

## RESEARCH ARTICLE

# Ending the HIV epidemic using National HIV Behavioral Surveillance (NHBS): Recommendations based on DC model

Suparna Das<sup>1</sup>\*, Richard Medina<sup>2</sup>\*, Emily Nicolosi<sup>2</sup>, Anya Agopian<sup>3</sup>, Irene Kuo<sup>3</sup>, Jenevieve Opoku<sup>1</sup>, Adam Allston<sup>1</sup>, Michael Kharfen<sup>1</sup>

**1** Strategic Information Division, HIV/AIDS, Hepatitis, STD, and TB Administration (HAHSTA), District of Columbia Department of Health, Government of the District of Columbia, Washington, DC, United States of America, **2** Department of Geography, University of Utah, Salt Lake City, Utah, United States of America, **3** Milken Institute of Public Health, Department of Epidemiology, George Washington University, Washington, DC, United States of America

\* These authors contributed equally to this work.

\* [Suparna.das@dc.gov](mailto:Suparna.das@dc.gov)



## OPEN ACCESS

**Citation:** Das S, Medina R, Nicolosi E, Agopian A, Kuo I, Opoku J, et al. (2021) Ending the HIV epidemic using National HIV Behavioral Surveillance (NHBS): Recommendations based on DC model. PLoS ONE 16(7): e0253594. <https://doi.org/10.1371/journal.pone.0253594>

**Editor:** Yury E. Khudyakov, Centers for Disease Control and Prevention, UNITED STATES

**Received:** January 8, 2021

**Accepted:** June 8, 2021

**Published:** July 22, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0253594>

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** Data cannot be shared publicly because of PHI and PII DC and federal/GW Institutional Data Access / Ethics Committee. There are legal and ethical restrictions

## Abstract

### Introduction

Social network strategies have been used by health departments to identify undiagnosed cases of HIV. Heterosexual cycle (HET4) of National HIV Behavioral Surveillance (NHBS) is a social network strategy implemented in jurisdictions. The main objectives of this research are to 1) evaluate the utility of the NHBS HET cycle data for network analysis; 2) to apply statistical analysis in support of previous HIV research, as well as to develop new research results focused on demographic variables and prevention/intervention with respect to heterosexual HIV risk; and 3) to employ NHBS data to inform policy with respect to the EHE plan.

### Method

We used data from the 2016 NHBS HET4 (DC). A total of 747 surveys were collected. We used the free social-network analysis package, GEPHI, for all network visualization using adjacency matrix representation. We additionally conducted logistic regression analysis to examine the association of selected variables with HIV status in three models representing 1) demographic and economic effects, 2) behavioral effects, and 3) prevention-intervention effects.

### Results

The results showed 3% were tested positive. Seed 1 initiated the largest networks with 426 nodes (15 positives); seed 4 with 273 nodes (6 positives). Seed 3 had 35 nodes (2 positives). All 23 HIV diagnoses were recruited from 4 zip-codes across DC. The risk of testing positive was higher among people high-school dropouts (Relative Risk (RR) (25.645); 95 CI % 5.699, 115.987), unemployed ((4.267); 1.295, 14.064), returning citizens ((14.319); 4.593, 44.645). We also found in the final model higher association of pre-exposure

to the data being made available publicly. The data can be made available upon request (contact via [manyadm@gwu.edu](mailto:manyadm@gwu.edu) or [acastel@gwu.edu](mailto:acastel@gwu.edu)).

**Funding:** The study was supported by the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

prophylaxis (PrEP) awareness among those tested negative ((4.783); 1.042, 21.944) and HIV intervention in the past 12 months with those tested positive ((17.887); 2.350, 136.135).

## Conclusion

The network visualization was used to address the primary aim of the analysis—evaluate the success of the implementation of the NHBS as a social network strategy to find new diagnoses. NHBS remains one of the strongest behavioral supplements for DC’s HIV planning activities. As part of the evaluation process our analysis helps to understand the impact of demographic, behavioral, and prevention efforts on peoples’ HIV status. We strongly recommend other jurisdictions use network visualizations to evaluate the efficacy in reaching hidden populations.

## Introduction

Identifying persons with undiagnosed Human Immunodeficiency Virus (HIV) infection and linking them to medical care and prevention services is a national priority [1, 2]. Fittingly, some present-day HIV research focuses on behavioral risks and new diagnoses using social network analysis (SNA) for knowledge extraction, as HIV transmits primarily through social interaction, either with direct personal contact (e.g., sexual activity), or through an intermediate activity (e.g., intravenous drug use) [3–5]. SNA has proven to be a valuable approach for HIV research with respect to HIV risk and drug use [6], transmission [7], influences on homeless youth [8], and venue based risk [9], among many others. This study first applies SNA and statistical regression models to identify HIV risk factors and evaluate the effectiveness of the National HIV Behavioral Surveillance (NHBS) strategy in heterosexual HIV diagnoses, and second provides recommendations to improve the implementation of the NHBS program across the United States to guide the Ending the HIV Epidemic—A plan for America (EHE).

We select the Washington, D.C. metro area for this study as it aligns well with national goals and trends. With an HIV prevalence of 1.8%, Washington, D.C. was selected as one of the target hotspots in the EHE. The NHBS program was developed by the Centers for Disease Control and Prevention (CDC), a component of the National HIV/AIDS Strategy for the United States in 2002, to help state and local health departments establish and maintain a surveillance system to monitor selected behaviors and prevention services among groups at the highest risk for HIV infection [10]. The NHBS uses social network data to assist in developing prevention efforts to reduce the spread of HIV [11]. It has played a major role in providing information on HIV related behavioral risks and HIV testing to U.S. jurisdictions. NHBS outcomes improve the understanding of HIV risk and testing behaviors, which are used to implement, as well as evaluate programs for communities [8].

The importance of this research is two-fold. First, there have been numerous studies using the *men who have sex with men* (MSM) and *people who inject drugs* (IDU) cycle data from NHBS surveys; however, research conducted on the *heterosexual* (HET) cycle has been limited to behavioral analyses alone, such as the anal intercourse risks associated with reporting [12], unprotected sex with casual/exchange partners [13], and substance abuse [13]. Heterosexual contact is the second most common route of HIV transmission, estimated to have accounted for approximately 24% of the new infections diagnosed in 2018 [14]. We add to the literature on heterosexual (HET cycle) HIV risk by including models of demographic and

prevention/intervention variables. Previous analyses have revealed the association of individual-level HIV-associated risk factors with increased risk of HIV [15], yet they fail to address the gaps in prevention-intervention, which is critical in planning the EHE. Second, the implications of NHBS as a significant policy tool for the EHE plan has been missing, which we offer here.

This research is the first effort to map NHBS HET data networks and their characteristics, to identify any evidence of HIV positive network clustering, and the first to use HET NHBS data to make recommendations for HIV policy implementation to achieve goals set out in the EHE. The main objectives of this research are to 1) evaluate the utility of the NHBS HET cycle data for network analysis, which is used by the CDC to identify new HIV positives; 2) to apply statistical analysis in support of previous HIV research, as well as to develop new research results focused on demographic variables and prevention/intervention with respect to heterosexual HIV risk; and 3) to employ NHBS data to inform policy with respect to the EHE plan. Our methods include social network visualization and regression analysis.

## Data and methods

### NHBS survey data

The National HIV Behavioral Surveillance (NHBS) program administered by the D.C. Department of Health and George Washington University under CDC directives, includes separate HIV based surveys and HIV testing among high-risk populations, which include men who have sex with men (MSM cycle), people who inject drugs (IDU Cycle), and heterosexuals with increased risk of HIV (HET Cycle).

We use 2016 NHBS HET Cycle data, specifically heterosexual cycle 4 (HET4), which was the fourth round of heterosexual data collection. In general, HET data were collected within 22 metropolitan statistical areas (MSAs) across the United States. The MSAs were selected based on their relatively high number of people living with HIV/AIDS. Survey participants were recruited by respondent-driven sampling (RDS). RDS is initiated with a limited number of “seed” participants who are purposefully chosen through formative research. These individuals were then given 3–5 coupons to recruit social connections into the study. Recruitment continued until the sample size was met, or the end of the data collection period was reached. A total of 747 surveys were collected through the NHBS HET cycle.

Eligibility for the heterosexual cycle of NHBS was restricted to men and women between 18 and 60 years old, who had not previously participated in 2016, were residents of a study MSA, were able to complete the survey in English or Spanish, were able to provide informed consent, and reported having vaginal or anal sex in the past 12 months with an opposite sex partner. Potential participants who reported ever injecting drugs or male-to-male sex in the past 12 months were not eligible to participate. For each survey cycle, an anonymous standardized questionnaire was used to collect information about HIV-associated behaviors, specifically sexual behaviors, substance use, HIV testing, and use of HIV prevention services [16].

For NHBS HET cycles, formative assessments include activities that identify and characterize High Risk Areas (HRAs). HRAs are geographic areas within the MSA where heterosexuals are at higher risk for HIV infection compared to other geographic areas within the MSA. HRAs are defined as areas with high rates of poverty. HRAs are used to identify appropriate locations for storefronts or van locations during survey implementation and to identify seeds. Locating the HRAs requires project sites to obtain geographic data, identify areas of high poverty, and map the HRAs using a Geographic Information System (GIS). Officials within the state health department are involved in the process of identifying HRAs, because it entails handling confidential data. HRA identification is described in greater detail in the NHBS Formative Assessment Manual [10]. Because the focus is on high-risk areas the data do not serve as a

random sample spatially or demographically. A large portion of the data are collected from a relatively smaller sample of locations from a large sample of Black/African Americans. While this can be an issue in interpreting the results of this research, we admit that our results are best applied within high-risk areas and high-risk populations.

Seeds must be residents of High-Risk Areas (HRAs). Since many social ties are formed among individuals who live on the same street or in the same neighborhood, it is important that seeds be residents of HRAs, so that recruitment begins in areas most likely to have a high proportion of low socio-economic status (SES) residents. HRA residency is assessed during the survey. Potential seeds who do not live in an HRA may participate in NHBS HET but will not be eligible to recruit others.

The NHBS funded by CDC and implemented by DC DOH with George Washington University. NHBS IRB for DC DOH was 2014 3 and GWU it is 12331.

### Digital network representation

To represent the connections collected between people in the NHBS HET survey, we used personal social network representations based on index individuals, or “seeds”. The SURID field contains ID numbers for each node in the dataset, including the seeds, and c1-c5 fields that identify social connections between each of the individuals in the dataset. Records are duplicated, such that survey responding nodes found in the c1-c5 fields are all represented in the SURID field. Null values in the dataset representing non-responses are marked -1. Using the SURID (node IDs) and c1-c5 (adjacent node connections and their IDs) fields, network visualizations were constructed for three seed nodes. Given seven initial seeds, only six returned connections, and only three returned meaningful network representations. The social network analysis package, GEPHI, was used for all network visualizations. The NHBS data contain many other fields representing demographic, behavioral, and health data, as described in the Center for Disease Control’s, NHBS HET4 CAPI Reference Questionnaire (CRQ) document [17].

The final compiled network dataset included selected data fields that were appended to the network node data. This allowed for the inclusion of node attributes, such as HIV Positive identifiers (‘EVERPOS’, ‘finrstl’), as shown in the network figures. The ‘EVERPOS’ variable in the data set represents whether or not the participant has ever tested positive for HIV. We tallied finrstl and EVERPOS variables to get the final HIV positive number for the analysis.

For networks initiated by Seeds 1, 3, and 4, a Louvain modularity algorithm for community detection was applied, such that the networks are segmented into “modules” [18]. This algorithm maximizes the within module connectivity and minimizes the between module connectivity.

### Multinomial logistic regression analysis

A multinomial logistic regression was applied to model outcome dependent variables and examine the associations between demographic, behavioral, and prevention-intervention variables and HIV. Multinomial logistic regression analysis is used when the dependent variable is comprised of more than two categories. We calculated a relative risk ratio (RR) for the associations to better interpret the results. We did not use imputation or other substitution methods were not used. In logistic regression, a logistic transformation of the odds (referred to as the logit) serves as the depending variable.

Where:

$p$  = the probability that a case is in a particular category,

$\exp$  = the exponential,

a = the constant of the equation and,

b = the coefficient of the predictor or independent variables

The likelihood ratio test is based on -2LL ratio. Significance at the 0.05 level or lower means the model with the predictors is significantly different from the one with the constant only (all 'b' coefficients being zero). It measures the improvement in fit that the explanatory variables make compared to the null model [19].

The variables were selected, and models were constructed in coordination with D.C.'s Ending the HIV Epidemic plan (EHE)

**Outcome variable.** The outcome variable for the regression model is HIV status, and is coded into the following three categories: 0 "did not test" 1 "tested negative" and 2 "tested positive."

**Predictor variables.** The predictor variables were coded into categories for the regression analysis. All predictor variables were obtained from the NHBS data set. The primary aim of the regression model was to evaluate variables that are being considered or implemented through various programs by DC for Ending the HIV Epidemic by 2030, so the variables for the regression analysis were selected based on priorities of policies that are implemented or in the process of being implemented. The predictor variables were divided into three categories: 1. demographic and economic variables, 2. behavioral variables, and 3. prevention-intervention variables for HIV. The variables were selected to evaluate the impact of policies that are currently implemented or in the process of being implemented in DC. The models measure the changes in the degree of association with the addition of each model's variable set at each of the three model iterations. The first model evaluates the association of HIV outcomes with demographic and economic variables. The second model (behavioral) adds variables including crack cocaine use, mental health conditions, jail term, sexual debut age and casual partners. These are priorities of DC's Ending the HIV Epidemic Plan for 2030. The final model (prevention-intervention) uses variables including technology, pre-exposure prophylaxis awareness and HIV interventions.

Multicollinearity tests were conducted on the models using variance inflation factors (VIFs) and tolerance. If a VIF value exceeded 4.0 or had a tolerance of less than 0.2, the variable was not included in the model as multicollinearity would be indicated [18]. Coefficients with 95% confidence intervals (CIs) were presented in the tables and the relationships were considered statistically significant at  $p\text{-value} < 0.05$ .

## Analysis and results

### Descriptive statistics

Table 1 shows the characteristics of the participants in the study. Out of 747 participants for the NHBS HET 4 cycle, approximately 73% (n = 545) tested negative, 24% (n = 178) had unknown status, which likely means they opted out of the test, and 3% (n = 24) were found HIV positive (Table 1). Out of those positives approximately, 42% (n = 10) were above ages 50 and 55% (n = 12) had education of grade 9 through 11. We combined two variables EVERPOS (Survey question: "Have you ever tested positive for HIV, that is, do you have HIV?") and finrst (Survey question: 'Final HIV result of those who were tested') to identify the HIV positives in the study. EVERPOS showed 15 people who knew they were HIV positive, and the total number of people identified as HIV positives were 24. The number of potential new diagnoses through the cycle were 8. The testing and reporting processes were anonymous, so we cannot confirm if they were new positives (Table 1).

Table 2 shows the characteristics of selected variables for seeds and connections. Approximately 71.43% (n = 5) of the seeds recruited in the sample were male and 28.57% (n = 2) were

Table 1. Descriptive characteristics of NHBS HET 4 cycle DC by HIV status.

Participants characteristics	Not Tested	Tested Negative	Tested Positive	Total
<b>Education</b>				
No information	139(78.09)	0(0.00)	0(0.00)	139(18.61)
Grades 1 through 8	14(7.87)	138(25.32)	5(25.32)	157(21.02)
Grades 9 through 11	18(10.11)	301(55.23)	12(55.23)	331(44.31)
Grade 12 and higher	7(3.93)	106(19.45)	7(19.45)	120(16.06)
Total	178(100)	545(100)	24(100)	747(100)
<b>Age</b>				
Age less than 20 (Ref)	23(12.92)	41(7.52)	1(4.17)	65(8.70)
Ages between 26 to 30	33(18.54)	166(30.46)	2(8.33)	201(26.91)
Ages between 31 and 40	14(7.87)	100(18.35)	4(16.67)	118(15.80)
Ages 41 and 50	27(15.17)	108(19.82)	7(29.17)	142(19.01)
Age more than 50	81(45.51)	130(23.85)	10(41.67)	221(29.59)
Total	178(100)	545(100)	24(100)	747(100)
<b>Gender</b>				
Male	105(58.99)	288(52.84)	12(50.00)	405(54.22)
Female	71(39.89)	257(47.16)	12(50.00)	339(45.38)
Transgender	2(1.12)	0(0.00)	0(0)	2(0.27)
Total	178(100)	545(100)	24(100)	747(100)
<b>Race</b>				
American Indian	2(1.12)	1(0.18)	0(0.0)	3(0.40)
Asian	0(0.00)	1(0.18)	0(0.0)	1(0.13)
Black/African American	168(94.38)	526(96.51)	24(100.00)	719(96.25)
Native Hawaiian	0(0.00)	1(0.18)	0(0.0)	1(0.13)
White	8(4.49)	14(2.57)	0(0.0)	22(2.95)
Total	178(100)	543(100)	24(100)	747(100)

<https://doi.org/10.1371/journal.pone.0253594.t001>

female. Roughly 57.14 (n = 4) were between ages 21 and 30 and 85.71% (n = 6) were Black/African American.

### Descriptive mapping

Fig 1 maps the distribution of 747 study participants and 23 HIV positives in the study based on residence zip codes. Geocoding was conducted using Maptitude, a proprietary GIS software developed by Caliper Corporation. All 23 HIV positives were in 4 zip codes—zip code number 20002 had one HIV case, 20030 had 11 HIV cases, 20019 had 6 HIV cases, and 20032 had 5 HIV cases. Based on the maps we found a higher number of participants were recruited from Wards 7 and 8 (Fig 1).

### Network visualizations

The network visualizations show each network segmented by community (module). Seeds 1, 3, and 4 had the most extensive connected networks, and also larger numbers of participants who tested positive. Seed 1, who was male, initiated a network through which 15 positives were identified, with one branch including 6 positives. A larger branch extending from Seed 1 recruited 3 clusters of positives. There was one transgender person identified in this network (Fig 2).

Seed 4, who was also male, initiated the second largest network (Fig 2). This network included 6 positives total and recruited one transgender person. We did not identify multiple



**Table 2. Descriptive characteristics of HET cycle 4 of seeds and associates.**

Gender	Associate	Seed	Total
Male	400(54.13)	5(71.43)	405(54.22)
Female	338(45.74)	2(28.57)	340(45.52)
Transgender	2(0.27)	0(0)	2(0.27)
Total	739(100)	7(100)	747(100)
<b>Age groups</b>			
Age less than 20	63(8.51)	2(28.57)	65(8.70)
Age between 21 and 30	197(26.62)	4(57.14)	201(26.91)
Age between 31 and 40	117(15.81)	1(14.29)	118(15.80)
Age between 41 and 50	142(19.19)	0(0)	142(19.01)
Age 50 and above	221(29.86)	0(0)	221(29.59)
Total	740(100)	7(100)	747(100)
<b>Race</b>			
American Indian	3(0.41)	0(0)	3(0.40)
Asian	1(0.14)	0(0)	1(0.13)
Black/African America	684(92.81)	6(85.71)	690(92.37)
Native Hawaiian	9(1.22)	1(14.29)	10(1.34)
Refuse to answer	40(5.43)	0(0)	40(5.35)
Unknown	0(0)	6(42.86)	6(0.40)
Total	737(100)	7(100)	747(100)

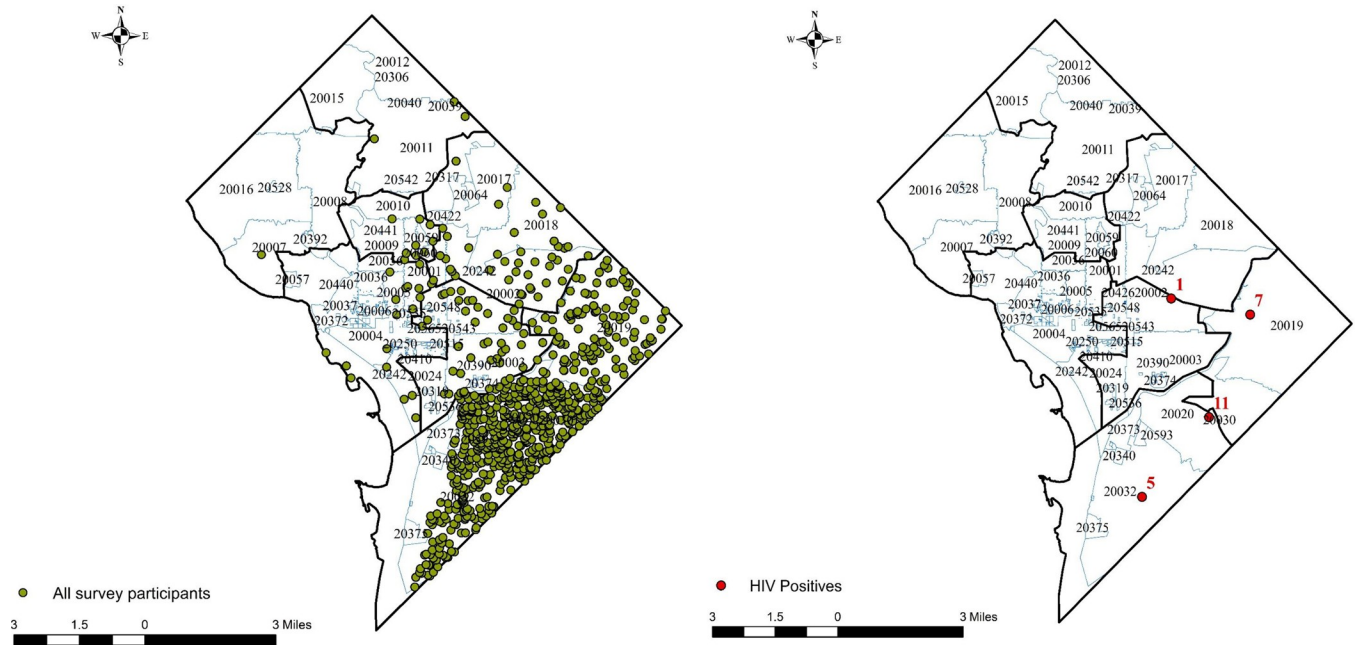
<https://doi.org/10.1371/journal.pone.0253594.t002>

network clusters of HIV positives for Seed 4. Seed 3 (Fig 2) was initiated by a male positive and had two other positives identified in the personal network. The size of the network (35 nodes) is considerably smaller compared to networks for Seeds 1 and 4. Seeds 2, 5, and 7 were substantially smaller. Seed 2 was the largest of the three, recruiting 9 individuals, primarily females. None of the three smaller networks included any HIV positive nodes.

### Characterizing the sample using multinomial logistic regression

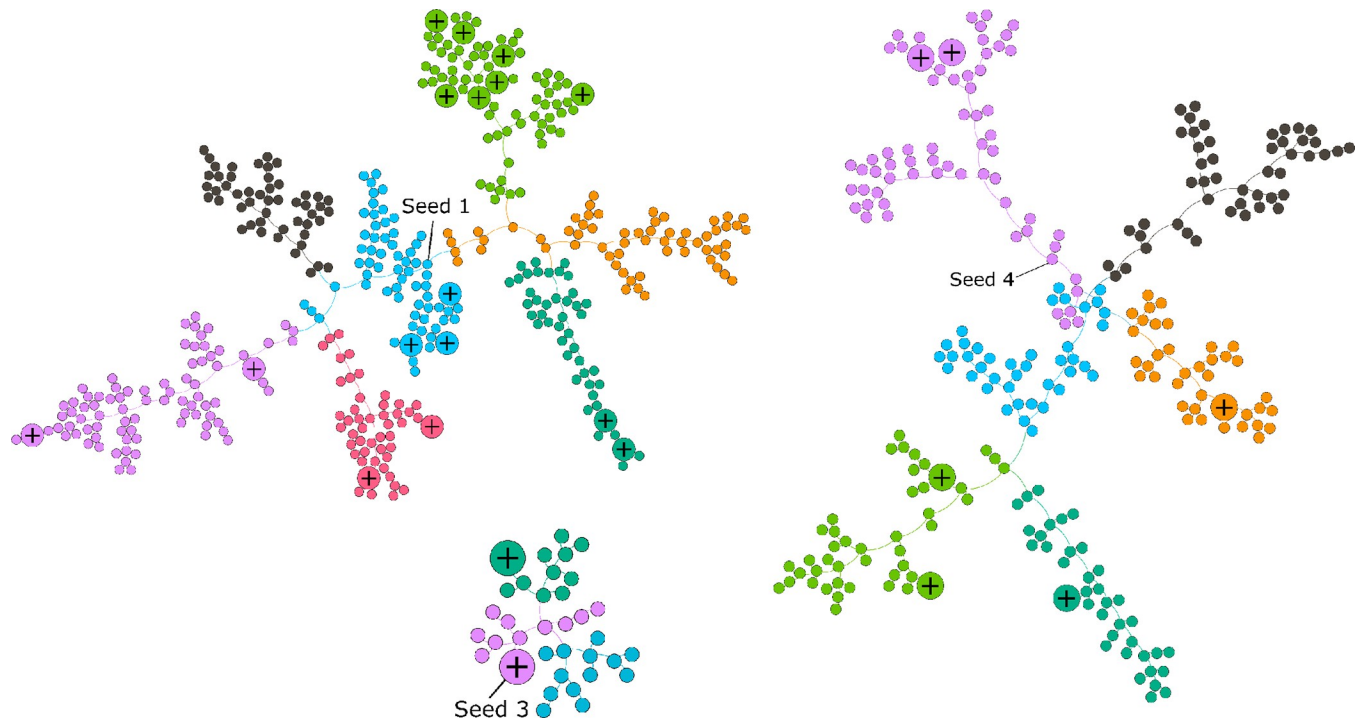
We constructed three iterative models to characterize the data and observe changes as variables were added. Model 1 showed ages 31–40 (RR: 6.902; 95% CI: 2.580, 18.460) were at higher risk of testing negative followed by ages 26–30 (RR: 4.445; 95% CI: 1.884, 10.486). Education with 9 and 11 grades (RR: 18.11; 95% CI: 7.525, 43.588) were also found to be at higher risk of HIV followed by those with education between 1<sup>st</sup> and 8<sup>th</sup> grade (RR: 17.511; 95% CI: 9.906, 30.952) for all who tested negative. Unemployment showed association (RR: 10.607; 95% CI: 5.233, 21.498) with those who tested negative. The likelihood ratio (LR) chi-square of Model 1, -352.29 with a *p*-value of 0.0001, suggests that our model was statistically significant, compared to the null model with no predictors (Table 3). The AIC for Model 1 is 744.587 (Table 3).

In Model 2 we added the behavioral variables to evaluate association with the risk of HIV. The association of risk of HIV with education and age decreased with the addition of the behavioral risks for both who tested positive and negative. Females (RR: 2.445; 95% CI: 1.444, 4.148) were found to be at higher odds of HIV for negatives. With the addition of the behavioral risks in Model 2 we found increased risk of testing positive with lower education—Grades 1 through 9 (RR: 15.913; 95% CI: 4.629, 54.699) and grades 9<sup>th</sup> through 11 (RR: 25.645; 95% CI: 5.699, 115.987). We also found increased risk of testing positive (RR: 4.267; 95% CI: 1.295, 14.064) with unemployment with the addition of behavioral variables. Increased risk of HIV for people who used crack cocaine (RR: 6.799; 95% CI: 1.714, 26.978) compared to those who tested negative and used cocaine (RR: 5.564; 95% CI: 2.040, 15.179). Our model also showed



**Fig 1. Distribution of HET4 cycle participants and HIV positives in DC.** The green dots show the distribution of the survey recruitment. The numbers in red for HIV positives designate the total number of cases in zip code.

<https://doi.org/10.1371/journal.pone.0253594.g001>



**Fig 2. Seeds 1, 3, and 4, the largest networks of HET4 cycle.** The positive signs indicate the HIV positives in the networks.

<https://doi.org/10.1371/journal.pone.0253594.g002>





increased risk among positive tested returning citizens (RR: 14.319; 95% CI: 4.593, 44.645) compared to those who tested negative (RR: 7.132; 95% CI: 3.485, 14.598) (Table 3).

In Model 3 we added prevention-intervention variables, which are currently implemented, or being considered for implementation in DC. While risk of demographic association with testing positive increased for high school drop-out groups (grades 1<sup>st</sup>–8<sup>th</sup> and 9<sup>th</sup>–11<sup>th</sup>). This may indicate the efficacy of prevention-intervention strategies. Pre-exposure prophylaxis awareness showed higher association to those who were at risk but tested negative (RR: 4.783; 95% CI: 1.042, 21.944), while HIV intervention showed association with those who tested positive (RR: 17.887; 95% CI: 2.350, 136.135). The AIC for Model 3 is -273.803 (Table 3).

## Discussion

The network visualization here addresses some of the utility of the NHBS as a survey mechanism, which has the capacity to recruit and test people who are at high risk of HIV using a social network strategy. Generalized testing strategies have become less effective in HIV diagnoses and the NHBS was implemented in jurisdictions to reach the “hidden” population (i.e., those for whom no sampling frame exists or whose members engage in stigmatized or illegal activities, making them reticent to divulge information that may compromise their privacy) [6]. There have been several analyses based on the NHBS data, unfortunately none of them address the success of the network implementation to find HIV diagnoses, which lies at the core of the implementation. Based on our network analysis we strongly recommend other jurisdictions employ network visualizations to evaluate the efficacy in reaching their own hidden populations. However, the structure of the social network data may offer little more than visualization. Still, visualizations often prove to be powerful tools. We recommend implementing CDC’s Social Network strategy to recruit hidden population through community-based organizations. The network visualization shows evidence of some recruitment clustering for HIV positives. We recommend re-evaluating the NHBS recruitment strategy to expand to areas that include more diverse populations in D.C.

This analysis identifies valuable results that can play an important role in HIV prevention-intervention planning and future data collection efforts. The HIV recruitment sample maps identify 24 positives residing in four zip codes, which suggests spatial clusters of HIV and supports previous studies that have identified HIV clusters in D.C. [19, 20]. That said, these recruits were primarily from High-Risk Areas (HRA) identified based on poverty and HIV diagnoses maps provided by the Health Department and were found based on spatially biased data. Three zip codes 20019, 20020, and 20032 had the highest number of recruits in the networks (see Table 1).

As part of the evaluation process our analysis helps to understand the impact of demographic, behavioral, and prevention efforts on peoples’ HIV status. It is reported that Black women aren’t always aware of their higher risk for HIV which stems from lack of public health awareness [20], we recommend higher recruitment and testing of Black women including strategies for PrEP awareness. Black communities traditionally have a high degree of social mixing between higher and lower risk individuals, which means that they are more likely to have a partner with a history of higher risk [20, 21]. Black women are relatively more susceptible to poverty, unstable housing, and unemployment, potentially increasing the likelihood of participation in sex trade, and making them more vulnerable to HIV [22]. We acknowledge that these situations are likely the result of structural racism. Black women also have a higher likelihood of experiencing violence or trauma. This is an important finding for D.C.’s EHE plan where increased engagement of Black women with HIV services and PrEP. We strongly recommend that the plan prioritizes Black women and their access to health care.

Among this HET sample we found that association of education did not change the HIV outcomes with behavior or interventions. The sample analysis also showed that crack cocaine, jail time, and a larger number of partners increased the risk of HIV for HETs in D.C. These behavioral issues may be significant as DC shapes the plan to end the HIV epidemic (EHE). As a part of the opioid use disorder and harm reduction initiative, has been planning and implementing several programs which include the Opioid Learning Institute led by the HealthHIV, opioid awareness campaign and education, and as well as Medication Assisted Treatment (MAT) and Substance Use Disorder Treatment (SUD). It is well established that substance use, including crack cocaine, can create a cycle in which people quickly exhaust their resources and turn to other ways to acquire the drug, including trading sex for drugs or money. This increases HIV risk, yet we lack programs that help mitigate it. We recommend collaboration with the Department of Behavioral Health (DBH) to develop behavioral treatment guidelines for people at risk or with HIV. The guidelines will assist providers with helping patients initiate abstinence and avoid relapse to cocaine use. These guidelines should include contingency management, cognitive behavioral therapy, and motivational interviewing. We also recommend integrating behavioral health screening for people who may be at risk of HIV and being prescribed PrEP.

Our results support that incarceration increases the risk of HIV. Unfortunately, this well-documented outcome often ignores the racial aspect. In the United States, Black and Latino/Hispanic men of lower socio-economic status are disproportionately incarcerated [23]. We recommend programs that may assist incarcerated individuals with prevention and treatment adherence in collaboration with the Metropolitan Police Department Detention Center. D.C. plans to incorporate programs to evaluate and monitor HIV treatment adherence among incarcerated individuals through the returning citizens program.

NHBS is planning on implementing home-based testing post COVID-19. Jurisdictions across the U.S. are considering home-based HIV testing, which will be shipped to those at-risk following completion of an online form. These jurisdictions must be aware of the digital divide that may have a significant impact on HIV outcomes, particularly during the current pandemic. We recommend that for prevention and intervention, we continue to engage with traditional community-based organizations and Disease Intervention Specialists in these areas of lower internet access.

## **NHBS limitations and recommendations for further sampling and research**

First, there is a possibility of sampling bias. This is clear in the spatial distribution of those surveyed. A chain-referral sampling method will be overly sampled within social networks, as defined by the seeds, while larger regional patterns of HIV infection will remain unknown.

Second, the identified positives were held anonymous; thus, they could not be confirmed as new HIV diagnoses. Owing to the CDC recommended pre-condition of selecting criteria of areas, the generalizability becomes difficult to assess. Our results also suggest that the HIV diagnoses may have been previous positives. This is not recommended for optimized social network strategy which has the capacity to identify new positives. We strongly recommend that CDC consider evaluating recruitment guidelines for NHBS.

We worry that current NHBS implementation strategies may not have enough considerations for Hispanic population, making them vulnerable to HIV. We recommend recruitment expansion to areas with higher Latinx population.

Social network analysis in this case is limited, given that traditional calculations would prove to be of little resource. Each node is connected to 1–6 other nodes, where direction through the network, also, does not provide much information. Network data collected in this

way will not have much clustering utility aside from visual, as their hierarchical nature restricts complex community identification. However, there remains an important use, to identify individual social network branches in which higher rates of infections are occurring. Characterizing those branches may prove valuable given that some spatial information exists within the database. Still, there are questions about sampling bias with respect to the non-random choices made to pass on coupons to the next person and its impact on finding the undiagnosed HIV.

Despite these limitations NHBS remains one of the strongest behavioral supplements for D.C.'s HIV planning activities. We recommended planning for a more widely sampled social network study to study sexual-sociospatial networks in D.C. as it can provide more information than solely social or spatial analyses alone. The study and results would be valuable in planning activities for Ending the HIV Epidemic—A plan for America.

## Supporting information

**S1 Table. Characteristics of networks.**  
(DOCX)

## Author Contributions

**Conceptualization:** Suparna Das.

**Data curation:** Suparna Das, Richard Medina, Anya Agopian, Jenevieve Opoku.

**Formal analysis:** Suparna Das, Richard Medina, Emily Nicolosi.

**Funding acquisition:** Irene Kuo, Jenevieve Opoku, Michael Kharfen.

**Investigation:** Suparna Das.

**Methodology:** Suparna Das.

**Project administration:** Suparna Das, Anya Agopian.

**Resources:** Suparna Das, Irene Kuo, Jenevieve Opoku, Adam Allston, Michael Kharfen.

**Software:** Suparna Das, Richard Medina, Emily Nicolosi.

**Supervision:** Suparna Das.

**Validation:** Suparna Das.

**Visualization:** Suparna Das, Richard Medina, Emily Nicolosi.

**Writing – original draft:** Suparna Das, Richard Medina.

**Writing – review & editing:** Suparna Das, Richard Medina, Irene Kuo.

## References

1. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV Epidemic: A Plan for the United States. *JAMA*. 2019; 321(9):844–5. Epub 2019/02/08. <https://doi.org/10.1001/jama.2019.1343> PMID: 30730529.
2. Janssen R, Onorato I, Valdiserr R, Durham T, Nichols W, Seiler E, et al. Advancing HIV prevention: new strategies for a changing epidemic—United States. Atlanta, Georgia: Centers for Disease Control and Prevention, 2003.
3. Rothenberg R. How a Net Works: Implications of Network Structure for the Persistence and Control of Sexually Transmitted Diseases and HIV. *Sexually Transmitted Diseases*. 2000; 28(2).
4. Jolly AM, Muth SQ, Wylie JL, Potterat JJ. Sexual Networks and Sexually Transmitted Infections: A Tale of Two Cities. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*. 2001; 78(3). <https://doi.org/10.1093/jurban/78.3.433> PMID: 11564847

5. Amirkhani YA. Social networks, sexual networks and HIV risk in men who have sex with men. *Curr HIV/AIDS Rep.* 2014; 11(1):81–92. Epub 2014/01/05. <https://doi.org/10.1007/s11904-013-0194-4> PMID: 24384832; PubMed Central PMCID: PMC3944100.
6. Burt RD, Thiede H. Evaluating consistency in repeat surveys of injection drug users recruited by respondent-driven sampling in the Seattle area: results from the NHBS-IDU1 and NHBS-IDU2 surveys. *Ann Epidemiol.* 2012; 22(5):354–63. Epub 2012/03/17. <https://doi.org/10.1016/j.annepidem.2012.02.012> PMID: 22420929; PubMed Central PMCID: PMC3338897.
7. Rothenberg RB, John J, Potterat JJ, Woodhouse DE, Muth SQ, Darrow WW, et al. Social network dynamics and HIV transmission. *AIDS.* 1998; 12. <https://doi.org/10.1097/00002030-199812000-00016> PMID: 9727575
8. Young SD, Rice E. Online social networking technologies, HIV knowledge, and sexual risk and testing behaviors among homeless youth. *AIDS Behav.* 2011; 15(2):253–60. Epub 2010/09/18. <https://doi.org/10.1007/s10461-010-9810-0> PMID: 20848305; PubMed Central PMCID: PMC3029467.
9. Oster AM, Wejnert C, Mena LA, Elmore K, Fisher H, Heffelfinger JD. Network analysis among HIV-infected young black men who have sex with men demonstrates high connectedness around few venues. *Sex Transm Dis.* 2013; 40(3):206–12. Epub 2013/02/14. <https://doi.org/10.1097/OLQ.0b013e3182840373> PMID: 23403601; PubMed Central PMCID: PMC4945956.
10. Prevention CfDCA. National HIV Behavioral Surveillance—Round 4: Injection Drug Use (NHBS-IDU4) and Heterosexuals at Increased Risk for HIV (NHBS-HET4). 2015.
11. Sanchez T, Finlayson T, Drake A, Behel S, Cribbin M, DiNenno E, et al. How a net works: implications of network structure for the persistence and control of sexually transmitted diseases and HIV. Georgia, Atlanta: Centers for Disease Control & Prevention (CDC), 2006.
12. Hess K, Johnson AS, Hu X, Li J, Wu B. Diagnoses of HIV Infection in the United States and Dependent Areas, 2016. Atlanta, GA: CDC, 2016.
13. Kuo I, Greenberg AE, Magnus M, Phillips G 2nd, Rawls A, Peterson J, et al. High prevalence of substance use among heterosexuals living in communities with high rates of AIDS and poverty in Washington, DC. *Drug Alcohol Depend.* 2011; 117(2–3):139–44. Epub 2011/02/15. <https://doi.org/10.1016/j.drugalcdep.2011.01.006> PMID: 21316871.
14. Dailey A, Gant Z, Johnson S, Li J, Wang S, Hawkins D, et al. Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 (Updated). CDC, 2018.
15. Magnus M, Phillips G 2nd, Kuo I, Peterson J, Rawls A, West-Ojo T, et al. HIV among women in the District of Columbia: an evolving epidemic? *AIDS Behav.* 2014; 18 Suppl 3:256–65. Epub 2013/05/25. <https://doi.org/10.1007/s10461-013-0514-0> PMID: 23702704; PubMed Central PMCID: PMC5519335.
16. Doyle K, Xia M, Finlayson T, Wejnert C. HIV Infection, Risk, Prevention, and Testing Behaviors Among Heterosexuals at Increased Risk for HIV Infection National HIV Behavioral Surveillance 17 U.S. Cities, 2016. Atlanta, GA: CDC, 2016.
17. Team BS. National HIV Behavioral Surveillance—Round 4: Injection Drug Use (NHBS-IDU4) and Heterosexuals at Increased Risk for HIV (NHBS-HET4). In: CDC, editor. 2015.
18. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment.* 2008.
19. Mohamad NA, Ali Z, Noor NM, Baharum A. Multinomial logistic regression modelling of stress level among secondary school teachers in Kubang Pasu District, Kedah. 2016.
20. DeCarlo P, Reznick OG. What are Black women's HIV prevention needs? 2009. <https://doi.org/10.2105/AJPH.2008.140541> PMID: 19372518
21. Adimora AA, Schoenbach VJ, Doherty IA. HIV and African Americans in the southern United States: sexual networks and social context. *Sex Transm Dis.* 2006; 33(7 Suppl):S39–45. Epub 2006/06/24. <https://doi.org/10.1097/01.olq.0000228298.07826.68> PMID: 16794554.
22. Hader SL, Smith DK, Moore JS, Holmberg SD. HIV Infection in Women in the United States Status at the Millennium. *JAMA.* 2001; 285(9).
23. Turney K. Incarceration and Social Inequality. *The ANNALS of the American Academy of Political and Social Science.* 2013; 651(1):97–101. <https://doi.org/10.1177/0002716213501273>