Title Page

Title:

A Peer-Led, Artificial Intelligence-Augmented Social Network Intervention to Prevent HIV among Youth Experiencing Homelessness

Authors:

Eric Rice,[1*†] Bryan Wilder,[2†] Laura Onasch-Vera,[1] Graham DiGuiseppi,[1] Robin Petering,[3]
Chyna Hill,[1], Amulya Yadav,[4] Sung-Jae Lee,[5] Milind Tambe[2]

[1]Suzanne Dworak-Peck School of Social Work and Center for Artificial Intelligence in Society,
University of Southern California
[2]Center for Research on Computation and Society,
John A. Paulson School of Engineering and Applied Sciences, Harvard University
[3] Lens Co
[4]College of Information Sciences and Technology, Pennsylvania State University

[5] Department of Psychiatry and Biobehavioral Sciences, Division of Population Behavioral Health, David Geffen School of Medicine, University of California, Los Angeles
[*]Correspondence to: ericr@usc.edu
[†]Co-lead author

Sources of support: California HIV Research Project

Running Head: Peer-Led, Prevention Intervention & AI

Abstract/Keywords Page

Background: Youth experiencing homelessness (YEH) are at elevated risk for HIV/AIDS and disproportionately identify as racial, ethnic, sexual, and gender minorities. We developed a new peer change agent (PCA) HIV prevention intervention with three arms: (1) an arm using an Artificial Intelligence (AI) planning algorithm to select PCAs; (2) a popularity arm, the standard PCA approach, operationalized as highest degree centrality (DC); and (3) an observation-only comparison group.

Setting: A total of 713 YEH were recruited from 3 drop-in centers in Los Angeles, CA.

Methods: Youth were consented and completed a baseline survey that collected self-reported data on HIV knowledge, condom use, and social network information. A quasi-experimental pre-test/post-test design was used; 472 youth (66.5% retention at one month post-baseline) and 415 youth (58.5% retention at three months post-baseline) completed follow-up. In each intervention arm (AI and DC), 20% of youth were selected as PCAs and attended a four-hour initial training followed by 7 weeks of half-hour follow-up sessions. Youth disseminated messages promoting HIV knowledge and condom use.

Results: Using generalized estimating equation models, there was a significant reduction over time ($p<.001$) and a significant time by AI arm interaction ($p<.001$) for condomless anal sex. There was a significant increase in HIV knowledge over time among PCAs in DC and AI.

Conclusions: PCA models that promote HIV knowledge and condom use are efficacious for YEH. Youth are able to serve as a bridge between interventionists and their community. Interventionists should consider working with computer scientists to solve implementation problems.

Keywords: youth experiencing homelessness; artificial intelligence; social networks; HIV prevention; prevention interventions

Introduction

Each year, approximately 4.2 million youth aged 13 to 25 in the United States experience some form of homelessness.[1] HIV prevalence of youth experiencing homelessness (YEH) far exceeds that of their housed counterparts.[2] HIV disparities for YEH are in part the result of systemic HIV disparities based on race/ethnicity, sexual orientation, and gender identity. In particular, Black and Latinx youth are more likely to experience homelessness as compared to their White peers.[1] Moreover, 20 to 40 percent of YEH identify as members of LGBTQ communities.[3]

One solution to HIV prevention is the *peer change agent* (PCA) model.[4-6] Given the central role that peers play in the HIV risk and protective behaviors of YEH,[7-10] researchers have suggested that a PCA model for HIV prevention should be developed for YEH.[7,9,11] PCA models identify a small number of individuals in a high-risk target population to become advocates in their community. These individuals are tasked with disseminating HIV prevention information and norm-changing messages to their peers.[4-6]

PCA models are effective for HIV prevention in many contexts in studies ranging from the mid 1990s to the present,[4-6,12] though there have been some notable failures.[13,14] Some failure has been attributed to focusing on health education rather than messages focused on changing norms.[14] Recently, however, Schneider and colleagues [15] suggested that PCA model failures may be due to *how* PCAs are selected. They argued that the change agents who are selected to do the PCA work can often be as important, if not more important, than the messages they convey. As developed by Kelly,[4,5] the standard method for selecting PCAs uses ethnographic methods to identify the most popular individuals in the social network. This can be operationalized more formally as selecting PCAs who have the greatest number of network connections to others in a population, a concept known in social network terminology as highest degree centrality. Several authors—particularly Valente and Pumpuang [16]—have described how network-driven prevention programs can benefit from explicitly modeling social networks and leveraging network methods in the context of intervention delivery.

Recent work has piloted the development of AI methods to improve the process of selecting PCAs.[17-22] This work was based on influence-maximization research in computer science. [22,23] While AI is still relatively new to many researchers in medicine, public health, and social work, techniques from AI, particularly machine learning approaches, have gained visibility and traction in recent years.[24-26]

The purpose of this paper is to present results of a quasi-experimental design that compares a PCA model delivered in drop-in centers with three study arms: (1) PCA selection based on AI; (2) PCA selection based on popularity, operationalized as YEH with highest degree centrality (DC); and (3) an observation-only control group (OBS).

Methods

Participants, Sampling, and Study Design

All study procedures were approved by the University of Southern California Institutional Review Board. This is a quasi-experimental, three-group (AI, DC, OBS), pre-test/post-test design. A total of 713 youth were recruited across nine networks of YEH (ages 16-24) from three different drop-in centers in different geographic locations across Los Angeles, CA, from September 2016 to October 2018. At each drop-in center, each arm of the study was conducted once, with recruitment separated by a six-month interval to allow for sufficient numbers of new youth to replace prior clients (see Figure 1). As outcomes are measured at the level of individuals within the network, at the level of the PCA, and diffusion of norms nearly guarantees contamination, randomization at the level of the individual within these 9 networks is not possible. All youth receiving services were eligible to participate and were informed of the study as they entered the drop-in center. One lead study staff recruited participants throughout the study to ensure participants were included only once.

[Figure 1 about here.]

Intervention Design and Delivery

Our intervention design was based on previous literature, community collaborations, youth input, and research team members' long-term experience working with YEH in both research and service delivery contexts. The intervention design was also informed by multiple theories, including the risk amplification and abatement model [27] and diffusion of innovations.[28]

In both AI and DC, PCAs were recruited over a three-week period and each training session was limited to a maximum of five participants. To achieve the desired total number of PCAs in the network, researchers conducted three subsequent trainings with smaller groups of participants. Training was delivered by two or three facilitators—an MSW staff lead and/or MSW student intern(s). The primary intervention training lasted approximately four hours (i.e., one half-day). The training was interactive and broken into six 45-minute modules on the mission of PCAs, sexual health, HIV prevention, communication skills, leadership skills, and self-care. Training minimized lecture-based learning and was designed to engage youth in a variety of learning activities, including group discussion, games, journaling and reflection, experiential learning, and role-playing. PCAs were asked to focus their communications on their social ties, particularly other youths at the drop-in center, and to promote regular HIV testing and condom use. PCAs were encouraged to focus on face-to-face interactions if possible but to use social media as well. The initial training was supported by seven weeks of 30-minute follow-up check-in sessions. Ninety-one percent of PCAs checked in at least once; the modal attendance for check-in sessions was four sessions. PCAs received $60 for the training and $20 for each check-in session.

PCA Selection Methods

Social Network Data. Assessment of whole (sociometric) networks followed an event-based approach,[29] wherein each network was composed of relational ties between youth receiving services within the defined boundaries of each of the three drop-in center at a given point in time. Assessment of sociometric networks was carried out by conducting an in-person interview with each participant that asked "Who are your friends at this agency?" Participants could list up to 10 friends (*alters* in social network terminology). A consistent team of two research assistants involved in participant recruitment at each agency then determined whether alters were also enrolled as research participants in the study. Given that network data collection was staggered for the three arms at three separate drop-in centers, this resulted in nine sociometric networks. Sociometric network ties at baseline were used to select peer leaders (discussed below).

PCA Selection Methods. In the DC arm, PCAs were selected based on popularity as defined by degree centrality—that is, the 20% of youths with the greatest number of ties to other youths were recruited.[16] In the AI arm, 20% of youth in the network were selected for PCA training based on the CHANGE algorithm. In each arm (AI and DC), agents are selected from a given network by only one method. Formal modelling details of the CHANGE algorithm and computational experiments of this AI selection procedure have been published previously in computer science outlets.[17-21] What follows is description of how the CHANGE algorithm works targeting a non-computer science readership.

The input for the CHANGE[20] algorithm are data on individuals and their network connections to one another (i.e. data on a face-to-face social network). The output of the algorithm is a set of individuals who should be trained to be PCA. These individuals, as a group, represent the persons who ought to be trained so as to maximize the influence of that small group with respect to the overall network, as determined purely by their relative position to one another and to full set of persons in the network.

To accomplish this task, CHANGE seeks to solve for three intervention implementation challenges.[20] First, collecting face-to-face social network data is very time consuming and difficult, so CHANGE seeks to reduce this burden. Second, when we consider the diffusion of innovations (or intervention messages) across an actual face-to-face social network, it is almost impossible to measure the exact rate (i.e. probability) of message transmission from person to person, which complicates the task of selecting exactly the right persons. And third, not every PCA that CHANGE selects as part of the group of most influential persons in the network may be available or willing to be trained, as homeless youth have very chaotic daily lives and have particular difficulty securing reliable transportation.

First, to solve the network data collection challenge, CHANGE takes as an initial input the roster of all persons in the network, absent any network tie data. CHANGE randomly selects one in ten persons and the field team then collects network tie information from that initial ten percent of persons. Then the algorithm randomly selects 1 person from the listed set of network ties collected from each of these initial individuals and the team collects network tie information from this second layer of persons. This generates a “sketch” of the network based on information collected only from one fifth of the network. In prior papers we have shown this “sketch” has sufficient detail to run the “influence maximization” step which follows.[20]

Second, the network “sketch” is entered as input. Ultimately, the intervention team trains 20% of the population to be PCA. This is done over a series of four weeks. CHANGE selects the 5% most influential youth each week to be trained. Rather than always selecting the persons with the largest number of ties (as is the case with degree centrality), CHANGE picks persons who are ideally positioned within sub-networks of the overall network who collectively maximize the reach of message dissemination.[22] Often the most connected persons in a network are also connected to each other and thus their potential reach is redundant; CHANGE takes these redundancies into account when generating the group of most influential persons to be trained as PCAs. As mentioned above, the exact rate of message dissemination is unknown and CHANGE treats this aspect of the diffusion process as probabilistic and uncertain, which adds substantially to the scale of the computational problem which CHANGE solves.

Finally, as the training proceeds in four sequential weeks, CHANGE keeps track of who failed to attend the trainings in prior rounds and searches for appropriate replacements for those particular individuals in subsequent weeks. Imagine a sub-group of 5 friends, where A is connected to B,C,D,and E, but where B is connected to A, C, D but not E. CHANGE would initially select Person A as the most locally influential person. If A does not attend the training, the algorithm will select person B the following week. CHANGE seeks to impact this subnetwork as much a possible given the initial failure. The algorithm is adaptive and “learns” about the network as the field team learns about the network and makes new decisions based on new information.

Data Collection

Participants completed a self-administered survey assessing their demographics, sexual behaviors, HIV knowledge, HIV testing behaviors, and their social networks. Three waves of survey data were collected: baseline and follow-ups at one and three months post-baseline. A flow diagram depicts the overall study design and participant retention for each arm (see Figure 1).

Measures

Demographics. Self-reports of age, birth sex, gender identity, race/ethnicity, sexual orientation, and current housing situation were assessed at baseline. Participants whose gender identity differed from their assigned sex at birth, or who reported gender identity as “trans-male,” “trans-female,” “gender queer/non-conforming,” or something else, were coded as transgender/non-binary. Current housing situation was assessed by self-report by asking participant to choose from a list of settings where they spent most of their nights during the past two weeks. Participants were then categorized as living in (1) an emergency shelter or transitional living placement, (2) unsheltered, or (3) unstably housed (i.e., “couch surfing” with friends or extended family).

Sexual Risk Behavior. Sexual risk behavior was operationalized using two variables: *condomless anal sex* (CAS) and *condomless vaginal sex* (CVS). Participants were asked to list their five most recent sexual partners in the past month. For each partner, participants were asked whether they had (a) vaginal sex without a condom, (b) anal sex without a condom, or (c) both anal and vaginal sex without a condom during their last sexual encounter. A binary variable was created indicating whether a participant had CAS (yes=1; no=0) or CVS (yes=1; no=0) with at least one sexual partner in the past month. Participants who reported no condomless sex or no sexual partners in the past month were coded as zero on these variables.

HIV Testing. HIV testing in the past six months was assessed by asking, “When was the last time you had an HIV test?” A binary variable was created such that participants who reported their last test occurring within the past six months were coded as 1; participants who reported their last test occurring more than six months ago or who were never tested were coded as 0.

HIV Knowledge. Participants completed a brief, six-question HIV knowledge quiz. True-or-false and multiple choice questions developed by our research team tested participant knowledge of HIV transmission, HIV testing, and local, population-relevant statistics (e.g., “How many homeless youth in Los Angeles are infected with HIV?”). Percentage of correct responses was used as the main outcome measure.

Data Analysis

Significant differences on descriptive characteristics between arms at baseline were tested using chi-square or ANOVAs. Generalized Estimating Equations (GEEs), a population-averaged extension of generalized linear models for repeated-measures data (Zeger et al., 1988), were used to test intervention effects for the outcome. GEEs predicting binary outcomes used a logit link function and a linear GEE model was specified for HIV knowledge (a continuous variable). We specified an unstructured working correlations matrix for all GEE models, given that differences in quasi-likelihood information criteria were negligible across different specifications. Bivariate logistic regression models indicted that participants who had greater odds of missingness at both follow-ups had lower HIV knowledge, were more likely to be in the DC group, were less likely to be residing in a shelter or transitional living placement, used drop-in centers less frequently and for shorter periods of time.

All GEE models included terms for AI (ref: OBS), DC (ref: OBS), time, and time-by-group interaction terms to assess whether change in outcomes differed for groups over time. Each model adjusted for demographic covariates. A binary term indicating whether a participant was a PCA was added, along with a PCA-by-time interaction term, to determine whether outcomes changed over time for PCAs relative to non-PCAs. With PCA-by-time included in the models, AI-by-time and DC-by-time interaction terms that remained significant would indicate that change in the outcome was not due to changes in behavior of the PCAs alone.

Results

[Table 1 about here.]

Descriptive statistics are presented in Table 1. Table 2 provides a detailed breakdown of the outcome measures at each time point. At baseline, a number of significant differences in participant characteristics were found across the three arms; namely, birth sex, LGBQ identity, living situation,; and which drop-in center participants were recruited from (all $p < .05$). Therefore, these and other characteristics were included as covariates in multivariable GEE models.

[Table 2 about here.]

Condomless Anal Sex (CAS). As shown in Table 3, a significant group-by-time interaction was found for the AI group (OR = 0.67, 95% CI: 0.46, 0.95, $p =$ 0.03), suggesting that improvements in CAS were driven by behavior changes among youth not selected to be PCAs. Post-hoc analyses of change at discrete time points showed a significantly greater reduction in CAS from baseline to the one-month follow-up in the AI group relative to the DC group (OR = 0.43, 95% CI: 0.19, 0.97, $p =$ 0.04). This significant reduction remained after controlling for PCAs in the model. Direct examination of the prevalence of CAS at each time-point (Table 2) shows that improvements in the AI group happened faster than the DC group. Most of the improvement for AI occured by the one-month survey, while improvements in the DC group were not fully realized until month three.

[Table 3 about here.]

Condomless Vaginal Sex (CVS). There was a marginally significant AI-by-time interaction (OR = 0.75, 95% CI: 0.55, 1.03, $p =$ 0.08). Post-hoc analyses at discrete time points showed significantly greater reductions in CVS from the one-month to the three-month follow-up in the AI group (relative to observation only) (OR = 0.39, 95% CI: 0.19, 0.79, $p = 0.01$).

HIV Knowledge and Testing. The PCA-by-time interaction term was significant (b = 0.11, SE = 0.02, $p < .001$), indicating that PCAs increased their HIV knowledge over time. Similarly, a PCA-by-time term was significant (OR = 1.82, 95% CI: 1.07, 3.09), indicating that PCAs had higher odds of testing for HIV over time.

Discussion

The findings from our study provide a unique and innovative strategy to optimize PCA selection by using the AI algorithm CHANGE. Our study utilized a quasi-experimental design with three arms: an observation only arm, a typical PCA selection method arm (i.e., namely popularity–highest degree centrality), and a PCA selection arm using the AI CHANGE algorithm. Results indicate using the AI algorithm is an efficacious intervention approach. First, and most importantly, there was a significant reduction in condomless anal sex for those in the AI arm, as indicated by the significant arm-by-time interaction. In the AI arm, there was a significant, 33% reduction in the odds of condomless anal sex over time compared to the observation only arm. By contrast, there were no statistically significant changes over time in the DC arm. Moreover, post-hoc analyses at discrete time points showed significantly greater reductions in CVS from the one-month to the three-month follow-up in the AI group. Together, these findings suggest that the AI algorithm does a better job in selecting PCAs than DC (i.e., popularity). As we mention in the results, there were some significant differences across the study arms with respect to population demographics, despite recruiting at the same three agencies across each study arm (one wave of recruitment per agency per arm for nine total recruitments). To account for this issue, we control for population demographics in the GEE models. Thus, our observed differences in behaviors over time across study arms are not due to differences in population characteristics.

Second, improvements in the CHANGE group happened faster than in the DC group; most of the improvement for CHANGE occurred by the one-month survey, while improvements in the DC group were not fully realized until month three. The purpose of the CHANGE algorithm is to identify a set of agents who have the most rapid and extensive reach. CHANGE identifies key persons throughout the network space, whereas popularity often results in selecting redundant change agents. Popular people often have overlapping ties, including to other popular individuals and/or to a shared set of others (i.e., multiple popular people are connected to the same others). So information may eventually diffuse through the entire network under popularity, but less efficiently compared to when CHANGE selects the agents. Fast results are important for two reasons. First, rapid adoption of protective behaviors helps to immediately curtail transmission in a high-risk population. Second, high transience among YEH means that a non-negligible portion of youth will have left the center by the time a three-month intervention is complete. We conclude that the AI-augmented intervention implemented with CHANGE has substantial advantages over an intervention in which PCAs are selected with the DC method.

Third, we observed improvements in HIV knowledge and HIV testing over time among the subset of youth who were trained as PCAs. The PCA's had extended contact with the research team and developed a level of commitment to HIV prevention. Thus, it is not surprising to see some HIV prevention behaviors change significantly for the PCAs, even if they were not effectively disseminated to others in the network.

The findings of this work suggest that AI may be a useful tool to augment public health and social work intervention design. Creating and testing new behavioral health interventions with human subjects is a time-consuming and costly endeavor. Computer science relies heavily on computational experiments to test and refine algorithms. This activity allows one to discard suboptimal solutions at relatively low cost. In our case, several alternatives were tested and discarded in early computational experiments [18,19,23] without ever reaching the stage of field tests with human subjects.

There are some limitations to the current study that must be acknowledged. First, this is a quasi-experimental design, not a randomized control trial. Because the intervention seeks to train PCAs to disseminate information within the networks in which they are embedded, randomization at the level of individual study participants was not possible. Second, all behavioral data come from self-reports. Third, YEH are a highly transient group and retention in the study over time is very challenging. Our study retention rate of 59% was somewhat lower than rates reported in other studies involving longitudinal follow-up of YEH.[30,31] However, this past research was conducted within more stable populations of new runaways (most all of whom returned home) [30] or youth recruited from shelter services.[31] Youth with chronic experiences of homelessness accessing drop-in centers are far more difficult to track and retain. Indeed, Bender and colleagues[32] worked with a similar drop-in-based population and reported similar retention rates.

Our hope is that this project provides an example for a broader research agenda exploring AI techniques to improve health and equity within our communities. In the past few years, we have observed a growing interest in how machine learning techniques in AI may be incorporated into health and behavioral health contexts. [24-26] We see the current study, however, as demonstrating how AI and social science can collaborate beyond the sphere of predictive analytics into enhanced behavioral health and prevention intervention design. The results from this intervention study provide evidence that AI can substantially improve the impact of services offered to the most vulnerable and disproportionately impacted communities.

One future challenge to be addressed is how to implement an AI method in community settings. Our intention is to make our algorithm available online for free to any community agency. We believe that the biggest obstacle to community organizations using this technology is not the technology itself, which we will give away, but rather the time and energy required to collect social network data. The CHANGE algorithm only needs to collect network information from 1 in 5 members of population, but this may still be difficult. In recent a paper we show that data on shared attendance in specific programming at the drop in center (for example a writing workshop) successfully approximates network connections for the purposes of our algorithm.[33] Linking our algorithm to attendance records, which could be a passive data collection linked to ongoing practice, seems a more useful solution rather than training staff to collect network data. Such an approach seems an important next step to test in order to increase community uptake of this intervention.

References

1. Morton MH, Dworsky A, Matjasko JL, et al. Prevalence and correlates of youth homelessness in the United States. *J Adolesc Health*. 2018;62:14-21.
2. Caccamo A, Kachur R, Williams SP. Narrative review: Sexually transmitted diseases and homeless youth-what do we know about sexually transmitted disease prevalence and risk? *Sex Transm Dis*. 2017;44:466-476.
3. McCann E, Brown M. Homelessness among youth who identify as LGBTQ+: A systematic review. *J Clin Nurs*. 2019;28:2061-2072.
4. Kelly JA, Murphy DA, Sikkema KJ, et al. Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities. Community HIV Prevention Research Collaborative. *Lancet*. 1997; 350:1500-5.
5. Latkin CA, Sherman S, Knowlton A. HIV prevention among drug users: outcome of a network-oriented peer outreach intervention. *Health Psychol*. 2003;22:332-9.
6. Schneider JA, Young L, Ramachandra A, et al. "A Pragmatic Randomized Controlled Trial to Increase PrEP Uptake for HIV Prevention: 55-Week Results From PrEPChicago." *J Acquir Immune Defic Syndr*. 2021; 86: 31-37.
7. Green HD Jr, de la Haye K, Tucker JS, et al. Shared risk: who engages in substance use with American homeless youth? *Addiction*. 2013;108:1618-24.
8. Rice E, Monro W, Barman-Adhikari A, Young SD. Internet use, social networking, and HIV/AIDS risk for homeless adolescents. *J Adolesc Health*. 2010;47:610-3.
9. Rice E, Barman-Adhikari A, Milburn NG, et al. Position-specific HIV risk in a large network of homeless youths. *Am J Public Health*. 2012;102:141-7.

10. Valente AM, Auerswald CL. Gender differences in sexual risk and sexually transmitted infections correlate with gender differences in social networks among San Francisco homeless youth. *J Adolesc Health*. 2013;5:486-91.

11. Arnold EM, Rotheram-Borus MJ. Comparisons of prevention programs for homeless youth. *Prev Sci*. 2009;10:76-86.

12. Medley A, Kennedy C, O'Reilly K, et al. Effectiveness of peer education interventions for HIV prevention in developing countries: a systematic review and meta-analysis. *AIDS Educ Prev*. 2009;21:181-206.

13. NIMH Collaborative HIV/STD Prevention Trial Group. Results of the NIMH collaborative HIV/sexually transmitted disease prevention trial of a community popular opinion leader intervention. *J Acquir Immune Defic Syndr*. 2010;54:204-14.

14. Kelly JA. Popular opinion leaders and HIV prevention peer education: resolving discrepant findings, and implications for the development of effective community programmes. *AIDS Care*. 2004;16:139-50.

15. Schneider JA, Zhou AN, Laumann EO. A new HIV prevention network approach: sociometric peer change agent selection. *Soc Sci Med*. 2015;125:192-202.

16. Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. *Health Educ Behav*. 2007;34:881-96.

17. Rice E, Yoshioka-Maxwell A, Petering R, et al.. Piloting the use of artificial intelligence to enhance HIV prevention interventions for youth experiencing homelessness. *J Soc Soc Work and Research* . 2018;9: 551–73.

18. Wilder B, Yadav A, Immorlica N, et al. Uncharted but not uninfluenced: Influence maximization with an uncertain network. Presented at: 16th Conference on Autonomous Agents and MultiAgent Systems; 2017; Sao Paulo.
19. Wilder B, Immorlica N, Rice E, et al. Maximizing influence in an unknown social network." Presented at: Thirty-Second AAAI Conference on Artificial Intelligence; 2018; New Orleans.
20. Wilder B, Onasch-Vera L, Hudson J, et al. End-to-end influence maximization in the field. Presented at: 17th International Conference on Autonomous Agents and MultiAgent Systems; 2018; Stockholm.
21. Yadav A, Wilder B, Rice E, et al. Influence maximization in the field: The arduous journey from emerging to deployed application. Presented at: 16th Conference on Autonomous Agents and MultiAgent Systems; 2017; Sao Paulo.
22. Kempe D, Kleinberg J, Tardos E. Maximizing the spread of influence through a social network. Presented at: Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2003; Washington D.C.
23. Yadav A, Marcolino L, Rice E, et al. Preventing HIV spread in homeless populations using PSINET. Presented at: 27th Conference on Innovative Applications of Artificial Intelligence; 2015; Austin.
24. Ahn WY, Ramesh D, Moeller FG, et al. Utility of machine-learning approaches to identify behavioral markers for substance use disorders: Impulsivity dimensions as predictors of current cocaine dependence. *Front Psychiatry*. 2016;10:34.
25. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - Addressing ethical challenges. *N Engl J Med*. 2018;378:981-983.

26. Verghese A, Shah NH, Harrington RA. What this computer needs Is a physician: Humanism and artificial intelligence. *JAMA*. 2018;319:19-20.

27. Milburn NG, Rice E, Rotheram-Borus MJ, et al. Adolescents exiting homelessness over two years: The risk amplification and abatement model. *J Res Adolesc*. 2009;19:762-785.

28. Rogers, Everett M. 2010. *Diffusion of Innovations*. New York: Simon & Schuster;2010.

29. Freeman L, Webster CM. Interpersonal proximity in social and cognitive space. *Soc Cognit.* 1994; 12: 223–47.

30. Milburn NG, Iribarren FJ, Rice E, et al. Family intervention to reduce sexual risk behavior, substance use, and delinquency among newly homeless youth. *J Adolesc Health.* 2012;50:358-64.

31. Rotheram-Borus MJ, Song J, Gwadz M, et al. Reductions in HIV risk among runaway youth. *Prev Sci*. 2003;4:173-87.

32. Bender KA, DePrince A, Begun S, et al. Enhancing risk detection among homeless youth: A randomized clinical trial of a promising pilot intervention. *J Interpers Viol.* 2018; 33:2945-2967.

33. Young LE, Mayaud J, Suen SC, et al. Modeling the dynamism of HIV information diffusion in multiplex networks of homeless youth. *Soc Net.* 2020; 63:112-121.

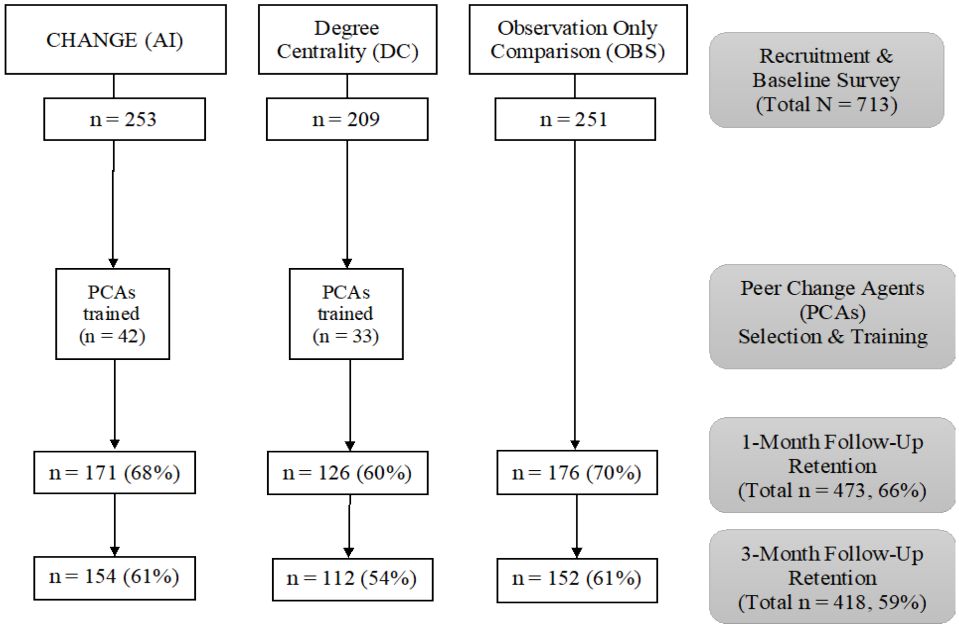

Figure 1. Participant Recruitment and Flow Diagram

Table 1. Participant Characteristics at Baseline (N = 713).

Variable	AI (*n* = 253) N (%) or M (SD)	DC (*n* = 209) N (%) or M (SD)	OBS (*n* = 251) N (%) or M (SD)	Full Sample N (%) or M (SD)
Age	22.0 (2.0)	21.8 (2.2)	21.8 (2.2)	21.9 (2.1)
Gender				
Male*	193 (76.3%)	164 (78.5%)	175 (69.7%)	533 (74.6%)
Female	60 (23.7%)	42 (20.1%)	76 (30.3%)	178 (24.9%)
Transgender/non-binary	32 (12.6%)	25 (12.0%)	37 (14.7%)	94 (13.2%)
Race/Ethnicity				
Black/African American	88 (34.8%)	64 (30.6%)	69 (27.5%)	221 (31.0%)
Non-Hispanic White				160 (22.4%)
Hispanic or Latino/a/x	33 (13.0%)	27 (12.9%)	45 (17.9%)	106 (14.8%)
Multiple	70 (27.7%)	50 (23.9%)	59 (23.5%)	179 (25.1%)
Other[a]	15 (5.9%)	14 (6.7%)	19 (7.6%)	48 (6.7%)
LGBQA*	120 (47.4%)	74 (35.4%)	112 (44.6%)	306 (42.9%)
Romantic Relationship (Current)	88 (34.8%)	66 (31.6%)[b]	96 (38.2%)	251 (35.2%)
Housing				
Shelter/Transitional Living	57 (22.5%)	56 (26.8%)[b]	47 (18.7%)	160 (22.4%)
Unstably Housed*	94 (37.2%)	99 (47.4%)[b]	94 (37.5%)	287 (40.2%)
Street (Unsheltered)*	102 (40.3%)	54 (25.8%)[b]	110 (43.8%)	267 (37.4%)
Drop-in Center*				
My Friend's Place	64 (25.3%)	74 (35.4%)	90 (35.9%)	228 (31.9%)
Youth Center on Highland	96 (37.9%)	81 (38.8%)	80 (31.9%)	257 (36.0%)
Safe Place for Youth	93 (36.8%)	54 (25.8%)	81 (32.3%)	228 (31.9%)

* Significant between-group differences at $p < .05$

[a] Other race includes Asian, Native, Pacific Islander, and "other" as a write-in category

[b] Data missing from first wave of DC (n = 54)

Table 2. Outcomes Over Time, Stratified by Intervention Group (N = 713).

Variable	**AI**		**DC**		**OBS**	
Condomless Anal Sex (CAS)	*n (%)*	*Total N*	*n (%)*	*Total N*	*n (%)*	*Total N*
Baseline	69 (27%)	253	69 (33%)	205	54 (22%)	251
1M	31 (18%)	171	43 (35%)	124	37 (21%)	176
3M	27 (18%)	154	26 (24%)	108	36 (24%)	152
Condomless Vaginal Sex (CVS)	*n (%)*	*Total N*	*n (%)*	*Total N*	*n (%)*	*Total N*
Baseline	90 (36%)	253	87 (42%)	206	116 (46%)	251
1M	51 (30%)	171	53 (43%)	124	62 (35%)	176
3M	31 (20%)	154	34 (32%)	108	61 (40%)	152
HIV Testing (past six months)	*n (%)*	*Total N*	*n (%)*	*Total N*	*n (%)*	*Total N*
Baseline	183 (74%)	246	139 (67%)	207	181 (72%)	249
1M	136 (82%)	166	101 (80%)	126	141 (81%)	174
3M	114 (75%)	152	79 (74%)	107	123 (83%)	149
HIV Knowledge Test (% correct)	*M (SD)*	*Total N*	*M (SD)*	*Total N*	*M (SD)*	*Total N*
Baseline	58% (20%)	251	56% (22%)	154	57% (20%)	249
1M	69% (22%)	167	64% (23%)	88	63% (20%)	174
3M	65% (23%)	152	63% (25%)	74	59% (19%)	150

Table 3. Including all outcomes in one table. Generalized Estimating Equations Predicting Study Outcome Variables, including PCA and PCA x Time Interaction.

Outcome Variable	**Predictor**	***b***	***SE***	**OR**	**95% CI**	***p***
Condomless Anal Sex (CAS)						
	Time	--	--	1.05	0.82, 1.33	0.72
	AI	--	--	1.44	0.89, 2.32	0.14
	DC	--	--	1.49	0.88, 2.52	0.14
	PCA	--	--	1.00	0.56, 1.78	0.99
	AI x Time	--	--	**0.67**	**0.46, 0.95**	**0.03**
	DC x Time	--	--	0.78	0.52, 1.15	0.20
	PCA x Time	--	--	1.17	0.73, 1.89	0.51
Condomless Vaginal Sex (CVS)						
	Time	--	--	0.87	0.71, 1.07	0.18
	AI	--	--	0.79	0.53, 1.18	0.25
	DC	--	--	1.09	0.68, 1.74	0.72
	PCA	--	--	0.86	0.49, 1.50	0.59
	AI x Time	--	--	*0.75*	*0.55, 1.03*	*0.08*
	DC x Time	--	--	0.85	0.60, 1.22	0.38
	PCA x Time	--	--	1.17	0.76, 1.81	0.47
HIV Testing						
	Time	--	--	**1.34**	**1.05, 1.71**	**0.02**
	AI	--	--	1.21	0.77, 1.91	0.40
	DC	--	--	0.95	0.58, 1.56	0.84
	PCA	--	--	0.79	0.41, 1.53	0.48
	AI x Time	--	--	*0.73*	*0.52, 1.03*	*0.07*
	DC x Time	--	--	0.91	0.58, 1.41	0.66
	PCA x Time	--	--	**1.82**	**1.07, 3.09**	**0.03**
HIV Knowledge						

Time	0.002	0.009	--	-0.02, 0.02	0.79
AI	-0.02	0.02	--	-0.05, 0.02	0.34
DC	*-0.04*	*0.02*	--	*-0.08, 0.01*	*0.09*
PCA	0.02	0.03	--	-0.03, 0.07	0.38
AI x Time	0.02	0.01	--	-0.01, 0.04	0.20
DC x Time	0.02	0.02	--	-0.01, 0.05	0.27
PCA x Time	**0.11**	**0.02**	**--**	**0.07, 0.16**	**< 0.001**

Note: Models adjusted for age, male birth sex, transgender, LGBQ identity, male X LGBQ, race, committed relationship, housing status, drop-in center.
Note: **Bolded** parameter estimates are significant at $p < .05$.
Note: Italicized parameter estimates are significant at $p < .10$