Effectiveness of Respondent Driven Sampling in Engaging Methamphetamine users in HIV Prevention Research in Cape Town, South Africa

by

Stephen Mburu Kimani

Duke Global Health Institute
Duke University

Date: ______________________

Approved:

___________________________
Christina S Meade, Supervisor

___________________________
Melissa H Watt

___________________________
Eric P Green

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Duke Global Health Institute in the Graduate School of Duke University

2014
ABSTRACT

Effectiveness of respondent driven sampling in engaging methamphetamine users in HIV prevention research in Cape Town, South Africa

by

Stephen Mburu Kimani

Duke Global Health Institute
Duke University

Date: ______________________

Approved:

___________________________
Christina S Meade, Supervisor

___________________________
Melissa H Watt

___________________________
Eric P Green

An abstract submitted in partial fulfillment of the requirements for the degree of Master of Science in the Duke Global Health Institute in the Graduate School of Duke University

2014
Abstract

South Africa has a substantial HIV epidemic as well as a rising methamphetamine use problem, particularly in Cape Town. Respondent driven sampling (RDS) may be a useful tool for engaging vulnerable and hard-to-reach populations in HIV research, although its effectiveness has not yet been examined among South African methamphetamine users. The aim of the current study was to describe the effectiveness of RDS as a method for engaging methamphetamine users in Cape Town into a HIV behavioral research study. RDS procedures were used to screen 374 potential participants from a peri-urban township in Cape Town. Measures of homophily, equilibrium and RDS-1 estimators were computed for key demographic and social variables.

Beginning with 8 seeds, 345 methamphetamine users were enrolled over a 6 month period, with a coupon return rate of 67%. The sample included 197 men and 148 women who were ethnically diverse (73% Coloured, 27% Black African) and had a mean age of 28.8 years (SD=7.2). Social networks were adequate (mean network size >5) and mainly comprised of close social ties. Equilibrium on race was reached after 11 waves of recruitment, and after ≤3 waves for all other variables of interest. There was little to moderate preference for either in- or out-group recruiting in all subgroups.

Results suggest that RDS is an effective method for engaging methamphetamine users into HIV prevention research in South Africa. RDS may be a useful strategy for seeking high risk methamphetamine users for HIV testing and linkage to HIV care in this and other low resource settings. We also discuss future directions for RDS studies.
Dedication

To my mother—my mentor and biggest fan. For those hard days you stood by me and taught me the importance of persistence, consistency and patience.

You are dearly missed.
Contents

Abstract ................................................................................................................................. iv

List of Tables .......................................................................................................................... viii

List of Figures ......................................................................................................................... ix

Acknowledgements ................................................................................................................ x

1. Introduction .......................................................................................................................... 1

   1.1 What is respondent driven sampling? .............................................................................. 1

   1.2 Methamphetamine use is on the rise in South Africa ....................................................... 3

   1.3 South Africa is home to the world’s largest HIV epidemic ............................................... 3

   1.4 Methamphetamine use is linked to HIV risk behavior .................................................... 4

   1.5 Role of RDS as a tool in addressing burden of HIV ....................................................... 5

   1.6 Relevance in global health .............................................................................................. 6

   1.7 Description of the field site ............................................................................................ 6

   1.8 Overview of the Delft Connections study ...................................................................... 7

   1.9 Personal fieldwork experience ....................................................................................... 7

2. Manuscript: Respondent driven sampling is an effective method for engaging methamphetamine users in HIV prevention research in South Africa ................................................................. 9

   2.1 Introduction ...................................................................................................................... 9

   2.2 Methods .......................................................................................................................... 12

       2.2.1 Setting and participants ......................................................................................... 12

       2.2.2 Procedures .............................................................................................................. 12

           2.2.2.1 Formative research ......................................................................................... 12

           2.2.2.2 RDS procedures ............................................................................................ 13

       2.2.3 Measures ............................................................................................................... 14
List of Tables

Table 1: Characteristics of the seeds (N=8) and resulting recruitment ........................................ 18

Table 2: Crude sample characteristics of recruited participants by race and gender (N=345, seeds are excluded) ........................................................................................................................................ 20

Table 3: Adjusted demographic, HIV risk, and substance use sample characteristics (N=345) .... 22
List of Figures

Figure 1: Recruitment network diagrams for the eight seeds (highlighted with bold rim)............ 19
Acknowledgements

I attribute the success of this study to the love, support and criticism that I received from many people, at Duke and beyond. First, utmost gratitude goes to my mentors; Dr. Christina Meade for cordially allowing me to work with members of her research team and for patiently reading and interpreting boundless heaps of articles with me; and Dr. Melissa Watt who graciously groomed me from the beginning of my studies at Duke to reach higher heights professionally and otherwise. From you, I have learnt the art of mentorship and true service. Additionally, I salute the team members of the Delft Connections project who consistently pruned this study and assisted with data management. To our partners in South Africa (Donald, Des, Albert, Tembie and Mariana), this study would not have succeeded without your contributions.

I also acknowledge other members of staff at the Duke Global Health Institute: Dr. Eric Green for keenly reviewing this thesis; Michael Russell for coordinating my learning experience; Sarah Martin and Lysa Mackeen without whom my studies at Duke and travel to South Africa may not have been possible; and the leadership of the Masters of Science program for supporting my scholarship at Duke. I salute my classmates and friends who shared in my celebrations and offered a shoulder to lean on when the going got tougher. Last, but not least, I wish to thank my wife, Raha, for providing me with much needed moral support, love and delicious cocoa that made the late nights, cold winters and adjustment to life in Durham memorable.
1. Introduction

1.1 What is respondent driven sampling?

Respondent driven sampling (RDS) is an innovative variant of chain-referral sampling that relies on peers to identify and engage members of a hard-to-reach groups.\textsuperscript{1,2} Hard-to-reach groups may be defined by virtue of being small and extensively dispersed geographically or being involved in stigmatized and/or illegal activities.\textsuperscript{1} Examples of such groups include illegal immigrants, sex workers, men who have sex with men (MSM) and drug users.\textsuperscript{3-7} RDS begins by identification of initial respondents who serve as “seeds”.\textsuperscript{1} These “seeds” are usually selected by a research staff through convenience sampling. After completing the study visit and receiving compensation for their participation (primary incentive), the seeds are asked to recruit their peers into the study. Specifically, they are given recruitment coupons, usually two to three, to recruit others from the target population. In this way, they act as ‘recruiters’ and are given (secondary incentive) if their ‘recruitees’ complete the study. Recruiters are offered the same dual incentives to serve as recruiters themselves. This process creates an expanding system of chain-referrals characterized by ‘waves’ of recruitment until the target community is saturated or the desired sample size is reached.\textsuperscript{1,8} Estimation methods are then applied to compensate for bias because of the initial nonrandom sample selection process.\textsuperscript{9,10}

Like other chain-referral methods, RDS reckons that those best able to access members of hidden populations are their peers. It is, however, based upon numerous assumptions that must be met to ensure representativeness of the sample recruited.\textsuperscript{1,2,11-13} First, it requires that the target population be socially networked with reciprocal relationships between recruiters and their recruitees. That is, members must have social contact with each other as friends, co-workers,
acquaintances, family and so on. The pattern and extent of social contact determines the size of the area within which RDS sampling can be effective. Second, RDS assumes that respondents recruit randomly from within their social network. This is because RDS analyses are based on a Markov chain model that characterizes recruitment as a random, “memoryless” process where characteristics of recruitees depends only on their immediate recruiter and not on other recruiters upstream. Third, respondents should be able to accurately report their network size. Finally, the trait defining membership in the target population must be verifiable to avoid misrepresentation. While misrepresentation is not a problem unique to RDS, dual incentives may exacerbate this problem especially in low income settings and when these incentives are improperly calibrated.

Recently, these assumptions have been under a lot of scrutiny, given their centrality in RDS analyses. It has been proposed that given poor reliability of self-reported network size and potential for preferential recruitment, RDS studies must collect additional “egocentric” data to compare characteristics of participants’ social network characteristics against that of actual sample population. Egocentric data is “data pertaining to a node’s attributes, direct personal ties, and attributes of those ties. It focuses on the individual as a unit of analysis and requires respondents to list the name of ties in a given category”. This data is used to assess “randomness” of the recruitment as well as account for any bias associated with over- or underestimation of one’s network size.

The primary advantage of RDS over other chain-referral methods is that it controls for peer-recruitment biases by limiting the number of recruits per participant and through statistical weighting. By collecting information about participants’ social networks, one can calculate sampling weights that account for differential network sizes and recruitment patterns, making it possible to generate accurate population estimates. Population estimates adjusted in this manner
are referred to as RDS-1 estimators. A participant’s network size is defined as the number of other members of the target group that the participant knows and they know him/her.¹

1.2 Methamphetamine use is on the rise in South Africa

Methamphetamine is a highly addictive, synthetic psychostimulant that increases energy and feelings of euphoria, among other physiological effects.²⁰ It is the second most widely abused drug worldwide.²¹ In South Africa, methamphetamine use emerged in the late 1990s, fuelled by economic, social and political changes after the end of apartheid and has increased steadily in the past decade.²² The prevalence of methamphetamine use is highest in the Western Cape Province, with its epicenter in the city of Cape Town.²³

In a recent survey among individuals recruited from alcohol serving venues in a Cape Town township, 6.4% of participants reported methamphetamine use within the past 4 months.²⁴ This observation is in line with admission data from substance abuse treatment centers in the city that show a 100 fold increase in the proportion of patients reporting methamphetamine as their primary substance of abuse during a four-year period (2002-06).²⁵ This dramatic rise in prevalence of methamphetamine use has been described as the “the fastest increase in admissions for any particular drug ever noted in South Africa”.²⁵ Methamphetamine use is more common among adolescents and young adults, males, and Coloured persons (an ethnic group of historically mixed race people).²³,²⁶,²⁷

1.3 South Africa is home to the world’s largest HIV epidemic

South Africa is home to the largest HIV epidemic in the world. In 2012, an estimated 6.4 million South Africans were living with HIV including 18% of adults aged 15-49 years.²⁸,²⁹ An estimated 350, 000 new infections occurred among those aged 15 years or older and 240, 000
lives were lost to AIDS. There is marked gender, racial and regional differences in the prevalence rates of HIV. Among adults aged 15-49 years, females had a 10% higher prevalence of HIV, that peaked earlier compared to their male counterparts. This highlights the disproportionate burden of the disease that is still borne by women in South Africa. Black South Africans bear the biggest brunt of this HIV epidemic with 15.0% of the population living with HIV; the rate among Coloureds is 3.1%, Whites (0.3%) and Indians (0.8%). At the regional level the province of KwaZulu-Natal had the highest prevalence at 17% while the province of Western Cape had the least at 5.

Policy changes, including universal access to antiretroviral therapy, have greatly benefitted public health in South Africa. However, with a total of 370,000 new infections, HIV prevention efforts are essential to containing the epidemic. In particular, little attention has been paid to preventing HIV among newly emerging at-risk populations that includes drug users.

**1.4 Methamphetamine use is linked to HIV risk behavior**

Methamphetamine is a concern for HIV incidence in the Western Cape because it increases sexual desire and is associated with increased prevalence of risky sexual behavior and HIV infection. Data from Cape Town confirms that methamphetamine users in this region are more likely to engage in risky sexual behaviors compared to non-users and that methamphetamine is commonly used with sex to augment the sexual experience. This dual epidemic of methamphetamine use and HIV is igniting fears that HIV infections in Cape Town may increase, especially in population groups with historically low HIV prevalence rates. This has led to a call to prioritize strategies that promote engagement of methamphetamine users...
in South Africa in research necessary for tracking the direction of the HIV epidemic and for planning effective responses.\textsuperscript{38}

\textbf{1.5 Role of RDS as a tool in addressing burden of HIV}

RDS studies have been extensively implemented among injection drug users (IDU), MSM, sex workers (SW) and heterosexual men and women with multiple partners.\textsuperscript{4,39,40} In these studies, RDS has been found to be an effective recruitment technique, when designed and implemented appropriately. This has led to research about its effectiveness in different socio-cultural settings and among specific at-risk populations.\textsuperscript{5,16,41,42} Findings from these effectiveness studies have shown that race and age composition in a target population markedly affect effectiveness of RDS recruitment. For instance, Wang et al noted that the dual incentive system did not work out as well as anticipated among young “Ecstasy” users in central Ohio.\textsuperscript{5} They postulate that the nature of pre-occupation of the respondents, and strong negative inter-ethnic and/or inter-age affiliation could have affected the RDS process in this setting. In their study on non-medical users of pharmaceutical opioids in the US, Daniulaityte et al failed to reach equilibrium in terms of ethnic composition and noted very strong in-group recruitment tendencies among White and African American respondents.\textsuperscript{41} Therefore, an expert panel reviewing use of RDS to recruit substance-using MSM agreed that implementation of RDS should incorporate socio-cultural elements to ensure effective recruitment.\textsuperscript{43} They recommended formative research using techniques such as focus-group discussions (FGD) and key informant interview in population where RDS has not been used before.\textsuperscript{43} In South Africa, RDS has been used to recruit high risk-populations such as heterosexual males with multiple partners, SWs and MSM.\textsuperscript{44-46} However, to the best of our knowledge, RDS has not been used with methamphetamine users in
The aim of this research therefore is to assess the effectiveness of RDS in recruiting and engaging methamphetamine users in HIV prevention research in South Africa.

### 1.6 Relevance in global health

RDS is becoming a crucial tool for HIV research and intervention given the emergence of data highlighting the central role of hard-to-reach populations in driving the rising incidence of HIV infections in parts of the world.\(^4^7\) Notably, there has been extensive application of RDS in HIV research among injection drug users (IDUs).\(^4^,4^8^,4^9\) This is probably because of the direct relationship between injection drug use and HIV that has led to the development of interventions such as needle exchange programs in this population. However, non-injection drug use is playing an important role in the HIV epidemic and linking these users to HIV prevention efforts remains paramount is mitigating the HIV epidemic especially in high prevalence settings.\(^3^2^,5^0\)

### 1.7 Description of the field site

This study was conducted in Delft, a township located approximately 15 miles from the downtown Cape Town, South Africa. Delft is unique in that it was established as a racially integrated township following the end of apartheid in 1994, with residents mainly drawn from Black African and Coloured races.\(^5^1\) According to South African census data, a majority of its 150,000 residents, are unemployed and living in poverty. This setting was ideal for the study given the study team’s history of community engagement and rapport that was based on past collaborative research. Consequently, this study utilized research infrastructure established for a concurrent study on HIV risk among patrons of alcohol-serving venues (R01-AA018074).

The study office was located inside the Delft South public library. This library was located in a larger compound that also housed a community hall where community members
frequently gathered for entertainment activities including games, movies, plays etc. This set up was ideal for the study to limit public scrutiny and possible identification of the study participants. Located along the Delft Main Road, it was easily accessible for participants. The study office was fitted with private spaces for clinical and in-depth qualitative interviews, desks (with computers and noise-cancelling headphones) for conducting computerized assessments, and toilet facilities for drug screens. Our field workers were well trained in the conduct of community-based research with substance users.

1.8 Overview of the Delft Connections study

This study was conducted as part of a larger HIV behavioral study (The “Delft Connections”) funded with the support of an AIDS-Science Track Award for Research Transition (A-START) grant (R03-DA03828, PI Dr. Christina Meade). “Delft Connections” was a mixed-methods study aimed at examining patterns of drug use and HIV risk behavior among methamphetamine users in a peri-urban township in Cape Town, South Africa. It was implemented in three phases comprised of: Formative research to inform the recruitment strategy and pilot the assessment battery (Phase 1); Quantitative survey to characterize the sample and examine multi-level correlates of HIV risk behavior (Phase 2); and in-depth qualitative interviews to obtain narrative accounts of how methamphetamine use and sexual risk behaviors intersect (Phase 3). This study used data drawn from Phases 1 and 2 of the “Delft Connections” project.

1.9 Personal fieldwork experience

Throughout the academic year (2012-2013), I assisted in creating the study protocol, identifying appropriate measures, and designing the recruitment and referral process. Particularly, I was tasked with reviewing literature on RDS techniques and ensuring that appropriate
procedures were incorporated in the study protocol for training field staff. Subsequently, I spent three months in Cape Town where I oversaw the implementation of the data collection process, provided day-to-day management of the project, and supported three study staff in debriefing after interviews with participants. My day-to-day tasks included management of the recruitment tracking process, data security and transfer to the US, and monitoring adherence to the study protocol. I also shadowed Dr. Donald Skinner, Research on Health and Society Unit, Stellenbosch University to several meetings and fora where I learned more about collaborative community-based public health intervention programs.
2. Manuscript: Respondent driven sampling is an effective method for engaging methamphetamine users in HIV prevention research in South Africa

2.1 Introduction

Respondent driven sampling (RDS) is a variant of chain referral sampling that relies on peers to recruit diverse samples from the target population.\(^1\)\(^,\)\(^2\) It is useful for engaging members of hard-to-reach groups, characterized by involvement in stigmatized and/or illegal behaviors.\(^1\)\(^,\)\(^12\) The primary advantage of RDS over other chain referral sampling strategies is that it employs statistical estimation methods to limit biases that may arise from peer-driven recruitment.\(^1\) In theory, RDS can generate unbiased and accurate point-prevalence estimates for the population of interest.\(^9\) While some recent evaluations of RDS have suggested that prevalence estimates can be biased with large design effects,\(^14\)\(^-\)\(^17\) others have concluded that RDS is an effective sampling method for HIV surveillance of hard-to-reach populations when appropriately designed and implemented.\(^3\)\(^,\)\(^9\)\(^,\)\(^40\)\(^,\)\(^52\)\(^-\)\(^54\) RDS has been successfully used in diverse settings internationally with numerous socially marginalized groups, including undocumented immigrants, sex workers, men who have sex with men (MSM) and illicit drug users.\(^3\)\(^-\)\(^6\)\(^,\)\(^10\)

RDS recruitment begins with purposive sampling of initial respondents (“seeds”) from the target population. Once a seed completes the study assessment, he/she is compensated for participation (“primary incentive”) and then asked to recruit a pre-determined number of peers (usually 2 to 3) using recruitment coupons. The seed is rewarded with a “secondary incentive” if their recruit is eligible and enrolls in the study. Enrolled participants then serve as recruiters and are offered the same primary and secondary incentives. This procedure creates an expanding
system of chain referrals characterized by “waves” of recruitment until the target community is saturated or the desired sample size is reached.\textsuperscript{1,8}

In many parts of the world, hard-to-reach and socially marginalized groups play a central role in the rising incidence of HIV infections.\textsuperscript{47} RDS methodology has been utilized extensively in HIV research among injection drug users.\textsuperscript{4,48,49} However, non-injection drug use is also driving the epidemic, contributing to HIV transmission via risky sexual behaviors.\textsuperscript{32,55} Therefore, linking non-injection drug users to HIV prevention efforts remains paramount, particularly in high prevalence settings.\textsuperscript{20}

South Africa is home to the largest HIV epidemic in the world, with an estimated 6.4 million residents living with HIV in 2012,\textsuperscript{29} and is experiencing an emerging epidemic of non-injection methamphetamine use. In the Western Cape Province, where the methamphetamine epidemic is concentrated, the proportion of admissions to drug treatment facilities due to methamphetamine has increased from 0.8\% to 52\% in 2011.\textsuperscript{23} Community-based surveys in Cape Town confirm the high prevalence of methamphetamine use, particularly in densely populated urban township communities. For example, in a community-based sample of \textgreater 3000 individuals recruited from alcohol serving venues in one township, 6.4\% of participants reported methamphetamine use within the past 4 months, with rates three times higher among persons who were Coloured (a recognized group of mixed ethnicities) compared to Black African.\textsuperscript{24} It is feared that this increase in methamphetamine use may contribute to a new wave of HIV infections in the Western Cape.\textsuperscript{56}

Methamphetamine is mainly smoked in South Africa, so risks associated with injection use remain low.\textsuperscript{23} However, as a stimulant, methamphetamine increases sexual desire and is
associated with increased prevalence of risky sexual behavior and HIV infection.\textsuperscript{30-34} Data from Cape Town confirms that methamphetamine smokers in this region are more likely to engage in risky sexual behaviors compared to non-smokers, and that methamphetamine is commonly used with sex to augment the sexual experience.\textsuperscript{24,26,36,57} Given that methamphetamine smoking is most prevalent in Coloured communities, while HIV continues to disproportionately affect Black Africans, there is concern that this dual epidemic of methamphetamine and HIV may increase HIV incidence.\textsuperscript{37} This has led to a call to prioritize strategies that promote engagement of methamphetamine smokers in research necessary for tracking the HIV epidemic and planning effective responses.\textsuperscript{38}

Identifying and engaging methamphetamine smokers in targeted HIV research and prevention programs South African setting remains difficult because methamphetamine-related stigma leads users to hide their addiction for fear of prosecution and rejection from family and friends.\textsuperscript{58,59} By utilizing social networks and providing financial incentives for recruitment, RDS has the potential to successfully identify and engage methamphetamine smokers. This study describes the effectiveness of RDS as a method for engaging a cross-section of methamphetamine smokers in an HIV behavioral research study in a racially integrated township in South Africa. Additionally, it explores the effectiveness of RDS in reaching various sub-groups of methamphetamine smokers stratified by HIV risk profile (self-reported HIV status, HIV testing history, perceived HIV risk, willingness to test) and substance use characteristics (frequency of methamphetamine smoking and concurrent other drug use).
2.2 Methods

2.2.1 Setting and participants

This study was conducted during a 6 month period (May-October 2013) in Delft, a township located 15 miles outside the Cape Town city center. Delft was established in the early 1990s as a racially integrated township, with Black African and Coloured residents. The majority of its 150,000 residents are unemployed and living in poverty. The target population was adults who smoked methamphetamine. Specific eligibility criteria were: ≥18 years of age, residence in Delft, and current methamphetamine use (verified by urine drug screen). Exclusion criteria were: acute intoxication, impaired mental status, and/or inability to provide informed consent. Except for seeds, all participants were required to present a valid recruitment coupon.

2.2.2 Procedures

2.2.2.1 Formative research

Formative research was used to adapt the RDS strategy for use with methamphetamine smokers living in a township community. We assessed feasibility by evaluating characteristics of the social network (size, sociometric depth and composition); acceptability of proposed incentives; and survey logistics such as study office location, hours of operation, and duration for the study visit. Four focus groups with members of the target population, stratified by race and gender, were conducted (N=31, 7-8 per group). Participants were recruited using convenience sampling based on relationships established during prior research in Delft. Findings revealed that methamphetamine smokers: (1) are from all race, gender, and age groups; (2) often smoke methamphetamine daily; and (3) have well-established social networks with other methamphetamine smokers (i.e. they socialize and smoke together, with reported network sizes
ranging from 2 to 60). We did not identify any group that existed in isolation or without existing “bridges” to other methamphetamine users. All participants reported that they would be willing to invite at least two peers from their social network. These results gave us confidence that RDS procedures would be feasible to implement.

2.2.2.2 RDS procedures

Recruitment started with eight seeds stratified by race and gender who were identified by our field staff during the formative phase. After completing the study visit, each seed was given coupons to recruit two peers. The coupons listed the address and operating hours of the study site and a telephone number for questions.

Recruited peers came to the study site with their coupons. Those who met preliminary eligibility provided written informed consent and completed a urine drug test. Only those who tested positive for methamphetamine were eligible to proceed. Participants completed an audio computer assisted structured interview (ACASI) on sexual risk behavior and mental health, a clinical interview on drug addiction, and additional face-to-face questions about their social network and relationship to their recruiter. The full study visit took approximately 2 hours and each participant was given two coupons and instructed how to recruit new peers. All participants received the primary incentive of a grocery gift card worth ZAR 70 (~US$7), and had the potential to receive the secondary incentive of a ZAR 20 (~US$2) gift card for each of a maximum of two successful recruits. Each recruitment coupon had a unique serial number that was used to track the relationship between participants. Participants received referral information for local support groups and treatment facilities for HIV and substance use.
To manage the flow of participants, seed enrollment was staggered throughout the study period and initiated one at a time. With the exception of seed 8, we used systemized reduction of recruitment coupons as the study progressed to control the flow of participants and to ensure that all coupons could be redeemed during the study period: participants in waves 9-11 received only one coupon and those in wave 12 received zero.

Participants were provided pamphlets containing referral information for local support groups and treatment facilities for HIV and substance use. Study approval was obtained from the ethical review board at Duke University Health System and Stellenbosch University’s Health Research Ethics Committee.

2.2.3 Measures

All study activities were conducted in the language of the participant’s choosing (Afrikaans, Xhosa or English).

2.2.3.1 Demographics

Participants reported their age, gender, race, marital status, employment status, and level of education.

2.2.3.2 Social network characteristics

Three questions were used to assess network size: “1) Think about the people in Delft, who you know by name and they know you by name. Of these people, think about the ones who use methamphetamine. How many people are these? 2) Of these people, how many are 18 years or older? 3) Of these people, how many have you seen in the last 1 month?” The answer to the third question was used as the measure of social network size. Recruits described their
relationship to their recruiter (friend, romantic partner, family member/relative, or other) and the perceived strength of the relationship (very close, somewhat close or not close). Participants were also asked: “Where did this person first ask you to join the study?” and “Did you give this person anything in exchange for the coupon?” Participants who answered affirmatively to the second question described in an open-ended format what was exchanged.

2.2.3.3 HIV testing and behavioral factors

The ACASI assessment asked about HIV testing history, HIV status, and willingness to test for HIV. The following item measured perceived HIV risk: “Based on your behavior over the past 3 months, how much do you think you are at risk for getting HIV?” Response options were “not at risk”, “a little bit at risk”, “somewhat at risk”, and “very much at risk”. For analyses, responses were dichotomized into either “no risk” or “any risk.” Participants were also asked about number of sex partners in the past 3 months, participation in any transactional sex involving methamphetamine in the past 3 months, concurrent use of marijuana (“dagga”) and methaqualone (“buttons” or “mandrax”) in the past 1 month, and frequency of methamphetamine use (defined as number of days of use in the past 1 month).

2.2.4 Data Analysis

Effectiveness of RDS recruitment was determined based on the following outcomes: adequacy of social network ties, recruitment tendencies (network homophily), and attainment of “equilibrium” for our key demographic variables of race and gender. We defined adequate social ties as mean network sizes ≥3. Network homophily values range from -1 (exclusive “out-group” recruitment, or tendency to recruit from outside their own groups) to +1 (exclusive “in-group” recruitment). Values close to 0 suggest that social ties among participants cross networks,
overcoming biases introduced by preferential in- or out-group recruitment.\textsuperscript{1,9} Equilibrium refers to the state in which distribution of sample population estimates converge and does not change during subsequent waves.\textsuperscript{1} For this study, sample population proportions were considered at equilibrium when the change in population proportions between waves was <2%.

We computed adjusted proportion estimates with 95% confidence intervals and adjusted mean network size. The RDS-adjusted confidence intervals were computed using enhanced data smoothing algorithm for bootstrapping with 15,000 bootstrap samples per interval estimate. We report the standard RDS-1 estimator that accounts for network size and recruitment pattern between subgroups.\textsuperscript{2,9} Respondent Driven Sampling Analysis Tool (RDSAT) version 7.1.38 (Cornell University, Ithaca, NY, USA) was used to calculate these measures.\textsuperscript{61} Recruitment diagrams were created using NetDraw 2.136 (Analytic Technologies, Harvard, MA). Stata version 12.1 (Stata Corporation, College Station, TX) was used to prepare the dataset for analysis and compute descriptive statistics.

\textbf{2.3 Results}

\textbf{2.3.1 Sample characteristics}

Table 1 describes the eight seeds, and Figure 1 illustrates the recruitment chains. The most productive seed (seed 6) resulted in 11 recruitment waves and 146 participants. Two others (seeds 1 and 2) recruited >50 peers. In total, 555 coupons were distributed, with 374 coupons returned (return rate of 67.4%). Of the 374 individuals who presented with a coupon, 29 were not eligible, leaving a final sample of 345. Reasons for ineligibility were: no reported methamphetamine use in the past week or negative drug screen (n=26), <18 years old (n=1), impaired mental status (n=1), and refusal to complete the assessment (n=1). Table 2 shows the
crude sample characteristics. The sample was 73% Coloured and 57% male with a mean age of 28.8 years (SD=7.2). Coloured participants used methamphetamine more frequently and had been using regularly for more years. Black African participants were younger and more likely to be male, employed, and unmarried. The majority of the participants used other drugs concurrently (64% reported methaqualone use; 78% reported marijuana use), with no differences by race.
Table 1: Characteristics of the seeds (N=8) and resulting recruitment

<table>
<thead>
<tr>
<th>Seed</th>
<th>Age</th>
<th>Race, Gender</th>
<th>Self-reported HIV status</th>
<th>Network size</th>
<th>Number of enrolled recruits</th>
<th>Number of waves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed 1</td>
<td>40</td>
<td>Coloured, Female</td>
<td>Negative</td>
<td>100</td>
<td>50</td>
<td>8</td>
</tr>
<tr>
<td>Seed 2</td>
<td>29</td>
<td>Coloured, Male</td>
<td>Negative</td>
<td>5</td>
<td>114</td>
<td>12</td>
</tr>
<tr>
<td>Seed 3</td>
<td>23</td>
<td>Black, Female</td>
<td>Negative</td>
<td>10</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Seed 4</td>
<td>41</td>
<td>Coloured, Female</td>
<td>Unknown</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Seed 5</td>
<td>18</td>
<td>Black, Female</td>
<td>Negative</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Seed 6</td>
<td>40</td>
<td>Coloured, Female</td>
<td>Negative</td>
<td>2</td>
<td>146</td>
<td>11</td>
</tr>
<tr>
<td>Seed 7</td>
<td>29</td>
<td>Black, Male</td>
<td>Negative</td>
<td>35</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Seed 8*</td>
<td>21</td>
<td>Black, Male</td>
<td>Negative</td>
<td>16</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

*Seed 8 was initiated at the very end of the study period when supplies were limited. To balance between number of enrolled recruits and number of recruitment waves, this seed and subsequent recruits were given only one coupon, and recruitment was not cut off after 12 waves.
Figure 1: Recruitment network diagrams for the eight seeds (highlighted with bold rim)
Table 2: Crude sample characteristics of recruited participants by race and gender (N=345, seeds are excluded)

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 345)</th>
<th>Coloured (n = 252)</th>
<th>Black (n = 93)</th>
<th>Pearson’s $\chi^2$ or t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, M (SD)</td>
<td>28.8 (7.2)</td>
<td>30.2 (7.2)</td>
<td>25.3 (5.6)</td>
<td>5.9106***</td>
</tr>
<tr>
<td>Male, n %</td>
<td>200 (57.1)</td>
<td>124 (48.8)</td>
<td>76 (79.5)</td>
<td>25.6322***</td>
</tr>
<tr>
<td>Completed primary school ed., n (%)</td>
<td>168 (48.7)</td>
<td>126 (50.0)</td>
<td>42 (45.2)</td>
<td>0.6366</td>
</tr>
<tr>
<td>Employed (part- or full-time), n (%)</td>
<td>61 (17.7)</td>
<td>34 (13.5)</td>
<td>27 (29.0)</td>
<td>11.2711***</td>
</tr>
<tr>
<td>Married, n (%)</td>
<td>47 (13.6)</td>
<td>40 (15.9)</td>
<td>7 (7.5)</td>
<td>4.0212**</td>
</tr>
<tr>
<td>Days of methamphetamine use in past 30, M (SD)</td>
<td>23.5 (8.9)</td>
<td>24.7 (8.3)</td>
<td>20.2 (9.9)</td>
<td>4.2224***</td>
</tr>
<tr>
<td>Years of regular methamphetamine use, M (SD)</td>
<td>7.1 (3.6)</td>
<td>7.8 (3.6)</td>
<td>5.3 (3.2)</td>
<td>5.8224***</td>
</tr>
<tr>
<td>Concurrent other use in past 30 days, n %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methaqualone</td>
<td>221 (64.1)</td>
<td>164 (65.8)</td>
<td>57 (61.3)</td>
<td>0.4236</td>
</tr>
<tr>
<td>Marijuana</td>
<td>269 (78.0)</td>
<td>192 (76.2)</td>
<td>77 (82.8)</td>
<td>1.7255</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, ***p < 0.001
2.3.2 Social network characteristics

Table 3 summarizes adjusted population proportions and network sizes, network homophily, and number of waves required to reach equilibrium. With the exception of race, equilibrium on all variables was reached in ≤3 waves of recruitment, well below the maximum number of waves in each group. Equilibrium proportion for race was reached in 9 waves for Coloured participants and 11 waves for Black African participants. With the exception of race, network homophily indices ranged from -0.23 to +0.32, indicating minimal preference for either in- or out-group recruiting. In contrast, there was moderate preference for in-group recruiting among Black African and Coloured methamphetamine smokers (homophily indices=0.69 and 0.50, respectively). After adjusting for over-sampling of participants with large networks and differential recruitment by network size, participants had an average of at least 5 social ties to peers across various demographic, HIV risk, and drug use sub-groups.

Participants were mostly recruited by friends (89%), with relatively few participants recruited by family members (8%) and romantic partners (3%). Very few of the recruiter-recruitee relationships were sexual (8%). Most participants reported having smoked methamphetamine with their recruiter (80%), and about half felt very close to their recruiter (52%). While a majority of the recruitees had known their recruiters for >2 years (67%) and saw them daily (68%), there was a notable number of recent relationships (10% were <6 months old). Twelve participants reported that they had exchanged something for the recruitment coupon, most commonly the ZAR 20 secondary incentive. Most participants (64%) reported being recruited from their homes, which was also where 42% of participants smoked methamphetamine, indicating that recruiters actively sought out participants. Only 10% reported that they were recruited on the streets.
| Table 3: Adjusted demographic, HIV risk, and substance use sample characteristics (N=345) |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                | n              | Adjusted population proportion (95% CI) | Mean Network Size | Homophily | Waves |
| Race                            |                |                                           |                  |            |       |
| Black                           | 93             | 18.3 (10.1, 28.3)                           | 8.8              | 0.69      | 11    |
| Coloured                        | 252            | 81.7 (71.7, 89.9)                           | 5.5              | 0.50      | 9     |
| Gender                          |                |                                           |                  |            |       |
| Female                          | 148            | 32.3 (24.9, 40.7)                           | 7.8              | 0.31      | 2     |
| Male                            | 197            | 67.7 (59.3, 75.1)                           | 5.3              | -0.01     | 2     |
| Age                             |                |                                           |                  |            |       |
| Young (<30 years)               | 234            | 64.1 (55.8, 72.8)                           | 6.5              | 0.22      | 1     |
| Old (>30 years)                 | 111            | 35.9 (27.2, 44.2)                           | 5.4              | 0.05      | 2     |
| Education                       |                |                                           |                  |            |       |
| ≤Grade 9                        | 177            | 51.5 (42.6, 59.7)                           | 5.8              | -0.01     | 1     |
| ≥Grade 10                       | 168            | 48.5 (40.3, 57.5)                           | 6.5              | 0.10      | 1     |
| Marital status                  |                |                                           |                  |            |       |
| Unmarried                       | 298            | 85.8 (80.0, 91.2)                           | 6.2              | 0.11      | 1     |
| Married                         | 47             | 14.2 (8.8, 20.0)                            | 5.8              | 0.05      | 1     |
| Any employment                  |                |                                           |                  |            |       |
| No                              | 284            | 85.1 (78.8, 90.6)                           | 5.9              | 0.03      | 2     |
| Yes                             | 61             | 14.9 (9.4, 21.2)                            | 7.2              | 0.20      | 3     |
| Ever tested for HIV             |                |                                           |                  |            |       |
| No                              | 70             | 27.2 (19.1, 35.6)                           | 4.5              | 0.01      | 2     |
| Yes                             | 275            | 72.8 (64.4, 80.9)                           | 6.7              | 0.33      | 1     |
| Self-reported HIV status        |                |                                           |                  |            |       |
| Negative/Unknown                | 205            | 93.2 (87.7, 97.3)                           | 6.7              | 0.01      | 1     |
| Positive                        | 16             | 6.8 (2.7, 12.3)                             | 7.2              | 0.08      | 2     |
| Willing to test for HIV         |                |                                           |                  |            |       |
| No                              | 60             | 19.1 (13.1, 26.8)                           | 6.0              | 0.09      | 2     |
| Yes                             | 255            | 80.9 (73.2, 86.9)                           | 6.1              | 0.11      | 1     |
| At risk for HIV (Perceived)     |                |                                           |                  |            |       |
| No                              | 172            | 54.8 (45.6, 64.5)                           | 6.0              | 0.15      | 2     |
| Yes                             | 142            | 45.2 (35.5, 54.4)                           | 6.1              | 0.17      | 2     |
| Multiple sexual partners        |                |                                           |                  |            |       |
| No                              | 237            | 70.3 (62.9, 77.2)                           | 6.0              | -0.02     | 0     |
| Yes                             | 108            | 29.7 (22.8, 37.1)                           | 6.4              | 0.03      | 0     |
| Any transactional sex           |                |                                           |                  |            |       |
| No                              | 197            | 54.5 (46.3, 63.1)                           | 6.4              | 0.16      | 1     |
| Yes                             | 148            | 45.5 (36.9, 53.7)                           | 5.8              | 0.08      | 1     |
| Daily methamphetamine use       |                |                                           |                  |            |       |
| No                              | 139            | 22.6 (17.3, 28.7)                           | 10.9             | 0.23      | 0     |
| Yes                             | 206            | 77.4 (71.3, 82.7)                           | 4.7              | -0.23     | 0     |
| Concurrent Mandrax use          |                |                                           |                  |            |       |
| No                              | 124            | 38.0 (29.6, 46.2)                           | 5.8              | 0.01      | 1     |
| Yes                             | 221            | 62.0 (53.8, 70.4)                           | 6.3              | 0.10      | 1     |
| Concurrent marijuana use        |                |                                           |                  |            |       |
| No                              | 76             | 17.9 (12.3, 23.9)                           | 7.5              | 0.05      | 0     |
| Yes                             | 269            | 82.1 (76.1, 87.7)                           | 5.8              | -0.05     | 0     |
2.4 Discussion

Responding to the emerging methamphetamine epidemic in South Africa, this study found that RDS is an effective strategy for engaging a large and diverse sample of active methamphetamine smokers, including those at high risk for HIV transmission, into HIV behavioral research. To date, studies in this setting have either surveyed broad cross-sections of the community or used admission data from treatment facilities to describe methamphetamine users. Yet, the vast majority of drug users remain hidden from the public and do not access substance use treatment, limiting representativeness of study findings. While RDS has been previously used to recruit high-risk populations in Cape Town, including MSM and heterosexual males and females with multiple partners, it had not been used to recruit methamphetamine users or other illicit drug users in South Africa. This study successfully recruited a community-based sample of active methamphetamine smokers who live in a low-income township community with a high prevalence of drug abuse, adding to our knowledge of the RDS process in a population central to the evolving HIV epidemic in South Africa.

RDS was an efficient recruitment strategy in this study. From just 8 seeds, we recruited 345 active methamphetamine smokers in 6 months. In a review of 128 studies that utilized RDS to engage high-risk populations internationally, including South Africa, the median sample size was 225 with an interquartile range of 152 – 360. Our sample size falls in the upper margin of this range, indicating that RDS was similarly effective in engaging methamphetamine smokers in Cape Town. Specific strategies were used to ensure the diversity of our sample. First, based on findings from our formative research, we purposively selected seeds who were representative of the racial composition in the township. Second, because seeds were staggered, we observed recruitment patterns in real time and were able to make adjustments. We quickly realized that,
despite being established as a racially integrated community, there were neighborhoods within Delft that were racially homogenous. In addition, given that McCreesh et al found that participants are more likely to refer peers who live near them,\textsuperscript{65} we selected subsequent seeds from geographically dispersed neighborhoods. Our staggered and adaptive implementation strategy resulted in a manageable flow of respondents at the study site, while still efficiently reaching our target sample size.

The ratio of referrals in this study (coupon return rate) was high (67.4%), indicating acceptance by the target population. As a result, 4 of the seeds resulted in more than 8 waves of recruitment, and all variables of interest reached equilibrium before the final wave. Theoretically, recruitment proceeded beyond wave 6 eliminates bias related to non-random selection of seeds.\textsuperscript{1} The long recruitment waves indicate that we reached deeper connections within the sampled networks, representing sufficient sociometric depth.\textsuperscript{66} Nevertheless, because equilibrium for race was reached in wave 11 for Black African participants, we recommend that future studies in mixed race communities in South Africa continue recruitment beyond wave 12 when possible.

Previous RDS studies in Cape Town have been implemented in racially homogenous townships.\textsuperscript{39,44,46} Our study is the first to implement RDS in a racially diverse township in South Africa. This gave us an opportunity to explore recruitment patterns between Black African and Coloured methamphetamine users. While participants tended to recruit peers from within their own racial group, our moderate homophily score for race implies that social ties did cross racial groups and that a single RDS sample is suitable.\textsuperscript{60,67} However, given potential selection bias by race, large samples are recommended when implementing RDS in racially diverse settings. Differences in recruitment by gender were also evident. While Coloured participants recruited equally from both genders, Black Africans were more likely to recruit males. Having observed
this in real time, we subsequently oversampled female seeds from the Black African community. This practice of “steering” recruitment has been described in previous studies.\textsuperscript{1,7} While the small number of Black African females in our final sample may suggest that we were unable to access this sub-group, it more likely reflects the fact that methamphetamine use is relatively less common among Black African women. In recent surveys conducted in township communities in Cape Town, Black African females had the lowest prevalence of methamphetamine use.\textsuperscript{24,68} However, given the high HIV prevalence rate among Black African females in South Africa, concerted efforts to engage this group of methamphetamine users, even if small in population size, are warranted. We recommend that future studies with methamphetamine users in this setting conduct focus groups to identify potential barriers to participation and employ a steering incentive to increase engagement of Black African women.

RDS yielded a sample of methamphetamine users diverse in HIV status, risk behaviors, and testing experiences. While a majority of participants had undergone HIV testing at some point in their lives, a quarter had never tested, 20\% of whom were unwilling to test. A third of enrollees reported having multiple partners and nearly half had exchanged sex for methamphetamine. In addition, our participants were typically unmarried and frequently used other illicit drugs. Recreational drug users have been identified as a key population urgently in need of targeted HIV prevention interventions.\textsuperscript{29} Given that RDS was able to reach this key population, it may be a useful strategy for enrolling a large sample of methamphetamine smokers into HIV prospective longitudinal studies and prevention trials.

Several limitations of the study should be noted. First, it is impossible to know how well our sample represents the population of methamphetamine users living in Cape Town townships. However, the demographic characteristics of participants in our study are very similar to that
reported in other studies, suggesting that RDS was able to capture a representative cross section of this population. Second, since we did not offer HIV testing, future studies are needed to estimate HIV seroprevalence in this population. Finally, we are unable to estimate the impact of preferential recruitment on our sample. Future studies should collect robust social network data on the “alters,” eligible individuals who refused study participation.

2.5 Conclusion

Methamphetamine smokers are critical to the evolving HIV epidemic in South Africa. This study demonstrates that RDS is an effective way to engage methamphetamine smokers into research aimed at understanding HIV risk behaviors. In addition, it highlights novel opportunities for harnessing established peer connections through RDS for the delivery of interventions to seek, test and link to treatment even the most at-risk methamphetamine smokers. Future implementation studies should examine the potential of RDS to deliver harm reduction interventions including HIV and addiction treatments. Thus far, methamphetamine users have not been fully targeted as a high risk group for HIV prevention and treatment in South Africa. RDS procedures offer a way to engage this group in order to understand their contributions to HIV transmission and to deliver tailored interventions to improve uptake of HIV testing and linkage to care.

3. Future directions

3.1 Standardized reporting in RDS studies

Substantial methodological heterogeneity among RDS studies has been observed. Many reports were also missing key data, which introduced uncertainty when computing pooled estimates such as design effects. To assess whether any RDS study generated representative data,
it may prove useful to establish certain key data that should be reported for each RDS study, as has been done for other quantitative research studies including randomized controlled trials. In general, RDS studies should report the following: (1) whether formative research was conducted, the quality and quantity of such research, and whether the population under study was found to be socially networked; (2) comprehensive description of eligibility criteria that differentiate self-report of group characteristics from verified ones; (3) how initial and replacement seeds were selected and how they were found; (4) the maximum number of allowable referrals per participant; (5) whether the recruiter-recruitee relationship was tracked; (6) length of time needed for data collection; (7) whether equilibrium was reached and for which variables; (8) how the sizes of participants’ social networks were measured; (9) whether survey data were adjusted using RDSAT or another statistical program; and (10) which RDS estimator was used. In addition, prevalence studies should report whether a design effect was used during calculation of sample size and the size of that design effect, as well as the sample size calculated versus the sample size attained. Now that RDS is being utilized in multiple international settings for routine Bio-Behavioral Surveillance studies (BBSS), there is increased need to assess its reliability in measuring HIV and risk behavior trends over time.  

### 3.2 Assessment of critical social network data for “alters”

RDS sampling transfers the sampling task (normally done by researchers) to the respondents. This means that the RDS data-generating process is largely outside the control and view of the researchers. As such, there is need to retrospectively observe the recruitment process by collecting information about how respondents’ recruited (or failed to recruit) peers. Because most participants are required to return to the interview site to collect an incentive for their success in recruiting others, RDS allows for these participants to provide data about the
individuals who refused to accept a coupon from a recruit. These data can also be used, at least to some degree, to measure rates of level of non-response or refusal to participate. However, specific analytical procedures are needed to identify and account for differential non-response in the results. These results may provide some insights that inform the interpretation of the RDS estimates. Therefore, future studies should explore analytical approaches that best utilize this data as well as the integration of those approaches into currently deployed statistical packages used in RDS analysis.

3.3 Integration of analytical codes into mainstream statistical software

RDS receives praise for being easy to implement and successful at producing representative samples. Proper RDS implementation requires adhering to numerous theoretical assumptions administered through well-planned and rigorous protocol. The procedures to analyze RDS data are not easily understood. In addition, analyzing data with RDSAT is not only challenging but also limiting in the sense that only a handful of estimates can be computed. This has made it difficult to conduct predictive analyses that employ methods such as regression and modeling using RDS-adjusted data points. Currently, codes for implementation of RDS analysis have been made available for Stata and R statistical packages. These codes make it give users more flexibility in data processing but still fall short of seamlessly allowing robust regression or modeling. Future research should endeavor to make RDS-adjusted weights more functional and user-friendly for predictive analysis.

3.4 Future role of RDS in implementation science

After nearly four decades of HIV research, a lot of evidence of best practices in HIV prevention and treatment has been generated. However, the population-level impact of these practices among marginalized and hard-to-reach populations remains minimal. This calls for
an urgent need to develop and evaluate effective mechanisms to deliver these interventions to the marginalized groups in a community setting. This form of implementation research will require a shift from traditional epidemiologic methods to ‘effectiveness’ trial studies at the population level. A novel application of RDS in the recruitment of participants for cluster-randomized trials (CRTs) in HIV research has been proposed. Utilized in this way, the goal of RDS-informed CRTs is to evaluate community-level effectiveness. We need to see more implementation research utilizing RDS in suitable target populations.

3.5 Utilization of geographic data in RDS

Participants in most RDS studies bear the responsibility of travelling to the assessment site (AS) and approaching other potential recruits. While the dual incentive serves to offset this cost, differences between members of the target population in their willingness and ability to travel to AS, and between recruits in their willingness and ability to recruit others, have the potential to bias the RDS results. In fact, some studies have reported little or no recruitment in certain parts of the target study area. In evaluating the role of spatial dimensions in RDS, McCreech and colleagues found that recruitment was concentrated around the AS and that distant contacts were less likely to be reported, and therefore recruited. While this did not bias estimates of their study, they observed that such recruitment pattern could result in bias in more geographically heterogeneous populations. Toledo and colleagues confirmed this observation in their study conducted in Brazil. Geographic analysis of RDS process provides an opportunity to assess the translation into practice of some basic assumptions of the method. Future studies should adapt geographic techniques as visualization tools for monitoring geographic representativeness, where feasible.
References


27. Wechsberg WM, Luseno WK, Karg RS, et al. Alcohol, cannabis, and methamphetamine use and other risk behaviours among Black and Coloured


60. Johnston LG, Whitehead S, Simic-Lawson M, Kendall C. Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations


