



Acceptability of electronic healthcare predictive analytics for HIV prevention: a qualitative study with men who have sex with men in New York City

Jennifer J. Mootz^{1,2}, Henry Evans³, Jack Tocco^{1,2}, Christian Vivar Ramon⁴, Peter Gordon⁴, Milton L. Wainberg^{1,2}, Michael T. Yin⁴

¹Department of Psychiatry, Columbia University Medical Center, New York, NY, USA; ²New York State Psychiatric Institute, New York, NY, USA;

³Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA; ⁴Division of Infectious Diseases, Department of Internal Medicine, Columbia University Irving Medical Center, New York, NY, USA

Contributions: (I) Conception and design: MT Yin, JJ Mootz, H Evans, J Tocco, P Gordon; (II) Administrative support: CV Ramon, H Evans; (III) Provision of study materials or patients: CV Ramon, H Evans; (IV) Collection and assembly of data: JJ Mootz, H Evans, J Tocco, CV Ramon; (V) Data analysis and interpretation: JJ Mootz, J Tocco, H Evans; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Michael T. Yin, MD, MS. Division of Infectious Diseases, Columbia University Irving Medical Center, 630 w168th Street, PH8-876, New York, NY 10032, USA. Email: mty4@cumc.columbia.edu.

Background: Large data sets, also known as “big data”, shared in health information exchanges (HIEs), can be used in novel ways to advance health, including among communities at risk for HIV infection. We examined values and opinions about the acceptability of using electronic healthcare predictive analytics (eHPA) to promote HIV prevention in men who have sex with men (MSM). Our aims were twofold: (I) to evaluate the perspectives of MSM with diverse race/ethnicity and age on the acceptability of predictive analytics to determine individual HIV risk and (II) to determine acceptability of having targeted prevention messaging based upon those risk estimates sent directly to the consumer.

Method: Two of the authors facilitated 12 focus groups (n=57) with adult MSM without HIV, living in NYC. Groups were divided by ethnicity (Black, Latino, and White) and age (under 35 and 35 and over). Participants were recruited through HIV prevention sites, community-based organizations, social media, and Internet sites that serve these communities. Grounded theory methods were used to analyze the data with Dedoose.

Results: We identified six main themes related to acceptability: (I) reach, relevance, and potential uptake of using predictive analytics to establish HIV risk and deliver targeted prevention messaging; (II) patient-provider communication; (III) public health and individual rights; (IV) perceptions of intervention effectiveness; (V) electronic health data security; and (VI) stigma. Within each thematic domain, MSM discussed concerns, benefits, and provided recommendations for implementation.

Conclusions: MSM in this study were supportive of the use of “big data” and technology to reach marginalized populations and improve public health, yet expressed concerns about the relevance, effectiveness, and security eHPA. Efforts to advance eHPA for HIV prevention should address these concerns, especially among the most-at-risk communities of color. Development of eHPA for HIV prevention should involve targeted messaging that addresses specific concerns regarding eHPA security, accuracy, and relevance.

Keywords: Predictive analytics; HIV prevention; electronic medical records; men who have sex with men (MSM)

Received: 10 July 2019; Accepted: 18 October 2019; Published: 05 April 2020.

doi: 10.21037/mhealth.2019.10.03

View this article at: <http://dx.doi.org/10.21037/mhealth.2019.10.03>

Introduction

The advent of sophisticated predictive analytics techniques that can be applied to electronic health records (EHR) and health information exchanges (HIEs), has set the stage for potential breakthroughs in managing and preventing various medical conditions. Several concerns have limited broad utilization including the following: information security concerns for large aggregated data sets where data were obtained for unrelated purposes; privacy concerns involved with targeting individuals for unsolicited prevention outreach; and lack of standards for the use of informed consent to use EHR for predictive analytics (1,2). However, the potential for individual and societal benefit by identifying critical health prevention opportunities is significant. Despite mounting concern over well-documented, unauthorized uses of collected, aggregated, and analyzed data, the general public's sharing of personal data has never been greater. Empirical data of patient or consumer attitudes are therefore necessary to help inform the evolving ethical framework around electronic healthcare predictive analytics (eHPA).

In 2014, New York State launched an ambitious campaign to end the HIV epidemic by 2020, the so-called "End the Epidemic" (EtE) initiative. The EtE initiative focuses upon: (I) identifying previously undiagnosed individuals and linking them to care; (II) keeping people with HIV linked to care in order to maintain virologic suppression, and (III) expanding the availability of preventive modalities such as pre-exposure prophylaxis (PrEP) to people at increased risk of infection (3). PrEP has been shown to significantly reduce the risk of HIV acquisition when taken consistently (4,5). Additionally, analyses have shown that PrEP could be a cost-effective prevention strategy when coupled with risk stratification and specific targeting of interventions toward most-at-risk populations (6,7). This highlights the need to both identify individuals at risk of HIV infection and provide them the information and resources to access PrEP and other preventive modalities.

Among those affected, men who have sex with men (MSM), and particularly MSM of color, carry a disproportionately large burden of new HIV infections. For example, in 2016, MSM accounted for 67% of the 40,324 new diagnoses in the United States, and 83% of those among males (8). Almost 40% of MSM affected were Black/African American (8). Multiple barriers have existed in reaching MSM, especially Black/African American MSM, with health interventions. Mistrust of medical

establishments and stigma surrounding sexual orientation are frequently cited barriers for Black/African American MSM, and fear of stigma in particular has been associated with longer lapses in either general healthcare appointments or appointments for treatment (9,10). Understanding the influence that perceived stigma and discrimination have on patients' interactions with health services is an important part of developing effective population-level HIV prevention strategies.

Recent publications on the test characteristics of different predictive analytic models applied to EHR have demonstrated the potential for this tool to help identify individuals at-risk of HIV infection (11,12). However, in addition to its efficacy, the acceptability of a particular health intervention among affected communities is an important consideration to its implementation (13). We defined acceptability among consumers as a "multi-faceted construct that reflects the extent to which people receiving a healthcare intervention consider it to be appropriate, based on anticipated cognitive and emotional responses to the intervention" (13). Our aims were twofold: (I) to evaluate the perspectives of MSM with diverse race/ethnicity and age on the acceptability of predictive analytics to determine individual HIV risk and (II) to determine acceptability of having targeted prevention messaging based upon those risk estimates sent directly to the consumer. We wanted to understand respondents' perspectives on the use of patient/consumer-facing applications—through the use of texts, smartphone applications, or linkage to online healthcare portals—to communicate a risk score and recommendations for linkage to care for PrEP, if the predicted individual risk was above a predetermined threshold.

Methods

Following ethical approval from the Institutional Review Board at the New York State Psychiatric Institute, we conducted 12 focus group discussions lasting approximately 90 minutes each with between four and seven gay and bisexual men ($n=57$ in New York City, disaggregated by age (18–34 and over 35) and ethnicity (Black, Latino, and White). Participants were recruited through advertisements posted at social media and gay dating sites/apps, lesbian gay bisexual transgender (LGBT) community organizations, health clinics, LGBT Pride events, and HIV research centers. To determine eligibility for the study, interested respondents either completed an online questionnaire through Qualtrics or contacted a member of the research

team who then determined eligibility over the phone. Once eligibility was established and enough participants qualified to hold a specific group, a member of the research team called respondents to confirm the date and time and provide directions to the focus group.

After obtaining informed consent and administering a brief demographic survey, and verbally reiterating expectations of confidentiality and mutual respect, two authors co-facilitated the focus groups by first presenting two short informational videos made specifically for this research that provided background information on PrEP, EHR, HIE's, and examples of predictive analytics available to consumers (e.g., shopping and movie recommendations). Facilitators then led semi-structured discussions about eHPA, HIV, and prevention, focusing on participants' prior knowledge, preferences, concerns, perception of benefits, and how they felt that cultural factors informed their perspectives. The discussions were audio recorded. As a form of analyst triangulation (14), prior to the conclusion, a facilitator provided a verbal summary of the group's discussion and invited the group members to revise or add to the summary. Pizza and nonalcoholic beverages were provided during the groups and each participant was given USD \$50 upon conclusion to cover transportation costs and in appreciation for their time.

Data analysis

Demographic survey data were entered into a spreadsheet and summarized using SPSS Version 25. All focus group discussions were transcribed by a professional transcription service and transcripts were uploaded to Dedoose, an online qualitative analysis software. Using grounded theory, two authors developed a thematic codebook by coding a transcript, comparing codes, and discussing discrepancies until reaching consensus. Data were coded line-by-line; the analysts assigned codes to each unit of meaning (open coding). These codes were then organized into themes, such as stigma and benefits (axial coding). The codebook organized into themes was uploaded into Dedoose, and three research team members coded another transcript in Dedoose together to improve interrater reliability, resolve coding disagreements through discussion, and iteratively update the codebook. The three coders then divided and coded the remaining transcripts. The analysis team met regularly to discuss coding questions, concerns, and observations. The team concluded by identifying overarching themes (selective coding) (15), a process that

involved analyzing patterns and frequency of co-occurring codes in Dedoose and systematically re-examining all quotations related to the construct of acceptability (e.g., subthemes of benefits, concerns, and recommendations).

Results

Demographics

All 57 participants (Black =21; White/Non-Latino =16; Latino =15; Other =5) identified as men (*Table 1*). Their ages ranged from 21 to 61 (median 37.5). Most men

Table 1 Demographics of focus group participants

Characteristics	n [%]
Age, years	
18 to 34	28 [49]
≥35	29 [51]
Race/ethnicity	
White	16 [28]
Hispanic or Latino	15 [26]
Black or African American	21 [37]
Other race	5 [9]
Sexual orientation (Kinsey Scale)	
0 straight	0 [0]
1	0 [0]
2	0 [0]
3 bisexual	9 [16]
4	4 [7]
5	6 [11]
6 gay/homosexual	38 [67]
Education	
Less than high school	1 [2]
Some high school	1 [2]
High school graduate (or GED)	3 [5]
Vocational/technical/trade school	0 [0]
Some college or associate's degree	21 [37]
Bachelor's degree	19 [33]
Master's or advanced degree	12 [21]

Table 1 (continued)

Table 1 (continued)

Characteristics	n [%]
Household income	
Less than \$15,000	7 [12]
\$15,000 to \$29,999	13 [23]
\$30,000 to \$59,999	17 [30]
\$60,000 to \$79,999	11 [19]
\$80,000 to \$100,000	3 [5]
More than \$100,000	6 [11]
Lifetime male sexual partners	
0	9 [16]
1 to 9	12 [21]
10 to 24	15 [26]
25 to 49	18 [32]
50+	1 [2]
Prefer not to answer	2 [4]
Seen provider in the past 12 months	
Yes	55 [96]
No	1 [2]
Don't know/prefer not to answer	1 [2]
HIV status	
Negative	53 [93]
Positive	1 [2]
Don't know/prefer not to answer	3 [5]

GED, general education diploma, or high school equivalency certificate.

(approximately 70%) were single. On a sexual orientation scale of 0 (straight/heterosexual) to 6 (gay/homosexual), participants' average score was 5.3. Most participants (91.2%) had some college or more education. Participants' household incomes ranged from less than \$15,000 annually (12%) to more than \$100,000 annually (11%). Most participants (72%), however, earned \$15,000–\$79,999 with roughly one-third of respondents reporting an annual income level ranging from \$30,000 to \$60,000.

Overall findings

Six themes of acceptability emerged from the focus group discussions: (I) reach, relevance, and potential uptake;

(II) patient-provider communication; (III) public health and individual rights; (IV) perceptions of intervention effectiveness; (V) electronic health data security; and (VI) stigma (Table 2). While participants across all demographic groups tended to focus discussion on concerns about predictive analytics, they also identified benefits when prompted and proposed novel recommendations for implementation.

Reach, relevance, and potential uptake

Participants often framed potential benefits as being relevant for others rather than themselves, a juxtaposition that rested on participants' widely shared beliefs that they were relatively well informed about their own risk of acquiring HIV. Respondents described people who could benefit from predictive analytics as marginalized or as having other psychosocial challenges related to demographic factors. Examples were those who are unaware, closeted, unconfident, dishonest with themselves, lack access to medical care, live in dangerous places, homeless, uneducated, perform transactional sex to survive, and without health insurance, among others.

“But it could be beneficial for other people, so it's kind of like going back to what I do—it's basically kind of goes both ways. It's like sometimes, you don't have people that know that you have these services, or you do this, or you X, Y, and Z, but have certain preventative measures.”—Younger Latino Man.

One group noted that predictive analytics have the potential of reaching LGBTQ populations that lack tailored sexual education training due to local politics. These responses could be collectively viewed as “altruistic” responses, highlighting the way that predictive analytics could help others in a less privileged situation:

“This could be really helpful in identifying, you know, like populations who don't have comprehensive sexual education, which is, I think, way more tricky to implement. Like as broadly as like, the analytics, because like there's so much local politics involved in like, what kinds of education children get. So it's like, if this helps a doctor...identify like, a kid who might, who like could use a specific like, queer-focused sexual education meeting, when they're meeting one on one with a doctor, like that would be great.”—Younger White Man.

While many focus groups discussed the potential of predictive analytics to reach at-risk populations, there were apprehensions that tempered these conversations. Questions about the logistics of how people without access to healthcare (i.e., not contributing data points for predictive

Table 2 Perceived benefits, concerns, and delivery recommendations of using eHPA for HIV prevention in MSM

Themes	Perceived benefits	Concerns	Delivery recommendations
Reach, relevance, and potential uptake	Relevant for people who have different demographics than participants; pertinent for people who lack access to healthcare; Transcends local politics of sexual education for LGBT populations	People and physicians already aware of risk status; will not reach most-at-risk; most-at-risk will not be receptive; consent issues for minors	Tailor messaging to different demographics (e.g., youth); ensure intervention carries no cost; develop in multiple languages; keep information simple; community education also needed (e.g., in schools); account for diversity in gender identity; use a public health approach with billboards and commercials
Patient-provider communication	Makes information sharing more efficient; promotes patient-provider communication	Loss of personal touch	Ensure there's messaging about and availability of local credible providers; educate providers about PREP and eHPA; risk status should be conveyed by a provider; warmth an important provider quality; previous patient-provider rapport helpful
Public health and individual rights	Strengthens HIV awareness efforts; facilitates community mobilization and collaboration	Infringes on autonomy; concerns of surveillance	Notifications should be general (e.g., "call your doctor") and not delivered via text or email; customize messaging to counteract problems with trust
Perceptions of intervention effectiveness	Encourages less risky behaviors	Doubts about predictive accuracy of analytics; strategy is ineffective/ incites riskier behaviors	Be transparent about pros and cons of PREP; create ways to continually update health and behavior status; include risk information on all health to allow for better decision-making; broaden scope beyond prevention to account for changes in HIV status; make predictive analytics more precise than demographics (e.g., where someone lives); adjust notification frequency to match risk status; communication must stand apart from frequency of other technological communication
Electronic health data security	–	Digital data are subject to security breaches; Corporate and individual interests may conflict; concerns about policy shifts related to healthcare, HIV status, and sexual orientation	Consent for eHPA should be considered, and preferred information delivery method could be provided during consent process
Stigma	Can potentially reduce stigma	Conveying risk status seems judgmental; targeting MSM gives message that others not at risk	Have messaging encompass broader health issues and not focus solely on HIV prevention; use the intervention to change the narrative related to HIV; also adapt and offer intervention for heterosexual people

eHPA, electronic healthcare predictive analytics; MSM, men who have sex with men; LGBT, lesbian gay bisexual transgender.

analytics to run) could receive accurate estimations of risk were a frequent topic of discussion.

"How does me sharing all this data help some teenager who does not have sex education get a text from a community health center that all of the teenagers in the area like, are signed up for, to say, hey, you know, they get a text when they turn like 14...

you should come in, so we can chat. And like, you can learn about stuff. Like that, it seems to be like what the point is."—Younger White Man.

Group members supposed that MSM who were accessing healthcare services and thereby contributing more data points for predictive analytics to consider already knew their

risk, qualifying the intervention's applicability. One group member stated, *"I feel—sorry. I just feel like people know if they're healthy or not."* Consent and assent complications for underage youth, a demographic identified as having higher need of predictive analytics, could also decrease reach.

Participants, however, offered several recommendations to offset these limitations. For relevance, respondents stressed the importance of tailoring the messaging to be meaningful for a range of people who have diverse demographic and cultural identities. Adapting messaging to be youth-friendly was a common recommendation, for instance.

"I feel like it also can be—the messaging can be tailored in a way that resonates more with the demographic that's using it. So, I'm now only conceptualizing it as like an app, but it could many things, but like, I don't know, it doesn't seem like it's dated for like, my grandma, I might relate to the information a little bit more."—Younger White Man.

Consideration of cost, language, and gender identity were other suggestions to extend reach to those with demographic vulnerabilities. Many respondents observed that predictive analytics alone would limit reach and should be paired with public health strategies, such as building community awareness and engagement.

Patient-provider communication

Respondents recognized that using technological interventions has the potential to facilitate communication between patients and providers. For example, communication can be made increasingly efficient. A younger Latino man shared, *"It's just pretty much understanding more or less how providers—it's easier for providers to communicate with each other."* Another respondent observed and group members concurred that predictive analytics could help providers use time during appointments more efficiently.

"I can see this as a tool for clinicians, because sometimes clinicians—you go see a clinician who has 10 minutes. And the clinician does not know what to do, and if he had this particular tool, he'd be able...he'd be better able to assist you."—Older Black Man.

With conversation about enhanced efficiency of conveying risk about HIV acquisition came suggestions about increasing training for providers about HIV risk prevention possibilities, such as PrEP. Establishing a local healthcare workforce that is knowledgeable about PrEP and other current guidelines for HIV prevention as well as

being LGBT friendly was paramount. Many respondents talked about attending healthcare clinics that were explicitly for LGBT populations, observing how availability of specialized clinics was a privilege of residing in a large urban city like NYC.

Despite the possibility that predictive analytics could facilitate patient-provider communication, an overarching concern was the presumed ways in which the use of technology could obstruct human interaction. Participants worried that summarizing patients' health statuses with statistics could be dehumanizing.

"You're a statistic. I mean, it just is what it is. You went from being (name) with the following symptoms to, you know, male, White, 40-49, you know, average weight, you know, average height, et cetera."—Older White Man.

Thus, corresponding recommendations implicitly underscored desire for and the importance of human connection when conveying HIV risk. In these narratives, group members often described technology as cold and human interaction as essential and fundamental.

"I feel that it's—like leading the information, it's a little bit more cold, more harsh. There's nothing behind it than being told by someone. Like, if I personally prefer hearing it from someone, 'Your risks are this, X, Y, and Z', but it's just the way it's said, it could be a little bit—it's more relatable..."—Younger Latino Man.

Respondents overwhelmingly expressed the desire for the intervention to be paired with a human provider, notwithstanding the potential benefits of having an objective, potentially more accurate (albeit "cold") intervention.

"Whereas humans, I think something about our spirit, something about who we are, having another person there being able to be that transparent, being that vulnerable, I think that makes it that much easier. I don't want a text message, it's impersonal. I don't want an email."—Older Black Man.

Usually group members identified the notifying professional as the primary care provider, but others were open to other qualified health, mental health, or social work providers who have a professional role in delivering information about HIV risk. Irrespective of professional title and training, what seemed to matter most to participants was personal characteristics of the provider, such as warmth, and the quality of the relationship between the intervention recipient and provider. When discussing preferences about notification, a participant shared the following:

"The person I feel closest to. It could be my social worker. It could be my doctor. It could be whoever. All of those (three?)"

parties, if I'm closest to—and, you know, who—the one—I would feel more comfortable with the closest person I'm closest to. Basically, that's it.—Older Black Man.

Feeling safe, accepted, and having an established relationship with a licensed and regulated health provider was important to participants and linked to the perception that receiving information about level of HIV risk (usually presumed to be high) would be distressing.

Public health and individual rights

Another commonly observed benefit of predictive analytics was its potential to improve public health through increasing awareness of HIV prevention. However, concerns about privacy and health data surveillance attenuated this greater good benefit. A respondent documented the tension between public health and individual privacy, while connecting this discourse to the broader issue of discrimination based on health status.

"It seems to me that it's a really good idea for everybody's healthcare records to be put in a common database and analyzed so that all kinds of things can be done on behalf of large groups of people, as opposed to each individual. But running counter to that is this notion that there are some individuals, many individuals, who would prefer not to participate in that because they risk having higher insurance rates if there's a data breach. Discrimination, that kind of thing."—Older White Man.

Most recommendations that referenced this theme were about protecting privacy. Notifications, for instance, should be generic and not convey individual health information. They instead would ideally be invitations to log in to secure and self-authenticating health portals. Messages related to predictive analytics also should not be delivered via text to a person's phone or via email to a computer screen because others might read these types of notifications. A participant explained the following:

"But text messages always show up—the notification. Or it shows a symbol that you have a text message. So for me, text, if someone just accidentally grabs your phone and goes through it, they could see the text."—Younger Latino Man.

Perceptions of intervention effectiveness

Respondents supposed that notifications stemming from predictive analytics could jar people into understanding the real possibility of acquiring HIV. This jarring was akin to providing a "wake-up call" to encourage behavior

change and PrEP uptake among most-at-risk individuals in particular.

"But it's great that they can have accountability of saying if they want to change, they're unhealthy or something. Because that will wake them up. Sometimes you need a wake-up call."—Younger Black Man.

The benefits of wake-up calls intertwined with narrative about stigmatization such that predictive analytics could reach beyond cultural barriers.

"Like having these types of systems, yes, it's good because it shows statistic-wise what's actually happening. (It) is like a wake-up call, and sometimes people do need the wake-up call, because you have cultural barriers. You have stigmas."—Younger Latino Man

The apprehensions about intervention effectiveness centered around whether a predictive analytics algorithm could accurately assess HIV risk levels. A younger Black man shared, *"It could just be running an algorithm off of incorrect information, and it wouldn't be very predictive for anyone."* These trepidations cited imperfections in technology, lack of (human) oversight of algorithms, dependence on transparent self-report of sexual risk behaviors, and population level data erroneously characterizing individual risk.

"Because you can live in a high-risk area, but don't, like you're not attracted to any of the people in your area, so you only sleep with the people outside. But because you live there, now you're on this prediction, so that still doesn't add up."—Younger Black Man.

Oftentimes, as is illustrated in the quotation below, participants confused conveying risk level for acquiring HIV with being informed of a positive HIV status. This misperception seemed to amplify apprehensions about intervention effectiveness by raising the stakes in terms of information that would be communicated about risk.

"But then in practice, machines aren't perfect...they're not that perfect. What happens if you're the one person that it tells you, it misread the information. Somebody pushed a button wrong and it tells you that you're dying of HIV and you're not even close. That changes your whole world until that mistake is rectified."—Older Latino Man.

To counteract gaps in self-report and lack of access to healthcare that could hinder predictive accuracy, principal suggestions were to build in opportunities for MSM to ensure their individual level data are current.

"Also you would need to update the information that this is running, because it could be running...I don't know how often anyone else in here goes to the doctor, but I don't really go to the doctor unless I'm like dying."—Younger Black Man.

Furthermore, respondents recognized that in order for

predictive analytics to be an effective wake-up call, the notifications must stand apart from the virtual noise created by the inundation of technological communications.

“Because I know I have the data somewhere in this device, I just don’t know how to find it. It’s problematic for me to receive push data, because I ignore so much of it.”—Older White Man.

One group discussion noted that a possible strategy to address the problem highlighted in the above illustration was to match frequency and intensity of notifications to risk level.

Electronic health data security

Security of health data was a prominent topic and focal point of discussion across all groups. While one member in a group of older White men recognized that predictive analytics could strengthen security of data by removing human error, most focus groups talked at length about concerns. In addition to divulging personal experiences with identity theft, these discussions frequently evoked current events (e.g., hackers’ breach of Equifax) where private data were illegally accessed and leveraged for profit.

“But I think health information is less on my mind than all my information now which is probably out there for whoever. So, I think it’s not a question of if you’re going to be hacked, but when they’re going to hack.”—Older White Man.

Particular worries were about how industries could leverage individual healthcare data for profit. Respondents conjectured that employers, for instance, could use information garnered through predictive analytics to enact discriminative hiring or workplace practices. Others noted how predictive analytics for HIV prevention crosscuts several contemporary policy level issues, such as health insurance, LGBT rights, and health status.

“Say I go sign up for life insurance, okay, and they go in there and find out that like I had a huge drinking problem years ago that was in like my doctor’s medical records and stuff. I mean, I wouldn’t tell the life insurance company before I applied for it, ‘You know, I was a total drunk at one point’. But if it’s my medical records and they have access to it, then you know, you’re screwed, basically. You know, be like if you were HIV positive years ago and try and get something. Denied. It’s preexisting condition. They’ll find ways to deny you or jack your rates up.”—Older White Man.

Given the multiple vulnerabilities associated with having a marginalized sexual orientation and historical HIV stigma, data breaches were perceived by focus group members as particularly threatening. Participants gave examples from

current events to illustrate that present policy assurances were tenuous and not guaranteed to be permanent. Protections established under an administration could be retracted when political agendas shift.

“I would be deathly afraid for my information to get out now. I mean, it’s political, but it isn’t. The president that we have right now, just look at what’s happening with healthcare. Look at the people that right now, how they’re battling policies in healthcare, and it’s not about HIV. It’s about simple things that we might all be predisposed to. And so, think about all the people that are dealing with that right now. Imagine adding HIV/AIDS to it.”—Older White Man.

Although participants connected concerns about security of data to policy level issues, their recommendations were at the individual level and involved obtaining written informed consent routinely and at multiple junctures.

Stigma

A benefit of a technological intervention like predictive analytics, according to participants, is that its impartiality obviates stigmatizing messages. This neutrality could especially benefit MSM who struggle with their sexual identity. By framing technology as less judgmental and “nagging,” the narrative below implicitly acknowledges how the use of technology can reduce the harm of human-generated stigma.

“But maybe there, for someone who’s maybe not as self-aware, not as confident in their ability to, you know, know what decisions they should be making or not, it would be extremely beneficial, or even to a closeted person, or someone that is having some difficulty in being honest with themselves about it, when they see this cold, hard information, and it’s coming from a source that’s non-judgmental, and really it’s not like another human being saying, you know, you got to do this or that. It’s not nagging.”—Younger White Man.

Focus group members recommended that broadening predictive analytics to include other health conditions and develop the intervention to reach all populations could help reduce stigma.

“You know, we’ve run a couple of stuff for you, and we think that this might be best for you. You know, the cholesterol that’s over there. Period. You had a jump. Call it a jump. You know, we would also put that into it [the notification]. Things like that. Make it more friendly.”—Younger Black Man.

To avoid sending a message that HIV is an illness that only MSM can acquire, some respondents suggested that the messaging broaden to include everyone and that

predictive analytics should be developed for all populations.

“It is a stereotype, and it’s unfortunate that in 2017, that it is still a stereotype, and we can’t get the proper channels out because people only see it as, it’s only within this (MSM) community that people get sick and get HIV. No. It’s anyone who has sexual intercourse with another individual is capable without protection to come up with HIV.”—Younger Black Man.

Discussion

Overall, MSM’s perceptions of the acceptability of using predictive analytics to determine HIV risk contained six themes about reach, communication, individual rights, effectiveness, data security, and stigma. In Chaudoir’s 2013 systematic review measuring factors affecting implementation of health innovations (16), a comprehensive multi-level framework posited five factors representing structural-, organizational-, patient-, provider-, and innovation-levels of analysis as a means of organizing and describing important sets of constructs that predict implementation outcomes. Our paper addresses patient-related factors to be considered to enhance implementation of eHPA for HIV prevention. To our knowledge, this is one of the first studies to examine the multi-faceted construct of eHPA acceptability among MSM as well as provide consumer-informed recommendations to enhance acceptability.

Respondents across all focus groups expressed concerns about digital security and potential negative consequences of being labelled at risk, especially with commercial healthcare insurers and in the current political climate. These concerns often dominated discussions and seemed to have the strongest link to perceptions of acceptability. Given that acceptability is considered a leading factor for implementation success (17), addressing these concerns at the policy (e.g., healthcare and LGBT rights protections) and individual (e.g., offering informed consent) levels is crucial.

However, participants identified potential advantages, such as reaching at-risk individuals without access to healthcare or those who feel stigmatized by healthcare providers. They saw the potential to overcome some of the limitations of sexual health education for marginalized populations. Another recent study among MSM and providers found similar results about mixed perceptions of intervention effectiveness (18). Topics of Gilkey *et al.*’s study focused on quantification and interpretation of risk and the perception of the capability of a predictive analytic

intervention to effect behavior change.

One unexpected finding was the majority’s desire for a healthcare provider to communicate the results of the predictive analytic and provide recommendations for intervention. Although participants recognized that involving provider interaction limited the potential reach of this information, since many of those at risk will not have regular contact with the healthcare system, many felt that it was important to have the “warmth” of a personal exchange with a healthcare professional. While the patient-provider interaction may be a source of potential external, internal, or anticipated stigma, feedback from many participants suggested that a provider-facing predictive analytic for HIV risk assessment is preferable compared to individual access via a patient portal or personal health record. It was also felt likely that if predictive analytic models were found to have high accuracy, there would be broader acceptance of a direct consumer-facing interface.

Whether informed consent is expected either by providers or patients/consumers for use of e-HPA for risk factor assessment and prevention messaging was not directly addressed within our focus groups. From an ethics perspective, there are no widely accepted standards regarding the specific use of informed consent with eHPA for individual patient use. Patient expectations and limits to the reasonable use of informed consent with e-HPA have not yet been sufficiently articulated. Some have argued for patient consent when eHPA implementation would expose patients to significant risks, but further boundaries on appropriate eHPA use have not been studied (1). A number of recent studies have begun to examine this tension in greater detail and have found evidence for both increasing patient acceptance for the unsolicited use of eHPA to promote health as well as ongoing concern for the potential for abuse or loss of privacy (19–22). Notably, there is a high level of consumer support for the unsolicited application of eHPA for genetic studies using stored biobank specimens and consumer notification of potentially important health information (23). Similarly, an HIV linkage to care project conducted in Louisiana found high levels of acceptability from the target provider and patient population for eHPA applied to claims data and the state HIV database to facilitate re-engagement for individuals who had fallen out of care (20). Other studies evaluating the potential of eHPA applied to large data repositories have also noted broad, but not universal, consumer interest for potential health benefits in prevention of cardiovascular disease, diabetes, cancer and genetically-linked diseases (24–26).

Given the stigma associated with HIV disease or disease risk, it is possible that HIV-related stigma might impact participants' understanding and acceptance of employing predictive analytics to predict HIV risk. A previous multicenter mixed methods study of acceptability of using HIEs in general to enhance care of HIV-positive patients found broad support despite potential concerns related to disease stigma (27). Another study found acceptability among a small group of HIV-positive patients with substance use disorders for the use of mobile technology coupled with peer health support to supplement HIV care (28). However, the use of predictive analytics to make individual risk predictions was not addressed in either case. Our participants expressed concern about the potentially stigmatizing impact of focusing upon one particular group (MSM) for predictive analytics and messaging concerning HIV risk. Suggestions included providing predictive analytics to different risk groups, wider demographics, and incorporation of other medical conditions to normalize the predictive analytic process. They also suggested using messaging that is part of a broader public health approach and tailored to particular demographics.

Several limitations of this study should be considered. First, study participants were recruited solely from the New York City greater metropolitan area, which limits the extension of our results to non-urban settings and other metropolitan areas in the US where predictive analytics could play an important role in HIV prevention. Additionally, participants were largely recruited through online advertising and as part of pre-existing HIV research groups, which may have selected for those who were more familiar with both health information technology and HIV prevention methods than the general population. Most participants in this study had at least some college education. The sociodemographics of the sample along with the recruitment method may have inflated perceptions of acceptability in comparison to populations who have less education, are from rural settings, and have less exposure to HIV prevention messaging and research. Finally, focus groups were separated by both age and race/ethnicity in order to encourage open dialogue and uncover differences based on demographics. Many participants felt that leading integrated groups could introduce new discussion points that may not have arose in the separated focus groups.

Conclusions

In conclusion, this is one of the first qualitative

investigations of a multi-faceted construct of acceptability for predictive analytics for HIV prevention, conducted in focus groups of MSM in New York City. Concerns about digital security and the stigmatization of being labeled "at risk" were common, as was the preference for a healthcare provider over anonymous electronic communication of risk determined by the predictive analytic model. Overall, the participants were supportive of such an approach as part of a public health effort. Development of a predictive analytic for HIV prevention should be applicable to a wide demographic with tailored messaging that is acceptable to diverse populations and be part of a larger public health effort in HIV prevention.

Acknowledgments

Funding: This research was supported by an HIV Center Pilot grant and an NIMH center grant (P30-MH43520; PI: Robert H. Remien, PhD).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the Institutional Review Board at the New York State Psychiatric Institute, and written informed consent was obtained from all patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Amarasingham R, Patzer RE, Huesch M, et al. Implementing Electronic Health Care Predictive

- Analytics: Considerations And Challenges. *Health Aff (Millwood)* 2014;33:1148-54.
2. Cohen IG, Amarasingham R, Shah A, et al. The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care. *Health Aff (Millwood)* 2014;33:1139-47.
 3. New York State Department of Health. Ending the Epidemic Progress Report, 2018. Available online: https://www.health.ny.gov/diseases/aids/ending_the_epidemic/docs/ete_summary_2018.pdf
 4. Grant RM, Lama JR, Anderson PL, et al. Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *N Engl J Med* 2010;363:2587-99.
 5. McCormack S, Dunn DT, Desai M, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. *Lancet* 2016;387:53-60.
 6. Drabo EF, Hay JW, Vardavas R, et al. A Cost-effectiveness Analysis of Preexposure Prophylaxis for the Prevention of HIV Among Los Angeles County Men Who Have Sex With Men. *Clin Infect Dis* 2016;63:1495-504.
 7. Juusola JL, Brandeau ML, Owens DK, et al. The cost-effectiveness of preexposure prophylaxis for HIV prevention in the United States in men who have sex with men. *Ann Intern Med* 2012;156:541-50.
 8. Centers for Disease Control and Prevention. HIV Surveillance Report, 2016. Available online: <https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html>
 9. Eaton LA, Driffin DD, Kegler C, et al. The role of stigma and medical mistrust in the routine health care engagement of black men who have sex with men. *Am J Public Health* 2015;105:e75-82.
 10. Thompson HM. Patient Perspectives on Gender Identity Data Collection in Electronic Health Records: An Analysis of Disclosure, Privacy, and Access to Care. *Transgend Health* 2016;1:205-15.
 11. Feller DJ, Zucker J, Yin MT, et al. Using Clinical Notes and Natural Language Processing for Automated HIV Risk Assessment. *JAIDS J Acquir Immune Defic Syndr* 2018;77:160-6.
 12. Krakower D, Gruber S, Menchaca JT, et al. Automated Identification of Potential Candidates for Human Immunodeficiency Virus Pre-exposure Prophylaxis Using Electronic Health Record Data. *Open Forum Infect Dis* 2016;3:860.
 13. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. Available online: <https://bmchealthservres.biomedcentral.com/track/pdf/10.1186/s12913-017-2031-8>
 14. Patton MQ. Qualitative evaluation and research methods. 4th ed. Newbury Park, CA: Sage Publications, 2015.
 15. Strauss A, Corbin J. Basics of qualitative research: Grounded theory procedures and techniques. London: Sage Publications, 1990.
 16. Chaudoir SR, Dugan AG, Barr CHI. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. *Implement Sci* 2013;8:22.
 17. Proctor E, Silmere H, Raghavan R, et al. Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda. *Adm Policy Ment Health* 2011;38:65-76.
 18. Gilkey MB, Marcus JL, Garrell JM, et al. Using HIV Risk Prediction Tools to Identify Candidates for Pre-exposure Prophylaxis: Perspectives from Patients and Primary Care Providers. *AIDS Patient Care STDS* 2019;33:372-8.
 19. Ahadzadeh AS, Pahlevan Sharif S, Ong FS, et al. Integrating Health Belief Model and Technology Acceptance Model: An Investigation of Health-Related Internet Use. *J Med Internet Res* 2015;17:e45.
 20. Bennett GG, Glasgow RE. The Delivery of Public Health Interventions via the Internet: Actualizing Their Potential. *Annu Rev Public Health* 2009;30:273-92.
 21. Gibbons MC. Personal health and consumer informatics. The impact of health oriented social media applications on health outcomes. *Yearb Med Inform* 2013;8:159-61.
 22. Smith RJ, Grande D, Merchant RM. Transforming Scientific Inquiry. *Acad Med* 2016;91:469-72.
 23. Lemke AA, Wolf WA, Hebert-Beirne J, et al. Public and Biobank Participant Attitudes toward Genetic Research Participation and Data Sharing. *Public Health Genomics* 2010;13:368-77.
 24. Glauber H, Karnieli E. Preventing Type 2 Diabetes Mellitus: A Call for Personalized Intervention. *Perm J* 2013;17:74-9.
 25. Sieh W, Rothstein JH, McGuire V, et al. The Role of Genome Sequencing in Personalized Breast Cancer Prevention. *Cancer Epidemiol Biomarkers Prev* 2014;23:2322-7.
 26. Shmueli G, Koppius OR. Predictive Analytics in Information Systems Research. *MIS Q* 2011;35:553.
 27. Maiorana A, Steward WT, Koester KA, et al. Trust, confidentiality, and the acceptability of sharing HIV-related patient data: lessons learned from a mixed methods

- study about Health Information Exchanges. *Implement Sci* 2012;7:34.
28. Westergaard RP, Genz A, Panico K, Surkan PJ, Keruly J, Hutton HE, et al. Acceptability of a mobile health

intervention to enhance HIV care coordination for patients with substance use disorders. *Addict Sci Clin Pract* 2017;12:11.

doi: 10.21037/mhealth.2019.10.03

Cite this article as: Mootz JJ, Evans H, Tocco J, Ramon CV, Gordon P, Wainberg ML, Yin MT. Acceptability of electronic healthcare predictive analytics for HIV prevention: a qualitative study with men who have sex with men in New York City. *mHealth* 2020;6:11.