UCLA
UCLA Previously Published Works

Title
The complex interplay of social networks, geography and HIV risk among Malaysian Drug Injectors: Results from respondent-driven sampling

Permalink
https://escholarship.org/uc/item/3b0282ms

Authors
Zelenev, Alexei
Long, Elisa
Bazazi, Alexander R
et al.

Publication Date
2016-11-01

DOI
10.1016/j.drugpo.2016.08.008

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Int J Drug Policy. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

Int J Drug Policy. 2016 November ; 37: 98–106. doi:10.1016/j.drugpo.2016.08.008.

The complex interplay of social networks, geography and HIV risk among Malaysian Drug Injectors: Results from Respondent-Driven Sampling

Alexei Zelenev, Ph.D.[1], **Elisa Long, Ph.D.**[2], **Alexander R. Bazazi, M.Phil.**[1,3], **Adeeba Kamarulzaman, FRACP**[1,4], and **Frederick L. Altice, M.D., M.A.**[1,3,4]

[1]Department of Internal Medicine, Section of Infectious Diseases, AIDS Program, Yale School of Medicine. 135 College St., Suite 323, New Haven, CT, 06510, USA

[2]UCLA Anderson School of Management, Decisions, Operations & Technology Management Department, Los Angeles, CA, USA

[3]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

[4]Centre of Excellence for Research in AIDS (CERiA), Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

Abstract

Background—HIV is primarily concentrated among people who inject drugs (PWID) in Malaysia, where currently HIV prevention and treatment coverage is inadequate. To improve the targeting of interventions, we examined HIV clustering and the role that social networks and geographical distance plays in influencing HIV transmission among PWID.

Methods—Data were derived from a respondent-driven survey sample (RDS) collected during 2010 of 460 PWID in greater Kuala Lumpur. Analysis focused on socio-demographic, clinical, behavioral, and network information. Spatial probit models were developed based on a distinction between the influence of peers (individuals nominated through a recruitment network) and neighbors (residing a close distance to the individual). The models were expanded to account for the potential influence of the network formation.

Results—Recruitment patterns of HIV-infected PWID clustered both spatially and across the recruitment networks. In addition, HIV-infected PWID were more likely to have peers and neighbors who were HIV-infected and lived nearby (<5 km), more likely to have been previously incarcerated, less likely to use clean needles (26.8% vs 53.0% of the reported injections, p<0.01), and have fewer recent injection partners (2.4 vs 5.4, p<0.01). The association between the HIV

Contact: Alexei Zelenev, Ph.D., **Address**: 135 College Street, New Haven, CT, USA 06510-2283, alexei.zelenev@yale.edu, **Phone**: +1.203.737-7907, **Facsimile**: +1.203.737.4051.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

COI: All Authors declare that they do not have any conflict of interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

status of peers and neighbors remained significantly correlated even after controlling for unobserved variation related to network formation and sero-sorting.

Conclusions—The relationship between HIV status across networks and space in Kuala Lumpur underscores the importance of these factors for surveillance and prevention strategies, and this needs to be more closely integrated. RDS can be applied to identify injection network structures, and this provides an important mechanism for improving public health surveillance, accessing high-risk populations, and implementing risk-reduction interventions to slow HIV transmission.

Keywords

HIV; Malaysia; Social Networks; Geography; Respondent-Driven Sampling; Addiction; People who inject drugs (PWID)

Introduction

With over 36.7 million people infected worldwide and 1.1 million deaths in 2015 alone, the HIV pandemic is the one of the most significant public health challenges of the 21st century (Joint United Nations Programme on HIV/AIDS (UNAIDS), 2016). Many countries struggle with developing and implementing effective HIV prevention and treatment strategies that target high-risk and hidden populations, including people who inject drugs (PWID), sex workers, transgender women and men who have sex with men (MSM). Illicit drug use, in particular, has a profound effect on the global burden of disease: among the 12 million PWID globally (United Nations Office on Drugs and Crime (UNODC), 2016), injection drug use as a risk factor for HIV accounts for 2.1 million Disability Life Adjusted Years (DALYs) (L. Degenhardt, Hall, W, 2012; L. Degenhardt, Whiteford, HA, Ferrari, AJ, Baxter, AJ, Charlson, FJ, Hall, WD, Freedman, G, Burstein, R, Johns, N, Engell, RE, Flaxman, A, Murray, CJ, Vos, T, 2013). Even in concentrated HIV epidemics, where total HIV prevalence in the population is <1%, effective prevention strategies are needed due to the salience of the "bridging ties" that create opportunities for HIV transmission from high-risk individuals to the lower-risk general population, increasing the odds that the HIV epidemic may become generalized (Doherty, 2006).

Bio-behavioral surveillance studies are often used to assess HIV prevalence and risk-behaviors in high-risk, hidden populations, and typically rely on either respondent-driven sampling (RDS) or time-space venue-based sampling (Kendall, 2008; Magnani, Sabin, Saidel, & Heckathorn, 2005) recruitment strategies. Each method, however, is fraught with challenges that undermine its ability to represent the intended population. Such limitations include non-response and selection bias due to differential recruitment (Amber, 2011), homophily (Mills, 2012), variability in geographical location (Bazazi, Crawford, et al., 2015; McCreesh, 2012; Toledo, 2011), and seed selection (i.e., who is initially recruited) (Heimer, 2005). Despite these methodological limitations, RDS remains a primary recruitment strategy for PWID by international public health authorities due to its efficiency in reaching hidden populations (Centers for Disease Control and Prevention, 2008; Goel, 2010; Malekinejad M, 2008).

Malaysia, a polycultural Southeast Asian country with a population of over 30 million, is home to an estimated 200,000 PWID, most of whom inject opioids (Bachireddy, et al., 2011; United Nations Office on Drugs and Crime (UNODC), 2016). HIV was primarily concentrated in PWID and HIV prevention and treatment efforts first focused on the introduction of needle/syringe exchange programs (NSPs) and in 2006 opioid agonist therapies (OAT) with buprenorphine and methadone (Kamarulzaman, 2009; Reid, Kamarulzaman, & Sran, 2007). Though there is nascent evidence of an emerging transitional epidemic, including transmission from PWID to their heterosexual partners (Ministry of Health & Malaysia, 2014; UNGASS, 2010), the majority of people living with HIV (PLH) are PWID. Yet, HIV prevention and treatment in Malaysia remains inadequately scaled to need (L. Degenhardt, et al., 2014; Kamarulzaman, 2009) with preventive measures reaching only a small fraction of the most-at-risk populations (Reid, et al., 2007). Based on recent 2013 surveillance data, Malaysia had a cumulative number of more than 100,000 HIV cases, including more than 85,000 PLH and more than 16,000 deaths related to HIV/AIDS (Ministry of Health & Malaysia, 2014).

In 2010, we conducted a bio-behavioral surveillance study in greater Kuala Lumpur using RDS to recruit opioid-dependent PWID (Bazazi, Crawford, et al., 2015; Bazazi, Zelenev, et al., 2015b). We analyzed how the spatial proximity of PWID to their peer network, influence HIV status and HIV risk behaviors in order to: a) inform improvements in sampling methods and b) guide the discussion for designing more optimal prevention strategies. Previous studies have demonstrated the importance of networks (S. Friedman, Curtis, Neaigus, Jose, & Des Jarlais, 2002; S. Friedman, Neaigus, A, Jose, B, Curtis, R, Goldstein, M, Ildefonso, G, Rothenberg, RB, Des Jarlais, DC, 1997; Latkin, Forman, Knowlton, & Sherman, 2003; Mustanski, 2014; R. Rothenberg, Long, DM, Sterk, CE, Pach, A, Potterat, JJ, Muth, S, Baldwin, JA, Trotter, RT 3rd 2000) for HIV transmission and geography for recruitment of populations most-at-risk for HIV (Jenness, 2014; R. Rothenberg, Muth, SQ, Malone, S, Potterat, JJ, Woodhouse, DE, 2005; Toledo, 2011), yet none of these studies have accounted for the influence of the network formation process, which can induce a non-causal pattern of observed correlations in the HIV outcomes. Findings from these analyses are relevant for future interventions that aim to target individuals most-at-risk and explore the potential for incorporating network, structural and spatial strategies in reducing HIV transmission.

Methods

Study Design and Recruitment

Recruitment methods have been previously described (Bazazi, Zelenev, et al., 2015a). In brief, from July to October in 2010, 460 PWID were recruited using RDS to examine a cross-sectional assessment of drug use behaviors, risk factors and health outcomes associated with drug use. Eligibility criteria included: (1) age ≥18 years; (2) residing in greater Kuala Lumpur; (3) drug injection in the previous 30 days, confirmed by physical examination of injection track marks and/or knowledge of drug preparation methods; and (4) willingness to undergo rapid HIV testing and counseling and urine toxicology testing. While positive urine toxicology tests for opioids represent use in the past 2–3 days, to avoid encouraging drug use to gain access to the study, urine test results were not used to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

determine eligibility. Respondent-driven sampling (RDS), a form of chain-referral sampling designed to efficiently recruit hidden populations (Heckathorn, 1997), was operated from three geographically distinct research sites. Outreach workers from each interview site recruited six "seeds" as initial participants; two were HIV-infected. Each participant, including seeds, was encouraged to recruit up to 3 PWID from their social network (peers) and received RM50 ($16 US) for their participation and RM25 ($8 US) for each eligible peer recruited. Trained interviewers administered the questionnaires in Bahasa Malaysia and conducted pre/post HIV counseling and testing and subsequent referral to services. This study was approved by Institutional Review Boards at the University of Malaya and Yale University School of Medicine.

Study Definitions and Indicators

For each study participant, the primary outcome was HIV-seropositive status, defined dichotomously as reactive on an initial HIV rapid test (OraQuick ADVANCE® Rapid HIV-1/2, OraSure Technologies, Inc.) and confirmed by a second rapid HIV test (ACON HIV 1/2/0 Rapid Test Device, ACON Laboratories, Inc). No discordance between test results were observed. We included the following covariates: 1) age; 2) gender, as a dichotomous variable with female being a referent category; 3) race/ethnicity indicators were defined in terms of binary variables based on self-reported categories: Malay, Chinese and Indian; 4) "unstable housing" was defined using a dichotomous variable based on a self-described living situation in the preceding 30 days that included self-reported homelessness, street residence, shelters or temporary residence at a partner's place or with family/friends, as well as short-term boarding, whereas "stable housing" included living arrangements such as one's own place, and having permanent residence either with family, friends or a partner; 5) relationship status was defined using a dichotomous variable, in which being married or having a partner constituted a "stable relationship", while being single, widowed or separated was used to define "unstable relationship" and was used as a referent group; 6) "network size" was defined in terms of the number of injection drug users, who were 18 years or older and living in the Klang Valley, whom the respondent was acquainted with and had seen within the past 3 months, a time frame that was selected to reduce problems with length-biased sampling; 7) number of incarcerations was based on a self-reported number of times an individual has been to prison; and previously incarcerated was defined in terms of a dichotomous variable based on whether the respondent reported a positive number of incarcerations; 8) "number of injection partners" was defined as the reported number of individuals with whom the respondent had injected drugs in the past 30 days, number of sharing partners was defined as the number of individuals with whom the respondent had shared either needles or syringes; 9) "years of injection" was calculated as a difference between the self-reported age and the age of first injection; 10) "number of times injected any substance in the past 30 days" was based on self-report; 11) "number of days injected heroin in the past 30 days" was based on self-report; 12) "percent of the time a respondent used clean needles" was based on number of times that the respondent reported using new or unused needle or syringe in the past 30 days divided by the number of times that the respondent reported injecting in that period; based on this definition we created a dummy variable for the "propensity to use clean needles" if the respondent reporting using new or unused needles more than 25% of the time, which constituted the 60th percentile of the

Author Manuscript

"percent of the time" variable; 13) "place of injection" was based on which type of place the respondent typically injected drugs in the previous 6 months, and included a distinction between a "private residence" vs "public places"; 14) "awareness of the individual's HIV status" was deduced from whether the respondent reported to have been previously tested for HIV, and whether the results of the previous test have been different from the HIV rapid test; 15) each individual reported his or her residential neighborhood location (among 44 distinct neighborhood locations) , which were geo-coded and used to calculate distances (in kilometers) among the residential locations for all the respondents.

Statistical Analysis

Measuring Spatial and Social Effects—First, we compared HIV-positive and HIV-negative individuals for several covariates and used a validated overlapping block bootstrap method to test whether the differences between HIV groups were statistically significant at $p<0.05$ (Lahiri, 2003). The block bootstrap is a non-parametric simulation based method that accounts for dependence among observations stemming from a network-based sampling design and provides an improvement to the poor asymptotic approximation of other statistical tests. Second, we explored the influence of spatial (neighbors) and social network (peer) effects on HIV risk by estimating a series of probit models with auto-correlations in the form:

$$Y_i=\rho_1 W_1 Y_i+\rho_2 W_2 Y_2+X_i \beta+\varepsilon_i \quad (1)$$

where Y is the dependent variable (HIV status) for individual i, W_1 and W_2 are two different contiguity matrices and X is a vector of explanatory variables, ε is an error term and ρ_1 , ρ_2 and β are parameters to be estimated. The specification of this model implies that each value of the dependent variable is a function of the explanatory variables and a weighted average of the dependent variable of the "nearby" observations (Anselin, 1988; J. LeSage, Pace RK, 2009). In the first specification, we use W_1 and W_2 to measure recruitment (social) network of individuals who were residing within a close distance (neighbors, <5 km) and more remote distance (5–10 km), respectively. Since the original recruitment matrix is measured with error due to missing links (Lyons, 2011), the social network was expanded to include both first and second degree of contacts with intent to capture a "small world" effect that other researchers have found in different populations, including PWID (Amato, Davoli, & Ferri, 2001; Rudolph, 2013; Watts, 1998). A first-degree recruitment contact for individual *i* includes everyone that *i* recruited (and the person who recruited *i*), while a second degree recruitment contact includes all contacts of those individuals whom *i* recruited, as well as contacts of *i*'s recruiter. We refer to such direct contacts as "peers" (and to the second degree as "peers of peers"). Here, we found that our models' estimates were robust to different definitions of W that included additional degree contacts beyond the first degree. After analyzing network effects, we focused on neighbors and we redefined the congruity matrices to measure the proximity to those individuals who were not within the individual's first and second degree recruitment network, yet who resided within a certain distance: (<5km) for W_1 and (5–10 km) for W_2.

Modeling HIV Status and Network Formation—One challenge facing studies that rely on social network data is the ability to draw causal inference regarding social influence when individuals may sort into groups in non-random ways (Bramoulle, 2009; Manski, 1993; Topa, 2015). Specific non-randomness can take on many forms including homophily, selection of recruits based on similar risk behaviors, and sero-sorting, which can give rise to correlated, but non-causally related, outcomes. We attempt to control for potential sources of non-randomness by including an additional term z_i to account for unobserved factors that may influence both the HIV status and network formation among PWID. This method follows closely recent developments in econometric methodology (Goldsmith-Pinkham, 2013; Hsieh, 2014). Equation (1) becomes:

$$Y_i=\rho_1^F W_1^F Y_i+\rho_2^N W_1^N Y_i+X_i\beta+\tau z_i+v_i \quad (2a)$$

In this model, the contiguity matrices capture the effects of HIV status of both peers (W^F) and neighbors (W^N) based on the results from two of the previous specifications. X is a matrix of covariates, z_i is the unobserved "random effect" that is related to the existence of ties among respondents through equation (2b), v_i is a residual and $\rho_1^F, \rho_2^N \beta, \tau$ are parameters to be estimated. If τ is different from zero, this provides evidence that the network is endogenous (due to such potential factors as sero-sorting). In this analysis, it is crucial to test whether the parameters ρ_1^F, ρ_2^N will remain non-zero, once we control for unobserved factors linked to network formation. Simultaneously to HIV status, we modeled the network formation process by estimating the probability that individual i is linked to individual j as a logistic function of differences of observed variables as well as unobserved terms:

$$P(\text{Link}=1 \left| x_i, x_j, z_i, z_j \right.)=\Lambda(\psi_\alpha+\Sigma_B \psi_B \left| x_{iB}-x_{jB} \right|+\psi_z \left| z_i-z_j \right|)=\Lambda(q'\psi) \quad (2b)$$

Equations (2a) and (2b) form a basis for a structural model, which we estimate using Bayesian methods. To estimate equation 2b, we formed all possible combinations of pairs of respondents in the sample. In the reported results, we used the definition of linkage based on a recruitment event. The estimation results were not found to be sensitive to the choice of parametric functions, as probit and logistic regression produced almost identical results. As a robustness check, we varied the definition of linkage by expanding the network to include second degree recruitment contacts and found that the results did not change significantly. The final model incorporated variables that seemed to be plausible controls to counter omitted variable bias and produced relative goodness-of-fit based on Akaike Information Criterion.

Bayesian Model Identification and Estimation—The specification of the model follows (Goldsmith-Pinkham, 2013; Hsieh, 2014) and is a variation of the sample-selection model developed in Econometrics (Heckman, 1979; Wooldridge, 2010). First we assume that conditional on observable variables, both the residual, ε_i from equation (1) and the unobserved variable, z_i from equation (2b), have a joint normal distribution:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

$$(\varepsilon_i, z_i) \sim \mathcal{N}\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_\varepsilon^2 & \sigma_{\varepsilon z} \\ \sigma_{\varepsilon z} & \sigma_z^2 \end{array}\right)\right) \quad (3)$$

For identification we check σ_z^2 and σ_ε^2 are set to equal 1, which are relatively standard assumptions. As a result of this normalization, the term $\boldsymbol{\sigma_{\varepsilon z}}$ will absorb σ_z^2. Given these assumptions, the model in equation 2a can be rewritten

$$Y_i = \rho_1 W_1 Y_i + \rho_2 W_2 Y_i + X_i \beta + z_i \sigma_{\varepsilon z} + v_i \quad (4)$$

Letting $\Theta = (\rho_1, \rho_2\ \beta', \sigma_{\varepsilon z}, \phi')$, we can write the joint probability of the HIV status and the network connections:

$$P(Y, W_1 | X, \Theta) = P(Y | W_1, W_2, X, \varepsilon, \Theta) P(W_1 | X, \varepsilon, \Theta) =$$
$$= (2\pi\sigma_v^2)^{-,5} | I - \rho_1 W_1 - \rho_2 W_2 \exp(-\frac{1}{2\sigma_v^2} vv') \prod_{i \neq j} \frac{\exp(q'\psi)}{1 + \exp(q'\psi)} \quad (5)$$

The application of Bayes theorem requires a complete specification of priors for all unobservable variables in the system:

$$\psi \sim N_d(\psi_0, \Psi_0) \text{ for } d = 1 \text{ to } D \text{ variables} \quad (6)$$

$$\rho_\mathrm{i} \sim U\left[\frac{1}{\lambda_{\min}}, \frac{1}{\lambda_{\max}}\right], \text{where} \lambda \text{ is an eigenvalue of } W_i \text{ and } i = 1, 2 \quad (7)$$

$$\beta \sim N_K(\beta_0, B_0), \text{ for } k = 1 \text{ to } K \text{ variables} \quad (8)$$

$$\sigma_{\varepsilon \mathrm{z}} \sim N(0, \Sigma) \quad (9)$$

$$z_i \sim N(0, 1) \quad (10)$$

Bayes theorem, the posterior distribution of each unobservable in the system is a product of the prior and the likelihood of the data:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$P(z_i\,|Y,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z}_{-i},\Theta)\propto\pi(\mathrm{z})P(\mathrm{Y},\mathrm{W}_1|\,\mathrm{z},\Theta) \quad (11)$$

$$P(\psi\,|\mathrm{W}_1,z)\propto\pi(\psi)P(\mathrm{W}_1|\,\mathrm{z},\Theta) \quad (12)$$

$$P(\rho\,|Y,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\beta,\sigma_{\varepsilon z})\propto\mathrm{P}(\mathrm{Y}|\,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\beta,\sigma_{\varepsilon z}) \quad (13)$$

$$P(\beta\,|Y,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\rho,\sigma_{\varepsilon z})\propto\pi(\beta)\mathrm{P}(\mathrm{Y}|\,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\rho,\sigma_{\varepsilon z}) \quad (14)$$

$$P(\sigma_{\varepsilon z}\,|Y,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\rho,\beta)\propto\pi(\sigma_{\varepsilon z})\mathrm{P}(\mathrm{Y}|\,\mathrm{W}_1,\mathrm{W}_2,\mathrm{z},\beta,\rho) \quad (15)$$

Because the distribution of β is Normal and has a closed-form solution, we apply the Gibbs sampler. For the other parameters, we employ Metropolis within Gibbs Sampling procedure, in which the proposed candidates from the target distribution are either accepted or rejected based on ratios of the posteriors (Gelman, 2014). In addition, because the difference in the unobservable variables is sign-invariant under absolute value sign in the network formation model, we apply a normalization to $\sigma_{\varepsilon z}$ by confining the sampling region to a non-negative domain, following (Hsieh, 2014) . We set the tuning parameters to arrive at an acceptance rate 30–50% for all the parameters as recommended in the Bayesian statistics literature (Gelman, 2014; J. LeSage, Pace RK, 2009). We run the Markov Chain Monte Carlo (MCMC) simulation for 50,000 iterations, and discard the first 5,000 as a burn-in period. In addition to visual inspection of the trace plots, we monitor convergence using several standard methods and are able to verify that the models converged (Geweke, 1992; Raftery, 1992). In our analysis, we employ the convergence diagnostics toolkit which is publicly available through the Spatial Econometrics Library for Matlab (J. LeSage, 1999). All computations were implemented and performed in Matlab, Release 2015b (“MATLAB Release 2015b,”).

Results

Geography of HIV and Recruitment Networks

Table 1 contains a summary of the sample. Most PWID were Malay men in their late 30’s, who on average injected 3 times per day, primarily with heroin. Most respondents had stable housing (82%) but were not involved in a stable relationship (69.3%). Compared to HIV-seronegatives, HIV-infected PWID were more likely to be homeless (31.5% vs 14.5%, $p<0.01$), have more prior incarcerations (5.2 vs 3.4, $p<0.01$), have fewer recent injection partners (2.4 vs 5.4, $p<0.01$), be less likely to use clean needles (26.8% vs 53.0%, $p<0.01$),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

have injected drugs for a longer period of time (18.8 years vs 14.3 years, $p<0.01$), and appear to be more aware of their HIV status (90% vs 81%, $p=0.03$).

Both the recruitment pattern as well as the distribution of HIV-infected PWID differed significantly across sites. The recruitment chains that originated in Shah Alam, covered a wider geographic distribution than other chains originating from Kampung Baru or Kajang, where the recruitment patterns remained close to the recruitment sites. Overall, the recruiters were more likely to bring individuals from their immediate or adjacent neighborhood (Figure 1a). More than 50% of participants lived within close proximity (<5 km) to their recruiters, while fewer than 10% lived further than 15 km from their recruiter's residence. Only 11 of 44 (or 25%) neighborhoods were penetrated by multiple recruitment chains, while in most neighborhoods, the recruited individuals were linked by the same chain. Focusing on the 11 neighborhoods, we did not find statistically significant differences in the probability of being HIV-positive based on site affiliation, implying that on average individuals from the same neighborhood were not statistically different from their neighbors, despite having been recruited through different sites (See Appendix I).

In addition, the three recruitment sites produced samples with remarkably different HIV infection prevalence: 37.3% in Kampung Baru, but 10.4% in nearby Kajang and 6.3% in Shah Alam (see Figure 2). The geographical concentration of HIV was found to be closely associated with the network-driven recruitment process. HIV-positive individuals were on average 7.2 times more likely to have been recruited by an HIV-positive PWID than by an HIV-negative one ($p<0.01$) (See Appendix I).

Spatial and Social Network Effects

The probit results (Table 1: Model 1) provide further evidence that on average, the HIV status of an individual was positively and significantly associated with the average HIV status of his social peers over close distance (≤5 km), but not statistically significant over longer distances (>5 km). We found a similar pattern for neighbors: the average HIV status of the individual was also positively and significantly associated with the average HIV status of the neighbors (non-peers) over close distances (≤5 km), and not statistically significant over longer distances (Table 2: Model 2). There is a stronger correlation in the HIV status among nearby neighbors (<5km) than among peers who reside further away. In our estimates of the network formation and recruitment (Table 3: Model 5), we found evidence for homophily: that is, individuals were more likely to recruit peers who were the same gender, race, had similar housing status and relationship status. Factors such as daily injection status, history of incarceration and the propensity to use clean needles, did not appear to be correlated in the recruitment network.

In our structural model (Table 3: Model 6), our estimates of the parameters and associated with unobservable variables are statistically different from zero and provide evidence for the "endogeneity" of the network, signaling that the HIV outcome of peers and neighbors is influenced by "omitted" variables that are also correlated with the network formation. We found that both the nearest neighbor as well as peer effects, however, remain positive and significant even after controlling for unobserved variation related to link formation (Table 3: Model 6). This implies that the correlation of HIV status among peers and neighbors

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

residing within a close distance is not due to unobserved common characteristics that drive the network formation process (that could include sero-sorting), but is related to HIV status, which is most likely occurring through transmissions over short distances.

Discussion

Our analysis of the interaction between the first and second degree social ties and geographical distances, underscores how a compact geospatial area can increase the risk of HIV transmission by facilitating close contact between HIV-infected PWID. We demonstrated that there is a gradient to spatial proximity depending on the type of relationship (peer vs neighbor). An increase in physical distance between social acquaintances is associated with a decline in HIV transmission risk, while, all things being equal, proximity to a HIV-infected neighbor is significantly associated with HIV status. The findings suggest that residence in a neighborhood with high HIV prevalence coupled with high turnover rate in injection partnership and population mixing can contribute to onward HIV transmission. We found evidence that is consistent with patterns of sero-sorting among PWID: most of the HIV-infected individuals appear to be aware of their HIV status and have lower numbers of injection sharing partners in their reported social networks. We also found that the parameters that link network formation and HIV status are statistically significant, so we cannot reject the hypothesis that sero-sorting is not occurring. We are, however, inclined to interpret the correlations in the HIV outcomes between peers and neighbors as possible evidence for transmissions, because the coefficients associated with the average HIV status of nearby peers and neighbors remains significant after we control for possible sero-sorting in the structural equations.

The analysis of the risk environment also emphasizes the need to strengthen and expand prevention programs geographically. Among HIV-negative individuals, sterile needles were used only 53% of time in the previous month, while the injection-partner networks are almost twice as large as the networks of HIV-infected PWID. A combination of preventive measures like expansion of NSPs, OAT with methadone and buprenorphine, HIV treatment as prevention and pre-exposure prophylaxis (PrEP) that target certain locations could be highly effective given spatial dispersion of HIV-negative networks. Similarly, targeting HIV-infected individuals to ensure adequate access to antiretroviral therapy along with adherence support and NSPs would also promote more effective prevention (Kamarulzaman & Altice, 2015).

RDS is an effective tool to reach hidden populations in a relatively short time, and its path dependency can be used as an advantage. This is because RDS can be used to recruit individuals into HIV testing and treatment programs, allowing more effective targeting of populations based on sero-status. RDS, despite being a cost-effective method for reaching hidden populations, is not without limitations. Careful seed selection is necessary, but not a sufficient condition, as seed diversification does not guarantee a sample free of selection bias. The ability to reach the HIV-positive population hinges on the capability to successfully identify the proper location and tap into the HIV-positive network. As this study demonstrates, the sampling process has a strong geographical component and location

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

matters a great deal both for the formation of social networks and the operation of the risk environment.

In the RDS literature, recruitment is modeled as a Markov process with a unique stationary equilibrium (Salganik, 2004; Volz, 2008), but in practice, nothing guarantees that the chain will converge to a single steady state in finite time. This is the case, especially, if not all the network nodes are reachable and certain groups are more prone to isolation than others. Our reported estimates of HIV prevalence would have been remarkably different if seeds at each site failed to accumulate local PWID. In addition, it is evident from respondents' interviews that some populations in our RDS study were under-sampled (including women and ethnic minorities, such as refugees) and even reweighting by network size would not correct for this bias, as network size itself is likely to be measured with error.

There are several limitations associated with this study. First, our estimates of social network effects are based on the recruitment network, which is not synonymous with the injection network (Bazazi, Zelenev, & Altice, 2013). In the survey, the type of relationship between recruiters and recruits (e.g., injecting partner, sexual partner, etc.) was not fully assessed. This could lead to a measurement error in the contiguity matrix, as well as introduce unobserved heterogeneity, especially if the peers residing over longer distances are remarkably different in their behavior from peers who live closer to one another. It is entirely possible that the neighbors are part of the injection network of the individuals and the distinction between neighbors and peers may be artificial. As a robustness check, we tried multiple definitions of social networks (including up to third-degree recruitment contacts), and our results pertaining to social network and spatial effects were fairly robust and did not vary significantly. Some degree of measurement error is also likely in the assessment of distances among respondents. Granularity in the data was imprecise and included residential neighborhood location rather than exact address, and within 44 distinct neighborhoods, there was variation in the size of the neighborhoods. We varied the sample by excluding large neighborhoods and did not find any significant change in our results. Finally, our structural models may suffer from misspecification errors, potentially yielding inconsistent estimates, and some of the variables in our models may not be strictly exogenous.

Conclusions

The clustering of HIV infections across networks and space in Kuala Lumpur underscores the importance of closely integrating surveillance and prevention strategies. RDS can be applied to identify injection network structures and this provides an important mechanism for improving public health surveillance, accessing high-risk populations, and implementing risk-reduction interventions to slow HIV transmission.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgments

Funding: This research was supported by NIH career development (K01 DA037826 for AZ, K24 DA017072 for FLA and F30 DA039716 for ARB), research (NIDA R01 DA025943 for FLA), and training (T32GM07205, T32MH020031 for ARB) grants as well as University Malaya's High Impact Research Grant (E-000001-20001; AK) and the Yale Downs Fellowship (ARB). OraSure Technologies, Inc. provided discounted rapid HIV tests. Funders had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

References

Amato, L.; Davoli, M.; Ferri, M. Measures of outcomes in opiate detoxification trials: an experience of the Cochrane drugs and alcohol group; 9th International Cochrane Colloquium; Lyon, France. 2001.

Amber T, Gile K. The Effect of Differential Recruitment, Non-response and Non-recruitment on Estimators for Respondent-Driven Sampling. Electronic Journal of Statistics. 2011; 5:899–934.

Anselin, L. Spatial Econometrics: Methods and Models. Boston, MA: Dordrecht: Kluwer Academic Publishers; 1988.

Bachireddy C, Bazazi AR, Kavasery R, Govindasamy S, Kamarulzaman A, Altice FL. Attitudes toward opioid substitution therapy and pre-incarceration HIV transmission behaviors among HIV-infected prisoners in Malaysia: implications for secondary prevention. Drug Alcohol Depend. 2011; 116:151–157. [PubMed: 21232882]

Bazazi AR, Crawford F, Zelenev A, Heimer R, Kamarulzaman A, Altice FL. HIV Prevalence Among People Who Inject Drugs in Greater Kuala Lumpur Recruited Using Respondent-Driven Sampling. AIDS Behav. 2015; 19:2347–2357. [PubMed: 26358544]

Bazazi AR, Zelenev A, Altice FL. Individual and neighborhood correlates of membership in drug-using networks with a higher prevalence of HIV in New York City (2006–2009). Ann Epidemiol. 2013; 23:664–665. [PubMed: 23972897]

Bazazi AR, Zelenev A, Fu JJ, Yee I, Kamarulzaman A, Altice FL. High prevalence of non-fatal overdose among people who inject drugs in Malaysia: Correlates of overdose and implications for overdose prevention from a cross-sectional study. International J Drug Policy. 2015a In press [Manuscript number: DRUGPO-D-14-246R241].

Bazazi AR, Zelenev A, Fu JJ, Yee I, Kamarulzaman A, Altice FL. High prevalence of non-fatal overdose among people who inject drugs in Malaysia: Correlates of overdose and implications for overdose prevention from a cross-sectional study. Int J Drug Policy. 2015b; 26:675–681. [PubMed: 25532449]

Bramoulle Y, Djebbari H, Fortin B. Indentification of peer effects through social networks. Journal of Econometrics. 2009; 150:41–55.

Centers for Disease Control and Prevention, U. Behavioral Surveillance: Introduction to Respondent-Driven Sampling: Participant Manual. 2008.

Degenhardt L, Hall W. Extent of illicit drug use and dependence, and their contribution to the global burden of disease. Lancet. 2012:379.

Degenhardt L, Mathers BM, Wirtz AL, Wolfe D, Kamarulzaman A, Carrieri MP, Strathdee SA, Malinowska-Sempruch K, Kazatchkine M, Beyrer C. What has been achieved in HIV prevention, treatment and care for people who inject drugs, 2010–2012? A review of the six highest burden countries. Int J Drug Policy. 2014; 25:53–60. [PubMed: 24113623]

Degenhardt L, Whiteford HA, Ferrari AJ, Baxter AJ, Charlson FJ, Hall WD, Freedman G, Burstein R, Johns N, Engell RE, Flaxman A, Murray CJ, Vos T. Global burden of disease attributable to illicit drug use and dependence: findings from the Global Burden of Disease Study 2010. Lancet. 2013; 382:1564–1574. [PubMed: 23993281]

Doherty I, Shiboski S, Ellen JM, Adimora A, Padian NS. Sexual bridging socially and over time: A simulation model exploring the relative effects of mixing and concurrency on viral sexually transmitted infection transmission. Sexually Transmitted Diseases. 2006:368–373. [PubMed: 16721330]

Friedman S, Curtis R, Neaigus A, Jose B, Des Jarlais D. Social Networks, Drug Injectors' Lives, and HIV/AIDS. 2002

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Friedman S, Neaigus A, Jose B, Curtis R, Goldstein M, Ildefonso G, Rothenberg RB, Des Jarlais DC. Sociometric risk networks and risk for HIV infection. American Journal of Public Health. 1997; 87:1289–1296. [PubMed: 9279263]

Gelman, A.; Carlin, JB.; Stern, HS.; Dunson, DB.; Vehtari, A.; Rubin, DB. Bayesian Data Analysis. 3rd. London, U.K.: Chapman and Hall; 2014.

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. Bayesian Statistics. Vol. 4. Oxford, UK: Clarendon Press; 1992.

Goel S, Salganik MJ. Assessing respondent-driven sampling. Proceedings of the National Academy of Sciences. 2010; 107:6743–6747.

Goldsmith-Pinkham P, Imbens GW. Social networks and the identification of peer effects. Journal of Business and Economic Statistics. 2013; 31:253–264.

Heckathorn D. Respondent-driven sampling: a new approach to the study of hidden populations. Social Problems. 1997; 44:174–199.

Heckman J. Sample selection bias as a specification error. Econometrica. 1979; 47:153–161.

Heimer R. Critical issues and further questions about respondent driven sampling: comment on Ramirez-Valles et al (2005). Aids and Behavior. 2005; 9:403–408. [PubMed: 16344920]

Hsieh C, Lee LF. A Social Interactions Model With Endogenous Friendship Formation and Selectivity. Journal of Applied Econometrics. 2014

Jenness S, Neaigus A, Wendel T, Gelpi-Acosta C, Hagan H. Spatial recruitment bias in respondent-driven sampling: Implications for HIV prevalence estimation in urban heterosexuals. Aids and Behavior. 2014; 18:2366–2373. [PubMed: 24122043]

Joint United Nations Programme on HIV/AIDS (UNAIDS). Global AIDS Update 2016. Geneva, Switzerland: 2016. at: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf [Accessed on May 28, 2016]

Kamarulzaman A. Impact of HIV prevention programs on drug users in Malaysia. J Acquir Immune Defic Syndr. 2009; 52(Suppl 1):S17–S19. [PubMed: 19858930]

Kamarulzaman A, Altice FL. Challenges in managing HIV in people who use drugs. Curr Opin Infect Dis. 2015; 28:10–16. [PubMed: 25490106]

Kendall C, Kerr LR, Gondim RC, Werneck GL, Macena RH, Pontes MK, Johnston LG, Sabin K, McFarland W. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. Aids and Behavior. 2008; 12:S97–S104. [PubMed: 18389357]

Lahiri, S. Resampling Methods for Dependent Data. New York: Springer; 2003.

Latkin CA, Forman V, Knowlton A, Sherman S. Norms, social networks, and HIV-related risk behaviors among urban disadvantaged drug users. Soc Sci Med. 2003; 56:465–476. [PubMed: 12570967]

LeSage J. Spatial Econometrics Toolbox. 1999

LeSage, J.; Pace, RK. Introduction to Spatial Econometrics. Boca Raton, FL: CRC Press/Taylor & Francis Group, LLC; 2009.

Lyons R. The Spread of Evidence Poor Medicne via Flawed Social Network Analysis. Statistics, Politics and Policy. 2011:2.

Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS. 2005; 19:S67–S72.

Malekinejad MJL, Kendall C, Kerr LR, Rifkin MR, Rutherford GW. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. Aids and Behavior. 2008; 12:S105–S130. [PubMed: 18561018]

Manski C. Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies. 1993; 60:531–542.

MATLAB Release. Natick. Massachusetts, USA: The MathWorks, Inc; 2015b.

McCreesh N, Frost SD, Seeley J, Katongole J, Tarsh MN, Ndunguse R, Jichi F, Lunel NL, Maher D, Johnston LG, Sonnenberg P, Copas AJ, Hayes RJ, White RG. Evaluation of respodent driven sampling. Epidemiology. 2012; 23:138–147. [PubMed: 22157309]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Mills H, Colijn C, Vickerman P, Leslie D, Hope V, Hickman M. Respondent driven sampling and community structure in a population of injecting drug users, Bristol, UK. Drug Alcohol Dependence. 2012; 126:324–332. [PubMed: 22728045]

Ministry of Health, & Malaysia. Global Aids Response Progress Report 2014 - Malaysia In HIV/STI SECTION & Disease Control Division. 2014

Mustanski B, Birkett M, Kuhns LM, Latkin CA, Muth SQ. The Role of Geographic and Network Factors in Racial Disparities in HIV Among Young Men Who have Sex with Men: An Egocentric Network Study. Aids and Behavior. 2014

Raftery, A.; Lewis, SM. How many iterations in the Gibbs sampler?. In: Bernardo, J.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. Bayesian Statistics. Oxford, UK: Clarendon Press; 1992. p. 4

Reid G, Kamarulzaman A, Sran SK. Malaysia and harm reduction: the challenges and responses. Int J Drug Policy. 2007; 18:136–140. [PubMed: 17689356]

Rothenberg R, Long DM, Sterk CE, Pach A, Potterat JJ, Muth S, Baldwin JA, Trotter RT 3rd. The Atlanta Urban Networks Study: a blueprint for endemic transmission. AIDS. 2000; 14:2191–2200. [PubMed: 11061661]

Rothenberg R, Muth SQ, Malone S, Potterat JJ, Woodhouse DE. Social and geographic distance in HIV risk. Sexually Transmitted Diseases. 2005; 32:506–512. [PubMed: 16041254]

Rudolph A, Crawford ND, Latkin C, Fowler JH, Fuller CM. Individual and neighborhood correlates of membership in drug using networks with a higher prevalence of HIV in New York City (2006–2009). Annals Of Epidemiology. 2013; 23:267–274. [PubMed: 23523090]

Salganik M, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology. 2004:193–239.

Toledo L, Codeço CT, Bertoni N, Albuquerque E, Malta M, Bastos FI. Brazilian Multicity Study Group on Drug Misuse Putting Respondent-Driven Sampling on the Map: Insights from Rio de Janeiro, Brazil. Journal of acquired immune deficiency syndromes. 2011; 57:S136–S143. [PubMed: 21857309]

Topa, G.; Zenou, Y. Neighborhood versus network effects. In: Duranton, G.; Henderson, V.; Strange, W., editors. Handbook of Regional and Urban Economics. Vol. 4. Amsterdam: Elsevier Publisher; 2015.

UNGASS. Malaysia. Geneva, Switzerland: 2010. UNGASS Country Progress Report (January 2008-December 2009).

United Nations Office on Drugs and Crime (UNODC). [Accessed on 27 June 2016] Word Drug Report 2016. 2016. at: http://www.unodc.org/doc/wdr2016/WORLD_DRUG_REPORT_2016_web.pdf). Vienna, Austria

Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. Journal of Official Statistics. 2008; 24:79–97.

Watts D, Strogatz SH. Collective dynamics of ‘small-world’ networks. Nature. 1998:393. [PubMed: 9450756]

Wooldridge, J. Econometric analysis of cross section and panel data. 2nd. Cambridge, MA: MIT Press; 2010.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

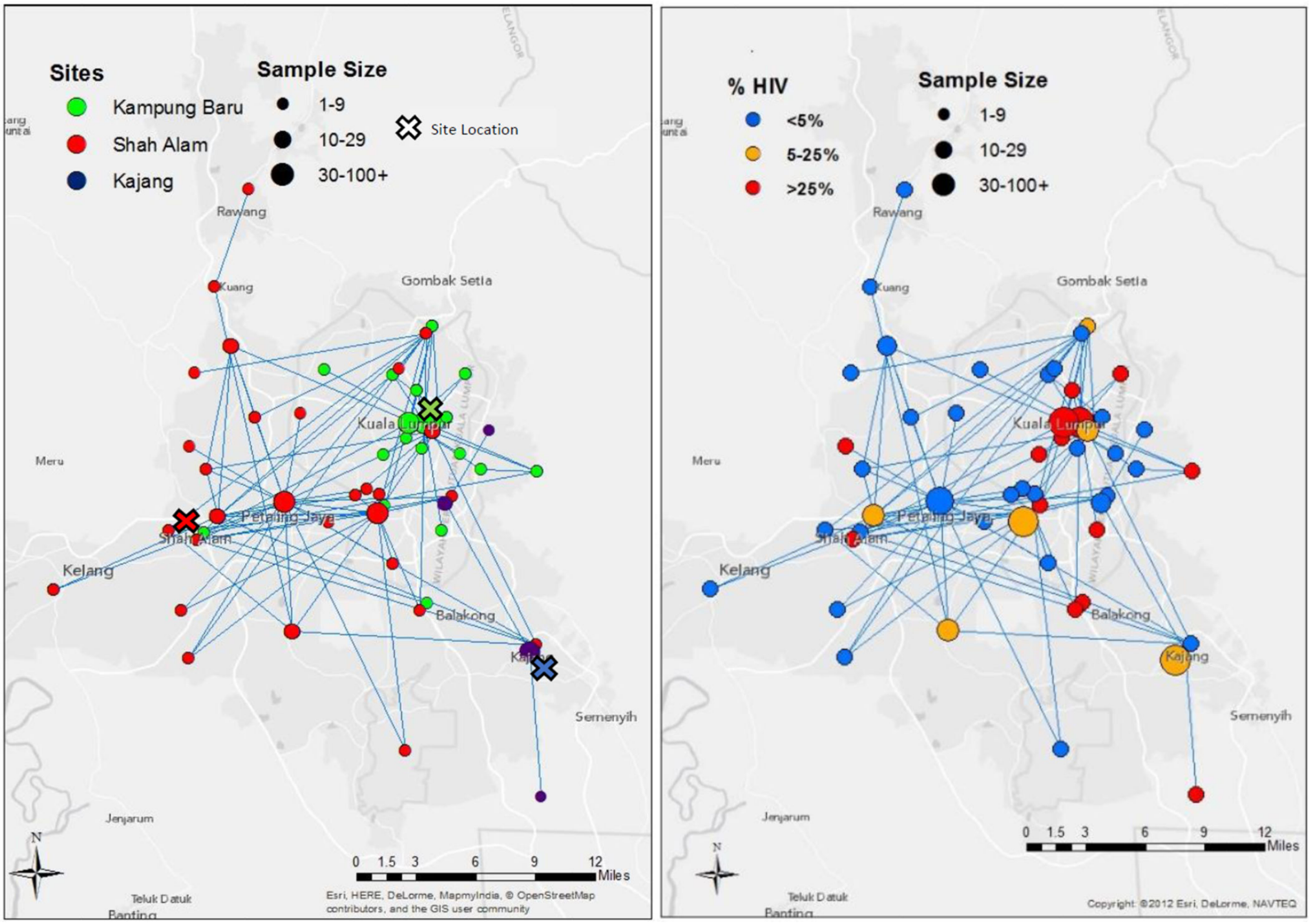

Figure 1.

a & b: Sampling Locations and HIV prevalence among 460 people who inject drugs in Greater Kuala Lumpur, Malaysia

Note: Nodes represent a group of sampled individuals at each residential location, edges represent recruitment links between individuals at each location, crosses indicate the location of the 3 recruitment sites (Kampung Baru, Shah Alam, Kajang). The size of the node corresponds to a specific group size consisting of a specific number of individuals (as indicated in the legend). In Figure 1a, the color of the node corresponds to different recruitment locations: red dots represent individuals that were recruited at the Shah Alam site; blue – individuals who were recruited at the Kajang site, and green –at the Kampung Baru site. In Figure 1b, the colors correspond to a specific HIV prevalence group: red dots are groups in which the prevalence is above 25%; yellow – HIV prevalence between 5 and 25%; and blue – HIV prevalence below 5%.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

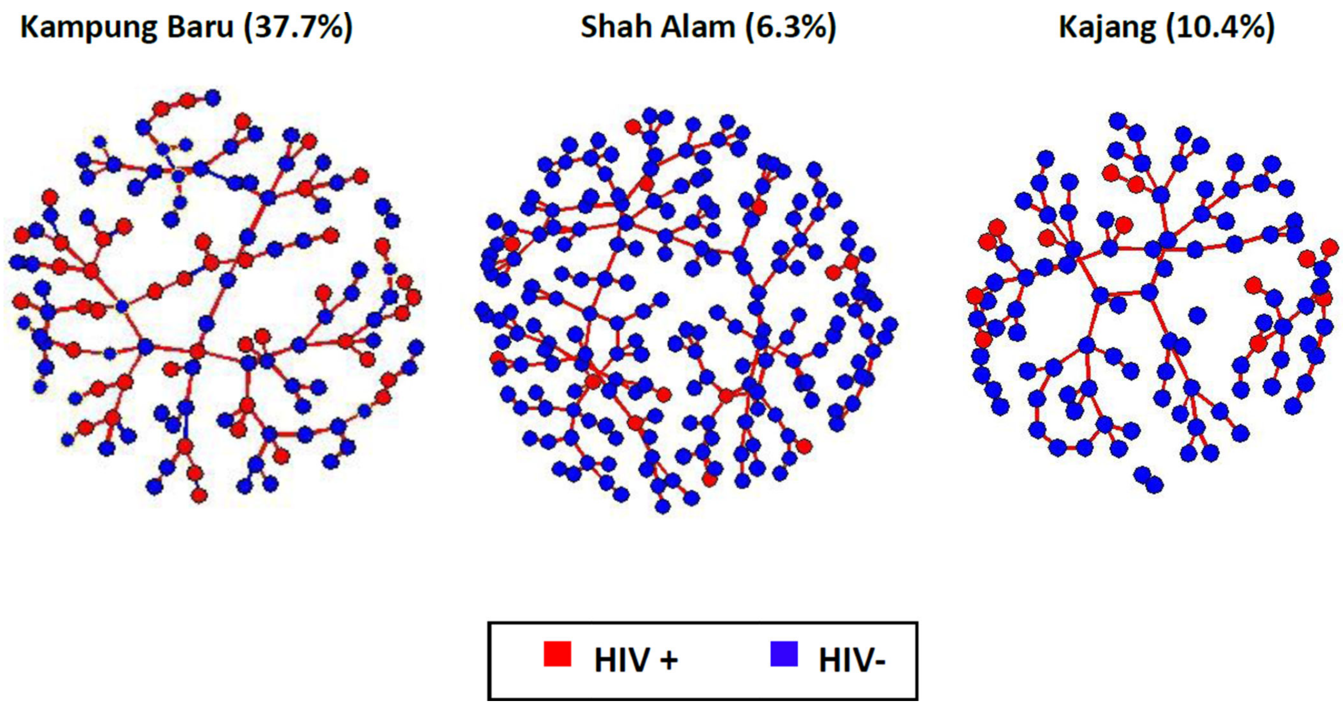

Figure 2.

Recruitment Network-based HIV Prevalence in Kuala Lumpur by Recruitment Site

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Comparison of demographic and risk behavior characteristics of people who inject drugs in Kuala Lumpur by HIV status (N=460)

Variables	**Total Sample**	**HIV positive**	**HIV negative**	**P value**
	N=460	**N=73**	**N=387**	
	N (%)	**N (%)**	**N (%)**	
Age - Mean (S.D.)	38.8 (9.2)	40.0 (7.8)	38.6 (9.5)	0.21
Gender				
Male	443 (96.3%)	377 (90.4%)	66 (97.4%)	0.01
Female	17 (3.7%)	10 (2.5%)	7 (9.6%)	Ref
Race/ Ethnicity				
Malay	416 (90.6%)	65 (89.0%)	351 (90.7%)	Ref
Chinese	12 (3.0%)	4 (5.5%)	28 (7.2%)	0.11
Indian	32 (7.0%)	4 (5.5%)	8 (1.7%)	0.59
Housing				
Stable Housing	381 (82.8%)	50 (68.5%)	331 (85.5%)	Ref
Unstable Housing	79 (17.2%)	23 (31.5%)	56 (14.5%)	<0.01
Relationship Status				
In a Stable Relationship	319 (69.3%)	65 (89.0%)	254 (65.6%)	Ref
Not in a Stable Relationship	141 (30.7%)	8 (11.0%)	133 (34.4%)	<0.01
Network Size - Mean (S.D)	20.4 (28.1)	16.4 (24.6)	21.2 (28.1)	0.19
Number of Incarcerations - Mean (S.D.)	3.6 (3.3)	5.2 (3.5)	3.4 (3.1)	<0.01
Injection Characteristics - Mean (S.D.)				
Mean Number of Injection Partners (past 30 days)	4.9 (7.9)	2.4 (1.9)	5.4 (8.5)	<0.01
Mean years of injection	15.1 (9.20)	18.8 (8.3)	14.3 (9.2)	<0.01
Mean Number of Times Injected Any Substance (past 30 days)	97.6 (51.2)	105.0 (54.8)	96.1 (50.5)	0.18
Mean Number of Days Injected Heroin (Past 30 days)	26.8 (8.7)	26.1 (8.6)	27.0 (9.3)	0.42
Percent of Time used Clean Needles (Past 30 days)	31.0 (32.3)	26.8 (29.4)	53.0 (37.9)	<0.01
Usual Place of Injection				
Public (Port or Shooting Gallery)	52.1 (50.0)	57.5 (50.0)	51.1 (47.8)	Ref
Private Residence	47.8 (50.0)	42.4 (50.0)	48.8 (47.8)	0.32
Recruitment Site				
Kampung Baru	127 (27.6%)	47 (64.4%)	80 (35.6%)	<0.01
Shah Alam	208 (45.2%)	13 (17.8%)	195 (50.4%)	<0.01
Kajang	125 (27.2%)	13 (17.8%)	112 (28.9%)	0.05
Aware of HIV Status	373 (81.1%)	66 (90.4%)	307 (79.3%)	Ref
Unaware of HIV Status	87 (18.9%)	7 (10.0%)	80 (20.7%)	0.03

Table 2

HIV Probit Regression Model with Peer and Neighbor Effects

Model 1 –Network Probit (Peer Effect)				Model 2 –Spatial Probit (Neighbor Effect)			
Outcome		**HIV**				**HIV**	
Covariates	**β**	**95% C.I.**	**P-value**		**β**	**95% C.I.**	**P-value**
Male Gender[1]	0.20	[−0.33,0.73]	0.46		0.46	[0.42,0.49]	0.18
Unstable Housing[2]	0.53	[0.18,0.88]	<0.01		0.55	[0.53, 0.57]	<0.01
In a Relationship[3]	−0.68	[−1.08, −0.27]	<0.01		−0.64	[−0.66, −0.62]	<0.01
Previously Incarcerated[4]	0.05	[0.01, 0.10]	0.02		0.05	[0.048, 0.053]	0.05
Number of Sharing Partners	−0.10	[−0.16, −0.03]	0.00		−0.10	[−.11, −.10]	<0.01
Propensity to Use Clean Needles[5]	0.16	[−0.16,0.48]	0.33		0.05	[0.03, 0.06]	0.79
Years of Injection Drug Use	0.02	[0.002, 0.04]	0.03		0.027	[0.026, 0.029]	0.01
Constant	−0.85	[−1.39,-0.30]	<0.01		−0.89	[−0.91, −0.86]	<0.01
Mean HIV status of Peers within a Close Distance (0–5 km) W^N_1*Y	0.42	[0.17, 0.67]	<0.01	Mean HIV status of Neighbors within a Close Geographical Distance (0–5 km) W^F_1*Y	0.44	[0.43, 0.45]	<0.01
Mean HIV status of Peers within a more Remote Distance (5–10 km) W^N_2*Y	−0.12	[−0.31,0.06]	0.19	Mean HIV status of Neighbors within a more Remote Geographical Distance (5 −10 km) W^F_1*Y	−0.06	[−0.07, −0.05]	0.17
Method	Maximum Likelihood			Method	Maximum Likelihood		

Referent group:

[1]Female;

[2]Stable Housing Referent;

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[3]"Not In a stable relationship";

[4]"Never Incarcerated"

[5]Propensity to use clean needles is defined as a dummy indicator that is equals to 1 if individual reported using clean needles more than 25% of the time, 0 otherwise.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Estimates of Endogenous Network-HIV Model with Peers and Neighbor Effects

Model 4 – HIV Probit (with Peer and Neighbors effects) (1 equation)				Model 6 – HIV Probit + Endogenous Network Formation (2 equations)			
Outcome:	**HIV**			**Equation 1 Outcome:**	**HIV**		
Covariates	**β^{probit}**	**95% C.I.**	**P-value**		**β^{probit}**	**95% C.I.**	**P-value**
Male Gender[1]	0.28	[−0.28, 0.85]	0.33		0.45	[−0.24, 1.12]	0.10
Unstable Housing[2]	0.49	[0.14, 0.84]	0.01		0.49	[0.88,0.094]	0.01
In a Relationship[3]	−0.66	[−1.00, −0.25]	<0.01		−0.68	[−1.14, −0.24]	0.00
Previously incarcerated[4]	0.05	[0.003, 0.10]	0.03		0.04	[−0.01, 0.093]	0.07
Number of Sharing Partners	−0.07	[−0.12, −0.01]	0.01		−0.08	[−0.15, −0.02]	<0.01
Propensity to Use Clean Needles[5]	−0.06	[−0.39, 0.28]	0.72		0.04	[−0.32, 0.43]	0.42
Years of Injection Drug Use	0.02	[0.004, 0.04]	0.02		0.03	[0.007, 0.048]	<0.01
Constant	−0.57	[−1.10, −0.04]	0.03		−0.74	[−1.20, −0.26]	<0.01
Mean HIV status of Peers residing within a Close Distance (0–5 km) $W^F Y$	0.29	[0.01, 0.57]	0.04		0.39	[0.167, 0.57]	<0.01
Mean HIV status of Neighbors residing within a	0.32	[0.06, 0.58]	0.02		0.17	[0.007, 0.42]	0.02
Close Distance (5–10 km) $W^N Y$							
Unobservables	NA				0.06	[0.005,0.19]	<0.01
Model 5: Exogenous Network Formation (w/ Logit link) (1 equation)							
Outcome	Link Presence			**Equation 2 Outcome:**	Link Presence		
Covariates	β^{logit}	95% C.I.	P-value		β^{logit}	95% C.I.	P-value
Same Gender *	0.73	[0.19,1.26]	0.01		0.36	[0.28,0.78]	<0.01
Same Race/Ethnicity *	0.45	[0.15,0.73]	0.00		0.44	[0.41,0.66]	<0.01

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Model 4 – HIV Probit (with Peer and Neighbors effects) (1 equation)				Model 6 – HIV Probit + Endogenous Network Formation (2 equations)			
Outcome:	**HIV**			**Equation 1 Outcome:**	**HIV**		
Covariates	**β^{probit}**	**95% C.I.**	**P-value**		**β^{probit}**	**95% C.I.**	**P-value**
Both Homeless *	0.28	[0.05,0.49]	0.01		0.33	[0.16,0.34]	<0.01
Both are in a Relationship *	0.15	[−0.0,0.33]	0.13		0.16	[0.04, 0.26]	0.01
Both with History of Incarceration *	0.24	[−0.0,0.51]	0.08		0.08	[−0.08,0.25]	0.16
Both Daily Injectors *	0.20	[−0.2,0.68]	0.42		0.20	[−0.21,0.37]	0.27
Both less likely to use Clean Needles [5,*]	0.16	[−0.0,0.38]	0.16		0.16	[−0.009,0.25]	0.03
Absolute Difference in the Number of Injection Partners	−0.01	[−0.0,0.00]	0.09		−0.01	[−0.01,-0.005]	<0.01
Absolute difference in years of Injection	−0.35	[−0.5,-0.1]	0.01		−0.32	[−0.41,-0.28]	<0.01
Constant	−6.86	[−7.4,-6.2]	<0.01		−5.50	[−5.90,-5.32]	<0.01
Absolute difference in the Unobservables	NA				−0.05	[−0.11,-0.03]	<0.01
Method Models 4,5	Maximum Likelihood Estimation			Method Model 6	Markov Chain Monte Carlo		

Referent group:

[1]Female;

[2]Stable Housing Referent;

[3]"Not In a stable relationship"

[4]"Never Incarcerated"

[5]Propensity to use clean needles is defined as a dummy indicator that is equals to 1 if individual reported using clean needles more than 25% of the time.

[*]If the pair of individual share the same characteristics, the variable takes on a value of 1, 0 otherwise