Supplemental Digital Content Description

Methods

A limitation of the logistic models is that they treat the network proportion variables as though they are measured without error, when in fact the 1–4 individuals in a respondent's egocentric referral network represent only a fraction of the respondent's total network, and an estimate of network proportion based on 4 network members is more precise than an estimate based on one. The resulting measurement error may be expected to bias the corresponding coefficient downward. To address this limitation, we also fit a generalized Structural Equation Model (Figure S1), combining a measurement model in which the respondent's true (unobserved) network proportion with a given characteristic affects the likelihood that each of his or her contact network members exhibit that characteristic, and a structural model in which the true network proportion affects the respondent's own status, controlling for individual-level covariates.[1] The resulting estimates of the association between network proportion and respondent's own status were compared to those from the logistic regression models.

Finally, although RDS's recruitment via social networks can decrease voluntarism and masking bias, concerns have been raised about whether the resulting recruitment network truly reflects respondents' social networks.[2,3] Participants may for instance sell recruitment coupons to strangers they do not know, or lie about their relationship to their recruiter in order to remain eligible. Such violations of the assumption that RDS reflects respondents' social networks may compromise the effectiveness of the technique in generating a representative sample.[4,5]

Results

Results from the generalized SEM models are shown in Table S4. As expected, the estimated association between the respondent's HIV serostatus and the proportion of his or her network that is HIV infected was greater than for the basic logistic regression model (AOR, 5.77; 95% CI, 1.60-20.8); the same was true for high frequency injecting (AOR, 2.51; 95% CI, 1.02-6.17)).

References

1. Skrondal A, Rabe-Hesketh S. Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. Boca Raton: Chapman & Hall/CRC; 2004.

2. Scott G. "They got their program, and I got mine": A cautionary tale concerning the ethical implications of using respondent-driven sampling to study injection drug users. Int J Drug Policy. 2008 Feb;19(1):42–51.

3. Heimer R. Critical Issues and Further Questions About Respondent-Driven Sampling: Comment on Ramirez-Valles, et al. (2005). AIDS Behav. 2005 Dec 1;9(4):403–8.

4. Heckathorn DD. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. Soc Probl. 1997 May 1;44(2):174–99.

5. Ramirez-Valles J, Heckathorn DD, Vázquez R, Diaz RM, Campbell RT. The Fit Between Theory and Data in Respondent-Driven Sampling: Response to Heimer. AIDS Behav. 2005 Dec 1;9(4):409–14.

Table S1. Characteristics of people who inject drugs in Athens, Greece recruited through RDS in 2012, (N=1030)[a].

INDIVIDUAL LEVEL VARIABLES	N (%)
Sample Characteristics	
Age	
18-30 yrs	285 (27.7)
31-40 yrs	506 (49.1)
41+ yrs	239 (23.2)
Sex	
male	884 (85.8)
female	146 (14.2)
Country of origin	
Greece	865 (84.0)
other	165 (16.0)
Highest level of education	
up to primary education	278 (27.1)
secondary/middle school	327 (31.9)
high school	285 (27.8)
university or equivalent	134 (13.1)
Homeless past 12 months	370 (36.0)
Ever been in Prison	545 (53.3)
Recruited by[b]	
family member	14 (1.3)
sex partner	14 (1.3)
drug using or drug buying partner	94 (9.1)
friend	384 (37.3)
acquaintance	376 (36.5)
stranger	213 (20.7)
HIV Seroprevalence	216 (21.0)
Risk Behaviors[c]	
Main drug	
heroin	851 (83.4)
cocaine or other	169 (16.6)
High frequency injecting	455 (44.3)
Shares syringes	
never	592 (57.9)
rarely	309 (30.2)
half the time or more	121 (11.8)

Table S1. Characteristics of people who inject drugs in Athens, Greece recruited through RDS in 2012, (N=1030)[a] (continued).

Share cookers, filters and/or water	
never	333 (32.5)
rarely	215 (21.0)
half the time	113 (11.0)
most of the time or always	365 (35.6)
Divide drugs	
never	620 (60.9)
rarely	267 (26.2)
half the time or more	131 (12.9)
Multiple sex partners	503 (48.8)
Unprotected sex	549 (54.0)
Alcohol or drugs with last sex partner	715 (70.0)
HIV Prevention Activities	
Ever in opiate substitution treatment	254 (25.4)
Ever in drug treatment	679 (66.3)
Receives condoms	605 (59.0)
Receives syringes	645 (62.8)

[a] First wave (seeds) and last wave have been excluded.
[b] Respondents could select more than one description for the relationship with their recruiter
[c] Results reflect behavior in past 12 months.

Table S2. Network proportions of selected characteristics[a] in successive RDS waves.

RDS wave	Networks	Network Members	% HIV+	P value	% High frequency injecting	P value	% Engaging in unprotected sex	P value
2	17	55	39.6	ref	56.3	ref	30.9	ref
3	38	99	36.2	0.51	55.7	0.25	52.2	0.86
4	61	158	37.2	0.94	58.6	0.17	52.3	0.64
5	97	229	30.5	0.17	53.6	0.13	52.9	0.53
6	132	306	29.0	0.19	44.5	0.06	58.1	0.64
7	174	391	19.1	0.02	36.0	0.01	57.3	0.75
8	217	511	15.8	0.01	41.8	0.04	53.0	0.48
9	294	673	16.7	0.01	44.6	0.03	33.1	0.67

[a] Average network proportion that is Human Immunodeficiency Virus seroprevalence (HIV+), high frequency injecting (HFI) and unprotected sex (US)

Table S3. Characteristics of people who inject drugs networks in Athens, Greece recruited through RDS in 2012, (N=1030)[a]

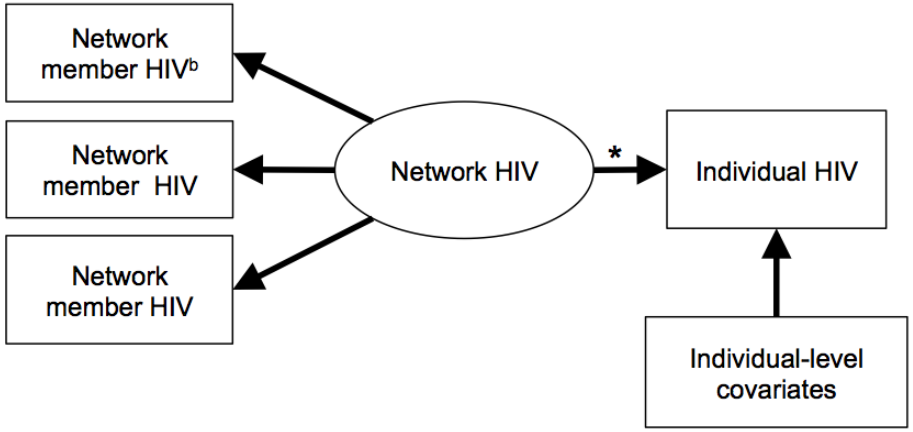
NETWORK LEVEL VARIABLES[b]	N (%)
Network HIV Seroprevalence (%)	
HIV infected	
Low (0)	671 (65.2)
Moderate (1 - 49)	109 (10.6)
High (≥50)	249 (24.2)
Network Risk Behaviors[c] (%)	
Main drug non-heroin	
Low (0)	717 (69.7)
Moderate (1 - 49)	115 (11.2)
High (≥50)	197 (19.1)
High frequency injecting	
Low (0)	405 (39.4)
Moderate (1 - 49)	97 (9.4)
High (≥50)	527 (51.2)
Shares syringes	
Low (0)	406 (39.5)
Moderate (1 - 49)	129 (12.5)
High (≥50)	493 (48.0)
Shares cookers, filters or water	
Low (0)	177 (17.2)
Moderate (1 - 49)	59 (5.7)
High (≥50)	793 (77.1)
Divides drugs	
Low (0)	446 (43.6)
Moderate (1 - 49)	113 (11.0)
High (≥50)	465 (45.4)
Multiple sex partners	
Low (0)	355 (34.5)
Moderate (1 - 49)	174 (16.9)
High (≥50)	501 (48.6)
Unprotected sex	
Low (0)	303 (29.5)
Moderate (1 - 49)	155 (15.1)
High (≥50)	570 (55.4)
Alcohol or drugs with last sex partner	
Low (0)	168 (16.3)
Moderate (1 - 49)	148 (14.4)
High (≥50)	713 (69.3)
Network Prevention Activities (%)	
Ever in opiate substitution treatment	
Low (0)	587 (57.8)
Moderate (1 - 49)	131 (12.9)
High (≥50)	297 (29.3)

Table S3. Characteristics of injecting drug user networks in Athens, Greece recruited through RDS in 2012, (N=1030)[a] (continued).

Ever in drug treatment	
Low (0)	205 (19.9)
Moderate (1 - 49)	56 (5.4)
High (≥50)	768 (74.6)
Receives condomsc,d	
Low (0)	263 (25.6)
Moderate (1 - 49)	61 (5.9)
High (≥50)	704 (68.5)
Receives syringes[c,d]	
Low (0)	247 (24.0)
Moderate (1 - 49)	62 (6.0)
High (≥50)	720 (70.0)
Other Network Characteristics (%)	
Male	
Low (0)	71 (6.9)
Moderate (1 - 49)	14 (1.4)
High (≥50)	944 (91.7)
Non-Greek	
Low (0)	784 (76.2)
Moderate (1 - 49)	71 (6.9)
High (≥50)	174 (16.9)
Primary school only	
Low (0)	583 (56.7)
Moderate (1 - 49)	138 (13.4)
High (≥50)	308 (29.9)
Homeless ‡	
Low (0)	470 (45.7)
Moderate (1 - 49)	137 (13.3)
High (≥50)	422 (41.0)
Ever in prison	
Low (0)	308 (30.0)
Moderate (1 - 49)	87 (8.5)
High (≥50)	633 (61.6)
Friend	
Low (0)	434 (42.1)
Moderate (1 - 49)	194 (18.8)
High (≥50)	402 (39.0)
Stranger	
Low (0)	678 (65.8)
Moderate (1 - 49)	158 (15.3)
High (≥50)	194 (18.8)

[a] First wave (seeds) and last wave have been excluded.
[b] Network-level variables were measured as network proportions. For example, if an individual had 4 network members and 3 were HIV infected, that respondent's HIV infected network proportion would be 75%.
[c] Results reflect behavior in past 12 months.
[d] Received condoms/syringes from an HIV prevention program

Figure S1. Sample general structural equations model[a] (GSEM)—the effect of network HIV seroprevalence on individual HIV infection.



[a] All variables are represented by rectangles or ovals. Rectangles represent *measured* variables, e.g. variables are directly measured through study. Ovals represent *latent* variables, e.g. unmeasured constructs that are responsible for correlations between measured variables in the model. For example, the proportion of an individual's network that is HIV infected (network HIV) is a latent variable that influences network members HIV serostatus (network member HIV). Arrows represent hypothesized pathways from independent variable to dependent variable.
[b] The actual number of network members depends on the number of study participants recruited by the given individual.
* Outcome of interest is the effect of network HIV seroprevalence on individual HIV serostatus.

Table S4. Generalized Structural Equations Models[a] (GSEM) relating network variables[b] and individual HIV seroprevalence among people who inject drugs in Athens, Greece 2012.

Network HIV Seroprevalence (%)	**Odds Ratio[e]**	**SE of Latent Variable**	**p value**	**95% CI**
HIV infected	5.30	0.08	0.01	(1.51,18.81)
Network Risk Behaviors[d] (%)				
Main drug non-heroin	4.31	0.08	0.28	(0.30,61.96)
High frequency injecting	2.31	0.05	0.05	(0.98,5.43)
Shares syringes	1.87E+5	0.04	0.86	(0.00,9.21E+61)
Shares cookers, filters or water	1.39	0.06	0.53	(0.49,3.94)
Divides drugs	4.06	0.05	0.11	(0.73,22.51)
Multiple sex partners	0.74	0.05	0.39	(0.39,1.38)
Unprotected sex	1.03	0.05	0.89	(0.67,1.55)
Alcohol or drugs with last sex partner	1.03	0.05	0.883	(0.66,1.62)
Network Prevention Activities (%)				
Ever in opiate substitution treatment	0.60	0.06	0.41	(0.17,2.04)
Ever in drug treatment	0.34	0.06	0.17	(0.07,1.56)
Receives condoms	2.49	0.05	0.10	(0.83,7.44)
Receives syringes	1.74	0.06	0.06	(0.99,3.08)
Other Network Characteristics (%)				
Non-Greek	0.94	0.15	0.58	(0.75,1.17)
Primary school only	1.44	0.06	0.82	(0.13,16.06)
Homeless[e]	1.50	0.05	0.50	(0.46,4.78)
Ever in prison	1.54	0.05	0.58	(0.34,7.08)
Friend	2.25	0.05	0.32	(0.45,11.23)
Stranger	0.18	0.08	0.18	(0.01,2.17)

[a] Each network variable represents an independent model. Multivariable models include the network variable indicated and individual-level covariates.

[b] Network variables refer to the latent variable true network proportion of a given characteristic. These latent network variables are estimated by the measured network characteristics of network members.

[c] Covariates for multivariate HIV models included: Covariates for multivariate HIV models included individual-level: age, sex, homelessness, highest level of education, history of incarceration, getting condoms from an HIV prevention activity, getting syringes from an HIV prevention activity, getting condoms from an HIV prevention activity, ever being in drug treatment, using divided drugs with a syringe that has already been used, frequency of injection, drug injected most often, sharing cookers/water or filters, sharing syringes, using alcohol or drugs with last sex partner and having unprotected sex.

[d] An increase of one unit in the latent variable increases the odds of the respondent being HIV by the listed odds ratio. For example, an increase in one unit in the latent variable network HIV seroprevalence increases the log odds ratio (log(5.30)=1.67) by 1.55 (1.67 x standard deviation = 1.67 x 0.93), corresponding to an odds ratio of 4.71 . Thus, an increase of one standard deviation in the true, underlying network proportion HIV infected increases the log odds of the respondent being HIV infected by an estimated 471%. This should be greater in magnitude than the effect estimated based on the observed proportion positive (among the 2-4 network members).

[e] Results reflect behavior in past 12 months.