

# HIV-1 Infection and Transmission Networks of Younger People in Chicago, Illinois, 2005-2011

Public Health Reports  
2017, Vol. 132(1) 48-55  
© 2016, Association of Schools and  
Programs of Public Health  
All rights reserved.  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0033354916679988  
journals.sagepub.com/home/phr



Ethan Morgan, BS<sup>1,2</sup>, Alexandra M. Oster, MD<sup>3</sup>,  
Stephanie Townsell, MPH<sup>4</sup>, Donna Peace, MPH<sup>4</sup>,  
Nanette Benbow, MS<sup>4</sup>, and John A. Schneider, MD, MPH<sup>1,2,5</sup>

## Abstract

**Objective:** Analysis of HIV nucleotide sequences can be used to identify people with highly similar HIV strains and understand transmission patterns. The objective of this study was to identify groups of people highly connected by HIV transmission and the extent to which transmission occurred within and between geographic areas in Chicago, Illinois.

**Methods:** We analyzed genetic sequences in the HIV-1 *pol* region in samples collected from people participating in the VARHS program in Chicago during 2005-2011. We determined pairwise genetic distance, inferred potential transmission events between HIV-infected people whose sequences were  $\leq 1.5\%$  genetically distant, and identified clusters of connected people. We used multivariable analysis to determine demographic characteristics and risk attributes associated with degree of connectivity.

**Results:** Of 1154 sequences, 177 (15.3%) were tied to at least 1 other sequence. We determined that younger people, men, non-Hispanic black people, and men who have sex with men were more highly connected than other HIV-infected people. We also identified a high degree of geographic heterogeneity—48 of 67 clusters (71.6%) contained people from >1 Chicago region (north, south, or west sides).

**Conclusion:** Our results indicate a need to address HIV transmission through the networks of younger non-Hispanic black men who have sex with men. The high level of geographic heterogeneity observed suggests that HIV prevention programs should be targeted toward networks of younger people rather than geographic areas of high incidence. This study could also guide prevention efforts in other diverse metropolitan regions with characteristics similar to those of Chicago.

## Keywords

HIV, networks, MSM

Phylogenetic and transmission network analyses can be used to examine networks of human immunodeficiency virus (HIV) infection, with the results highlighting differences in transmission among subtypes,<sup>1</sup> over time,<sup>2</sup> and by presence of drug resistance.<sup>3</sup> These methods also can be used to assess differences across geographic regions,<sup>4</sup> to understand the characteristics of people in social clusters,<sup>5</sup> and to examine the extent to which transmission occurs among demographic or risk groups.<sup>6,7</sup> Perhaps most important, transmission network data can be used to prioritize interventions to disrupt transmission and, ultimately, to control outbreaks of infectious diseases,<sup>8,9</sup> including HIV.

In the United States, the annual rate of new HIV diagnoses is 15.0 cases per 100 000 population, but black people, who have a rate of 55.9 cases per 100 000 population, are disproportionately affected. Two-thirds of new diagnoses occur

<sup>1</sup> Department of Public Health Sciences, University of Chicago, Chicago, IL, USA

<sup>2</sup> Chicago Center for HIV Elimination, University of Chicago, Chicago, IL, USA

<sup>3</sup> Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA

<sup>4</sup> Chicago Department of Public Health, Chicago, IL, USA

<sup>5</sup> Department of Medicine, University of Chicago, Chicago, IL, USA

## Corresponding Author:

John A. Schneider, MD, MPH, Global Health Programs, Department of Medicine, University of Chicago, 5841 S Maryland Ave, MC 5065, Chicago, IL 60637, USA.

Email: jschnei1@medicine.bsd.uchicago.edu

among men who have sex with men (MSM), and that proportion increased from 77.5% in 2010 to 82.7% in 2014.<sup>10</sup> Compared with the burden in the United States as a whole, the burden of HIV infection in Chicago is more severe (40.4 per 100 000 population). In 2013, the rate of new HIV diagnoses in Chicago was 2.5 times that of the United States overall. However, since 2009, the rate of new diagnoses has declined among all risk categories except MSM, who have experienced a 5% increase in new diagnoses. The rate of new HIV diagnoses in Chicago among black people (64.0 cases per 100 000 population) is double that of white people (28.1 cases) and Hispanic people (28.0 cases).<sup>11</sup>

The geography of Chicago is a key factor in HIV risk and transmission. According to American Community Survey estimates,<sup>12</sup> demographic characteristics vary widely across Chicago's 200 neighborhoods. The neighborhoods on the north side of Chicago are mostly white, with the percentage of white people in some neighborhoods exceeding 80% of the population, although certain neighborhoods, particularly those on the far north side, have a diverse population of black, white, and Hispanic people. The neighborhoods on the south side are mostly black; at least 80% of the population in most of these neighborhoods is black. Those on the west side of Chicago are divided between majority black and majority Hispanic populations, although a few neighborhoods have a single racial/ethnic group exceeding 80% of the population.<sup>12</sup> The south and west side neighborhoods have the highest percentage of households living below the federal poverty level and the lowest per capita income.<sup>13</sup> Neighborhoods with high HIV diagnosis rates are located throughout Chicago and include Edgewater (north, 100 cases per 100 000 population), Uptown (north, 132 cases), West Garfield Park (west, 100 cases), Washington Park (south, 98 cases), and Pullman (south, 89 cases).<sup>11</sup>

Studies have examined the dynamics of HIV transmission and the characteristics of people who are potential transmission partners because of their highly similar HIV strains.<sup>6</sup> However, given that geographic differences exist in demographic characteristics, risk, and HIV transmission, analyzing data at a finer geographic level can aid public health efforts. Understanding transmission among different groups and geographic areas in Chicago may assist with targeting local interventions and guiding public policy. The objective of this study was to identify groups of people highly connected by HIV transmission and the extent to which transmission occurred within and between geographic areas in Chicago, Illinois.

## Methods

We analyzed HIV-1 *pol* sequences collected from January 2005 through March 2011 from newly diagnosed cases among people in Chicago. From these data, we (1) determined genetic distances between pairs of sequences, (2) inferred potential transmission events (ie, ties) between

people whose sequences were  $\leq 1.5\%$  genetically distant, (3) identified clusters of connected people, (4) determined the degree of connectivity among people, (5) examined the demographic characteristics and risk attributes associated with degree of connectivity, and (6) established the extent to which transmission occurred within and between geographic areas in Chicago.

## Study Population

We generated nucleotide sequences through routine HIV testing as part of the Variant, Atypical, and Resistant HIV Surveillance (VARHS) system,<sup>14</sup> a drug-resistance surveillance program conducted by the Centers for Disease Control and Prevention (CDC) from 2004 through 2012. The VARHS program was conducted during 2 periods, 2004-2007 and 2008-2012, in several cities and states as described elsewhere.<sup>7</sup>

We limited site selection in Chicago to locations that had the facilities required for specimen handling and storage. Sites chosen for specimen collection included all Chicago Department of Public Health clinics, department associates (at Jackson Park), and Chicago-area county facilities (CORE Center of Cook County, John H. Stroger Jr Hospital of Cook County, Provident Hospital of Cook County, and Winfield Moody Health Center in Chicago). The VARHS program collected remnant blood specimens taken at least 3 months after a new diagnosis of HIV infection in Chicago. All people at these facilities were eligible to participate in the VARHS program, and sensitivity analyses compared VARHS participants with nonparticipants. In certain instances, specimens were collected from people who were believed to be newly diagnosed but were later determined to have been diagnosed earlier in another state; these specimens were excluded from this study. A CDC-contracted laboratory used polymerase chain reaction, which included protease and at least the first 900 base pairs of reverse transcriptase, to sequence a portion of the *pol* region. We restricted analysis to sequences from people who were antiretroviral naïve and at least 13 years of age. Each sequence represented 1 person.

We used all specimens collected during the study period for this analysis, with 2 exclusions: (1) those with an HIV diagnosis before 2005, because these people were previously diagnosed with HIV and thus were not new HIV diagnoses, and (2) those who were diagnosed with HIV in another state.

We studied sequence data from Chicago only, and we obtained all sequences from VARHS from 2005 to 2011. CDC and the Chicago Department of Public Health Institutional Review Board determined VARHS to be exempt from human subjects protection review, and the University of Chicago considered the study exempt from institutional review board review. Because the data in this analysis were collected for the purpose of public health surveillance, consent was not required.

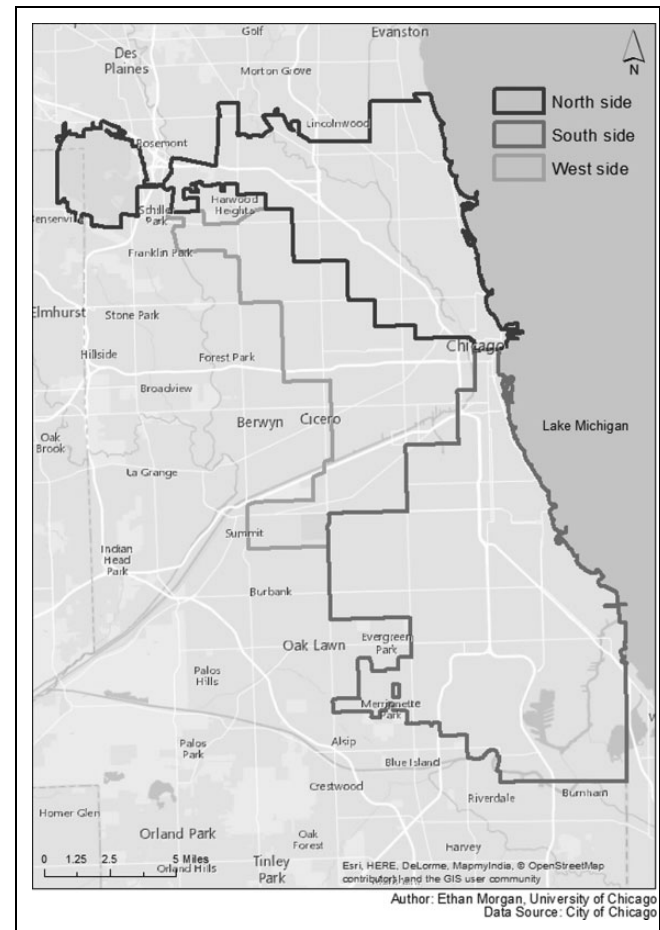
### Transmission Network Inference

We aligned all sequences to the HXB2 reference sequence through MUSCLE (Multiple Sequence Comparison by Log-Expectation; European Bioinformatics Institute, Hinxton, UK) multiple-sequence alignment<sup>15</sup> and SeaView software (Rhône-Alpes Bioinformatics Center, Pôle Rhône-Alpes de Bioinformatique).<sup>16</sup> We performed phylogenetic tree analyses via the neighbor-joining method,<sup>17</sup> with distance calculated by Kimura's 2-parameter analysis,<sup>18</sup> and PHYLIP 3.696 software (PHYLogeny Inference Package; University of Washington, Seattle),<sup>19</sup> bundled in the SeaView software package. As previous studies did to identify recent transmission events, we defined a transmission event as a genetic distance  $\leq 1.5\%$  between *pol* sequences.<sup>5,20</sup> We defined a node as a single person and a tie as a single connection between 2 people; in our analysis, a tie constituted a potential transmission event. We defined the degree of connectivity for a person as the total number of ties to all other nodes (eg, a person with 3 ties would have a degree of 3). We defined a cluster as a group of at least 2 people linked by at least 1 potential transmission event. We performed all cluster visualizations through NodeXL 1.0.1.340.<sup>21</sup>

Similar to past studies, ours utilized a network framework, rather than a phylogenetic tree, to visualize inferred clusters of HIV-infected people. Although a phylogenetic tree would also be suitable, the network framework allows for easier visualization of clusters and for the use of network software that can quickly analyze the degree of connectivity for each person. Note that a phylogenetic tree must still be used to determine the genetic distance between individual viral genetic sequences.

### Demographic and Risk Characteristics

We combined race and ethnicity into a categorical variable, defined as non-Hispanic white, non-Hispanic black, Hispanic, and "other." We categorized age at diagnosis as  $\leq 24$ , 25-29, 30-39, 40-49, and  $\geq 50$ . We categorized year of HIV diagnosis as 2000 or earlier, 2001-2005, 2006-2010, and after 2010. We determined location of residence in Chicago (Figure 1) through resident ZIP code at the time of HIV diagnosis and classified it as north side, south side, west side, or "other," which included surrounding suburbs and non-Chicago ZIP codes. We defined transmission category as 5 mutually exclusive groups: MSM, person who injects drugs (PWID), heterosexual contact, MSM and PWID, and "other." The "other" category included people who had received clotting factor or had received a transfusion or transplant, those who had a perinatal exposure with HIV diagnosed at age  $\geq 13$ , and those with no identified or reported risk. We used the Pearson  $\chi^2$  test to assess differences in demographic characteristics of those in a cluster and those not in a cluster.



**Figure 1.** Region of residence (north, south, and west) in Chicago, Illinois, defined by ZIP code at the time of human immunodeficiency virus (HIV) diagnosis for a sample of 1154 people: Variant, Atypical, and Resistant HIV Surveillance system, 2005-2011. People with a ZIP code not in one of these regions were classified as "other." Source of data on regions of residence: City of Chicago.

### Statistical Analyses

The outcome for this analysis was degree of connectivity to other people in a cluster. We first analyzed the association of all variables with degree of connectivity through unadjusted Poisson regression analysis.<sup>22</sup> We included all people, including those with a degree of 0, in all models. We then used multivariable Poisson regression to estimate the association between the degree of connectivity as the dependent variable and the transmission category as the main independent variable. We included all covariates identified as significant at  $P \leq .05$ , using the Wald test statistic, in the multivariable regression model. We assessed effect modification through cross-products between transmission category and each covariate. To assess potential misclassification because of self-reported transmission categories, we also performed a sensitivity analysis in which we classified those in the "other" category as MSM, repeating the aforementioned analyses. We performed all analyses with the R statistical software package, version 3.1.1.<sup>23</sup>

**Table 1.** Characteristics of a sample of HIV-infected people (n = 1154) participating in the Variant, Atypical, and Resistant HIV Surveillance system, stratified by presence in a cluster, Chicago, Illinois, 2005-2011

Characteristic	No. of Participants (Column %)	No. of Participants, by Cluster <sup>a</sup> (Row %)		P Value
		In a Cluster	Not in a Cluster	
Total	1154 (100.0)	177 (15.3)	977 (84.7)	
Age, y				
≤24	87 (7.5)	27 (31.0)	60 (69.0)	<.001
25-29	188 (16.3)	53 (28.2)	135 (71.8)	
30-39	348 (30.2)	53 (15.2)	295 (84.8)	
40-49	263 (22.8)	30 (11.4)	233 (88.6)	
≥50	268 (23.2)	14 (5.2)	254 (94.8)	
Sex				
Male	880 (76.3)	152 (17.3)	728 (82.7)	.001
Female	274 (23.7)	25 (9.1)	249 (90.9)	
Race/ethnicity				
Non-Hispanic black	764 (66.2)	122 (16.0)	642 (84.0)	.51
Non-Hispanic white	106 (9.2)	13 (12.3)	93 (87.7)	
Hispanic	266 (23.1)	41 (15.4)	225 (84.6)	
Other	18 (1.6)	1 (5.6)	17 (94.4)	
Transmission category				
MSM	497 (43.1)	115 (23.1)	382 (76.9)	<.001
Heterosexual	30 (2.6)	1 (3.3)	29 (96.7)	
PWID	57 (4.9)	3 (5.3)	54 (94.7)	
MSM and PWID	229 (19.8)	22 (9.6)	207 (90.4)	
Other <sup>b</sup>	341 (29.5)	36 (10.6)	305 (89.4)	
Year of diagnosis				
2000 or before	41 (3.6)	2 (4.9)	39 (95.1)	.003
2001-2005	73 (6.3)	2 (2.7)	71 (97.3)	
2006-2010	939 (81.4)	156 (16.6)	783 (83.4)	
After 2010	101 (8.8)	17 (16.8)	84 (83.2)	
Region <sup>c</sup>				
North side	268 (23.2)	44 (16.4)	224 (83.6)	.10
South side	401 (34.7)	67 (16.7)	334 (83.3)	
West side	285 (24.7)	47 (16.5)	238 (83.5)	
Other	200 (17.3)	19 (9.5)	181 (90.5)	

Abbreviations: HIV, human immunodeficiency virus; MSM, men who have sex with men; PWID, person who injects drugs.

<sup>a</sup>Defined as a group of at least 2 people linked by at least 1 potential transmission event. A transmission event was defined as a genetic distance  $\leq 1.5\%$  between *pol* sequences.

<sup>b</sup>“Other” categories include received clotting factor, received transfusion/transplant, perinatal exposure with HIV diagnosed at age  $\geq 13$ , and no identified or reported risk.

<sup>c</sup>Defined by ZIP code of residence at HIV diagnosis.

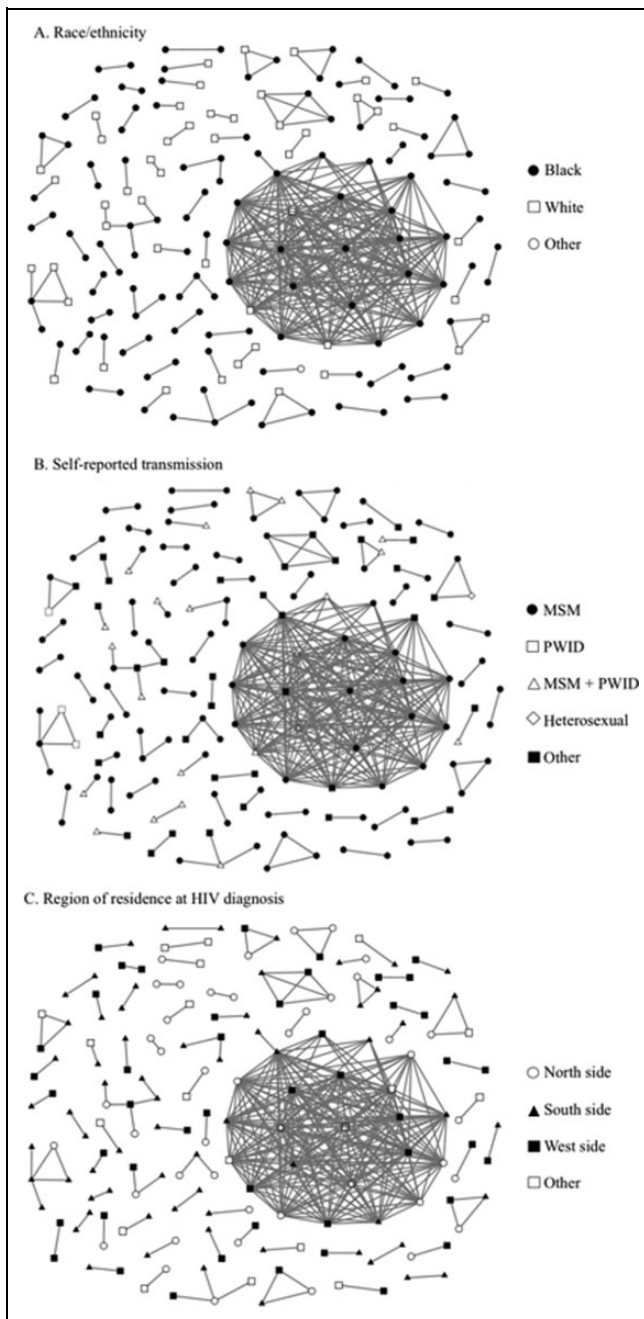
## Results

We collected specimens from 1154 people with a new HIV diagnosis; these people represented 14.0% of people with newly diagnosed infections (n = 8247) in Chicago. We found no significant difference between VARHS program participants and nonparticipants in sensitivity analyses. Our sample primarily included men (n = 880, 76.3%) and non-Hispanic black people (n = 764, 66.2%). The median age was 38 years (range, 16-89). People in our sample were evenly distributed among the north side, south side, west side, and other regions combined. Nearly half of the people were MSM (n = 497, 43.1%), and most diagnoses occurred during 2006-2010 (n = 939, 81.4%; Table 1).

Of 1154 genetic sequences, 177 (15.3%) had a tie to at least 1 other person. People in clusters had a median degree

of 1.0 and a mean degree of 3.5 (range, 1-22; Figure 2). The mean genetic distance among all people was 7.9% (standard deviation [SD] = 1.8%), and the mean genetic distance among people in a cluster was 0.9% (SD = 0.4%).

We identified 67 clusters overall. The median age of people in a cluster was 31 years (range, 18-76). Of people <29 years of age, 29.1% (80 of 275) were in a cluster, but only 11.0% (97 of 879) of those aged  $\geq 30$  were in a cluster (Table 1). A higher percentage of men (152 of 880, 17.3%) than women (25 of 274, 9.1%) were in a cluster. By race/ethnicity, clustering was highest among non-Hispanic black people (122 of 764, 16.0%), followed by Hispanic people (41 of 266, 15.4%) and non-Hispanic white people (13 of 106, 12.3%). By transmission category, the percentage of those in a cluster was highest among MSM (115 of 497, 23.1%), followed by those in the “other” category (36 of



**Figure 2.** Inferred human immunodeficiency virus (HIV) transmission network among people in clusters (15.3% of total people with sequences) in Chicago, Illinois, 2005-2011: A, race/ethnicity; B, self-reported transmission; C, region of residence at HIV diagnosis. Edge connection between nodes represents an inferred potential transmission between people, as assessed through phylogenetic analyses of the *pol* region with a maximum genetic distance of 1.5% ( $n = 1154$ ). Figures created by NodeXL, version 1.0.1.340 (Social Media Research Foundation).<sup>21</sup> Abbreviations: MSM, men who have sex with men; PWID, person who injects drugs.

341, 10.6%). Of those receiving an HIV diagnosis during 2006-2010, 16.6% (156 of 939) were in a cluster.

The percentage of people in clusters differed significantly by age group ( $P < .001$ ), sex ( $P < .001$ ), transmission

category ( $P < .001$ ), and year of diagnosis ( $P = .003$ ). Of 67 clusters, 48 (71.6%) contained only men; 61 (91.0%) contained at least 1 non-Hispanic black person (Figure 2A); 61 (91.0%) contained at least 1 person in the MSM transmission category; and only 2 (3.0%) contained at least 1 person in the PWID category (Figure 2B). Additionally, in 39 (58.2%) clusters, all people were in the same transmission category. Those in the “other” transmission category clustered mainly with those in the MSM category, indicating that these people might have been primarily MSM. Forty (59.7%) clusters contained people from only 1 racial/ethnic category.

In adjusted Poisson regression analyses (Table 2), differences by age were significant for those aged 30-39 (rate ratio [RR] = 0.33; 95% confidence interval [CI], 0.25-0.42), 40-49 (RR = 0.22; 95% CI, 0.16-0.31), and  $\geq 50$  (RR = 0.07; 95% CI, 0.04-0.12), all of whom had a significantly lower degree of connectivity than those  $< 24$  years. We also determined that, as age increased, the degree of connectivity decreased as compared with younger people. Compared with men, women had a significantly lower degree of connectivity (RR = 0.39; 95% CI, 0.26-0.59). Differences by race/ethnicity were significant, with non-Hispanic white people (RR = 0.59; 95% CI, 0.41-0.83) and Hispanic people (RR = 0.52; 95% CI, 0.41-0.66) having a significantly lower degree of connectivity than non-Hispanic black people. Among transmission categories, PWID (RR = 0.17; 95% CI, 0.05-0.52) and MSM and PWID (RR = 0.41; 95% CI, 0.27-0.63) had a significantly lower degree of connectivity than MSM.

We also identified a large number of people in clusters, most of whom were non-Hispanic black and male, who were missing data on risk behavior or sexual identity. The sensitivity analysis in which we recategorized those in the “other” category as MSM did not change the results from our multivariable model.

Clustering was consistent by region of residence, with 16.4% (44 of 268) of north-side residents in clusters, 16.7% (67 of 401) of south side, and 16.5% (47 of 285) of west side. We also found a high degree of heterogeneity among people in clusters, with 48 of 67 clusters (71.6%) including people from  $> 1$  Chicago region (Figure 2C). In clusters of at least 3 people, 10 of 15 clusters included people from  $> 2$  Chicago regions. Among clusters that were homogeneous for region of residence, none contained  $> 2$  people.

## Discussion

From a large sample of people with newly diagnosed HIV infection residing in Chicago, we determined that people who were younger, male, non-Hispanic black, and MSM were more highly connected than other people. Moreover, as age increased, degree of connectivity decreased. Supplementing another recent study conducted in Chicago,<sup>24</sup> we determined that these characteristics were also substantial indicators of connectivity with other people and not merely indicators of being in a cluster. Additionally, our findings revealed that transmission clusters were mainly homogeneous

**Table 2.** Unadjusted and adjusted Poisson regression models of association of selected characteristics with degree of connectivity in a sample of HIV-infected participants (n = 1154) in the Variant, Atypical, and Resistant HIV Surveillance system, Chicago, Illinois, 2005-2011

Variable	Unadjusted Model <sup>a</sup>		Adjusted Model	
	RR (95% CI)	P Value	RR (95% CI)	P Value
Age, y				
≤24	1 [Reference]		1 [Reference]	
25-29	1.24 (1.00-1.53)	.050	1.18 (0.95-1.46)	.136
30-39	0.27 (0.21-0.35)	<.001	0.33 (0.25-0.42)	<.001
40-49	0.14 (0.10-0.20)	<.001	0.22 (0.16-0.31)	<.001
≥50	0.04 (0.03-0.07)	<.001	0.07 (0.04-0.12)	<.001
Sex				
Male	1 [Reference]		1 [Reference]	
Female	0.18 (0.13-0.25)	<.001	0.39 (0.26-0.59)	<.001
Race/ethnicity				
Non-Hispanic black	1 [Reference]		1 [Reference]	
Non-Hispanic white	0.49 (0.35-0.69)	<.001	0.59 (0.41-0.83)	.003
Hispanic	0.45 (0.35-0.57)	<.001	0.52 (0.41-0.66)	<.001
Other	0.08 (0.01-0.59)	.013	0.24 (0.03-1.74)	.160
Transmission category				
MSM	1 [Reference]		1 [Reference]	
Heterosexual contact	0.04 (0.01-0.26)	.001	0.16 (0.02-1.18)	.073
PWID	0.06 (0.02-0.18)	<.001	0.17 (0.05-0.52)	.002
MSM and PWID	0.14 (0.09-0.20)	<.001	0.41 (0.27-0.63)	<.001
Other <sup>b</sup>	0.44 (0.36-0.53)	<.001	0.88 (0.72-1.08)	.231
Year of diagnosis				
2000 or before	1 [Reference]		1 [Reference]	
2001-2005	0.56 (0.08-3.99)	.564	0.44 (0.06-3.12)	.409
2006-2010	12.03 (3.00-48.22)	<.001	3.93 (0.97-15.93)	.055
After 2010	14.82 (3.64-60.37)	<.001	4.40 (1.07-18.14)	.040
Region <sup>c</sup>				
North side	1 [Reference]			
South side	1.10 (0.90-1.33)	.325		
West side	1.02 (0.82-1.26)	.854		
Other	0.25 (0.16-0.36)	<.001		

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus; MSM, men who have sex with men; PWID, person who injects drugs; RR, rate ratio.

<sup>a</sup>Each variable is a separate model; constant not provided because of individual modeling of each variable.

<sup>b</sup>"Other" categories include received clotting factor, received transfusion or transplant, perinatal exposure with HIV diagnosed at age ≥ 13, or no identified or reported risk.

<sup>c</sup>Defined by ZIP code of residence at time of HIV diagnosis.

with respect to sex and transmission category and highly heterogeneous with respect to region of residence. These results indicate that HIV transmission does not occur primarily within one's region of residence; rather, people move about the city when engaging in high-risk behaviors.

These analyses show that interventions might be most effective if targeted to young non-Hispanic black MSM throughout Chicago. The average age of people in clusters was lower than those not in clusters, indicating a higher degree of connectivity and perhaps a higher degree of transmission among younger people, similar to what previous analyses identified.<sup>7,24</sup> Furthermore, as people age, connectivity decreases, indicating that age is the main driver of connectivity and, thus, recent transmission.

A previous study compared characteristics of people in phylogenetic transmission clusters with characteristics of people not in such clusters.<sup>24</sup> Compared with past analyses,

our study had data on a higher percentage of diagnosed cases for analysis; we found more people to be in a cluster and approximately 3 times as many clusters. Compared with our previous analysis, in our multivariable analyses we identified similar characteristics (young, non-Hispanic black, and MSM) as those most associated with connectivity.<sup>24</sup> Our results indicate that not only are these characteristics associated with being in a cluster, as previously found,<sup>24</sup> but they are also associated with the degree of connectivity within these clusters. Phylogenetic methods in combination with traditional epidemiologic approaches can inform and identify changes in local outbreaks, including tracking the potential spread of HIV infection to new demographic risk groups. Although our analysis confirmed the spread of HIV infection through previously identified populations, it highlights the utility of genetic sequencing data to identify patterns in local transmission networks.

Our findings suggest a need for broadly targeted HIV prevention efforts throughout Chicago. Similar to the findings of previous studies,<sup>6</sup> our findings showed that transmission events had a high degree of geographic heterogeneity, suggesting mixing across geographic areas. Such mixing could result, for example, from events that draw people from multiple parts of the city or from online sex-seeking behaviors. Themed parties or special events at bars or clubs can draw people from all over the city, including black MSM who reside on the south, north, and west sides of the city. Moreover, online geospatial “hookup” applications facilitate sexual encounters throughout the city and outside one’s neighborhood. Of the clusters that were homogeneous for region, none contained >2 people, indicating that movement throughout the city might be highly prevalent among those at risk for acquiring and transmitting HIV infection. Previous research found that migration among communities is associated with HIV transmission risk<sup>25</sup>; however, given our data, we were unable to assess migration throughout the Chicago region. Although our analysis was not able to assess migration among community areas, the observed geographic heterogeneity might be a result of the increase in the ease of meeting people online and quickly engaging in sexual activity throughout the region. Previous studies reported that a substantial number of MSM engage in online sex seeking and that those who do are more likely to engage in HIV-risk behaviors.<sup>26,27</sup> Future studies should assess whether one’s region of residence and region of sexual encounter are identical and whether the latter is associated with HIV-risk behaviors.

### Limitations

Our study had several limitations. First, we had access to the sequences of only 14.0% of people who were HIV positive during the study period; thus, the inferred transmission network was incomplete. Additionally, we could not determine whether these transmission links were indirect or direct, nor could we determine a direction of transmission.

### Conclusion

Our study points to the usefulness of phylogenetic and transmission network analyses to identify transmission networks; it also points to a need to follow and address the movement of HIV infections through networks of younger people, rather than geographic areas. This study could also guide prevention efforts in other highly diverse metropolitan regions with characteristics similar to those of Chicago.

As demonstrated during a recent outbreak of newly diagnosed cases of HIV infection in Indiana, when combined with traditional surveillance methods, phylogenetic and transmission network analyses can aid public health officials in identifying and tracking changes in a localized outbreak.<sup>28</sup> Separately, these analyses provide only partial information on Chicago’s HIV burden, but when considered together,

they can guide the development of better-targeted prevention strategies. Future work should assess and understand the reasons for the pattern of geographic mixing. Finally, incorporating the use of phylogenetic analyses into surveillance methods can help determine the changing structure of the local transmission network, provide a means of proactively assessing current and future outbreaks of HIV infection, and offer a basis for tailoring local prevention policy.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the National Institutes of Health (grants R21GM113694, R01DA033875, and R01DA039934). VARHS was funded through a cooperative agreement with CDC. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of CDC.

### References

1. Chalmet K, Staelens D, Blot S, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect Dis*. 2010;10(1):262.
2. Hué S, Pillay D, Clewley JP, Pybus OG. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci U S A*. 2005; 102(12):4425-4429.
3. Yerly S, Junier T, Gayet-Ageron A, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. *AIDS*. 2009;23(11):1415-1423.
4. Louwagie J, McCutchan FE, Peeters M, et al. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS*. 1993;7(6): 769-780.
5. Smith DM, May SJ, Tweeten S, et al. A public health model for the molecular surveillance of HIV transmission in San Diego, California. *AIDS*. 2009;23(2):225-232.
6. Oster AM, Pieniazek D, Zhang X, et al. Demographic but not geographic insularity in HIV transmission among young black MSM. *AIDS*. 2011;25(17):2157-2165.
7. Oster AM, Wertheim JO, Hernandez AL, Ocfemia MCB, Saduvala N, Hall HI. Using molecular surveillance data to understand transmission between subpopulations in the United States. *J Acquir Immune Defic Syndr*. 2015;70(44):444-451.
8. Lee S, Rocha LE, Liljeros F, Holme P. Exploiting temporal network structures of human interaction to effectively immunize populations. *PLoS One*. 2012;7(5):e36439.
9. Little SJ, Kosakovsky Pond SL, Anderson CM, et al. Using HIV networks to inform real time prevention interventions. *PLoS One*. 2014;9(6):e98443.
10. Centers for Disease Control and Prevention. Diagnoses of HIV infection in the United States and dependent areas, 2013. <http://>

- www.cdc.gov/hiv/library/reports/surveillance/2013/surveillance\_Report\_vol\_25.html. Accessed December 1, 2015.
11. Chicago Department of Public Health. HIV/STI surveillance report, Chicago. [https://www.cityofchicago.org/content/dam/city/depts/cdph/HIV\\_STI/2014HIVSTISurveillanceReport.pdf](https://www.cityofchicago.org/content/dam/city/depts/cdph/HIV_STI/2014HIVSTISurveillanceReport.pdf). Published 2014. Accessed January 10, 2016.
  12. US Census Bureau. 2005-2009 American Community Survey 5-year estimates. <http://www.census.gov/data/developers/datasets/acs-survey-5-year-data.html>. Published 2011. Accessed September 16, 2016.
  13. City of Chicago. Census data—selected socioeconomic indicators in Chicago, 2008-2012. <https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2>. Accessed January 10, 2016.
  14. Wheeler WH, Ziebell RA, Zabina H, et al. Prevalence of transmitted drug resistance associated mutations and HIV-1 subtypes in new HIV-1 diagnoses, US—2006. *AIDS*. 2010;24(8):1203-1212.
  15. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792-1797.
  16. Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol*. 2010;27(2):221-224.
  17. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4(4):406-425.
  18. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16(2):111-120.
  19. Felsenstein J. PHYLIP—Phylogeny Inference Package (version 3.2). *Cladistics*. 1989;5:164-166.
  20. Volz EM, Koelle K, Bedford T. Viral phylodynamics. *PLoS Comput Biol*. 2013;9(3):e1002947.
  21. Hansen DL, Shneiderman B, Smith MA. *Analyzing Social Media Networks with NodeXL*. Burlington, MA: Elsevier; 2011.
  22. Weisberg S. *Applied Linear Regression*. Hoboken, NJ: John Wiley & Sons; 2005.
  23. R Core Team. *R Version 3.1.1*. Vienna, Austria: R Foundation for Statistical Computing; 2005.
  24. Lubelchek RJ, Hoehnen SC, Hotton AL, Kincaid SL, Barker DE, French AL. Transmission clustering among newly diagnosed HIV patients in Chicago, 2008 to 2011: using phylogenetics to expand knowledge of regional HIV transmission patterns. *J Acquir Immune Defic Syndr*. 2015;68(1):46-54.
  25. Beyrer C, Razak MH, Lisam K, Chen J, Lui W, Yu XF. Overland heroin trafficking routes and HIV-1 spread in south and south-east Asia. *AIDS*. 2000;14(1):75-83.
  26. Liao A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sex Transm Dis*. 2006;33(9):576-584.
  27. Rosser BRS, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL. HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's Internet Sex Study—II (MINTS-II). *AIDS Behav*. 2009;13(3):488-498.
  28. Spiller MW, Shields J, Bradley H, et al. Network analysis of a contact network from an investigation of a community outbreak of HIV infection linked to injection drug use of oxycodone—Indiana, 2015 [abstract]. [http://www.ias2015.org/WebContent/File/IAS\\_2015\\_\\_MED2.pdf](http://www.ias2015.org/WebContent/File/IAS_2015__MED2.pdf). Accessed January 10, 2016.