Published in final edited form as:

Int J Soc Res Methodol. 2013 ; 16(4): . doi:10.1080/13645579.2012.661204.

Community understanding of Respondent-Driven Sampling in a medical research setting in Uganda: importance for the use of RDS for public health research

Nicky McCreesh¹, Matilda Nadagire Tarsh², Janet Seeley^{1,2,3}, Joseph Katongole², and Richard G White¹

¹London School of Hygiene and Tropical Medicine

²MRC/UVRI Uganda Research Unit on AIDS

³School of International Development, University of East Anglia

Abstract

Respondent-driven sampling (RDS) is a widely-used variant of snowball sampling. Respondents are selected not from a sampling frame, but from a social network of existing members of the sample. Incentives are provided for participation and for the recruitment of others. Ethical and methodological criticisms have been raised about RDS. Our purpose was to evaluate whether these criticisms were justified. In this study RDS was used to recruit male household heads in rural Uganda. We investigated community members' understanding and experience of the method, and explored how these may have affected the quality of the RDS survey data. Our findings suggest that because participants recruit participants, the use of RDS in medical research may result in increased difficulties in gaining informed consent, and data collected using RDS may be particularly susceptible to bias due to differences in the understanding of key concepts between researchers and members of the community.

Introduction

Respondent-driven sampling (RDS) is an increasingly widely used variant of snowball sampling for recruiting hidden populations (Heckathorn, 1997; Salganik & Heckathorn, 2004), including populations at risk of HIV infection (Heimer, 2005b; Ramirez-Valles et al., 2005; Semaan et al., 2009). Respondents are selected not from a sampling frame but from a social network of existing members of the sample. The sampling process starts with a small number of individuals who are selected by convenience. These individuals, the seeds, then recruit others to participate in the study. The process of existing sample members recruiting others to join the sample continues until the desired sample size, or some measure of data quality, is reached. While this sampling method has been found to be efficient by many researchers (Malekinejad et al., 2008), others have expressed concerns about the accuracy and precision of the RDS estimates (Goel & Salganik, 2010; Heimer, 2005a) and the ethical implications of using this sampling approach (Scott, 2008).

Scott undertook a qualitative evaluation of the RDS used to recruit the sample for an HIV behaviour surveillance study among injection drug users. He suggested that his ethnographic study 'indicated several RDS implementation issues' (Scott 2008: 43) which demanded

Corresponding author: Professor Janet Seeley, MRC/UVRI Uganda Research Unit on AIDS, P.O. Box 49, Entebbe, Uganda; Or, School of International Development, University of East Anglia, Norwich NR4 7TJ. janet.seeley@mrcuganda.org or j.seeley@uea.ac.uk.

scrutiny: in particular that participants ‘quickly learned various methods for subverting the study’s screening process’ which created tensions between participants, and that there was evidence of coercion in recruitment practices. The veracity of Scott’s study has been criticised by several commentators (Broadhead, 2008; Lansky & Mastro, 2008; Ouellet, 2008) who question both his methods and the ethical conduct of his own research. In view of Scott’s findings, and the criticism of his research, we set out to investigate whether RDS is indeed a reliable and ethical method of sampling and data collection in medical research settings.

It has been asserted that under certain conditions RDS may be able to generate representative data on hidden populations (Volz & Heckathorn, 2008). This claim is methodologically difficult to evaluate because the representative or total-population data that are required for robust evaluation are usually unavailable for these populations. We dealt with this problem by evaluating RDS in a non-hidden/non-stigmatised population on which high quality total population data were available: an annually-surveyed open cohort in southern Uganda. We were able to compare the characteristics of the RDS sample with the known characteristics of the target population from the total population data. The findings of this evaluation are published in full elsewhere (McCreesh et al., In review), but in brief we found that RDS produced a sample that was largely representative of the total population for most variables, but was biased such that men who were younger, of higher socioeconomic status, and with unknown sexual activity or HIV status were under-represented. Our preliminary analysis suggested that these biases may have been linked to the community understanding of RDS. A-priori, we also had concerns that the RDS incentive scheme may reduce participation in other, ongoing, survey activities, in the same population, that did not give such a direct incentive for participation as RDS (White & Seeley, 2010).

Drawing on concerns raised by Scott, our own a-priori concerns about the potential of RDS to reduce participation in other survey activities, and the biases we observed in our own evaluation of RDS, the aims of this study were to describe the community understanding and experience of RDS, and explore how these may have affected the quality of the RDS data and participation in other ongoing survey activities.

Background

Respondent-Driven Sampling

In RDS, as in snowball sampling, respondents recruit their peers. The seeds are given coupons, typically three, which they give to the others from the target population to invite them to be interviewed. If they are eligible and they consent, they themselves can become recruiters. Recruits are given an incentive, often money, for taking part in the survey and also for recruiting others. This process continues in recruitment ‘waves’ (with wave zero being defined as the seeds, wave one as their recruits, etc.) until a pre-determined sample size is reached or the distribution of participant characteristics becomes similar between waves. A record is kept of who has recruited whom and the reported number of social contacts of each participant (their ‘network size’). Statistical estimation methods are then applied to account for the non-random sample selection in an attempt to generate unbiased estimates for the target population.

The study setting

The study area is in southern Uganda. The Medical Research Council/Uganda Virus Research Institute established a General Population Cohort (GPC) there in 1989 in order to study the epidemiology of HIV. Annual demographic and serological surveys have been conducted with this population for the last 22 years, resulting in a wealth of data on the population characteristics of approximately 20,000 people in approximately 5000

households. People living in the study area are largely subsistence farmers who produce small amounts of cash crops such as beans, bananas and coffee. The majority, 70%, of the population is ethnically Baganda. The rest of the population is made up of immigrants from Rwanda (15%) and Tanzania (4%), and a mixture of other tribes. The main local language is Luganda which is spoken and understood by all the tribes. The community is predominantly Roman Catholic (60%) with 23% Muslim and 17% Protestant. Just over 50% of the population is under 15 years old. Most households have less than five acres of land. There are only a few sizeable land owners and relatively few households are landless.

Methods

Quantitative methods

The primary aim of the quantitative evaluation of RDS was to evaluate whether RDS could be used to generate representative data by comparing field-collected RDS data with total population data on the same population. The selected target population was male household heads in 25 villages in rural southern Uganda. Many potential target groups were considered, but the selected group was male household heads because it was the best match of the following criteria: data were routinely collected on this group during the GPC annual survey; group membership would be easily verifiable by an interviewer; group membership would not vary much between annual survey rounds; there was only one group member in each household; the group was large enough to ensure a sample of 250 was not more than around 10% of the group, and group membership was known to other members of the village population (a requirement for RDS to succeed). Full details of the quantitative methods can be found in (McCreesh et al., In review), but a summary follows.

The data used to define the RDS target population were available from the ongoing GPC (Shafer et al., 2008). Each year, after obtaining consent, a total-population household census and an individual questionnaire and HIV-1 serosurvey are administered to this population.

The target population consisted of 2402 men who were recorded as a head of a household within the GPC villages between February 2009 and January 2010. The characteristics of the target population (male household heads) are shown in Table 1 (population proportion column). In brief, 49% were aged under 40 years old, 60% were Catholic, 23% were Muslim and 17% were Protestant. 70% were from the Baganda tribe and the percentage in each village varied between 1.7% and 8.6%. 42% reported one sex partner in the last year and 3.7% reported four or more. 6.3% were known to be HIV infected.

To maximise the generalisability of our results, where possible we employed currently used RDS data collection methods (Malekinejad et al., 2008). Ten seeds, who stated that they were confident that they could recruit other male household heads, were purposively selected from the target population to vary in age, tribe, and village. Seeds and subsequent recruits were given three coupons to recruit other men into the study. Seeds and recruits were offered incentives for participation and recruitment, either soap, salt or school books to the value of ~\$1US. One incentive was offered for completing the first interview and another incentive for each person successfully recruited. The incentives were called 'tokens of appreciation' by the MRC staff and community during the study.

The eligibility criteria for the RDS survey were (1) having a RDS study coupon with a valid coupon number (except seeds) and (2) on the GPC list of male heads of household in any of the 25 GPC study villages at any time in the year prior to the start of the RDS (8th March 2010).

If the man was eligible, information about the study was given and consent requested. Confidentiality was emphasized.

The main measure of a recruit's network size (the number of other male household heads known to the recruit) was determined from their response to the question: *"Baami bameka b'omanyi nga (i) mu myezi kkumi n'ebiri egiyise baali ba nannyinimu mu byalo bya MRC, (ii) era ng'obamanyi nabo bakumanyi, (iii) ng'obalabyeko mu week ewedde?"* "How many men do you know who (i) were head of a household in the last 12 months in any of the MRC villages, and (ii) you know them and they know you, and (iii) you have seen them in the past week". Recruits were also asked to provide detailed information on the name, village, age, tribe and religion of all men they knew who met the criteria, what their relationship was to each one, how long they had known them, and how often they saw them. In addition, they were asked why they had agreed to participate in the study and data were collected on the recruit's relationship with his recruiter. If recruiters returned to collect their incentives for recruitment, data were also collected on who they did and did not offer their coupons to, and who accepted or refused coupons.

The possible effects of the RDS survey on participation in the annual GPC survey were explored using data on 4790 men and 5367 women aged 13 or over who were censused as resident in the annual GPC census round that had been completed before the start the RDS study (round 20, November 2008 to September 2009) and who were also censused as resident in the annual GPC census round that was ongoing when the RDS study took place (round 21, December 2009 to October 2010). The RDS survey took place in March and April 2010. As such, GPC round 21 survey interviews were completed in six GPC villages before the start of the RDS study (defined as the 'before' villages) and started in 13 GPC villages after the end of the RDS study (defined as the 'after' villages). 'New participants' were defined as residents who were eligible for the GPC in both round 20 and 21 who did not participate in round 20, but did participate in round 21. 'New drop-outs' were defined as residents who were eligible for the GPC in both round 20 and 21 who did participate in round 20, but did not participate in round 21. We compared the proportions of new participants and new drop-outs between villages surveyed before and after the RDS.

The RDS survey questionnaire was programmed in Access 2003 VBA (Microsoft Corporation, 2003) on Samsung Q1 UMPCs. Pre-processing and data analysis of the data were performed using Stata v11 (StataCorp, 2010). Geographic plots were performed in ArcGIS 9.2 (Environmental Systems Research Institute). Village names were replaced with letters for confidentiality. Full details of the methods can be found in (McCreesh et al., In review).

Qualitative methods

In preparation for the qualitative study, selected community members or staff were approached and asked for their consent to participate in a short follow up to the RDS quantitative study. If they consented, arrangements were made to meet them for a 30-60 minute conversation. This took the form of a semi-structured interview, guided by a checklist that was memorised by the interviewers prior to the interviews. The interviews were conducted in Uganda. The interviewers, two men and one woman, were experienced in qualitative data collection.

Fifty four members of the population in the study villages were selected for interview. The groups sampled, sample sizes, and sampling methods were 1) RDS Recruit: 10 RDS recruits were randomly selected from the 917 eligible recruits (excluding seeds), 2) RDS Refuser: 10 men who were reported by recruiters as having refused coupons and we knew had not enrolled in the RDS quantitative study were randomly selected from 29 eligible, 3) Village

Resident: 10 male or female village residents over the age of 18 years who were not RDS recruits or refusers were randomly selected from 8695 eligible, 4) Key Informant: 10 key informants (for example, local leaders) from the study population were selected purposively, 5) RDS Interviewer: All 10 RDS interviewers were selected for interview, 6) GPC Interviewer: four GPC survey staff were randomly selected from 25 eligible.

For the qualitative study, the selected participants were approached by the interviewers and were given information sheets which they went through with the interviewers, after which consent was sought. Confidentiality was emphasized.

The interviews were written up in English by the interviewers on the day of interview. All data were edited manually before being uploaded into Nvivo 8 (QSR International Pty Ltd, 2008), which was used for the data management for analysis. The data collection team developed a list of themes emerging from the data which was used to code the interview scripts. Scripts were exchanged during the theme development stage so that the team consolidated themes from them. This gave the team a broader view of people's opinions. Two members of the team separately coded each of the scripts and then checked the coded scripts of the other to make sure that their discrepancies were resolved.

Results

All ten RDS seeds successfully recruited men into the RDS survey, with the number of waves originating from each seed ranging from three to 16. A total of 1141 people (including the 10 seeds) were assessed for eligibility over a period of 54 days (8th March – 30th April 2010). No new coupons were distributed after day 47. 196 men attended but were ineligible, 16 were eligible but had already been recruited, two were eligible but did not give consent, and 927 were eligible, consented and were recruited.

Fifty three of the 54 members selected for qualitative interview were interviewed, consisting of 10 out of 10 RDS Recruits, 10 out of 10 RDS Refusers, 10 out of 10 Village Residents (three men and seven women), 10 out of 10 Key Informants, nine out of 10 RDS Interviewers (the declining RDS interviewer was 'too busy') and four out of four GPC Interviewers. During analysis of the qualitative interview data and cross-checking with the quantitative RDS survey data, four of the 10 RDS Refusers selected and interviewed were found to be ineligible (they had in fact been recruited during the RDS survey) and therefore the data on these men was removed from the analysis, resulting in data from six interviews on RDS Refusers being available for analysis. The final qualitative sample size was 49.

The results below are split into five sections: the communities understanding of the purpose of the study, community understanding of the eligibility criteria, communities' views of the RDS interviews, communities' views on the incentive scheme, and community experiences of the RDS method of recruitment and subversion or coercion. Throughout these sections we evaluate (where possible) how people's understanding and experiences may have affected the quality of the RDS survey data and participation in other ongoing survey activities in this community.

Community understanding of the purpose of the RDS study

The data from the RDS Recruits suggested that their understanding of the purpose of the RDS study was very limited, even after study information was provided and consent obtained. Only two out of 10 RDS Recruits (from recruitment waves four and five) interviewed in the qualitative study showed any understanding of the purpose of the RDS study. Both said that they thought that the MRC was testing new ways of surveying people. Two of the Key Informers were seeds in the RDS study. Both said that they thought that the

MRC/UVRI was looking for new ways to recruit people. For instance, one said that he thought that the *“MRC had widened a way of getting information from people that could be left out by the GPC team by using a new method”* (Key informant #4, male, 41). Seven out of the nine RDS Interviewers said that the majority of recruits did not understand the purpose of the study when they came to the interview and five said that at least half of recruits had only been told about getting an incentive by their recruiters. However four also said that the seeds and the first zero to two waves of recruits had a good understanding of the study. When asked about what they thought the purpose of the RDS quantitative study was Recruits, RDS Refusers and Village Residents gave a wide variety of answers. These included that the MRC/UVRI was rewarding GPC participants, that the MRC/UVRI was testing a new method of recruiting people that may be used by the GPC in future as a way of reducing costs, that the study was about health-related issues and medical treatment, and that the study was about exploring people’s livelihoods.

The communities’ understanding of eligibility criteria

Comments made by a Key Informant and a Village Resident suggested that some RDS recruiters may not have thought of unmarried men as household heads, e.g. *“they [teenagers below the age of 18] were being left out because some of the older men didn’t take them as household heads because they didn’t have any wives at the time”* (Key Informant #1, male, 47). This may help to explain the under-recruitment of younger men in the RDS survey (14% of recruits were <30 years old compared to 22% of the target population, $p<0.0001$) (Figure 1).

There was a (mistaken) belief among the target population that only men who had participated in the last GPC survey round were eligible to participate in the RDS survey. Three of the RDS Interviewers, two RDS recruits and one Key Informant suggested that some previous GPC-refusers would now participate in future rounds in order for them to be eligible for any future studies that gave incentives. For instance, one RDS Recruit said that *“such people would fight hard to get registered in future”* (RDS Recruit #4, male, 40). This was supported by the common view reported by people interviewed for the qualitative study (4/10 Key Informants, 3/10 RDS Recruits, 2/6 RDS Refusers) that the MRC was rewarding people for having given blood as part of the GPC. This is likely to be responsible for the belief expressed by some interviewees that only men who had given blood in the previous GPC survey round were eligible for the study. For instance, when asked why he had not participated in the RDS study one RDS Refuser said that *“he had not given blood that round”* and that *“People told him that he could not participate if he had not given blood to the GPC”* (RDS Refuser #1, male, 33). This is supported by the fact that 74.2% of recruits had participated in the previous GPC survey round compared to 53.7% of the target population ($p<0.0001$).

Some people who were interviewed (1/9 RDS Interviewers, 2/10 RDS Village Residents, 5/10 Key Informants, 1/10 RDS Recruits, 1/6 RDS Refusers) mentioned that women felt that it was unfair that they were not eligible for the RDS study and were therefore unable to receive incentives. One Village Resident said that *“...maybe some women would not participate in future studies due to the fact that they felt left out”* (Village Resident #7, female, 62). Further, two RDS Interviewers, one Village Resident, and one GPC Interviewer reported that some men who believed that they were eligible for the RDS study were not allowed to participate, either because they were not household heads or due to errors in the MRC’s census database. One interviewer named three such men and said that *“... they got so bitter missing the study and swore never to participate in MRC activities”* (RDS Interviewer #7, female, 27).

Overall, our a-priori concerns that the RDS study may affect participation in the ongoing GPC survey activities may have been realised. Although the before-and-after study of participation in the annual GPC survey showed that there was no evidence for a difference in the proportion of new participants in the GPC between rounds 20 and 21 in villages surveyed before or after the RDS study (36.8% in 'before' villages, compared to 38.4% in 'after' villages, $p=0.5$), there was very good evidence that the proportion of new drop-outs of the GPC survey between rounds 20 and 21 was larger in villages surveyed after the RDS study than in villages surveyed before the RDS (total = 27.6% dropped out in 'after' villages compared to 22.0% in 'before' villages, $p<0.0001$) (Table 2). This suggests that overall the RDS study may have slightly reduced participation in the annual GPC.

The communities' views of the RDS interviews

Eight out of nine RDS Interviewers said that they and/or the RDS Recruits found the questions on contacts monotonous and tiring. During the RDS survey, RDS Recruits provided the detailed information on up to 50 contacts, with a mean network size of 12.1. One RDS Recruit, who provided details of 13 contacts, said that the interview was *"really long"* because *"the people he knew were too many and that it would take some good time to finish it"*. He said that he *"spent more than an hour at the site and felt a bit exhausted"* (RDS Recruit #1, male, 52).

One RDS Interviewer said that *"at first participants didn't comprehend that the more contacts they gave the more time they would spend during the interview but they realized later that those who gave few contacts would finish in a short time and they started giving few contacts"* (RDS Interviewer #6, male, 35), and three others gave similar opinions. RDS Interviewers reported using a variety of methods to deal with this problem. One said that *"they became strict and asked participants to give a minimum of ten contacts to receive a token of appreciation"* (RDS Interviewer #6, male, 35). Another said she told one recruit who gave details of only two contacts *"to go and think of any other people and come back later"* and the recruit then *"immediately remembered some others"* (RDS Interviewer #1, female, 26).

One RDS Interviewer and one Key Informant said that the questions asked about recruits' and contacts' tribes led to concerns in the target population, either because *"they thought non-Baganda by tribe will not be allowed to vote"* (RDS Interviewer #9, female, 26) or because *"they were worried and thought that at one point they may be eliminated from their villages because most of them are non-Baganda"* (Key Informant #4, male, 41). However, there was no evidence from the quantitative RDS study that this concern led to an under-recruitment of non-Baganda; overall Non-Baganda made up a slightly higher proportion of recruits compared to the target population (33.3% compared to 30.4%, $p=0.06$).

Communities' views on the incentive scheme

Most members of the community were positive about the incentives (2/10 Village Residents, 7/10 Key Informants, 9/10 RDS Recruits, 5/6 RDS Refusers), particularly the soap, and said that it greatly encouraged themselves and/or others to participate in RDS. The incentives were called 'tokens of appreciation' by the MRC/UVRI staff and community. One Key Informant said that *"The people really participated because they wanted to get the tokens of appreciation"* (Key informant #3, male, 39) and one Village Resident said that *"the token might look small to some people and big to others"* (Village Resident #2, female, 42). This was not universal however, and one community member said that he *"had been given a coupon but decided not to participate because he could buy the soap that was being given out"* (Village Resident #10, male, 27). This difference in the motivation to participate caused by the incentive may be the main reason why men in the highest socioeconomic group were

under-represented in the RDS sample compared to the target population, and men in the lowest socioeconomic group were overrepresented (men in the lowest group made up 21.4% of the target population and 26.6% of the sample [$p < 0.0001$] and men in the highest group made up 25.7% of the target population and 17.8% of the sample [$p < 0.0001$]).

The main reasons for agreeing to participate in the study reported by RDS recruits during the RDS survey itself (as opposed to the subsequent qualitative interviews) was that they *'like the MRC'* (49.3%, 452/917) or that they *'participate in MRC studies'* (27.9% 256/917) (Table 3). A further 12.1% (111/917) said that they agreed to participate because they *'wanted the token of appreciation'*.

Community experiences of the RDS method of recruitment and subversion or coercion

Seven of the 49 interviewees (3/10 Key Informants, 1/6 Refusers, 3/4 GPC Interviewers) said they thought the RDS method of recruitment was time saving and convenient for participants. They said that it would be easier to use than the current method used by the MRC/UVRI to mobilise and interview the villagers in the annual GPC survey where the team goes door to door, thus interfering with the villagers' work and other activities. One Key Informant said that, *"in this study, a person was willing to go at his convenient time which was good because you could plan on when to go and participate"* (Key informant #10, male, 75). Another said *"some local officials cannot convince people to participate in the study because people don't believe in them, yet friends believe in friends i.e. casual workers recruited casual workers and drunkards recruited drunkards too"* (Key Informant #7, male, 38). Seventeen out of 49 of the interviewees (2/10 RDS Interviewers, 5/10 Village Residents, 5/10 RDS Recruits, 4/6 RDS Refusers, 1/4 GPC Interviewers), however, said that they and/or others did not like having to travel to the RDS interview sites and therefore preferred to be found by interviewers at their homes.

A number of interviewees (5/9 RDS Interviewers, 3/10 Village Residents, 3/10 RDS Recruits and 3/6 RDS Refusers) expressed the opinion that the RDS method had led/would lead to an under-recruitment of people who were old, weak, or ill. For instance, one interviewer said that *"those who could not walk long distances could fail to participate, especially the old and ill"* (RDS Interviewer #8, male, 25). This was supported by three interviewers reporting that the sons or wives of sick men would turn up at the study sites to try (unsuccessfully) to participate and collect the initial incentive on their husband/father's behalf. Of recruiters who reported that men they offered coupons to had refused to accept them, 16% (8) said that the men had refused because they were *"too old"* or *"too ill"*. Ill health or old age may also have led to recruits refusing to become recruiters. 46.8% (36 out of 77) of recruits who refused to become recruiters said that they refused because they were old, weak and/or ill. However, this potential bias against the old was not seen in the RDS sample. Older men were over rather than under-represented in the RDS sample compared to the target population (aged 50+: RDS sample = 37.3%, target population = 29.7%, $p < 0.0001$, Figure 1). This was probably because there was a bias towards over-recruiting older household heads, and this bias was larger than any bias towards under-recruiting ill or weak older men.

Not surprisingly, RDS Recruits recruited later in the RDS study found it harder to find recruits. One Key Informant who had taken part in the RDS study said that *"they were almost the last people to know about the study and getting other male households to recruit was hard for them"* (Key Informant #5, male, 68). This is supported by the fact that the mean proportion of a recruiter's reported network who had already been recruited at the time of their interview increased from 13.8% in the first two weeks of the study to 38.9% in the last two weeks that coupons were given out, and that average number of recruits per

recruiter decreased steadily from 2.57 in the first week of the RDS survey to 0.62 in the last week that coupons were given out (McCreesh et al., In review).

There was little evidence of deals between recruiters and their recruits to share the incentives they were given. Only one person interviewed in the qualitative study said that he had heard of it happening. He said that it had occurred amongst very few people and that it was *“done as an appreciation for being given a coupon”* as *“some saw it as a favour being given a coupon and felt like returning the favour to such a person by at least giving him a quarter bar of soap they had received”* (RDS Refuser #8, male, 48).

Three RDS Interviewers reported that in three villages (villages G, L and U), some recruiters gave their coupons to the village chairman or local councillor to distribute. One RDS Interviewer suggested the recruiters (in village U) did this because *“maybe these recruits feared to return coupons to our RDS/MRC offices after having failed to get recruits”* (RDS Interviewer #4, male, 29) and another suggested that the recruiters (in village G) did this because *“they knew the chair[man] would get the people easily to give the coupons and also due to the fact that he was really respected”* (RDS Interviewer #2, male, 28). There was weak evidence from the quantitative RDS survey that the proportion of recruits who reported that they had no relationship with their recruiter was higher for recruits whose recruiters came from villages L and U, than for recruits whose recruiters came from village G or the other villages (village L: 20.5% (9/35), village U: 19.2% (9/38), village G: 4.0% (1/24), all other villages: 11.4% (91/710), $p=0.07$). There was no evidence, however, for any difference between villages in the proportion of recruits who incorrectly named their recruiter (village G: 4.0% (1/24), village L: 2.3% (1/43), village U: 0% (0/47), all other villages: 2.9% (23/778), $p=0.7$). All of these three RDS Interviewers reported that they did not know if the chairmen or local government official were being paid or rewarded in any way for distributing the coupons. No Village Residents, RDS Recruits, RDS Refusers, Key Informants, or GPC Interviewers reported having heard about the trading of coupons or coupon markets.

RDS Interviewers reported three cases where impersonation had occurred and been discovered. In one case, a young man tried (unsuccessfully) to be recruited, claiming to be a man who was listed as being around 60 years old in the GPC database. When asked about it, he said that the older man was his father. In another, an eligible man gave a friend permission to use his name. This was discovered when the man himself later tried to participate. In the third, *“someone came in from outside the study villages and gave them another person’s name that was in the study area. By the time the owner of the name came in, they had already taken his token”* (RDS Interviewer #9, female, 26). Four RDS Interviewers also mentioned that relatives would sometimes try to collect incentives on the behalf of eligible men, but without trying to deceive the RDS interviewer as to their identity. No Village Residents, RDS Recruits, RDS Refusers, Key Informants, or GPC Interviewers reported having heard about impersonation occurring. Ten of the 49 qualitative interviewees suggested that people would have been too afraid of discovery to attempt impersonation.

Nobody interviewed as part of the qualitative study reported having heard of ‘unreasonable’ threats or sanctions being used by recruiters to ‘force’ individuals to be recruited. However, one interviewer did report that *“some people when they came back to get their tokens on a re-visit, whenever they could find out that one of his recruits did not turn up, they would be angry saying that they were going to confront them and quarrel”* (RDS Interviewer #7, female, 27). Two RDS Interviewers also suggested that recruiters tried to recruit people that they could *“easily influence”* (RDS Interviewer #7, female, 27), either *“relatives whom they could easily push to participate because they had authority over them”* (RDS Interviewer #6,

male, 25) or “*drunkards*” who they would “*escort ... up to the office to see to it that they have participated*” (RDS Interviewer #7, female, 27).

Discussion

Our aim in this paper was to investigate, through an evaluation of RDS, whether the ethical, implementation and methodological criticisms levelled at this sampling method are justified. We used qualitative methods to gather people’s opinions of the method. Participant, and wider community members’, views in any study might influence reliability; the work of Scott (2008), cited above, suggests that this may be the case for RDS if certain forms of participant behaviour affected the conduct and use of the method.

Whether people involved in the study had a good understanding of what it was about, and therefore whether they had given informed consent, was one area of particular interest. Our findings suggest that the understanding of the study purported to be held by community members after their recruitment, a time when they were expected to have been given full information on the study, was limited. This finding raises questions about the best way in which to provide information so that the decision on whether to consent to participate may be based on accurate knowledge about the nature and purpose of the study. Semaan et al. (Semaan et al., 2009) in their review of HIV prevention studies which have used RDS, highlight the difficulties of obtaining informed consent from people who are a part of a marginal group. They stress the importance of the content and timing of informed consent when dealing with populations engaged in activities that put them, and their partners, at risk of HIV-infection. The informed consent process in our study consisted of providing information on our reason for undertaking the study, the methods employed, and the plans for sharing our findings. It is interesting to note that even with that seemingly straightforward procedure the interviewers reported that RDS participants, with the exception of seeds and the early recruits, often did not understand the essence of the study, nor were they particularly interested in obtaining more information. Often, it seemed that recruits provided reasons for the study based on their view of what they may have thought the study might be about (or hoped it was about), rather than the information provided in the information sheet. Similar concerns have been raised about informed consent comprehension in other studies (Flory & Emanuel 2004).

Our interest in participant behaviour was a direct result of concerns raised by Scott (Scott, 2008) that the approach was open to abuse by participants who could work the system for their own gain. He noted the market in RDS coupons which represented ‘a fiduciary instrument, a capital-producing commodity that creates its own market of buyers and sellers’ (p. 49). Some commentators (Broadhead, 2008; Lansky & Mastro, 2008; Ouellet, 2008) have criticised Scott for the way his research was conducted and the results interpreted, but the fact remains that the possibility of obtaining an incentive might be perceived to be a trade-able item that may distort the conduct of the research. Our results suggest that there was some evidence of subversion (impersonation) and that the distribution of coupons by someone other than the recruiter occurred. This may have affected the quality of data collected. However there was no evidence of the occurrence of coercion during the RDS study (which Scott reported was common in his study). Indeed, it seems that people willingly took part and the appreciation shown for the access to the incentives corroborates Semaan et al. (Semaan et al., 2009) observation that providing adequate remuneration shows ‘respect for participants’ time and effort in recruitment’. This observation is supported by Grady (2001) who gives her support for providing financial compensation to participants, finding no adverse effects on informed consent. However, our study was carried out in a population that is atypical for RDS studies, and it is possible that coercion may be more likely to occur in other populations, for instance injection drug users. Scott, Seeman, and

their colleagues were focusing on RDS used with injection drug users, and concerns that remuneration might in some way cause harm by enhancing access to drugs; the risk of promoting this type of harm was not a factor in our study.

When we consider the effectiveness of the method for recruitment of the target population, comments made by participants suggested that younger men, men who had not taken part in the previous survey round, men of higher socioeconomic status, non-Baganda, and old, weak or ill men may have been under-recruited into RDS. Quantitative data supported the views that there was under-recruitment of younger men, non-participants in the previous survey round, and men of higher socioeconomic status. However, the quantitative findings did not support the view that there was under-recruitment of non-Baganda or older men (although it is likely that ill older men were under-recruited). Collecting detailed information on each contact given by recruits may have led to later recruits naming fewer contacts. Strategies used by interviewers to reduce this problem varied. Both the fact that some recruits under-reported their network sizes and the varying strategies used by interviewers to try to reduce this may have biased the RDS adjusted estimates.

These biases are likely to be generalisable to other RDS studies. The bias in recruitment by socio-economic status is likely to be generalisable to most if not all RDS studies because different sub-groups of the target population are likely to be differentially incentivised by whatever (range of) incentives are offered. Men who had not taken part in the previous survey rounds are likely to represent a sub-population that is less willing to take part in medical research. Differential recruitment of individuals in the population by willingness to participate in surveys is likely to be a generalisable finding, but it is not limited to RDS. However, it is difficult to estimate the size of this bias using RDS data, as information on people who refuse to participate can only be obtained indirectly from the subset of recruiters who return to collect their secondary incentives. The bias in recruitment by age may not exist in other RDS studies, but it does highlight another challenge for RDS if the community understanding of target group membership does not correspond with the researcher's definition. As in this case, the bias may be quite subtle and difficult to detect. Quantification of the size of the bias would require triangulation with other sources of quantitative data, and the explanation for the bias may only become clear if qualitative data are collected.

A-priori concerns about the effect of the RDS on other studies, particularly GPC recruitment, were partially realised. Members of staff and community members gave reasons both why participation may increase and why it may decrease. Although there was no evidence for the RDS increasing participation in the GPC, there was very good evidence that participation in the GPC decreased more in villages surveyed after the RDS study than in villages surveyed before. This may have been caused by the RDS study, however it does not take into account confounding by village or confounding by the dates on which round 21 surveys took place. It is unclear what the longer-term effects of the RDS study on participation in future MRC studies will be. This is an area for further investigation.

In all research the possibility exists that participants do not provide accurate information to researchers. We detected some mis-information on numbers of contacts, as some participants tried to shorten the interview time. It is possible that some people did understand the purpose of the study but chose to say something else when they were asked what the study was about. This may have been because they thought MRC/UVRI had an undisclosed ulterior motive or they were simply suggesting a reason for the study based on their own expectations. As the work of many commentators shows (Bernard & Killworth, 1977; Bleek, 1987; Herzeld, 1983; Mays & Pope, 2000, among others) one needs to be aware of this tendency and seek to find ways to triangulate data, but it is unlikely that it can be eradicated in any research approach where people are asked to provide information.

Our aim was to use quantitative and qualitative methods to investigate community members' understanding and experience of respondent-driven sampling, and explore how these may have affected the quality of the RDS survey data, and participation in other ongoing survey activities in this community. While overall community members were positive about RDS recruitment and of incentives for participation, our findings suggest that because members of the community are the first point of contact between researchers and potential recruits, the use of RDS in medical research may result in increased difficulties in gaining informed consent, and data collected using RDS may be particularly susceptible to bias due to differences in the understanding of key concepts between researchers and members of the community.

Acknowledgments

The authors would particularly like thank the study participants and staff at the MRC/UVRI Uganda Research Unit on AIDS without whom this study would not have been possible. We are grateful to the Medical Research Council (UK), the Bill and Melinda Gates Foundation (19790.01), and the EU FP7 (242061) for funding.

References

- Bernard HR, Killworth PD. Informant accuracy in social network data II. *Human Communication Research*. 1977; 4:3–18.
- Bleek W. Lying Informants: A Fieldwork Experience from Ghana. *Population and Development Review*. 1987; 13:314–322.
- Broadhead RS. Notes on a cautionary (tall) tale about respondent-driven sampling: A critique of Scott's ethnography. *International Journal of Drug Policy*. 2008; 19:235–237. [PubMed: 18424109]
- Environmental Systems Research Institute. ArcGIS. Version 9.2. Redlands, CA:
- Goel S, Salganik MJ. Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences*. 2010; 107:6743–6747.
- Heckathorn DD. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems*. 1997; 44:174–199.
- Heimer R. Critical issues and further questions about respondent-driven sampling: comment on Ramirez-Valles, et al. (2005). *Discussion. AIDS Behav*. 2005; 9:409–413.
- Heimer R. Critical issues and further questions about respondent-driven sampling: comment on Ramirez-Valles, et al.(2005). *AIDS Behav*. 2005; 9:403–408. [PubMed: 16344920]
- Herzeld M. Looking both ways: the ethnographer in the text. *Semiotica*. 1983; 46:151–166.
- Lansky A, Mastro TD. Using respondent-driven sampling for behavioural surveillance: response to Scott. *The International journal on drug policy*. 2008; 19:241. [PubMed: 18439811]
- Malekinejad M, Johnston L, Kendall C, Kerr L, Rifkin M, Rutherford G. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. *AIDS and Behavior*. 2008; 12:105–130.
- Mays N, Pope C. Assessing quality in qualitative research. *BMJ*. 2000; 320:50–52. [PubMed: 10617534]
- McCreesh N, Frost SD, Seeley J, Katongole J, Ndagire-Tarsh M, Ndungutse R, et al. An empirical evaluation of Respondent-Driven Sampling. *Proceedings of the National Academy of Sciences*. In review.
- Microsoft Corporation. Microsoft Access 2003. Washington: 2003.
- Ouellet LJ. Cautionary comments on an ethnographic tale gone wrong. *International Journal of Drug Policy*. 2008; 19:238–240. [PubMed: 18434125]
- QSR International Pty Ltd. NVivo qualitative data analysis software. QSR International Pty Ltd; 2008.
- Ramirez-Valles J, Heckathorn DD, Vázquez R, Diaz RM, Campbell RT. From networks to populations: the development and application of respondent-driven sampling among IDUs and Latino gay men. *AIDS and Behavior*. 2005; 9:387–402. [PubMed: 16235135]

- Salganik MJ, Heckathorn DD. Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology*. 2004; 34:193–240.
- Scott G. “They got their program, and I got mine”: A cautionary tale concerning the ethical implications of using respondent-driven sampling to study injection drug users. *International Journal of Drug Policy*. 2008; 19:42–51. [PubMed: 18226516]
- Semaan S, Santibanez S, Garfein RS, Heckathorn DD, Des Jarlais DC. Ethical and regulatory considerations in HIV prevention studies employing respondent-driven sampling. *International Journal of Drug Policy*. 2009; 20:14–27. [PubMed: 18243679]
- Shafer LA, Biraro S, Nakiyingi-Miir J, Kamali A, Ssematimba D, Ouma J, et al. HIV prevalence and incidence are no longer falling in southwest Uganda: evidence from a rural population cohort 1989-2005. *AIDS*. 2008; 22:1641–1649. [PubMed: 18670225]
- StataCorp. Stata Statistical Software. Release 11.0. Stata; College Station, Texas: 2010.
- Volz E, Heckathorn D. Probability Based Estimation Theory for Respondent Driven Sampling. *Journal of Official Statistics*. 2008; 24:79–97.
- White, RG.; Seeley, J. Protocol for qualitative component of evaluation of respondent-driven sampling. London: 2010.

Figure 1. Map of study area showing location of target population and seed households and RDS interview sites

Different shades of grey are used to represent households in different villages. Each village has been labelled with a letter for confidentiality.

Table 1
Target population proportions and sample proportions

		Target population	RDS sample (n=927 including seeds)
Age group (years)	0-19	0.020	0.004
	20-29	0.202	0.133
	30-39	0.275	0.250
	40-49	0.206	0.240
	50+	0.297	0.373
Tribe	Muganda	0.697	0.667
	M'rwanda/kole	0.179	0.210
	Mukiga	0.017	0.021
	Murundi	0.047	0.061
	Other known/ unknown	0.060	0.040
Religion	Catholic	0.598	0.624
	Protestant	0.170	0.171
	Muslim	0.227	0.202
	Other known/ none/unknown	0.005	0.003
Socio-economic status	Highest	0.257	0.179
	Higher	0.249	0.242
	Lower	0.229	0.275
	Lowest	0.214	0.266
	Unknown	0.052	0.038
Village	A	0.033	0.021
	B	0.017	0.017
	C	0.042	0.070
	D	0.032	0.047
	E	0.027	0.027
	F	0.067	0.016
	G	0.025	0.028
	H	0.031	0.010
	I	0.060	0.045
	J	0.028	0.034
	K	0.031	0.045
	L	0.040	0.047
	M	0.026	0.033
	N	0.033	0.038
	O	0.049	0.062
	P	0.034	0.023
	Q	0.086	0.047
	R	0.038	0.055

		Target population	RDS sample (n=927 including seeds)
	S	0.038	0.038
	T	0.050	0.061
	U	0.050	0.065
	V	0.039	0.045
	W	0.040	0.033
	X	0.043	0.035
	Y	0.041	0.059
Number of sexual partners in the last years	0	0.113	0.148
	1	0.419	0.577
	2-3	0.114	0.140
	4+	0.037	0.035
	Unknown	0.316	0.100
HIV status	Positive	0.063	0.079
	Negative	0.600	0.817
	Unknown	0.337	0.105

Table 2
Participation in annual GPC before and after the RDS study

<i>New participants</i>					
	<u>Villages surveyed in R21 before RDS study</u>		<u>Villages surveyed in R21 after RDS study</u>		
	Number of R20 non-participants	Number (percent) of new participants	Number of R20 non-participants	Number (percent) of new participants	p-value
Men	301	109 (36.2%)	1337	541 (40.5%)	0.2
Women	286	107 (37.4%)	843	296 (35.1%)	0.5
Total	587	216 (36.8%)	2180	837 (38.4%)	0.5
<i>New drop-outs</i>					
	<u>Villages surveyed in R21 before RDS study</u>		<u>Villages surveyed in R21 after RDS study</u>		
	Number of R20 participants	Number (percent) of new drop-outs	Number of R20 participants	Number (percent) of new drop-outs	p-value
Men	622	141 (22.7%)	1337	384 (28.7%)	0.005
Women	814	175 (21.5%)	2098	564 (26.9%)	0.003
Total	1436	316 (22.0%)	3435	948 (27.6%)	<0.0001

Notes: R20 = GPC survey round 20. R21 = GPC survey round 21. 'New participants' were defined as residents who were eligible for the GPC in both round 20 and 21 who did not participate in round 20, but did participate in round 21. 'New dropouts' were defined as residents who were eligible for the GPC in both round 20 and 21 who did participate in round 20, but did not participate in round 21.

Table 3
RDS recruits (excluding seeds) responses to the question “*Why did you agree to participate in this [the RDS] study?*”

	Number	Percent
Likes the MRC	452	49.3%
Participates in MRC studies	256	27.9%
Wanted the token of appreciation	111	12.1%
Wanted to know more about the study	34	3.7%
Recruiter is a friend	18	2.0%
Other known reason	46	5.0%
Total	917	