Published in final edited form as:

Sex Transm Dis. 2017 February ; 44(2): 71–78. doi:10.1097/OLQ.0000000000000561.

The HIV endemic: maintaining disease transmission in at-risk urban areas

Richard Rothenberg, MD MPH¹, Dajun Dai, PhD², Mary Anne Adams, MSW¹, and John Wesley Heath, MPH¹

¹School of Public Health, Georgia State University, Atlanta, GA

²Department of Geosciences, Georgia State University, Atlanta GA

Abstract

Objectives—A study of network relationships, geographic contiguity, and risk behavior was designed to test the hypothesis that all three are required to maintain endemicity of HIV in at-risk urban communities. Specifically, a highly interactive network, close geographic proximity, and compound risk (multiple high risk activities with multiple partners) would be required.

Methods—We enrolled 927 participants from two contiguous geographic areas in Atlanta, GA: a higher risk area and lower risk area, as measured by history of HIV reporting. We began by enrolling 30 “seeds” (15 in each area) who were comparable in their demographic and behavioral characteristics, and constructed 30 networks using a chain-link design. We assessed each individual’s geographic range; measured the network characteristics of those in the higher and lower risk areas; and measured compound risk as the presence of two or more (out of six) major risks for HIV.

Results—Among participants in the higher risk area, the frequency of compound risk was 15%, compared to 5% in the lower risk area. Geographic cohesion in the higher risk group was substantially higher than that in the lower risk group, based on comparison of geographic distance and social distance, and on the extent of overlap of personal geographic range. The networks in the two areas were similar: both areas show highly interactive networks with similar degree distributions, and most measures of network attributes were virtually the same.

Conclusions—Our original hypothesis was supported in part. The higher and lower risk groups differed appreciably with regard to risk and geographic cohesion, but were substantially the same with regard to network properties. These results suggest that a “minimum” network configuration may be required for maintenance of endemic transmission, but a particular prevalence level may be determined by factors related to risk, geography, and possibly other factors.

Introduction

The annual HIV diagnosis rate in the United States declined by 33.2% from 2002–2011,¹ and by 5% from 2011 to 2014.² This overall change has been paralleled in other high income countries (for example, the number of new diagnoses in the United Kingdom fell by 20% between 2005 and 2012). Among 58 low and middle income countries assessed in the UNAIDS Global Report for 2012,³ 49 had decreasing (39) or stable (10) rates and nine exhibited increases between 2001 and 2011.³ Although this overall change hides important

variation by age, sex, race, and risk group, with considerable increases among some subgroups and within some countries, the era of unfettered epidemic spread of HIV appears to be over. As we enter a phase of endemic transmission, understanding the new dynamics is of increasing importance.

Some years ago, we postulated that endemic transmission of HIV in at-risk urban areas was maintained by the interaction of geographic contiguity, compound risk (multiple risks with multiple people), and network structure.^{4,5} This hypothesis rested on data from diverse sources. We have since conducted a study of the personal geographic space, risk-taking, and network structure of two groups at different levels of HIV risk in communities in the inner city of Atlanta GA. We present here data that confirm some features of the hypothesis.

Methods

The Atlanta Metropolitan Area, defined by the 2010 US Census as a 20-county area, has a population of 4,228,492. The City of Atlanta (hereafter, Atlanta) houses less than 10% (420,003) of the metro area.⁶ We conducted this study in 10 Atlanta ZIP codes—five in higher risk areas (total population, 147,938) and five in lower risk (total population, 197,195),⁷ between 2006 and 2011. The higher risk ZIP codes have an area of 67 square miles; the lower risk ZIP codes, 173 square miles. They share a border that is 16.7 miles in length. These groupings were selected based on their HIV reporting history at the start of the study; the higher risk area accounted for 10 times more reported cases than did the lower risk area (Figure 1).

Study design

In order to construct a full network (that is, a sociogram that connects interviewed persons, as opposed to an egocentric design, that only identifies contacts to a respondent, but does not connect respondents), we used a chain-link design.⁸ This approach uses one of the contacts that a respondent names as the next link in a chain. We chose this design in preference to the major alternative—a snowball design wherein all contacts are interviewed, then all the contacts of contacts are interviewed—because of logistical and some statistical advantages.

We spent 6 months conducting ethnographic assessment of the neighborhoods in these areas and recruited 30 persons to act as “seeds” for the study—three in each of five ZIP Codes in each of the two areas. We recruited an equal number of men and women who qualified by being 18 years of age or older, being involved in HIV risk-taking, either through use of drugs or sexual activity, and who communicated a willingness to name and discuss their partners. Each seed was interviewed using a standard survey instrument, tested for HIV and STDs, and information about their contacts was elicited. We sought demographic and risk information about all partners and interviewed as many as we could locate. In addition, we sought specific geographic information on a variety of sites that helped define the respondents’ geographic range (home, if any; usually places for sleeping; centers of activity (meaning where the respondent spent a lot of time and could be usually found); sites at which risk activity took place; specific sites of contact with each partner; other frequented places). Sites outside Atlanta were not named commonly and were not included in the construction of geographic range. The latitude and longitude of each site named was

obtained using either automated procedures or direct measurement using a global positioning system device at the site. In all, 3,771 unique sites were identified and geocoded.

We asked each seed to nominate a contact as the next person in a chain. That person was interviewed, tested, and asked to name a third person in the chain. We maintained a running tally of persons named, deduplicating “on the fly” by matching persons with the same demographics and using the field team’s knowledge of the clients for final confirmation. For the large number of contacts named but not interviewed, we used an algorithmic matching program based on contact characteristics. The final network from each seed was thus a connected component of three major respondents, all their interviewed contacts, and all their unvisited contacts. These 30 networks contained 927 interviewed persons, and a total population of 11,323 distinct individuals.

Risk assessment

We defined compound risk as performing multiple different risky acts with multiple partners. We chose six variables as representative of such risks: having sex with 10 or more different partners of either gender in the past 6 months; having sex with six or more male partners in the past 6 months; any history of injection drug use; history of ever engaging in sex work; history of ever having had sex with an IDU; history of anal sex in past 6 months. Compound risk was a binary variable: negative for zero or one of these six risk factors, positive for two or more risk factors. We assessed the frequency of personal and dyadic characteristics in the comparison of the high-burden with the lower-burden areas. These analyses were conducted using SAS, version 9.3.⁹

Geographic assessment

For each individual who was interviewed, we constructed a polygon of personal geographic range based on the coordinates of the sites named. The centroid of this polygon was used to calculate that individual’s geographic locus for computation of geographic distance between individuals and for comparison of geographic distance with social distance, measured as the geodesic (shortest path) between two persons. For each of the 30 networks, the polygons of each member were overlaid on a map of Atlanta and the extent of overlap assessed. We used the extent of overlap to estimate geographic compactness of the network (for example, the size of the area that contained overlap of 50% of network members). These calculations and visualizations were conducted using Arc GIS.¹⁰

Network structure

We measured the structural characteristics of the 30 networks, each of which was a connected component by design, using UCINET 6.¹¹ The major differences among them would then center on the number of persons named in common by those interviewed and the types of connections. All ties were considered undirected, on the assumption that sexual contact, drug use, and acquaintanceship had to be mutual activities. Time of first and last contact was obtained, but the networks were temporally compressed for this analysis. We calculated the major network parameters, using several measures of degree, structure, and cohesion. In particular, we used point connectivity—defined as the number of non-overlapping paths between persons, equivalent to the number of intervening persons who

would need to be removed to “disconnect” two persons—as a measure of the extent of interconnectedness in a subgroup. We examined the distribution of the number of contacts named per person by calculating the power law fit (where “a” is the power law coefficient in the formula $y = bx^{-a}$) using the method described by Newman.¹²

For this study, we use “dyad” to refer to any two people who are connected to each other by a pathway of any length. We combined network and geographic structure by examining the joint distribution of geodesic network distance (the minimum number of nodes separating each connected dyad) and geographic distance (the distance in kilometers separating each possible dyad, measured, as noted above, from the centroid of each person’s polygon).⁴ The content of each cell of the geodesic-geographic matrix is the number of dyads whose members were at that specific social and geographic distance from each other (Supplemental Tables 1 and 2). The matrix of social distance and geographic distance permitted simultaneous assessment of differences in compactness of networks in lower and higher risk areas. Because the higher risk geographic area was greater than that of the lower risk area, we compared the proportional distribution of geographic distances, summed over all geodesic distances to determine if the greater land mass contributed to differences in the joint geographic-geodesic distributions.

Results

Participant characteristics

The prevalence of HIV in the lower risk area was 12% and, in the higher risk area, 19%. In general, there were few major differences between the higher and lower risk groups. Greater than 90% of study participants were African American, with a male to female ratio of 1.1 (Table 1). Approximately 70% were single, and about 40% had a high school diploma or GED. The unemployment proportion was similar in men and women (49–59%), but homelessness was considerably less in the lower risk areas and among women (male: lower, higher: 10%, 27%; female: lower, higher: 6%, 21%). Over 90% of men in both areas had been incarcerated; among women, the proportion varied from 63% (lower risk) to 77% (higher risk).

Less than 30% of men and women in both areas classified their health as excellent (Table 2). The majority of both sexes and persons in both areas self-identified as heterosexual, but a greater proportion of women stated that they were bisexual, especially in the higher risk area (15.8% vs. 1.3%). A history of gonorrhea, syphilis, chlamydia, and (among women) trichomoniasis was common in both areas. The vast majority of participants smoked (88% to 95%). Crack use was about twice as great in the higher risk areas and was similar in men and women. Heroin use, and drug injection were highest among men in the higher risk area (22%), and anal sex in the past 6 months was highest among women in the higher risk areas (12%).

Compound risk characteristics

Of the six components of compound risk (Table 3), 6 or more male sex partners (counting men and women together, since men who had sex with men were a small proportion of the

total) was the most common (26.2%). Among the 797 participants for whom full information was available (398 in the lower risk area; 397 in the higher risk area), 10.9% exhibited two or more of the components of compound risk. There were 24 such persons in the lower risk area (5.7%) compared to 73 persons in the higher risk area (15.5%) (z-score for difference of proportions = 4.6; exact p-value = 3.1×10^{-5}).

Geographic characteristics

The matrix of social distance (shortest number of steps between any two participants) and geographic distance (length in kilometers between the centroids of any two participants) for the lower risk area was 10×21 , and for the higher risk area was 8×21 . For the lower risk area, 19.4% of dyads were within 1.0 km or less of each other; for the higher risk area, 29.1% were within 1.0 km of each other. Comparing successively larger square subsets of each matrix, from a social distance of 1 (direct contacts) and a geographic distance of 1, to a social distance of 8 and a geographic distance of 8, at every point the higher risk areas demonstrates greater compactness (Figure 2). The 1×1 cell in both the higher and lower risk areas both contain 0.8% of the dyads. The 8×8 square contains 66% of the dyads in the higher risk area and 57% of the dyads in the lower risk area. The proportional distributions of geographic distance demonstrated a ratio of 3:2 (higher to lower) for distances of 1.0 km or less, but the distributions were almost identical over the remainder of the distances (Figure 3).

The amount of overlap of the geographic range of participants in higher and lower areas also differed substantially. On average, the area that contained overlap of 50% of lower risk group members was 10.6 kms; for higher risk group members, 50% overlap occurred over 3.8 kms. When the 15 groups in each area are ranked by the size of the overlap of 50% of participants, the difference in distributions is evident (Figure 4), and demonstrates a markedly greater compactness within the higher risk area.

Network characteristics

The 30 networks varied in size from 111 to 512 nodes, with an average of 248.5 in the lower risk area, and 288.3 in the higher risk area. The screening set of network measures that we used to compare the higher and lower configurations yielded several small differences (Table 4). On average, there were more components per network in the higher risk areas (7.27) compared to the lower risk areas (4.33) but the proportion of persons in the largest component in higher and lower networks were almost identical (lower, 0.79; higher, 0.77). The degree mean and variance were virtually the same in both areas as was concurrency, transitivity, the average distance between nodes and the network diameters (the largest geodesic). Using the degree (number of contacts) of each node as the measure of centrality, the average degree was marginally larger in the lower risk group (9.32 (lower) vs. 8.73 (higher)). Average point connectivity (the number of disjoint paths between any two connected nodes) was slightly higher in the higher risk area (0.98 vs. 0.91) (Table 4, legend).

As with other network measures, the distributions of numbers of contacts, calculated as a power law model and displayed as the log of the number of contacts against the log of the proportion of contacts, were similar (Figure 5). The power law coefficients were 2.4 for the

lower risk area and 2.3 for the higher risk area. We did not attempt to determine if the power law fit was the best modeling approach^{13,14} but used it as an heuristic for comparison of the two distributions.

Discussion

The motivating premise for this study—that endemic HIV transmission in at-risk communities is maintained through the interaction of intense compound risk taking, geographic compactness, and a conducive network structure—is based on the idea that, within such a configuration, new partners are persons who are likely to be part of the same network, and therefore face a higher prevalence than would be present in a broader, less confined population. The hypothesis is supported in part by the data. We had postulated that differences in all three elements would be apparent in communities at differing risk for HIV transmission. In comparing a higher risk area (with a 19% prevalence in this study) to a lower risk community (with a 12% prevalence) we demonstrated that the former had a three times greater frequency of compound risk (15.5% vs. 5.7% for the occurrence of two or more major risks in individuals), and evidence of substantially higher geographic compactness (greater overlap of personal geographic space and more intense joint social and geographic proximity). The social network structure in both groups were similar, however. Both higher and lower risk areas displayed evidence of highly connected and interactive groups, with large connected components that contained on average three-fourths of the group members in each of the 30 networks examined, high levels of concurrency, and a long tail to the right in the distribution of numbers of contacts per person.

These data suggest an alternative view of the maintenance of HIV endemicity in at-risk communities than the one originally suggested. Though the data are not presented here in detail, most of the network parameters we calculated varied over a substantial range, but that variation was similar in the two risk areas. The networks, then, form an infrastructure for transmission. The amount of transmission depends on other factors—here, the intensity of risk and geographic compactness—that can promote or obstruct viral acquisition. A third major factor would be the effect of treatment programs on community viral load,¹⁵ but this study was conducted prior to major policy changes that recommended universal HIV treatment regardless of clinical status.¹⁶

The importance of networks in disease transmission is firmly established,¹⁷ though considerable work is currently devoted to understanding the specific relationship of network configuration and transmission.^{18,19} The importance of the interaction between risk taking and network configuration is also well established, especially to highlight the independent role that network structure, such as concurrency,^{20–23} plays in transmission dynamics. Geographic aspects of network relationships have received increasing attention in recent years.²⁴ Techniques for visualizing networks in real space have been developed, and considerable theoretical and modeling endeavors have attempted to distinguish the effects of geographic proximity and network relationships. A number of empirical studies have explored the general relationship of geographic distance to social ties, and have examined the joint geographic-social effect in criminal activity, communications networks and

economic activity (documentation of these observations is provided in supplemental references 30s–53s).

For empirical studies of infectious disease, joint social and spatial analysis has received less attention. Heimer et al.²⁵ mapped the location of 788/900 participants in a study of IDU and HIV infection in St. Petersburg, Russia and found that the HIV positives (29.9%) were tightly clustered in 5% of the populated areas of the city. Network ties among participants were not reported. In an analysis of diarrheal illness in Ecuador, Bates et al.²⁶ found that diarrheal risk was higher in less dispersed communities and lower in communities with lower social connectedness (lower degree). Giebultowicz and colleagues²⁷ used kinship networks and the distance between persons in different neighborhoods to examine cholera distribution in Matlab, Bangladesh over a 21-year period. Their result suggested that social ties had a less consistent relationship to cholera clustering than did proximity and (probably) unmeasured environmental variables. In a subsequent analysis,²⁸ the same group using female kinship networks and the configuration of roads to demonstrate the clustering of diarrheal illness near roads, and the lesser effect of networks. Several studies have used geographic information to examine disease clustering (with the presumption of network relationships). In the single study directly relevant to this one, Hixson et al.²⁹ demonstrated marked clustering (60%) of prevalent HIV cases in downtown Atlanta, with a prevalence within the cluster area more than four times greater than the prevalence outside it. Nearly half of the identified HIV providers in the Atlanta area were situated within the area of highest clustering, and their geo-position was used to determine presumed travel times. Like several previously cited, however, this study did not examine actual network connections.

There is, then, little available in the extant literature for comparison with our results. Given the sample of 30 networks, it is tempting to assume that these networks are representative of those that would be encountered in disadvantaged at-risk communities. It would be injudicious, however, to assume that our joint examination of risk, network attributes, and geographic configuration is generalizable. Several other limitations constrain conclusive statements. These data are cross-sectional comparisons, and it is well understood that behavior at the time of acquisition of HIV may not be reflected in current behavior. The assumption for a stronger conclusion would have to be that the transmission milieu reflects that of an earlier time. Another caveat for these results is that the temporal nature of contact relationships is not considered, and chronology may play a critical role in understanding network contributions to transmission.³⁰ Because the lower risk geographic area was substantially larger than that of the higher risk area (a ratio of approximately 3:1), some of the distance differences could be attributable to a population that was more “spread out.” The distribution of geographic distances in each area, summed over all social (geodesic) distances, suggests that a greater proportion of dyads were at 1 km or less from each other, but that distance between dyads was otherwise similar. It is likely that neighborhood aggregation had a greater effect than overall land mass. It is possible that seeds in the lower risk areas had less risky behavior than seeds in the higher risk area—a self-fulfilling prophecy. Though we made every effort to enroll similar people in both areas (and their demographics and behavioral characteristics were similar; data not shown) an intrinsic difference in seeds may still be a possibility. Such differences could potentially affect the results if respondents have systematic differences in willingness to communicate

information about partners. A segment of our interview includes a post-interview assessment of accuracy and veracity and we detected no difference between the higher and lower prevalence areas (data not shown).

The alternative hypothesis that these data suggest should not be interpreted to mean that networks are not important. Rather, they might indicate that a certain “minimum” network is required to maintain endemicity, with, as noted, other factors that determine the prevalence set point. This concept lends itself well to modeling and simulation, since each of the factors invoked—risk configuration, geographic relationships, (and treatment)—can be modeled in the context of the empirical network observations. Further exploration of the interaction of the multiple factors that affect transmission is likely to provide greater insight into the dynamics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by NIH/NIDA grant R01 DA019393. The authors wish to thank Mr. Kelvin Parker and Ms. Eleanor Hillman for their contributions to this study and to the community, Mr. Stephen Muth for his assistance in earlier analyses of this work, to Ms. Evelyn Olansky for her assistance with current analyses, and to Professor James Moody for advice on an earlier version of this manuscript.

Reference List

1. Johnson AS, Hall HI, Hu X, Lansky A, Holtgrave DR, Mermin J. Trends in diagnoses of HIV in the United States, 2002–2011. *JAMA*. 2014; 312:432–4. [PubMed: 25038362]
2. HIV Surveillance Report. 2014; 26 Accessed August 10, 2016, at <http://www.cdc.gov/hiv/library/reports/surveillance/>.
3. UN AIDS report on the Global AIDS Epidemic. 2012. Accessed 8/12/2016, 2015, at http://www.unaids.org/en/resources/campaigns/20121120_globalreport2012/globalreport/
4. Rothenberg R, Muth SQ, Malone S, Potterat JJ, Woodhouse DE. Social and geographic distance in HIV risk. *Sex Transm Dis*. 2005; 32:506–12. [PubMed: 16041254]
5. Rothenberg RB. Maintenance of endemicity in urban environments: a hypothesis linking risk, network structure and geography. *Sex Transm Inf*. 2007; 83:10–5.
6. Atlanta Regional Commission: Census 2010. 2014. Accessed 1/13/2015, 2015, at <http://www.atlantaregional.com/info-center/2010-census>
7. Atlanta, Georgia (GA) Zip Code Map - Locations, Demographics. 2014. at <http://www.city-data.com/zipmaps/Atlanta-Georgia.html>
8. Rothenberg RB, Long D, Sterk C, et al. The Atlanta urban networks study: a blueprint for endemic transmission. *AIDS*. 2000; 14:2191–200. [PubMed: 11061661]
9. What’s new in SAS 9.3. Cary, NC: SAS Institute Inc; 2012.
10. ESRI. ArcGIS Desktop: release 10. Redlands, CA: Environmental Systems Research Institute; 2011.
11. Borgatti SP, Everett M, Freeman L. Ucinet 6 for Windows: Software for Social Network Analysis. Natick MA: Harvard: Analytic Technologies. 2002
12. Newman MEJ. Power laws, Pareto distributions and Zipf’s law. *www arXiv org arXiv:condmat/0412004 v22005*.
13. Clauset A, Cosma RS, Newman MEJ. Power law distributions in empirical data. *arXiv*. 2009; 2:1–43. 0706.1062v.

14. Hamilton DT, Handcock M, Morris M. Degree distributions in sexual networks: A framework for evaluating evidence. *Sex Transm Dis.* 2008; 35:30–40. [PubMed: 18217224]
15. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *N Engl J Med.* 2016
16. Watts DJ, Dodds PS, Newman ME. Identity and search in social networks. *Science.* 2002; 296:1302–5. [PubMed: 12016312]
17. Danon L, Ford AP, House T, et al. Networks and the Epidemiology of Infectious Disease. *Interdisciplinary Perspectives on Infectious Diseases.* 2011; 2011:28.
18. Keeling MJ, Eames KT. Networks and epidemic models. *J R Soc Interface.* 2005; 2:295–307. [PubMed: 16849187]
19. Eames KTD, Keeling MJ. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *PNAS.* 2002; 99:13330–5. [PubMed: 12271127]
20. Goodreau SM. A decade of modelling research yields considerable evidence for the importance of concurrency: a response to Sawers and Stillwaggon. *Journal of the International AIDS Society.* 2013; 14doi: 10.1186/758-2652-14-12
21. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnership, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS Behavior.* 2013; doi: 10.1007/s10461-010-9858-x
22. Gorbach PM, Drumright LN, Holmes KK. Discord, discordance, and concurrency: comparing individual and partnership-level analyses of new partnerships of young adults at risk of sexually transmitted infections. *Sex Transm Dis.* 2005; 32:7–12. [PubMed: 15614115]
23. Gorbach PM, Stoner BP, Aral SO, Whittington WLH, Holmes KK. It takes a village: understanding concurrent sexual partnerships in Seattle, WA. *Sex Transm Dis.* 2002; 49:453–62.
24. Adams J, Faust K, Lovasi GS. Capturing context: Integrating spatial and social network analyses. *Social Networks.* 2012; 34:1–5.
25. Heimer R, Barbour R, Shaboltas AV, Hoffman IF, Kozlov AP. Spatial distribution of HIV prevalence and incidence among injection drug users in St Petersburg: implications for HIV transmission. *AIDS.* 2008; 22:123–30. [PubMed: 18090400]
26. Bates SJ, Trostle J, Cevallos WT, Hubbard A, Eisenberg JNS. Relating diarrhea disease to social networks and geographic configuration of communities in rural Ecuador. *American Journal of Epidemiology.* 2007; 166:1088–95. [PubMed: 17690221]
27. Giebultowicz S, Ali M, Yunus M, Emch M. The Simultaneous Effects of Spatial and Social Networks on Cholera Transmission. *Interdisciplinary Perspectives on Infectious Diseases.* 2011:1–6. Article ID 604372.
28. Perez-Heydrich C, Furgurson JM, Giebultowicz S, et al. Social and spatial processes associated with childhood diarrheal disease in Matlab, Bangladesh. *Health & Place.* 2013; 19:45–52. [PubMed: 23178328]
29. Hixson BA, Omer SB, del Rio C, Frew PM. Spatial clustering of HIV prevalence in Atlanta, Georgia and population characteristics associated with case concentrations. *Journal of Urban Health.* 2011; 88:129–41. [PubMed: 21249526]
30. Moody J. The Importance of Relationship Timing for Diffusion. *Social Forces.* 2002; 81:25–56.

Summary

Endemic HIV transmission in Atlanta is posited to require a conducive network configuration, but the prevalence set point is determined by compound risk and geographic contiguity.

Figure 1.
Distribution of cumulative AIDS cases, 1998–2003, Atlanta (Fulton County), GA.

Figure 2. Comparing the higher and lower risk areas, the proportion of dyads enclosed within succeeding larger square subsets of the matrix of social distance times geographic distance.

Figure 3.
Distribution of geographic distance in the higher and lower prevalence areas (all geodesic (social) distances combined)

Figure 4. Comparison of higher and lower risk areas with regard to the size in kilometers of overlap of 50% of network members. The 15 networks in each of the two areas are ranked by size of the overlap and displayed as pairs. For example, the 12th pair consists of the group in each area that was ranked 12th largest.

Figure 5.

Log-log plot of number of contacts (k) and proportion of the number of contacts ($p[k]$): high risk and low risk areas. The exponent (α) associated with the power law curve ($y = bx^{-\alpha}$) for the higher risk area was 2.3, and for the lower risk area was 2.4.

Table 1

Sociodemographic characteristics of the higher risk and lower risk areas

	Male		Female	
Socio-demographics	Lower risk	Higher risk	Lower Risk	Higher Risk
African American (%)	100.0	95.8	97.4	93.7
Age (mean)	34.2	32.0	41.1	36.2
Marital Status (%)				
Single	73.5	65.4	73.3	73.8
Married	10.0	9.3	10.5	9.1
Divorced	10.0	13.9	6.8	11.8
Separated	5.2	9.7	6.8	2.3
Widowed	1.3	1.7	2.6	3.2
Education (%)				
K to 8	3.0	4.6	3.1	3.2
9 to 11	48.7	40.9	43.5	38.5
GED	6.1	7.2	5.8	4.5
12th	30.9	32.1	35.1	34.8
Bachr Degree	1.3	0.8	.	2.7
Religious affiliation (%)				
Christian, NOS	46.1	27.4	31.9	23.5
Catholic	0.43	2.53	1.57	3.17
Baptist	34.3	47.7	47.6	56.6
Methodist	3.04	1.69	1.57	1.81
Muslim	0.87	3.38	.	.
None	12.6	11.8	13.1	8.1
Atheist	.	.	0.5	0.5
Unemployed (%)	55.0	52.0	59.0	49.0
Homeless (%)	10.0	27.0	6.0	21.0
Sources of Income (%) (not exclusive categories)				
Paid job	44.0	39.0	40.0	33.0
Benefits	13.0	15.0	13.0	8.0
Family	48.0	35.0	62.0	54.0
Partner	43.0	34.0	61.0	67.0
Friend	37.0	40.0	34.0	47.0
Other	38.0	44.0	29.0	38.0
Adverse interactions (%)				
Ever incarcerated or detained	90.0	94.0	63.0	77.0
Currently on probation	18.0	19.0	10.0	15.0
Currently on parole	1.0	1.0	0.0	1.0
Threatened with a weapon, past 6 m	13.0	17.0	10.0	20.0

Table 2

Health Status and Risks of persons in the higher risk and lower risk areas

	Male		Female	
General Health Status	Lower risk	Higher risk	Lower Risk	Higher Risk
Excellent	29.1	20.3	25.1	23.5
Good	48.7	53.2	45.5	49.8
Fair	21.3	21.9	25.7	22.2
Poor	0.87	4.64	2.62	4.52
HIV-positive	2.2	5.5	1.1	9.1
Sexual orientation				
Heterosexual	90.4	96.6	84.3	81.0
Bisexual	8.7	1.3	12.6	15.8
Gay	0.9	1.7	1.1	1.4
Lesbian	.	.	1.1	1.4
Transgender	1 person		11 persons	
Ever diagnosed with:				
Gonorrhea	24.0	40.0	29.0	29.0
Syphilis	5.0	14.0	8.0	15.0
Chlamydia	13.0	12.0	39.0	29.0
Genital Herpes	1.0	3.0	6.0	5.0
Hepatitis C	2.0	10.0	2.0	5.0
Trichomonias	3.0	4.0	28.0	30.0
Other STD	2.0	3.0	2.0	4.0
Drugs and Sex				
Currently smoke cigarettes	92.0	94.0	88.0	95.0
Ever used crack	34.0	62.0	29.0	56.0
Ever used heroin	7.0	30.0	3.0	10.0
Ever injected any drug	4.0	22.0	3.0	11.0
Had anal sex, 6 mos	6.0	8.0	9.0	12.0
Had oral sex, 6 mos	65.0	68.0	62.0	69.0

Table 3

Components of compound risk (total N = 894)

Components	N	%
10 or more total sex partners, 6 m	67	7.5
6 or more male sex partners, 6 m	234	26.2
Ever injected drugs, lifetime	14	1.6
Ever engaged in sex work, lifetime	37	4.1
Ever had sex with an IDU, lifetime	24	2.7
Anal sex, 6 m	80	9
Distribution of component frequency		
(the percent of persons with a given number of components)	0	64.1
	1	25.1
	2	7.4
	3	2.8
	4	0.6
	5	0.1
	6	0.0
Comparison of compound risk*		
	N	%
Lower Risk Area	24	5.7
Higher Risk Area	73	15.5
Areas combined	97	10.9

Table 4

Comparison of network characteristics in higher and lower risk areas

Network characteristics	Lower risk	Higher risk
HIV PREVALENCE	0.12	0.17
Number of nodes (mean)	248.53	288.27
Number of Ties (mean)	282.20	318.60
Number of components (mean)	4.33	7.27
Size of Largest component (mean)	196.40	218.93
Proportion of persons in the largest component	0.79	0.77
Degree (mean)	2.19	2.13
Degree (variance)	12.71	12.73
Concurrency (mean)	7.00	7.10
Network Centrality (based on degree)	9.32	8.73
Transitivity	0.02	0.01
Betweenness	17.70	24.73
Average distance between nodes (mean)	4.22	4.19
Diameter (largest average distance) (mean)	7.53	7.40
Point connectivity (mean) *	0.91	0.98

* Point connectivity is the number of nodes that would need to be removed to disconnect one node from another. It is equivalent to the number of disjoint paths between two nodes. The average point connectivity is calculated for each person in the network with 1 or more disjoint paths. The average of those persons is the mean point connectivity for that network, and we report here the average of those means for the lower risk and higher risk networks. Persons with no disjoint connections to others are omitted from this calculation.