Implementation and evaluation of web-based respondent driven sampling as a method to incorporate ethnic minorities in survey research:
A national online survey among men who have sex with men

Fayaaz A M. Joemmanbaks, Method and Statistics, Faculty of Social Sciences, Utrecht University, f.a.m.joemmanbaks@uu.nl
Edith D. de Leeuw, Method and Statistics, Faculty of Social Sciences, Utrecht University, e.d.deleeuw@uu.nl
Peter G M. van der Heijden, Method and Statistics, Faculty of Social Sciences, Utrecht University, p.g.m.vanderheijden@uu.nl
Wim Zuilhof, STI AIDS Netherlands, Program MSM, Amsterdam, wzuilhof@soaaids.nl

13.08.2018

Abstract

Men who have sex with men (MSM) are a high-risk population in the global HIV epidemics. Therefore, an up-to-date study is needed to investigate the recent dynamics of MSM epidemics in the Netherlands. Earlier research using convenience sampling methods has shown that MSM with non-Western background are often underrepresented in survey research. This study focusses on sampling subgroups of MSM (e.g., non-Western and younger) by implementing Respondent Driven Sampling: A chain referral peer-driven sampling method. In the first phase of recruitment 80 initial respondents from non-Western backgrounds were invited to fill out the questionnaire and to invite other peers through a web-based RDS system. In total, 95 eligible respondents have completed the questionnaire. Most of these respondents are from 'other' non-Western countries (CBS standards), from a younger age group and higher educated. The men in RDS differ on key variables (prevention attitudes and HIV test behavior) from those in the national study. With many improvements, RDS is a useful sampling method to combine with convenience methods to sample a more representative group of MSM.

Keywords

Respondent Driven Sampling, Convenience methods, non-Western MSM, web-survey, HIV studies,

Acknowledgement

1. Introduction

On a yearly basis the National Institute for Public Health and the Environment (RIVM) reports on new HIV- and Sexually Transmitted Infection (STI) diagnoses in the Netherlands. In their most recent report they indicated that the diagnoses of Dutch men and women with STI's increased from 17.2% in 2015 to 18.4% in 2016 (Visser et al., 2017). Also, in the same study 285 new HIV diagnoses were detected, where 93% were among Men who have Sex with Men (MSM).

MSM are considered a global risk population in HIV epidemics (Baral, Sifakis, Celghorn & Beyrer, 2007). Beyrer et al. (2012) showed that the high probability of HIV transmission through receptive anal intercourse plays a central role in these epidemics. HIV transmission factors are not only explained by the interaction between two or more individuals, but they also depend on the structures of social networks (Bengtsson, Lu, Liljeros, Thanh & Thorson, 2014). Early evidence suggests that also racial differences in HIV infection rates among MSM may depend on multiple group factors (Millett, Flores, Peterson & Bakerman, 2007).

A lot is still unknown about the social structures of HIV transmission factors. Therefore, there is an urgent need to study the underlying dynamics of MSM epidemics on individual level, as well as within social networks. In order to conduct a study, it is necessary to have a representative sample of the MSM population. Having accurate prevalence estimates of risk behaviors is key to understanding risks and developing research priorities to promote health and prevent disease within this population. In this paper we will further elaborate on appropriate ways of sampling 'hard-to-reach' minority groups, such as MSM.

1.1. Convenience sampling

In 2017-2018 STI AIDS Netherlands in cooperation with researchers of Utrecht University and RIVM are conducting a national online survey among the MSM population in the Netherlands. The goal of this project is to study potential risk factors regarding HIV- and STI transmission and also to collect other relevant data, regarding sexual attitudes and sexual behaviour. In order to reach this 'hard-to-reach' MSM population, sampling will take place through online forums (e.g., web-pages and web-communities) and off-line venues (e.g., personal contacts of healthcare professionals). However, applying convenience sampling methods has potential disadvantages that might lead to an unrepresentative sample of MSM.

Convenience sampling is a form of nonprobability sampling that does not use a random selection procedure (de Leeuw, Hox & Dillman, 2008). The chance of inclusion is unknown and inferences based on nonprobability samples cannot depend on statistical probability theory. As there is no well-defined sampling frame of MSM, sampling error is unknown.

Next to that are also other (survey) errors that may affect the estimated statistic of interest (Groves et al., 2009; de Leeuw, Hox & Dillman, 2008). The first is coverage error in which respondents without internet access are not included in the study, because an online survey requires respondents to have Internet access. Secondly, with nonprobability sampling it is not appropriate to apply statistical inference to generalise to the target population (de Leeuw, Hox & Dillman, 2008). Finally, from the ones who receive the invitation, not all are likely to participate, due to refusal or inability. This results in non-response error (de Leeuw, Hox & Dillman, 2008; Groves et al., 2009). After taking these errors of representation (coverage, sampling, and nonresponse) into account, it is uncertain whether the final group of MSM who participated in the study is representative for the MSM population in the Netherlands.

In particular, MSM with non-Western backgrounds and younger people were underrepresented when van Empelen, van Berkel, Roos & Zuilhof (2011) conducted a similar study on potential risk factors regarding HIV- and STI transmission, using convenience sampling methods. In general, the errors described above are even amplified in survey research among individuals with non-Western backgrounds (Kappelhof, 2017). Non-Western people are poorly understood because of the heterogeneity of immigrant populations, the fear to participate in epidemiological research and/or perceived ethnic/racial discrimination (Dovidio, Gluszek, John, Ditlmann & Lagunes, 2010). Therefore, we pay special attention in this study to incorporate

subgroups of the MSM population, who are otherwise not sampled with traditional convenience methods.

1.2. Respondent Driven Sampling

As an alternative sampling method to reach 'hard-to-reach' groups, Respondent-Driven Sampling (RDS) was developed in 1997 by Heckathorn. The main reason for developing RDS was to overcome sampling challenges for populations in which the sampling frame was difficult or impossible to define, such as for MSM, drug-users and sex-workers (Heckathorn, 1997). The RDS method was developed as part of an AIDS prevention intervention that targeted injective drug users in Connecticut. Nowadays many researchers increasingly implement RDS as a sampling method to reach hard-to-reach groups.

RDS is a variation on collecting and analyzing chain-referral data. Just as other chain-referral methods, RDS assumes that those best able to access other members of hidden populations are their own peers (Heckathorn, 1997). However, RDS is to distinguish from other chain-referral methods. One important feature is that RDS combines traditional snowball sampling with a mathematical model that weights for the fact that the sample was collected in a non-random way (Wejnert, 2010). In RDS each person sampled does not have the same probability of being included in the sample, because persons with larger network sizes have a greater chance of inclusion in the study. RDS takes this into consideration by weighting the data based on the reported network size (degree) of each respondent.

During the recruitment process of RDS, respondents are initially invited by researchers to participate in the study. These initial respondents are referred to as 'seeds' (Heckathorn, 1997). After participation, these seeds are in turn asked to recruit other eligible members (usually three or four) within their social network to participate in the study. These members are referred to as 'recruits' (Stein et al. 2015). The number of recruits is limited, so respondents with larger network sizes are not overrepresented in the study (Heckathorn, 2002). The process of recruitment continues until several waves of respondents are created and (ideally) long referral chains are formed.

However, so far only a few researchers have managed to reach long recruitment waves (Strömdahl, Lu, Bengtsson, Liljeros & Thorson, 2015; Bengtsson et al., 2012). It remains a challenge to encourage participants to invite others for (online) participation. Those who reached long recruitment waves, pointed out that factors such as adequate incentives, adaptive surveys to smartphones and tablets, and a heterogenous group of seeds were crucial in maintaining online referrals (Strömdahl, Lu, Bengtsson, Liljeros & Thorson, 2015; Bengtsson et al., 2012). Heckathorn (2002) pointed out that both ethnicity and gender of respondents affected within-recruitment and stimulated online referral. Many of these factors will be taken into consideration when implementing online RDS.

1.3. Current RDS study

In this study we will partially focus on the improvement of participant referral within RDS. In order to do so, we will conduct a methodological experiment. In the experiment the number of (requested) recruits will be manipulated. A random half of the seeds will be asked to invite two people, whereas the other half will be asked to invite four people. By doing so we expect to observe unequal patterns of online recruitment and thereby hope to reach long recruitment waves.

The current study has two main purposes: 1. Improving the recruitment process of RDS and thereby tapping deeper into the network of MSM subgroups 2. Evaluate whether RDS is a useful sampling method to reach (relevant information from) MSM subgroups, who are otherwise not sampled through the national study using convenience methods.

To pursue these goals a web-based RDS survey will be implemented with special attention given to sampling MSM with non-Western backgrounds from a young age group. These men will then be compared to the men sampled in the national study using

convenience methods and also to population registers of non-Western men living in the Netherlands.

In both the national online survey as well as RDS, respondents will receive similar questions. However, the RDS survey will be a shorter version of the questionnaire from the national online survey. The main reason to implement a shorter questionnaire in RDS is to lower the burden on respondents, as they need to be motivated to invite other peers for participation after having completed the questionnaire.

Through implementing RDS online, we aim to cross social boundaries of social networks. By tapping deeper into the social networks of individuals and (hopefully) reaching long recruitment waves, we expect to reduce underrepresentation of the target group. Finally, we tend to make empirically based recommendations on how to improve survey procedures using RDS and also on ways to incorporate ethnic minorities and younger people when applying RDS for health-related studies.

In the next section we provide a literature overview of RDS studies and the research questions of the current study.

2. Literature review

Web-based RDS studies among MSM were previously implemented in Vietnam and Sweden. These studies mainly sampled young and well-educated MSM of diverse socio-economic backgrounds (Strömdahl, Lu, Bengtsson, Liljeros & Thorson, 2015; Bengtsson et al., 2012; Bengtsson, Lu, Liljeros, Thanh & Thorson, 2014). However, some of these researchers failed to reach adequate sample sizes and long recruitment chains due to technical challenges with the web-based RDS system and relatively low incentives (Strömdahl, Lu, Bengtsson, Liljeros & Thorson, 2015).

Several methodological studies compared traditional (off-line) RDS with Internet surveys using convenience sampling methods (Johnston, Trummal, Lõhmus & Ravalepik, 2009; Evans et al., 2011). These studies reported that RDS reached a less diverse group of MSM considering the characteristics age, sexual orientation and HIV testing (Johnston, Trummal, Lõhmus & Ravalepik, 2009) and that RDS also failed to reach the required sample sizes and stable recruitment waves (Evans et al., 2011). Some explanations for this were a lack of trained recruiters, inadequate incentives and language problems.

2.1. RDS Online recruitment

A crucial element within RDS is for respondents to recruit other (eligible) peers for participation after filling out the questionnaire. Respondents need to recruit at least one other peer for a recruitment wave to generate (Wirtz et al., 2016). Until now there has been no common agreement or guideline on an appropriate number of recruits to ask for. The number of recruits is generally restricted to three or four to (1) avoid overrepresentation of people with larger networks, (2) limit the burden on respondents, and (3) force the sample recruitment chain to penetrate into the social network of seeds (Steijn et al., 2014). Inviting a large number of recruits could theoretically lead to a larger sample size and potentially a representative sample. On the other hand, it also requires more cognitive and practical effort, because respondents have to think of who they want to invite and select the best way to invite their peers.

Previous RDS studies that restricted the number of recruits to four, reached chains that varied in length between 1-29 waves (Strömdahl, Lu, Bengtsson, Liljeros & Thorson, 2015; Bengtsson, Lu, Liljeros, Thanh & Thorson, 2014; Stein et al., 2015; Bengtsson et al., 2012). To investigate the relation between the requested number of recruits and the length of recruitment chains, we will implement an experiment (2 versus 4 requested referrals) within the RDS study. This leads to the following research question:

1. Does a small (2) or a larger (4) request for referrals lead to an improvement of RDS recruitment?

2.2. Representativeness of RDS sample

For our second research question, we focus on the representativeness of the respondents sampled. According to Statistic Netherlands (CBS) (2017) around 10% of the Dutch population has a non-Western background. This population is divided over multiple subgroups, namely Turkish (18.5%), Moroccan (18%), Surinamese (16%), Antillean (7%) and other non-western (40.5%). In previous studies using convenience sampling in the Netherlands, researchers sampled less than 5% of non-Western MSM (Empelen, van Berkel, Roos & Zuilhof, 2011). On the other hand, the first RDS study that focussed on non-Western MSM in the United States, sampled a representative group of Latin and Black MSM with a large sample size (Murrill et al., 2016)

In 2011 van Empelen, van Berkel, Roos & Zuilhof employed nonprobability methods (convenience sampling) to sample MSM for an online survey. They reported that more than 33% of the sample was older than 45 years. This points to selectivity in the realized sample, possibly due to self-selection of older MSM during the recruitment and/or more non-response among younger MSM. On the other hand, Bengtsson, Lu, Liljeros, Thanh & Thorson (2014) were successful in sampling younger MSM by implementing web-based RDS among the Vietnamese population. These researchers enabled a much larger geographical coverage and allowed (online) participation of MSM who would otherwise not risk to physically access a survey office.

Finally, CBS (2017) described in their latest year report 'Integration' that on average the educational level of non-Western people living in the Netherlands is lower than the educational level of the local Dutch people. These findings are partially explained by the low educational level and socioeconomic status of their parents. Heath & Brinbaum (2007) confirm that ethnic minorities are more disadvantaged within the Western educational system. In 2011 van Empelen, van Berkel, Roos & Zuilhof mostly sampled well-educated MSM with Western backgrounds. By specifically focussing on non-Western MSM, the expectation is to reach respondents who are on average lower educated than those in the national study.

To investigate the representativeness of the RDS sample we will compare the estimates to population registers from Statistics Netherlands and also to estimates from the national study using convenience methods. This leads to the following research question:

2. Does the RDS approach lead to a representative sample of MSM when considering the background characteristics ethnicity, age and educational level?

2.3. Estimates of risk factors

In 2015 the Center Sexual Health (CSG) did HIV-testing among the Dutch population. They reported that HIV rates were relatively higher among some MSM groups with non-Western backgrounds, namely from the Netherlands Antilles and Latin America. In terms of HIV epidemiology, generally people with non-Western backgrounds from HIV endemic areas have a higher risk profile (van Veen, 2010). The higher risks are explained by a combination of individual factors, cultural norms and demographic factors. As a consequence, the groups in RDS might differ to the groups in the convenience sample on several key variables. This information is urgently needed, because less is known about immigrant subpopulations who remain 'hidden' from both researchers and practitioners (Rosario, Schrimshaw & Hunter, 2010).

In the present study we specifically want to evaluate whether RDS can reach respondents who differ on key variables from those who respond to convenience sampling methods. Therefore, our third research question focusses on whether there are differences on key variables between non-Western respondents in the RDS and in the national study (while controlling for the background variables age and educational level). The key variables that are investigated are sexual risk behaviours, HIV knowledge and test

behaviours. This enables us to assess whether RDS is a useful and necessary method next to the standard convenience sampling method. This leads to the third research question of the current study:

3. Are there differences in the substantive answers of the questionnaire (key variables) when comparing non-Western men sampled through RDS and sampled through convenience methods, taking into account potential differences in age and education?

In sum we will investigate the following research questions:

Research question 1: Does a small (2) or a larger (4) request for referrals lead to an improvement of the RDS recruitment? or more specifically:

1a. Does the instruction regarding number of referrals (2 versus 4) influence the number of people actually invited to the survey?

1b. Does the instruction regarding number of referrals (2 versus 4) influence the final response rate of the survey?

Research question 2: Does the RDS approach sample a representative group of MSM when considering the background characteristics ethnicity, age and educational level? or more specifically:

2a. Does RDS result in more non-Western MSM who are younger and less educated, compared to the convenience sample in the national online survey?

2b. Does RDS result in a similar distribution of males regarding ethnicity, age and education as the population registers (of non-Western men) from Statistics Netherlands?

Research question 3: Are there differences in the substantive answers of the questionnaire (key variables) when comparing non-Western men sampled through RDS and sampled through convenience methods, taking into account potential differences in age and education?

3. Methods

3.1. Design

In previous years and also in 2018 STI Netherlands took the initiative to conduct a national online survey on the sexual health of MSM. In order to reach this hard-to-reach group the researchers used convenience sampling methods. In the current study RDS was implemented as a smaller part of the national survey, in order to sample subgroups among MSM who are generally underrepresented in the larger study. In the next section, we will describe the global design of the national study and the design of the RDS in more detail.

3.1.1. National study using Convenience sampling

A national cross-sectional online survey was conducted in the Netherlands by STI AIDS Netherlands and researchers of Utrecht University and RIVM using convenience sampling methods. The survey gathered information about behavior, health and sexual experiences of MSM in the Netherlands. The recruitment period was between March 1, 2018 and June 1, 2018. An active campaign was set up, both online -and offline, to invite members of the MSM community to participate. The link to the survey was available on several websites and dating apps. For a detailed recruitment schedule, see Additional material: appendix A[1].

The questionnaire contained 150 questions and the duration of filling in was between 25-30 minutes. Furthermore, the questions were adaptive to smartphones and tablets and most questions were mandatory. Participants could win one out of twenty €25 bol.com gift vouchers by means of a lottery system when they completed the questionnaire.

3.1.2. Smaller study among MSM subgroups using RDS methods

3.1.2.1. General description

A smaller study was conducted mainly focusing on sampling MSM with non-Western backgrounds. This part applied a web-based RDS system for online recruitment and participation. The recruitment period was between March 26, 2018 and May 28, 2018. In the RDS study, researchers provided an online questionnaire that was a shorter version of the questionnaire provided in the national survey. The questionnaire consisted of a total of 50 questions. The duration of filling in this questionnaire was between 10-15 minutes. The questionnaire was divided over multiple pages and the questions were also adaptive to tablets and smartphones. In order to avoid missing data, most questions were mandatory (i.e., participants had to answer the questions, otherwise they could not proceed to the next page). Some personal questions, such as postal code, were optional (i.e., respondents could skip the question). Furthermore, the text for RDS was provided in Dutch, English, Arabic, Farsi, Turkish and French. Participation was entirely voluntary and anonymous. Respondents had the opportunity to unsubscribe from the survey and give reasons of doing so.

Participants were eligible for an incentive when they (1) completed the questionnaire *and* (2) successfully recruited eligible peers for participation (i.e., the peers also had to complete the questionnaire). A monetary incentive was offered by means of drawing a lottery from one out of ten €50 gift vouchers.

3.1.2.2. Population and inclusion criteria RDS

In the Netherlands almost 94.4% of the population from 12 years and older has access to the Internet (Statistics Netherlands, 2017). Therefore, access to the online survey was assumed not to be a major concern for most respondents. The target group of the RDS study is MSM with non-Western backgrounds in The Netherlands. Non-Western is defined as 1st and 2nd generation migrants from Latin America, Africa, Asia and the Middle East. Eligibility criteria are men, 16 years and older, who have ever had sexual intercourse (oral- and/or anal sex) with other men, and who currently live in the Netherlands.

3.1.2.3. RDS questionnaire

Content of the questionnaire
The RDS questionnaire contained fewer, but similar questions as the national online survey. Questions consisted of several demographic variables (e.g., ethnic background, age, educational level, and sexual orientation and – identity) and several key variables concerning sexual health: attitudes about prevention strategies (e.g., 'What do you think of the following ways to prevent HIV transmission'), sexual relations and behaviors (e.g., 'With how many men did you have condomless anal sex during the last 6 months') and HIV testing (e.g., 'When were you last tested for HIV'). See additional material appendix B[2] for the RDS questionnaire.

Questionnaire reduction
As mentioned earlier the questionnaire for RDS is a shorter version of the national online survey. The main reasons to shorten the RDS questionnaire was to lower the burden on respondents and at the same time keep them motivated to invite other MSM for participation. The questionnaire was shortened based on the rating of relevant theme's that were important to the researchers. Two researchers rated each theme, using a scale from 1 (May be omitted without (large) problems) –

10 (Should be kept in questionnaire at all costs). The themes (and corresponding items) with an average score of 9 and higher were selected for the final RDS questionnaire. Additional material appendix C[3] contains more detailed information about the shortened questionnaire.

Questionnaire translation

As mentioned above, the text for the questionnaire was provided in six different languages. The method used for translation was: Back Translation. This method is used since the 70's and involves translating the 'actual' questionnaire back into the source language (Behr & Shishido, 2016). Subsequent comparisons are made between the two source- language versions, in order to identify potential discrepancies. The translators were all local speakers who translated the questionnaire to their mother tongue (with the exception of English). The same questionnaire was translated back by an external person to the source language.

Furthermore, the questionnaire was also translated in context. This means that background information was given of the study purposes and the intention of providing the questions was explained to the translators. Also, the translators were provided with a brief introduction to the study with the possibility to contact the researchers if any questions would arise during translation.

3.1.2.4. Experimental manipulation

In order to optimize peer recruitment, we manipulated the number of referrals that participants were asked to invite. The independent variable (Y_i) is the experimental condition (2 versus 4 requested referrals). A random half of the seeds was asked to invite 2 eligible peers for online participation, whereas the other half was asked to invite 4 eligible peers.

$$Y_i \quad \begin{cases} Y_0 = 2 \text{ referrals} \\ Y_1 = 4 \text{ referrals} \end{cases}$$
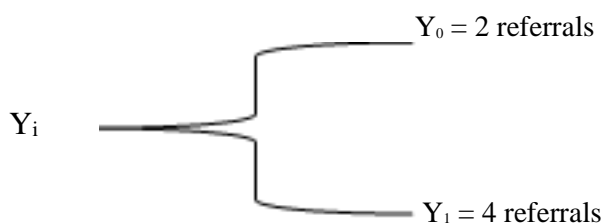
3.2. Implementation web-based RDS

3.2.1. Sampling and recruitment

Steijn et al. 2015 describe the recruitment strategy as following:

1. Initial participants are indicated as 'seeds'.

These seeds are eligible respondents from the target population. RDS proceeds by approaching seeds who are (mostly) familiar to the researchers (see section 3.2.2). Seeds are instructed about the study purposes, because they are the ones who initiate the RDS recruitment pattern. After they have been instructed by the researcher, they receive an invitation to fill out the questionnaire. Once they have completed the questionnaire, they are asked to recruit a number of eligible contact persons within their network to participate in the study.

2. These contact persons are indicated as 'recruits'.

Recruits are members of the target population who are in the same network as the seed who invited them. Each recruit can receive the invitation to the survey through a different route (see section 3.2.3). The sampling patterns repeats itself: after the recruits have completed the questionnaire, each of them in turn is asked to invite other eligible members within their network to participate in the study.

3. 'Waves' refer to consecutive subsamples, with seeds in wave 0 and recruits invited by seeds in wave 1, and so forth. After each new invitation has been send out, a new wave of respondents is created.

4. 'Recruitment trees' refer to chains of participant connected via recruitment.

A seed and his corresponding recruits are all connected within the same recruitment tree. The RDS sample is expected to contain many recruitment trees, whereby respondents in one recruitment tree are connected through a unique chain. For a graphical presentation of the RDS sampling structure, see Figure 1.
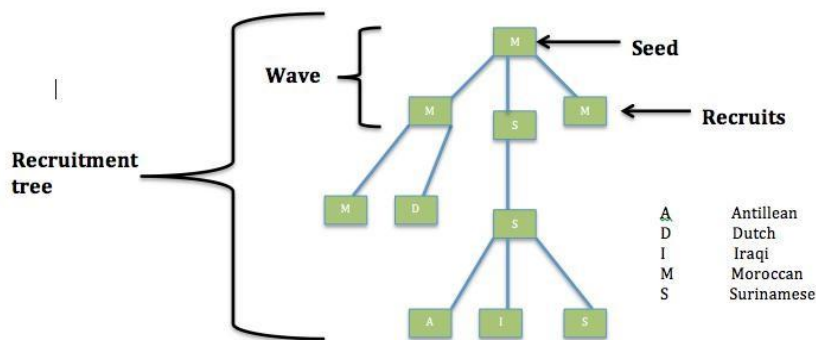

Figure 1. RDS Recruitment tree
Note. This is an example of one recruitment tree that has been generated through a unique seed. In RDS multiple seeds are invited and therefore multiple recruitment trees are generated.

3.2.2. Data collection

To start the RDS data collection in total eighty MSM seeds with diverse non-Western backgrounds were selected through lesbian, gay, bisexual, transgender, queer and intersex and other (LGBTQI+) organizations in the Netherlands and through personal contacts of the researchers. The recruitment of seeds occurred in 4 rounds; 20 seeds invited per round. In between every round there was a period of two weeks. After every round we evaluated the recruitment and response of RDS.

Seeds were mainly selected based on their ethnic backgrounds. As there are no population estimates on MSM, the total amount of seeds is based on the proportion of *all* non-Western men living in the Netherlands. Seeds were approached through different recruitment methods (i.e. Facebook, e-mail, phone call and face-to-face) and each were given a standard instruction letter that contained background information about the study purposes and a request to participate. Additional materials: appendix D[4] contains a detailed recruitment schedule and the instruction letter for the seeds. Finally, the ones who agreed to participate were sent an invitation email with a request to fill out the questionnaire.

3.2.3. Web-based RDS system

For facilitating and tracking online recruitment, we used a web-based RDS survey system that was provided by RIVM and originally implemented by Bengtsson et al., 2012. Participants had access to the survey through a unique link that was send through an invitation email. These links were based on personal codes that were automatically generated by the system for each participant. The invitation email contained background information about the study, a request to participate and a personal link to the survey.

After completing the questionnaire, participants were given the opportunity to invite other MSM friends within their network. There were multiple options for inviting contacts: 1. Sending an invitation email indirectly through the system, 2. Invite people through Facebook private messages and 3. Invite people through Whatsapp. Participants who did not complete the survey were sent a reminder email after five days. Those who finished the survey but did not recruit other members were sent a reminder email after 7 days and after 14 days. After completing the (short) RDS questionnaire, a link was attached to the larger national survey.

3.2.4. Changes during fieldwork

Content of the RDS questionnaire
After evaluation of the first round of data collection several changes were applied to the RDS questionnaire for two main reasons: (1) prevent high-dropout rate of respondents and (2) Improve RDS recruitment. Changes were applied based on additional paradata of the survey (i.e., pages with the highest dropout rates). One example is that many respondents dropped-out when they had to fill in their mother's-, fathers- and own birth country via a dropdown menu. Asking for personal information in combination with a dropdown menu consisting of 223 response categories and mandatory to answer, has most likely resulted in high drop-out rates on this particular page. In order to prevent this we added clarifying instructions (e.g.,"you can type in the first three letters, so you do not have to scroll all the way to the bottom") and removed 73 uncommon categories to make the dropdown menu much shorter. Secondly, we also ensured anonymity and confidentiality by explicitly adding clarifying sentences (e.g., "your data will be treated confidentially and will not be shared with others") to some personal questions. For an overview of all the changes, see Additional materials: appendix E[5].

Interviews with seeds
We also conducted two types of interviews with seeds to further improve the RDS data recruitment. The first type of interview is known as a cognitive (post) interview (Campanelli, 1997). During this interview the researcher fills out the questionnaire together with a respondent, in which he or she focuses on the cognitive processes that the respondent uses to answer survey questions. The aim is to evaluate sources of response errors and find out potential reasons of drop-out. Two strategies of cognitive post interviews are applied: (1) 'think aloud' in which the seed is explicitly instructed to think aloud as he answers the survey questions and (2) 'verbal probing' in which the interviewer 'probes' further into the basis of a response by asking additional questions.

The second type of interview were open interviews. These were conducted with seeds who did not complete the questionnaire and with seeds who completed the questionnaire but did not invite others. From each group respectively 5 seeds were randomly selected and asked the same questions, namely "Why did you not finish the survey" or "Why did you not invite others"? Based on the outcomes on these interviews, the cognitive-post interviews and the additional paradata final changes were applied to the questionnaire before the second round of data collection started (for all changes see, Additional material: Appendix E[5]).

3.3. RDS data analysis

3.3.1 Data cleaning

Linking respondents
Due to technical challenges within the web-based RDS system, a large amount of manual data cleaning had to be done. Firstly, the registration of completes had to be checked. Respondents who opened the questionnaire but finished at a later moment received a new personal code from the system. The system indicated 'not completed' for those respondents. We had to link their data to their initial code to keep track of the RDS recruitment. Secondly, a common mistake in RDS is for respondents to send their own link or a singular link to every peer they invite. The RDS system only tracks unique links that are generated. As a result, every respondent who participated through a same (seed) link is not linked to the person who actually invited them. To link every respondent within one recruitment tree, we applied these changes manually.

Participant removal
Similar to the national study using convenience methods, every person who completed all of the demographic variables and at least one of the key variables in the RDS study was included for data analyses. From the total amount of participants who were included, the ones who did not fit the inclusion criteria (e.g., Dutch men, women and people younger than 16 years old) were excluded from further analysis. Finally, respondents who took less than 3 minutes to answer the questionnaire were also excluded from the analysis, because it seemed impossible to read and answer all of the questions in such a short time.

3.3.2. Statistical computations

Dependency of the data
RDS has a hierarchical data structure: respondents are nested under recruitment trees. Thus, people within one recruitment tree are expected to be more similar than people from other recruitment trees. To account for this dependency the Intra Class Correlation Coefficient (ICC) was computed. The values of the ICC were lower than 0.10, as only 12.5% (N = 8) of the seeds generated recruitment waves. Therefore, for statistical analysis we analyses the data of all respondents on one level instead of applying multilevel analysis.

Network size (Degree)
In RDS each person sampled does not have the same probability of being included in the sample, because persons with larger personal networks have a greater chance of being sampled than those with smaller personal networks. RDS takes this into consideration by weighting data based on reported network size. Respondents were asked: "How many men could you invite within your network to participate in this study".

   Respondents reported having a network size ranging from 0-100. Only one person reported a network size of 1000. This data point was considered an outlier, as it highly influenced the estimates in the regression model(s). To account for potential bias, we recoded this value to the value of the person with the second highest network size of 100. To control for the (average) network size in the experimental condition, we added the network size as a dummy variable (0 = reported a network size of zero, and 1 = reported a network size higher than zero). Next to that we created a second variable where 0 was recoded to the average of the network size ($M = 7.042$) and the remaining respondents had their initial value.

Experiment
To investigate peer recruitment we build up two poisson regression models.
For research question 1a: the dependent variable is the number of people invited to the questionnaire (ranging from 0-4). The independent variable is the number of requested referrals (2 versus 4). For research question 1b: the dependent variable is

the number of people (invited by others) who responded to the survey. The independent variable is the number of referrals. In both regression models the variables age and ethnicity were included as covariates

Comparison between RDS, National study and CBS
To answer research questions 2a for the comparison between respondents in RDS and population registers from Statistics Netherlands (CBS) we computed descriptive analyses (total amount and proportions) of the variables ethnicity, age and educational level. Similar characteristics of men living in the Netherlands were retrieved from CBS Statline (Statistics Netherlands, 2017).

To answer research questions 2b for the comparison between RDS and the national survey using convenience sampling we computed similar descriptive analyses of the same background variables and used Chi-square (χ^2) tests to compare the differences between the categories in both studies.

To measure the (dis)similarity of the distributions from the background variables, we computed the Duncan Segregation Index. This index score can be interpreted as the percentage of respondents in the two studies that would have to move in order to produce a uniform distribution. The number ranges from 0 – 1, with 0 = perfect similarity and 1 = perfect dissimilarity.

To answer our third research question, we compared key variables 'sexual behavior', 'HIV prevention' and 'HIV testing' between RDS and the national study using linear and (multinomial) logistic regression models, while taking into account the background variables ethnicity, age and educational level.
Statistical analyses were performed in R (version 3.4.2).

4. Results

4.1 RDS Descriptive Statistics

In total 95 eligible participants completed the questionnaire. The age of the respondents ranged from 20 to 48 years old, with an average of $M = 29.53$, $SD = 5.91$. Almost 59% (N = 56) of the sample had a higher vocational education or University degree. Furthermore, the final sample consisted of solely (biological) men with non-Western backgrounds. Also, 8 Dutch men and 3 women were removed from the sample as they did not fit the inclusion criteria.

Figure 2 presents the response rates of seeds and invites in the RDS study. The response rate is computed as the completed and partial survey's divided by the total number of survey requests send out. This calculation is based on Response Rate 2 by the American Association for Public Opinion Research (2016).

RR2 = C + P / C + P + (N+D+R)
C = Completed, P = Partial, N = Non-Response, D = Declined, R = Remaining

The total response rate in RDS is (70 + 24) / (70 + 24 + 94) = 50%. Among the seeds the response rate is 87.64% and among the invites the response rate is 39%. Figure 1 gives an overview of the response type from the total requested survey's in RDS.
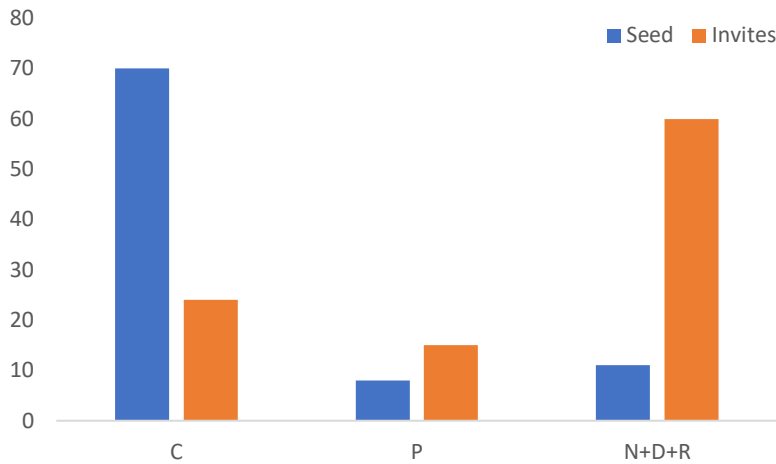
Figure 2. Number of survey requests RDS
Note. C = Completed, P = Partial, N+D+R = Non-Response, Declined and Remaining.

4.2. Experimental condition

4.2.1. Number invited

The experiment was conducted to test whether there was a relation between the requested number of invites and the number of people who were invited to the survey. To investigate this relationship, we used two Poisson regression models (with and without covariates). In both models we measured the 'number of requested referrals' as dependent variable.

-Model 1a: One unit increase in the requested number of invites results in an increase of $\exp(\beta) = 2.30$ ($\beta = .83$, SE = 0.25, $p < .05$) on the number of people invited to the survey. The model accounts for by 14% of the variance (Nagelkerke $R^2 = .1423$)
-Model 1b: One unit increase in the requested number of invites results in an increase of $\exp(\beta) = 1.90$ ($\beta = .64$, SE = 0.31, $p < .05$) on the number of people invited to the survey while taking the variables age, ethnicity and network size into account. The model is accounted for by 81% of the variance (Nagelkerke $R^2 = .8126$)

4.2.2. Number responded

The second part of the experiment was to investigate the relation between the requested number of invites and the number of invites who actually responded to the survey. To investigate this relationship, we also used two Poisson regression models (with and without covariates), but this time measuring 'number of invites responded' as the dependent variable.

-Model 2a: There is no significant effect off the experimental manipulation on the number of responded people in the survey, $\exp(\beta) = 1.33$ ($\beta = .29$, SE = 0.39, $p = .46$).
-Model 2b: There is no significant effect off the experimental manipulation on the number of responded people in the survey, $\exp(\beta) = 1.04$ ($\beta = .04$ SE = 0.47, $p = .93$, while taking into account the variables age, ethnicity and network size.

4.3. Comparison between RDS, National study and Statistics Netherlands on background variables

Data from the RDS study were compared with figures from both the national online study and from Statistics Netherlands (CBS). The variables that are studied are ethnicity, age and educational level. In the next sections we will describe the comparison of each background variable separately.

4.3.1. Ethnic background

Table 1 gives an overview of the variables measuring ethnic background. Respondents in both studies were asked to report their own birth country and both their parents birth country. The countries presented in table 1 are based on proportions of non-Western men living in the Netherlands.

Both RDS as well as the national online study generally reached high proportions of respondents from 'other non-Western' countries.

Country born
In RDS, the proportion of respondents born in most non-Western countries is relatively higher than in the national study, with the exception of Netherlands-Antilles ($\chi^2 = 83.25$, p $< .001$, $\varphi = .29$). The dissimilarity index indicates that 41.63% of the RDS study has to move in order to achieve a uniform distribution of the population by the country they were born in (D = .4164).

Mother Country
When considering the mother's birth country, the national study has a much higher proportion of 'other non-Western' men (86.62%) compared to the RDS study (62.10%), $\chi^2 = 247.61$, p $< .001$, $\varphi = .51$. The dissimilarity index indicates that 35% of the RDS study has to move to in order to achieve a uniform distribution of the population by the country they were born in (D = .3450).

Father Country
Finally, looking at Father's birth country, we also see a higher proportion of non-Western men in the RDS sample, with the exception of the 'other non-Western' countries ($\chi^2 = 229.86$, p $< .001$, $\varphi = .49$). The dissimilarity index indicates that 32% of the RDS study has to move to the National study in order to achieve a uniform distribution of the population by the country they were born in (D = .3227).

In summary, RDS has reached proportionally more men from Morocco, Turkey and Suriname, whereas the national study is more successful in reaching MSM from 'other non-Western' countries.

Table 1.

Comparison between RDS and national study on birth countries of non-Western men

	BornC		MoC		FaC	
Country	*RDS*	*NS*	*RDS*	*NS*	*RDS*	*NS*
	N (%)	N (%)	N (%)	N (%)	N (%)	(N, %)
Morocco	6	4	9	0	10	1
	(6.31)	(0.51)	(9.47)	(0)	(10.52)	(0.13)

Turkey	3	5	5	0	5	0
	(3.16)	(0.64)	(5.26)	(0)	(5.26)	(0)
Suriname	6	26	11	0	11	0
	(6.31)	(3.35)	(11.58)	(0)	(11.58)	(0)
Antilles	4	51	5	8	4	8
	(4.21)	(6.56)	(5.26)	(1.03)	(4.21)	(1.02)
Other-NW	55	214	59	673	61	654
	(57.89)	(27.54)	(62.10)	(86.62)	(64.21)	(84.17)
NL	19	336	4	2	2	2
	(20.00)	(43.24)	(4.21)	(0.26)	(2.10)	(0.25)
Other-W	2	141	2	94	2	112
	(2.11)	(18.15)	(2.11)	(12.10)	(2.11)	(14.41)

Note. The countries are based on categories of non-Western men living in the Netherlands (retrieved from Statistics Netherlands, CBS). The countries are respectively Morocco, Turkey, Suriname, Dutch-Antilles and Other non-Western. The Netherlands and other-Western countries are included as there are also second generation and multi-ethnic migrants included in these categories. The comparison is made between RDS and National Study (NS) on the variables Country Born (BornC), Mother Country (MoC) and Father Country Born (FaC)

Table 2 gives an overview of the ethnic background of respondents in the RDS sample and proportions of non-Western men retrieved from Statistics Netherlands. We can observe that the distribution of non-Western men in RDS is more similar (D = .2937) to the distribution of Statistics Netherlands when compared to distribution of the national study to Statistics Netherlands (D = .4993). Furthermore, we there is an underrepresentation of Antillean people in the RDS sample and an overrepresentation of the category 'other non-Western' men in the national study, both compared to proportions of Statistics Netherlands.

Table 2

Comparison of ethnicity from non-Western men between RDS, national study and CBS

Ethnicity	Sur	Moroc	Turk	Ant	O-NWest
	N (%)	N (%)	N (%)	N (%)	N (%)
RDS	11	10	5	5	64
	(11.58)	(10.53)	(5.26)	(5.26)	(67.37)
NS	26	4	5	51	691
	(3.34)	(.51)	(.64)	(6.56)	(88.93)
CBS	137 000	135 000	158 000	58 000	312 000
	(17.13)	(16.88)	(19.75)	(7.25)	(39.00)

Note. The criteria of defining ethnicity is similar to the definition of Statistics Netherlands. When a person is born outside the Netherlands he is considered 1st generation migrant and when at least one of the parents is born in another country he is considered 2nd generation migrant. The groups are respectively Surinamese (Sur), Moroccan (Moroc), Turkish (Turk), Antillean (Ant) and other non-Western (O-NWest)

4.3.2 Educational level

Table 3 gives an overview of the educational level of non-Western men in the RDS and the National study. To compare these categories, we used chi-square tests. The table shows that RDS has reached a higher proportion of non-Western men from primary school and MAVO, whereas the national study is more successful in reaching non-Western men with an MBO degree ($\chi^2 = 18.89$, p < .05, $\varphi = .14$). The dissimilarity index indicates that 12.90% of the

RDS study has to move to in order to achieve a uniform distribution of the population by educational level (D = .1290)

Table 3
Comparison of educational level between RDS, National study and Statistics Netherlands Note.

	RDS (N %)	NS (N %)	CBS (N %)
Primary school	4 (4.21)	7 (.91)	
Primary Education	4 (4.21)	7 (.91)	291000 (36.39)
MAVO	5 (5.26)	21 (2.71)	
VMBO	5 (5.26)	19 (2.45)	
HAVO/VWO/ Gymnasium	14 (14.74)	90 (11.58)	
MBO	10 (10.53)	178 (22.91)	
Secondary Education	34 (35.79)	308 (39.75)	309 000 (38.63)
HBO/WO/PhD	56 (58.95)	450 (57.92)	
Higher Education	56 (58.95)	450 (57.92)	166 000 (20.75)
Other	1 (1.05)	8 (1.03%)	34 000 (4.25)

The categories are measured similarly in both studies. For the comparison with Statistics Netherlands they were recoded into: primary education (primary school, MAVO/MULO and VMBO), secondary education (HAVO/VWO/Gymnasium and MBO) and higher education (HBO/WO/PhD)

The third column of table 3 presents the educational level of non-Western men retrieved from Statistics Netherlands. The results show that the national study has been more successful in reaching non-Western men with secondary education, corresponding more to the proportions of Statistics Netherlands. In both RDS as well as the National study there is an overrepresentation of higher educated non-Western MSM when compared to the proportions of Statistics Netherlands. In order to achieve a uniform distribution of population by educational level, 51.71% of men in RDS would need to move (D = .5171). Concluding that the RDS distribution is much more similar to the National study, then to the distribution of Statistics Netherlands.

4.3.3 Age

To compare the age distribution of non-Western men in both studies, we categorized the variable according to standard categories of Statistics Netherlands (Table 4) and used a chi-square test to compare the categories. RDS has been successful in reaching a higher proportion of MSM from the category 25-35 years old, whereas the national study has mostly reached non-Western MSM between 15-25 years old ($\chi^2 = 66.60$, p < .001, $\varphi = .26$). In order to achieve a uniform distribution of the population by age, 38% of men in RDS would need to move (D = .3842)

Table 4

Comparison of age categories between RDS, National study and Statistics Netherlands

Age categories (year)	RDS N (%)	National study N (%)	CBS N (%)
15-25	23 (24.21)	395 (50.84)	189 000 (23.63)
25-35	59 (62.11)	200 (25.74)	180 000 (22.50)
35-45	11 (11.58)	114 (14.67)	167 000 (20.88)
45-55	2 (2.10)	90 (11.58)	144 000 (18.00)
55-65	0 (0)	44 (5.66)	87 000 (10.88)
65-75	0 (0)	13 (1.67)	24 000 (3.00)
>75	0 (0)	2 (0.25)	10 000 (1.25)

Note. The age categories are divided similar to the categories of Statistics Netherlands

Table 4 also presents the population registers of age categories from non-Western men, retrieved from Statistics Netherlands. The proportions of men in the age category 15-25 years in RDS (24.21%) is almost equal to proportions of CBS (23.63%), whereas there is an overrepresentation of the age category 25-35 years old. Also, RDS has not been able to successfully reach men from an higher age groups older than 45 years old (2.10%). In order to achieve a uniform distribution of population by age, 40% of men in RDS would need to move (D = .3972). In sum, the distribution of age categories from RDS is similar to both the national study as well as the distribution from CBS.

4.4 Comparison between RDS and National study on key variables

To compare the differences between the theme's 'sexual behavior' and 'HIV prevention' we modelled multiple linear and logistic regression models. Table 5 shows that there is no significant difference between the RDS and National study in terms of the sexual risk behaviors. However, people in RDS are less aware of HIV prevention methods than people in the national study.

Table 5

Regression models comparing the difference between RDS and CS on key variables

	No covariates				Covariates			
Dependent variables	ß	SE	P-value	R^2	ß	SE	P-value	R^2
Sexual behavior								
First time sex	-3.18	7.92	.69	.00	-.88	8.15	.91	.01
Anal sex	7.42	35.19	.83	.00	25.94	35.23	0.46	.06

Steady anal sex	-.22	.37	.55	.00	-.14	.39	.72	.07
Steady anal sex 6m	.18	.49	.71	.00	.28	.51	.58	.05
HIV Prevention								
Informed	-.46	.13	.00055**	.02	-.42	.13	.00122**	.10
Condom use	-.79	.13	3.02e-09**	.05	-.84	.14	8.61e-10**	.07
Ejaculation	-.92	.18	4.09e-07**	.04	-.99	.19	1.35e-07**	.07
PrEP use	-1.12	.14	1.31e-15**	.10	-1.12	.14	5.24e-15**	.11
Regular testing	-.97	.16	3.56e-09**	.05	-1.03	.17	7.65e-10**	.09

Note. The estimates in the table are from the predictor variable that has been coded as 0 = National study and 1 = RDS
** significant at alpha = .05

The final theme studied was HIV testing behaviour. An indicator for this theme is the last time people had an HIV test (measured in 4 categories). To compare the difference between HIV testing in both studies a multinomial logistic regression was modelled. The results indicate that people in RDS (compared to the national study) generally get tested more often, relative to people never having tested before. Specific to each category:

-Respondents in the RDS compared to the National study are 19.84 times more likely to have been tested 0-3 months ago, relative to those never having an HIV test before (RR = 19.84, SE = 1.03, p < .05)

-Respondents in the RDS compared to the National study are 10.30 times more likely to have been tested 3-6 months ago, relative to those never having an HIV test before (RR = 10.30, SE = 1.06, p < .05)

-Respondents in the RDS compared to the National study are 13.44 times more likely to have been tested 6-12 months ago, relative to those never having an HIV test before (RR = 13.44, SE = 1.09, p < .05)

-Respondents in the RDS compared to the National study are 16.03 times more likely to have been tested 6-12 months ago, relative to those never having an HIV test before (RR = 13.44, SE = 1.09, p < .05)

In summary, men in the RDS simple do not differ in their sexual behaviours, but all less informed on HIV prevention methods. On the other hand, in RDS men are more likely to get tested often, relative to those never having been tested before.

5. Discussion

5.1. Implementation of RDS online experiment

The first part of this study mainly focussed on improving the recruitment strategy of RDS. An online experiment was implemented to investigate the relation between the (requested) number of referrals and the number of invites. The outcome of the experiment showed that respondents who were requested to invite four peers also managed to recruit more people on average, as opposed to those who were requested to invite two other peers. Thus, increasing the number of requested referrals in RDS results in higher number of invitations. Note that this number cannot be infinite, otherwise there will be an overrepresentation of people with larger network sizes.

On the other hand, we found no direct relation between the number of requested referrals and the number of completed questionnaires. In the condition of four referrals, the number of completed questionnaires was not higher than those in the condition of two referrals. Regardless of the experimental condition, the non-response rate was extremely high under invites. This might indicate that the non-response under invites is influenced by factors other than the requested number of referrals.

One explanation for the high non-response under invites is the (lack of) instructions that were given to initial seeds. These seeds were solely instructed about the study goals and purposes. However, there was no additional instruction on the recruitment methods. One example is that invites only received a link to the questionnaire without any further information. Not knowing what is behind the link or whether it is save or not to open it, made most of them to not click on the link at all. This could have been prevented if seeds were instructed on how to invite their peers and what they had to communicate.

Secondly, technical challenges within the web-based RDS system also created barriers for respondents to participate. An example is that peers only received the invitation to the study in the language that the seed filled out the questionnaire. This made cross-language referral almost impossible.

Despite the partial success of the experiment, it also has its limitations that we want to address. The first limitation is that respondents were designated to one condition (2 referrals versus 4 referrals) during different time points. There was one month in between the condition of two referrals (round 1 and 2) against four referrals (round 3 and 4). As the two conditions were not at the same timepoint, there could have been history effects: intervening events that have the potential to influence the dependent variable (Christ, 2007). These effects are a threat to the internal validity. A second limitation is based on the changes that were implemented during fieldwork. These changes were applied to improve the recruitment strategy of RDS. As most changes were implemented in the content of the questionnaire, we could not control for their (direct) effect on the dependent variables of the experiment. The higher online referrals could also have been the cause of the changes that were implemented during round 3.

Implementation of the experiment was successful on one hand: online referral has been improved. However, an increase of requested referrals did unfortunately not lead to more completed questionnaires. This tells us that the response (of invites) in RDS is quite complex and may depend on other factors (e.g., instruction of the seeds and technical limitations) that need to be taken into account for further research.

5.2 Similarities and differences between RDS, national study and Statistics Netherlands

The second part of the study mainly focussed on sampling a representative group of MSM by considering the background characteristics ethnicity, age and educational level. Proportions of non-Western MSM in the RDS study were compared to those in the national study and to population registers (of all non-Western men) from Statistics Netherlands.

In terms of ethnic background: the RDS study has been more successful in reaching non-Western men from Suriname, Morocco and Turkey, whereas the national study has reached higher proportions of Antillean and 'other non-Western' men. The distribution of RDS is also more similar to the distribution of Statistics Netherlands, when compared to the distribution of the national study. Thus, RDS is capable of sampling specific ethnic minority groups, who are otherwise not sampled through the national study using convenience methods.

When comparing the proportions of both studies to the population registers of Statistics Netherlands, we can observe an overrepresentation of the category 'other non-Western' men. An explanation for this is the recent arrival of refugees and asylum seekers in the Netherlands. In 2017 Vluchtelingenwerk Nederland reported that 14.716 people asked for asylum in the Netherlands and 18.161 people in 2016. As Statistics Netherlands only registers status holder, a large proportion of these people are not covered in their registers. We can observe that 35% of

the men in the RDS study and 21% of men in the national study are from Middle-Eastern and African countries; areas where most refugees arrive from.

Regarding the educational level in both studies: the proportions of non-Western men in RDS are more similar to those in the national study, in comparison to the population registers of Statistics Netherlands. Namely, there is an overrepresentation of higher educated non-Western men in both studies. This in contrast to earlier evidence suggesting that non-Western men in the Netherlands are on average lower educated (Statistics Netherlands, 2017). The opposing findings might indicate sampling bias for the higher educated in both RDS as well as the national study using convenience methods.

Previous literature reports that there is a systematic difference between those who respond to surveys and those who do not respond in terms of their academic attainment; namely higher educated participate more often (Richardson, 2005). Another explanation for the overrepresentation of higher educated is the familiarity with internet technology. Internet users are generally younger, richer, better educated, and more urban than non-Web users (Reips, 2002).

Finally, Yetter and Capacciolo (2010) report that people with primary and secondary (in comparison to higher educated) are less likely to respond to Internet surveys than to questionnaires administered via traditional postal methods. According to the social exchange analysis: individuals least likely to participate in survey research are those who see themselves as least obligated to provide resources in return for resources received (Groves, Cialdini & Couper, 1992). As the sponsors of the current RDS survey are representatives of major societal institution, it may also explain the low rates of survey participation among the lower educated.

Taking into consideration the age of respondents: RDS an equal proportion of non-Western men 15-25 years old, compared to population registers of non-Western men. However, there is an overrepresentation of non-Western men between 25-35 years old. Furthermore, men older than 45 years are almost entirely absent in the RDS study. From the beginning of the study, we mainly focused on sampling (non-Western) MSM from a younger age group. By doing so we did not actively reach out to older non-Westen men in the initial recruitment phase. Almost 93% of the seeds were under 40 years old. Stein et al. (2015) report that respondents are expected to recruit other members with similar background characteristics. As a consequence, most younger seeds in the initial recruitment phase did not refer to older men.

On a general note, previous literature reports that MSM survey's do not have a normal distribution of age due to factors such as coming-out, sexual presentation and internet access (Johnson, Mulongeni, Marr & Lane, 2018). Older (non-Western) MSM are mostly affected by these factors, resulting in even higher barriers for them to participate in survey research. Finally, Kappelhof (2015) suggests that illiteracy is another (direct) cause for the underrepresentation of older non-Western people in survey research.

The final part of the study focussed on differences in key variables between non-Western men in RDS and non-Western men in the national study using convenience methods, while controlled for the background variables ethnicity, age and educational level. The results have shown that men in RDS do not differ on sexual risk behaviours compared to men in the national study. However, we did find a difference in the knowledge of HIV prevention methods. Namely, men in RDS were less informed and also less aware of effective prevention strategies, compared to those in the national study. At the same time, men in RDS got tested more often than men in the national study. These results imply that different types of men can be reached depending on the kind of sampling method. Intervention methods can be applied targeted at these MSM subgroups in RDS who differ from those in the national study. This shows that the national study by its own is not sufficient enough in sampling all types of men. Next to that, RDS is a necessary sampling method to implement among MSM with non-Western backgrounds. Combining the two studies would be the most optimal strategy to sample a representative group of MSM.

In summary, RDS is a useful sampling method to reach additional specific ethnic minority groups. These groups differ on HIV awareness and testing behavior compared to those sampled

in the national study. However, similar to previous research it still remains a challenge to reach some subgroups, such as older and lower educated non-Western MSM. These groups require special attention in survey research, potentially through other modes of data collection. Combining RDS with the national study in a mixed mode design could be a potential solution to balance the out the homogeneity in each sample.

5.3. Limitations of current research

Our study has several limitations that we want to address. The first limitation lies on the content of the questionnaire. Tourangeau, Rips & Rasinski (2000) report that (item) non-response is more likely to occur for questions involving some kind of psychological threat, such as questions about sexual behavior. As all of the theme questions in the current study were compulsory and respondents could not proceed without answering, high drop-out was likely to have been the result. Another reason for high drop-out is when the questions are too difficult for respondents to answer (Dillman, Eltinge, Groves and Little, 2002). Findings from the qualitative interviews (additional material appendix E) confirm that drop-outs had difficulties answering certain types of questions containing complicated health-related terms, such as 'sero-sorting', 'strategic positioning' and 'viral load sorting'. The inability to answer such questions was a reason for respondents to stop filling out the questionnaire.

The second limitation of the current RDS study was the accessibility to the target group. Seeds were informed and instructed about the study purposes, but not actively included to be part of the study. As a result, there were some barriers to understanding the content of the questionnaire (e.g., questions were perceived as too personal) and also towards communicating with the seeds (e.g., not clearly instructed to invite others). The current frame is targeted at well educated, younger white men. The input of the target group could increase the quality of the study and also result in higher response rates. One way could be to organize focus groups specifically aimed at non-Western men. The goal of these focus groups is to receive input from the target group by actively engaging respondents to participate (further discussed in section 5.4)

A third limitation is the use of population registers from Statistics Netherlands. As the Netherlands has no population registers of specifically MSM or gay men, we assumed that the distribution of non-Western MSM is equal to the distribution of *all* non-Western men living in the Netherlands. This assumption is based on estimates of 7-10% of the general population worldwide being non-heterosexual (Lim & Bernstein, 2012; Felten & Boss, 2017). In addition, regardless of these estimates, the population registers are not completely reliable. The first reason is because of the large influx of many migrants in the last two years (Vluchtelingenwerk Nederland, 2017). Migrant have the opportunity to grant asylum based on sexual orientation (Immigratie-en Naturalisatiedienst, 2016). As a consequence, higher proportions of MSM migrants choose The Netherlands as potential country to ask for asylum. The second reason of unreliable population estimates results from stigma towards sexual minorities (Reitman, 2015). For years, many of these people hide their feelings or do not admit having sex with people from the same gender. The acceptance and tolerance is country specific and therefore the assumption might underestimate the number of MSM specifically living in the Netherlands.

5.4. Improvements for further research

Implementation of RDS can be useful to a wide range of researchers in order to sample hard-to-reach (minority) groups. But first, further methodological research needs to be conducted on ways how to improve the recruitment strategy of RDS. It might be beneficial to set up a development research in which theoretically the development of the RDS sampling method is tested (empirical evidence for its effectives) and practically methodological directions are generated for the design and evaluation of the sampling method (Van den Akker, 1999). In this section we discuss potential (useful) strategies that can be implemented to set up a development research for the improvement of the RDS sampling method.

1. In the current RDS study we observed a high response under initial seeds and low response under the invites. These results are partially explained by the direct connection between the seeds and the researcher. Initial seeds have a crucial role in the RDS recruitment and therefore need special training to motivate their invites to participate in the study and also for them to invite their peers, and so on. Reisner et al. (2010) found that key variables regarding sexual health (i.e. risky sexual behavior and stimulant drug use during sex) have an influence in generating productive seeds. Meyer and Wilson (2009) suggest that individuals of lower socioeconomic position (SEP) and/or those who regularly use drugs may be more likely to successfully recruit others into a study, perhaps because they are more motivated by the levels of monetary incentives offered. These findings suggest that seed characteristics are of high importance to facilitate collecting a sample that is closer to reflecting the general population. In order to reach this, it is important for researchers to include a heterogeneous group (both on background characteristics as well as key variables) of seeds by setting up a screening procedure to qualify their eligibility. A second way is by incorporating qualitative interviews with seeds that provide future insights as to why some seeds are generative and others are not. Finally, other factors that also have to be taken into account are social networks, the strength of social ties or the frequency of participation in previous research projects (Reisner et al., 2010). If seeds are strictly selected bases on these procedures, it can increase their productivity to successfully recruit others.

2. Besides the productivity of the seeds, there are many other strategies researchers themselves can implement to motivate invites to fill out the questionnaire. A well-known motivational technique is the use of (monetary) incentives. Earlier research has shown that several types of incentives can increase the response rate (Perneger, Etter & Rougemont, 1993). However, the major cost of monetary incentives is that the kind of research is greatly restricted, and their use will influence both what questions are set out to answer, and the way they are answered (Vahidov & He, 2010). Thus, the use of (monetary) incentives can have its benefits, but only serves as extrinsic motivation. In order to motivate survey respondents intrinsically there are many useful (cognitive) strategies one can take into consideration (derived from the persuasion theory (Groves, Cialdini & Cooper, 1992).

   The first strategy for a persuasive effort is the principle of social validation – more willing to comply with a request to the degree that one believes similar others would comply with it. One way to implement it within RDS is to state the number of friends, or the number of people with similar (ethnic) background who are also willing to fill out the questionnaire. This may result in more respondents also filling out the questionnaire, as similar others have done the same.

   A second strategy is the principle of reciprocation – directing individuals to provide to others the general form of behavior that they have received from those others. In RDS, the researchers benefit from respondents filling out the questionnaire. In turn it would be helpful to exchange the contact details of the researchers for more information or to give respondents the opportunity to receive the final results of the research. It also leads to appreciation from the side of the respondent.

   Another principle is (investigating in) authorities – more willing to comply with a request of someone who one perceives as a legitimate authority. When combined with the first principle and stating that authorities within the community have agreed to fill out the questionnaire or are part of the research project, people are expected to be more willing to participate.

   Finally, Chaiken (1980) has shown that high levels of issue involvement lead recipients to imply a systematic information processing strategy in forming their opinion judgement. The initial opinion judgement of respondents who are highly involved is strongly influenced by the amount of argumentation provided in the message (instead of the communicator's likeability). Thus, respondents who are actively included in the research project can be persuaded by the argumentation through extra instruction letters or briefing. Another way to involve seeds is by means of focus groups. The goal of these focus groups would be to get the respondents closely in touch with the research project by not only

discussing the goals and purposes of the study, but also the content of the questionnaire and the sampling strategy.

3. The final recommendation is based on the content of the questionnaire. The order in which questions are asked (depending on the context) can have large effects on survey results. In the current RDS study we there was a relatively high drop-out rate at the very beginning of the questionnaire. These drop-out could potentially have been avoided if the order of the questionnaire was structured differently. Blair, Czaja & Blair (2013) mention several characteristics of what the first question(s) should accomplish to keep respondents motivated to continue.

    Most importantly, the beginning of the questionnaire should at least contain questions that are relevant to the central topic. In this study respondents were in advance motivated to answer a questionnaire about sexuality and health but received questions about demographic characteristics at the beginning of the survey. The link between demographics and key variables is not always understood by respondents and the beginning of the survey should be consistent with the theme. Therefore, it is not advised to start with demographic variables.

    Secondly, the questionnaire should be applicable to and answerable by most respondents. Lee, Jones, Mineyama & Zhang (2002) have shown differences between cultures in responding to questions containing Likert scales. Some of the eastern cultures had difficulties answering them and others frequently skipped these types of questions. For a cross-cultural survey (or when surveying ethnic minorities in general) it is highly recommended to avoid Likert scales and proceed with simple, close-ended questions.

The recommendations above can be executed through qualitative cognitive interview or by means of focus groups. These interviews and focus groups can be used as a tool to gain relevant information about the productivity of the seeds and also to identify potential problems (such as wording or question order) within the content of the questionnaire. Furthermore, when trying to motivate respondents to answer the questionnaire, one can take into consideration many cognitive principles that are derived from the persuasion theory.

6. Conclusion

Despite the small sample size, we can conclude that RDS is a sampling method able to reach specific ethnic minority groups from a younger age. These men differ on several key variables, such as testing behavior and knowledge on HIV prevention methods, when compared to men from the national study using convenience methods. In the national study there are also more men where (one of) the parents originate from Western countries.

Under the right conditions RDS can reach subgroups of MSM who are different from MSM reached through convenience methods. Further researchers who want to implement RDS as a sampling method among (ethnic) minority groups have to take into consideration the accessibility to the target group and account for technical challenges that might occur during the recruitment process. We have advised to set up a development research program in which empirically based recommendations are given on how to improve the recruitment of RDS and thus increase the response rate (among invites). The most important factors to take into account are the heterogeneity of the seeds, communicating through cognitive principles and set up a questionnaire applicable to the target group.

In summary, with many improvements RDS can implemented among specific hard-to-reach groups. The combination of factors, such as high incentives, multiple languages, accessibility to the target group and the ones recommended for the development program are relevant to stimulate the recruitment strategy within RDS. One relevant finding to stimulate online recruitment is to increase the number of referrals to a limited number, so overrepresentation of people with larger network sizes can be avoided. When proceeding through these steps, RDS has the potential to lead to large sample sizes and reach long recruitment waves. Combining RDS

with convenience sampling can be a potential solution reach a representative sample of minority populations.

[1] Additional material Appendix A: Sampling and Recruitment plan National online survey

[2] Additional material Appendix B: RDS online questionnaire

[3] Additional material Appendix C: Reasons to shorten the RDS questionnaire

[4] Additional material Appendix D: Seed Recruitment schedule and communication letter

[5] Additional material Appendix E: Changes applied during RDS fieldwork

Abbreviations

RDS = Respondent Driven Sampling
MSM = Men who have Sex with Men
HIV = Human Immunodeficiency Virus
STI = Sexually Transmitted Infection
CBS = Statistics Netherlands
RIVM = National Institute for Public Health and the Environment
CSG = Center Sexual Health

References

Baral, S., Sifakis, F., Cleghorn, F., & Beyrer, C. (2007). Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000–2006: a systematic review. *PLoS Medicine* , *4*(12): e339.

Behr, D., & Shishido, K. (2016). The Translation of Measurement Instruments for Cross-Cultural Survey In C. Wolf, D. Joye, T.W. Smith & Y. Fu (Eds), *The SAGE Handbook of Survey Methodology* (pp. 278-297). SAGE Reference: Los Angeles.

Bengtsson, L., Lu, X., Liljeros, F., Thanh, H. H., & Thorson, A. (2014). Strong propensity for HIV transmission among men who have sex with men in Vietnam: behavioural data and sexual network modelling. *BMJ open*, *4*(1): e003526.

Bengtsson, L., Lu, X., Nguyen, Q. C., Camitz, M., Le Hoang, N., Nguyen, T. A., & Thorson, A. (2012). Implementation of web-based respondent-driven sampling among men who have sex with men in Vietnam. *PLoS One*, *7*(11): e49417.

Beyrer, C., Baral, S. D., Van Griensven, F., Goodreau, S. M., Chariyalertsak, S., Wirtz, A. L., & Brookmeyer, R. (2012). Global epidemiology of HIV infection in men who have sex with men. *The Lancet*, *380*(9839): 367-377.

Blair, J., Czaja, R. F., & Blair, E. A. (2013). *Designing surveys: A guide to decisions and procedures*. Sage Publications.

Campanelli, P. (1997). Testing survey questions: New directions in cognitive interviewing. *Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique*, *55*(1), 5-17.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, *39*(5), 752-766.

Christ, T. J. (2007). Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. *Psychology in the Schools*, *44*(5), 451-459.

Dillman, D.A., Eltinge, J.L., Groves, R.M., & Little, R.J.A. (2002). Survey Nonresponse in Design, Data Collection, and Analysis. In R.M. Groves, D.A. Dillman, J.L. Eltinge, R.J.A. Little (Eds.), *Survey Nonresponse* (pp. 3-26). New York: John Wiley & Sons.

Dovidio, J. F., Gluszek, A., John, M. S., Ditlmann, R., & Lagunes, P. (2010). Understanding bias toward Latinos: Discrimination, dimensions of difference, and experience of exclusion. *Journal of Social Issues*, *66*(1), 59-78.

van Empelen, P., van Berkel, M., Roos, S., & Zuilhof, W. (2011). SchorerMonitor. Amsterdam: Schorer.

de Leeuw, E. D., Hox, J.J. & Dillman, D. A. (2008*).* The cornerstones of survey research. In E.D. de Leeuw, J.J. Hox & D.A. Dillman (Eds.), *International Handbook of Survey Methodology* (pp. 1-17). UK: Taylor & Francis Group.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1): 1-4.

Evans, A. R., Hart, G. J., Mole, R., Mercer, C. H., Parutis, V., Gerry, C. J., & Burns, F. M.(2011). Central and East European migrant men who have sex with men in London: a comparison of recruitment methods. *BMC medical research methodology*, *11*(1): 69

Felten, H., & Boss,E.M. (2017). Handreiking LHBT emancipatie. Retrived from https://www.movisie.nl/sites/movisie.nl/files/publication-attachment/Handreiking feiten cijfers-lhbti%20%5BMOV-13351204-1.1%5D.pdf

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1991). Toward a useful theory of survey participation. In *Proceedings of the Section on Survey Methods Research, American Statistical Association* (pp. 88-97).

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). Inference and Error in Surveys In R.M. Groves, G. Kalton, J.N.K. Rao, N. Schwarz, C. Skinner (Eds.), *Survey Methodology* (pp. 39-65). Hoboken, New Jersey: John Wiley & Sons.

Heath, A., & Brinbaum, Y. (2007). Guest editorial: Explaining ethnic inequalities in educational attainment. *SAGE publications, 7*(3): 291-204.

Heckathorn, D. D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social problems*, *44*(2): 174-199.

Heckathorn, D. D. (2002). Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Social problems*, *49*(1), 11-34.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Immigratie- en Naturalisatiedienst. Asielzoeker. Retrieved from https://ind.nl/asiel/Paginas/Asielzoeker.aspx

Johnson, L. F., Mulongeni, P., Marr, A., & Lane, T. (2018). Age bias in survey sampling and implications for estimating HIV prevalence in men who have sex with men: insights from mathematical modelling. *Epidemiology & Infection*, 1-7.

Johnston, L. G., Trummal, A., Lõhmus, L., & Ravalepik, A. (2009). Efficacy of convenience sampling through the internet versus respondent driven sampling among males who have sex with males in Tallinn and Harju County, Estonia: challenges reaching a hidden population. *AIDS care*, *21*(9): 1195-1202. DOI: 10.1080/09540120902729973

Kappelhof, J. (2015). The effect of different survey design on nonresponse in surveys among non-Western minorities in the Netherlands. In *Surveying ethnic minorities: the impact of survey design on data quality* (pp. 35-65). Den-Haag: Textcetera

Kappelhof, J. (2017). Survey Research and the Quality of Survey Data Among Ethnic Minorities. In P.P. Biemer, E.D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, N.C. Tucker & B.T. West (Eds.), *Total Survey Error in Practice* (pp. 235-252). New Jersey: John Wiley & Sons.

Keuzenkamp, S., & Oudejans A. (2011). Gewoon aan de slag: De sociale veiligheid van de werkplek voor homoseksuele mannen en vrouwen. *Sociaal en Cultureel Planbureau*. Den-Haag: SCP-publicatie.

Laguilles, J. S., Williams, E. A., & Saunders, D. B. (2011). Can lottery incentives boost web survey response rates? Findings from four experiments. *Research in Higher Education*, *52*(5): 537-553.

Lim, F. A., & Bernstein, I. (2012). Promoting awareness of LGBT Issues in Aging in a Baccalaureate Nursing Program. *Nursing Education Perspectives*, *33*(3), 170-175.

Meyer, I. H., & Wilson, P. A. (2009). Sampling lesbian, gay, and bisexual populations. *Journal of Counseling Psychology*, *56*(1), 23-31.

Millett, G. A., Flores, S. A., Peterson, J. L., & Bakeman, R. (2007). Explaining disparities in HIV infection among black and white men who have sex with men: a meta-analysis of HIV risk behaviors. *Aids*, *21*(15), 2083-2091.

McCreesh, N., Frost, S., Seeley, J., Katongole, J., Tarsh, M. N., Ndunguse, R., Jichi, F., Lunel.N.L.,Maher, D., Johnston, L.G., Sonnenberg, P., Copas, A.J., Hayes, J.H., & White, R.G. (2012). Evaluation of respondent-driven sampling. *Epidemiology, 23*(1), 138-147. doi: 10.1097/EDE.0b013e31823ac17c

Nikkelen, S., de Graaf, H., van Lisdon, J., & Bakker, B. (2016). Het meten van seksuele oriëntatie en genderidentiteit in bevolkingsstudies: Operationalisaties en schattingen in Nederland. *Rutgers:*1-8. Retrieved from: https://www.rutgers.nl/sites/rutgersnl/files/PDF Onderzoek/Whitepaper_SO_en_genderidentiteit_DEF.pdf.

Perneger, T. V., Etter, J. F., & Rougemont, A. (1993). Randomized trial of use of a monetary incentive and a reminder card to increase the response rate to a mailed health survey. *American Journal of Epidemiology*, *138*(9), 714-722.

Reips, U.D. (2002). Standards for Internet-based experimenting. Experimental Psychology, 49, 243-256. doi:10.1026//1618-3169.49.4.24.

Reisner, S. L., Mimiaga, M. J., Johnson, C. V., Bland, S., Case, P., Safren, S. A., & Mayer, K. H. (2010). What makes a respondent-driven sampling "seed" productive? Example of finding at-risk Massachusetts men who have sex with men. *Journal of Urban Health*, *87*(3), 467-479.

Reitman, D.S. (2015). Sexual orientation. *MedScape*. Retrieved from https://emedicine.medscape.com/article/917792-overview.

Rosario, M., Schrimshaw, E. W., & Hunter, J. (2010). Cigarette smoking as a coping strategy: negative implications for subsequent psychological distress among lesbian, gay, and bisexual youths. *Journal of Pediatric Psychology*, *36*(7), 731-742.

Sell, R. L., Wells, J. A., & Wypij, D. (1995). The prevalence of homosexual behavior and attraction in the United States, the United Kingdom and France: Results of national population-based samples. *Archives of sexual behavior*, *24*(3): 235-248

Statistics Netherlands. (2016). Jaarrapport integratie. Den-Haag: Centraal Bureau voor de Statistiek.

Statistics Netherlands. (2017). Bevolking; kerncijfers [Database]. Retrieved from http://statline.cbs.nl/statweb/publication/?vw=t&dm=slnl&pa=37296ned&d1=0-2,8-13,19-21,25-35,52-56,68&d2=0,10,20,30,40,50,60,64-65&hd=151214-1132&hdr=g1&stb=t

Stein, M., van Steenbergen, J., Chanyasanha, C., Tipayamongkholgul, M., Buskens, V., van der Heijden, P., Sabaiwan, W., Bengtsson, L., Lu, X., Thorson, A. and Kretzschmar, M. (2018). *Online Respondent-Driven Sampling for Studying Contact Patterns Relevant for the Spread of Close-Contact Pathogens: A Pilot Study in Thailand*, *9*(1): 1-11. doi: 10.1371/journal.pone.0085256

Stein, M. L., van der Heijden, P. G., Buskens, V., van Steenbergen, J. E., Bengtsson, L., Koppeschaar, C. E., & Kretzschmar, M. E. (2015). Tracking social contact networks with online respondent-driven detection: who recruits whom? *BMC infectious diseases*, *15*(1): 522.

Strömdahl, S., Lu, X., Bengtsson, L., Liljeros, F., & Thorson, A. (2015). Implementation of Web-based respondent driven sampling among men who have sex with men in Sweden. *PloS one, 10*(10): e0138599.

Tourangeau R., Rips, L. J., and Rasinski, K. (2000). *The psychology of survey response*. Cambrdige, UK: Cambridge University Press.

Vahidov, R., & He, X. (2010). Situated DSS for personal finance management: Design and evaluation. *Information & management*, *47*(2), 78-86.

Van den Akker, J. (1999). Principles and methods of development research. In *Design approaches and tools in education and training* (pp. 1-14). Springer Netherlands.

van Veen, M. (2010). HIV- en soaepidemiologie bij hoogrisicogroepen in Nederland. *Bulletin*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieuhygiëne.

Visser, M., van Aar, F., van Oeffelen, A.A.M., Op de Coul, E.L.M., Hofstraat, S.H.I., Heijne, J.C.M., den Daas, C., Hoenderboom, B.M., van Wees, D.A., Basten, M., Woestenberg, P.J., Götz, H.M., van Sighem, A.I., de Hoon, S., & van Benthem, B.H.B. (2017). Sexually transmitted infections including HIV, in the Netherlands. *Centre for Infectious Disease Control.* Bilthoven: National Institute for Public Health and the Environment. DOI:10.21945/RIVM-2017-0003

VluchtelingenWerk Nederland (2017). Vluchtelingen in getallen. Retrieved from https://www.vluchtelingenwerk.nl/sites/public/u152/Vluchtelingeningetallen2017compleet-1.pdf

 Wejnert, C. (2010). Social network analysis with respondent-driven sampling data: A study of racial integration on campus. *Social Networks*, *32*(2): 112-124. doi: 10.1016/j.socnet.2009.09.002

Williams, R. (2015). Measurement Error 2: Scale Construction. *Department of sociology, pp.* 1-6. Indiana: University of Notre Dame. Retrieved from https://www3.nd.edu/~rwilliam/stats2/l23.pdf

Wirtz, A., Mehta, S., Latkin, C., Zelaya, C., Galai, N., Peryshkina, A., Mogilnyi, V., Dzhigun, P., Kostetskaya, I. & Beyrer, C. (2018). Comparison of Respondent Driven Sampling Estimators to Determine HIV Prevalence and Population Characteristics among Men Who Have Sex with Men in Moscow, Russia. *PLoS one 11*(6): 1-15. doi: 10.1371/journal.pone.0155519

Yetter, G., & Capaccioli, K. (2010). Differences in responses to Web and paper surveys among school professionals. *Behavior Research Methods*, *42*(1), 266-272.

.