positive participants unaware of status (%) (range)</td>
<td>61.1 (7.6–100)</td>
<td>63.6 (7.6–100)</td>
<td>56.7 (7.8–97.2)</td>
</tr>
<tr>
<td>Weighted median site-level proportion of HIV-positive participants unaware of status (%) (range)*</td>
<td>64.1 (7.2–100)</td>
<td>70.0 (10.1–100)</td>
<td>59.2 (7.2–97.6)</td>
</tr>
<tr>
<td>Median site-level prevalence of HIV viremia (%) (range)</td>
<td>9.0 (1.8–33.8)</td>
<td>5.0 (1.8–13.2)</td>
<td>14.2 (2.8–33.8)</td>
</tr>
<tr>
<td>Weighted median site-level prevalence of HIV viremia (%) (range)*</td>
<td>7.5 (0.8–29.7)</td>
<td>3.6 (0.8–10.1)</td>
<td>11.4 (2.7–29.7)</td>
</tr>
<tr>
<td>Median site-level HCV prevalence (%) (range)</td>
<td>—</td>
<td>—</td>
<td>45.5 (4.1–68.4)</td>
</tr>
<tr>
<td>Weighted median site-level HCV prevalence (%) (range)*</td>
<td>—</td>
<td>—</td>
<td>41.1 (4.9–64.9)</td>
</tr>
<tr>
<td>Median site-level proportion of HCV-positive participants unaware of status (%) (range)</td>
<td>—</td>
<td>—</td>
<td>95.4 (75.6–99.7)</td>
</tr>
<tr>
<td>Weighted median site-level proportion of HCV-positive participants unaware of status (%) (range)*</td>
<td>—</td>
<td>—</td>
<td>94.4 (70.2–99.9)</td>
</tr>
<tr>
<td>Median site-level prevalence of HCV viremia (%) (range)</td>
<td>—</td>
<td>—</td>
<td>37.0 (3.5–58.0)</td>
</tr>
<tr>
<td>Weighted median site-level prevalence of HCV viremia (%) (range)*</td>
<td>—</td>
<td>—</td>
<td>34.1 (4.5–53.9)</td>
</tr>
</tbody>
</table>

*Weighted using RDS-II estimator.

HCV, hepatitis C virus; MSM, men who have sex with men; PWID, people who inject drugs.

https://doi.org/10.1371/journal.pmed.1002460.t001

Prevalence of unaware and viremic individuals by recruitment wave

The prevalence of HIV-infected persons who were unaware of their status in the overall sample increased with sampling depth, from 7.9% in participants recruited in waves 1 through 5 to 12.8% among those recruited in waves 26 and above (p-value for trend < 0.001; Table 2). Among HIV-positive persons, the proportion who were unaware of their HIV infection increased from 46.5% in waves 1–5 to 77.7% in waves 26 and above (p-value for trend < 0.001; Fig 2). The odds ratio of identifying persons unaware of their HIV infection increased significantly by recruitment wave (odds ratio per 5-recruitment-wave increase: 1.31; 95% CI: 1.24, 1.38; Table 2). The prevalence of HIV viremic individuals recruited increased from 11.3% in waves 1–5 to 14.6% in waves 26 and above. Among HIV-positive persons, the proportion with
viremia increased from 67% to 88.4% between waves 1–5 and waves 26 and above; in this group, there was a 3.73 higher odds of identifying a viremic person in waves 26 and above compared to waves 1 through 5 (95% CI: 2.32, 6.01). With respect to HCV, the odds of detecting HCV-infected PWID who were unaware of their HCV infection was 1.32-fold higher (95% CI: 1.22, 1.44) per 5-recruitment-wave increase, but there was no association between recruitment depth and identification of HCV viremic PWID.

Rate of detection of unaware and viremic individuals

The detection rate of HIV-infected individuals with viremia overall was 0.7 persons per day, with median detection rates across PWID and MSM sites of 0.7 and 0.4 per day, respectively. However, there was substantial variability, ranging from 1 unaware PWID identified every 10 days in Gangtok to 3.2 detected each day in Churachandpur (Table 3). Among the MSM sites, Hyderabad was the only site where the rate exceeded 1 per day (1.2 unaware HIV-infected MSM per day; Table 4). The detection rate of HIV-infected PWID with viremia ranged from 1 per 3.3 days in Gangtok and Lunglei to 2.7 per day in Churachandpur. The median detection rate of HIV viremic MSM was 0.5 per day (site range: 0.2–1.1 per day). The detection rate of
HIV viremic individuals was positively associated with underlying community HIV prevalence and prevalence of HIV viremia (linear regression coefficient per 1-percentage-point increase in prevalence: 0.05 and 0.07, respectively).

The median detection rate of HCV-infected individuals who were unaware of their status across all sites combined was 2.5 per day, but the rate was as high as 10.2 per day in Churachandpur (Table 3). The detection rate of HCV viremic PWID ranged from approximately 1 every 3 days in Gangtok to 9 per day in Churachandpur. Underlying HCV prevalence and prevalence of HCV viremia in the community were positively associated with a more rapid detection rate for HCV viremic PWID (linear regression coefficient per 1-percentage-point increase in prevalence of HCV and HCV viremia was 0.13 and 0.16, respectively; \( p < 0.004 \) for both).

### Cost per unaware and viremic person identified

The cost of identifying 1 unaware HIV-infected individual ranged from US$51 to US$2,072 across PWID sites (Table 3) and from US$189 to US$5,637 across MSM sites (Table 4). The median cost of identifying an HIV viremic individual was US$367 (site range: US$123–US$1,406). At PWID sites, the cost ranged from US$123 to US$858, and the range extended from US$305 to US$1,406 at the MSM sites. The cost of identifying viremic individuals was strongly associated with underlying prevalence of viremia in the community (linear regression coefficient per 1-percentage-point increase in prevalence of viremia: \(-US$31.8; p < 0.001; \) Fig 3).

Incorporating all site operational and RDS costs without taking into account HIV antibody and HIV RNA testing costs, the median site cost per unaware HCV-infected PWID identified was US$50 (site range: US$27–US$549). The median site cost to identify 1 HCV viremic individual was US$194 (range: US$158–US$723). The cost of identifying 1 unaware HCV-infected individual was associated with underlying community HCV prevalence (linear regression coefficient per 5-percentage-point increase in HCV prevalence: \(-US$29.9; p < 0.001; \) Fig 3). If HCV testing is incorporated into an existing RDS program framework (as in this study for HIV), the additional cost to identify 1 unaware HCV-infected PWID is US$11 (site range: US$7–US$150).

### Table 2. Association of respondent-driven sampling depth and prevalence of unaware and viremic persons.

<table>
<thead>
<tr>
<th>Sampling depth</th>
<th>Odds ratio (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prevalence of unaware HIV+ (full sample; ( n = 26,444 ))</td>
</tr>
<tr>
<td>Per 5-wave increase</td>
<td>1.18 (1.14, 1.22)</td>
</tr>
<tr>
<td>Wave of recruitment</td>
<td>1</td>
</tr>
<tr>
<td>1–5</td>
<td>0.85 (0.72, 1.01)</td>
</tr>
<tr>
<td>6–10</td>
<td>1.29 (1.09, 1.53)</td>
</tr>
<tr>
<td>16–20</td>
<td>1.50 (1.24, 1.81)</td>
</tr>
<tr>
<td>21–25</td>
<td>1.34 (1.05, 1.70)</td>
</tr>
<tr>
<td>26+</td>
<td>1.71 (1.37, 2.15)</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pmed.1002460.t002

HIV viremic individuals was positively associated with underlying community HIV prevalence and prevalence of HIV viremia (linear regression coefficient per 1-percentage-point increase in prevalence: 0.05 and 0.07, respectively).

The median detection rate of HCV-infected individuals who were unaware of their status across all sites combined was 2.5 per day, but the rate was as high as 10.2 per day in Churachandpur (Table 3). The detection rate of HCV viremic PWID ranged from approximately 1 every 3 days in Gangtok to 9 per day in Churachandpur. Underlying HCV prevalence and prevalence of HCV viremia in the community were positively associated with a more rapid detection rate for HCV viremic PWID (linear regression coefficient per 1-percentage-point increase in prevalence of HCV and HCV viremia was 0.13 and 0.16, respectively; \( p = 0.004 \) for both).
Discussion

In this study, RDS was able to identify a substantial number of HIV-infected and HCV-infected individuals in key populations who were unaware of their infection and out of care. Moreover, the deeper RDS penetrated within a community, the more likely it was to identify individuals with HIV and HCV viremia, the groups at the highest risk for transmission. Costs were relatively low, particularly in settings with high disease burden. RDS was also able to reach communities across the city via a single site. Combining RDS (or other network-driven recruitment approaches) with strategies focused on linkage to care may be a viable option for achieving the 90-90-90 targets in key populations in resource-limited settings.

RDS was originally conceived as an HIV prevention strategy and as a means to arrive at representative estimates of disease burden in hidden populations, populations with no obvious sampling frame [13,14]. Consequently, RDS is most commonly used in surveillance studies among MSM and PWID and has become the cornerstone for accrual of samples of MSM and PWID globally. The majority of RDS samples comprise ~400–600 participants, and samples are typically terminated either when the sample achieves equilibrium or the desired sample size has been accrued. However, in recent years, RDS has been incorporated into trials particularly among MSM, PWID, and female sex workers to recruit the study population [20], identify high-risk individuals [15,16], or measure the impact of an intervention on the community [21,22]. Further, data from a study among 722 MSM in Nigeria suggested that deeper RDS waves were associated with lower likelihood of awareness of HIV status and being on ART [23]. Currently, there is a trial in progress (HPTN 078) that is assessing the ability of “deep-chain” RDS (defined as recruitment beyond the sixth wave) to identify viremic MSM in the US [24], reflecting that the potential for RDS to be used as an intervention is increasingly being recognized.

The data in this report represent one of the largest samples, to our knowledge, of MSM and PWID systematically collected using RDS across diverse settings at different epidemic stages within the same country. Our findings have implications for potential tailoring of RDS as an intervention in these 2 key populations. It has been hypothesized that the deeper RDS penetrates within a community (measured by the number of waves), the more likely it is to recruit hidden populations [13]. These distal recruits are also hypothesized to have relatively smaller networks (i.e., to be less connected) compared to those recruited in the early waves. Indeed, this was supported in our data by ethnographic work at the sites, which suggested that more participants recruited in later waves had never before visited the local non-governmental organization through which the RDS was operating, and demonstrated that those recruited in later waves were significantly less likely to be engaged in HIV and HCV services. This was particularly apparent for HCV, for which testing but not treatment services were available in India at the time of this study. For HCV, people who were recruited in earlier waves were significantly more likely to be aware of their HCV status, suggesting that those who are more connected within their community have more access to testing services, but there was no difference in HCV viremia by wave, which may reflect the near absence of treatment availability at the time of this study. By contrast, for HIV, both diagnostic and treatment services are widely available, and the earlier recruits were more likely to be aware of their status and virologically suppressed compared to later recruits, suggesting that the deeper the RDS runs, the more likely it will be...
to identify persons at highest risk for transmission of HIV. This has also been demonstrated among MSM in Nigeria [23].

It is important to note that there was extensive variability in the efficiency of RDS in identifying unaware and viremic persons by site. At sites with low HIV prevalence, the yield tended
to be lower and the costs significantly higher (e.g., MSM in Chennai and Mangalore and PWID in Gangtok and Bhubaneswar). By contrast, among PWID in Churachandpur and Imphal, more than 2 viremic persons were being detected each day the RDS was operational. Among MSM sites in general, the yield was lower relative to PWID, likely attributable to the lower HIV prevalence of HIV in MSM compared to PWID [18,19]. Yet, it is important to note that even in MSM populations in settings such as Hyderabad and Vijayawada in the high-prevalence state of Andhra Pradesh, at least 1 individual with HIV viremia was detected each day RDS was ongoing.

We observed similar variability with respect to the costs per unaware or viremic individual identified. Sites with high disease prevalence had low costs and sites with low disease burden had higher costs, and this was consistent both at MSM and PWID sites. A recent cost-effectiveness analysis suggested that screening all persons in India for HIV once every 5 years compared to the current standard of testing (at presentation with AIDS-defining illness or at presentation to an integrated counseling and testing center, antenatal center, or sexually transmitted infection clinic) would be “cost-effective,” and a one-time screening of the national population would be “very cost-effective” as per WHO criteria for cost-effectiveness [25]. The prevalence of HIV in India is 0.26% [26]; using this prevalence and the market value of US$3.1 for an HIV test, it would cost US$1,192 to find 1 infection in the national population, without accounting for staffing, infrastructure, community mobilization, etc. Yet this approach was deemed to be “very cost-effective.” While we do not present a cost-effectiveness model in this paper, the cost of identifying 1 person with HIV viremia was lower than this threshold at 25 out of the 27 sites, highlighting the potential cost-effectiveness of combining RDS with other

![Image](https://doi.org/10.1371/journal.pmed.1002460.t004)

Table 4. Respondent-driven sampling recruitment metrics and HIV characteristics across MSM sites.

<table>
<thead>
<tr>
<th>Metric/characteristic</th>
<th>Andhra Pradesh</th>
<th>Karnataka</th>
<th>Tamil Nadu</th>
<th>Central and north India</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HYD</td>
<td>VJW</td>
<td>VZG</td>
<td>BLR</td>
</tr>
<tr>
<td>Total number recruited</td>
<td>998</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
</tr>
<tr>
<td>Number of seeds</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Recruitment time (days)</td>
<td>118</td>
<td>97</td>
<td>104</td>
<td>121</td>
</tr>
<tr>
<td>Number of waves (by seed)</td>
<td>22</td>
<td>(22, 17)</td>
<td>17</td>
<td>(12, 2)</td>
</tr>
<tr>
<td>Number of recruits by seed</td>
<td>619, 379</td>
<td>997, 3</td>
<td>291, 709</td>
<td>126, 874</td>
</tr>
<tr>
<td>Coupon response rate (%)</td>
<td>52</td>
<td>51</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Median network size*</td>
<td>10</td>
<td>20</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>Prevalence of unaware HIV-infected persons (%)</td>
<td>13.7</td>
<td>8.4</td>
<td>4.3</td>
<td>5.3</td>
</tr>
<tr>
<td>Rate of detection of unaware HIV-infected persons per day</td>
<td>1.16</td>
<td>0.87</td>
<td>0.41</td>
<td>0.44</td>
</tr>
<tr>
<td>Cost per unaware HIV-infected person identified (USD)</td>
<td>200.7</td>
<td>300.5</td>
<td>424.0</td>
<td>455.5</td>
</tr>
<tr>
<td>Prevalence of HIV viremia (%)</td>
<td>13.2</td>
<td>10.5</td>
<td>4.9</td>
<td>6.4</td>
</tr>
<tr>
<td>Rate of detection of HIV viremic persons per day</td>
<td>1.12</td>
<td>1.08</td>
<td>0.47</td>
<td>0.53</td>
</tr>
<tr>
<td>Cost per viremic person identified (USD)</td>
<td>304.7</td>
<td>328.1</td>
<td>533.7</td>
<td>473.5</td>
</tr>
</tbody>
</table>

*BNumber of MSM seen in prior 30 days.

BGM, Belgaum; BHO, Bhopal; BLR, Bengaluru; CHE, Chennai; CMB, Coimbatore; DEL, New Delhi; HYD, Hyderabad; LKW, Lucknow; MAD, Madurai; MLR, Mangalore; MSM, men who have sex with men; VJW, Vijayawada; VZG, Vishakhapatnam; USD, US dollars.

https://doi.org/10.1371/journal.pmed.1002460.t004
RDS to identify HIV- and HCV-infected key populations in India
strategies to improve linkage to care and viral suppression. Further, cost-effectiveness would be significantly improved by incorporating screening for other diseases that impact mortality and disease transmission (e.g., tuberculosis and HCV in PWID, and syphilis and HSV-2 in MSM). The additional cost of identifying 1 HCV-infected PWID with viremia across the 15 PWID sites was only US$136 (and this includes market value for HCV antibody and HCV RNA testing, which amount to US$105). Given the low price of HCV therapy in India (~US $300 per 12-week regimen of sofosbuvir and daclatasvir) and the potential to prevent cirrhosis and hepatocellular carcinoma, this could potentially be very cost-effective.

These data collectively highlight the potential of RDS (if combined with other linkage-to-care strategies) to reach individuals who could benefit from treatment of HIV and are at highest risk for transmission. Yet questions remain on how best to optimize RDS. One might argue that running RDS indefinitely at a site may provide the highest yield. An alternate argument could be to continue RDS until certain programmatic targets have been achieved (e.g., recruit 500 HIV viremic PWID). Likely, these targets would need to be tailored to the setting based on underlying population size, disease prevalence, and the proportion of infected individuals with viral suppression. As a compromise, RDS could be allowed to run until sample proportions have stabilized (e.g., equilibrium is achieved). Generally, this occurs around 300–400 persons, as was the case in all our 27 samples; HIV prevalence remained stable between this point and when the final sample size was achieved. The decision to continue the RDS or not could be based on the HIV prevalence within a site sample once equilibrium is attained. As clearly highlighted in the data, settings with historic epidemics and good access to ART such as Chennai will not be good candidate sites for RDS or other network-driven recruitment approaches to find viremic persons. In contrast, emerging sites with limited diagnostic and treatment access such as Kanpur are potentially good candidate sites for such an approach. The optimal number of coupons handed out is also worth further consideration. In this study, all participants were provided with 2 coupons as RDS was being used for surveillance, and depth was critical for generalizability of the sample. However, if RDS was being used to find unaware and viremic individuals, providing more coupons could result in identification of cases more rapidly, but at the same time the site would need to be suitably staffed to handle an increased client load. Alternatively, more coupons could be provided to select participants with characteristics suggestive of being “good” recruiters—participants more likely to hand out coupons to unaware and viremic persons. Future studies will need to explore these areas to optimize the efficiency of RDS to identify unaware and viremic persons.

A number of limitations need to be acknowledged. This study was conducted exclusively in India and may not be generalizable to other settings. However, we did have data on a range of epidemic stages across diverse geographic settings within India. It is challenging to compare the costs of this RDS-based strategy with those of the standard of care (targeted interventions) in India because these costs are not widely available. Further, while we used biometric data to exclude duplicates in our setting, this is not the standard across other targeted interventions in India. Our study was unique in that we recruited such large samples with only 2 seeds. This is in contrast to most studies, which often use upwards of 6 to 12 seeds. It is possible that our networks were more interconnected than others or that there was a large number of eligible individuals within the target area. It is possible that RDS is only effective in settings where the
population is one large interconnected network. However, in other samples that have required more seeds (e.g., 7 in the study among MSM in Nigeria) [23], the same association between sample depth and proportion unaware of their HIV status has been seen. Participants were classified as unaware of their HIV and HCV status based on self-report, which is subject to bias/misclassification. Ideally, verification of prior diagnoses would be conducted with clinical records or electronic medical records. However, the availability of HIV RNA and HCV RNA data provided an objective outcome for the analyses, and these data are possibly the most important data if control or elimination is to be achieved.

In conclusion, in this large study among MSM and PWID across diverse settings in India, RDS was effective in reaching a large sample of unaware and viremic HIV-infected and HCV-infected individuals at relatively low cost, reflecting its potential for use as an intervention to improve detection of individuals with the potential to transmit HIV and/or HCV infection. In LMIC settings with high disease burden among key populations, combining RDS (or other network-driven recruitment approaches) with strategies focused on linkage to care may be a viable option for achieving the 90-90-90 targets.

Supporting information

S1 Fig. Flow diagram of study population.
(PDF)

S1 Text. STROBE-RDS checklist.
(PDF)

S2 Text. Institutional review board approvals.
(PDF)

Acknowledgments

We thank the National AIDS Control Organization, India, and all of our partner non-governmental organizations throughout India that assisted with the recruitment of the study sample. Most importantly, we would like to thank our participants, without whom this research would not have been possible.

Author Contributions

Conceptualization: Sunil S. Solomon, Allison M. McFall, Gregory M. Lucas, Shruti H. Mehta.

Data curation: Allison M. McFall, Santhanam Anand.

Formal analysis: Sunil S. Solomon, Allison M. McFall, Shruti H. Mehta.

Funding acquisition: Sunil S. Solomon, Gregory M. Lucas, Thomas C. Quinn, David D. Celentano, Shruti H. Mehta.


Methodology: Sunil S. Solomon, Allison M. McFall, Shruti H. Mehta.


Software: Sunil S. Solomon, Santhanam Anand.

Writing – original draft: Sunil S. Solomon, Allison M. McFall.

Writing – review & editing: Sunil S. Solomon, Allison M. McFall, Gregory M. Lucas, Aylur K. Srikrishnan, Muniratnam S. Kumar, Santhanam Anand, Thomas C. Quinn, David D. Celentano, Shruti H. Mehta.

References


