BMJ Open

Strong propensity for HIV transmission among men who have sex with men in Vietnam: behavioural data and sexual network modelling

Linus Bengtsson,1,2 Xin Lu,1,2,3 Fredrik Liljeros,2,4 Hoang Huy Thanh,5 Anna Thorson1

ABSTRACT

Objectives: Survey data from men who have sex with men (MSM) in Asian cities indicate ongoing and drastic increases in HIV prevalence. It is unknown which behavioural factors are most important in driving these epidemics. We aimed to analyse detailed sexual behaviour data among MSM in Vietnam and to model HIV transmission using improved assumptions on sexual network structure.

Setting: Vietnam.

Participants: Internet-using men who had ever had sex (any type) with a man, aged ≥18 years and living in Vietnam. The study was cross-sectional, population-based and performed in 2012, using online respondent-driven sampling. The Internet-based survey instrument was completed by 982 participants, of which 857 were eligible. Questions included sociodemography and retrospective sexual behaviour, including number of unprotected anal sex (UAS) acts per partner.

Primary and secondary outcome measures: Estimated basic reproductive number over 3 months as a function of transmission risk per UAS act; frequency distributions of number of UAS partners and UAS acts during last 3 months.

Results: 36% (CI 32% to 42%) reported UAS at least once during the last 3 months. 36% (CI 32% to 41%) had ever taken an HIV test and received the result. UAS partner numbers and number of UAS acts were both highly skewed and positively correlated. Using a weighted configuration model, taking into account partner numbers, frequency of UAS and their correlations, we estimated the basic reproductive number (R0) over 3 months. The results indicated rapid transmission over a wide range of values of per-act transmissibility.

Conclusions: Men with multiple partners had unexpectedly high UAS frequency per partner, paired with low HIV testing rates. The study highlights the importance of collecting data on frequency of UAS acts and indicates the need to rapidly scale-up HIV prevention services and testing opportunities for MSM in Vietnam.

INTRODUCTION

Men who have sex with men (MSM) are at high risk of HIV infection in all regions globally. Rapid increases in HIV prevalence rates have recently been documented from cities in Southeast Asia. HIV prevalence rates in cross-sectional surveys among MSM in Bangkok rose from 17% to 31% between 2003 and 2007. Similar increases have been recorded among MSM in Vietnam, where HIV prevalence rates in Hanoi and Ho Chi Minh City (HCMC) were reported at 9% and 5%, respectively, in 2005 and at 15% and 17%, respectively, in 2009. Sexual HIV transmission dynamics in these MSM populations are not well known and need to be characterised to inform HIV prevention strategies.

HIV epidemics are however highly complex phenomena where population-level transmission depends both on the probabilities of HIV transmission between individuals and on the structures of sexual networks. Several challenges exist to gain knowledge about the dynamics of HIV transmission among MSM and we address three of the most important ones in this paper. First, representative sampling of MSM is very difficult to attain due to lack of a sampling frames and high levels of stigma. Second, although HIV transmission probabilities per anal sex act are highly heterogeneous, the frequency of unprotected anal sex...
(UAS) per sexual partner is likely to have a large impact on the basic reproductive number and total population-level transmission of HIV.\textsuperscript{13} \textsuperscript{14} Despite this, traditional sexual behaviour surveys among MSM typically do not report the frequency of UAS per partner.\textsuperscript{4} The few studies reporting data on anal sex frequency are based on convenience samples, do not report data on the proportion of acts that are unprotected or study MSM groups with low or uncertain representativeness.\textsuperscript{12} \textsuperscript{14} \textsuperscript{16} We found no published information on the correlation between partner numbers and UAS frequency among MSM, which is crucial knowledge for correctly targeted prevention measures. Third, valid information on the structures of sexual networks is difficult to collect due to the considerable practical and ethical difficulties of tracing sexual connections from person to person in a population. Instead researchers rely on individuals to report the numbers and types of sexual contacts they have had, gathering data on people’s so called “ego-networks.”\textsuperscript{17} To model HIV transmission based on this data, researchers then connect the ego-networks to one complete sexual network. The number of ways in which ego-networks can be connected-up is however very large, with widely different implications for population-level transmission and the effectiveness of prevention interventions.\textsuperscript{18} \textsuperscript{19} We seek to meet these challenges by using new ways to collect and analyse detailed sexual behaviour data among Internet-using MSM in Vietnam. Respondent-driven sampling (RDS) is a chain-referral sampling method, which has important advantages for improving representative sampling in populations without a sampling frame.\textsuperscript{11} \textsuperscript{20} \textsuperscript{21} We have in earlier work described the successful development and implementation of an online RDS (webRDS) system for sampling of MSM in Vietnam.\textsuperscript{22} The webRDS draws on the fact that MSM increasingly communicate online and allows respondents to participate anonymously and with little effort, while covering a large and diverse geographical area.

We aimed to analyse partner numbers, frequency of UAS acts and their correlation among Internet-using MSM in Vietnam. We further aimed to model the sexual network using a weighted configuration model in order to estimate the propensity for HIV transmission among MSM in Vietnam.

\textbf{METHODS}

We performed a cross-sectional online survey between 26 March and 22 April 2012, applying online respondent driven sampling (webRDS). Internet use as a proportion of the population in Vietnam was 34\% in February 2012 (31 million persons).\textsuperscript{23} Internet use among urban MSM may be considerably higher.\textsuperscript{24} Ninety-four per cent of MSM in an offline RDS in Ha Noi stated that they used the Internet.\textsuperscript{25}

\textbf{Sampling and participants}

The study was performed in collaboration with a local research organisation in Vietnam (iSEE), working for non-discrimination of lesbian, gay, bisexual and transgender persons. Thirty-six 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. Participants logged into the survey with a code received from their recruiter, accessed information about the survey, approved participation and confirmed eligibility, before answering a questionnaire. After the survey, participants were encouraged to provide their email or Yahoo! Messenger address and were sent four invitation messages to forward to MSM friends. Bengtsson et al\textsuperscript{22} provide a detailed description of the sampling method. Nine hundred and sixty-nine non-seed submissions were received during the study. The maximum number of waves in the recruitment trees was 29 (see online supplementary figure S1).

\textbf{Data collection}

The questionnaire collected information on sociodemographics, frequency of Internet use, sexual preference, social network size and type of relationship between the participant and his recruiter. Sexual behaviour data included, for the last 3 months, the number of male sexual partners (including manual, oral and anal sex), and the number of partners with whom the respondent practiced UAS. For each of the three most recent UAS partners, data were collected on the number of UAS acts during the last 3 months. Qualitative research\textsuperscript{26} 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. Additionally, the Vietnamese language lacks a word for “relationship” which could be used appropriately in this context. Specifically we therefore asked if the respondent was “emotionally involved or mutually in love” (quan hê tình cam hoắc tình yêu) with his sexual partner, excluding that the partner was casual. Henceforth, we use the terms “casual” and “non-casual” partners, while keeping in mind that casual partners can include, for example, long-term transactional partners.

\textbf{Data cleaning}

Seeds were excluded as is customary in RDS analysis. Questionnaires were analysed for ineligible or duplicate participation. Nine per cent 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 submissions with no answers on sociodemographic 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 analysed the sensitivity of the descriptive estimates to inclusion and exclusion of these flagged submissions. Differences were present for the variables province and Internet use but were otherwise small (see online supplementary table S1). We also checked selected analyses (table 1, figures 1 and 2), with and without flagged submissions and there were no major changes to the results.

**Data analysis**

Equilibrium was reached after a maximum of 11 waves for all variables reported in this paper (see online supplementary figure S5). Ninety-five per cent CIs for proportions were calculated using bootstrapping, according to Salganik, and for numeric estimates a design effect of 4 was used. We asked participants for the number of MSM they had interacted with in any way during the past week. We then asked how many of these persons they believed used the Internet. We used the answer to the latter question to define the participants’ personal network sizes. For submissions with missing network size we used the average, RDSII adjusted, network size of submissions with non-missing data (12 persons).

We performed multivariate logistic regression to explore the associations between reporting UAS at least once during last 3 months (dependent variable) and the 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 analyses were also repeated with the dependent variable two or more UAS partners last 3 months. We followed the recommendation from Ramirez-Valles et al to perform regression analysis on unweighted data and include in the model all

| Independent variables | Unadjusted | Adjusted |
|-----------------------|------------|----------|
| Age (years)           |            |          |
| 18–20                 | 36         | Ref      |
| 21–22                 | 34         | 0.91 (0.62 to 1.34) | 0.99 (0.64 to 1.52) |
| 23–25                 | 43         | 1.32 (0.88 to 1.98) | 1.12 (0.69 to 1.79) |
| 26–30                 | 53         | 1.95 (1.20 to 3.18)* | 1.66 (0.92 to 2.97) |
| >30                   | 46         | 1.49 (0.79 to 2.82) | 1.20 (0.56 to 2.59) |
| Education (highest level started) |            |          |
| 12 years or less      | 41         | Ref      |
| 2-year college        | 46         | 1.21 (0.64 to 2.29) | 0.81 (0.39 to 1.70) |
| 3-year college        | 40         | 0.97 (0.58 to 1.60) | 0.78 (0.43 to 1.42) |
| Bachelor level, university | 37     | 0.85 (0.55 to 1.31) | 0.73 (0.43 to 1.24) |
| Masters/doctorate level university | 56      | 1.84 (0.83 to 4.09) | 1.11 (0.44 to 2.76) |
| Income (million VND all sources) |            |          |
| 0–less than 1         | 25         | Ref      |
| 1–less than 3         | 40         | 1.98 (1.14 to 3.45)* | 2.37 (1.28 to 4.36)* |
| 3–less than 5         | 46         | 2.48 (1.40 to 4.39)* | 2.54 (1.33 to 4.87)* |
| Above 5               | 44         | 2.31 (1.27 to 4.20)* | 2.04 (1.02 to 4.08)* |
| Province              |            |          |
| HCMC                  | 33         | Ref      |
| Ha Noi                | 50         | 1.98 (1.41 to 2.79)* | 2.30 (1.57 to 3.38)* |
| Hai Phong             | 51         | 2.08 (1.15 to 3.78)* | 2.25 (1.14 to 4.44)* |
| Khat Hoa              | 53         | 2.22 (0.88 to 5.63) | 1.96 (0.68 to 5.62) |
| Can Tho               | 40         | 1.33 (0.22 to 8.10) | 1.21 (0.19 to 7.63) |
| Da Nang               | 43         | 1.50 (0.51 to 4.43) | 3.18 (0.76 to 13.23) |
| Other                 | 31         | 0.91 (0.60 to 1.38) | 1.16 (0.72 to 1.87) |
| Internet use (days per month) |            |          |
| 1–7                   | 23         | Ref      |
| 8–14                  | 41         | 2.26 (1.03 to 4.95)* | 1.69 (0.68 to 4.21) |
| 15–21                 | 49         | 3.17 (1.57 to 6.40)* | 2.06 (0.90 to 4.73) |
| 22–30                 | 39         | 2.07 (1.09 to 3.95)* | 1.51 (0.69 to 3.32) |
| Partner preference    |            |          |
| Prefer only men       | 45         | Ref      |
| Prefer men to women   | 43         | 0.90 (0.64 to 1.28) | 0.77 (0.53 to 1.11) |
| Prefer women to men   | 18         | 0.26 (0.07 to 0.91)* | 0.18 (0.05 to 0.66)* |
| Prefer only women     | 2          | 0.02 (0.00 to 0.16)* | 0.02 (0.00 to 0.17)* |

*p<0.05 (not adjusted for multiple comparisons).
CI, confidence interval; HCMC, Ho Chi Minh City; UAS, unprotected anal sex.
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. Poisson regression was performed to investigate the correlation between the number of anal sex partners and number of UAS acts. As above, these analyses included the variables education and frequency of Internet use, which were associated to network size.

Modelling
The RDS-adjusted sample was assumed to provide a representative frequency distribution of ego-network types in the population. A weighted configuration model\(^3\) was used to model the connections between the ego-networks in the population. No sexual connections were thus assumed to exist within the recruitment network. Using the weighted configuration model, analytical estimates of the basic reproductive number\(^3\) was calculated for varying per-act transmission probabilities. Note that the basic reproductive number in this paper, henceforth denoted \(R_0^{3M}\), refers to a 3-month period instead of the entire infectious period. A 3-month period was chosen to minimise recollection bias and to approximate the highly infectious primary infection period. The basic reproductive number is an abstract measure of transmission propensity after introduction of HIV in a previously non-infected population. The actual reproductive number at a given time and area will be dependent on the proportion already infected in that area. The weighted configuration model is a development of the standard configuration model\(^3\) (see online supplementary figure S2 and SI section four, for details). In the standard model a sexual network is recreated from egonetwork 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 thus connects links with the same UAS frequency. This method produces a more realistic sexual network compared to what is achieved using the standard configuration model (see online supplementary figure S2). It is important to keep in mind that the method will not correctly model the sexual network if UAS frequency distributions differ between spatially distinct groups.

For men with more than three partners we imputed data on the number of UAS acts for partner four and above, before estimating \(R_0^{3M}\) 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 therefore used two different imputation strategies

Figure 1  Probability density distributions for men reporting UAS last three months. a) Number of partners last three months; b) total number of UAS acts with three most recent partners last three months. Insets show corresponding cumulative log-log plots.

Figure 2  a) Total number of UAS acts last three months stratified on the number of UAS partners reported for the same period; b) Number of UAS acts per partner, last three months stratified on the number of UAS partners reported for the same period (RDSII-unadjusted).
(a low and a high sexual frequency scenario) for partners four and upwards. We imputed for each individual with >3 UAS partners, in the low frequency scenario, one UAS act per partnership four and upwards (the logically smallest possible value) and for the high frequency scenario the smallest reported number of UAS acts among his three partners for which UAS act data were reported. To investigate the transmission effects stemming from correlations between partner numbers and frequency of UAS acts, we compared R₀ calculated from the original data (ie, with correlations present) and when the values for the number of UAS acts per partnership had been randomly shuffled between all partnerships (ie, with no correlations present). This latter way of modelling transmission on the network is equivalent to using the standard configuration model. Both shuffled and non-shuffled estimations were performed for the low and high sexual frequency scenarios.

**Ethics**

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 communication between the users and the server was encrypted.

**RESULTS**

**Sociodemographic characteristics**

The mean age was 24 years (CI 23.0 to 24.4). Eighty-one per cent had some education after high-school and 53% used the Internet 22 days or more per month. Seventy per cent came from the two large metropolitan areas of HCMC and Ha Noi. Forty-two of 63 provinces were represented (table 2).

**Sexual risk behaviour**

Sixty-two per cent (CI 54% to 65%) of the men reported sex (any type) at least once with another man during...
the last 3 months. Fifty-five per cent (CI 50% to 60%) reported anal sex at least once and 36% (CI 32% to 42%) had UAS at least once during the same period (see online supplementary figure S3). Among the men in the latter group, only 40% (CI 32% to 47%) had ever tested themselves for HIV.

Altogether, the men in the sample reported 2291 UAS acts (non-RDSII adjusted). Almost half of all UAS acts (45%; CI 44% to 49%) involved a respondent who had never tested for HIV, while almost all acts (92%; CI 91% to 93%) were reported to have taken place in non-casual relationships.

**Associations between sexual risk behaviour and sociodemographic factors**

Having a monthly income over one million VND (US $48) was associated with significantly increased odds of reporting UAS, as did living in two large provinces in the north, Ha Noi and Hai Phong. Preferring women to men or only women as sexual partners was associated with a significant decrease in the odds of reporting UAS (table 1).

Repeating the multivariate analyses with the outcome two or more UAS partners last 3 months (HCMC as reference) showed increased ORs for three large provinces: Ha Noi (OR=4.0, p<0.000), Hai Phong (OR=4.7, p=0.002) and Khang Hoa (OR=5.2, p=0.018).

**Number of UAS partners and acts**

In the group of MSM reporting UAS during the last 3 months, 88% reported having had UAS with one or two male partners (71% and 18%, respectively). The distribution of UAS partner numbers was highly skewed with 5% reporting five or more UAS partners (figure 1A, see online supplementary figure S3). The distribution of partner numbers when including all types of sex (with and without condom) was similarly right skewed (see online supplementary figure S4).

For those reporting UAS during the last 3 months the median number of UAS acts was four. The distribution of the number of UAS acts per person during the last 3 months was even more right skewed than the distribution of partner numbers (figure 1B). Forty-six per cent of those practicing UAS reported five or more UAS acts.

**Relationship between number of UAS partners and number of UAS acts**

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 2A). 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; figure 2B). 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, all during the last 3 months.

The Pearson correlation between the number of UAS acts within each man’s reported partnerships was significantly positive (+0.32), indicating that men with multiple partnerships who reported many acts in one partnership, tended to do the same for his other partnerships.

**Estimating population-level transmission**

The results of modelling transmission, using the weighted configuration model, show that both the low and high sexual frequency scenarios (see figure 3, top and bottom green and brown lines, assuming different numbers of UAS acts for missing data) produced very high R0SM. Results are, as expected heavily dependent on assumed values of per-act transmission probabilities.

To contextualise these values, estimates from current transmission studies during the asymptomatic stage show average per-act transmission risk of 0.6% for insertive sex (uncircumcised men) and 1.4% for receptive anal sex. 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 to 34.6).

The low sexual frequency scenario, that very likely underestimates UAS frequency, yielded an R0SM of one already at a per act transmission probability of 3.3%. Assuming a per-act transmission probability of 5% (e.g., for UAS during primary infection) yielded, in the worst-case scenario, an R0SM of 9 for this 3-month period.

We then investigated the contribution to transmission stemming from the observed positive correlation between partner numbers and numbers of UAS acts per partner. The results show that the high R0SM remains after shuffling and is thus not an artefact of the weighted configuration model. Additionally, R0SM 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 R0SM that we observe when we move from the low sexual frequency scenario (moving from green filled line to brown filled line), is due to the formation of a network structure that greatly facilitates transmission.

**DISCUSSION**

More than one-third of Internet-using Vietnamese MSM reported UAS at least once during the last 3 months. Among these men only 36% had ever tested themselves for HIV. Additionally we show, to the best of our knowledge, a previously undocumented positive correlation between UAS frequency and UAS partner numbers among MSM. This is clearly different from the trends reported in studies of heterosexual persons, where persons with many sex partners tend to report fewer sex acts thus reducing their contribution to population-level
transmission. Our modelling indicates that the positive correlations between partner numbers and unprotected sex acts found in this study can fuel very rapid HIV transmission.

Our results highlight the significance of collecting data on UAS frequency per partner to more accurately estimate 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 modelling shows how 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 results overall highlight the importance of rapidly scaling-up HIV prevention programmes for MSM in Vietnam. Harm reduction programmes for MSM are concentrated to the big cities but our study highlights the need to stretch these nationwide. In this lower-middle-income setting, the overwhelming number of UAS acts among participating MSM, takes place with partners whom the respondent is emotionally involved with. This is consistent with studies from Amsterdam where an estimated 86% of new infections occurred in stable relationships during the early 2000s. These results imply that condom campaigns focusing on casual sexual encounters are not enough. Interventions and prevention communication need to take into account the complexities of sexual practices within relationships.

Right-skewed frequency distributions of the number of anal sex acts and anal sex partners have been reported elsewhere. Our findings indicate however that this was also the case for UAS acts and UAS partnerships and there was a positive correlation between the two. These findings imply the need for local prevention programmes to target high-risk individuals within the MSM population as they may contribute to a substantial share of new infections. These high-risk men come however from diverse backgrounds. Men reporting more than five UAS acts last 3 months and at least two different UAS partners during the same period, all preferred only men or men to women as sexual partners. These men were also older and had better incomes, although these differences were not statistically significant. Frequent Internet users reported more UAS than infrequent users, potentially indicating a window of opportunity for Internet-based prevention strategies. Contact tracing and potentially peer-driven interventions may be efficient way to reach these individuals, but should be evaluated in the local context. Pre-exposure prophylaxis and treatment-as-prevention strategies may have a role but requires that access barriers to testing and care are removed and that high-risk individuals can be targeted. Our data show low testing rates and earlier studies suggest a high level of stigma attached to both identifying as gay and to being associated with HIV/AIDS, indicating a wider need for structural interventions in addition to biomedical approaches.

Figure 3  R03M for different per act transmission probabilities. Top yellow line: high-transmission scenario. Green bottom line: low-transmission scenario. Yellow and green dashed lines: correlations between frequency of UAS acts and partner numbers removed (shuffled UAS frequencies). Grey horizontal dashed line shows where R03M = 1. Vertical grey lines indicate per-act probabilities of 0.6% (insertive anal sex asymptomatic period) and 1.4% (receptive anal sex asymptomatic period). Per-act probability during primary infection period (18%) is outside the figure.
recent proposal by the Minister of Justice to legalise same-sex marriage in Vietnam is particularly laudable.

This study has limitations. Data on receptive and insertive UAS acts would have added further realism. In 48% of partnerships the first UAS act within the partnership took place less than 3 months 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 analysed, as was carried out in this study. Over the very long-time frame it is likely that the fast transmission in the beginning of the epidemic turns much slower at later stages, as Colgate et al suggest. The modelling performed here thus describes estimated transmission during early epidemic stage. Modelling the long-term evolution of the epidemic requires additional data sources, including, for example, inflow of non-infected individuals to the population and estimated effects of interventions. The model contains a relatively homogenous network structure, while real networks are likely to contain spatially distinct clusters. However, as long as the frequency distribution of ego-network types are similar in geographically distinct network clusters, very similar results would likely apply. Transmission into new clusters combined with sexual behaviour patterns such as those reported here can potentially also explain observed clustering in HIV sequencing studies. Poisson regression of UAS frequency and partner numbers, may contain bias if intra-personal dependencies between sexual acts are substantial. Figure 2A,B provides the actual distributions.

Our survey sampled mostly young and relatively well-educated MSM in the major Vietnamese cities and should not be generalised to non-Internet using MSM. However, compared to many self-selected online samples with response rates of a few in a thousand to a few in a hundred, the response rate in this study is likely to have been at substantially higher, as RDS recruitment stops if participants on average recruit less than one other person with his four invitation coupons (exact participation rates in RDS is not possible to determine as researchers do not know the total number of invited individuals). For methodological issues related to webRDS among MSM see, for example, refs. In comparison with a standard coupon-based RDS, the present study enabled a much larger geographical coverage and allowed participation of MSM who would not risk to physically access a survey office.

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

Author affiliations
1Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden
2The Institute for Futures Studies, Stockholm, Sweden
3College of Information System and Management, National University of Defense Technology, Changsha, People’s Republic of China
4Department of Sociology, Stockholm University, Stockholm, Sweden
5Institute for the Study of Society, Economy and the Environment, Ha Noi, Vietnam

Acknowledgements The authors would like to thank iSEE, Vietnam for dedicated support in the data collection.

Contributors All authors participated in research design. LB, XL and HHT collected the data. LB and XL performed analyses. LB wrote the first draft. All authors approved the final manuscript.

Funding The Swedish International Development Cooperation Agency (Sida) supported the study financially.

Competing interests None.

Ethics approval The study was approved by the ethics review board at Hanoi Medical University, Vietnam and the Regional Ethical Review Board for Stockholm (EPN).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES

1. Baral S, Sifakis F, Cleghorn F, 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:e339.
2. Smith AD, Tapsoba P, Peshu N, et al. Men who have sex with men and HIV/AIDS in sub-Saharan Africa. Lancet 2009;374:416–22.
3. Mumtaz G, Hilmi N, McFarland W, 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:e1000444.
4. Beyrer C, Baral SD, van Griensven F, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet 2012;380:367–77.
5. van Griensven F, van Wijngaarden JWD. A review of the epidemiology of HIV infection and prevention responses among MSM in Asia. AIDS 2010;24:S30–40.
6. van Griensven F, Varangrat A, Wimonsate W, et al. Trends in HIV prevalence, estimated HIV incidence, and risk behavior among men who have sex with men in Bangkok, Thailand, 2003–2007. J Acquir Immune Defic Syndr 2009;53:234–9.
7. Ministry of Health Vietnam. Integrated biological and behavioral surveillance round 2–2009. Hanoi: Ministry of Health Vietnam, 2011.
8. Ministry of Health Vietnam. Results from the HIV/STI Integrated Biological and Behavioral Surveillance (IBBS) in Vietnam 2005– 2006: a systematic review. WHO Regional Office for South-East Asia: e77. – 2006: a systematic review. PLoS Med 2007;4:e339.
9. Magnani R, Sabin K, Saida L, et al. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS 2005;19(Suppl 2):S67–72.
10. Volk JE, Liu A, Vittinghoff 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;61:112–15.
11. The Commission on AIDS in Asia. Redefining AIDS in Asia. Crafting an Effective Response. New Delhi; 2008.
12. Crawford JM, Kippax SC, Mao L, 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;10:325–31.
15. Kobrill BA, Chesney MA, Husnik MJ, et al. High-risk behaviors among men who have sex with men in 6 US cities: baseline data from the EXPORE Study. Am J Public Health 2003;93:926–32.
16. van Griensven F, Thienkrua W, Sukwicha W, 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. Sex Transm Dis 2006;33:342–7.
17. Lijeros F. Human sexual networks. Encyclopedia of Complexity and System Science. Springer Science and Business Media, 2009.
18. Newman ME. Assortative mixing in networks. Phys Rev Lett 2002;89:208701.
19. Newman ME. Mixing patterns in networks. Phys Rev E Stat Nonlin Soft Matter Phys 2003;67:026126.
20. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Prob1997;44:174–99.
21. Malekinejad M, Johnston LG, Kendall C, et al. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. AIDS Behav 2008;12:S105–30.
22. Bengtsson L, Lu X, Nguyen QC, et al. Implementation of web-based sexual-risk behavior sampling among men who have sex with men in Vietnam. PLoS ONE 2012;7:e49417.
23. Internet World Stats. Vietnam. Internet Usage Stats and Marketing Report, In; 2012.
24. Ngo DA, Ross MW, Phan H, et al. Male homosexual identities, relationship practices among young men who have sex with men in Vietnam: implications for HIV prevention. AIDS Educ Prev 2009;21:251–65.
25. Nguyen QC. Sexual risk behaviors of men who have sex with men in Vietnam. Chapel Hill: North Carolina State University, 2010.
26. Bengtsson L, Thorson A, Than VN, et al. Sexual relationships among men who have sex with men in urban Vietnam: a qualitative interview study. BMC Public Health 2013; 13:108.
27. Heckathorn D, Semaan S, Broadhead R, et al. Extensions of respondent-driven sampling: a new approach to the study of injection drug users aged 18–25. AIDS Behav 2002;6:55–67.
28. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. J Urban Health Bull N Y Acad Med 2006;83:198–112.
29. Wejnert C, Pham H, Krishna N, 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;16:797–806.
30. Ramirez-Valles J, Garcia D, Campbell RT, et al. HIV infection, sexual risk behavior, and substance use among Latino gay and bisexual men and transgender persons. Am J Public Health 2008;98:1036–42.
31. Britton T, Deijen M, Lijeros F. A weighted configuration model and inhomogeneous epidemics. J Stat Phys 2011;145:1368–84.
32. Diekmann O, Heesterbeek JA, Metz JA. On the definition and the calculation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. J Math Biol 1990;28:365–82.
33. Andersson H. Limit theorems for a random graph epidemic model. Ann Appl Probability 1998:1331–49.
34. Jin F, Jansson J, Law M, et al. Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART. AIDS 2010;24:307–13.
35. Baggaley RF, White RG, Boly M-C. HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention. Int J Epidemiol 2010;39:1048–63.
36. Leynaert B, Downs AM, de Vincenzi I. Heterosexual transmission of human immunodeficiency virus variability of infectivity throughout the course of infection. Am J Epidemiol 1998;148:88–96.
37. Nordvik MK, Lijeros F. Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. Sex Transm Dis 2006;33:342–9.
38. Morris M, Epstein H, Wawer M. Timing is everything: international variations in historical sexual partnership concurrency and HIV prevalence. PLoS ONE 2010;5:e14092.
39. Harrison A, Cledand J, Frohlich J. Young people’s sexual partnerships in KwaZulu-Natal, South Africa: patterns, contextual influences and HIV risk. Stud Fam Plann 2008;39:295–308.
40. Blower SM, Boe C. Sex acts, sex partners, and sex budgets: implications for risk factor analysis and estimation of HIV transmission probabilities. J Acquir Immune Defic Syndr 1996;6:1347–52.
41. Xirdou M, Geskusa R, de Wit J, et al. The contribution of steady and casual partnerships to the incidence of HIV infection among homosexual men in Amsterdam. AIDS 2003;17:1029–38.
42. Schneeberger A, Mercer CH, Gregson SAJ, et al. Scale-free networks and sexually transmitted diseases—a description of observed patterns of sexual contacts in Britain and Zimbabwe. Sex Trans Dis 2004;31:380–7.
43. Lin H, He N, Ding Y, et al. Tracing sexual contacts of HIV-infected individuals in a rural prefecture, Eastern China. BMC Public Health 2012;12:505.
44. Zhang H, Wu Z, Zheng Y, et al. A pilot intervention to increase condom use and HIV testing and counseling among men who have sex with men in Anhui, China. AIDS J Acquir Immune Defic Syndr 2010;53:S88–92.
45. Amirkhanian YA, Kelly JA, Kabakchieva E, et al. A randomized social network HIV prevention trial with young men who have sex with men in Russia and Bulgaria. AIDS 2005;19:1897–905.
46. Amirkhanian YA, Kelly JA, Kabakchieva E, 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;15:205–20.
47. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365: 509–20.
48. Van Damme L, Cornell M, Ahmed K, et al. Preexposure prophylaxis for HIV infection among African women. N Engl J Med 2012;367: 411–22.
49. Wilson D. HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med 2012;9:e1001231.
50. Baltiel AD, Freedberg KA, Scott CA, et al. HIV preexposure prophylaxis in the United States: impact on lifetime infection risk, clinical outcomes, and cost-effectiveness. Clin Infect Dis 2014;58:1763–74.
51. USA Today. Vietnam considers same-sex marriage. 2012.
52. Colgate SA, Stanley EA, Hyman JM, et al. Risk behavior-based model of the cubic growth of acquired immunodeficiency syndrome in the United-States. Proc Natl Acad Sci USA 1989;86:4793–7.
53. Lewis F, Hughes B, Thomas D, et al. Transmission of HIV related to sexual activity in Sydney, Australia. Trans Dis 1997;44:174–82.
54. Brenner BG, Roger M, Stephens D, 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;204:1115–19.
55. Jakopanec I, Schimmer B, Grijbovski AM, 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 Infect Dis 2010;10:261.
56. Zhang D, Bi P, Lv F, et al. Internet use and risk behaviours: an online survey of visitors to three gay websites in China. Sex Transm Infect 2007;83:571–6.
57. Lu X, Malmros J, Lijeros F, et al. Respondent-driven sampling on directed networks. Electron J Stat 2013;7:329–322.

Bengtsson L, Lu X, Lijeros F, et al. BMJ Open 2014;4:e003526. doi:10.1136/bmjopen-2013-003526.
Supporting Material

High Sexual Risk Behavior Indicates Explosive HIV Transmission among Men who have Sex with Men in Vietnam

Authors: Linus Bengtsson\textsuperscript{a,b}, Xin Lu\textsuperscript{a,b,c,d}, Fredrik Liljeros\textsuperscript{b,c}, Hoang Huy Thanh\textsuperscript{d}, Anna Thorson\textsuperscript{a}

\textsuperscript{a}Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden;
\textsuperscript{b}The Institute for Futures Studies, Stockholm, Sweden;
\textsuperscript{c}Department of Sociology, Stockholm University, Stockholm, Sweden;
\textsuperscript{d}College of Information Systems and Management, National University of Defense Technology, Changsha, China;
\textsuperscript{e}Institute for the Study of Society, Economy and the Environment, Ha Noi, Vietnam.

Corresponding author:
Linus Bengtsson
Dept. of Public Health Sciences
Karolinska Institute
171 77 Stockholm
Sweden
linus.bengtsson@ki.se
Supporting Material

1. Supporting figures

Figure S1: Recruitment trees. Each color represents a separate recruitment tree originating from a single seed. Three seeds did not generate further participant
Figure S2: Recreating the sexual network based on individually reported partnerships. Red balls (persons) are connected to each other with thick black links (partnerships with many UAS acts) or by thin grey links (partnerships with few UAS acts); a) The real network contains six persons in stable partnerships and eight individuals with several one-night stands; b) Partner data gathered from the 14 respondents in the network; c) In the standard configuration model the network is re-created by randomly connecting links; d) 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. In the example the method correctly recreates a network with three stable, monogamous couples, disconnected from the rest of the network.
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 bar chart shows the proportion of these UAS acts that took place with non-casual (92%) and casual partners (7%). Within both categories separate numbers are shown for acts in which the respondent had ever tested for HIV vs. tested for HIV at least once. E.g. four percent of all reported UAS acts were reported to have taken place with casual partners by respondents who had never tested for HIV. Due to rounding, the four groups in 2b sum to 99%.

Figure S3: Proportions of all MSM practicing different sexual behaviors last three months.
Figure S4: Distribution of anal sex partners during the past three months (with and without condom).
2. Supporting table

In addition to the checks in the main paper we checked all surveys in the cleaned sample for other signs of duplication or presence of non-valid data. We flagged surveys containing a repeated IP number, short completion times or other deviations (as described below). IP numbers are generally shared by all customers at an Internet café, which are common in Vietnam as many people do not own their own computer. Repeated IP-numbers could thus indicate repeated submission by the same person but also participation from an Internet café catering to MSM. We analyzed the sensitivity of the estimates to inclusion and exclusion of these flagged submissions. Specifically, we compared the RDS estimates generated from the cleaned sample with the RDS estimates generated from groups with progressively stricter inclusion criteria. These groups were generated as follows:

**Full sample (not shown)**
All seeds and non-seed respondents

**Cleaned sample**
- All non-seed respondents with valid age (≥18 years);
- Exclusion of 4.3% of surveys, which had not provided answers to the socio-demographic items;
- Exclusion of submissions with an email, Yahoo! Chat ID or telephone number, which had previously been registered in the system;

**Strict sample**
As the cleaned sample, but additionally excluding submissions with an IP address that had previously been registered in the system;

**Extreme:**
As the strict sample but additionally excluding submissions with short completion times (< 3 minutes) and submissions stating no education (rare in Vietnam).

RDS estimates of socio-demographic variables and sexual partner preferences for the above groups with progressively stricter inclusion criteria are listed in Table S1. The differences between groups are generally small but noticeable differences are present for province and Internet-use. The main change between in these two estimates takes place when we remove participants with a repeated IP address (moving from the cleaned to the strict sample). A smaller proportion of the strict sample consists of persons with infrequent Internet-use and of people from outside the large cities. This is also what we would expect if we removed persons who accessed the survey from Internet cafés.

The average absolute differences in proportional estimates when comparing the cleaned sample to the two other groups, is three percent (maximum difference 21%), see Table S1.

---

1 This is the cleaned sample discussed in the paper.
We checked selected analyzes (main paper: Table 2, Figure 1 and Figure 2), using the different subgroups and there were no important changes to the reported results.

Table S1. Descriptive statistics of socio-demographic variables and sexual partner preferences for samples with progressively stricter inclusion criteria

| Category | Correlation | (loosely sample n=551) | (strict n=399) | (extreme n=592) |
|----------|-------------|------------------------|---------------|-----------------|
| Age      |             |                        |               |                 |
| 18-20    | 210         | 210                    | 210           | 210             |
| 21-22    | 200         | 200                    | 200           | 200             |
| 23-26    | 200         | 200                    | 200           | 200             |
| 26-30    | 100         | 100                    | 100           | 100             |
| >30      | 90          | 90                     | 90            | 90              |
| Don't want to answer | 0         | 0                      | 0             | 0               |

| Education | (highest level) |                        |               |                 |
|-----------|-----------------|------------------------|---------------|                 |
| 12 years or less | 117         | 117                    | 117           | 117             |
| 3-year college | 63          | 63                     | 63            | 63              |
| Bachelor level, univ | 156        | 156                    | 156           | 156             |
| Master's/doctoral level | 34        | 34                     | 34            | 34              |
| Don't want to answer | 5         | 5                      | 5             | 5               |

| Income | (million VND) |                        |               |                 |
|--------|----------------|------------------------|---------------|                 |
| 0- less than | 87         | 87                     | 87            | 87              |
| 1- less than | 355        | 355                    | 355           | 355             |
| 3-less than | 206        | 206                    | 206           | 206             |
| Above 5 | 154         | 154                    | 154           | 154             |
| Don't want to answer | 30        | 30                     | 30            | 30              |

| Province |                        |               |                 |                 |
|----------|------------------------|---------------|-----------------|-----------------|
| HCM      | 339                    | 339           | 339             | 339             |
| Ha Noi   | 256                    | 256           | 256             | 256             |
| Hai Phong | 55          | 55                     | 55             | 55              |
| Khanh Hoa | 20          | 20                     | 20             | 20              |
| Can Tho  | 5                      | 5                      | 5              | 5               |
| Da Nang  | 15                     | 15                     | 15             | 15              |
| Other    | 160                    | 160           | 160             | 160             |
| Don't want to answer | 30        | 30                     | 30            | 30              |

| Internet use | (days/month) |                        |               |                 |
|--------------|--------------|------------------------|---------------|                 |
| 1-3 | 84          | 84                     | 84             | 84              |
| 4-14 | 72          | 72                     | 72             | 72              |
| 15-21 | 190         | 190                    | 190            | 190             |
| 22-30 | 559         | 559                    | 559            | 559             |
| Don't want to answer | 12        | 12                     | 12            | 12              |

| Partner Preference |                        |               |                 |                 |
|--------------------|------------------------|---------------|-----------------|-----------------|
| Prefer only men   | 544                    | 544           | 544             | 544             |
| Prefer men to women | 181        | 181                     | 181            | 181             |
| Prefer women to men | 20         | 20                      | 20             | 20              |
| Prefer only women | 56          | 56                      | 56             | 56              |
| Don't want to answer | 38        | 38                     | 38            | 38              |

3. Sample equilibrium analysis

To get an overview of whether the sample reached equilibrium, we plotted both the RDSII estimates along with the increased sample size for most variables surveyed in this study (see Figure S5). During the last 200 respondents, the changes of estimates for all variables are very small. Let $\Delta p_{RDSII} = |p_{RDSII}^n - p_{RDSII}^{n-200}|$ be the absolute change of estimates for the variable, on average the difference is only 0.87% for proportional estimates with a maximum of 3.6%. For numeric estimates, the maximum difference is only 0.02.
4. Weighted configuration model for R0

The weighted configuration model by Britton et al [1] provides means for analysis infectious disease transmission on weighted networks. More specifically, given a degree distribution with a degree-weight dependency, their approach allows an analytical solution for the basic reproduction number, R0, which is defined as a function of the model parameters such that the probability for a large outbreak to occur with any initial infection in the network is positive if and only if R0>0.

Let $P(D = d) = p(d)$ be the probability distribution for nodes with degree $d$ in the network, and let $P(W_j = w | D_i = d) = q(w | d), j = 1, \ldots, d$ be the distribution of edge weights associated with a node of degree $d$. Assuming that the network is formed by connecting links with the same weight at random (see also Fig 1a-d, main paper), R0 is given by the largest eigenvalue of the mean offspring matrix from the theory of multi-type branching process:

$$M_2 = (m_{jk})_{d,k} \geq 2$$
where $m_{dk}$ is the element of the mean offspring matrix, which is the expected number of $k$-nodes that an infected $d$-node infects:

$$m_{dk} = (d-1) \sum_w \pi(w)q(w|d)\tilde{p}_w(k)$$

where $\pi(w) = 1 - (1 - s)^w$ is the infection probability for a connection to an infected node with intensity $w$, and $\tilde{p}_w(k) = \frac{q(w|k)kp(k)}{\sum_j q(w|j)jp(j)}$

Note that the definition of $R_0$ we used in this paper (denoted $R_0^3M$), is slightly different from the standard epidemiological definition, which refer to the average number of new infections over an infected person’s entire infectious period, while we in this paper use a three-month period.

**Reference**

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