Published in final edited form as:

Ann Epidemiol. 2013 October ; 23(10): 664–665. doi:10.1016/j.annepidem.2013.07.018.

Letter to the Editor: Regarding “Individual and neighborhood correlates of membership in drug using networks with a higher prevalence of HIV in New York City (2006–2009)”

Alexander R. Bazazi^{1,2}, Alexei Zelenev², and Frederick L. Altice^{1,2}

¹Department of Epidemiology of Microbial Diseases, Yale School of Public Health

²Department of Medicine, Section of Infectious Diseases, AIDS Program, Yale School of Medicine

Keywords

respondent-driven sampling; drug use; HIV; social network analysis

Rudolph et al. (2013) present data on the distribution of HIV infection in a respondent-driven sampling (RDS) recruitment network of drug users in New York City. Their article highlights the importance of integrating network and spatial perspectives into the study of HIV in marginalized populations. The effort is commendable as one of the first articles to explore network effects beyond the level of dyadic or ego-centric analysis using only standard RDS recruitment data. There are, however, several issues worth considering that challenge the authors' conclusions. We discuss analytical limitations and highlight key challenges that should be addressed when using RDS data. Resolving these issues will ultimately advance our understanding of the role of networks in HIV transmission.

Thinking clearly about ties

First, the distinction between the RDS recruitment process and relational features of the network are blurred in the article, resulting in a lack of conceptual clarity in the analysis and interpretation. From the article, it is not clear what recruitment ties among drug users actually mean and whether these ties translate into behaviors associated with HIV transmission. Furthermore, not all recruitment ties that the authors use to define their network are equal, representing the same kind of tie in the underlying social network. For example, in higher HIV prevalence recruitment networks, the ties sampled through recruitment may be more likely to represent relationships that facilitate HIV transmission. The article's definition of the main outcome (membership in higher HIV prevalence recruitment networks) appears artificial and somewhat ad hoc due to a lack of accounting for different kinds of ties and interactions between individuals. It is not clear that lower or higher prevalence recruitment networks represent meaningful or coherent groups. Gaining a richer understanding of the nature of ties among drug users is important for better understanding the HIV epidemic in this population.

Second, the authors make claims about far-reaching effects of HIV in social networks based on inference on partially observed networks and do not adequately justify their choice of outcome. Even if all recruitment ties represent similar relationships (which they may or may

not), the “recruitment” network does not provide complete information on the “social” network. From the observation of higher HIV prevalence among those separated by as many as six recruitment waves from HIV-infected individuals, the conclusion is made that “the influence of HIV may extend beyond six degrees of separation” (Rudolph 2013). It is more likely that individuals separated by many recruitment waves are socially connected, but their connections are not observed directly. Though this limitation is acknowledged in the discussion section, the observed clustering over six recruitment waves is used as justification for choosing membership in a higher-prevalence recruitment network as the primary outcome, rather than a more intuitive outcome, such as HIV infection. Furthermore, the random graph simulation offered to support this choice does not serve as a valid statistical test of the “reach” of HIV in the social network. The simulation simply confirms what we would expect to find and what is readily apparent from visual inspection of the article’s first figure: HIV is not randomly distributed in recruitment networks.

What about space?

The authors argue convincingly that integrating network and spatial analysis is important, yet they miss an opportunity to explore the interrelationship between spatial distance and social ties. The clustering observed over many waves is primarily driven by two extremely high-prevalence recruitment networks that constitute 9% (34/378) of the sample. This observed effect may, however, be due to unique individual characteristics in these recruitment networks or the areas from which they were sampled rather than a network effect of HIV extending from one individual to the next. In the context of explaining an association between minority composition of neighborhoods and membership in higher HIV prevalence networks, the authors mention that most participants from the Bronx were members of higher prevalence networks. This suggests that spatial clustering of HIV may be another reasonable explanation for the observed clustering of HIV in the recruitment network, but this potential explanation is not mentioned.

RDS dynamics and selection bias

The authors examine neighborhood-level correlates of membership in higher-prevalence recruitment networks, but interpretation of their findings is complicated by potential sample selection bias arising from recruitment dynamics. If individuals with a specific outcome have a different probability of being observed at different levels of a covariate, then estimates of the association between a given covariate and the outcome will be biased. For example, if HIV-infected individuals living in wealthy rather than poor neighborhoods are more likely to be recruited, then the association described between neighborhood wealth and HIV would be over-estimated. Selection bias would be of particular concern for estimating neighborhood effects if neighborhood characteristics strongly influence recruitment. Such selection biases are problematic in all RDS studies and remain a significant challenge for valid analysis.

Closing

The article raises numerous important issues, addresses them in novel ways, and opens the door for more rigorous examination of these issues in the next generation of methodological inquiries. Uniting network and spatial analysis to better understand the distribution of HIV infection in marginalized populations is innovative, though spatial and network effects can be difficult to disentangle. If network effects coincide too closely with spatial effects, then it may be extremely difficult to identify these effects separately. Network analysis using RDS recruitment data poses unique challenges, and analysts should distinguish “recruitment” networks from underlying “social” networks, extending conclusions from one to the other cautiously. Advancing inquiry in this area will require new techniques to more fully utilize RDS recruitment network data.

Acknowledgments

This work was supported by NIH grants NIDA K24 DA017072 and T32 GM07205.

References

Rudolph AE, 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(5):267–74. [PubMed: 23523090]