Published in final edited form as:

J Acquir Immune Defic Syndr. 2015 January 1; 68(1): 46–54. doi:10.1097/QAI.0000000000000404.

Transmission clustering among newly diagnosed HIV patients in Chicago, 2008 to 2011: using phylogenetics to expand knowledge of regional HIV transmission patterns

Ronald J. Lubelchek, MD^{*,1,2,3,4}, Sarah C. Hoehnen, MD^{1,4}, Anna L. Hotton, PhD, MPH^{2,3}, Stacey L. Kincaid, MPH², David E. Barker, MD^{1,2,3,4}, and Audrey L. French, MD^{1,2,3,4}

¹Division of Infectious Diseases, John H. Stroger, Jr. Hospital of Cook County, Chicago, IL, United States of America (USA)

²Ruth M. Rothstein CORE Center, Chicago, IL, USA

³Chicago Developmental Center for AIDS Research (D-CFAR), Chicago, IL, USA

⁴Rush University Medical Center, Department of Medicine, Chicago, IL, USA

Abstract

Introduction—HIV transmission cluster analyses can inform HIV prevention efforts. We describe the first such assessment for transmission clustering among HIV patients in Chicago.

Methods—We performed transmission cluster analyses using HIV *pol* sequences from newly diagnosed patients presenting to Chicago's largest HIV clinic between 2008 and 2011. We compared sequences via progressive pairwise alignment, using neighbor joining to construct an un-rooted phylogenetic tree. We defined clusters as >2 sequences among which each sequence had at least one partner within a genetic distance of 1.5%. We used multivariable regression to examine factors associated with clustering and used geospatial analysis to assess geographic proximity of phylogenetically clustered patients.

Results—We compared sequences from 920 patients; median age 35 years; 75% male; 67% Black, 23% Hispanic; 8% had a Rapid Plasma Reagin (RPR) titer 1:16 concurrent with their HIV diagnosis. We had HIV transmission risk data for 54%; 43% identified as men who have sex with men (MSM). Phylogenetic analysis demonstrated 123 patients (13%) grouped into 26 clusters, the largest having 20 members. In multivariable regression, age < 25, Black race, MSM status, male gender, higher HIV viral load, and RPR 1:16 associated with clustering. We did not observe geographic grouping of genetically clustered patients.

*Corresponding author contact information (reprint requests to same address): Ruth M. Rothstein CORE Center, 2020 W. Harrison, Chicago IL, 60612, rlubelchek@cookcountyhhs.org, Phone: 312-572-4739, FAX: 312-572-4718.

Meetings at which preliminary data has been presented:

1. ID Week, San Francisco, Oct. 2–6, 2013.
2. CROI, Boston, March 3–6, 2014

Conflicts of Interest:

No authors have conflicts of interest to report.

Discussion—Our results demonstrate high rates of HIV transmission clustering, without local geographic foci, among young Black MSM in Chicago. Applied prospectively, phylogenetic analyses could guide prevention efforts and help break the cycle of transmission.

Keywords

HIV epidemiology; HIV risk; phylogenetics

Introduction

While incident HIV infections in the United States (US) have remained relatively stable since the mid-1990s, disproportionate increases in HIV incidence in key demographics necessitate development of innovative HIV prevention efforts. Such efforts can help meet key goals laid out by the National HIV/AIDS Strategy to decrease HIV incidence, improve access and quality of care for people living with HIV/AIDS (PLWHA) and reduce HIV-related health disparities¹.

Diverse geographic regions have reported marked increases in new HIV diagnoses in young, black men who have sex with men (MSM)^{2–4}. While higher rates of new HIV diagnoses could reflect intensified testing efforts, incidence data confirm an expansion of the HIV epidemic in the Black young men who have sex with men (YMSM) population. Nationally, HIV incidence increased 48% among Black YMSM between 2006 and 2009⁵. Importantly, Torrone et al. demonstrated that, rather than a few, large metropolitan statistical areas driving this increase, 85% of statistical areas with population > 500,000 demonstrated an average increase of 69% in HIV incidence in this population⁶.

Just as the Black MSM population has endured disproportionate increases in HIV incidence, this group also has worse outcomes at each point along the HIV treatment continuum. Surveillance data reveal that significantly more Black MSM with HIV remain unaware of their serostatus, 26% vs. 20% for the overall population of PLWHA⁷. Even more concerning, 59% of PLWHA between the ages of 13–24 remain unaware⁷. At the other end of the treatment continuum, only 21% of Black PLWHA have achieved virologic suppression vs. 30% of White PLWHA, while for Black vs. White MSM an estimated 28% vs. 47% have achieved virologic suppression^{8,9}.

Phylogenetics represents the study of the evolutionary relationships among organisms delineated via comparison of molecular sequences. Compared to more traditional epidemiologic methods, such as contact tracing, HIV phylogenetics may provide data to more effectively assess the local dynamics of the HIV epidemic. Despite the promise of contract tracing to identify previously undiagnosed HIV, the unknown duration of infection at the time of HIV diagnosis and the low rate of transmission per coital act, reduce the yield of traditional contact tracing for HIV¹⁰. In addition, while contact tracing may allow epidemiologic mapping of sexual networks, molecular analyses often refute epidemiologic linkages¹¹.

By comparing relatedness of HIV *pol* genes obtained from individual patients, investigators have identified HIV transmission clusters, i.e. groups of HIV-infected individuals between

whom HIV has been transmitted^{12–20}. HIV transmission cluster analyses may provide a means for identifying segments of the population at highest risk for incident HIV infection. These data, then, may be used to more efficiently deploy HIV prevention resources. Researchers in Quebec used molecular phylogenetics to compare the rate of clustering for patients identified as having acute HIV infection (AHI) vs. patients with chronic infection. While samples from the AHI patients represented fewer than 10% of their cohort, they accounted for half of the onward transmission events²¹. Similarly, in British Columbia, Canada researchers determined transmission clustering for those newly diagnosed between 2002 and 2005. They found a third of the newly diagnosed patients composed transmission clusters, which were enriched with aboriginal and injection drug using patients. Importantly, they found that the clusters that expanded most between '02 and 05 had been initially seeded by recently infected individuals²².

The Chicago Eligible Metropolitan Area (EMA) represents the fifth most HIV-affected EMA nationwide²³. Chicago PLWHA experience the same HIV-related disparities seen nationally. In Chicago, between 2003 and 2011, the number of new HIV diagnoses declined among White MSM, while the number of new diagnoses increased by 47% among Black MSM, with a nearly 60% increase among Black MSM aged 20–29²⁴. While surveillance data reveal these regional HIV-related disparities, to date, no HIV phylogenetic analysis has been presented for the Chicago region. We report our group's HIV molecular transmission cluster analysis for the Chicago area, demonstrating how this type of analysis may provide vital data to supplement basic HIV surveillance with the potential to inform region-specific HIV prevention efforts.

Methods

We undertook a molecular phylogenetic study of a sample of newly diagnosed HIV patients in Chicago. The Cook County Health and Hospitals System (CCHHS) Institutional Review Board approved this study.

Setting

For newly diagnosed PLWHA receiving care at the Ruth M. Rothstein (RMR) CORE Center in Chicago, we collected specimens for baseline HIV genotype testing as part of the Center for Disease Control and Prevention's (CDC) Variant, Atypical and Resistant HIV Surveillance (VARHS) program between January, 2008 and May, 2011. The RMR CORE Center provides primary care to 5000+ PLWHA, or roughly a quarter of all PLWHA in the Chicago area.

Source of samples

Between January, 2008 and May, 2011, all patients presenting to the RMR CORE Center within 90 days of their HIV diagnosis contributed specimens, which staff aliquoted off confirmatory HIV Western blot sera, for HIV genotyping to direct local care decisions and contribute to the CDC's VARHS program. Our staff sent a single specimen for baseline genotyping per patient, and the department of health also employed strict de-duplication algorithms to prevent multiple sample submissions from the same individual.

As part of VARHS surveillance a reference lab (Stanford University, Palo Alto, CA) carried out bulk HIV *pol* sequencing encompassing a 1200 base pair region of the HIV *pol* gene inclusive of the HIV reverse transcriptase and HIV protease coding regions. Sequences from RMR CORE Center patients have been stored in encrypted files.

Phylogenetic analysis for HIV transmission clusters

First we carried out progressive pairwise alignment of all collected HIV *pol* sequences (Geneious version 6.1.6, Biomatters, Auckland, New Zealand). Next, we constructed an unrooted neighbor-joining phylogenetic tree under the Tamura-Nei evolutionary model. We employed bootstrap re-sampling with 500 replicates to construct a consensus phylogenetic tree. We evaluated for transmission clusters among sequences that clustered around common proximal nodes with 99% bootstrap support. Of clusters containing > 2 sequences that met this criterion, we identified final transmission clusters as those in which each sequence had at least one neighbor within a patristic distance of 1.5% substitutions per site as measured via the length between branch tips on the originally generated phylogenetic tree. We employed a conservative 1.5% cut-off to define clusters to reduce the possibility of including unrelated sequences in clusters, as this cut-off is significantly less than the estimated > 5% genetic distance between sequences from unrelated infections in the United States^{14,20,25}. This cut-off has been used in multiple similar analyses^{16,21,26–30}. We considered clusters as groups of greater than two members since transmission pairs may have entered care as part of partners testing and may less likely represent sources of continuous, ongoing transmission^{17,27}.

Variable collection and classification

We extracted demographic, laboratory, and transmission risk information from the electronic medical record for all newly diagnosed patients with available sequences. We included CD4 count, grouped by typical CD4 strata; log₁₀ HIV viral load, considered as a continuous variable; and rapid plasma reagin (RPR) titer if it had been measured within 90 days of HIV diagnosis. For patients with multiple tests, we used values closest to the date of diagnosis. We based HIV transmission risk information on self-reported risk at the time of diagnosis, and included MSM, MSM/intravenous drug use (IDU), IDU, perinatal infection, transfusion, and heterosexual risk. We also created a composite three category variable for analysis using gender and transmission risk: Male, MSM; Male, not MSM, and Female. Given the small numbers with the combined MSM/IDU risk factor, those with MSM/IDU transmission risk were included in the MSM category for this analysis. We collected age as a continuous variable and analyzed as < 25 vs. ≥ 25 since we were specifically interested in determining whether youth vs. adults were more likely to be included in transmission clusters. We dichotomized race/ethnicity for multivariable analyses as Black vs. non-Black because there were too few Whites, Hispanics, and persons of other race/ethnicity in transmission clusters to obtain valid parameter estimates when analyzed using finer categorization. We classified syphilis infection based on RPR titer confirmed by *Treponema pallidum* particle agglutination assay (TPPA). While some reports have used a RPR titer of 1:8 to define active, early syphilis, these reports often corroborate active syphilis status and stage with additional clinical data, which we did not possess³¹. Considering that HIV patients may more likely be sero-fast following syphilis treatment and given our desire for a

more specific definition for active syphilis in order to reduce chances of over-estimating a possible association between syphilis and HIV transmission clustering, we defined active syphilis as patients with an RPR titer 1:16 and a positive TPPA^{32,33}. In addition, we examined year of diagnosis to assess for trends in phylogenetic clustering over time.

Analysis for associations with transmission clustering

We considered membership in a phylogenetic transmission cluster as the dependent variable for our analysis. We used Pearson Chi-Square, Fisher's exact tests and Student's t-tests to compare clustered vs. non-clustered patients. We then entered variables either with p-values < 0.2 on preliminary analysis, or those variables that we considered conceptually important based on previous literature, into univariable and multivariable regression models to examine for associations with transmission clustering¹⁶. We conducted the multivariable analysis using the full dataset, and again on the sample with available transmission risk data. Among men with known transmission risk, we further stratified the regression analysis to identify differences in associations with clustering among MSM vs. non-MSM. We based multivariable model building on an iterative approach and we assessed model fit using the Hosmer-Lemeshow goodness-of-fit test. We analyzed data using SAS version 9.2 (SAS Institute, Cary, NC).

Geographic analysis

We obtained patient addresses from the electronic medical record and geocoded the addresses in SAS using 2010 Census data. We used Kulldorff's spatial scan statistic (SaTscan version 9.0, Boston, MA) to ascertain whether phylogenetically clustered vs. non-clustered patients grouped geographically within Cook County. We used home address census tract as the primary geographic unit of measurement, and aggregated the number of phylogenetically clustered and non-clustered patients within each census tract. The spatial scan method imposes multiple, circular windows on the map, allowing the center of the windows to move across the map area such that the windows include different sets of neighboring geographic areas. At each position, the radiuses of the windows vary continuously from zero to a maximum radius such that each window never includes more than 50% of the total sample population. For each window we used a Bernoulli model for binary data to test the alternative hypothesis that there were larger numbers of observed phylogenetically clustered cases within vs. outside the windows. We obtained p-values for the test of the null hypothesis via Monte Carlo simulation using 9,999 replicates³⁴⁻³⁶. We developed maps using ArcGIS (version 10.1, Esri, Redlands, CA).

Results

We included 920 patients' HIV *pol* sequences for analysis; 77% were 25 years old (median age 34.5 years); 75% were male; 67% were Black, 23% Hispanic, and 9% White. Eight percent of the sample had an RPR titer 1:16 at the time of HIV diagnosis (Table 1).

Phylogenetic analysis demonstrated that 123 of 920 patients (13%) grouped into 26 clusters. Clusters were comprised of three to twenty members: twelve clusters contained three

members, four clusters contained four, five, and six members, one contained seven members, and one large cluster contained twenty members (Figure 1).

Factors associated with membership in phylogenetic cluster: full dataset analysis

In the initial exploratory analysis clustered vs. non-clustered patients were younger (median age 25 vs. 37 years) and were more likely to be Black (88% vs. 61%), and male (92% vs. 73%); p -value < 0.001 for all comparisons (Table 1).

We had access to baseline laboratory data for a majority of the cohort, with total number for each category denoted in Table 1. Clustered vs. non-clustered patients more frequently had RPR titers 1:16 and CD4 counts > 350 cells/ ml^3 . Conversely, we found that significantly fewer clustered patients had HCV co-infection (Table 1).

In multivariable analysis, age < 25 , Black race, male gender, RPR 1:16, CD4 counts > 350 cells/ ml^3 , and higher $\log_{10}$ HIV RNA (adjusted odd's ratio of 1.57 per unit $\log_{10}$ HIV RNA increase) associated with cluster membership (Table 2).

Factors associated with clustering for patients with known HIV transmission risk

The clinical record revealed HIV transmission risk factor data for 495 of 920 (54%), with no differences in risk factor data availability for clustered vs. non-clustered patients. Among men with available risk factor data, MSM or MSM/IDU transmission risk was more common among clustered vs. non-clustered patients (Table 1).

Our additional multivariable analysis, inclusive of the subgroup for which we had HIV transmission risk data, revealed an association between clustering and age < 25 , Black race, male gender, MSM status, RPR 1:16, and higher $\log_{10}$ HIV viral loads (Table 3). For men with known HIV transmission risk, when we stratified the multivariable analysis by MSM status, age < 25 (odd's ratio [OR] 2.92, 95% CI 1.26 – 6.77, $p = 0.013$) and Black race (OR 5.74, 95% CI 2.02 – 16.30, $p = 0.001$) associated with clustering for MSM. For non-MSM, RPR 1:16 (OR 13.90, 95% CI 2.01 – 52.40, $p = 0.008$) and higher $\log_{10}$ HIV RNA (OR 2.81, 95% CI 1.04 – 11.20, $p = 0.043$) significantly associated with cluster membership.

Geographic analysis of phylogenetically clustered patients

We had addresses on 830/920 (90%) of patients at the time of diagnosis; 782/830 (94%) resided in Cook County and were included in our geographic analysis. We did not find statistically significant geographic groupings of genetically clustered vs. non-clustered patients; the ratio of observed to expected cases was 3.65 for the most likely geographic grouping of phylogenetically clustered patients ($p = 0.232$) (Figure 2a). Geographic analysis limited to clusters containing greater than five members revealed similar results; the observed number of patients in the most likely geographic groupings was not greater than what would be expected due to chance (geographic cluster rate ratio 2.45, $p = 0.199$) (Figure 2b).

Discussion

We compared HIV *pol* gene sequences from the 920 newly diagnosed patients presenting for care at the RMR CORE Center in Chicago between Jan, 2008 and May, 2011. We found that 13% of those patients grouped into 26 molecularly defined HIV transmission clusters. On multivariable analysis age < 25, Black race, male gender, MSM risk, higher HIV viral loads and having an RPR titer 1:16 correlated with clustering. For MSM, Black race and age < 25 correlated with cluster membership, while for non-MSM men RPR titer 1:16 and higher HIV viral loads correlated with transmission clustering. Our geographic information systems analysis revealed that members of specific clusters did not group geographically.

This report represents the first description of molecularly defined HIV transmission clustering in Chicago, the US's fifth most HIV-affected metropolitan area²³. Our findings, that young, Black MSM more likely compose active HIV transmission clusters, corresponds to HIV incidence surveillance that shows the highest incidence for this subset of patients³⁻⁵. Importantly, this report also corroborates others findings that HIV and syphilis transmission correlate³⁷. This relationship may reflect the effect that primary syphilis has on increasing susceptibility to acquiring HIV, as well as the parallel risk behaviors that promote syphilis and HIV acquisition, and it underscores the importance of focusing prevention resources toward both these co-epidemics³⁸.

In our subgroup analysis examining risk factors for HIV transmission clustering among men, the fact that syphilis significantly correlated with clustering only for non-MSM vs. MSM may have resulted from high rates of syphilis among both clustered and non-clustered MSM. The high published rate of syphilis in Chicago's MSM population supports this interpretation³⁹. Given the association between MSM status and syphilis in Chicago, the correlation we found between active syphilis and transmission clustering among non-MSM men may be from increased HIV transmission risk directly related to syphilis co-infection vs. syphilis acting as a marker for non-disclosed MSM status³⁹.

We also would like to highlight the significance of not finding geographic foci for HIV transmission clusters within the Chicago region. This finding echoes the results of researchers in Mississippi, who carried out a similar analysis. Oster et al, found more young, Black, MSM patients in transmission clusters and they noted that despite the clustered patients' demographic homogeneity, they distributed evenly throughout Mississippi's geographic regions³⁰. Researchers have demonstrated increasing use of internet resources for identification of sex partners among MSM which may have contributed to the lack of geographically grouped clusters in our study⁴⁰⁻⁴³. This outcome seems to recommend against geography-bounded HIV prevention efforts, at least on the local/regional level.

Our conservative definition of transmission clusters likely selected clustered patients with early HIV infection, some of whom may have had acute HIV infection, and may account for the correlation between higher baseline HIV viral loads, higher CD4 counts and transmission cluster membership. While we found higher CD4 counts among clustered vs. non-clustered patients in our initial, exploratory analysis (Table 1) and in the multivariable analysis inclusive of the entire dataset (Table 2), our subset regression analysis inclusive of

only those with transmission risk data did not support this association (Table 3). For this subgroup analysis, which had a smaller sample size, other factors more strongly associated with clustering and outweighed the association with CD4 count. Similarly, HCV status fell out on the regression analysis, which may have been due to co-variation between lower HCV infection rates and MSM status.

Our findings have several important limitations. Notably, the sample included in our analysis only represents a portion of the newly diagnosed HIV patients residing in the Chicago region between 2008 and 2011. While this may limit the generalizability of our findings, we believe the large size and diverse nature of the patient population we included allows our findings to retain relevance. According to Illinois Department of Public Health surveillance reports, between Jan, 2008 and May, 2011 2782 persons were newly diagnosed with HIV in the Chicago EMA⁴⁴. Therefore, our sample of 920 patients represented a third of all newly diagnosed patients for the region during the time period in question. The demographics of the RMR CORE Center patients mirror the population of PLWHA diagnosed in Illinois during the study period, other than the CORE Center sample having more racial and ethnic minorities⁴⁴. While sampling only from the RMR CORE Center catchment area may bias our analysis, we believe this small bias toward including more racial/ethnic minorities may be justified given this segment of the population suffers greater HIV-related disparities.

That we only had HIV transmission risk factor data for 54% of our sample represents another limitation. Incomplete risk factor data led to smaller samples sizes in sub-analyses and more complete risk factor data may have altered our findings. In spite of this limitation, our finding that clustered vs. non-clustered patients were more likely to be younger, Black and MSM corresponds with other reports that highlight the active nature of the HIV epidemic in these sub-populations^{4,5,45–47}.

In addition to having incomplete risk factor data, we also did not have socioeconomic data on our patients and we acknowledge that associations between race/ethnicity and clustering may have been driven more by socio-economic, than race/ethnicity-related, forces. A meta-analysis examining disparities in HIV risk among Black vs. non-Black MSM noted that while Black MSM practice preventive measures more often than MSM in general, they also suffer more from structural barriers such as lower income, lower education levels, higher rates of incarceration and unemployment⁴⁸.

We acknowledge that we used a tight definition for transmission clustering in which inclusion in a cluster required each sequence to have another sequence within a genetic distance of 1.5%. This definition likely selected for patients with early HIV infection and our desire to exclude unrelated infections from clusters may have led chronically infected, but related patients, to be excluded from clusters. We believe the context of using transmission cluster data to direct public health resources towards subpopulations with the highest rates of active HIV transmission justifies the definition for transmission clustering we used and others have employed similar, if not tighter, definitions^{16,20,26–30,49}.

With this study we use phylogenetic analysis of HIV *pol* sequences to elucidate patterns of HIV transmission clustering in the Chicago region. We demonstrate high levels of clustering among young, Black MSM and we confirm the link between syphilis and HIV transmission that likely warrants ongoing research and policy-related attention. Moreover, our data suggest that HIV transmission in the Chicago region is not bound by local geography. We believe the combination prevention strategy of identifying transmission clusters in a timely manner and targeting clusters with focused prevention efforts, such as sexual network based HIV testing and sub-group validated evidence-based behavioral interventions merits further study.

Acknowledgments

Sources of support/funding:

Co-author Anna L. Hotton received partial salary support from the NIH via the Chicago Developmental Center for AIDS Research via P30 AI 082151. This funding provided partial salary support for her role as the Chicago D-CFAR clinical core bio-statistician, in which capacity she worked on the research presented in this manuscript.

This work was supported by National Institute of Health funding to the Chicago Developmental Center for AIDS Research [P30 AI 082151 salary support to ALH]. The authors would like to thank both the Chicago Department of Public Health and the Illinois Department of Public Health for supporting the collection of baseline genotypes for transmitted drug resistance surveillance, which enabled this work.

References

1. Millett GA, Crowley JS, Koh H, et al. A Way Forward: The National HIV/AIDS Strategy and Reducing HIV Incidence in the United States. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2010; 55:S144–S147. doi:10.1097/QAI.1090b1013e3181fbc1004.
2. CDC. HIV Infection Among Young Black Men Who Have Sex with Men - Jackson, Mississippi, 2006–2008. *Morbidity and Mortality Weekly Report*. 2009; 58(4):77–81. [PubMed: 19194368]
3. CDC. Increase in Newly Diagnosed HIV Infections Among Young Black Men Who Have Sex with Men - Milwaukee County, Wisconsin, 1999–2008. *Morbidity and Mortality Weekly Report*. February;2011 60(4):99–102. [PubMed: 21293324]
4. Pathela P, Braunstein SL, Schillinger JA, Shepard C, Sweeney M, Blank S. Men Who Have Sex With Men Have a 140-Fold Higher Risk for Newly Diagnosed HIV and Syphilis Compared With Heterosexual Men in New York City. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2011; 58(4):408–416. doi:10.1097/QAI.1090b1013e318230e318231ca.
5. Prejean J, Song R, Hernandez A, et al. Estimated HIV Incidence in the United States, 2006–2009. *PLoS ONE*. 2011; 6(8):e17502. [PubMed: 21826193]
6. Torrone EA, Bertolli J, Li J, et al. Increased HIV and Primary and Secondary Syphilis Diagnoses Among Young Men - United States, 2004 – 2008. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2011; 58(3):328–335. doi:10.1097/QAI.1090b1013e31822e31075.
7. Chen M, Rhodes PH, Hall HI, Kilmarx PH, Branson BM, Valleroy LA. Prevalence of Undiagnosed HIV Infection Among Persons Aged 13 Years --- National HIV Surveillance System, United States, 2005 – 2008. *Morb Mortal Wkly Rep*. 2012; 61(2):57–64.
8. Hall HI, Holtgrave D, Tang T, Rhodes P. HIV Transmission in the United States: Considerations of Viral Load, Risk Behavior, and Health Disparities. *AIDS and Behavior*. 2013; 17(5):1632–1636. [PubMed: 23456577]
9. Hall HI, Holtgrave DR, Mausbly C. HIV transmission rates from persons living with HIV who are aware and unaware of their infection. *AIDS*. Apr 24; 2012 26(7):893–896. [PubMed: 22313960]
10. Golden MR, Stekler J, Kent JB, Hughes JP, Wood RW. An evaluation of HIV partner counseling and referral services using new disposition codes. *Sexually Transmitted Diseases*. Feb; 2009 36(2): 95–101. [PubMed: 19131906]

11. Resik S, Lemey P, Ping L, et al. Limitations to contact tracing and phylogenetic analysis in establishing HIV type 1 transmission networks in Cuba. *AIDS Research and Human Retroviruses*. 2007; 23(3):347–356. [PubMed: 17411367]
12. Lewis F, Hughes G, Rambaut A, Pozniak A, Leigh Brown A. Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLoS Medicine*. 2008; 5(3):e50. [PubMed: 18351795]
13. Yerly S, Vora S, Rizzardi P, et al. Acute HIV infection: impact on the spread of HIV and transmission of drug resistance. *AIDS*. Nov 23; 2001 15(17):2287–2292. [PubMed: 11698702]
14. Pao D, Fisher M, Hué S, et al. Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. *AIDS*. 2005; 19(1):85. [PubMed: 15627037]
15. Hué S, Pillay D, Clewley J, Pybus O. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(12):4425. [PubMed: 15767575]
16. Aldous JL, Pond SK, Poon A, et al. Characterizing HIV transmission networks across the United States. *Clinical Infectious Diseases*. 2012; 55(8):1135–1143. [PubMed: 22784872]
17. Dennis AM, Hua S, Hurt CB, et al. Phylogenetic insights into regional HIV transmission. *AIDS*. 2012; 26(14):1813–1822. doi:1810.1097/QAD.1810b1013e3283573244. [PubMed: 22739398]
18. Bao L, Vidal N, Fang H, et al. Molecular tracing of sexual HIV Type 1 transmission in the southwest border of China. *AIDS Res Hum Retroviruses*. May; 2008 24(5):733–742. [PubMed: 18462079]
19. Lubelchek RJ, Beavis KG, Gonzalez M, Kendrick SR, Roberts RR, Barker DE. Can we broaden the applicability of HIV transmission cluster analyses? *AIDS*. 2012; 26(8):1043–1044. doi: 10.1097/QAD.1040b1013e3283522d3283581. [PubMed: 22552480]
20. Smith D, May S, Tweeten S, et al. A public health model for the molecular surveillance of HIV transmission in San Diego, California. *AIDS*. 2009; 23(2):225. [PubMed: 19098493]
21. Brenner BG, Roger M, Routy JP, et al. High rates of forward transmission events after acute/early HIV-1 infection. *J Infect Dis*. Apr 1; 2007 195(7):951–959. [PubMed: 17330784]
22. Ragonnet-Cronin M, Ofner-Agostini M, Merks H, et al. Longitudinal Phylogenetic Surveillance Identifies Distinct Patterns of Cluster Dynamics. *J Acquir Immune Defic Syndr*. 2010; 55(1):102–108. [PubMed: 20622676]
23. CDC. Diagnosis of HIV infection in the United States and dependent areas, 2011. HIV Surveillance Report. Mar 7. 2013 23:23. available at: <http://www.cdc.gov/hiv/topics/surveillance/resources/reports/>.
24. Prachand, N. Overview of the HIV epidemic and continuum of care in Chicago. Chicago Area HIV Integrated Services Council Meeting; Chicago. 2013.
25. Brooks JT, Robbins KE, Youngpairoj AS, et al. Molecular analysis of HIV strains from a cluster of worker infections in the adult film industry, Los Angeles 2004. *AIDS*. 2006; 20(6):923–928. [PubMed: 16549978]
26. Brown AE, Gifford RJ, Clewley JP, et al. Phylogenetic reconstruction of transmission events from individuals with acute HIV infection: toward more-rigorous epidemiological definitions. *Journal of Infectious Diseases*. Feb 1; 2009 199(3):427–431. [PubMed: 19133810]
27. 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 Infectious Diseases*. 2010; 10(1):262. [PubMed: 20822507]
28. Hue S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS*. Mar 26; 2004 18(5):719–728. [PubMed: 15075506]
29. Mehta SR, Kosakovsky Pond SL, Young JA, Richman D, Little S, Smith DM. Associations Between Phylogenetic Clustering and HLA Profile Among HIV-Infected Individuals in San Diego, California. *Journal of Infectious Diseases*. May 15; 2012 205(10):1529–1533. [PubMed: 22448013]
30. 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. [PubMed: 21866038]

31. Huhn GD, McIntyre AF, Broad JM, et al. Factors associated with newly diagnosed HIV among persons with concomitant sexually transmitted diseases. *Sexually transmitted diseases*. 2008; 35(8):731–737. [PubMed: 18607308]
32. Dowell ME, Ross PG, Musher DM, Cate TR, Baughn RE. Response of latent syphilis or neurosyphilis to ceftriaxone therapy in persons infected with human immunodeficiency virus. *The American Journal of Medicine*. 93(5):481–488. [PubMed: 1442850]
33. Lynn WA, Lightman S. Syphilis and HIV: a dangerous combination. *The Lancet Infectious Diseases*. 2004; 4(7):456–466. [PubMed: 15219556]
34. Kulldorff M. A spatial scan statistic. *Communications in Statistics--Theory and methods*. 1997; 261:1481–1496.
35. Kulldorff M, Feuer EJ, Miller BA, Freedma LS. Breast cancer clusters in the northeast United States: a geographic analysis. *American Journal of Epidemiology*. 1997; 146(2):161–170. [PubMed: 9230778]
36. Kulldorff M, Nagarwalla N. Spatial disease clusters: Detection and inference. *Statistics in Medicine*. 1995; 14:799–810. [PubMed: 7644860]
37. Torrone EA, Bertolli J, Li J, et al. Increased HIV and Primary and Secondary Syphilis Diagnoses Among Young Men—United States, 2004–2008. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2011; 58(3):328.
38. Zetola NM, Klausner JD. Syphilis and HIV infection: an update. *Clinical Infectious Diseases*. 2007; 44(9):1222–1228. [PubMed: 17407043]
39. Chicago Department of Public Health. HIV/STI Surveillance Report, 2013. Chicago, IL: City of Chicago; Dec. 2013
40. Hightow LB, MacDonald PD, Pilcher CD, et al. The unexpected movement of the HIV epidemic in the Southeastern United States: transmission among college students. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2005; 38(5):531–537.
41. Benotsch EG, Kalichman S, Cage M. Men who have met sex partners via the Internet: Prevalence, predictors, and implications for HIV prevention. *Archives of sexual behavior*. 2002; 31(2):177–183. [PubMed: 11974643]
42. Berg RC. Barebacking among MSM Internet users. *AIDS and Behavior*. 2008; 12(5):822–833. [PubMed: 17676278]
43. Bull S, McFarlane M, Rietmeijer C. HIV and sexually transmitted infection risk behaviors among men seeking sex with men on-line. *American Journal of Public Health*. 2001; 91(6):988. [PubMed: 11392947]
44. Illinois Department of Public Health. [Accessed February 27, 2014, 2014] Monthly HIV/AIDS/STD Surveillance Reports. Illinois AIDS/HIV/STD Surveillance Reports. 2013. http://www.idph.state.il.us/aids/surveillance_reports.htm
45. Increase in newly diagnosed HIV infections among young black men who have sex with men--Milwaukee County, Wisconsin, 1999–2008. *MMWR Morb Mortal Wkly Rep*. Feb 4; 60(4):99–102.
46. Hall H, Frazier EL, Rhodes P, et al. Differences in human immunodeficiency virus care and treatment among subpopulations in the united states. *JAMA Internal Medicine*. 2013; 173(14):1337–1344. [PubMed: 23780395]
47. Hurt CB, Beagle S, Leone PA, et al. Investigating a Sexual Network of Black Men Who Have Sex With Men: Implications for Transmission and Prevention of HIV Infection in the United States. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2012; 61(4):515–521. doi:510.1097/QAI.1090b1013e31827076a31827074.
48. Millett GA, Peterson JL, Flores SA, et al. Comparisons of disparities and risks of HIV infection in black and other men who have sex with men in Canada, UK, and USA: a meta-analysis. *The Lancet*. 2012; 380(9839):341–348.
49. Wertheim JO, Brown AJL, Hepler NL, et al. The global transmission network of HIV-1. *Journal of Infectious Diseases*. 2014; 209(2):304–313. [PubMed: 24151309]

Figure 1. Phylogenetically clustered HIV patients diagnosed 2008 to 2011; RMR CORE Center, Chicago

Phylogenetic tree depicting HIV patients newly diagnosed in Chicago between Jan, 2008 and May, 2011; Of 920 patients, the 123 clustered patients are shown here in 26 transmission clusters – grouped by color and labeled A to Z, with each tip representing an individual patient. Each cluster depicted satisfied the 99% support on bootstrap resampling cut-off. The scale line indicates genetic distance.

Figure 2. Geographic distribution of patients in HIV transmission clusters

Shaded map area represents Cook County, IL; blue outline represents the boundaries of the city of Chicago. Points represent approximate locations of (A) phylogenetically clustered and non-clustered patients, and (B) patients in clusters with five members. Points displayed on the maps have been randomly shifted within their associated census tract boundaries to protect patient anonymity while maintaining approximate geographic distribution of cases.

Table 1

Patient Demographics and Laboratory Studies at Diagnosis

	Total ^a n (%)	In a cluster n (%)	Not in a cluster n (%)	p-value
Total, row %	920	123 (13.4)	797 (86.6)	---
Age				
<25	209 (22.7)	58 (47.2)	151 (19.0)	<0.001
≥25	711 (77.3)	65 (52.9)	646 (81.0)	
Race/ethnicity				
White	78 (8.5)	1 (0.8)	77 (9.7)	<0.001
Black	619 (67.3)	108 (87.8)	511 (64.1)	
Hispanic	209 (22.7)	13 (10.6)	196 (24.6)	
Other	14 (1.5)	1 (0.8)	13 (1.6)	
Gender				
Male	693 (75.3)	111 (90.2)	582 (73.0)	<0.001
Female	227 (24.7)	12 (9.8)	215 (27.0)	
Risk factor data available				
Yes	495 (53.8)	69 (56.1)	426 (53.5)	0.584
No	425 (46.2)	54 (43.9)	371 (46.5)	
HIV transmission risk^b				
<i>Male (n=352)</i>				0.022
MSM	206 (58.5)	46 (74.2)	160 (55.2)	0.999
MSM/IDU	6 (1.7)	2 (3.2)	4 (1.4)	
IDU	27 (7.7)	3 (4.8)	24 (8.3)	
Heterosexual	112 (31.8)	11 (17.7)	101 (34.8)	
Perinatal	1 (0.3)	0 (0.0)	1 (0.3)	
Transfusion	0 (0.0)	0 (0.0)	0 (0.0)	
<i>Female (n=143)</i>				0.999
IDU	5 (3.5)	0 (0.0)	5 (3.7)	
Heterosexual	137 (95.8)	7 (100.0)	130 (95.6)	
Perinatal	0 (0.0)	0 (0.0)	0 (0.0)	
Transfusion	1 (0.7)	0 (0.0)	1 (0.7)	
RPR (n=839)				
Non-reactive	696 (83.0)	84 (73.0)	612 (84.5)	<0.001
Reactive, < 1:16	67 (8.0)	9 (7.8)	58 (8.0)	
Reactive, ≥ 1:16	76 (9.1)	22 (19.1)	64 (7.5)	
CD4 count, cells/mm³ (n=815)				
< 200	329 (40.4)	31 (27.9)	298 (42.3)	0.018
200–349	219 (26.9)	31 (27.9)	188 (26.7)	
350–499	146 (17.9)	27 (24.3)	119 (16.9)	
≥ 500	121 (14.9)	22 (19.8)	99 (14.1)	
Log₁₀ HIV RNA, copies/mL (n=739)	4.18 (0.80)	4.26 (0.73)	4.17 (0.81)	0.266
HCV status (n=810)				

	Total ^a n (%)	In a cluster n (%)	Not in a cluster n (%)	p-value
Positive	72 (8.9)	2 (1.8)	70 (10.0)	0.005
Negative	738 (91.1)	108 (98.2)	630 (90.0)	
Year of diagnosis ^c				
2008	263 (28.6)	33 (26.8)	230 (28.9)	0.640
2009	279 (30.3)	45 (36.6)	234 (29.4)	
2010	278 (30.2)	33 (26.8)	245 (30.7)	
2011	100 (10.9)	12 (9.8)	88 (11.0)	

Abbreviations: IQR, interquartile range; IDU, injection drug use; HCV, Hepatitis C virus; RPR, rapid plasma reagin; MSM, men who have sex with men

^aTotals may not sum to 920 due to missing data.

^b Among those with available data on transmission risk. Information on transmission risk was unknown or unavailable for 341/693 (49%) males and 84/227 (37%) females.

^c p-value by Cochran-Armitage trend test

Table 2

Factors Associated with Cluster Membership: Full Sample Analysis (n=920)

	Univariable OR (95% CI)	p-value	Multivariable OR ^a (95% CI)	p-value
Age				
<25	3.82 (2.57–5.67)	<0.001	2.50 (1.53–4.09)	<0.001
25	1.0 (Ref)	---	1.0 (Ref)	---
Race				
Black	4.03 (2.30–7.05)	<0.001	5.01 (2.54–9.91)	<0.001
Non-Black	1.0 (Ref)	---	1.0 (Ref)	---
Gender				
Male	3.42 (1.85–6.33)	<0.001	2.92 (1.43–5.97)	0.003
Female	1.0 (Ref)	---	1.0 (Ref)	---
RPR				
Reactive, 1:16	2.97 (1.72–5.12)	<0.001	2.31 (1.18–4.51)	0.014
Reactive, < 1:16	1.13 (0.54–2.37)	0.745	1.14 (0.48–2.71)	0.772
Non-reactive	1.0 (Ref)	---	1.0 (Ref)	---
CD4 cell count, cells/mm³				
< 200	1.0 (Ref)	---	1.0 (Ref)	---
200–349	1.59 (0.93–2.69)	0.089	1.38 (0.73–2.60)	0.324
350–499	2.18 (1.25–3.81)	0.006	2.59 (1.27–5.27)	0.009
500	2.14 (1.18–3.86)	0.012	2.84 (1.35–5.98)	0.006
Log₁₀ HIV RNA	1.17 (0.89–1.53)	0.266	1.57 (1.12–2.22)	0.010
HCV Infection				
Positive	0.17 (0.04–0.69)	0.013	0.30 (0.07–1.35)	0.116
Negative	1.0 (Ref)	---	---	---

Abbreviations: OR, odds ratio; CI, confidence interval; HCV, Hepatitis C virus; RPR,

^a Multivariable OR is adjusted for all variables for which estimates are presented.

Table 3

Factors Associated with Cluster Membership among Patients with Available transmission risk data (n=495)

	Univariable OR (95% CI)	p-value	Multivariable OR ^a (95% CI)	p-value
Age				
<25	4.35 (2.57–7.37)	<0.001	2.87 (1.44–5.71)	0.003
25	1.0 (Ref)	---	1.0 (Ref)	---
Race				
Black	4.09 (2.09–8.01)	<0.001	4.04 (1.88–8.70)	<0.001
Non-Black	1.0 (Ref)	---	1.0 (Ref)	---
Gender/risk category				
Male, MSM	5.69 (2.49–13.0)	<0.001	3.18 (1.25–8.10)	0.015
Male, non-MSM	2.16 (0.84–5.52)	0.108	2.85 (1.02–7.97)	0.046
Female	1.0 (Ref)	---	1.0 (Ref)	---
RPR				
Reactive, 1:16	3.17 (1.58–6.37)	0.001	2.52 (1.08–5.87)	0.033
Reactive, < 1:16	1.00 (0.34–2.98)	0.996	1.10 (0.31–3.92)	0.871
Non-reactive	1.0 (Ref)	---	1.0 (Ref)	---
CD4 cell count, cells/mm³				
< 200	1.0 (Ref)	---	1.0 (Ref)	---
200–349	1.29 (0.65–2.54)	0.470	0.98 (0.44–2.19)	0.951
350–499	1.80 (0.88–3.68)	0.106	2.05 (0.85–4.91)	0.109
500	1.63 (0.75–3.53)	0.219	2.03 (0.77–5.40)	0.155
Log₁₀ HIV RNA	1.47 (1.03–2.11)	0.035	1.63 (1.09–2.46)	0.019
HCV Infection				
Positive	0.18 (0.02–1.32)	0.092	0.44 (0.05–3.52)	0.436
Negative	1.0 (Ref)	---	1.0 (Ref)	---

Abbreviations: OR, odds ratio; CI, confidence interval; HCV, Hepatitis C virus; RPR, rapid plasma reagin; MSM, men who have sex with men

^aMultivariable OR is adjusted for all variables for which estimates are presented.