OPEN

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

Eric Rice, PhD,[a] Bryan Wilder, PhD,[b] Laura Onasch-Vera, MSW,[a] Graham DiGuiseppi, ScM,[a] Robin Petering, PhD,[c] Chyna Hill, PhD,[a] Amulya Yadav, PhD,[d] Sung-Jae Lee, PhD,[e] and Milind Tambe, PhD[b]

Background: Youth experiencing homelessness (YEH) are at elevated risk of HIV/AIDS and disproportionately identify as racial, ethnic, sexual, and gender minorities. We developed a new peer change agent (PCA) HIV prevention intervention with 3 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 consented and completed a baseline survey that collected self-reported data on HIV knowledge, condom use, and social network information. A quasi-experimental pretest/posttest design was used; 472 youth (66.5% retention at 1 month postbaseline) and 415 youth (58.5% retention at 3 months postbaseline) completed follow-up. In each intervention arm (AI and DC), 20% of youth was selected as PCAs and attended a 4-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 ($P < 0.001$) and a significant time by AI arm interaction ($P < 0.001$) for condomless anal sex act. There was a significant increase in HIV knowledge over time among PCAs in DC and AI arms.

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.

Key Words: youth experiencing homelessness, artificial intelligence, social networks, HIV prevention, prevention interventions

(*J Acquir Immune Defic Syndr* 2021;88:S20–S26)

INTRODUCTION

Each year, approximately 4.2 million youth aged 13–25 years 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 partly 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 when compared with their White peers.[1] Moreover, 20%–40% YEH identify as members of lesbian, gay, bisexual, transgender or queer 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] although 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 et al[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

From the [a]Suzanne Dworak-Peck School of Social Work and Center for Artificial Intelligence in Society, University of Southern California Los Angeles, CA; [b]Center for Research on Computation and Society, John A. Paulson School of Engineering and Applied Sciences, Harvard University Cambridge, MA; [c]Lens Co; [d]College of Information Sciences and Technology, Pennsylvania State University College Station, TX; and [e]Department of Psychiatry and Biobehavioral Sciences, Division of Population Behavioral Health, David Geffen School of Medicine, University of California, Los Angeles Los Angeles, CA.

Supported by California HIV Research Project.

The authors have no funding or conflicts of interest to disclose.

E.R., and B.W. are colead authors.

Correspondence to Eric Rice, PhD Suzanne Dworak-Peck School of Social Work, University of Southern California, 669 W 34th St, Los Angeles, CA 90089 (e-mail: ericr@usc.edu).

selecting PCAs who have the greatest number of network connections with others in a population, a concept known in social network terminology as the highest degree centrality (DC). 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 artificial intelligence (AI) methods to improve the process of selecting PCAs.[17–22] This work was based on influence maximization research in computer science.[22,23] Although 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 article is to present results of a quasi-experimental design that compares a PCA model delivered in drop-in centers with 3 study arms: (1) PCA selection based on AI; (2) PCA selection based on popularity, operationalized as YEH with highest DC; and (3) an observation-only comparison 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, 3-group (AI, DC, and OBS), pretest/posttest design. A total of 713 youth were recruited across 9 networks of YEH (aged 16–24 years) from 3 different drop-in centers in 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 6-month interval to allow for sufficient numbers of new youth to replace previous clients (Fig. 1). Because 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.

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 3-week period, and each training session was limited to a maximum of 5 participants. To achieve the desired total number of PCAs in the network, researchers conducted 3 subsequent trainings with smaller groups of participants. Training was delivered by 2 or 3 facilitators—an masters in social work staff lead and/or masters in social work student intern(s). The primary intervention training lasted approximately 4 hours (ie, 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 7 weeks of 30-minute follow-up check-in sessions. Ninety-one percentage of PCAs checked in at least once; the modal attendance for check-in sessions was 4 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 3 drop-in centers at a given point in time. Assessment of sociometric networks was performed 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 2 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 3 arms at 3 separate drop-in centers, this resulted in 9 sociometric networks. Sociometric network ties at baseline were used to select peer leaders (discussed further).

PCA Selection Methods

In the DC arm, PCAs were selected based on popularity, as defined by DC—that is, the 20% of youths with the greatest number of ties to other youths was recruited.[16] In the AI arm, 20% of youth in the network was 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. Details of the development and computational experiments of this AI selection procedure are available elsewhere.[17–19]

What follows is the logic of the CHANGE algorithm.[19] The population of YEH is modeled as nodes of a graph $G = (V, E)$. We seek to recruit a set of youth S to be PCAs, where $|S| \leq k$. This budget constraint reflects the fact that PCAs are given a resource-intensive training and support process. The objective is to maximize the total expected number of youth who receive information about HIV prevention, given by the function $f(S)$. Here, f represents the number of youth receiving information in expectation over a

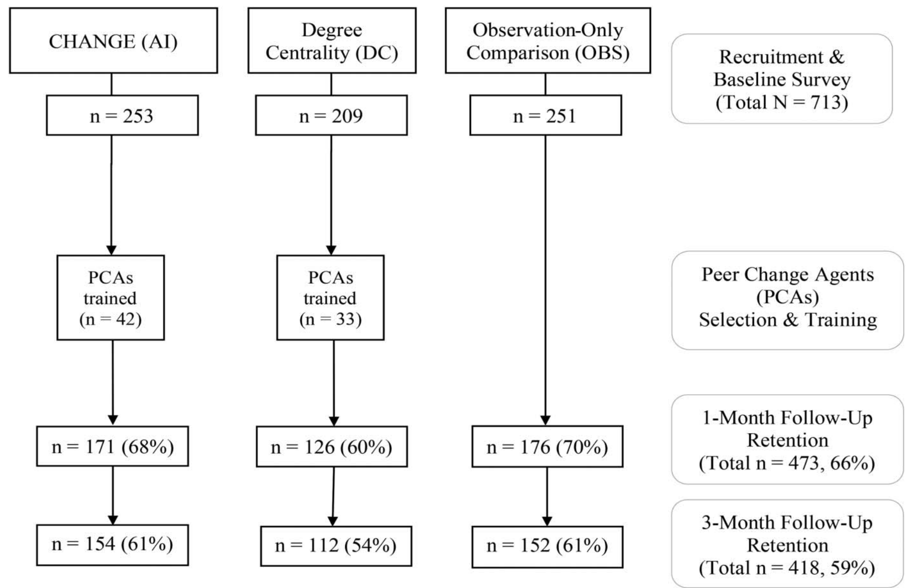

FIGURE 1. Participant recruitment and flow diagram.

probabilistic model of information diffusion across the network. We use the standard independent cascade model[22] in which each node who receives information transmits it to each of their neighbors with probability P. The cascade model[22] has become the standard computational model used in computer science to examine the influence dynamics and diffusion in networks. Within computer science, much work has been done to explore how to spread information in networks most efficiently (or optimally), an area known as influence maximization. The optimization problem $max_{|S|\leq k}f(S)$ is the subject of this well-known influence maximization problem. In this study, we detail 3 steps for deploying an influence maximization intervention in the field and provide some high-level description of our proposed solutions, which build on our previous work.[18,19]

First, information about the network structure G must be gathered. The relevant network is that of face-to-face interactions. Collecting such data requires time-intensive interviews with youth about their connections. Accordingly, the first stage of our algorithmic problem is to decide which nodes to query for network information (ie, which youth to interview about their network ties). Ideally, only a small fraction of nodes is queried, substantially reducing the cost of the intervention. CHANGE uses a simple heuristic to select nodes to query. Queries are made in pairs. The first query in the pair is made to a node selected uniformly at random from the network. The second query in the pair selects a uniformly random neighbor of the first node. These steps are repeated until approximately 20% of the network has been sampled. Thus, only one in 5 youth need to be interviewed for network information, increasing the ease of data collection.

Second, this network information is used to select an initial set of PCAs. This stage more closely resembles the standard influence maximization problem.[22] However, there is an additional complication that the propagation probability P is not known. Indeed, there is no data source from which it could be inferred. Instead, we formulate an uncertainty set $\mathscr{P}$ containing a set of possible values for P, which are consistent with previous knowledge (in CHANGE, we took $\mathscr{P}$ to be a discretization of the interval (0,1), reflecting limited previous knowledge). The aim is to find a set S that performs near-optimally for every scenario contained in $\mathscr{U}$. We developed a robust optimization algorithm that obtains a provable approximation for this problem.[18,19]

Third, after an initial set of PCAs S is identified, recruitment proceeds in an adaptive manner. Not all youth invited to become PCAs will actually attend the training session because of a number of potential barriers (eg, transportation access, schedule conflicts, trust in research teams, etc). Formally, we model that each youth who is invited will actually attend with probability q (based on pilot studies, we took $q = 0.5$). Because of this variation in attendance and capacity limits for the initial training, PCAs are recruited over multiple rounds. At each round, CHANGE selects PCAs through an objective function that accounts for which PCAs were successfully recruited in previous rounds and the probability that prospective PCAs will attend in the current round.

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 1 and 3 months post-baseline. A flow diagram depicts the overall study design and participant retention for each arm (Fig. 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/nonconforming," or something else, were coded as transgender/nonbinary. 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 2 weeks. Participants were then categorized as living in (1) an emergency shelter or transitional living placement, (2) unsheltered, or (3) unstably housed (ie, "couch surfing" with friends or extended family).

Sexual Risk Behavior

Sexual risk behavior was operationalized using 2 variables: condomless anal sex (CAS) and condomless vaginal sex (CVS) acts. Participants were asked to list their 5 most recent sexual partners in the past month. For each partner, participants were asked whether they had (1) vaginal sex act without a condom, (2) anal sex act without a condom, or (3) both anal and vaginal sex acts without a condom during their last sexual encounter. A binary variable was created indicating whether a participant had CAS act (yes = 1; no = 0) or CVS act (yes = 1; no = 0) with at least one sexual partner in the past month. Participants who reported no condomless sex act or no sexual partners in the past month were coded as zero on these variables.

HIV Testing

HIV testing in the past 6 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 6 months were coded as 1; and participants who reported their last test occurring more than 6 months ago or who were never tested were coded as 0.

HIV Knowledge

Participants completed a brief, 6-question HIV knowledge quiz. True or false and multiple choice questions developed by our research team tested participants' knowledge of HIV transmission, HIV testing, and local, population-relevant statistics (eg, "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 the χ^2 test or analysis of variances 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 correlation matrix for all GEE models, given that differences in quasi-likelihood information criteria were negligible across different specifications. Bivariate logistic regression models indicated that participants who had greater odds of missing data 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, and 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 was 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

Descriptive statistics are summarized in Table 1. Table 2 summarizes 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 3 arms; namely, birth sex, LGBQ identity, living situation, and which drop-in center participants were recruited from (all $P<0.05$). Therefore, these and other characteristics were included as covariates in multivariable GEE models.

CAS

As summarized in Table 3, a significant group-by-time interaction was found for the AI group (OR = 0.67, 95% CI: 0.46 to 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 1-month follow-up in the AI group relative to the DC group (OR = 0.43, 95% CI, 0.19 to 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 in the DC group. Most of the improvements for AI occurred by the 1-month survey, whereas improvements in the DC group were not fully realized until month 3.

CVS

There was a marginally significant AI-by-time interaction (OR = 0.75, 95% CI, 0.55 to 1.03, $P = 0.08$). Post hoc analyses at discrete time points showed significantly greater reductions in CVS act from the 1-month to the 3-month follow-up in the AI group (relative to observation only) (OR = 0.39, 95% CI, 0.19 to 0.79, $P = 0.01$).

HIV Knowledge and Testing

The PCA-by-time interaction term was significant (b = 0.11, SE = 0.02, $P<0.001$), indicating that PCAs

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/nonbinary	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†	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)‡	96 (38.2)	251 (35.2)
Housing				
Shelter/transitional living	57 (22.5)	56 (26.8)‡	47 (18.7)	160 (22.4)
Unstably housed*	94 (37.2)	99 (47.4)‡	94 (37.5)	287 (40.2)
Street (unsheltered)*	102 (40.3)	54 (25.8)‡	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 $P < 0.05$.
†Other race includes Asian, Native, Pacific Islander, and "other" as a write-in category.
‡Data missing from first wave of DC (n = 54).

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

Variable	AI n (%)	AI Total N	DC n (%)	DC Total N	OBS n (%)	OBS Total N
CAS						
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
CVS						
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 6 mo)						
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)						
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

increased their HIV knowledge over time. Similarly, a PCA-by-time term was significant (OR = 1.82, 95% CI, 1.07 to 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 used a quasi-experimental design with 3 arms: an observation-only arm, a typical PCA selection method arm (ie, namely popularity–highest DC), and a PCA selection arm using the AI CHANGE algorithm. Results indicate that using the AI algorithm is an efficacious intervention approach. First, and most importantly, there was a significant reduction in CAS act 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 CAS over time compared with the observation-only arm. In 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 1-month to the 3-month follow-up in the AI group. Together, these findings suggest that the AI algorithm does a better job in selecting PCAs than DC (ie, popularity).

Second, improvements in the CHANGE group happened faster than in the DC group; most of the improvement for CHANGE occurred by the 1-month survey, whereas improvements in the DC group were not fully realized until month 3. 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 (ie, multiple popular people are connected with the same others). So information may eventually diffuse through the entire network under popularity but less efficiently compared with when CHANGE selects the agents. Fast results are important for 2 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 nonnegligible portion of youth will have left the center by the time a 3-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.

TABLE 3. Including all Outcomes in 1 Table

Outcome Variable	Predictor	b	SE	OR	95% CI	P
CAS						
	Time	—	—	1.05	0.82 to 1.33	0.72
	AI	—	—	1.44	0.89 to 2.32	0.14
	DC	—	—	1.49	0.88 to 2.52	0.14
	PCA	—	—	1.00	0.56 to 1.78	0.99
	AI × time	—	—	0.67	0.46 to 0.95	**0.03**
	DC × time	—	—	0.78	0.52 to 1.15	0.20
	PCA × time	—	—	1.17	0.73 to 1.89	0.51
CVS						
	Time	—	—	0.87	0.71 to 1.07	0.18
	AI	—	—	0.79	0.53 to 1.18	0.25
	DC	—	—	1.09	0.68 to 1.74	0.72
	PCA	—	—	0.86	0.49 to 1.50	0.59
	AI × time	—	—	0.75	0.55 to 1.03	*0.08*
	DC × time	—	—	0.85	0.60 to 1.22	0.38
	PCA × time	—	—	1.17	0.76 to 1.81	0.47
HIV testing						
	Time	—	—	1.34	1.05 to 1.71	**0.02**
	AI	—	—	1.21	0.77 to 1.91	0.40
	DC	—	—	0.95	0.58 to 1.56	0.84
	PCA	—	—	0.79	0.41 to 1.53	0.48
	AI × time	—	—	0.73	0.52 to 1.03	*0.07*
	DC × time	—	—	0.91	0.58 to 1.41	0.66
	PCA × time	—	—	1.82	1.07 to 3.09	**0.03**
HIV knowledge						
	Time	0.002	0.009	—	−0.02 to 0.02	0.79
	AI	−0.02	0.02	—	−0.05 to 0.02	0.34
	DC	−0.04	0.02	—	−0.08 to 0.01	*0.09*
	PCA	0.02	0.03	—	−0.03 to 0.07	0.38
	AI x time	0.02	0.01	—	−0.01 to 0.04	0.20
	DC × time	0.02	0.02	—	−0.01 to 0.05	0.27
	PCA × time	0.11	0.02	—	0.07 to 0.16	**<0.001**

Generalized estimating equations predicting study outcome variables, including PCA and PCA × time interaction.

Models adjusted for age, male birth sex, transgender, LGBQ identity, male X LGBQ, race, committed relationship, housing status, and drop-in center.

Bolded parameter estimates are significant at $P < 0.05$.

Italicized parameter estimates are significant at $P < 0.10$.

There are some limitations to this 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 came 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 et al[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 this 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 affected communities. One future challenge to be addressed is on 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, however, that some additional implementation research may be needed to most effectively facilitate the transfer of this technology to community settings.

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;350:1500–1505.

5. Latkin CA, Sherman S, Knowlton A. HIV prevention among drug users: outcome of a network-oriented peer outreach intervention. *Health Psychol*. 2003;22:332–339.
6. Schneider JA, Young L, Ramachandran 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: 1618–1624.
8. Rice E, Monro W, Barman-Adhikari A, et al. Internet use, social networking, and HIV/AIDS risk for homeless adolescents. *J Adolesc Health*. 2010;47:610–613.
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; 102:141–147.
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; 53:486–491.
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;54:204–214.
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;16:139–150.
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;34:881–896.
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 Res*. 2018;9:551–573.
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, Brazil.
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, LA.
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, Sweden.
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, Brazil.
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, DC.
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, TX.
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;7:34.
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 EM. *Diffusion of Innovations*. New York, NY: Simon & Schuster; 2010.
29. Freeman L, Webster CM. Interpersonal proximity in social and cognitive space. *Soc Cognit*. 1994;12:223–247.
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;50:358–364.
31. Rotheram-Borus MJ, Song J, Gwadz M, et al. Reductions in HIV risk among runaway youth. *Prev Sci*. 2003;4:173–187.
32. Bender KA, DePrince A, Begun S, et al. Enhancing risk detection among homeless youth: a randomized clinical rrial of a promising pilot intervention. *J Interpers Violence*. 2018;33:2945–2967.