From
Karolinska Institutet, Stockholm, Sweden

Assessing Sexual Risk Behavior among Men who Have Sex with Men in Vietnam

Development and Use of a New Internet-Based Sampling
Method

Stockholm 2012

ABSTRACT

Background: Men who have sex with men (MSM) are at high risk of HIV infection in all regions globally. Surveys among MSM in Hanoi and Ho Chi Minh City in Vietnam have shown prevalence rates of 15% and 17% respectively. Representative data about MSM and other hidden populations are extremely difficult to collect, due to stigma. Much available data may thus not be representative of the MSM population as a whole, preventing an evidence-based response to the HIV epidemic. Respondent-driven sampling (RDS) was developed to enable representative sampling of hidden groups but relies on untested assumptions. In addition, an Internet-based version of RDS (webRDS) can potentially improve several key limitations of the RDS method.

Aims: To develop a webRDS system for representative sampling of internet-using MSM in Vietnam and to increase knowledge about sexual risk behavior for HIV among MSM in Vietnam.

Methods: RDS studies were simulated on anonymized data from a large web-community for lesbian, gay, bisexual and transgender persons. The assumptions under which RDS is known to be unbiased were systematically violated and the effects were studied. In-depth interviews (IDIs) were performed with 17 purposively sampled MSM in Hanoi, Vietnam and latent content analysis was used. A webRDS system, designed for MSM in Vietnam, was developed and used in two cross-sectional surveys in Vietnam.

Results: Simulations indicate that bias using RDS is large if respondents choose to invite people based on characteristics that are correlated with study outcomes. Bias and variance increased if participants preferentially invited frequent contacts. Diverse seed selection was highly beneficial. IDIs showed that male-male relationships were perceived to be short and unstable. Faithfulness was highly valued but largely seen as unobtainable. The webRDS system enabled sustained recruitment of 676 and 983 submissions respectively in the two surveys. Maximum wave length was 29 waves and equilibrium was reached for all but one variable. The sample was younger and of higher education than the Vietnamese average. Thirty-six percent (CI: 32-42%) of MSM in the second webRDS survey had unprotected anal sex (UAS) at least once during the last three months. Thirty-six percent (CI: 32-41%) had ever taken an HIV test and received the result. Ninety-two percent of all UAS acts were reported to have taken place with non-casual partners. The numbers of UAS partners and UAS acts were positively correlated. Modeling showed that this sexual behavior pattern is likely to lead to explosive HIV transmission.

Conclusions: Simulations showed that RDS is relatively robust but empirical studies of non-random recruitment and a better understanding of real-life variance is needed. WebRDS is a highly promising method. Further evaluation and development is warranted. HIV prevention services niched for MSM needs to be aggressively scaled up in Vietnam.

Keywords: Epidemiology, HIV, Internet, Men who have Sex with Men, Sampling Methods, Sexual Networks, Sexual Risk Behavior, Respondent-Driven Sampling, Vietnam

LIST OF PUBLICATIONS

I.   Lu X, Bengtsson L, Britton T, Camitz M, Kim BJ, Thorson A, *et al.* The sensitivity of respondent-driven sampling. *Journal of the Royal Statistical Society Series A - Statistics in Society* 2012,**175**:191-216.

II.   Bengtsson L, Thorson A, Thanh VPN, Allebeck P, Popenoe R. Sexual Relationships among Men who have Sex with Men in Urban Vietnam: a Qualitative Interview Study. *Revisions ongoing in BMC Public Health.*

III.   Bengtsson L, Lu X, Nguyen QC, Camitz M, Hoang NL, et al. (2012) Implementation of Web-Based Respondent-Driven Sampling among Men Who Have Sex with Men in Vietnam. *PLoS One* 7: e49417.

IV.   Bengtsson L, Lu X, Liljeros F, Thanh H H, Thorson A. High Sexual Risk Behavior Indicates Explosive HIV Transmission among Men who have Sex with Men in Vietnam. *Manuscript.*

These articles will be referred to in the text by their Roman numerals (I-IV).

LIST OF ABBREVIATIONS

AHI	Acute HIV Infection
AIDS	Acquired Immunodeficiency Syndrome
FTE	Full Time Equivalent
GUD	Genital Ulcer Disease
HCMC	Ho Chi Minh City
HIV	Human Immunodeficiency Virus
HIV-	HIV Negative
HIV+	HIV Positive
IBBS	Integrated Bio-Behavioral Survey
IDI	In-Depth Interview
LGBT	Lesbian, Gay, Bisexual and Transgender
LMIC	Low- and Middle-Income Countries
MAE	Mean Absolute Error
MSM	Men who have sex with men
PrEP	Pre-Exposure Profylaxis
RDS	Respondent-Driven Sampling
SD	Standard Deviation
STI	Sexually Transmitted Infection
UAS	Unprotected Anal Sex
WebRDS	Web-Based Respondent-Driven Sampling

TABLE OF CONTENTS

INTRODUCTION

Ever since the HIV epidemic emerged in the news in 1981, gay men have been at the center of the epidemic in high-income countries. When I came to Vietnam in 2006 the public health efforts in the HIV sector focused however squarely on injecting drug users, female sex workers and young men and women in the general population. At that time, little was known about HIV among men who have sex with men in Vietnam. A few studies indicated however an alarming situation. In deciding to focus on HIV epidemiology among men who have sex with men in Vietnam I imagined myself as the secret agent in the service of the suppressed, darting into and out of dark parks, basements and other secret meeting places for closeted gay men in Hanoi. I did not know then that those dark meeting places had moved into mobile phones and brightly colored online social networking sites. Instead of the dark parks, I took my studies online.

To understand anything about these men however, it was important for me and my colleagues to figure out how to reach them online, how well (or badly) we would expect the chosen method to work, and finally we had to implement the chosen sampling method. Then we could start asking questions.

The most promising way I could see to reach these men was to develop an online version of a sampling method called Respondent-Driven Sampling (RDS). Using this method, the researcher starts the study by contacting a few men who have sex with men who answer a questionnaire. The researcher then asks these men to recruit other men who have sex with men, who in turn recruit others, forming long recruitment chains penetrating deep into the social network. As a statistic method, it was however new and relied on a number of untested assumptions. Together with my fantastic colleague Xin Lu we first investigated the RDS method by simulating RDS studies on a real and anonymized online social network of homosexual men in the Nordic countries (Qruiser on qx.se). These men, represented by zeroes and ones in a computer at the Sociology Department at Stockholm University, were however a far cry from the real men in Vietnam. Together with the sociologist Vu Pham Thanh in Hanoi we therefor performed in-depth interviews with men who have sex with men in Hanoi to understand their social lives and relationships.

In parallel, we developed an Internet-based version of RDS (webRDS) for sampling of MSM in Vietnam. The first version was highly unsuccessful with 25 participants. Using the second version we managed to recruit 84 men. This was promising but not enough. We redesigned the site and increased the incentives for the participants. Then the recruitments started to jump from person to person across Vietnam. That was a big relief and after the first successful webRDS survey focusing on the method itself, we could in the fourth launch finally start asking questions about HIV risk behavior.

This thesis largely follows the chronological order sketched out in this introduction. In the background section I have included an overview of important aspects of the HIV situation and HIV research among MSM, especially focusing on the challenges of representative sampling and the current knowledge of the determinants of HIV transmission on a population level. The method section gives an introduction to the Vietnamese study context and the methods used in the papers I-IV. The result section summarize selected findings in the four papers. In the discussion section I have included a broader perspective on the study findings, with a focus on the webRDS method as implemented in Vietnam. The thesis finishes with a a summary on future research and implications for HIV

prevention policy. The four papers contain a high level of detail. For the reader who is already familiar with the papers, the background section and the discussion will be most interesting.

A core theme in these studies is that of networks. Many homosexual men in Vietnam are incredibly lonely. By connecting to other men, socially and sexually, they create the social network links which can improve their lives. At the same time these links give rise to the sexual networks through which HIV is transmitted. In addition, the social connections between these men also allow us researchers to sample these men and start to understand at least a part of what is going on, in order to develop appropriate public health policies. Ultimately, those same social links between the men will form the pathways through which changing norms, information and behaviors will spread, which can ultimately stop the epidemic.

Background

The HIV Epidemics

The Global Picture

HIV-1 group M, the most common variant of the HIV virus, was in all likelihood transmitted at a single occasion from the Chimpanzee, *Pan troglodytes troglodytes,* to humans somewhere in Central Africa, around the Congo River [1], sometime during the first three decades of the 20th century. This event lead to a global epidemic that so far has infected more than 60 million people [2-6] among which an estimated 34 million were alive at the end of 2010 [7].

As a proportion of the global population only 0.5% are infected [8, 9]. In certain populations however, the epidemic has reached devastating prevalence rates. Eighteen percent of the population of ages 15 to 49 years in South Africa is infected [9] and more than half of injecting drug users in many parts of Indonesia carry the virus [10].

MSM and HIV

MSM is a term used to define men based on behaviors rather than sexual identity or social roles. MSM includes men who self-identify as gay, homosexual, bisexual or heterosexual. Heterosexual MSM include, for example heterosexual men selling sex to other men, or men who perceive themselves as heterosexual but have sex with other men in certain setting, such as an all-male prison environment [11]. The social context of male-male sex in Vietnam is described further in study setting, under methods.

In most high-income countries MSM is the group with the highest, or one of the highest prevalence rates of HIV. Approximately half of all new HIV infections in the US and Sweden takes place among MSM [12, 13].

MSM face an equally devastating HIV epidemic in low- and middle income counties (LMICs) [14-16], where access to prevention, treatment and health care is often absent. Odds ratios for HIV infection among MSM, as compared to the general population, are estimated at eight and 23 in low and middle-income countries, respectively [14]. As results from surveys have started to come in from a number of countries in sub-Saharan and North Africa, as well as from the Middle-East, it has become clear that MSM also in these regions face very high risks of HIV infection [17-23]. The situation in Asia is similarly daunting [24]. Forty-two percent of new HIV infections in Asia is forecasted to occur among MSM in 2020 [25]. Although it is clear that HIV infection is an extremely serious problem for MSM globally, specific estimates and predictions, such as the one for Asia is based on data and assumptions with a very high level of uncertainty, due the enormous challenges in collecting representative data and to model HIV transmission among MSM populations [26, 27].

Sampling of MSM

The Challenge of Surveying Hidden Groups

Lack of valid and representative data on HIV prevalence and risk behavior among MSM globally is primarily caused by two related phenomena.

First, men who practice male-male sex are heavily stigmatized in most countries globally. Many MSM therefor hide their sexual preferences and behaviors, making it almost impossible for researchers to sample these men representatively. Stigmatization is most apparent in low- and middle-income countries [28] and human rights abuses against MSM has been widely documented globally [29-32]. In about a third of countries globally male-male sex is criminalized and in many countries, e.g. Iran, Nigeria, Uganda, Sudan and many others, MSM risk sentences of more than ten years in prison and in some countries death [33, 34]. In these settings it is thus no surprise that representative sampling of MSM will form a great challenge to researchers.

The second difficulty in representative sampling of MSM is that there is obviously no list (sampling frame) of MSM from which a random sample can be drawn. Sampling instead from a list of people in the general population would be highly inefficient, even in the absence of stigmatization, since the proportion of men practicing male-male sex is in most settings small [35]. For example, if the aim is to reach a nationally representative sample of say 1000 MSM in Sweden and two in a hundred in the Swedish population are sexually active MSM, one needs to contact more than 50,000 randomly selected persons to reach the required sample size.

Traditional Methods and their Shortcomings

Instead of random population samples, traditional methods to sample hard-to-reach populations, such as MSM, have involved contacting especially knowledgeable persons, known as key informant sampling [36], targeted sampling where participants are recruited from locations which MSM are known to visit [37], or snowball sampling, where members of a group are asked to give researchers contact details of others in the same group who can subsequently be contacted and interviewed [38].

However, these methods all introduce a considerable selection bias, which impairs generalization of the findings from the sample to the study population [27, 39]. Key informants have access to a limited part of the network and will typically only be able to reliably report sensitive information about their close contacts [39]. Targeted sampling ignores individuals who do not frequent known meeting places [37, 39]. Snowball sampling do not take into account that some individuals have a large number of social connections and thus higher likelihood of being invited to the study compared to those with smaller social networks. In addition, asking persons in a stigmatized populations to give out contact details to their friends without the consent of those persons, may be ethically unacceptable. Finally, snowball sampling typically does not proceed in more than a few steps from the starting persons. The sample may thus have significant similarities to the persons starting the survey, who in turn may not be representative of the MSM population as a whole [39].

Respondent-Driven Sampling

Overview

Respondent-driven sampling (RDS) is a method which was developed to overcome these challenges in sampling hidden populations [40-43]. An RDS study starts out by purposively selecting a handful of participants who are members of the study population (usually five to fifteen persons). These persons are called 'seeds'. These seeds are given a fixed number of invitation coupons (usually three) to distribute to friends and acquaintances within the study population. If these friends decide to participate, they are in turn given the same number of coupons to invite further friends in the group. Participants are rewarded for their personal participation in the study, as well as for each peer they invite and who also participates. The invitation coupon contains a serial number that enables the researchers to follow the recruitment chains in the sample. Additionally, each participant is asked for the number of people he or she knows within the study population, known as his or her 'personal network size' or 'degree'. The degree of a participant is important to collect as participants with large degrees are oversampled and participants with small degrees are under sampled. Knowing the degree of each participant hence allows adjustment for this bias.

In essence, the method relies on recruitments moving randomly through the links of the entire social network of the group, collecting a sample of participants as the recruitment process continues.

When the sample has been collected, the proportion of people with characteristic A in the population can be estimated by the RDS estimator (RDSII) [43], according to:

$$\hat{P}_A = \left(\frac{n_A}{n}\right)\left(\frac{\hat{\delta}_U}{\hat{\delta}_A}\right)$$

where n_A is the number of individuals with characteristic A, n is the total sample size, $\hat{\delta}_U$ is the mean degree of the individuals in the sample (the mean number of social contacts of the individuals in the sample) and $\hat{\delta}_A$ is the mean degree of individuals with characteristic A. The RDS method will thus not be able to estimate the size of the population but it will be able to estimate the proportion of the population with a certain characteristic, e.g. the proportion of MSM who are HIV positive or the proportion who are younger than 30 years etc.

Volz and Heckathorn [43] proved that the RDSII estimator provides asymptotically unbiased population estimates if the following assumptions are fulfilled:

1. Reciprocity: individuals in the population maintain and recruit peers through reciprocal relationships, i.e. the network within which recruitment happens is undirected.
2. Connectedness: each individual in the population studied has a chance of being invited to participate, i.e. the network forms a single component.
3. Sampling is with replacement: individuals are allowed to be recruited into the sample more than once.
4. Degree: respondents can accurately report their degree in the network.
5. Random recruitment: peer recruitment is a random selection from the respondents' personal networks.
6. Each respondent recruits a single peer, i.e. the number of recruitment coupons is one.

A consequence of these assumptions is that the sampling will conform to a so called first order Markov process [44], meaning that the sample composition will rapidly converge to a stable level that is independent of the characteristics of the seeds (figure 1).
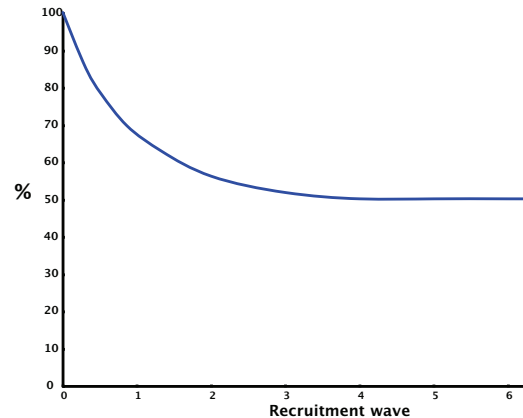

Figure 1: This ideal RDS study started with a handful of participants who were all HIV infected (100% HIV+ at wave 0, the start of the sampling). After three to four waves (steps from the seeds), the composition of the sample stabilized at 50%. If the RDS assumptions are fulfilled the final composition of the sample (50% HIV+) will always be the same, irrespectively of the proportion of the starting participants who were HIV+.

Homophily
The speed with which the composition of the sample stabilizes is highly dependent on how often people have connections with others who have characteristics that are different from themselves. On average people with high incomes are often friends with other people with high incomes and an immigrant from Vietnam, living in Sweden, is more likely than the average citizen in Sweden to be friends with another person from Vietnam. The extent to which people of a certain group tend to be connected with persons who are similar to themselves is called "homophily". Homophily, as defined by Heckathorn [40], can vary from -1 to +1. If persons with characteristic A on average choose friends completely at random with respect to characteristic A, then their homophily with respect to characteristic A is zero. If their homophily is 0.5, it means that their friendship composition with regard to characteristic A is as if these persons first chose 50% of their friends only among people with characteristic A and then chose the remaining 50% of their friends randomly in the population (where people have both characteristic A and not A). As we shall see, homophily is one important determinant of success in RDS studies.

Bias and Variance when RDS Assumptions are Fulfilled
The RDS estimator has been shown analytically be asymptotically unbiased when all the mentioned assumptions are fulfilled [43]. This means that as sampling proceeds, the average RDS II estimate will get closer and closer to the true population value. Variance of RDS estimates has been shown to depend heavily on the structures of the social network in which recruitment takes place. Most importantly, if networks have a high degree of homophily, variance can become very high [42]. In addition, networks can also have weak connections between parts of the network (bottlenecks), independently of homophily which can drive up variance to very high levels [45]. Possibly, high variance in

degree distribution also increase variance, but this assumption has not been rigorously investigated in the literature.

RDS in Real-Life

As expected, research has shown a large number of real-life deviations from the ideal RDS assumptions. Most social networks contain directed links, or links that do not have the same strength in both directions [46-49]. Second, to prevent participants from colluding to recruit each other back and forth to gain rewards, real life RDS studies sample without replacement, meaning that respondents can participate only once. Third, it is difficult for respondents to accurately report their number of social contacts in the group (degree) [50]. Fourth, to avoid recruitment chains stopping early, researchers use more than one coupon (most often three coupons) [51, 52]. Finally, there is mixed evidence on wether participants on average invite social connections randomly from within their social networks [26, 53-56]. For example, participants may pass their coupons to peers with whom they have a close rather than a more distant relationship, which is not a random selection among peers [53, 57].

The level of variance may be equally as important as the level of bias. One study used real-world network data from a hidden population of injecting drug users and sex workers in the U.S. as well as from empirical social networks among several U.S. high school students, to simulate RDS studies [58]. That assessment showed that RDS studies within these populations could have very high variance.

Overall assessments of the effects of real-life deviations from the RDS assumptions are so far inconclusive, primarily due to lack of valid data on the true compositions of hidden populations with which estimates can be compared.

Improving the RDS Method: the Contribution of WebRDS

Non-random recruitment in RDS can be thought of as a type of selection bias. There are a number of potential causes of non-random recruitment in standard RDS, which may be possible to mitigate by transferring RDS to an online environment.

First, stigma is the primary reason for using RDS as a sampling method. While we know that MSM often are stigmatized, it seems unrealistic to assume that stigma would be perceived and enacted equally among all MSM in a society. It thus seems reasonable that some individuals have higher privacy concerns than others and that individuals' willingness to access survey offices physically will differ. Individuals who perceive a high level of stigma may thus be under represented in standard RDS studies.

Second, most RDS studies use financial incentives to compensate for the time and effort to travel to, and take part in, an RDS study. It seems likely that persons from middle- and upper income levels will be less incentivized by these financial rewards than persons of lower incomes. If the effort to participate can be reduced and non-financial incentives made to appeal to middle and upper income levels, bias is likely mitigated.

Third, like many other sampling methods, standard RDS involves relatively high costs since it requires strategically placed survey offices and skilled staff for extended periods of time.

Finally, the geographic area of study in a normal RDS study needs to be small enough to allow participants to travel to the study site. Participants who do not live close to a survey office are thus excluded or under represented.

14

All these potential deviations from the ideal RDS assumptions may be possible to mitigate by the use of webRDS. During the work with this thesis, three webRDS surveys have been published, two involving students at Cornell University [56, 59] and one among users of Facebook in the U.S [60]. The results of these studies showed that RDS estimates agreed relatively well with the true characteristics study population with the exception, in the Facebook survey, of under-sampling of participants who self-identified as Hispanic/Latino, African American/Black American and were of lower education levels. These surveys did not, however, target hidden populations, which is the subject of Paper III and IV in this thesis.

Sexual Transmission of HIV

Overview: Three Determinants of Epidemic Spread

In order to understand the drivers of the HIV epidemics among MSM, it is not enough to perform representative surveys. We need to understand the details of sexual spread between men and what types of behavior would be most important to target for HIV prevention interventions.

To understand population level sexual spread of HIV we can conceptually distinguish between three determinants [61]: I) the contagiousness of HIV during sex, II) the duration of infectiousness of HIV-infected persons, and III) the sexual contact structure of the population. The first and second area of inquiry consist roughly of questions that belong to the medical and biological sciences while the third area deals mostly with issues related to the behavioral sciences.

Transmission Probabilities during Sexual Intercourse

There is limited knowledge of the biological determinants of HIV transmission due to the inability to carry out randomized experiments in humans. It is known that the probability of an HIV- person being infected by HIV when having unprotected sex with an HIV+ person is dependent on a large number of factors which lie within the infectious person, the susceptible person and potentially also within the virus [62-65].

Pooled estimates across studies globally indicate a baseline risk of 0.07% for a person to acquire HIV during a single penile-vaginal intercourse with an infected partner [65]. However, under certain important circumstances the risk of infection is known to be considerably different from this baseline rate.

First, for penile-vaginal sex, male circumcision decreases the risk to acquire HIV with approximately 60% for the male [66-68].

Second, the period of time since an HIV+ person was infected influence the risk of transmitting the virus. During acute HIV infection (AHI), the first weeks to months after transmission has taken place, the body's immune system has not yet been able to efficiently combat the virus and the bodily fluids contain high concentration of the virus [69]. Also during late stage HIV infection, when the body's immune system is weak, the viral concentrations are again high, increasing the probability of transmission during sex [70]. The risk of transmission during penile-vaginal sex between a person during early and late stage infection is estimated to be respectively about nine and seven times higher than during the asymptomatic stage [65]. Risk of infection per act of receptive anal sex during the primary as well as the late stage infection has been estimated to 18% (95% CI 2.08–34.6) [70].

15

Third, the presence of genital ulcer disease (GUD) increases the risk of HIV transmission. In penile-vaginal sex the risk of transmission is estimated to be five times higher with GUD compared to without GUD [65]. Among men practicing anal sex, the presence of gonorrhoea and anal warts have been estimated to increase the risk of acquiring HIV approximately seven and four times, respectively [71].

Finally, the type of sex practiced greatly influence the probability of transmission. Oral sex carries little risk although transmission through this route can occur [72]. Anal sex on the other hand carries a much larger risk than penile-oral or penile-vaginal sex. A pooled estimate of 1.4% (CI: 0.2–2.5) has been reported for the probability of infection during a single act of receptive anal sex a risk. The risk of acquiring HIV through insertive anal sex is lower [64]. Jin et al [73] recently estimated per-contact probabilities for anal sex among men in Sydney, Australia. For unprotected anal intercourse they confirmed a per-contact risk of 1.4% if ejaculation occurred, and 0.65% if withdrawal occurred without ejaculation. They estimated the per-contact risk for insertive unprotected anal sex at 0.11% in men who were circumcised and 0.62% in uncircumcised men. Interestingly estimates of transmission risks per anal sex act and per male-male partnership, in which anal sex is practiced, are difficult to reconcile. This may be due to differences in infectiousness over time and potentially also to differences between couples. [64, 74].

Period of Infectiousness
The net median survival after HIV infection, in the absence of treatment, is estimated to 11 years. In countries where HIV subtype E accounts for most infections median survival time is estimated to nine years [75].

The period of infectiousness is determined largely by the person's average viral load (viral concentration) during the long asymptotic period of infection. The higher the viral load, the shorter the time until the infected person develops AIDS. For the virus there is thus a trade-off between the period of infectiousness and the probability of transmission during sex. There is evidence that the average survival times is a consequence of viral evolution, maximizing its long-term transmission success [62].

Sexual Contact Structures
Many infectious diseases, e.g. childhood diseases and the flu are easily transmitted during every day activities. The spread of these diseases can be predicted quite accurately by assuming that people are like molecules in a gas, randomly bouncing against each other, and in each encounter having a fixed probability of transmitting the infectious agent [76].

Sexual contacts patterns are very different from this random model. People do not choose their sex partners randomly on a daily basis, there is a large variation in people's number of sexual partners [77, 78], people differ substantially in how often they have sex with their partners [79] and individuals with a large number of partners tend to have sex with other individuals who have a large number of partners [80]. These heterogeneities have several important consequences for the spread of HIV. Some important consequences are covered below:

Studies have shown that HIV spread is much slower than what may be assumed from random mixing between people [81]. This is presumably because the virus will first spread rapidly among people with the highest level of risk behavior. When the virus then continues to people with lower level of risk behavior, the spread in the population becomes progressively slower. One type of risk-behavior is the number of sexual partners an individual has per unit of time. Assortativity is the extent to which people with similar

16

partner numbers have sex with each other. All things equal, in a highly assortative sexual network, spread is fast in the beginning and then slows down as the virus reaches people with less partners [61].

Transmission is much more efficient in a population with concurrent sexual partnerships compared to serially monogamous partnerships [82]. Considering two populations with the same number of lifetime sexual partners as well as the same number of lifetime intercourses, one can imagine one population in which people never have sex with a previous partner after switching to a new one, and another population that is free of such restrictions and where sex happen concurrently with several partners. It can be shown that populations having concurrent relationships will experience a higher incidence and prevalence of HIV than a serially monogamous population [83]. It has been argued that this mechanism accounts for the huge differences in HIV prevalence between Southern Africa and other parts of the world [84], although the strength of the empirical evidence for the hypothesis has been debated [85].

Sexual spread is slower in networks with a high degree of clustering. Clustering, also called transitivity, is the extent to which a person's sexual partners also have sex with each other. Clustering in a sexual network is usually measured by the relative number triangles in the sexual network, (figure. 2). The more triangles in a network, the slower the spread of the virus. The reason is that infected persons in a network with many triangles will tend to have sex more often with persons who are already infected rather than with uninfected persons who would then run the risk of acquiring HIV and spreading it onwards. Triangles can per definition not occur in sexual networks where people partner only with the opposite sex [61].
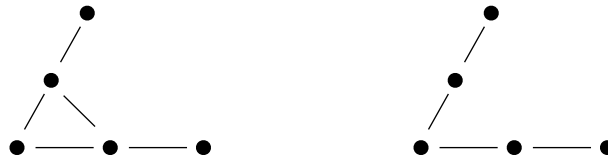

Figure 2: Two small sexual networks. One with a triangle and one without.

Populations likely also differ from each other in terms of their number of sexual partners per time period but also in their frequency of intercourse. There is likewise a huge variation between individuals within populations [86-90]. The large variance of partner numbers per time period is likely the reason why HIV and other STIs can circle within subpopulations with high numbers of partners, despite the fact that the infection would have died out if the same number of unprotected sexual contacts were distributed evenly in the population [76]. However, since people with many partners are unlikely to have the same number of intercourses per partner as those who have few partners, both the number of sex acts and the number of partners need to be taken into account to understand transmission patterns [79]. This issue is investigated in depth among MSM in Vietnam in Paper IV.

Summary of Background and Knowledge Gaps

Sampling of Men who Have Sex with Men

Representative sampling of MSM is extremely difficult due to stigma and discrimination of men who have sex with men as well as the lack of sampling frames from which to draw random samples. Respondent-driven sampling is a relatively new sampling method that may be able to circumvent these problems. There is however limited information about the performance of RDS among hidden populations.

Standard RDS methodology involve relatively high cost, it may violate the assumption that men recruit friends randomly from within their social networks and sampling can only cover relatively small geographic areas. Internet-based RDS (webRDS) may potentially mitigate these problems but has never been implemented among MSM or other hard-to reach populations.

Sexual Networks and HIV Transmission among MSM

Conceptually we can distinguish between three determinants of population-level transmission of HIV: I) the contagiousness of HIV during sex, II) the duration of infectiousness of HIV-infected persons, and III) the sexual contact structure of the population.

Available data on HIV prevalence and risk behaviors among MSM in Vietnam and many other LMICs indicate a highly worrisome picture but the drivers of these epidemics are not well understood. There is lack of representative data about several behavioral indicators such as distributions of partner numbers and frequencies of unprotected anal intercourse. Little is also known about the real-world structures of sexual networks of MSM, despite their potentially profound influence on transmission.

AIMS AND RESEARCH QUESTIONS

Main Aim

To develop a system for representative sampling of internet-using men who have sex with men (MSM) in Vietnam and to increase knowledge about sexual risk behavior for HIV among MSM in Vietnam.

Research Questions

1. What is the reliability and validity of the RDSII estimator?

2. Is webRDS a feasible method for sampling internet-using MSM in Vietnam?

3. What are the characteristics of sexual contact patterns among MSM in Vietnam and how conducive are these to HIV transmission?

OVERVIEW OF STUDIES

Research Questions	Methods	Articles
What is the reliability and validity of the RDSII estimator?	Computer Simulations	I
Is webRDS a feasible method for representative sampling of Internet-using MSM in Vietnam?	webRDS surveys (2011&2012)	III&IV
What are the characteristics of sexual contact patterns among MSM in Vietnam and how conducive are these to HIV transmission?	In-depth Interviews	II
	webRDS survey (2012)	IV

The three main research questions in this thesis span four papers. Three types of methods are used: computer simulations, in-depth interviews and webRDS surveys. The four papers are numbered in the order they were implemented. Paper II contains also other information besides what is relevant for the last research questions and may therefor be interesting to read before the two webRDS studies (Paper III and IV).

METHODS

Study Setting: Vietnam

Overview
Vietnam has an estimated population of 92 million people (2012) and is the 13th most populous country on earth [91]. It is situated in South-East Asia bordering China to the north, Laos and Cambodia to the West and the South China Sea to the east.

The area of present day Vietnam has historically been influenced and ruled foremost by Chinese dynasties, which has heavily influenced Vietnamese culture. Vietnam became an independent vassal state to China in 938 AD. The French occupied the country in 1858 and incorporated it into French Indochina. A communist North Vietnam emerged after the second world war. Military conflict with France, U.S.A. and South Vietnam continued until reunification in 1975.

Starting in 1986, Vietnam moved towards market economy, a political path called doi moi ("renovation"). The reforms led to rapid economic growth and an enormous social change affecting virtually every aspect of Vietnamese social life. These developments include e.g. increased income disparities, rapid urbanization, increased internal migration, reduced fertility rates and an increased influence of foreign cultures [92].

Demographics of Vietnam
The Vietnamese population is young (figure 3) and grows at 1.1% annually, closely paralleling that of the global average. The total fertility rate (live births per woman) has fallen sharply and is 1.7 [91], thus below the long-term replacement fertility of 2.1 children per woman.

Urbanization has been rapid but the urban population still encompass only a quarter of the population [93]. People are heavily concentrated to the Red River Delta in the north, the Mekong Delta in the south as well as along coastal settlements (map 1).

Fifty-four ethnic groups are recognized by the government, out of which the Kinh group constitutes 86% of the population [94]

Administrative, Political, Social, and Economic Situation
Vietnam is divided into 58 provinces and five centrally controlled municipalities (map 2) [95].

Vietnam is a one-party state and the communist party is the dominant political institution. The party leads numerous political and social associations, it directly or indirectly appoints the president, prime-minister and candidates for seats in the National Assembly [92].

Vietnam has a Confucian heritage, placing a high value on the family, which is seen as carried on by the male lineage. Sons are highly valued [92].

Vietnam has recently become a middle-income country. The GDP per capita is $3,549 (current international dollar, PPP) (figure 4) [96].

Vietnam's Gini coefficient (2008) is low to moderate at 36 (2008), not far from Sweden's 33 (2010) [97]. Literacy is high at 93% of the population above 15 years of age (2009) [98].
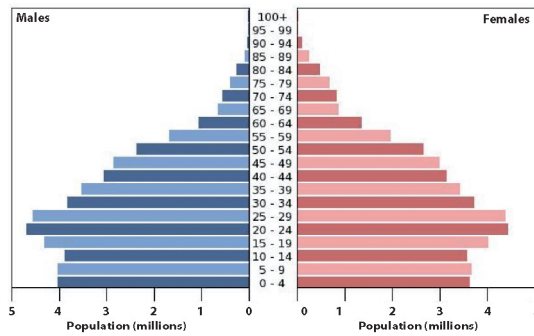

Figure 3: Population pyramid for Vietnam 2012. Adapted from U.S. Census Bureau [99].

Health Situation

Average life expectancy in Vietnam has seen tremendous growth and is now estimated at 75.5 years [100]. The proportion of children dying before the age of five years has fallen to 2.3%. Twenty percent of children under five years of age are moderately to severely underweight [101].
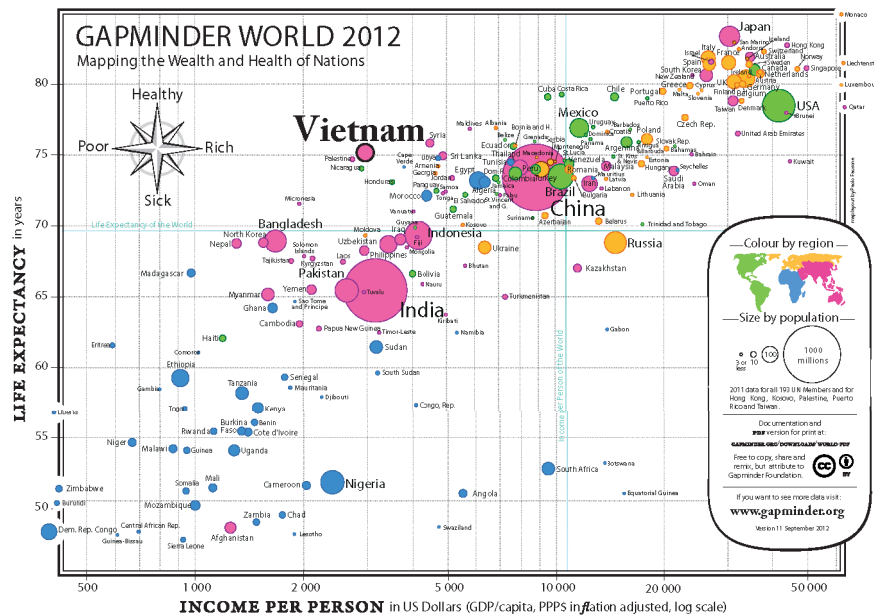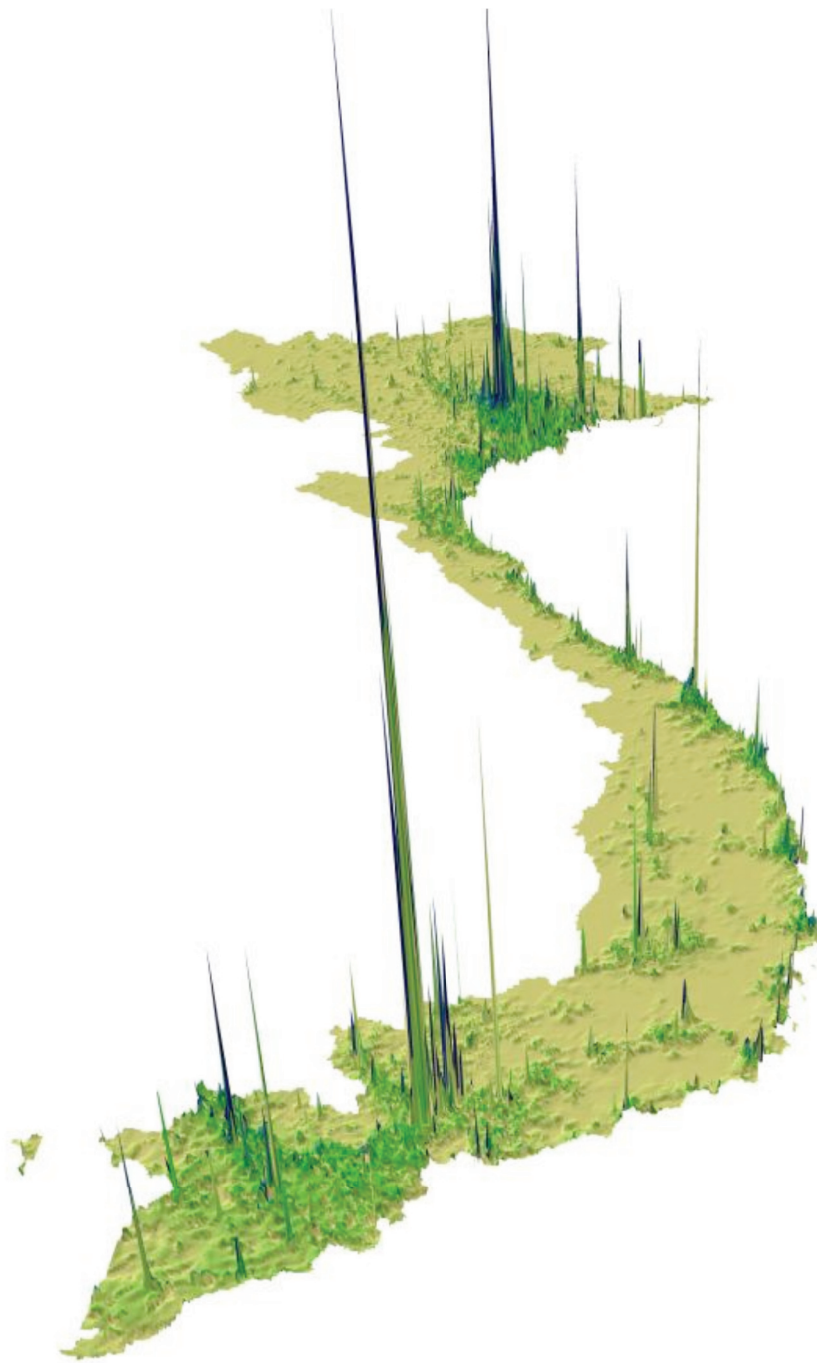

Figure 4: Gapminder World Map 2012.Vietnam highlighted [102].

22

Map 1: Relative population density in Vietnam based on 100 meter square satellite images [103]. Reproduced with permission of Andrew Tatem, Florida State Univ..

Map 2: Administrative division of Vietnam [104]. *Ha Tay has ceased to exist and is now part of Hanoi municipality.*

24

MSM in Vietnam

Gender Roles, Stigma and Societal Attitudes
In Vietnam as elsewhere in Asia, identities of MSM are based not only on whether a man has sex with men or women, but also on a man's gender role. A man may thus have a masculine role in society and perceive himself as a man but nevertheless have sex with other men, without acquiring a homosexual identity. Other MSM perceive themselves as fully or partly female and seek male sex partners who are masculine [105]. The terms used by MSM themselves describe MSM according to the person's degree of femininity, as well as according to how closeted or open he is and his appearances. "*Bong lo*" is the term for men who dress and present themselves as women and are open about their sexual preferences for men. "*Bong kin*" refers to masculine-looking MSM who are generally not open about their sexual preferences outside the MSM group [105, 106]. Identities of MSM in Vietnam are, however, more nuanced and varied than these overarching categories indicate, and Ngo et. al [105] have documented a wide range of terms and categories in use. These include, for example, terms for someone who could be classified as either a man or a woman (*người mang hai dòng máu, hột vịt lộn*) and someone who has sex with both men and women (*hi fi, supersim*). European and American homosexual identities have also impacted on the way young homosexual men construct their identities and the term "gay" has become increasingly common [105].

Though not illegal, homosexuality is still heavily stigmatized in Vietnam [106, 107]. Until recently homosexuality was not acknowledged publicly at all in Vietnam, but over the last decade effeminate men have become increasingly visible, as have bars and venues specifically catering to MSM. Homosexuality has also been given increased attention in the media but is often portrayed as an import from the West [108].

MSM and Internet-Use
An estimated 34% of the Vietnamese population (31 million persons) used the internet in February 2012 [109]. Internet provides an important environment in which MSM communicate with each other and meet partners. Internet-use among urban MSM may thus be considerably higher than in the general population [105]. Ninety-four percent of MSM in an offline RDS in Hanoi stated that they used the Internet [110]. Internet access in Vietnam costs approximately 0.15 USD per hour at Internet cafés.

HIV in Vietnam

National Estimates
According to the preliminary results of the 2011 Viet Nam HIV/AIDS Estimates and Projection report the estimated HIV prevalence in Vietnam is 0.45% among people aged 15 to 49 years. The HIV epidemic is concentrated primarily among men who inject drugs (MWID), men who have sex with men (MSM) and female sex workers (FSW) [111]. However, detailed description of sampling methods and data collection makes it difficult to assess the level of precision in estimates.

The proportion of women among people newly diagnosed with HIV has been rising and is now 31% [111]. It is believed that most of these women have been infected through sex within long-term relationships [112].

According to the 2011 sentinel surveillance, the HIV prevalence among men who inject drugs and female sex workers was 13% and 3% respectively [111]. Reported prevalence rates are however highly heterogenous across the country. Among men who inject drugs

the HIV prevalence in sentinel surveillance surveys in Quang Ninh, Hai Phong and HCMC was 56%, 48%, 46% respectively, while only one percent in Da Nang [111].

HIV among Men who have sex with Men

Reported HIV prevalence rates among MSM in Vietnam has increased from 2005 to 2009. HIV prevalence rates in Ha Noi and Ho Chi Minh City was reported at nine and five percent respectively in 2005 [113] and at 15% and 17% respectively in 2009 [111, 114]. A number of other surveys have been performed in large Vietnamese cities during 2005 to 2011 [115-121]. Prevalence rates in these surveys have ranged from two to 37% with most point estimates lying between five and 15%. Values higher than 30% have been found among male sex workers in HCMC [122] and Hanoi [118]. Among many of these men, injecting drug is however the likely primary transmission route. In the latest large survey of MSM (IBBS round II) there were no major differences in HIV prevalence rates between MSM who sold and who did not sell sex [114]. Two surveys have been performed outside large cities, both in Khanh Hoa province. Prevalence rates in these surveys were zero and 1.9 percent respectively [121, 123].

Mean age at first sex has been estimated to approximately 18 years among MSM in the four largest cities. Depending on location, 66-92% reported sex with a male partner during the last month [114].

Anal sex is relatively common. In the IBBS round II, approximately 50% to 80% reported anal sex in the past month and around 25% of respondents reported anal sex with two or more partners during this period. The proportion reporting four or more anal sex partners last month ranged from 4.6% to 7.1% [114]. Forty-seven percent of rural MSM in Khanh Hoa province reported having engaged in anal intercourse with a casual male partner in the previous six months [121].

The proportion of surveyed MSM selling sex in the past month varied between seven and 52% and mean number of anal sex clients ranged from 0.5 to 3.6 clients during the same period [114]. One study of male sex workers in Hanoi reported that 79% of men who sell sex in Hanoi came from other provinces and 74% of the men reported being exclusively attracted to women [117].

Consistent condom use was reported by 30 to 54% during the last month in the four biggest cities, according to the latest IBBS in 2009. For men selling sex the corresponding figures with clients were 20-47% [114]. Colby and Mimiaga has presented preliminary data from a cohort of male sex workers in HCMC, where 36% and 22%, in 2010 and 2009, respectively, reported UAS during the past month [115, 119].

The IBBS round II reported 19-28% of MSM in the four biggest cities to ever have had an HIV test and getting the result [114].

As evidenced by the considerable variability in estimates, there is substantial uncertainty both regarding the prevalence of risk behaviors and the prevalence of HIV. Data on anal sex frequency or studies of sexual networks have not been reported. No modeling of HIV transmission within the MSM group has been performed.

Study Design, Data Collection and Analysis

Terminology

To draw graphs is a way of visualizing and formalizing interactions between objects. In studies of sexual and social networks the objects or nodes in a graph are in most cases persons. A link on a sexual network denote the presence of one or several defined sexual acts, during a specific period, e.g. the number of anal intercourses during the last three months. Links in social networks may denote for example that a person knows another person by name or have communicated with that person during the last seven days. A number of concepts in network science have received specialist terms (table 1).

Term	Meaning
Vertex or Edge	A link between two nodes
Node	A connection point in a network. In sexual and social network science it refers to a person
Directed link	A link that has a direction, e.g. person A knows person B
Reciprocal link	A link that goes in both directions, e.g. if person A knows person B, B also knows A.
Irreciprocal link	A link that does not go in both directions, e.g. person A have heard person B on the radio
Connected Component	A group of connected nodes that are not connected to other nodes in the larger graph
Giant Connected Component (GCC)	A connected component consisting of a large fraction of the nodes in the larger graph

Table 1: Network terminology

Computer Simulations of RDS Studies *(Paper I)*

Overview

In Paper I we simulated RDS studies on a real online network of homosexual men. In these simulations we systematically, and to varying extent, violated the assumptions behind the RDS method and then described the effects on the RDS II estimates.

Data Collection

The data comes from a real social network called Qruiser (www.qx.se). This web site was at the time of the study the Nordic region's largest and most active web community for homosexual, bisexual, transgender, and queer persons. Contacts between members on the web site are maintained mainly by a so called "favorite list", on which each member can add any other member without approval from that member. Members can attend clubs (web pages about specific topics) and send messages to each other [124].

We collected, in collaboration with the web site, anonymized information on personal profiles and records of each messages sent within the web-community from December 15, 2005, to January 18, 2006. The content of messages were not recorded and we did not collect any other personal identity data. During the data collection period, 12,590,911 messages were recorded and there were 184,819 distinct members registered on the web site.

Generation of Social Networks

Based on selected variables in the membership profiles, we extracted a network that contained only members characterizing themselves as homosexual males. We defined an outgoing edge to be formed if a member had another member on his favorite list. We called an edge reciprocal if a connected pair of members had both an ingoing and an outgoing edge between each other. If a pair did not have both an ingoing and an outgoing edge between each other, it was called irreciprocal. To avoid the inclusion of inactive persons, we required members to have sent at least one message to any other person on the site during the data collection period.

For our research purpose, we kept only members of the Giant Connected Component (GCC) as nodes (16,082 active, gay men). We defined the GCC as the largest component connected with only reciprocal edges. Keeping only the reciprocal edges in the GCC, we obtained an undirected network (G1) with an average degree of 6.74 edges per member. Keeping both the reciprocal and irreciprocal edges, we obtained a directed network (G2) with an average degree of 17.2. Note that our definition of the GCC ensures that all nodes had a chance to be recruited in both G1 and G2. Degree distributions for both G1 and G2 are plotted in Figure 5. The degree distributions were very skewed, e.g., half of the members in G2 have no more than 10 outgoing edges, while a small proportion of members had a large number of outgoing edges.
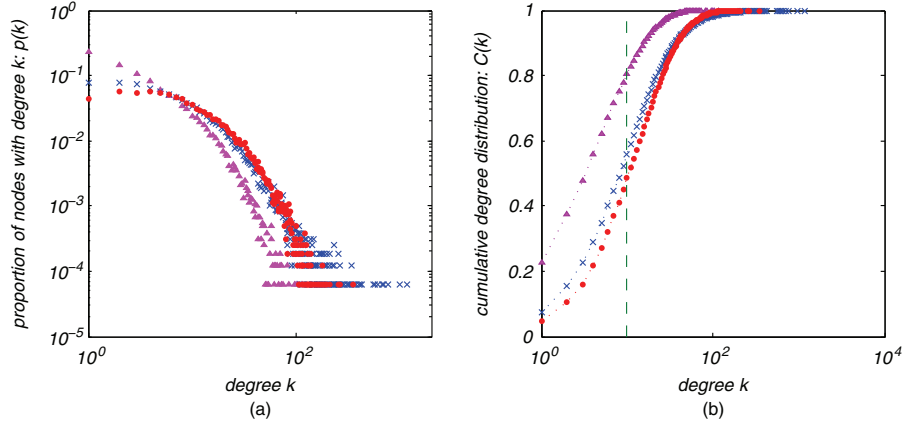
Figure 5: (a) Degree distribution and (b) Cumulative degree distribution. The out-degree is the number of outgoing edges that leave a node and the in-degree is the number of ingoing edges that point to the node. Pink triangles: degrees in G1; blue crosses: out-degree in G2; red circles: in-degree in G2.

The detailed procedures for generating these networks and variants of these networks can be found in Appendix A of Paper I. To test the effects of preferential recruitment in RDS, we weighted each reciprocal edge in G1 in two ways: by the maximum number of messages sent in any one direction, and by the minimum number of messages sent in any one direction. For example, if node A sent ten messages to node B and received five back from B, the weight on edge $e_{A,B}$ ($e_{B,A}$) was set to ten for the maximum-weighted network (G1$_{max}$) and five for the minimum-weighted network (G1$_{min}$). In these two weighted networks, respondents were supposed to recruit peers with probability proportional to the edge weights.

Simulations

We first ran simulations on the undirected network, G1, with all the RDS assumptions satisfied. We started each simulation with a single randomly selected seed and we restricted the number of coupons to one, that is, each recruiter could only recruit one other person (RDS assumption six, see background). All respondents were selected randomly from the recruiters' personal network (assumption five), and nodes could be selected multiple times (sampling with replacement, assumption two). Since all participants' degrees were assumed to be known by the participants themselves, and G1 is a single connected component with only reciprocal edges, assumptions one, two and four were also satisfied. We kept recruiting participants until the sample size reached 10,000 participants.

After establishing the performance of the RDS estimator under these ideal conditions we repeated simulations under six types of real-life-like violations of the RDS assumptions, according to the following:

1. We performed simulations on the directed network (G2) instead of the undirected network (G1)
2. We performed sampling with and without replacement.
3. We allowed participants to reject invitations as well as letting participants ignore peers when inviting. For each invited member, the probability of rejecting an invitation is called P_r. A rejected coupon was discarded and not reused for recruiting

a new member. Additionally, we let the recruiters ignore edges when inviting members from his or her personal network, since recruiters presumably will not remember all people they could potentially invite. In the simulations, we thus gave each edge a probability of being ignored. We call this probability P_i. An ignored edge was given a zero probability of being selected and was not included in the network size of the participant when we calculated the RDSII estimates. We set the number of seeds to ten and coupons to three to make sure that a sample could still be recruited when P_r and P_i became large.

4. We let P_r and P_i depend on the characteristics of the participants. Specifically, we let P_r and P_i be the probability that a member in the group of interest will be ignored by his friends, and P_r' and P_i' the corresponding ignore and reject probabilities for members who are not in groups of interest.

5. We allowed friends with whom the participant communicated often with to have higher probability of being recruited.

6. We varied the number of seeds and coupons.

Variables and Homophily

We estimated in each simulation four different variables: age (born before 1980 versus others), County (living in Stockholm versus others), civil status (being single versus others) and profession (employed versus others). The two variables age and county of residence had high homophily while civil status and profession had low homophily. Taking county within the undirected network G1 as an example, members who lived in Stockholm formed edges with members who also lived in Stockholm 50% of the time, whereas they formed edges randomly among all cities (including Stockholm) the remaining 50% of the time. The homophily for people living in Stockholm was thus 0.5. Homophilies for civil status and profession were very small, indicating that edges were formed as if members, regarding civil status and profession, chose randomly among other members (table 2).

	Age		County		Civil status		Profession	
	Before 1980	Others	Stockholm	Others	Single	Others	Employed	Others
P* (%)	77.77	22.23	38.79	61.21	40.39	59.61	38.19	61.81
H G1	0.40	0.37	0.50	0.40	0.05	0.08	0.13	−0.05
H G2	0.23	0.34	0.50	0.28	0.03	0.07	0.06	0.02

Table 2. Proportions P* and homophilies H, in the undirected (G1) and directed networks (G2).

Measures

Four measurements are used throughout the study: *the average estimate* (AE), defined by the mean of the *RDSII* estimates, where est_{ij} is the *RDSII* estimate at the ith simulation when sample size is *j*; *the bias*, defined by the absolute difference between AE and the true population composition, $Bias_j = |AE - P*|$; *the standard deviation* (SD) of estimates for a given sample size *j*, $SD_j = \sqrt{\sum_{i=1}^{m}(est_{ij} - AE_j)^2 / (m-1)}$; and finally *the mean absolute error*

(MAE) of estimates for a given sample size *j*, $MAE_j = \sum_{i=1}^{m}(|est_j - P*|)/m$.

30

In-Depth Interviews with MSM in Hanoi (Paper II)

Sampling and Study Population

Sampling and study population: We conducted purposive sampling of 17 men to reach MSM from a wide range of backgrounds. Interviews took place between November 2008 and September 2009. Two-thirds of the informants had previously participated in a respondent-driven sampling (RDS) study among MSM in Hanoi [110]. They had, at that time, expressed interest in participating in other research studies and were thus contacted. In order to round out the data with information from MSM who were poorer, older, or had migrant and rural backgrounds, an additional six informants were obtained through snowballing from the first group. Inclusion criteria were person biologically born as males, living in Vietnam, 18 years and above, who had ever had sex (any type) with another man. Eight men from the previous RDS study declined to participate when contacted on the phone. We did not ask for their reason to do so. Participants ranged in age from 19 to 48 years. Half of them were born in Hanoi, and the other half were born in villages or smaller towns in northern Vietnam. Six men lived with their parents, one with his sister, one with his wife, two with their boyfriends, three with friends and one man lived alone. Three men had children. The men worked as building workers, porters, shop assistants, hair dresser, teacher, medium level managers in private companies, as government employee, computer programmer, two ran their own shops and one was a student.

Interviews

The interviews, which lasted between 40 and 90 minutes, were carried out primarily at the Institute for Studies of Society, Economy and Environment (iSEE), a Hanoi-based Vietnamese research organization working for LGBT rights. When preferred by participants, interviews took place at a café chosen by the informant and in the case of one informant, in his home. 15 interviews were carried out by TPV, a female Vietnamese sociologist, with LB (male Swedish medical doctor) present. LB carried out two interviews with an interpreter, and AT (female Swedish medical doctor) was present at one interview conducted by TPV. All interviews were conducted in Vietnamese. The men were given time to freely express their experiences and thoughts. During most interviews the atmosphere was warm and most men were eager to share their experiences and thoughts. The interviews were tape-recorded, transcribed verbatim, and translated into English. Two translators were used and the translations were checked for accuracy by TPV.

A thematic interview guide was used, focusing on the men's sexual and love relationships with other men, though the interviews in most cases evolved as informal conversations. In practical terms this meant the informant was asked to relate his experiences of sexual and love relationships from the time when he realized that he was attracted to other men, and up until the present. When the interviewees had been in a large number of relationships, the emphasis was put on those relationships that the participant viewed as being of longer duration or of a deeper emotional significance. Topics emerging from one interview were explored in subsequent interviews.

Analysis

Latent content analysis was used. LB and RP performed the initial analysis of the interview transcripts focusing on both preconceived research questions and emergent themes. Research questions included beliefs about, experiences of, and attitudes towards male-male relationships; relationship preferences and perceived advantages and disadvantages to MSM relationship patterns, as well social norms around male-male relationships. An emergent issue during the analyses was the inconsistent statements in many of the interviews. Notes taken during and after the interviews on non-verbal aspects of communication and the "feel" of the interview helped to sort out what these many contradictions meant. During and after the initial analysis of the data the other authors read and commented on the analysis, based on their own readings of the transcripts. Together the authors reached a consensus about how to make sense of and best represent the men's narratives.

Implementation of Internet-based RDS (Paper III)

Study Design

This survey was cross-sectional, performed online and carried out between February 18 and April 12, 2011, applying web-based respondent driven sampling (webRDS).

Inclusion Criteria

Eligible participants were adult men (18 years and above) who had ever had any type of sex (including oral sex and mutual masturbation) with another man, had not previously participated in the survey, and were living in Vietnam at the time of the study. The Internet-using part of this group formed the population to which the sample aimed to generalize.

Sampling

The study was performed in collaboration with a local research organization in Vietnam working to promote LGBT and ethnic minority rights (iSEE). iSEE has an extensive knowledge and contact network among MSM community groups and a close collaboration with web administrators of Vietnamese LGBT web sites. Fifteen seeds, who were recruited through these networks, initiated the survey and a further five seeds were added two weeks later to increase the speed of recruitment. Six seeds came from Ho Chi Minh City, ten from Hanoi and four from Hoa Binh. Nineteen out of the 20 seeds had attended some kind of education after high school (vocational training, college or university). Participants received, from their recruiter, an invitation message with a login code and a web address. They logged in, accessed detailed information about the study, approved participation and eligibility and answered a written questionnaire. Participants could then compare their own answers to aggregated answers of earlier participants, displayed in informative bar charts. On the last page the participants were encouraged to recruit MSM friends by providing an e-mail or Yahoo! Messenger address (popular for communications in Vietnam), and being automatically sent four invitation messages, which could be forwarded to MSM friends. The messages were also displayed on the screen and could be copied for sending by other preferred means. Text both on the web site and in the email/Yahoo! chat messages emphasized that only MSM living in Vietnam and of age 18 years or above were allowed to participate. A warning was included saying that advanced checks were applied and that failure to follow the recruitment rules would mean loss of compensation. No restriction was given as to whether the recruiter knew each other in real life or only through the Internet. Reminders to recruit were sent out two and four days after completing the survey. Participants were informed that they had seven days to recruit and were given rewards for recruitments that took place during that time.

Some participants took the survey at a later time point. They were retained in the sample and the persons they recruited were given standard compensation.

Web Site
The graphic design of the web site aimed at giving a professional and friendly impression (figure 6).

Figure 6: Screenshots form the English and Vietnamese versions of the survey site (2012). Color tones and certain graphic details differ from the live web site.

Incentives and Recruitment Stimuli
1) 2.45 USD (50,000 VND) as credit on the participant's SIM card and the same amount for each successful recruitment of an MSM friend (maximum four); 2) the option of donating the monetary reward to an MSM community organization chosen by the participant; 3) a lottery with the possibility of winning an iPad; 4) text emphasizing participation in order to support MSM in Vietnam; and 5) being able to compare one's own answers to those of other participants in simple, informative and anonymous charts.

Eight questions were included in the questionnaire specifically to stimulate the participants' interest in comparing themselves with other participants.

Piloting and Early Versions of the System
The web site and recruitment system was extensively pilot tested. Interviews and focus-group discussions among MSM were performed to understand social networks among MSM, online interaction and to decide on appropriate incentives. Two versions of the webRDS site were used for sampling before the study described in this paper was carried out. These webRDS systems differed in that they had a less advanced graphic design and smaller incentives. In the first survey in 2009, recruitment died out after a maximum of 5 waves (25 participants, 15 seeds). The second time, recruitment improved but stopped after 5 waves (84 participants, 15 seeds).

Data Collection
The questionnaire contained 17 questions, including number of sexual partners in the past six months, sexual partner preferences (prefer as sexual partners only men, men to women, women to men or only women), the duration of the respondent's longest relationship, opinion on legalizing same-sex marriage in Vietnam (for or against), frequency of Internet-use, socio-demographic characteristics, network size (see separate heading), relationship between the participant and his recruiter (stranger, acquaintance, friend, close friend, lover/ex-lover, or relative), and the social context in which the participant got to know his recruiter. Logical checks with error messages were used for interdependent questions. Only positive integers were allowed for numeric answers. All questions included a "don't want to answer" option and all questions needed to be answered. Participants who wanted to receive rewards filled out contact details and a personal identifier (telephone number, email or Yahoo! Messenger address, and the last three digits of their nine-digit ID number). Time points at which each participant loaded the web pages was stored to facilitate identification of ineligible submissions, including unserious attempts to answer the questionnaire or the same person trying to answer more than one questionnaire to receive additional rewards.

Analyses of Duplicated Submissions, Data Cleaning and Analysis
9.6% of completed surveys (65 surveys) included a stated age below 18 years, or a telephone number, e-mail or Yahoo! Chat address that had previously been registered in the system. We defined these as "invalid". We excluded seeds (customary in RDS analysis [42]) together with the aforementioned invalid submissions to produce a cleaned sample. From this sample we estimated, in Matlab©, population proportions using the current state of the art estimator, RDSII, which requires only information on the sample compositions and the social network sizes of the participants [43]. We have not included confidence intervals in this paper since there is currently no consensus on how to best estimate RDS design effects.
We checked all surveys for other signs of duplication or invalidity by flagging surveys containing a repeated IP number, deviating answers (as described below), or short completion times. We analyzed the sensitivity of the estimates to inclusion and exclusion of these flagged submissions. Specifically we compared the RDS II estimates generated from the full sample of non-seed submissions with valid age with the RDS II estimates generated from groups with progressively stricter inclusion criteria according to the following: 1) exclusion of submissions with a repeated email, Yahoo! Chat ID or telephone number (forming the cleaned sample above); 2) additionally excluding repeated IP numbers; and 3) additionally excluding submissions with short completion times (<

three minutes), submissions stating no education (rare in Vietnam), or submission stating six-month partner numbers above 1,000. Differences were small between the groups. Details are included in the supplementary material. For all estimates in the supplementary material the maximum absolute differences when comparing the full sample to the groups with progressively stricter inclusion criteria were 6.6%.

Personal Network Size
We asked participants for the number of MSM they had interacted with in any way during the past seven days (including on the phone, Internet, or in person). We then asked how many of these persons they believed used the Internet. We chose the seven-day time-frame to reflect the potential high frequency of contacts online. We used the second network question to define the participants' personal network. We replaced missing personal network size data with the RDSII-estimated average network size from submissions with non-missing network data. The average network size was 5.5 persons.

Evaluation and Analyses of Equilibrium
As there is no gold standard by which to validate the sampling, we evaluated the system in terms of its success in generating sustained recruitment, the degree to which the sample compositions stabilized with increasing sample size (independence of the sample from the seeds), and finally, in the discussion, we contrast the sample compositions with results from other surveys.
We analyzed whether equilibrium was achieved in two ways. We first used the standard criterion from the RDS study literature [41]. This criterion requires the sampling process to have reached a certain number of waves. The number of waves required is, for each variable, determined by the number of steps required by a first-order Markov process to reach a less than a two percent relative difference between its value at a given step and its value after an infinite number of steps. The transition probabilities used to calculate the values of the Markov process are the averaged transition probabilities in the study's recruitment chains. Second, we produced plots of the changes in the sample compositions as sample size increased.

Sexual Behavior and HIV Transmission Modeling *(Paper IV)*

Study Design, Inclusion Criteria and Sampling:
The survey was cross-sectional, performed online and carried out between 26 March and 22 April, 2012, applying web-based respondent driven sampling (webRDS). Eligible participants were adult men (18 years and above) who had ever had sex (any type) with another man, had not previously participated in the survey, and were living in Vietnam at the time of the study. Fourteen seeds, were recruited through MSM contacts of iSEE. Seven seeds came from Ho Chi Minh City, two from Hanoi and five from the provinces Khan Hoa, Bac Ninh, Dien Bien, Ha Tinh and Son La. Nine-hundred-sixty-nine non-seed submissions were received during the study. The maximum number of waves in the recruitment trees was 29 and five trees exceeded five waves. The rest of the sampling methodology was identical to the survey described in Paper III.

Data Collection
The questionnaire collected socio-demographic information (age, province, education, income, frequency of Internet-use, sexual preference, social network size and type of relationship between the participant and his recruiter). Sexual behavior data included, for the last three months, the number of male sexual partners (including manual, oral and anal sex), number of anal sex partners and number of partners with whom the respondent

practiced unprotected anal sex (UAS). For each of the three most recent UAS partners, data was collected on the number of UAS acts during the last three months and whether the partner was casual or respondent was emotionally involved with his partner. Regarding the latter question, qualitative research [125] and piloting in preparation for the study indicated that it was difficult to capture the dichotomy of "stable" versus "casual" sexual relationships, due to that many sexual relationships were short and unstable but respondents could nevertheless perceive certain relationships as non-casual, based on love. Additionally, Vietnamese language lacks a word for "relationship" which could be used appropriately in this context. Specifically we therefor asked in Vietnamese if the respondent was "emotionally involved or mutually in love" (*quan hệ tình cảm hoặc tình yêu*) with his sexual partner, excluding that the partner was casual. Henceforth, we use the terms "casual" and "non-casual" partners, keeping in mind those casual partners can include e.g. long-term transactional partners.

Data Cleaning

Seeds were excluded as is customary in RDS analysis. Questionnaires were analyzed for ineligible or duplicate participation. Eight point seven percent of questionnaires contained an email, Yahoo chat addresses or telephone numbers that had earlier been registered in the system. These questionnaires, together with entries with a stated age under 18 years or no answers for all socio-demographic questions (4.3%) were excluded. The remainder of the sample (88.4%) formed the cleaned sample. In addition we checked all surveys for other signs of duplication or invalidity by flagging surveys containing a repeated IP number, deviating answers, or short completion times. We analyzed the sensitivity of the descriptive estimates to inclusion and exclusion of these flagged submissions. Differences were present for the variables province and frequency of internet-use but were otherwise small (Paper IV, Table. S1). We also checked selected analyzes (see Paper IV, Table 2, Figure 3a and 3b and Figure 4a and b), with and without flagged submissions and there were no qualitative changes to the results.

Data Analysis

Equilibrium [41] was reached after a maximum of 11 waves for all variables reported in this paper. In the supplementary material we also show the change in the sample compositions with increasing sample size. Ninety-five percent confidence intervals for proportions were calculated using bootstrapping, according to Salganik [126] and for numeric estimates a design effect of four was used [59]. RDSII-adjusted network size of submissions with non-missing data (12 persons) was calculated as in the previous study.

We performed multivariate and univariate logistic regression to explore the associations between UAS at least once during last three months (dependent variable) and the socio-economic variables age, income, education, province, frequency of Internet-use and sexual partner preferences (prefer as sexual partners only men, men to women, women to men or only women). The same analyzes were also repeated with the dependent variable two or more UAS partners last three months. We followed the recommendation from Ramirez-Valles et al. [127] to perform regression analysis on unweighted data and include in the model all variables associated to network size (education and frequency of Internet-use, according to Kruskal-Wallis rank test). All results in the article are RDSII adjusted if not otherwise specified. Age, income, education and province include collapsed categories for clarity. Poisson regression was performed to investigate the correlation between the number of anal sex partners and number of UAS acts. As above, these analyzes included the variables education and frequency of Internet, which were associated to network size.

Modeling

A weighted configuration model [128] was used to model the network structure and to estimate the basic reproductive number [129] for varying per act transmission probabilities. Note that the basic reproductive number in this study, henceforth denoted $R0^{3M}$, refers to a three-month period instead of the entire infectious period. This period was chosen to minimize recollection bias and to approximate the highly infectious primary infection period. The weighted configuration model is a development of the standard configuration model [130]. In the latter model a sexual network is recreated from ego-network data by randomly connecting reported links (partnerships). The weighted configuration model takes advantage of the fact that both partners in any sexual partnership will have the same number of UAS acts and connects links with the same UAS frequency. Although there are scenarios, which will not be correctly modeled by the weighted configuration model, for example when UAS frequency distributions differ between spatially distinct groups, many realistic cases yields a much more plausible network structure than the standard configuration model. The rational behind the model is further explained in Figure 1a-d, Paper IV.

For men with more than three partners we imputed data on the number of UAS acts for partner four and above, before estimating $R0^{3M}$ in the weighted configuration model, since observations with more than three partners were censored. Differences in imputed data on the number of UAS acts for persons with high numbers of partners can potentially influence outcomes heavily. We therefor used two different imputation strategies (a best and a worst case scenario) for partners four and upwards. We imputed for each individual with >3 UAS partners a) one UAS act per partnership four and upwards (the logically smallest possible value) and b) the smallest reported number of UAS acts among his three partners for which UAS act data was reported. To investigate the transmission effects stemming from correlations between partner numbers and frequency of UAS acts, we compared $R0^{3M}$ calculated from the original data (i.e. with correlations present) and when the values for the number of UAS acts per partnership first had been randomly shuffled between all partnerships (i.e. with no correlations present). This way of modeling transmission on the network is equivalent to using the standard configuration model. Both shuffled and non-shuffled estimations were done for the best and worst case scenarios.

Ethics

Ethical considerations are especially important when studying hidden and stigmatized groups.

In Paper I, the data were extracted and anonymized with approval from the Regional Ethical Review Board in Stockholm (EPN). It was not possible, based on the data, to identify real individuals.

In paper II, all participants were informed about the content of the study on the phone before coming to the interview location. After reading detailed information about the study, all participants provided written consent or, if they did not want to write their name for reasons of anonymity, they provided clear verbal consent. Participants were not required to reveal any personally identifying information. The study was approved by the Hanoi Medical University Review Board for Bio-Medical Research.

In paper III and IV, we put a lot of effort into making the site and the recruitment system completely safe and confidential for the participants. We also provided the users with extensive information about the study. All information about the survey was available on

all web pages except the log-in page and could be accessed at any time. All pages included a log-out button, which automatically removed traces of the survey from the computer and transported the user to Google.com. Browser history has to be deleted manually by the user and participants were given detailed instructions on how to do so. Telephone and chat support were available. IP addresses were converted to a unique anonymous code using a one-way encryption algorithm, and the original IP numbers were deleted. Login passwords were only valid for a single session and could not be used on two computers simultaneously. Communication between the users and the server was encrypted. Graphs with aggregated survey results were updated with data from a large number of participants at a time and displayed so that it was impossible to understand what others answered. The studies were approved by the Hanoi Medical University Review Board for Bio-Medical Research and the regional ethical board in Stockholm.

RESULTS

Performance of the RDSII Estimator *(Paper I)*

Overview
Paper I includes the results a very large number of combinations of violations of the RDS assumptions and describes the results of several million simulated RDS surveys. Below I therefore focus on key findings and the interested reader can find the complete results in Paper I.

Ideal Scenario
Simulations of RDS studies with all assumptions fulfilled gave an SD of around 0.05 and the an MAE of around 0.04 when sample sizes were between 500 and 1000 participants. The design effects were around 13 and ten for age and county respectively, and five for civil status and profession (see figure 7).
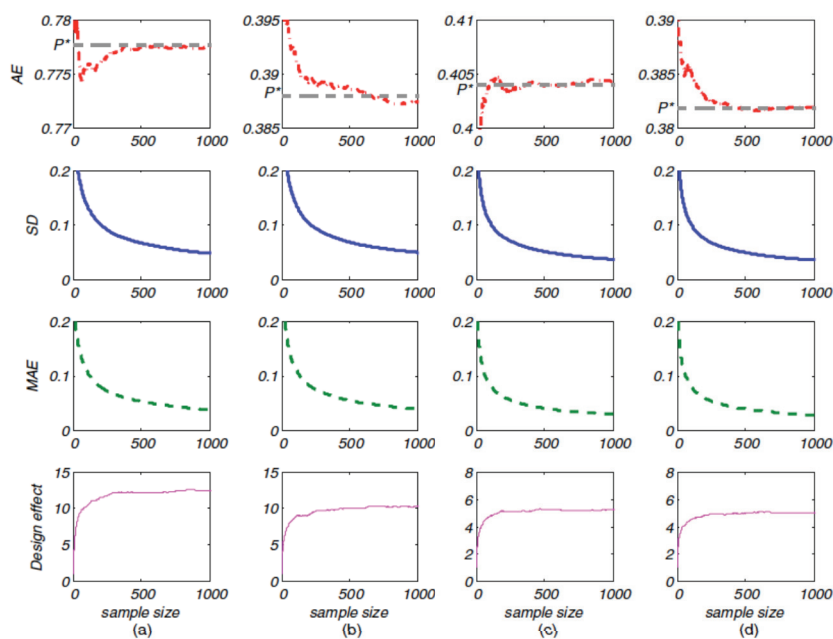
Figure 7. RDSII estimations on the undirected network G1: a) age, b) county, c) civil status and d) profession. Sample size: 500.

Violations of Assumptions

Directed network: For the two variables with high homophily (age and county), the MAE:s of RDSII in the directed network (G2) were at least twice as high as those in the undirected network for age and county. This was not the case for the estimates of profession and civil status, which had low homophily.

Sampling without replacement: Sampling without replacement always had smaller SD and MAE than those for sampling with replacement. This was especially apparent in G1.

Rejecting invitations and forgetting peers: Bias, SD and MAE changed only little when varying P_r and P_i.

Rejecting invitations and forgetting peers depending on member characteristics: When the ignore or reject probabilities depended on the characteristics of the members, the RDS estimates gave very large bias and error. For example, when members who were born before 1980 rejected half of the invitations that were given to them and the members who were born after 1980 did not reject any invitations (P_r and P_i both set to zero), the bias was over 0.3 for age (Figure 8).

Preferential recruitment of close friends: Biases were 0.01, 0.02, 0.04 and 0.03 for the four groups. When we compare the SD and MAE of preferential recruitment with non-preferential recruitment, the former had larger SD and MAE values for all groups.

Effect of seeds and coupons: The number of seeds and coupons had very small effects on bias. Both the SD and the MAE increased when more coupons were used. Increased number of seeds gave low SD and MAE even when the number of coupons was high.
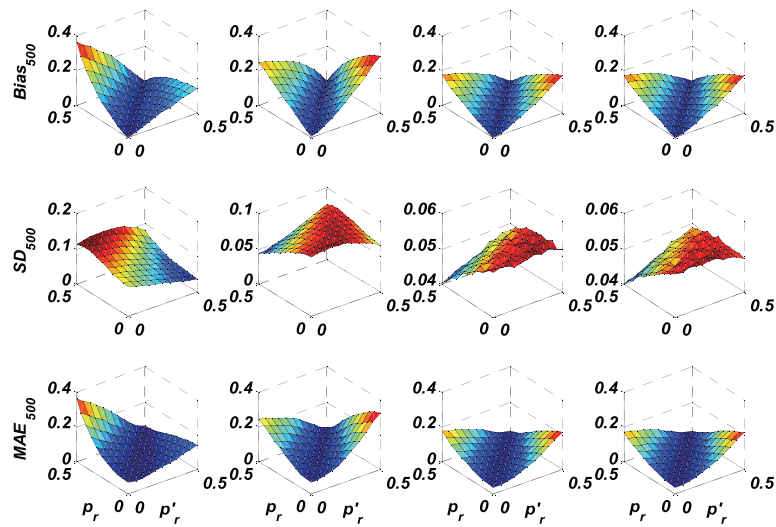

Figure 8: Estimate of age on G1 with ignore and reject probabilities dependent on the individuals' characteristics. Simulations were repeated 10,000 times for each combination; number of seeds: ten; number of coupons: three; sampling with replacement; sample size: 500; Seeds were randomly selected at the beginning of each simulation $P_i=0$ and $P_i'=0$.

40

Implementation of WebRDS *(Paper III&IV)*

Sampling Dynamics

676 and 983 submissions respectively were recorded in the studies of Paper III and IV. The maximum wave lengths were 24 and 29 waves (excluding seed waves). The study in Paper III stopped by itself while the study in Paper IV were stopped by us after recruitment had slowed down considerably.

Equilibrium

Using the standard criteria in the literature [41], equilibrium was reached for all variables in both surveys. In Paper III, equilibrium was reached after a maximum of seven waves and in Paper IV after a maximum of 11 waves. We plotted the study sample compositions over increasing sample sizes. All variables reached satisfactory stabilization in paper IV. In paper III the curve for province never stabilized before the study stopped. Variables in study IV are shown in Figure 9. Similar curves are available for Paper III (see Suppl. Inform.).
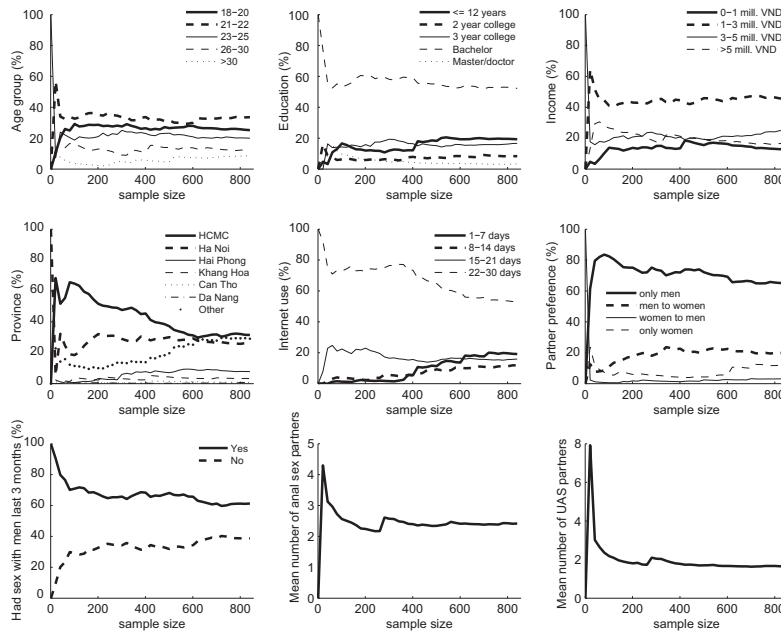
Fig 9: Sample compositions over increasing sample sizes for variables in Paper IV

Characteristics of the Samples

Both samples consisted of young persons with an estimated mean age of 22 and 24 years respectively. The estimated proportion with education of vocational school (post-high school), college or university was 87% and 81% respectively. An estimated 98% and 86% preferred only men or preferred men to women as sexual partners. An estimated 82% and 58% came from the two metropolitan provinces of HCMC and Hanoi. Seventeen and 24% respectively had an income above five million VND (US$240). Thirty-two and 42 provinces respectively were represented in the two surveys.

Contact Patterns and Modeling *(Paper II&IV)*

Sexual relationships described in the in-depth interviews (IDIs) varied greatly in terms of emotional attachment, commitment, trust, relationship ideals, sexual satisfaction and exchange of money or gifts.

The men could not report any terms specific for MSM relationships. Instead they used words that are also used for heterosexual relationships in Vietnam. Among the sexually active, four informants had solely casual sexual relationships where emotional attachment to their partners was limited and the sexual encounters took place at a single or at a limited number of occasions with each partner. For these informants the sexual partnerships could be both concurrent as well as serial and of short duration. One man related:

"I myself am acquainted with many other men. And as I see it, there is the truth that one is dying for something but gets fed up with it very quickly...When I achieve my purpose, I feel bored for some reason. I am 39 years old now and I don't remember how many loves I have had so far in total."

On the other end of this continuum were four informants who reported that they actively sought stable, monogamous relationships with strong emotional attachment and trust. For them, the sex in itself was less important than the emotional attachment and they were not interested in purely casual sexual encounters. Two of these men were quite well-off, articulate and seemed content with their lives. Two other men were manual laborers from the countryside and were less satisfied with their circumstances. One of these men was depressed and socially isolated while the other man lived in a stable relationship in which his partner had recently become unfaithful.

Half the informants' relationship lives were located between these extremes of solely practicing casual sex and on the other hand searching only for stable long-term relationships. These men described their relationship experiences as often being of a few months duration interspersed with casual sexual encounters. Frequently these men had occasional, casual sexual encounters also during the time when they were in a stable relationships. Seven of the participants had at some point in their lives been in a relationships lasting from one to four years.

One 22 year-old student captured the experiences of men who often had short but intense and emotionally attached relationships, saying of one love, *"We were in love for exactly 18 days."* When asked why the relationship was so short, he answered,

"In our world, there are loves which lasts only 1 or 2 hours...Regarding real love, I haven't seen any couples in love for over 2 years in Hanoi. Mostly, they are in love for just about 1 or 2 months."

Many informants who sought, or had previously sought, stable relationships described faithfulness (*lòng chung thủy, sự chung thủy, chung thủy*) as a highly valued quality of a partner, but one which was largely unobtainable. The concept of faithfulness itself seemed however to be used in the same way as among heterosexuals in Vietnam. Many men had experienced their partners being unfaithful, many men had themselves been unfaithful and they knew that many of their friends were unfaithful while in relationships. For several of these informants the search and longing for a long-term committed and monogamous relationship had ended in repeated disappointments through a painful process of unfaithfulness and break-ups. This group of men entered

the MSM world in search of a romantic, emotionally close and stable relationship but eventually either lost their belief in this goal or chose not to pursue it. As one man commented:

"Regarding all the relationships, I no longer hope for long lasting love, no longer dream of long lasting relationship, I just take it as it comes.
Q: In your opinion, what is an ideal relationship in your world like?
A: The most important thing is faithfulness. But it is very difficult to be faithful in this world, it just exists for a certain period, it cannot exist the lifetime."

The men described no clear social norms against having too many or too few sexual partners. One man noted, "*only those who are too ugly and unattractive have no choices, and don't have many lovers*". Informants did not express any strong norms against unfaithfulness in relationships, accepting it instead as the way things were. "*People in our world are very unfaithful. There are many choices so everyone becomes unfaithful. It is common. One is also a bit angry ... but he has to accept.*"

Study IV provided quantitative data on sexual behavior. An estimated 62% (CI: 54-65%) had sex (any type) at least once with another man during the last three months. Some 55% (CI: 50-60%) had anal sex at least once and 36% (CI: 32-42%) had unprotected anal sex at least once during the same period (Figure 10). Among the men in the latter group, only 40% (CI: 32-47%) had ever tested themselves for HIV.

Altogether, the men in the sample reported 2,291 UAS acts (non-RDSII adjusted). Looking at the types of partnerships in which these acts took place, almost all acts (92%; CI: 91-93%) were reported in relationships where the respondent was emotionally involved. In addition, almost half of all UAS acts (45%; CI: 44-49%) involved a respondent who had never tested for HIV (Figure 10).
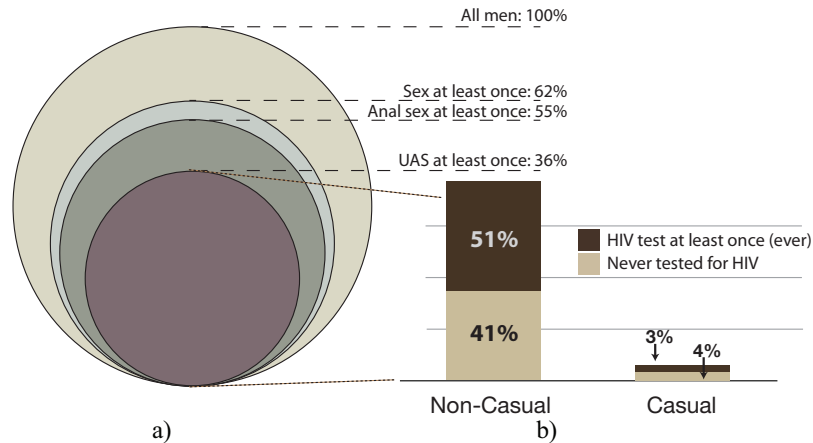

Figure 10: a) Proportions of all MSM practicing different sexual behaviors last three months. Sixty-two percent of all MSM reported having had sex and 36% of all MSM reported having had at least one unprotected anal sex act; b) Among the group of MSM who had unprotected anal sex during the last three months, a total of 2291 UAS acts took place (non-RDSII adjusted). The right side of the figure shows the proportion of these UAS acts that took place with non-casual (92%) and casual partners (7%). Separate numbers are shown for acts in which the respondent had ever tested for HIV vs. tested for HIV at least once. Due to rounding, the four groups in 2b sum to 99%.

The in-depth interviews (Paper II) provide some more contextual data around these quantitate estimates. One example is a 27 year-old respondent from Hanoi who talks about his condom-use, saying:

"Yes, of course. I use 99% of the times I have sex. In my world, we can have sexual relationships very promptly, we can have both sexual and love relationships promptly, so I have to protect myself, I have to use condom. Sometimes I don't use condom when I have sex with people who I feel that I can trust, who I have been in love for some months. I don't need to use it when I trust or love one very much, it is ok to die because we sleep together."

In Paper IV the distribution of UAS partner numbers and number of UAS acts were highly skewed (fig 11). Large differences in the number of lifetime sexual partners was also present in the small in-depth interview sample of 17 men (Paper II). Among these men lifetime partner numbers ranged from one to about 200.
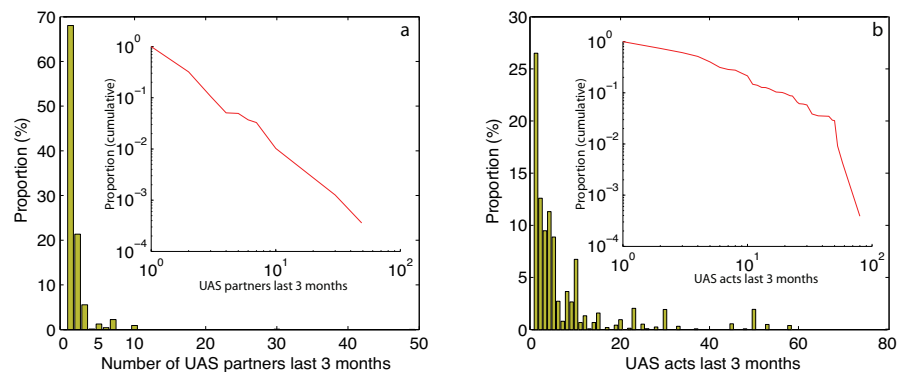

Figure 11: Probability density distributions for men reporting UAS last three months. a) Number of partners last three months; b) total number of UAS acts with the three most recent partners last three months. The few individuals with many partners or many acts are poorly visualized on a linear scale. The insets therefor show corresponding cumulative log-log plots.
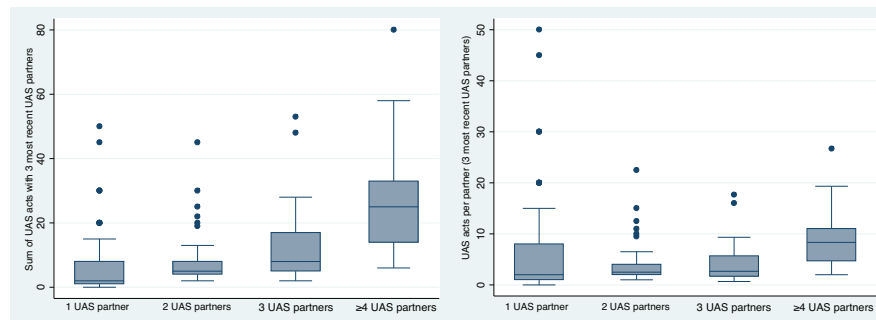

Figure 12: a) Total number of UAS acts (y-axis) last three months stratified on the number of UAS partners reported for the same period (1, 2, 3 and >3 partners); b) Number of UAS acts per partner, stratification and period the same (RDSII-unadjusted).

The sum of reported UAS acts with the three most recent UAS partners increased with increasing number of reported UAS partners (p<0.000, Poisson regression) (Figure 12a). The number of reported UAS acts per partner for the three most recent UAS partners also increased with increasing number of UAS partners (p<0.000) (Fig 12b). The latter result was significant also when we removed participants with eight or more partners. For participants with one to four UAS partners the trend was negative (p=0.013), although the trend for sum of UAS acts with the last three partners remained significantly positive. Note that we collected data on the total number of UAS partners but the number of UAS acts refers to the three most recent UAS partners, both during the last three months.

Using the weighted configuration model, we modeled HIV transmission (figure 13). The results show that both the best and worst-case scenarios (top and bottom green and brown lines, which assumes different numbers of UAS acts for missing data) produced very high $R0^{3M}$. The best-case scenario that very likely underestimates UAS frequency, yielded an $R0^{3M}$ of one already at a per-act transmission probability of 3.3%. Assuming a per act transmission probability of five percent (i.e. a modest assumption for UAS and primary infection transmission) yielded, in the worst-case scenario, an $R0^{3M}$ of nine for the three-month period.

We then investigated the contribution to transmission stemming from the observed positive correlation between partner numbers and number of UAS acts per partner. We did this by shuffling the UAS frequencies and then estimating $R0^{3M}$. The results showed that the high $R0^{3M}$ remains after shuffling and is thus not an artifact of the weighted configuration model. Additionally, $R0^{3M}$ is indeed higher in the scenario with higher number of imputed acts but the relative increase compared to the non-shuffled data is much less pronounced. Thus, most of the very large increase in $R0^{3M}$ that we observed when we moved from the best to the worst case scenario (green filled line to brown filled line), is due to the formation of a network structure that greatly facilitates transmission.
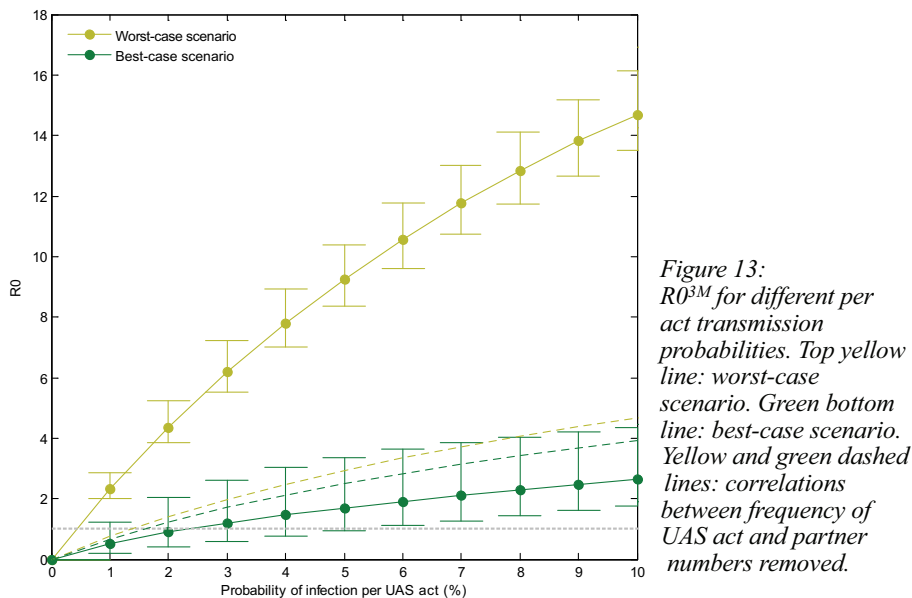

Figure 13: $R0^{3M}$ for different per act transmission probabilities. Top yellow line: worst-case scenario. Green bottom line: best-case scenario. Yellow and green dashed lines: correlations between frequency of UAS act and partner numbers removed.

DISCUSSION

Performance of the RDSII Estimator *(Paper I)*

Key Findings and Contributions

When all the assumptions underlying the estimator are fulfilled, the RDSII estimates were asymptotically unbiased but had high design effects, especially for the two groups with high homophily. Directed edges increased the mean absolute error (MAE) for the two variables with high homophily. Sampling without replacement, which is used in practice, gave a lower design effect, standard deviation (SD) and MAE compared to sampling with replacement. The RDSII estimator showed strong resistance to bias when probabilities that individuals ignored contacts and rejected invitations were independent of the characteristic that were estimated. However, if these probabilities were made dependent on the individuals' characteristics the bias and MAE became very large. As participants in RDS studies are rewarded for successful recruitments, rational and self-interested participants would ignore contacts who are considered to be less likely to accept invitations. Simulations show that such a combination of a group having a high probability of rejecting invitations and a high probability of being ignored can give rise to very large bias and error. When recruitment choices are weighted by the frequency with which people have communicated with each other, the bias, SD, MAE and design effect all increased. Both the SD and the MAE increased when more coupons were used. On a positive note, when more than ten seeds were selected randomly (either through simple random sampling or proportionally to degree), SD decreased to very acceptable levels. Presumably this way of selecting seeds allowed for diverse starting points in parts of the network that were poorly connected to each other.

Comparison to Previous Studies

We showed that non-random recruitment can lead to large bias. The extent to which recruitment is on average non-random in real-life RDS studies needs empirical study. A recent study [54] assessed the random recruitment assumption by comparing participants' reported social networks with their actual recruitments and found that recruitment was random with respect to age, marital status, and drug use mode, but not to gender and education levels. Deviations were however not very large for the latter. Wang et al did not find statistically significant non-random recruitment for gender, age and ethnicity in a study among NMDA users [53]. Wejnert on the other hand found non-random recruitment for gender, age and education in his studies but bias was modest [55, 56]. In a paper by me and Dr. Anna Thorson in AIDS, 2010 [26] (not part of this thesis), we discuss another potentially serious deviation from the random recruitment assumption, namely that short-term sex partners are under-recruited, potentially leading to serious underestimation of partner numbers. To summarize, studies investigating the random recruitment assumption have so far been few in numbers and results have been mixed. A large number of further studies is needed.

Variance was also a concern. Goel and Salganic [45] examines artificially generated networks and shows that if a network has poorly connected parts, this can generate very large design effects. The extent to which real networks contain such bottlenecks remained an open question. Our results indicated that our network indeed contained some such poorly connected parts, but also that a relatively modest number of well-placed seeds could overcome these bottlenecks. Goel and Salganic has also assessed design effects

based on empirical network data. One of these data sets came from a population with a high proportion of IDUs and sex workers and one other data set consisted of friendship networks at middle- and high-schools, both in the U.S [58]. They found very high design effects (median design effects eleven and five respectively) despite using ten seeds, which in our study drastically reduced variance. However, their simulations also include limitations. First, they use sampling with replacement, which is not used in practice and which we showed increased variance [131, 132]. Second, both their data sources are likely to include a large amount of missing data. In the IDU/Sex worker network it is highly likely that many social connections were not reported by the participants [133] and in the high school network each participant could report a maximum of five friends. Overall, it is thus possible that very close friendships were preferentially reported. This would make recruitment chains less likely to connect to new parts of the network and variance could have been driven up.

Another form of data on real-life variance includes data from repeated cross-sectional surveys that have been performed under similar conditions in the same population. A study reporting results from three consecutive RDS studies among MSM in Beijing provides such interesting data [134]. This Beijing study reports two study variables to be very similar in the three surveys while substantially differences existed for three other variables when comparing the three surveys. A comprehensive review of such cross-sectional surveys has not been performed. I return to this issue in the next section where I compare the estimates in two cross sectional webRDS studies in this thesis (Paper III and IV), which contained partly identical questions.

Limitations

The deviations from the assumptions that we simulated in Paper I can be modeled in different ways, which affect the conclusions.We have opted for deviations that we consider relevant to RDS studies in different contexts, but the simulations still represent subjective choices and do not cover all situations relevant to real life RDS studies. Moreover, although we have tried to make results more generalizable by varying the properties of the original network, the network characteristics that we actually picked for simulations do not reflect all types of networks, which could impact on the interpretations of the results.

Summary and implications

We have shown the effects of a large number of deviations from assumptions. We showed that the bias, MAE, SD and design effects were all affected by a large number of parameters. In further work, it would be valuable to run simulations of RDS studies with realistic combinations of violations taken from real-life studies as well as running simulations on other complete real-life networks from diverse settings.

Non-random recruitment gave the most serious bias in these simulations. Real-life studies evaluating this assumption have been few and the results have been mixed [53-57]. When possible, RDS studies should routinely compare participants' reported network composition with actual recruitments to provide further empirical evidence on this assumption from a wide variety of contexts.

Diversity of seeds sharply reduced variance. We recommend researchers who implement RDS studies to carefully select diverse sets of seeds, especially with regard to variables that can be assumed to have high homophily.

Implementation of WebRDS *(Paper III&IV)*

Key Findings and Contributions

We showed for the first time that it was possible to implement an automatic webRDS system to sample men who have sex with men. Recruitment was rapid and robust in both surveys despite questions on a range of sensitive topics. We evaluated the independence of the starting participants from the sample by showing that sample composition stabilized well for all variables, except home province in the first survey.

The webRDS Recruitment System

What Enabled Robust Recruitment?

To achieve sustained recruitment, each participant needs to recruit on average at least one other participant. While the addition of a specific incentive or feature on the site may be what suddenly makes the recruitment roll, the last addition does not necessarily need to be the most important part of the package. It is also important to remember that surveys with close to one recruitment per participant can die out early, solely due to random fluctuations in recruitment.

The first two versions of the webRDS system gave insufficient recruitment. The great leap forward happened in the third launch. In this version of the system, we made the site more professional looking, while retaining a friendly and warm appearance, and we included a varied set of incentives. Perhaps the most important difference between the second and third version was the addition of mobile phone air-time given to each participant. Most people are risk averse (or consider money to have a decreasing marginal utility) [135] and a small incentive, such as air-time, that the person is sure to receive, may be more attractive than a lottery with a large price. The downside of giving out airtime to people's phones is that the survey may attract participants who try to cheat the system by participating more than once or the survey may attract non-MSM persons. I discuss these issues further in the next section. Incentives other than air-time was probably also important to get strong recruitment, as evidenced by recruitment taking place over several waves also in the first two launches, although these recruitment chains subsequently died out. In particular, the possibility for participants to compare their own answers to the answers of others was very positively perceived by MSM testing the survey in Hanoi. Several questions were included only for this purpose. Additionally, we emphasized extensively the respondent's positive contribution to the situation of MSM in Vietnam, although we do not know how much this contributed to recruitment. It was also possible for participants to give away their price to a charity of their choice, which they could chose from a list. Only a handful of participants did so, although it is possible that the survey was nevertheless better perceived by users through making this option available.

How Representative were the Samples?

There is no gold standard with which our estimates can be compared as we do not know the true composition of Internet-using MSM in Vietnam. However, comparing national statistics and other published research data to our estimates show interesting similarities and dissimilarities that may reflect sampling bias, variability between data collection instruments, variance, and systematic differences between the sexually active Internet-using MSM population and the general population. Using the RDSII estimator, 97% of the Internet-using MSM population in the first study was estimated to be below 30 years of age and the sample mean and median ages were 22 years. In the second survey, the mean age was 24 years. By comparison, 43% of the adult male population in Vietnam is between 18 and 29 [99]. The lower mean age of sampled MSM compared to the national

age distribution for men is however consistent with an offline RDS study of MSM in Khanh Hoa, Vietnam, which reported a median ages of 24 years [121] and an RDS in Hanoi with median age of 20-24 years [110]. One online survey among visitors to Vietnamese MSM websites has been published and had a median age of 23 years with 18% stating an age above 30 years [110]. Income distribution (Figure 4b, Paper III) is broadly consistent with the national average monthly per capita income for urban areas (2,130,000 VND, 2010 [136]). It is also comparable to data from the online survey among visitors to Vietnamese MSM websites [110] and an offline RDS in Hanoi 2008 [110], although inflation, economic growth and differential categorization of income levels precludes an exact comparison. An estimated 88% in the first survey had some type of post-secondary education, including vocational training and the corresponding estimate in the second survey was 81%. This can be compared with 68% in the offline RDS in Hanoi [110] and 79% in the survey among visitors to Vietnamese MSM websites [110]. The sample in the first survey was heavily concentrated to the two large metropolitan areas of Ho Chi Minh City (HCMC) and Hanoi, with a population estimate of 84% for these cities combined, while the estimate in the second survey was 51%. The most likely explanation for this difference between the two estimates is that equilibrium, as evidenced by fig 3, Paper III, was not reached for this variable in the first survey. HCMC and Hanoi constitute approximately 55% of the urban population in Vietnam and about 16% of the national population [93, 137]. In the results of an online banner survey on Vietnamese MSM websites 74% lived in Hanoi or HCMC [110]. Potential explanation for the observed differences compared with national statistics may include migration of young MSM to large cities, urban-rural differences in prevalence of male-male sex and different levels of access to the Internet.

We did not find evidence of the men's social networks forming geographically isolated groups, which otherwise would have been a source of bias. The recruitment chains in our sample frequently crossed over between provinces. In total, 30% of all recruitment events took place between persons in different provinces. Additionally, like other social networks, MSM networks in Vietnam are most likely small-world networks [138], with short numbers of steps between provinces. However, equilibrium was attained much more slowly for province than for other variables and the results from the second survey are likely more accurate than those from the first. This highlights the need to select seeds with particular attention to diverse geographic distribution.

One percent in the first survey stated that they preferred only women or preferred women to men as sexual partners. In the second survey the proportion was 15%. The banner survey on MSM sites [110] and an offline RDS in Hanoi with similar question [110] recorded 15% and 1.9% respectively for the same responses. A middle option ("Prefer women and men equally") was available in these studies in contrast to our study, with 14% and 8% of answers respectively.

It is important to consider wether the participants truly were MSM. This is also an issue for standard RDS surveys and to some extent also for other sampling strategies of MSM, especially if they use material incentives. We are confident that the seeds belonged to the population. Given the design and information of the site and our piloting work with the system it seems very likely that MSM participants in general did not invite non-MSM persons on a massive scale. However, given that the known number of recruitment events was more than 1,500, recruitment of non-MSM could be assumed to have occurred at several occasions. The important question then is wether the survey would have caught on among these non-MSM in a way that would seriously have biased the results.

Although, there is no firm evidence on this issue it seems plausible that the survey would on average appeal to MSM much more than non-MSM. Male-male sex is heavily stigmatized and many men in Vietnam would not like to be associated with a same-sex

web site. The aggregate answers that were shown after each participant had answered the questionnaire was likely also considerably more interesting to MSM than non-MSM. Thus, given that the average number of recruitments per participant in the surveys stayed relatively close to one, it seems plausible that the same value for non MSM would have been below one. This would further have meant that recruitment among non-MSM, when taking place, would potentially have died out relatively quickly. Second, although highly anecdotal, we asked around among our test persons wether they had heard about misuse of the survey but we did not get any such reports.

Third, it is plausible that non-MSM participants would have mostly have participated in order to acquire the air time incentive. If they participated once to acquire this, the bias in the results should be low. If they participated multiple times they would have had a good chance of being flagged during our data checking procedures. Fourth, we can compare the compositions of answers in the two surveys. Several estimates were very similar in the two webRDS surveys. It appears plausible that if the surveys would contain large clusters of non-MSM participants, these clusters would appear in different parts of the network leading to divergent results in the two surveys. Finally, some of the core findings in the second webRDS survey (Paper IV) show important similarities with findings from other MSM studies, i.e. the right-skewed frequency distributions of number of UAS acts and number of UAS partners [87-89, 130].

This theoretical reasoning and the available empirical data available thus points towards that the results were not heavily biased by non-MSM participants. To establish webRDS as a valid and robust method over the long term, research on the frequency and detection of fraudulent participation should be high on the agenda. I return to this issue below when discussing areas of further development for webRDS.

To conclude this section on the representativeness of webRDS sampling of Internet-using MSM, our two surveys succeeded in reaching a varied sample of largely young men concentrated in the two major cities of Vietnam, with an education higher than the average for the country. The extent to which these results indicate sampling bias, or reflect differences between the general Vietnamese population and the Internet-using MSM population is difficult to assess, although circumstantial evidence and theoretical reasoning supports a relatively positive picture.

Variance and Design Effects
As evidenced by the results of Paper I, the size of the variance is obviously extremely important. What indications on the size of variance can we see in the results of the two surveys?

First, we can see that most variables reached equilibrium very fast. This empirical result points to low homophily for most variables and consequently supports a relatively low variance. For a few variables the approach to equilibrium was however slow. This was most apparent for the variable "province" which in the first study never seemed to reach a stable level. In the second survey (Paper IV) the province variable did however reach such a stable level. As a side issue, it is interesting to see that when the survey in Paper III stopped, the variable Province was actually moving towards the steady state level that was achieved with a large sample size in Paper III and it thus seems plausible that a larger sample already in Paper III would have resulted in similar estimates as in Paper IV.

Second, for variables that were included at both occasions we can compare the results of the two consecutive surveys. Among five easily comparable variables three came close to each other in the two surveys (age, education and income). Two were more different (province and sexual preferences). Province is however slightly unfair to compare

between the surveys given that this variable did not achieve equilibrium in the first survey. In summary, although it is difficult to judge the variance to any precise degree we can at least relatively satisfactorily falsify the hypothesis that variance and design effects were huge and that estimates would effectively have been random.

Comparison to Previous Studies

In comparison to standard RDS studies, these two webRDS samples achieved extremely long, and potentially the longest recruitment trees ever recorded in RDS studies, which in general should serve to reduce bias and variance.

Web-RDS will in most cases entail a lower costs than a standard RDS study. Cost for monetary incentives were in our study on average 5.9 USD per participant in the cleaned sample (3,353 USD in total). Staff hours to interact with seeds, deliver incentives, monitoring invalid submissions etc, totaled one month full-time equivalents (FTE). Adjustment of the site to appeal to the local target group is technically easy but requires formative research. For comparison, an offline RDS would have shared similar costs for incentives and formative research (see e.g. [139]) but would also require a survey office and at least five months staffing (conservative FTE estimate).

Current online recruitment of hidden groups is based on self-selected samples of persons who access certain Internet sites and click banner ads for a study. These surveys often have participation rates of a few in a thousand to a few in a hundred [140, 141]. As compared with such online samples, these webRDS surveys are likely to have achieve considerably reduced self-selection bias because sustained recruitment is likely to be dependent on high participation rates.

Limitations of WebRDS

In addition to the issues discussed so far a number of limitations should be noted. In Paper III we excluded 13% (n=85) of the submissions because of duplicated personal information or an age below 18 years and in study. In Paper IV the corresponding figure was 8.7%. While this shows that the recruitment system did not work perfectly, it also shows, on a positive note, the potential for eliminating duplicate submissions. Out of the completed surveys in Paper III 17.5% included an IP number that had previously been registered in the system, which may signal duplicated submissions. However, it is important to note that IP-numbers are shared by all users at Internet cafés. Further, we opted for removal of ineligible participants after the study was concluded in order to observe recruitment behavior without outside involvement and to avoid running the risk of inadvertently stopping the survey. This procedure should not produce bias in the RDSII estimation if removal of ineligibles is made in a correct way. However, there will surely always remain questions as to the extent to which ineligibles have been fully removed or not. If future studies show that ineligibles can safely be removed without stopping global recruitment this is preferable.

The network size question did not exclude MSM living outside Vietnam and those of ages under 18 years. Potentially this may have underestimated e.g. the proportion of young persons.

Eight percent of participants were recruited by a stranger. Other RDS studies among MSM have recorded similar proportions [57, 134, 142]. We do not think this caused serious bias in this study (see e.g [143]) but the issue should be monitored in future webRDS studies. Although this study aimed to sample Internet-using MSM, access to the Internet, including literacy, will always be a limiting factor for representative sampling of MSM online. Bio-markers will obviously also be challenging to collect with webRDS.

Further Development of WebRDS

Many areas of development exist for webRDS among MSM and other hidden groups. Some key questions include the following:

1. What incentives work best for different groups while not encouraging participation of fraudulent participants?
2. Can a successful webRDS system based purely on non-material incentives be successfully implemented?
   Currently we collaborate with Mart Stein and Mirjam Kretzschmar at Utrecht University, who apply our system for studies of contact patterns for the spread of influenza and other communicable diseases. Mart has in pilot studies been successful in enabling recruitment among young people without any material incentives, which if sustainable would be a great leap forward.
3. What algorithms, control questions and monitoring systems work best for eliminating fraudulent or non-serious participants?
4. Presently RDS includes adjustment only for network size (and we kept to this tradition). Potentially future adjustment in webRDS may want to include adjustment based on frequency of internet-use, especially if recruitment is very fast.
5. Randomizing the design and questions on the site during a survey can be used to determine well-functioning features as well as provide empirical evaluation of the RDS assumptions.
6. Performing RDS for groups with known characteristics will be important to investigate bias and variance.
7. Developing RDS for mobile phones, preferably for the cheapest types, will enable broad participation in LMICs.

Summary

We developed a webRDS system to sample men who have sex with men in Vietnam and showed that is was possible to survey participants on a range of sensitive issues, including sexual behavior, while sustaining recruitment and achieving equilibrium. The results indicate that the method could potentially be implemented at low cost among Internet-using MSM. More research is needed to validate and develop the method in diverse settings. Care should be taken in future studies to ascertain that seed characteristics should be varied, especially with regard to geographic distribution. With further evaluation and development, webRDS may fill an important gap in the current epidemiological tool box.

Contact patterns and modeling *(Paper II&IV)*

Key Findings and Contributions
More than one third of the MSM population in Paper IV reported at least one episode of unprotected anal sex (UAS) during the last three months. Among these men only 36% had ever tested themselves for HIV. The empirical data showed, to the best of our knowledge, a previously undocumented positive correlation between UAS frequency and UAS partner numbers, which is clearly different from the trends reported in studies of heterosexual persons [79, 144-146]. Our modeling results indicate that this sexual behavioral pattern can fuel extremely rapid HIV transmission.

The study also highlights the significance of collecting data on UAS frequency per partner to better estimate the potential for population level transmission. This may be especially important as UAS acts, depending on the population, may preferentially take place among men with high numbers of UAS partners as in our study. Additionally, our analyses show that UAS frequency per partner may be an important source of information on sexual network structure, which in turn may profoundly affect population level transmission.

The in-depth interviews in paper II, revealed a very diverse landscape of sexual relationships behind these quantitative estimates. Short and unstable relationships were the norm among these men in Hanoi. Most men, who had at least some economic means, could socialize freely within the MSM social networks but few could be open to outsiders. Sexual contacts were easy to establish for most respondents. Faithfulness was highly valued in relationships but largely unobtainable in the long run. Repeated failures in relationships had made several of the men lose hope of establishing long-term relationships and many men were under strong family pressure to marry a woman.

Relationship to other research
High variability in partner numbers and intercourse frequency has been observed in other settings among MSM [87-89, 130]. The absolute number of partners and acts of MSM in Vietnam would be important to compare to studies from other settings. Differences in sampling methodology and our focus specifically on UAS partners and UAS acts, however makes such comparisons difficult. However, our results on partnership numbers seem to be relatively well in line with sexual partner distributions in China [86], which has previously been pointed out to potentially be quite a lot lower than among MSM in western high-income countries [147-153]. A larger set of variables including, condom-use, sero-sorting, UAS frequency and sexual practices, would however be needed to understand the real-life impact from such potential differences in partner numbers. As is clear from the modeling in Paper IV, the level of risk behavior recorded by the Vietnamese study participants is easily enough to fuel rapid HIV spread.

Paper IV showed a highly worrying pattern of high number of UAS acts among men with high number of UAS partners. Relationship between UAS partner numbers and UAS frequency does not seem to be well investigated in previous studies. It seems thus largely unknown to what extent MSM with many partners have more or less unprotected sex acts in total than those with few partners. Since the probability of becoming infected with HIV is likely to be positively correlated to the number of unprotected sex acts a person has with an infected person, this type of correlations may have important consequences for population level transmission.

There is one study in a large representative sample of heterosexual persons on the island of Gotland in Sweden [79] showing that the more sexual partners a woman reported, the

lower her total number of reported intercourses. For men the relationship was similar but the total reported number of intercourses first decreased and then remained stable over increasing partner numbers. A similar study [146] reported data among non-married heterosexuals in a so called high-risk area in San Francisco, U.S. This study used data from the late 1980s and found a slightly increasing or stable number of intercourses with increasing partner numbers, with a clear negative trend for the number of intercourse per partner when partner numbers increased. These studies did however not differentiate between protected and unprotected intercourses. Finally, there are studies among heterosexuals using much more imprecise measures of sexual frequency but nevertheless reporting lower frequency of sex with casual or secondary partners as compared with a primary or stable partner [144, 145].

Studies on the correlation between the number of unprotected anal intercourses and UAS partner numbers can clearly not be extrapolated from studies of heterosexual persons. Rigorous investigation in reasonably representative samples of MSM in different contexts is warranted to investigate wether the pattern observed in Paper IV exists also in other settings.

Relationship Stability and Causes of Short Relationships
One fundamental factor for HIV spread is the rate of partner change. Below a critical level of partner change, HIV will eventually disappear. As we saw in Paper IV, the vast majority of UAS acts took place within non-casual relationships. At the same time many men in the in-depth interviews (Paper II) talked about a high prevalence of short and unstable relationships with frequent sexual concurrency and unfaithfulness. The same theme of unstable male-male relationships also surfaced in my discussions with other homosexual men in Hanoi as well as in studies among urban MSM in Shanghai and Shenzhen, China [154, 155]. Many MSM in Shenzhen described a strong desire to establish long-term committed and loving relationships but felt that sexual relationships were overemphasized in the MSM community at the expense of more loving ones.

Without judging any one type of relationship as in itself preferable to other types of relationships, it is possible that failure to establish long term relationships among MSM who wants them, increases the frequency of partner change and leads to increased HIV transmissions. There are a number of potential causes of high prevalence of short relationships among MSM in Vietnam.

First, both the formation and the resolution of relationships seemed to carry low costs. Sexual relationship could form quickly and the cost of breaking up was often low. There were no children, houses, mortgages or any extended family that could disapprove. In addition there was no legally binding marriage contracts. This lack of ties was also repeatedly brought up by the interviewees us a reason for relationship instability.

While couples could easily meet, the hidden nature of their relationships plausibly makes relationships hard to sustain. For one, hidden meetings presumably become impractical when everyday chores and long working days takes its toll. Some men also said it was difficult to trust their sexual partners, which seem understandable if one never gets an external validation of one's lover's stories.

Family pressure to marry was very strong. Getting children and preferably at least one male child is highly valued in Vietnamese culture and denying parents this future was not easy for some of the men. It is not hard to imagine that this situation puts a potential male partner in an even more difficult situation. If a man does not even know himself if he will

be able to withstand the pressure to marry a woman, it will be very difficult for the partner to feel safe and committed.

Many interviewees had experienced repeated painful breakups caused by unfaithfulness, leading to loss of faith in long-term relationships. With falling expectations one may hypothesis that the energy these men are likely to invest in the next potential partner is likely to be low. Interestingly, although unfaithfulness seemed to be extremely common there did not seem to be any tendency for these men to accept non-monogamous long-term relationships.

Finally some men in our interviews stated that their partner had been "lured away" from them. On a more speculative note, relationship stability may also be influenced by a type of herd immunity. If a population consists only of a few couples who continuously have to socialize with singles looking for partners, the likelihood of unfaithfulness and breakup may be considerable. On the other hand, if most people have already formed couples, those people in couples will mostly meet other people who are in couples, reducing the probability that relationships breaks up. Thus, along this line of reasoning, an unintended consequence of the formation of each new couple is the reduced risk of breakups that this confer on other couples.

As may be clear from this list of potential causes, I lean more towards explanations focusing on limited opportunities to form stable relationships, rather than limited preferences for having them. The issue is however open for debate (and a lot more empirical data).

Limitations

Paper II
The issues discussed in the interviews were sensitive and although the interviewers experienced a good and open atmosphere, social desirability in the discussions may have been present. No local MSM were present at the interviews. This may have led to misinterpretation of verbal and non-verbal messages. It may however also have led to that the men could speak freely about their lives without needing to conform to prevailing norms or fear that their statements would be passed on to others. The findings were also discussed with local key informants to avoid misinterpretations of the data. The sampling was not probabilistic and the number of informants was limited. The data analysis process relied on a number of subjective choices regarding the most important findings to bring forward to the reader and other researchers may have made other choices.

Paper IV
Although we included a high level of detail regarding the frequency of UAS in study IV, collection of other sexual behavior data would have been useful, including what acts were receptive and insertive. Using the weighted configuration model we assumed a uniform infection probability per UAS act. In reality transmission probabilities are different over the course of the infection and may differ between individuals and in relation to sexual practices [64]. The short time interval assumed in the analyses will however decrease this bias. Forty-eight percent of partnerships had their first UAS act less than three months ago (84% less than a year ago). We have not included this partner change rate in the configuration model. The bias that can arise from a non-dynamic model is however sharply reduced when short time spans are analyzed, as was done in this study. Over the very long time frame it is also possible that the fast transmission in the beginning of the epidemic turns much slower at later stages, as Colgate et al. suggests [81].

55

The model also leads a relatively homogenous network structure, while real networks are likely to contain spatially distinct clusters. Spread into new clusters or new geographic areas combined with sexual behavior patterns such as those reported here can potentially explain observed clustering in HIV sequencing studies [156, 157]. Caution should however be made in generalizing the Vietnamese findings to other settings. Recollection bias will be present but the analyzed variables were limited to the most recent three-month period. Sampling of hidden and vulnerable populations are challenging. The present study excluded 8.7% of the submitted surveys because of suspected duplicated submissions. Comparing results when including and excluding potentially duplicated submissions did not show important differences. The survey sampled mostly young and relatively well-educated MSM in the major cities and should not be generalized to non-Internet using MSM. Low-income and older men is likely to use the Internet less and are likely under sampled. Interestingly, however, significantly less men in this study, who belonged to the lowest income level (<1 million VND/<US$48) as compared to the higher income strata practiced UAS, possibly due to lack of means to meet and establish new relationships [125, 158]. The study may also under sample male sex workers who may not use the Internet and who may not all be part of the social network in which recruitments flowed. However, compared to many self-selected online samples [140, 159], the response rate in this study is likely to have been at least one to two magnitudes higher. In comparison with a standard coupon-based RDS, the present study enabled a much larger geographic coverage and allowed participation of MSM who would not risk to physically access a survey office. Additionally, analyzes focusing on relationships between variables are less sensitive to bias from non-representative sampling than descriptive statistics. Finally, some of the core findings in this study show important similarities with findings from other studies, supporting the validity of the results. These findings include e.g. the right-skewed frequency distributions number of UAS acts and UAS partners among MSM [87-89, 130] (van Griensven et al refers to all types of sex; Crawford et al to UAS with and without condom). It is important to remember that the study population was Internet-using MSM. It is possible that important differences in sexual behavior exists for non-Internet using MSM.

Summary and Implications

The results overall highlights the importance of rapidly scaling up HIV prevention programs for MSM in Vietnam. Access to non-stigmatizing health services and outreach programs, specifically catering to MSM, needs to be implemented on a national scale. This seems to be particularly urgent for MSM outside Ho Chi Minh City. The finding that the overwhelming number of UAS acts takes place with partners whom the respondent is emotionally involved with implies that condom use campaigns may not be sufficiently effective within emotionally close relationships. This echoes modeling results from the Amsterdam Cohort Study showing 86% of new infections in Amsterdam occurring in stable relationships during the early 2000s [160]. The findings of a positive correlation between partner numbers and UAS frequency indicates the need for future research and local prevention programs to better characterize and target high-risk groups within the MSM population as they may contribute to a substantial share of new infections. Socio-economic variables included in our study seemed not be enough as markers of risk. Men reporting more than five UAS acts last three months and also at least two different UAS partners, all preferred only men or men to women as sexual partners. They tended to be somewhat older and have slightly better incomes, although these differences were not statistically significant.

Contact tracing [161] and potentially peer-driven interventions [162-164] may be efficient ways to reach these potentially highly connected individuals, but should be evaluated in the local context. Pre-exposure prophylaxis (PrEP) and treatment-as-prevention strategies have been the focus of much recent discussion [165-168] and may have a role if very

high-risk individuals can be selectively reached and become adherent. However, our data may indicate a high proportion infected through primary transmission and shows low testing rates in addition to other challenges, such as cost and the recent discussions about real-world impact of PrEP [167, 168]. Frequent Internet-users reported more UAS than infrequent users, potentially indicating a window of opportunity for internet-based prevention strategies. Outreach- and harm-reduction activities including stigma reduction need to be prioritized, so that uptake of prevention and testing activities drastically increases. The recent proposal by the Minister of Justice [169] to legalize same-sex marriage is particularly laudable and might also introduce more stability and permanence into some MSM relationships, thus potentially contributing to decreased HIV transmission.

In summary, risk behavior among surveyed MSM in Vietnam was very high, conducive to very rapid HIV transmission and viciously paired with low HIV testing rates among risk takers. The data showed a previously undocumented positive correlation between UAS frequency and UAS partner numbers. The study highlights the need to incorporate UAS frequency data into sexual behavior surveys. Overall, the reported risk behavior is unacceptably high and the results provide support for aggressively scaling-up HIV prevention interventions for MSM in Vietnam.

CONCLUSIONS

- Simulations of RDS studies on a real online social network of homosexual men indicated that design effects in RDS can be very high, but may be possible to reduce drastically by selecting diverse seeds.

- Simulations indicated that sampling without replacement, which is used in practice, gave lower design effect, variance and mean absolute error than sampling with replacement

- Simulations showed that non-random recruitment in RDS can cause severe bias.

- WebRDS implemented among MSM in Vietnam could generate sustained recruitment while achieving equilibrium.

- WebRDS is a very promising method for online sampling of hidden populations.

- In-depth interviews indicated that many MSM in Hanoi perceived relationships with other men to be short and unstable, and that sexual concurrency and unfaithfulness was common.

- The vast majority (an estimated 92%) of all unprotected anal sex acts among surveyed Internet-using MSM in Vietnam took place in non-casual relationships.

- Results showed, to the best of our knowledge, a previously undocumented positive correlation between UAS frequency and UAS partner numbers among Internet-using MSM.

- Overall, risk behavior among surveyed MSM was very high, while HIV testing was very low.

- Modeling indicates that sexual networks of Internet-using MSM in Vietnam, are highly conducive to HIV transmission.

POLICY AND RESEARCH IMPLICATIONS

- RDS researchers and funders should focus more efforts on empirical investigation of the random recruitment assumption.

- Further research is warranted to evaluate and develop webRDS among MSM and other suitable populations.

- Ways to enable sustained recruitment in webRDS without material incentives and methods to improve removal of ineligible participants should be prioritized.

- HIV prevention interventions among MSM in Vietnam should target HIV transmission risks in emotionally close relationships.

- Studies should investigate the proportion of unprotected anal sex acts taking place in casual and non-casual relationships, also among non-internet using MSM.

- Positive correlations between UAS frequency and UAS partner numbers among MSM in Vietnam warrants further investigation in other settings and may imply more rapid HIV transmission among MSM than previously thought.

- HIV prevention interventions, including testing services, needs to be aggressively scaled up for MSM in Vietnam

ACKNOWLEDGEMENTS

REFERENCES

1. Pepin, J., *The origins of AIDS* 2011, Cambridge, UK ; New York: Cambridge University Press. xiv, 293 p.

2. Yusim, K., et al., *Using human immunodeficiency virus type 1 sequences to infer historical features of the acquired immune deficiency syndrome epidemic and human immunodeficiency virus evolution.* Phil. Trans. R. Soc. Lond. B 2001. **356**(1410): p. 855-866.

3. Sharp, M.P., et al., *The origins of acquired immune deficiency syndrome viruses: where and when?* Phil. Trans. R. Soc. Lond. B, 2001. **356**(1410): p. 867-876

4. Mokili, J. and B. Korber, *The spread of HIV in Africa.* J Neurovirol, 2005. **11 Suppl 1**: p. 66-75.

5. Weiss, R.A., *Natural and iatrogenic factors in human immunodeficiency virus transmission.* Philos Trans R Soc Lond B Biol Sci, 2001. **356**(1410): p. 947-53.

6. UNAIDS, *UNAIDS report on the global AIDS epidemic 2010*, 2010: Geneva.

7. UNAIDS, *UNAIDS World AIDS Day Report*, 2011: Geneva.

8. U.S. Census Bureau International Data Base. *World POPClock Projection*. 2012; Available from: http://www.census.gov/population/popclockworld.html.

9. UNAIDS. *AIDS Info*. 2012; Available from: http://www.unaids.org/en/dataanalysis/datatools/aidsinfo/.

10. National AIDS Commission Indonesia, *Country report on the follow-up to the Declaration of Commitment on HIV/AIDS*, 2008: Jakarta.

11. Jenkins, C., *Male Sexuality, Diversity and Culture: Implications for HIV Prevention and Care*, 2004.

12. Hall, H.I., et al., *Estimation of HIV incidence in the United States.* JAMA, 2008. **300**(5): p. 520-9.

13. Smittskyddsinstitutet. *Statistik för hivinfektion*. 2012 [cited 2012 15 October]; Available from: http://www.smittskyddsinstitutet.se/statistik/hivinfektion/.

14. Baral, S., et al., *Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review.* PLoS Med, 2007. **4**(12): p. e339.

15. Beyrer, C., et al., *Global epidemiology of HIV infection in men who have sex with men.* Lancet, 2012. **380**(9839): p. 367-377.

16. van Griensven, F., et al., *The global epidemic of HIV infection among men who have sex with men.* Curr Opin HIV AIDS, 2009. **4**(4): p. 300-7.

17. Smith, A.D., et al., *Men who have sex with men and HIV/AIDS in sub-Saharan Africa.* Lancet, 2009. **374**(9687): p. 416-422.

18. Chapman J Fau - Koleros, A., et al., *High HIV risk behavior among men who have sex with men in Kigali, Rwanda: making the case for supportive prevention policy.* (1360-0451 (Electronic)).

19. Fay, H., et al., *Stigma, Health Care Access, and HIV Knowledge Among Men Who Have Sex With Men in Malawi, Namibia, and Botswana.* Aids and Behavior, 2011. **15**(6): p. 1088-1097.

20. Mumtaz, G., et al., *Are HIV Epidemics among Men Who Have Sex with Men Emerging in the Middle East and North Africa?: A Systematic Review and Data Synthesis.* PLoS Med, 2011. **8**(8): p. e1000444.

21. Baral, S., et al., *A cross-sectional assessment of population demographics, HIV risks and human rights contexts among men who have sex with men in Lesotho.* J Int AIDS Soc, 2011. **14**: p. 36.

22. Beyrer, C., et al., *Bisexual concurrency, bisexual partnerships, and HIV among Southern African men who have sex with men.* Sex Transm Infect, 2010. **86**(4): p. 323-7.

23. WHO UNICEF UNAIDS, *Global HIV/AIDS response: epidemic update and health sector progress towards universal access, in Progress Report 2011*, 2011: Geneva, Switzerland.

24. van Griensven, F. and J.W.D. van Wijngaarden, *A review of the epidemiology of HIV infection and prevention responses among MSM in Asia.* Aids, 2010. **24**: p. S30-S40.

25. The Commission on AIDS in Asia, *Redefining AIDS in Asia. Crafting an Effective Response.*, 2008: New Delhi.

26. Bengtsson, L. and A. Thorson, *Global HIV surveillance among MSM: is risk behavior seriously underestimated?* AIDS, 2010. **24**(15): p. 2301-3.

27. Magnani, R., et al., *Review of sampling hard-to-reach and hidden populations for HIV surveillance.* AIDS, 2005. **19 Suppl 2**: p. S67-72.

28. World Value Survey. 2012 [cited 2012 23 October]; Available from: http://www.worldvaluessurvey.org/index_html.

29. Baral, S., et al., *HIV prevalence, risks for HIV infection, and human rights among men who have sex with men (MSM) in Malawi, Namibia, and Botswana.* PLoS One, 2009. **4**(3): p. e4997.

30. Thomas, B., et al., *The influence of stigma on HIV risk behavior among men who have sex with men in Chennai, India.* 2012.

31. Millett, G.A., et al., *Common roots: a contextual review of HIV epidemics in black men who have sex with men across the African diaspora.* The Lancet, 2012.

32. Berry, M., et al., *Social environment and HIV risk among MSM in Hanoi and Thai Nguyen.* 2012.

33. Wikipedia. *LGBT rights by country or territory.* 2012 [cited 2012 27 October]; Available from: http://en.wikipedia.org/wiki/LGBT_rights_by_country.

34. Semugoma, P., S. Nemande, and S.D. Baral, *The irony of homophobia in Africa.* Lancet, 2012. **380**(9839): p. 312-4.

35. Caceres, C.F., et al., *Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003-2007 estimates.* Sex Transm Infect, 2008. **84 Suppl 1**: p. i49-i56.

36. Deaux, E. and J.W. Callaghan, *Key informant versus self-report estimates of health behavior.* Evaluation Rev, 1985. **9**: p. 365-368.

37. Watters, J.K. and P. Biernacki, *Targeted Sampling: Options for the Study of Hidden Populations.* Social Problems, 1989. **36**(4): p. 416-430.

38. Erickson, B.H., *Some Problems of Inference from Chain Data.* Sociological Methodology, 1979. **10**: p. 276-302.

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

40. Heckathorn, D.D., *Respondent-driven sampling: a new approach to the study of hidden populations.* Social problems, 1997. **44**(2): p. 174-99.

41. Heckathorn, D., et al., *Extensions of Respondent-Driven Sampling: A New Approach to the Study of Injection Drug Users Aged 18–25.* Aids and Behavior, 2002. **6**(1): p. 55-67.

42. Salganik, M.J. and D.D. Heckathorn, *Sampling and estimation in hidden populations using respondent-driven sampling*, in *Sociological Methodology, Vol 34* 2004, Wiley-V C H Verlag Gmbh: Weinheim. p. 193-239.

43. Volz, E. and D.D. Heckathorn, *Probability Based Estimation Theory for Respondent Driven Sampling.* Journal of Official Statistics, 2008. **24**(1): p. 79-97.

44. Kemeny, J.G. and L.J. Snell, *Finite Markov chains* 1960: Van Nostrand.

45. Goel, S. and M.J. Salganik, *Respondent-driven sampling as Markov chain Monte Carlo.* Statistics in Medicine, 2009. **28**(17): p. 2202-2229.

46. Lu, X., et al., *Respondent-driven Sampling on Directed Networks.* arXiv:1201.1927v2, 2012.

47. South, J.S., *Friendship Networks of Mobile Adolescents.* Social Forces, (2004) . **83** (1): p. 315-350. .

48. Wallace, W.L., *Student culture: Social structure and continuity in a liberal arts college,* 1966, Chicago: Aldine Publishing Company.

49. Feld, S.L. and W.C. Carter, *Detecting measurement bias in respondent reports of personal networks.* Social Networks, 2002. **24**(4): p. 365-383.

50. Marsden, P.V., *Recent Developments in Network Measurement*, in *Models and Methods in Social Network Analysis*, P.J. Carrington, John Scott, and S. Wasserman, Editors. 2005, Cambridge University Press: New York. p. 8-30.

51. Johnston, L.G., et al., *Implementation challenges to using respondent-driven sampling methodology for HIV biological and behavioral surveillance: field experiences in international settings.* AIDS Behav, 2008. **12**(4 Suppl): p. S131-41.

52. Malekinejad, M., et al., *Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: A systematic review.* Aids and Behavior, 2008. **12**(4): p. S105-S130.

53. Wang, J., et al., *Respondent-driven sampling to recruit MDMA users: a methodological assessment.* Drug Alcohol Depend, 2005. **78**(2): p. 147-57.

54. Liu, H., et al., *Assessment of Random Recruitment Assumption in Respondent-Driven Sampling in Egocentric Network Data.* Social Networking, 2012 Published Online October 2012.

55. Wejnert, C., *An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and out-of-Equilibrium Data.* Sociol Methodol, 2009. **39**(1): p. 73-116.

56. Wejnert, C. and D.D. Heckathorn, *Web-based network sampling - Efficiency and efficacy of respondent-driven sampling for online research.* Sociological Methods & Research, 2008. **37**(1): p. 105-134.

57. Frost, S.D., et al., *Respondent-driven sampling of injection drug users in two U.S.-Mexico border cities: recruitment dynamics and impact on estimates of HIV and syphilis prevalence.* J Urban Health, 2006. **83**(6 Suppl): p. i83-97.

58. Goel, S. and M.J. Salganik, *Assessing respondent-driven sampling.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(15): p. 6743-6747.

59. Wejnert, C., et al., *Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States.* AIDS Behav, 2012. **16**(4): p. 797-806.

60. Bauermeister, J.A., et al., *Innovative Recruitment Using Online Networks: Lessons Learned From an Online Study of Alcohol and Other Drug Use Utilizing a Web-Based, Respondent-Driven Sampling (webRDS) Strategy.* J Stud Alcohol Drugs, 2012. **73**(5): p. 834-8.

61. Liljeros, F., *Human sexual networks*, in *Encyclopedia of Complexity and System Science* 2009, Springer Science and Business Media.

62. Fraser, C., et al., *Variation in HIV-1 set-point viral load: Epidemiological analysis and an evolutionary hypothesis.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(44): p. 17441-17446.

63. Powers, K.A., et al., *Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis.* Lancet Infectious Diseases, 2008. **8**(9): p. 553-563.

64. Baggaley, R.F., R.G. White, and M.C. Boily, *HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention.* Int J Epidemiol, 2010. **39**(4): p. 1048-63.

65. Boily, M.C., et al., *Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies.* Lancet Infect Dis, 2009. **9**(2): p. 118-29.

66. Auvert, B., et al., *Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial.* PLoS Med, 2005. **2**(11): p. e298.

67. Bailey, R.C., et al., *Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial.* Lancet, 2007. **369**(9562): p. 643-56.

68. Gray, R.H., et al., *Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial.* Lancet, 2007. **369**(9562): p. 657-66.

69. Pilcher, C.D., et al., *Brief but efficient: acute HIV infection and the sexual transmission of HIV.* J Infect Dis, 2004. **189**(10): p. 1785-92.

70. Leynaert, B., A.M. Downs, and I. de Vincenzi, *Heterosexual transmission of human immunodeficiency virus: variability of infectivity throughout the course of infection. European Study Group on Heterosexual Transmission of HIV.* Am J Epidemiol, 1998. **148**(1): p. 88-96.

71. Jin, F., et al., *Anal sexually transmitted infections and risk of HIV infection in homosexual men.* J Acquir Immune Defic Syndr, 2010. **53**(1): p. 144-9.

72. Vittinghoff, E., et al., *Per-contact risk of human immunodeficiency virus transmission between male sexual partners.* Am J Epidemiol, 1999. **150**(3): p. 306-11.

73. Jin, F., et al., *Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART.* AIDS, 2010. **24**(6): p. 907-13.

74. DeGruttola, V., et al., *Infectiousness of HIV between male homosexual partners.* J Clin Epidemiol, 1989. **42**(9): p. 849-56.

75. UNAIDS Reference Group on Estimates, M., and Projections,, *Improving parameter estimation, projection methods, uncertainty estimation, and epidemic classification. Report of a meeting of the UNAIDS Reference Group on Estimates, Modelling, and Projections, Prague, Czech Republic, 29 Nov—1 Dec*, 2006.

76. Anderson, R.M. and R.M. May, *Infectious diseases of humans : dynamics and control.* Oxford science publications. 1991: Oxford University Press.

77. Liljeros, F., et al., *The web of human sexual contacts.* Nature, 2001. **411**(6840): p. 907-908.

78. Johnson, A.M., et al., *Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours.* Lancet, 2001. **358**(9296): p. 1835-42.

79. Nordvik, M.K. and F. Liljeros, *Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections.* Sexually Transmitted Diseases, 2006. **33**(6): p. 342-9.

80. Newman, M.E., *Assortative mixing in networks.* Physical Review Letters, 2002. **89**(20): p. 208701.

81. Colgate, S.A., et al., *Risk Behavior-Based Model of the Cubic Growth of Acquired Immunodeficiency Syndrome in the United-States.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(12): p. 4793-4797.

82. Morris, M. and M. Kretzschmar, *Concurrent Partnerships and Transmission Dynamic in Networks.* Social Networks, 1995. **17**(3-4): p. 299-318.

83. Morris, M. and M. Kretzschmar, *Concurrent partnerships and the spread of HIV.* AIDS, 1997. **11**(5): p. 641-648.

84. Halperin, D.T. and H. Epstein, *Concurrent sexual partnerships help to explain Africa's high HIV prevalence: implications for prevention.* Lancet, 2004. **364**(9428): p. 4-6.

85. Lagarde, E., et al., *Concurrent sexual partnerships and HIV prevalence in five urban communities of sub-Saharan Africa.* AIDS, 2001. **15**(7): p. 877-84.

86. Zhang, L., E.P.F. Chow, and D.P. Wilson, *Men who have sex with men in China have relatively low numbers of sexual partners.* Infectious Disease Reports, 2011. **3**(1): p. e10.

87. Volk, J.E., et al., *Sexual Frequency and Planning Among At-Risk Men Who Have Sex With Men in the United States: Implications for Event-Based Intermittent Pre-Exposure Prophylaxis.* J Acquir Immune Defic Syndr, 2012. **61**(1): p. 112-115.

88. van Griensven, F., et al., *Sex frequency and sex planning among men who have sex with men in Bangkok, Thailand: implications for pre- and post-exposure prophylaxis against HIV infection.* J Int AIDS Soc, 2010. **13**: p. 13.

89. Schneeberger, A., et al., *Scale-free networks and sexually transmitted diseases - A description of observed patterns of sexual contacts in Britain and Zimbabwe.* Sexually Transmitted Diseases, 2004. **31**(6): p. 380-387.

90. Anderson, R.M. and R.M. May, *Infectious diseases of humans : dynamics and control.* Oxford science publications1991, Oxford ; New York: Oxford University Press. viii, 757 p.

91. U.S. Census Bureau's International database. *Demographic Overview - Custom Region - Vietnam.* 2012 [cited 2012 17 August, 2012]; Available from: http://www.census.gov/population/international/data/idb/region.php.

92. Encyclopædia Britannica Online Academic Edition, *Vietnam*, 2012.

93. World Bank, *Open data: Urban Population*, 2010.

94. CIA World Fact Book. *Vietnam.* 2012 [cited 2012 17 August]; Available from: https://www.cia.gov/library/publications/the-world-factbook/geos/vm.html.

95. International Organization for Standardization, *ISO 3166-2*, 2012.

96. IMF. *World Economic Outlook Database.* 2012 [cited 2012 14 November]; Available from: http://www.imf.org/external/pubs/ft/weo/2012/01/weodata/weorept.aspx?pr.x=84&pr.y=9&sy=2009&ey=2012&scsm=1&ssd=1&sort=country&ds=.&br=1&c=582&s=NGDPD%2CNGDPDPC%2CPPPGDP%2CPPPPC%2CLP&grp=0&a=.

97. SCB. *Hushållens ekonomi (HEK).* 2012 [cited 2012 10 November]; Available from: http://www.scb.se/Pages/TableAndChart____163550.aspx.

98. Bank, W. *Literacy rate, adult total (% of people ages 15 and above).* 2012 [cited 2012 10 November]; Available from: http://data.worldbank.org/indicator/SE.ADT.LITR.ZS.

99. U.S. Census Bureau's International database. *Vietnam.* 2012 [cited 2012 17 August, 2012]; Available from: http://www.census.gov/population/international/data/idb/region.php.

100. Asiapop.org, *Estimated population density in Vietnam*, 2012.

101. Nga Viet, *Administrative division of Vietnam*, 2009, Wikimedia Commons,.

102. United Nations. Department of Economic and Social Affairs. Population Division, *World Population Prospects: The 2010 Revision*, 2011.

103. United Nations Statistics Division. *Children under 5 moderately or severely underweight, percentage.* 2012 [cited 2012 14 November]; Available from: http://mdgs.un.org/unsd/mdg/SeriesDetail.aspx?srid=559.

104. Gapminder, *Modified from Gapminder World Map*, 2012.

105. Ngo, D.A., et al., *Male Homosexual Identities, Relationships, and Practices among Young Men Who Have Sex with Men in Vietnam: Implications for Hiv Prevention.* Aids Education and Prevention, 2009. **21**(3): p. 251-265.

106. Vu, M.L., et al., *MSM in Viet Nam- Social Stigma and Consequences*, 2009, SHAPC: Hanoi.

107. Vu, B.N., et al., *Male sexuality in Vietnam: the case of male-to-male sex.* Sexual Health, 2008. **5**(1): p. 83-88.

108. Colby, D., N.H. Cao, and S. Doussantousse, *Men who have sex with men and HIV in Vietnam: A review.* Aids Education and Prevention, 2004. **16**(1): p. 45-54.

109. Internet World Stats. *Vietnam. Internet Usage Stats and Marketing Report.* 2012 [cited 2012 9 September]; Available from: http://www.internetworldstats.com/asia/vn.htm.

110. Nguyen, Q.C., *Sexual risk behaviors of men who have sex with men in Viet Nam*, 2010, North Carolina State University: Chapel Hill.

111. National Committee for AIDS, D.a.P.P.a.C., Vietnam, *Viet Nam AIDS Response Progress Report 2012*, 2012: Hanoi.

112. UNAIDS, U.W., *Measuring Intimate Partner Transmission of HIV in Viet Nam, A data triangulation exercise*, 2012: Ha Noi.

113. Ministry of Health Vietnam, *Results from the HIV/STI Integrated Biological and Behavioral Surveillance (IBBS) in Vietnam 2005 – 2006*, 2006: Hanoi.

114. Ministry of Health, V., *Integrated Biological and Behavioral Surveillance round 2*, 2009: Hanoi.

115. García, C.M., B.S. Meyer, and P. Ward, *Elevated HIV prevalence and risk behaviours among men who have sex with men (MSM) in Vietnam: a systematic review.* BMJ Open, 2012. **2012**(2).

116. Nguyen, T.A., et al., *Prevalence and risk factors associated with HIV infection among men having sex with men in Ho Chi Minh City, Vietnam.* AIDS Behav, 2008. **12**(3): p. 476-82.

117. Giang, L. and M. Clatts, *Men selling sex to other men in Hanoi: findings from an ethno-epidemiological study.* , 2009, UNAIDS: Hanoi.

118. Clatts, M., et al., *Male sex work and HIV risk among young heroin users in Hanoi, Vietnam.* Sexual Health, 2007. **4**: p. 261–267.

119. Colby, D. and M. Mimiaga, *Results of research on male sex workers in Ho Chih Minh City. (Unpublished).*

120. Nguyen, C., *HIV epidemic among men who have sex with men in Ho Chi Minh City. (Unpublished).*

121. Colby, D., T.T. Minh, and T.T. Toan, *Down on the farm: homosexual behaviour, HIV risk and HIV prevalence in rural communities in Khanh Hoa province, Vietnam.* Sex Transm Infect, 2008. **84**(6): p. 439-443.

122. Nguyen, T.A., et al., *Prevalence and risk factors associated with HIV infection among men having sex with men in Ho Chi Minh City, Vietnam.* Aids and Behavior, 2008. **12**(3): p. 476-482.

123. Lowe, D. and P. Thien, *Rapid situation and response assessment of HIV and AIDS programs for men who have sex with men in Vietnam. DRAFT REPORT*, 2010, Vietnam Ministry of Health—Vietnam. Administration for AIDS Control and Prevention (VAAC): Hanoi.

124. Rybski, D., et al., *Scaling laws of human interaction activity.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(31): p. 12640-12645.

125. Bengtsson, L., et al., *Sexual Relationships among Men who have Sex with Men in Urban Vietnam: a Qualitative Interview Study.* Submitted to BMC Public Health, 2012.

126. Salganik, M.J., *Variance estimation, design effects, and sample size calculations for respondent-driven sampling.* Journal of Urban Health-Bulletin of the New York Academy of Medicine, 2006. **83**(6): p. I98-I112.

127. Ramirez-Valles, J., et al., *HIV infection, sexual risk behavior, and substance use among Latino gay and bisexual men and transgender persons.* Am J Public Health, 2008. **98**(6): p. 1036-42.

128. Britton, T., M. Deijfen, and F. Liljeros, *A Weighted Configuration Model and Inhomogeneous Epidemics.* Journal of Statistical Physics, 2011. **145**(5): p. 1368-1384.

129. Diekmann, O., J.A. Heesterbeek, and J.A. Metz, *On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations.* J Math Biol, 1990. **28**(4): p. 365-82.

130. Crawford, J.M., et al., *Number of risk acts by relationship status and partner serostatus: Findings from the HIM cohort of homosexually active men in Sydney, Australia.* AIDS Behav, 2006. **10**(3): p. 325-31.

131. Lu, X., et al., *The sensitivity of respondent-driven sampling.* Journal of the Royal Statistical Society Series a-Statistics in Society, 2012. **175**: p. 191-216.

132. Gile, K.J. and M.S. Handcock, *Respondent-driven sampling: An assessment of current methodology.* Sociological Methodology, 2010: p. no. doi: 10.1111/j.1467-9531.2010.01223.x.

133. Klovdahl, A.S., et al., *Social networks and infectious disease: the Colorado Springs Study.* Soc Sci Med, 1994. **38**(1): p. 79-88.

134. Ma, X.Y., et al., *Trends in prevalence of HIV, syphilis, hepatitis C, hepatitis B, and sexual risk behavior among men who have sex with men - Results of 3 consecutive respondent-driven sampling surveys in Beijing, 2004 through 2006.* Jaids-Journal of Acquired Immune Deficiency Syndromes, 2007. **45**(5): p. 581-587.

135. Elster, J., *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*2007: Cambridge University Press.

136. General Statistics Office of Vietnam, *Result of the Vietnam Household living standards survey 2010*, 2010: Hanoi.

137. General Statistics Office of Vietnam, *The 2009 Vietnam Population and Housing census: Completed results*, 2009: Hanoi.

138. Watts, D.J. and S.H. Strogatz, *Collective dynamics of 'small-world' networks.* Nature, 1998. **393**(6684): p. 440-442.

139. Johnston, L.G., et al., *Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations in HIV behavioral and biological surveillance: lessons learned from four case studies.* AIDS Care, 2010. **22**(6): p. 784-92.

140. Zhang, D., et al., *Internet use and risk behaviours: an online survey of visitors to three gay websites in China.* Sexually Transmitted Infections, 2007. **83**(7): p. 571-576.

141. Jakopanec, I., et al., *Self-reported sexually transmitted infections and their correlates among men who have sex with men in Norway: an Internet-based cross-sectional survey.* Bmc Infectious Diseases, 2010. **10**(1): p. 261.

142. Iguchi, M.Y., et al., *Simultaneous Recruitment of Drug Users and Men Who Have Sex with Men in the United States and Russia Using Respondent-Driven Sampling: Sampling Methods and Implications.* Journal of Urban Health-Bulletin of the New York Academy of Medicine, 2009. **86**: p. S5-S31.

143. Lu , X., et al., *The Sensitivity of Respondent-driven Sampling.* Journal of the Royal Statistical Society, Series A (Statistics in Society), 2011.

144. Morris, M., H. Epstein, and M. Wawer, *Timing Is Everything: International Variations in Historical Sexual Partnership Concurrency and HIV Prevalence.* Plos One, 2010. **5**(11).

145. Harrison, A., J. Cleland, and J. Frohlich, *Young People's Sexual Partnerships in KwaZulu-Natal, South Africa: Patterns, Contextual Influences, and HIV Risk.* Studies in Family Planning, 2008. **39**(4): p. 295-308.

146. Blower, S.M. and C. Boe, *Sex acts, sex partners, and sex budgets: implications for risk factor analysis and estimation of HIV transmission probabilities.* J Acquir Immune Defic Syndr, 1993. **6**(12): p. 1347-52.

147. Koblin, B.A., et al., *Risk factors for HIV infection among men who have sex with men.* AIDS, 2006. **20**(5): p. 731-739.

148. Koblin, B.A., et al., *High-risk behaviors among men who have sex with men in 6 US cities: baseline data from the EXPLORE Study.* Am J Public Health, 2003. **93**(6): p. 926-32.

149. Dukers, N.H.T.M., et al., *Risk Factors for Human Herpesvirus 8 Seropositivity and Seroconversion in a Cohort of Homosexula Men.* Am J Epidemiol, 2000. **151**(3): p. 213-224.

150. O'Brien, T.R., et al., *Evidence for concurrent epidemics of human herpesvirus 8 and human immunodeficiency virus type 1 in US homosexual men: rates, risk factors, and relationship to Kaposi's sarcoma.* Journal of Infectious Diseases, 1999. **180**(4): p. 1010-1017.

151. Fogarty, A., *The Health in Men and Positive Health cohorts: a comparison of trends in the health and sexual behaviour of HIV-negative and HIV-positive gay men, 2002-2005*2006: National Centre in HIV Social Research, Faculty of Arts and Social Sciences, the University of New South Wales.

152. Adam, P., I. Teva, and J. de Wit, *Balancing risk and pleasure: sexual self-control as a moderator of the influence of sexual desires on sexual risk-taking in men who have sex with men.* Sexually Transmitted Infections, 2008. **84**(6): p. 463-467.

153. Bissessor, M., et al., *Delay in the diagnosis of early syphilis among men who have sex with men: need for greater community and health provider education.* International journal of STD & AIDS, 2009. **20**(1): p. 52-53.

154. Chapman, J., et al., *Sex and sexuality in the Shenzhen tongzhi circle: HIV risk context and migrant men who have sex with men in China.* Culture Health & Sexuality, 2009. **11**(7): p. 689-702.

155. Zhongxin, S., J. Farrer, and K.-h. Choi, *Sexual Identity Among Men Who Have Sex with Men in Shanghai.* China perpectives, 2006. **64**.

156. Lewis, F., et al., *Episodic sexual transmission of HIV revealed by molecular phylodynamics.* PLoS Med, 2008. **5**(3): p. e50.

157. Brenner, B.G., et al., *Transmission clustering drives the onward spread of the HIV epidemic among men who have sex with men in Quebec.* J Infect Dis, 2011. **204**(7): p. 1115-9.

158. Mishra, V., et al., *HIV infection does not disproportionately affect the poorer in sub-Saharan Africa.* Aids, 2007. **21**: p. S17-S28.

159. Jakopanec, I., et al., *Self-reported sexually transmitted infections and their correlates among men who have sex with men in Norway: an Internet-based cross-sectional survey.* Bmc Infectious Diseases, 2010. **10**.

160. Xiridou, M., et al., *The contribution of steady and casual partnerships to the incidence of HIV infection among homosexual men in Amsterdam.* Aids, 2003. **17**(7): p. 1029-1038.

161. Lin, H., et al., *Tracing sexual contacts of HIV-infected individuals in a rural prefecture, Eastern China.* BMC Public Health, 2012. **12**(1): p. 533.

162. Zhang, H., et al., *A Pilot Intervention to Increase Condom Use and HIV Testing and Counseling Among Men Who Have Sex With Men in Anhui, China.* JAIDS Journal of Acquired Immune Deficiency Syndromes, 2010. **53**: p. S88-S92 10.1097/QAI. 0b013e3181c7de63.

163. Amirkhanian, Y.A., et al., *A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria.* AIDS, 2005. **19**(16): p. 1897-905.

164. Amirkhanian, Y.A., et al., *Evaluation of a social network HIV prevention intervention program for young men who have sex with men in Russia and Bulgaria.* AIDS Educ Prev, 2003. **15**(3): p. 205-20.

165. Cohen, M.S., et al., *Prevention of HIV-1 infection with early antiretroviral therapy.* N Engl J Med, 2011. **365**(6): p. 493-505.

166. Van Damme, L., et al., *Preexposure prophylaxis for HIV infection among African women.* N Engl J Med, 2012. **367**(5): p. 411-22.

167. Wilson, D., *HIV Treatment as Prevention: Natural Experiments Highlight Limits of Antiretroviral Treatment as HIV Prevention.* PLoS Med 2012. **9**(7): p. e1001231.

168. Paltiel, A.D., et al., *HIV Preexposure Prophylaxis in the United States: Impact on Lifetime Infection Risk, Clinical Outcomes, and Cost-Effectiveness.* Clinical Infectious Diseases, 2009. **48**(6): p. 806-815.

169. USA Today, *Vietnam considers same-sex marriage*, 2012. http://www.usatoday.com/news/world/story/2012-07-29/vietnam-gay-marriage/56573384/1

Appendices

Appendix I

In-Depth Interviews – Question Guide

About this Document
The question guide will serve to create a framework for themes that will be discussed, and possible questions that may be posed. All areas and probes will not be brought up during each interview but the interview guide rather servers to show the type of areas and issues that will be explored during the interviews. The order of the themes will not be predetermined but brought up in a sequence that is suitable for the particular context of the individual interview. Unanticipated issues not previously conceived of will be explored, when encountered. The interview will differ substantially depending on if the interviewee has stable relationship(s) or only casual contacts.

The interview will start with providing information about the interviewers including the purpose of the study. Possibilities are given to pose all questions that the participant wants to ask. Written or clear verbal consent is given by the participant before start of the interview.

The Interview

General information
(serving as an introduction, starting up of the interview situation and provide a general background for the ensuing talk)

Can you tell me a bit about yourself and your background?
-… and what does your family look like?
-… What about your friends?
-… How old are you?
-… Are you born in Hanoi?
-… Can you tell me briefly about what you have done before, for example education and work?
-…What do you want to do in the future?

Sexual relationships
Exploring ongoing relationship(s), what is important in the relationship(s) and understand the reasons for concurrent partnerships/multiple sex partners.
Painting a picture of the person's ongoing and previous sexual relationships:
How is your relationship life now? Can you tell me about your relationships that you have now?

Do you have sex with other people too?

In your relationship with person X:
-… how is your relationship?
-… what things are important in your relationship with him?

-… what is good and bad in your relationship?
-… is there anything you would like to be different in your relationship?

How about the other relationship you had with person Y:
-… how is your relationship?
-… how did you meet?
-… what things are important in your relationship with him?
-… what is good and bad in your relationship?
-… is there anything you would like to be different in your relationship?

How are the other relationships you talked about earlier?

How have your previous relationships been? Why did they end?

Do you go out to other places to search for someone to have sex with?

What do you want then? What is the reasons for that?

- …Do you sometimes feel pressured to have sex or that you "should" have sex when you do not really want to? In what way do you feel that?

Among your friends, what kinds of people are they impressed by? What kind of sexual relationships do they have?

Imagine that you could not have sex anymore for some medical reason, how would that be? How would that make you feel? How would other people see you?

Reflecting on the described relationships
You have described different kinds of relationships with different persons

What is different between them?

How come you have these different relationships?

How would it be to have just one relationship/one person to have sex with?

When you are in a relationship like now, you also seek out sex with other people, tell me more about why you do that?
In the relationship you have with person X how much do you talk about your other relationships?

How do you think he feels about your other relationships?

How do you feel about person X having another relationship?

What kind of commitment is there between you? Are there things he expects you to do or not to do? What do you expect him to do for you?

What do feel if you have sex with someone else now?

What would you feel if he had sex with someone else?

Concept of relationship
Do you or other people have names for different kinds of relationships? What?

How are they different?

How would your ideal relationship be?

How does that differ from the relationship you have today?

Why do you think this relationship you described before, is different from your Ideal relationship?

In marriage between men and women, many people think that both people should be faithful to each others and not have any other sexual relations. Would you like to have such a relation with a man/your partner X/someone?

Why don't you have such a relationship now?

Do your friends also think like this about relationships? What is different?

Tying it all together: HIV and multiple relationships
You have described about a lot of relationships you have had. Do you ever think about HIV when you have these relationships?
Do you think that you have a risk of getting HIV?
What way do you think is best for you to reduce your risk of getting HIV?
If everybody kept to one partner, it could be one way to prevent HIV from spreading. What do you think about that? How would that work for you?
What are your friends thinking about that?
Evaluation of the interview by the participant
What can be improved in the interview? What was good? What was bad?

Appendix II

Question Number	Question English	Response Alternatives English	Question Vietnamese	Response Alternatives Vietnamese
Q1	Do you have a boyfriend now? ("boyfriend" here indicates a stable relationship with emotional attachment)	Yes No Don't want to answer	Hiện tại bạn có người yêu là nam giới không?	Có Không Không muốn trả lời
Q2	What is the longest time that you have maintained a ("love") relationship with a man?	Never had a ("love") relationship Less than 1 month 1 to 6 months 6 months to 1 year 1 to 3 years More than 3 years Don't want to answer	Mối quan hệ tình cảm dài nhất với một người đàn ông mà bạn từng có từ trước đến nay kéo dài trong bao lâu?	Chưa bao giờ Dưới 1 tháng Từ 1 đến 6 tháng Từ 6 tháng đến 1 năm Từ 1 đến 3 năm Hơn 3 năm Không muốn trả lời
Q3	If you meet someone for sex, what characteristic of that person is most important to you?	Humorous Rich Intelligent Good looking Understanding Rich Respect me Good at sex Don't care/Don't want to answer	Nếu bạn gặp một người đàn ông chỉ để quan hệ tình dục, bạn coi đặc điểm gì của người ấy là quan trọng nhất?	Hài hước Thông minh Đẹp trai Hiểu tôi Giàu có Tôn trọng tôi Làm tình giỏi Không quan tâm/ không muốn trả lời
Q4	If you are looking for a man for a long-term relationship, what characteristic of that person is most important to you?	Humorous Intelligent Good looking Understanding Rich Faithful Respect me Good at sex Don't care/Don't want to answer	Nếu bạn tìm kiếm một người đàn ông làm người yêu lâu dài, bạn coi đặc điểm gì của người ấy là quan trọng nhất?	Hài hước Thông minh Đẹp trai Hiểu tôi Giàu có Chung thủy Tôn trọng tôi Làm tình giỏi Không quan tâm/ không muốn trả lời
Q5	Do you think same sex marriage should be permitted in Vietnam?	Yes No No opinion/ Don't want to answer	Bạn có nghĩ luật pháp Việt Nam nên cho phép kết hôn đồng giới không?	Có Không Không có ý kiến/ không muốn trả lời
Q6	During the last 6 months, how many men have you had sex with (anal, oral or masturbation)?	Text	Trong 06 tháng qua, bạn đã quan hệ tình dục (đường miệng, đường hậu môn hoặc thủ dâm cho nhau) với bao nhiêu nam giới? (Nếu không nhớ chính xác, hãy đưa ra một con số ước lượng gần nhất)	Text
Q7	Which of the following statements best describe your preferences?	Prefer only men as sexual partners Prefer men to women as sexual partners Prefer women to men as sexual partners Prefer only women as sexual partners Don't want to answer	Trong những câu sau, câu nào miêu tả sở thích chọn bạn tình của bạn chính xác nhất?	Chỉ thích bạn tình là nam giới Thích bạn tình là nam giới hơn nữ giới Thích bạn tình là nữ giới hơn nam giới Chỉ thích bạn tình là nữ giới Không muốn trả lời
Q8	In the last 6 months, have you ever had sex in a public place like a sauna, gym, swimming pool, public toilet or park?	Yes No Don't want to answer	Trong vòng 06 tháng qua, bạn có quan hệ tình dục ở những chỗ công cộng như phòng tắm hơi, phòng tập thể hình, bể bơi, nhà vệ sinh công cộng hay công viên không?	Có Không Không nhớ/ không muốn trả lời
Q9	In what year were you born?	Don't want to answer 1999 … 1940	Bạn sinh năm nào?	Tôi không muốn trả lời 1999 … 1940

Q10	What is your highest level of education? Select only one option.	No schooling	Trình độ học vấn cao nhất của bạn?	Chưa bao giờ đi học
		Primary (Grade 1-5)		Tiểu học (Lớp 1-5)
		Secondary school (Grade 6-9)		Cấp 2 (Lớp 6-9)
		High school (Grade 10 – 12)		Cấp 3 (Lớp 10 – 12)
		University, college or vocational training		Đại học hoặc cao đẳng, dạy nghề
		Postgraduate		Sau đại học
		Don't want to answer		Không muốn trả lời
Q11	During the last 12 months, what was the average amount of money you received per each month, from all sources. (include money from salary, parents, interests and all other sources)	<1,000,000 VND	Trong 12 tháng vừa qua, số tiền thu nhập trung bình mỗi tháng từ tất cả các nguồn (ví dụ: lương, tiền bố mẹ cho, từ bạn bè, tiền lãi trong kinh doanh...) của bạn là bao nhiêu?	Dưới 1,000.000 đồng
		1,000,000 – 3,000,000 VND		Từ 1.000.000 Đồng đến dưới 3.000.000 Đồng
		3,000,000 – under 5,000,000 VND		Từ 3.000.000 Đồng đến dưới 5.000.000 Đồng
		5,000,000 – under 10,000,000 VND		Từ 5.000.000 Đồng đến dưới 10.000.000 Đồng
		>10,000,000 VND		Trên 10.000.000 Đồng
		Don't want to answer		Không nhớ/ không muốn trả lời
Q12	In which province are you living?	Don't want to answer/outside Vietnam	Bạn đang sống ở tỉnh/ thành phố nào?	Không muốn trả lời/ ở bên ngoài lãnh thổ Việt Nam
		Hồ Chí Minh		Hồ Chí Minh
		Hà Nội		Hà Nội
		Hải Phòng		Hải Phòng
		Khánh Hoà		Khánh Hoà
		Cần Thơ		Cần Thơ
		----------------		----------------
		An Giang		An Giang
		…		…
		Yên Bái		Yên Bái
Q13	During the last month, how many days did you use the Internet? If you do not remember exactly, please give your best guess.	Don't want to answer	Bạn sử dụng Internet bao nhiêu ngày trong vòng 30 ngày vừa qua? (Nếu không nhớ chính xác, hãy đưa ra một con số ước lượng gần nhất)	Không nhớ / không muốn trả lời
		1 or less than 1 day per month		1 ngày hoặc dưới 1 ngày
		2 days		2 ngày
		…		…
		30 days		30 ngày
Q14	During the last 7 days, how many people in your world have you had any type of contact with (in person, on the phone, on chat, facebook, mail or in some other way)?If you do not remember exactly, please give your best guess.	Text	Trong vòng 7 ngày vừa qua, bạn nói chuyện với bao nhiêu người trong giới, dưới bất kỳ hình thức nào như nói chuyện trực tiếp, qua điện thoại, chat, facebook, email, gửi thư hoặc một cách nào khác?	Text
Q15	Out of these in question 14, how many use the internet?If you do not know exactly, please give your best guess.	Text	Trong số những người bạn đề cập ở câu 14 (những người bạn nói chuyện với dưới bất kể hình thức nào trong vòng 7 ngày vừa rồi), có bao nhiêu người từ 18 tuổi trở lên và sử dụng Internet? (Nếu không biết chính xác, hãy đưa ra một con số ước lượng gần nhất)	Text
Q16	What is your relationship to the person who invited you to this study? (You can choose a maximum of two alternatives)	He is a stranger (I have not communicated with him before I got this invitation)	Bạn có quan hệ như thế nào với người mời bạn tham gia nghiên cứu này? (Bạn có thể chọn tối đa 2 phương án trả lời)	Người ấy là người lạ (tôi không có liên lạc gì với người ấy trước khi tôi nhận được lời mời này)
		He is an acquaintance		Người ấy là người quen biết
		He is a friend		Người ấy là bạn
		He is a close friend		Người ấy là bạn thân
		He is a lover/ex lover		Người ấy là người yêu/ người yêu cũ
		Relative		Họ hàng
		Don't want to answer		Không muốn trả lời

Q17	In what context did you get to know that person? How did you get to know this person? (Choose more than one alternative when appropriate)	Through an MSM web page, chat room, facebook or other Internet site	Bạn biết người ấy trong trường hợp nào? (Bạn có thể chọn nhiều phương án trả lời)	Thông qua các trang web giành cho người đồng tính, phòng chat, facebook hoặc các trang web khác trên internet
		Through people I know (friends, relatives, lovers etc)		Qua người quen (bạn bè, họ hàng, người yêu...)
		Through an MSM club		Qua câu lạc bộ MSM
		Through work		Qua công việc
		Through school, university or other type of education		Ở trường học hoặc các cơ sở đào tạo
		Through a leisure activity		Qua các trò vui chơi, giải trí
		At an MSM venue (bar, disco, sauna, park, street for MSM etc)		Tại các tụ điểm cho MSM (bar, sàn nhảy, phòng tắm hơi (sauna), công viên hay đường phố)
		At a non-MSM venue.		Các tụ điểm khác không dành riêng cho MSM
		Other		Khác
		Don't remember/ Don't want to answer		Không nhớ/ không muốn trả lời

Appendix III

Question Number	Eng Question	Vn Question	Question Type
	Part 1/5: Background	**Phần 1/5: Thông tin chung**	
1.1	**What year were you born?** Don't want to answer 1999-1940	**Bạn sinh năm bao nhiêu?** Không muốn trả lời 1999-1940	Drop down
1.2	**What is the highest level of education you have reached?** Please use the highest level you have attended, even if you have/did not finish it. Never been to school Level 1 (grades 1 to 5) Level 2 (grades 6 to 9) Level 3 (grades 10 to 12): not yet graduated Level 3 (grades 10 to 12): graduated Trung cấp [2-year college] Cao đẳng [3-year college] Bachelor level Master level Doctoral level Don't want to answer	**Trình độ học vấn của bạn?** Hãy chọn bậc cao nhất mà bạn đã từng học, dù có thể bạn chưa học xong. Chưa từng đi học Cấp 1 (từ lớp 1 đến lớp 5) Cấp 2 (từ lớp 6 đến lớp 9) Cấp 3 (từ lớp 10 đến lớp 12): chưa tốt nghiệp Cấp 3 (từ lớp 10 đến lớp 12): đã tốt nghiệp Trung cấp [hệ 2 năm] Cao đẳng [hệ 3 năm] Đại học Cao Học Tiến Sĩ Không muốn trả lời	Radio button
1.3	**On average, what is your monthly income?** Count both the money you make and the money you are given by family or others (if any). 0 to less than 1 million VND 1 million – less than 2 million 2 million – less than 3 million 3 million – less than 4 million 4 million – less than 5 million 5 million – less than 10 million 10 million – less than 15 million 15 million – less than 20 million Over 20 million Don't want to answer	**Trung bình mỗi tháng thu nhập của bạn khoảng bao nhiêu?** Tính cả tiền tự làm ra và tiền được gia đình hoặc người khác cho (nếu có). Từ 0 đến dưới 1 triệu đồng 1 triệu - dưới 2 triệu 2 triệu - dưới 3 triệu 3 triệu - dưới 4 triệu 4 triệu - dưới 5 triệu 5 triệu - dưới 10 triệu 10 triệu - dưới 15 triệu 15 triệu - dưới 20 triệu Trên 20 triệu Không muốn trả lời	Radio box
1.4	**In which province are you living?** Don't want to answer/Outside Vietnam Hồ Chí Minh Hà Nội Hải Phòng Đà Nẵng Khánh Hoà Cần Thơ ---------------- All provinces including the ones above in alphabetic order: An Giang to Yên Bái	**Bạn đang sống ở tỉnh nào?** Không muốn trả lời/Ở nước ngoài Hồ Chí Minh Hà Nội Hải Phòng Đà Nẵng Khánh Hoà Cần Thơ ---------------- An Giang...Yên Bái	Drop down
1.5	**During the last 30 days, how many days did you use the Internet?** If you do not remember exactly, please give your best guess. 1-7 days 8-14 days 15-21 days 22-30 days Don't want to answer	**Bạn sử dụng Internet bao nhiêu ngày trong vòng 30 ngày vừa qua?** Nếu không nhớ chính xác, hãy đưa ra một con số ước lượng gần nhất 1-7 ngày 8-14 ngày 15-21 ngày 22-30 ngày Không muốn trả lời	Radio button

	Eng question	Vn question	Question type
1.6	During the last 7 days, how many people in your world have you communicated with (in person, on the phone, on chat, email or in some other way)? If you do not remember exactly, please give your best guess. Fill in a number in the below box.	Trong vòng 7 ngày vừa qua, bạn đã nói chuyện với bao nhiêu người trong giới của bạn (chẳng hạn như nói chuyện trực tiếp, qua điện thoại, chat trên mạng, facebook, email, gửi thư hoặc một cách nào khác? Nếu không nhớ chính xác, hãy đưa ra một con số ước lượng gần nhất. Điền câu trả lời bằng chữ số vào ô dưới đây.	Box for text input
1.7	Among the people you mentioned in question 1.6 (the people you communicated with in the last 7 days by any means), how many are 18 years or above and use the Internet? If you do not know exactly, please give your best guess. Fill in a number in the below box.	Trong số những người bạn đề cập ở câu 1.6 (những người bạn đã nói chuyện dưới bất kể hình thức nào trong vòng 7 ngày vừa rồi), có bao nhiêu người từ 18 tuổi trở lên và sử dụng Internet? Nếu không nhớ chính xác, hãy đưa ra một con số ước lượng gần nhất. Điền câu trả lời bằng chữ số vào ô dưới đây.	Box for text input
1.8	What is your relationship to the person who invited you to this study? (You can tick choose up to two alternatives) He is a stranger (I have not communicated with him before I got this invitation) He is an acquaintance He is a friend He is a close friend He is a lover/ex lover Relative Don't want to answer	Bạn có quan hệ như thế nào với người mời bạn tham gia nghiên cứu này? (Bạn có thể chọn tối đa 2 phương án trả lời) Người ấy là người không quen biết (tôi không có liên lạc gì với người ấy trước khi tôi nhận được lời mời này) Người ấy là người quen biết Người ấy là bạn Người ấy là bạn thân Người ấy là người yêu/ người yêu cũ Họ hàng Không muốn trả lời	Check boxes
1.9	Which of the following statements best describe your preferences? Prefer only men as sexual partners Prefer men to women as sexual partners Prefer women to men as sexual partners Prefer only women as sexual partners Don't want to answer	Trong những câu sau, câu nào miêu tả sở thích chọn bạn tình của bạn chính xác nhất? Chỉ thích bạn tình là nam giới Thích bạn tình là nam giới hơn nữ giới Thích bạn tình là nữ giới hơn nam giới Chỉ thích bạn tình là nữ giới Không muốn trả lời	Radio button
	Eng question	**Vn question**	**Question type**
	Part 3/5: Sex with men last 3 months	**Phần 3/5: Quan hệ tình dục với nam giới trong vòng 3 tháng trở lại đây**	
3.1	During the last <u>3 months</u> have you had sex with a <u>man</u> (manual, oral or anal)? Yes No Don't want to answer	Trong vòng <u>3 tháng</u> qua, bạn có quan hệ tình dục (tay, miệng, hậu môn) với <u>nam giới</u> không? Có Không Không muốn trả lời	Radio buttons
3.2	During the last three months, with <u>how many different men</u> did you have <u>anal intercourse</u> (top or bottom)?	Trong vòng 3 tháng qua, bạn đã <u>quan hệ tình dục qua đường hậu môn</u> với <u>bao nhiêu người nam giới khác nhau</u> (trên (top) hoặc dưới (bottom))?	Text input
3.3	With how many of these men did you have anal intercourse <u>without a condom</u> during the last 3 months?	Trong số những người này có bao nhiêu người mà bạn đã quan hệ qua đường hậu môn mà <u>không sử dụng bao cao su</u> trong vòng 3 tháng qua?	Text input
3.3.1-A	Now we will ask you some questions about this man with whom you had anal sex without condom sometime during the last 3 months. During the last 3 months, <u>how many anal intercourses</u> did you have with this man <u>without using condom</u> (counting both when you were top and when you were bottom)? If you don't know exactly, try your best to estimate.	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về người nam giới mà bạn đã có quan hệ tình dục qua đường hậu môn mà không sử dụng bao cao su trong 3 tháng qua. Trong vòng 3 tháng qua, bạn đã <u>bao nhiêu lần quan hệ tình dục qua đường hậu môn</u> với người này mà <u>không sử dụng bao cao su</u> (tính cả khi bạn ở vị trí trên (top) và khi bạn ở vị trí dưới (bottom))? Nếu bạn không nhớ chính xác, hãy ước lượng một con số gần đúng nhất.	Text input

3.3.1-B	**When was the first time you had <u>anal intercourse</u> with this man <u>without using condom</u>?** less than 3 months ago less than 6 months ago less than 12 months ago less than 3 years ago More than 3 years ago Don't want to answer	**Lần đầu tiên bạn có <u>quan hệ tình dục qua đường hậu môn</u> với người này <u>mà không sử dụng bao cao su</u> là khi nào?** Trong vòng 3 tháng trở lại đây Trong vòng 6 tháng trở lại đây Trong vòng 1 năm trở lại đây Trong vòng 3 năm trở lại đây Từ hơn 3 năm trước Không muốn trả lời	Text input
3.3.1-C	**Were you in a <u>romantic/emotional relationship</u> with him at any time during the last 3 months?** Yes No Don't want to answer	**Bạn có ở trong một <u>mối quan hệ tình cảm hoặc tình yêu</u> với người này vào một thời điểm nào đó trong thời gian 3 tháng qua hay không?** Có Không Không muốn trả lời	radio buttons
3.4.1-A	Now we will ask you some questions about these two men with whom you had anal sex without condom sometime during the last 3 months. Think about the first of these two men: *Repeat: 3.3.1-A to C*	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về hai người nam giới mà bạn đã có quan hệ tình dục qua đường hậu môn mà không sử dụng bao cao su trong 3 tháng qua. Hãy nghĩ đến người đầu tiên trong hai người nam giới này: Repeat: 3.3.1-A to C	Text input
3.5.1-A	Now think about the second of these two men: *Repeat: 3.3.1-A to C*	Bây giờ hãy nghĩ đến người thứ hai trong hai người nam giới này: Repeat: 3.3.1-A to C	Text input
3.6.1-A	Now we will ask you some questions about these three men with whom you had anal sex without condom sometime during the last 3 months. Think about the first of these three men: *Repeat: 3.3.1-A to C*	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về ba người nam giới mà bạn đã có quan hệ tình dục qua đường hậu môn mà không sử dụng bao cao su trong 3 tháng qua. Hãy nghĩ đến người đầu tiên trong ba người nam giới này: Repeat: 3.3.1-A to C	Text input
3.7.1-A	Think about the second of these three men: *Repeat: 3.3.1-A to C*	Bây giờ hãy nghĩ đến người thứ hai trong ba người nam giới này: Repeat: 3.3.1-A to C	Text input
3.8.1-A	Think about the third of these three men: *Repeat: 3.3.1-A to C*	Bây giờ hãy nghĩ đến người thứ ba trong ba người nam giới này: Repeat: 3.3.1-A to C	Text input
3.9.1-A	Now we will ask you some questions about the three men with whom you <u>most recently</u> had anal sex without a condom during the last 3 months. First think about the man with whom you <u>most recently</u> had sex without a condom: *Repeat: 3.3.1-A to C*	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về ba người nam giới <u>gần đây nhất</u> mà bạn đã có quan hệ tình dục qua đường hậu môn mà không sử dụng bao cao su trong 3 tháng qua. Trong ba người này, hãy nghĩ đến người <u>gần đây nhất</u> mà bạn đã có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 3.3.1-A to C	Text input
3.10.1-A	Now think about the second man with whom you <u>most recently</u> had sex without a condom: *Repeat: 3.3.1-A to C*	Bây giờ hãy nghĩ đến người thứ hai trong số ba người nam giới <u>gần đây nhất</u> mà bạn có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 3.3.1-A to C	
3.11.1-A	Now think about the third man with whom you <u>most recently</u> had sex without a condom: Repeat: 3.3.1-A to C	Bây giờ hãy nghĩ đến người thứ ba trong số ba người nam giới <u>gần đây nhất</u> mà bạn có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 3.3.1-A to C	Text input

	Part 4/5: Sex with women last 3 months	Phần 4/5: Quan hệ tình dục với phụ nữ trong vòng 3 tháng qua	
4.1	**Have you had vaginal or anal sex with a <u>woman</u> during the last 3 months?** Yes No Don't want to answer	Trong vòng <u>3 tháng</u> qua, bạn có quan hệ tình dục qua đường hậu môn hoặc âm đạo với <u>phụ nữ</u> không? Có Không Không muốn trả lời	radio buttons
4.2	**During the last 3 months, with <u>how many</u> different women did you have vaginal or anal intercourse <u>without using condom</u> ?**	Trong vòng 3 tháng qua, bạn đã quan hệ tình dục qua đường hậu môn hoặc âm đạo với bao nhiêu người phụ nữ mà <u>không dùng bao cao su</u>?	text input
4.2.1-A	Now we will ask you some questions about this woman with whom you had vaginal or anal sex <u>without condom</u> sometime during the last 3 months. **During the last 3 months, how many vaginal or anal <u>intercourses without using condom</u> did you have with this woman?** If you don't know exactly, try your best to estimate.	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về người phụ nữ mà bạn đã có quan hệ tình dục qua đường hậu môn hoặc âm đạo mà <u>không sử dụng bao cao su</u> trong 3 tháng qua. **Trong vòng 3 tháng qua, bạn đã <u>bao nhiêu lần quan hệ tình dục qua</u> đường hậu môn hoặc âm đạo với người này mà <u>không sử dụng bao cao su</u>?** Nếu bạn không nhớ chính xác, hãy ước lượng một con số gần đúng nhất.	text input
4.2.1-B	**When was the first time you had <u>vaginal or anal intercourse</u> with this woman <u>without using condom</u>?** Less than 3 months ago Less than 6 months ago Less than 12 months ago Less than 3 years ago More than 3 years ago Don't want to answer	**Lần đầu tiên bạn có <u>quan hệ tình dục qua đường hậu môn hoặc âm đạo</u> với người này <u>mà không sử dụng bao cao su</u> là khi nào?** Trong vòng 3 tháng trở lại đây Trong vòng 6 tháng trở lại đây Trong vòng 1 năm trở lại đây Trong vòng 3 năm trở lại đây Từ hơn 3 năm trước Không muốn trả lời	radio buttons
4.2.1-C	**Were you in a <u>romantic/emotional relationship</u> with her at any time during the last 3 months?** Yes No Don't want to answer	**Bạn có ở trong một <u>mối quan hệ tình cảm hoặc tình yêu</u> với người này vào một thời điểm nào đó trong thời gian 3 tháng qua hay không?** Có Không Không muốn trả lời	radio buttons
4.3.1-A	Now we will ask you some questions about these two women with whom you had vaginal or anal sex <u>without condom</u> sometime during the last 3 months. Think about the first of these two women: Repeat: 4.2.1-A to C	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về hai người phụ nữ mà bạn đã có quan hệ tình dục qua đường hậu môn hoặc âm đạo mà <u>không sử dụng bao cao su</u> trong 3 tháng qua. Hãy nghĩ đến người đầu tiên trong hai người phụ nữ này: Repeat: 4.2.1-A to C	text input
4.4.1-A	Think about the second of these two women: Repeat: 4.2.1-A to C	Hãy nghĩ đến người thứ hai trong hai người phụ nữ này: Repeat: 4.2.1-A to C	text input
4.5.1-A	Now we will ask you some questions about these three women with whom you had vaginal or anal sex <u>without condom</u> sometime during the last 3 months. Think about the first of these three women:	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về ba người phụ nữ mà bạn đã có quan hệ tình dục qua đường hậu môn hoặc âm đạo mà <u>không sử dụng bao cao su</u> trong 3 tháng qua. Hãy nghĩ đến người đầu tiên trong ba người phụ nữ này: Repeat: 4.2.1-A to C	text input
4.6.1-A	Think about the second of these three women: Repeat: 4.2.1-A to C	Hãy nghĩ đến người thứ hai trong ba người phụ nữ này: Repeat: 4.2.1-A to C	text input
4.7.1-A	Think about the third of these three women: Repeat: 4.2.1-A to C	Hãy nghĩ đến người thứ ba trong ba người phụ nữ này: Repeat: 4.2.1-A to C	text input

4.8.1-A	Now we will ask you some questions about the three women with whom you <u>most recently</u> had vaginal or anal sex without a condom during the last 3 months. First think about the woman with whom you <u>most recently</u> had sex without a condom: Repeat: 4.2.1-A to C	Bây giờ chúng tôi sẽ hỏi bạn một đôi điều về ba người phụ nữ <u>gần đây nhất</u> mà bạn đã có quan hệ tình dục qua đường hậu môn hoặc âm đạo mà <u>không sử dụng bao cao su</u> trong 3 tháng qua. Trong ba người phụ nữ này, hãy nghĩ đến người phụ nữ <u>gần đây nhất</u> mà bạn có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 4.2.1-A to C	text input
4.9.1-A	Now think about the second woman with whom you <u>most recently</u> had sex without a condom: Repeat: 4.2.1-A to C	Bây giờ hãy nghĩ đến người thứ hai trong ba người phụ nữ <u>gần đây nhất</u> mà bạn có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 4.2.1-A to C	text input
4.10.1-A	Now think about the third woman with whom you <u>most recently</u> had sex without a condom: Repeat: 4.2.1-A to C	Bây giờ hãy nghĩ đến người thứ ba trong ba người phụ nữ <u>gần đây nhất</u> mà bạn có quan hệ tình dục mà không sử dụng bao cao su: Repeat: 4.2.1-A to C	text input
	Part 5/5: Male Sex Partners during Last 12 Months	**Phần 5 / 5: Các bạn tình là nam giới trong vòng 12 tháng qua**	
5.1	**During the <u>last 12 months</u>, with how many different men did you have <u>anal sex without a condom?</u> (top or bottom)**	**Trong vòng <u>12 tháng qua</u>, bạn đã có <u>quan hệ tình dục qua đường hậu môn mà không sử dụng bao cao su</u> với <u>bao nhiêu người nam giới?</u> (trên (top) hoặc dưới (bottom))**	Text input
5.2	**With how many of the these men in did you have a emotional/romantic relationship?**	**Bạn đã có quan hệ tình cảm hoặc tình yêu với bao nhiêu người trong số những người nam giới đó?**	Text input
5.3	**When was the last time you took an HIV test and also received the results?** I never tested for HIV or I never received the results < 3 months ago < 6 months ago < 12 months ago < 3 years ago > 3 years ago Don't want to answer	**Lần cuối cùng bạn làm xét nghiệm và nhận kết quả kiểm tra HIV là khi nào?** Tôi chưa bao giờ làm xét nghiệm hoặc chưa bao giờ nhận kết quả xét nghiệm HIV < dưới 3 tháng < dưới 6 tháng < dưới 12 tháng < dưới 3 năm > trên 3 năm Không muốn trả lời	
5.4	**Would you like to tell us the results of the test?** I had HIV I didn't have HIV Don't want to answer	**Bạn có thể vui lòng cho chúng tôi biết kết quả xét nghiệm lần cuối cùng của bạn không?** Tôi đã bị nhiễm HIV Tôi không bị nhiễm HIV Không muốn trả lời	