



HHS Public Access

Author manuscript

Ethics Behav. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Ethics Behav. 2020 ; 30(7): 481–495. doi:10.1080/10508422.2019.1692663.

“They know what they are getting into:” Researchers confront the benefits and challenges of online recruitment for HIV research

Elise Bragard^a, Celia B. Fisher^b, Brenda L. Curtis^c

^aDepartment of Psychology, Fordham University, Bronx, NY

^bCenter for Ethics Education and Department of Psychology, Fordham University, Bronx, NY

^cTechnology and Translational Research Unit, Intramural Research Program, National Institute on Drug Abuse, Baltimore, MD

Abstract

Online research has become a critical recruitment modality for understanding and reducing health disparities among hidden populations most at risk for HIV infection. There is a lack of consensus and guidelines for the responsible conduct of online recruitment for HIV risk populations. Using semi-structured phone interviews, this study drew on the experiences of principal investigators (PIs) engaged in online HIV research to illuminate scientific and ethical benefits and challenges of social media recruitment. Using Thematic Analysis five major themes emerged: sampling advantages and disadvantages; challenges of data integrity; control of privacy protections; researcher competence and responsibility; and resources.

Keywords

HIV; online recruitment; social media; ethics Internet research; privacy

Introduction

Ninety percent of US adults use the Internet, with seventy percent using social media sites (Anderson, Perrin, Jiang, & Kumar, 2019). Types of Internet use has grown and diversified, and technological advances in social networking have changed the way different populations can connect (Rendina & Mustanski, 2018). The rapid growth of online advertising opportunities over the past twenty years, including sponsored searches (Fain & Pedersen, 2006), contextual ads (Martín-Santana & Beerli-Palacio, 2012), and behavioral advertising (Ur, Leon, Cranor, Shay, & Wang, 2012), has been accompanied by a parallel increase in the use of online paid posts for research recruitment (Gelinas, Pierce, & Cohen, 2017). Online recruitment has been especially effective for the research recruitment targeting socially stigmatized and difficult to locate diverse samples of “hidden populations” for HIV epidemiological, prevention and intervention research (Bowen, Williams, & Horvath, 2004;

Chiasson et al., 2006; Du Bois, Johnson, & Mustanski, 2012; Franks et al., 2018; B. S. Mustanski, 2001; Sanchez et al., 2018; Saxton, Dickson, & Hughes, 2013; Simon Rosser et al., 2009). For example, HIV investigators have focused recruitment strategies for men who have sex with men infected with or at HIV risk (“HIV by Group | HIV/AIDS | CDC,” 2018) based on their usage of online chat rooms and mobile geosocial networking applications (i.e. “apps”) to identify potential partners (Bowen et al., 2004; Buckingham et al., 2017; Burrell et al., 2012; Grov, Breslow, Newcomb, Rosenberger, & Bauermeister, 2014; Iribarren et al., 2018; Jones & Salazar, 2016; Liau, Millett, & Marks, 2006; B. S. Mustanski, 2001; Rendina & Mustanski, 2018; Zou & Fan, 2017).

There are similarities between traditional recruitment (hard-copy flyers, mailed materials, in-person meetings) and Internet recruitment (Gelinas et al., 2017; Raymond et al., 2010), yet use of online tools involves fundamentally different scientific and ethical challenges. For example, online recruiting raises concerns about data validity, specifically the difficulties of assuring participant identity, eligibility and avoiding multiple entries by individual respondents (Grey et al., 2015; Pequegnat et al., 2007). Data validity checks may be particularly challenging when anonymity is promised as a means of protecting participant privacy for research on socially sensitive issues making it quite easy for respondents to mislead investigators about their geographical location, age, gender, race, sexual orientation and behaviors (Curtis, 2014; Pequegnat et al., 2007).

In addition, researchers have to be wary of fraudulent survey responses in the form of automated bots and ineligible participants seeking compensation. Strategies to prevent and detect fraudulent responses have potential challenges and disadvantages including the need for costly technological expertise, continuous monitoring to delete suspicious responses, or additional procedures that require participation identification (Chen et al., 2018; Grey et al., 2015; Pequegnat et al., 2007; Simon Rosser et al., 2009; Teitcher et al., 2015). If such methods are too stringent, researchers may eliminate valid data for their target population, resulting in a sample that does not represent the communities they are researching or include invalid data. Scientific integrity is also challenged by unequal access to the Internet, known as the digital divide (Du Bois et al., 2012; Lorence, Park, & Fox, 2006; Luque et al., 2013). If marginalized groups within target populations are less likely to have private Internet access, samples may be not be representative leading to inaccurate conclusions.

There are also ethical considerations tied to the ability to identify minors seeking access to studies (Markham & Buchanan, 2012; Pequegnat et al., 2007). This is especially problematic for HIV research which often is directed at participants over the age of 18 to avoid IRB requirements for guardian permission of under aged youth for epidemiological or intervention research involving collection of information regarding sexual behavior or interventions that include HIV and STI testing and biomedical preventions or interventions (Fisher, Arbeit, Dumont, Macapagal, & Mustanski, 2016; Fisher & Mustanski, 2014; Macapagal, Coventry, Arbeit, Fisher, & Mustanski, 2017; B. Mustanski, Coventry, Macapagal, Arbeit, & Fisher, 2017; B. Mustanski & Fisher, 2016). Social networking sites offer investigators the use of targeted behavioral advertising techniques based on aggregate personal data obtained from users, their families, and friends to create individual user profiles that contain sensitive and personal information (Curtis, 2014; Gelinas et al., 2017).

Typical information available to researchers includes, but is not limited to, standard demographic data, education, employment histories, interests, location histories, sexual orientation, personal and sexual relationships, and users' browser histories. Participants in online communities where researchers are conducting observational research may not reasonably expect that their comments will be analyzed or published in journals as the semi-public community affords a sense of privacy (Dawson, 2014; Gelinias et al., 2017). In addition, many potential participants are unaware that the act of showing interest in a research study through clicking on an online recruitment advertisement is providing data to third-party companies and leaving an identifiable digital trail (Gelinias et al., 2017; Rendina & Mustanski, 2018). Although participants may have signed terms of service agreements that explain how their data may be made available, most Internet-users do not read these long, verbose contracts in detail nor do they understand most of the details (Galbraith, 2017).

Some authors have raised concerns that researchers and IRBs are also unaware of these privacy and confidentiality risks (Curtis, 2014; Dawson, 2014). The steps researchers take to protect participant data may not always be effective. In some published studies, identifiable participant data was found by searching the Internet for participants' quotes that the authors had taken efforts to anonymize (Dawson, 2014). Participants have less trust in social media or geosocial dating apps (Rendina & Mustanski, 2018), especially given instances such as the large data breach on a popular geosocial dating app where millions of users' private data was exposed (Burns, n.d.). This is of particular concern for HIV research where privacy and confidentiality are key elements of the relationship of trust and respect that exist between the researcher and the participant.

As popular as online recruitment tools have become for HIV and other types of socially sensitive research, to date there is little regulatory or professional guidance on how to identify and resolve challenges to data integrity and participant protections and data integrity (Bruckman, 2014; Gelinias et al., 2017; Shilton & Sayles, 2016). The available guidance is primarily based on theory, anecdotal experiences of individuals, and expert panels (British Psychological Society, 2017; Gelinias et al., 2017; Gupta, 2017; Hills, 2013; Kraut et al., 2004; Markham & Buchanan, 2012). It is also difficult for guidance to keep up with rapidly emerging technologies, therefore recommendations published recently may not be relevant to Internet researchers today. For example, existing guidelines do not address ethical issues involved in recruiting via online crowdsourcing platforms (Law, Gajos, Wiggins, Gray, & Williams, 2017). To date there is an absence of research on the actual experiences of HIV investigators, thus we do not know whether the critical literature identify these issues reflect the real world advantages and challenges of online recruitment for HIV research. To begin to contribute to critical dialogue on these issues, this study drew on the experiences of principal investigators (PIs) engaged in online HIV research to illuminate scientific and ethical benefits and challenges of using social media tools for recruitment.

Methods

Recruitment

Participants were recruited, following IRB approval, via the construction of a sampling frame using the NIH RePORTER, AIDS Clinical Trial Information Services (ACTIS), HIV Prevention Trials Network (HPTN), and Center for AIDS Research (CFAR) to search for researchers who had received PHS funds from 2007-2014 to conduct HIV/AIDS prevention and intervention research—including biomedical and social-behavioral research—but excluding career awards, fellowships, and small business awards. In addition, PIs must have been from US academic institutions and have used the internet to recruit participants into an HIV prevention study in the past five years. Although ethical challenges concerning the use of online recruitment and data collection are both nationally and internationally relevant, we focused on a national sample since Internet privacy regulations differ widely across different countries. Additionally, across international institutions there are different ethical guidelines that must be followed for the protection of human subjects. Principal investigators (PIs) were invited to participate via a direct email invitation which ascertained whether they met inclusion criteria. Of the 17 investigators contacted who met the inclusion criteria, 82% agreed to be interviewed. Following recruitment, researchers were provided an overview of the study via email or over the phone and an informed consent was obtained through email. The third author conducted semi-structured telephone interviews with 14 PIs who met inclusion criteria.

Interviews

An interview schedule was designed using key themes identified from issues prevalent in the research literature (Boyatzis, 1998; Braun & Clarke, 2006). The purpose of the study was described to all participants before the interviews took place; verbal consent for participation and audio recording were obtained before each interview. Interviews were conducted by the third author, an investigator with substantial experience with online research involving vulnerable populations. In addition to prepared questions, prompts were used to verify interpretations of answers and to explore emerging themes. The qualitative method was iterative in nature (DiCicco-Bloom & Crabtree, 2006). The interviews resembled a conversation between two professionals. However, the interviewer worked to achieve a rapport with the interviewee, assuming the role of a listener who directed the conversation to cover the main themes. In this way, new ideas emerged that were not anticipated with the initial schedule; these themes were integrated into subsequent interviews. After 14 interviews, no new themes emerged, indicating the saturation of the data. The duration of the interviews varied between 45 and 65 min.

Participants

As demonstrated in Table 1 the majority of participants were male (57.1%), white (85.7%) and non-Hispanic (92.9%). The interviewees were spread across the US with most residing in the Northeast (35.7%) or Midwest (42.9%). The majority of participants were in their 30s (42.9%) or 40s (35.7%), had been at their current university for an average of seven years, and over half were tenure track (50% assistant professor; 7.1% associate professor). All 14

interviewees worked at universities with a doctoral program with a majority specializing in Epidemiology (28.6%) or Public Health (28.6%).

Content analysis

The interviewer listened to all the audio recordings and verified the precision of transcription. All identifying information was removed from the transcripts. The transcripts were entered into Dedoose a qualitative analysis software package (“Dedoose Version 5.0.11, web application for managing, analyzing, and presenting qualitative and mixed method research data,” n.d.). All data were examined line-by-line, and the main categories and themes identified and coded using thematic analysis and the constant comparison method. The second author generated the initial thematic codes after reading all the transcripts. Five major themes emerged. The first and second authors independently applied the guide to seven transcripts, discussed disagreements, and modified criteria where appropriate, creating a final coding guide. The first and second authors then applied the final coding guide to 7 interviews yielding good inter-rater agreement (Kappa = .88 – 1.00). The first author then coded the remaining 7 interviews.

Results

Overview

The interview goals were to elicit the perspectives of principal investigators (PIs) regarding strategies employed and challenges that have arisen for ensuring scientific integrity and participant protections during the online recruitment of participants into HIV related research. Five main themes emerged: (1) Sampling advantages and disadvantages; (2) challenges for data integrity; (3) control of privacy protections; (4) researcher and participant responsibility; and (5) resources. Below we describe the themes with quotes from respondents. Table 2 includes longer quotes that illustrate the range of experiences and opinions reflecting each theme and Kappa values for each theme.

Theme 1: Sampling Advantages and Disadvantages: “Fast, easy, cheap, great way to find hard to reach populations”

“Hidden populations are really tough to find. And so epidemiologically I think the Internet is the most efficient way to get them...I think the Internet is the common connection mechanism for many of these groups”

Theme 1 reflected principle investigators’ perspectives on the benefits of online recruitment to sample hidden and socially stigmatized populations and to expand national and international reach at relatively low cost. But along with these gains were challenges faced by continuing low numbers of racial ethnic populations and oversampling.

Low cost access to hidden and geographically dispersed populations.—A majority of researchers interviewed agreed that a benefit of online recruitment for HIV research was the ability to use targeted advertising to sample hidden or socially stigmatized populations most at HIV risk, including MSM and people who use illicit drugs. Researchers use popular social media sites and geosocial dating apps because “[name of social media

platform] allows you to pick age, gender and location to target ads to people” as well as “men interested in men.” The Internet provides a space where socially stigmatized and “geographically disparate” populations such as “[sexual minority] guys in rural and conservative areas,” can safely connect and for whom “a community...exists only virtually.” However, accessing these communities raise ethical questions whether research staff should “join” these online communities solely for the purpose of recruitment. One PI reported that they had once discovered an assistant had created a “fake profile” to recruit on dating sites when the assistant “kept getting kicked off” the site. Another PI sought to ethically address this issue by having staff join a chat room stating that they were part of a research study “in the profile and on the image.” However, the PI still “wonder[ed] if we are legit” to create user accounts and recruit from chat rooms in this way.

Even with the expanded reach of online recruitment, some researchers found that as with in-person or ground mail recruitment, they “have a hard time reaching men of color—Black and Hispanic.” Population tailored methods required to sample these “minorities within minorities” included “placing actors who look Black or Hispanic in the ads,” “selecting for Black or Hispanic when buying the [name of social media platform] ad,” and “buy[ing] banner ads in English and Spanish.”

Cost effectiveness, oversampling and site restrictions.—Compared to in-person recruitment, online advertising is relatively inexpensive given the large numbers of people who might view the ad. As noted by one PI, “I would not have been able to reach as many people... with the budgets that I had, because I would not have been able to travel to all these different places.” However, some researchers learned difficult lessons around oversampling. If an ad was not sufficiently specific, large numbers clicked on the ad, but failed the eligibility screener, and researchers had to “pay for all those clicks... eating up our small recruitment budget.” This problem was sometimes caused by restrictions imposed by the host website limiting sexual language or photos depicting sexual minority men. Site word count restrictions of “140 characters” also meant “banner ads become more generic than we would like them,” leading to oversampling.

Many of the researchers spoke about how corporate regulations limited the control they had over their own recruitment. Researchers had to abide by the company’s rules, even when it made their procedures less effective. One researcher realized that their recruitment methods may have been against one website’s terms of service but continued since “so far [name of geosocial dating app] hasn’t kicked [them] out.” There was some desire that “companies and the researchers and the practitioners... sit down and say alright, we all have slightly different goals with some common goals here, how do we work together.”

Theme 2: Data Integrity: “You never know if people are who they say they are”

“There may be some enticements, but it is not the same feeling of someone who is standing right there with the survey saying...Hey please, please, please... I really think the Internet actually helps to diminish that either false responses or bad responses”

Theme 2 reflected how the anonymity of online recruitment can be advantageous for the quality and honesty of participant responses, while simultaneously presenting new challenges for detecting fraudulent or duplicate responses and ensuring data is valid.

Anonymity encourages honest responses.—Many PIs mentioned that the anonymity of online recruitment led to more truthful responses compared to “someone sitting in front of them potentially judging them,” especially when questions involved sensitive information such as HIV status or sexual activity. If participants felt uncomfortable, they could “turn off their web browser in a heartbeat” rather than give false responses.

Incentives and fraudulent responses.—Researchers had to balance the benefits of anonymity with the difficulties with data integrity. These included “bots or computerized algorithms,” fraudulent responses or repeat participants. Although, researchers often limited inclusion criteria to 18 and older to avoid IRB guardian permission requirements, validating age was often difficult and researchers often felt “there is nothing I can do to prevent that” had to “trust that they say they are over 18.” Although certain social media sites paid postings purportedly allowed targeting individuals of a certain age, PIs reported that individuals who “have said in their [name of social media platform] that they are eighteen” indicated they were younger during recruitment. Some investigators ran multiple questions on age or other inclusion criteria worded in different ways to attempt to exclude fraudulent responses.

Many attributed the mention of monetary incentives for participation as a primary cause of fraudulent responses: People tried to get around the eligibility criteria “in order to get paid, doing it repeatedly.” Incentives caused such a problem that a couple of the researchers had “quit paying people,” or were least considering it. Including in the ad that participants would receive financial incentives was typically the cause of such threats to validity: “if you put the word research or university and a dollar sign, then you get the robots.”

Strategies for data integrity.—To guard against bots or repeat participants, researchers used manual or automatic deduplication protocols. Although cross-referencing with data from social media pages was fairly common, there were limitations, e.g. “someone could set up a fake [name of social media platform] page.” Researchers doing intervention follow-ups found that when they spoke to the participants, “there were so many people who we found out had taken the survey that were not MSM, that we ended up tossing all of that data.” Another strategy that researchers used for ensuring data integrity was to “not give away any of our criteria” in recruitment materials. This strategy sometimes failed because of people posting on websites the “eligibility criteria you need to get through.” To overcome this challenge, one researcher “searches for places that may be advertising our study.”

Theme 3: Control of Privacy Protections: “We assume they know that [name of social media platform] is collecting all that information.”

“I do not believe that the level of security afforded by [name of social media platform] is suitable for online research. I mean, the only thing that we do is

observation type of ethnographic research on [name of social media platform] ... I would not necessarily want to communicate with participants directly.”

Theme 3 reflected the limits of control that researchers had over participant data during online recruitment when using social media companies, online survey companies, or other third-parties. This theme also reflected as participant perspectives of this control of protections.

Limited control over privacy and data protection.—Many researchers discussed the limited control over the protections of their participants’ data when recruiting on social media sites. There was consensus that some websites would use participant data for direct marketing, which was concerning because HIV studies involve sensitive information: “we did not want to make it when somebody clicks it... they would somehow be disclosed in some way.” There were “reports that [name of social media platform] has outed individuals” by “following where a person visits on the web using cookies.” Cookies are small text files that allow websites to recognize you and track your preferences. Using third-parties for “intervention development” could also lead to “unintentional disclosure” and “once those stories leak to the press, that can be incredibly damaging to people in the community.” Many of the researchers were uneasy with collecting identifying information about the participants if “the data could be housed externally” on these companies’ servers. As a result, researchers limited the types of research on certain social media sites because “[name of social media platform] is definitely walking a very razor thin kind of ethical line.” Citing language from federal regulations, minority felt that this “minimal risk” was “consistent with a risk that we encounter in daily life” when browsing the Internet or using our credit card online. One interviewee reasoned that they were comfortable with the “big data” that “large sites like [name of search engine]” collected, because “it really is about the vast quantity of data collected on a global level, rather than this one person’s personal story.”

Participant understanding of privacy protections. A couple of researchers spoke about how much control their participants believed the researchers had. When participants understood the limits of control, they would prefer researchers “develop your own platform” and “don’t use [name of social media platform], don’t use [name of search engine]” for recruitment. One researcher believed participants’ views were affected by the “huge [public] discourse about what is [name of social media platform] doing with my data.”

Strategies for control of privacy protections.—There was a lot of variation between where researchers collected and stored survey data. Some researchers used a secure server, rather than a web-based server because “revealing or risky” data would end up in “someone else’s cloud.” Some had a faculty “computer scientist develop our own...web survey” that was housed on institutional servers, but this was not economically feasible for everyone. Whilst there were still data risks involved with secure servers, this was a better option than “offline data collection” which involved traveling with “tapes and pieces of paper with people’s personal information.” One of the researchers decided not to use online data collection at all because they hadn’t figured out “how to make sure that everybody can keep their data safe.”

Theme 4: Researcher and Participant Responsibility: “I think our job as researchers and as IRBs is to do everything we can to mitigate and then it is the best we can do”

“We know these Tech companies are collecting mounds of personal data on individuals. Thus we need to protect them by building safeguards into our projects... We try not to collect any data on social media sites or on sites I don’t have an agreement with. So, I have control of all the data I collect.”

Theme 4 reflected the researcher’s responsibility to understand privacy and informational risk during online recruitment, implement protections, train staff, and communicate these risks with participants. Theme 4 also considers the participant’s responsibility to have basic understanding of how their online data can be used.

Researcher responsibility to understand, implement and train staff on data protection.—Some researchers articulated an ethical responsibility for “the researcher to know what data is being collected and what their use is.” A couple of researchers raised doubts as to whether the majority of researchers were taking this responsibility seriously and had “seen a lot of Internet-based researchers not give enough thought on those issues.”

Many researchers explained the privacy protections they had implemented in their own studies to protect participant data. These “safeguards” varied in their comprehensiveness. One researcher believed that “the more sensitive the information, the more incumbent it is on us to have both measures that are of extremely high confidentiality.” An example of this was a researcher who, after asking about HIV status, would “download the data multiple times during the day and...remove the names.”

A few researchers acknowledged it was “important... to train staff” to follow privacy protections. Simple procedures included “do not have data on their laptops.” Some intervention researchers had to “train our people very extensively” because of “risks to him, to the institution and... risks to participants.” A few researchers had dealt with staff who had acted inappropriately during recruitment. This could cause companies to block researchers: “you have one person out there that’s being very aggressive...then [name of geosocial dating app] says...no more researchers.”

Researcher responsibility to consider participant understanding and communicate data protection limits to participants.—Some researchers indicated that they included language describing the limits of their ability to protect privacy, for example, “there is an uncertain likelihood that [name of online survey software] may or may not use this information for something.” One researcher “outline[d] different protection measures,” but participants said it was “just too long.” Some researchers felt it had been necessary to give guidance on how participants could protect their privacy, for example “not to use public computers” and “clear the cache.” These efforts didn’t always work because “participants didn’t listen.”

Participant responsibility to understand privacy protections on websites they use.—Many of the researchers felt “that if someone is using [name of social media platform] and they are an adult they know the risks” because “the assumption of privacy is

really different now.” One researcher believed “the user is aware that they were being tracked, they have consented by the privacy policies of that group then they know that.” Even when collecting sensitive data such as HIV status, one researcher felt that because the participant’s status had been publicly posted in their screenname, they did not have to worry about privacy protections: “when you post something on the Internet it is a public space.”

Theme 5: Resources: “Borrowing other people’s work, using computer scientists for expert knowledge, and then using what we call expert advisory panels for stakeholder input”

“The IRB... understand all of the issues. They are very strict compared to other people... I actually appreciate that. I feel like they are protecting me as a person and a professor... our school and our studies. They really do understand the issue of social networking.”

Theme 5 reflected the various resources researchers used to conduct quality online recruitment, as well as the difficulties they faced with either a lack of guidance, or inhibitive guidance.

Institutional review boards.—There was a wide range of experiences with IRBs across the researchers interviewed. Many researchers spoke of “collaborative” and “well-informed” IRBs that worked with them to make sure their online recruitment procedures were ethically sound. IRBs acted as a referee when members of target communities were unhappy with research recruitment in private online spaces: “we were following IRB to the letter and the researcher had disclosed their status... the complaint was found to have no basis.” One researcher felt that their IRB had “challenged me... to do things that I would not have otherwise done because I thought they were unethical.”

In contrast, other researchers felt that IRBs put unnecessary restrictions on their recruitment practices because they were not “well-informed” about the nature of online research. One researcher wasn’t able to do her own online data collection because her IRB was “overly protective” and instead did secondary data analysis in collaboration with a “colleague at a different institution who has a wonderful IRB.” Some researchers felt it was important that the IRBs included IT specialists or that online research be “reviewed by the IT department at the institution.”

National or regulatory guidelines.—Many of the researchers wished there were “guidelines by the NIH and by all the IRBs that are separate and specific to Internet research” because “people are still winging it.” One of the researchers wanted more guidelines because “right now we do not even go there because we are so worried about it.” Another believed that by “decoupling the review of data security and informational risk from other risk is a very good idea.” In contrast, other researchers felt this “wouldn’t work... each study is so different and it’s gonna mean we have to jump through a lot of hoops. I think the local IRB should decide risk.”

IT experts, other researchers, and community stakeholders.—Many of the researchers learned best practices for online recruitment from other academics. Sometimes

this was structured such as an organized meeting, and sometimes this involved asking “other people doing research in this area. I find out what they are doing.” Another category of resource were “expert advisory panels of stakeholders” from the community. A few of the researchers spoke highly of the university “IT person” or “computer science specialist” who supported their research by making sure participant data was secure. One researcher claimed that their department is “so far ahead of the curve” within online recruitment and attributed this to their “strong health information technology group.”

Discussion

Our semi-structured interviews with principal investigators highlight tension between the benefits of online recruitment for HIV research and the difficulties or challenges. Our study also revealed a range of viewpoints regarding the responsibilities of PIs to protect participant data. The emergent themes illustrate that without clear and consistent guidelines across IRBs, researchers in this field have had to employ a ‘learn-as-they-go’ approach to online recruitment. Many interviewees had experienced similar difficulties, so researchers who share their successes and failures with colleagues from different institutions prevent investigators from repeating their mistakes.

Online recruitment is a cost-effective way of obtaining large, nationally representative samples and reaching hidden or marginalized populations, however, the desired samples aren’t going to be easily obtained without employing certain strategies. Recruiting racial and ethnic minority populations who are at increased risk for HIV remains a challenge (Beymer, Holloway, & Grov, 2018; Hirshfield, Grov, Parsons, Anderson, & Chiasson, 2015; Madkins et al., 2018). Seeking advice from community stakeholders about language and images that will work, while complying with restrictions set by the recruitment website, is an effective strategy that can prevent oversampling and paying for large numbers of ineligible ad clicks. Advice from community advisory boards in past HIV research studies have included suggestions about advertising images/language and the social media platforms best suited for recruiting online (Franks et al., 2018; Raymond et al., 2010; Yuan, Bare, Johnson, & Saberi, 2014). There is also need for empirical research on the opinions of participants from target populations about online recruitment for HIV research specifically. A recent study about participant perspectives on data privacy is a good example of this (Rendina & Mustanski, 2018), but it would be useful to know more from participants from hidden populations about what kind of recruitment materials are or are not effective.

The anonymity of online recruitment can encourage more honest responses but poses significant challenges for data integrity. Researchers have found effective ways to minimize invalid responses, many of which are consistent with suggestions from published guidelines based on theory or individual experience (British Psychological Society, 2017; Gelinis et al., 2017; “Guidance Regarding Social Media Tools,” 2016; Hills, 2013; Kraut et al., 2004; Teitcher et al., 2015; Young, 2012). It is critical to establish validation protocols prior to recruitment, however the current research sheds light on the successes and difficulties researchers have had using these protocols, as well as highlighting that researchers have different ideas about how stringent these checks should be. There is a need for empirical research that assesses the efficacy of validation and deduplication checks used by a large and

diverse sample of HIV researchers. Some HIV researchers have published research that analyzes the differences between valid and invalid data detected by their manual and automatic protocols (Grey et al., 2015). Guidelines from professional organizations should be updated based on this empirical research. This would ensure that the research community can be confident about the data integrity of the work they read and cite.

Researchers feel they cannot completely control the privacy protections afforded by social media sites or online survey software, which is interesting given the high levels of trust in online/mobile research reported by participants in another study (Rendina & Mustanski, 2018). There are measures investigators can take to reduce informational risk and make participants aware of the limit of their control. Using institutional servers where possible, rather than cloud-based servers, is an approach recommended by many researchers dealing with sensitive data. The existing guidelines are not clear on the best approach regarding servers and some guidelines do not even mention this issue (Kraut et al., 2004; Young, 2012). There is a need for further research examining the experiences of a larger sample of HIV researchers regarding the data protection of their participants using different types of servers.

There needs to be more discussion within the research community regarding the possibility that clicking on online ads may give the company potentially identifying or sensitive information about that participant. This is a serious concern given risk of social stigmatization for marginalized populations. The informed consent stage may be too late to give prospective participants the necessary information about informational risk since participants have already clicked on the ad. It may be necessary for the research community to speak with social media companies about this issue and develop solutions now that social media is becoming a primary platform for research companies. Researchers should at least understand the business goals of social media companies they use and how this may affect their recruitment procedures (Young, 2012). There is also need to understand the perspectives of target populations about informational risk. Existing research suggests that while participants have relatively high levels of trust in researchers, they have less trust in apps or websites that may be storing their data (Rendina & Mustanski, 2018).

There was little consensus about whose responsibility it is during the research process to protect participant data. The viewpoint that adults nowadays should know the risks of being online, that Internet companies are constantly mining their data, and therefore research doesn't pose additional risk, is one that merits further debate. Researchers who assume that adults "know the risks" may be overestimating what the reasonable person understands. Not everyone may be up-to-date on news stories about privacy protections and the ways in which user data can be tracked. Participants may have a higher level of trust in researchers' abilities to protect their data than what is actually feasible. Further research should explore the extent to which target populations for HIV research understand online privacy, as this may illuminate what needs to be communicated at the recruitment stage. Participant focus groups have previously been used to explore participant understanding of privacy implications of social media recruitment within cancer research (Bender, Cyr, Arbuckle, & Ferris, 2017).

One reason why researchers may hold such a broad range of perspectives about the ethics of online recruitment, is that IRBs at different universities clearly have different understandings about informational risk and hold researchers to different standards. The current study focused on US researchers but this appears to be an issue internationally (Beddows, 2008). National, empirically-based guidelines may increase consistency across research studies, especially since researchers have found existing guidelines to be vague and sometimes conflicting (Gelinas et al., 2017; Kosinski, Matz, Gosling, Popov, & Stillwell, 2015). Previous research surveying IRB members themselves discovered conflicting opinions on the strengths and weakness of existing online research guidelines (Buchanan & Hvizdak, 2009). It's possible that for some researchers, such guidelines may hinder the types of research they have been permitted to conduct for years, whilst for other researchers, guidelines may open doors for studies they have previously struggled to get approved by their IRBs. A clear recommendation from our analysis, which echoes suggested/theoretical guidelines (Hills, 2013; Kraut et al., 2004; Young, 2012), is that IRBs should include a technology expert who understands what is and is not possible when conducting online recruitment, so that research that may help marginalized communities is not hindered by false assumptions, and so that research that may cause unintended harm to such groups is prevented.

Conclusion

This study speaks to the online recruitment experiences of HIV researchers sampling hidden populations, addressing data integrity challenges, and managing participant privacy protections. In so doing this study contributes information on the real-world experiences of HIV principle investigators conducting online recruitment to recommendations based on individual authors and guidelines. The benefits and challenges of online recruitment will change with the evolving nature of technology and online media, and therefore guidelines should be updated regularly to reflect the evolving experiences of HIV researchers.

Acknowledgements:

This work was supported by NIDA under grant #R25 DA031608-01; the Fordham GSAS Summer Assistantship, and the Intramural Research Program of the NIH, NIDA.

References:

- Anderson M, Perrin A, Jiang J, & Kumar M (2019). 10% of Americans don't use the Internet. Who are they? Pew Research Center Retrieved from <https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/>
- Beddows E (2008). The methodological issues associated with Internet-based research. *International Journal of Emerging Technologies & Society*, 6(2), 124–139.
- Bender JL, Cyr AB, Arbuckle L, & Ferris LE (2017). Ethics and privacy implications of using the Internet and social media to recruit participants for health research: A privacy-by-design framework for online recruitment. *Journal of Medical Internet Research*, 19(4), e104 10.2196/jmir.7029 [PubMed: 28385682]
- Beymer MR, Holloway IW, & Grov C (2018). Comparing self-reported demographic and sexual behavioral factors among men who have sex with men recruited through Mechanical Turk, Qualtrics, and a HIV/STI clinic-based sample: Implications for researchers and providers. *Archives of Sexual Behavior*, 47(1), 133–142. 10.1007/s10508-016-0932-y [PubMed: 28332037]

- Bowen A, Williams M, & Horvath K (2004). Using the Internet to recruit rural MSM for HIV risk assessment: Sampling issues. *AIDS and Behavior*, 8(3), 311–319. 10.1023/B:AIBE.0000044078.43476.1f [PubMed: 15475678]
- Boyatzis RE (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. SAGE.
- Braun V, & Clarke V (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. 10.1191/1478088706qp0630a
- British Psychological Society. (2017). *Ethics Guidelines for Internet-mediated Research*. Retrieved from <https://www.bps.org.uk/news-and-policy/ethics-guidelines-internet-mediated-research-2017>
- Bruckman A (2014). Research ethics and HCI In Olson JS & Kellogg WA (Eds.), *Ways of Knowing in HCI* (pp. 449–468). 10.1007/978-1-4939-0378-8_18
- Buchanan EA, & Hvizdak EE (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics*, 4(2), 37–48. 10.1525/jer.2009.4.2.37
- Buckingham L, Becher J, Voytek CD, Fiore D, Dunbar D, Davis-Vogel A, ... Frank I (2017). Going social: Success in online recruitment of men who have sex with men for prevention HIV vaccine research. *Vaccine*, 35(27), 3498–3505. 10.1016/j.vaccine.2017.05.002 [PubMed: 28526330]
- Burns J (n.d.). Report says Grindr exposed millions of users' private data, messages, locations. Retrieved 6 21, 2019, from Forbes website: <https://www.forbes.com/sites/janetwburns/2018/03/29/report-says-grindr-exposed-millions-of-users-private-data-messages-locations/>
- Burrell ER, Pines HA, Robbie E, Coleman L, Murphy RD, Hess KL, ... Gorbach PM (2012). Use of the location-based social networking application Grindr as a recruitment tool in rectal microbicide development research. *AIDS and Behavior*, 16(7), 1816–1820. 10.1007/s10461-012-0277-z [PubMed: 22851153]
- Chen Y-T, Bowles K, An Q, DiNenno E, Finlayson T, Hoots B, ... Wejnert C (2018). Surveillance among men who have sex with men in the United States: A comparison of web-based and venue-based samples. *AIDS and Behavior*, 22(7), 2104–2112. 10.1007/s10461-017-1837-z [PubMed: 28667468]
- Chiasson MA, Parsons JT, Tesoriero JM, Carballo-Diequez A, Hirshfield S, & Remien RH (2006). HIV behavioral research online. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 83(1), 73–85. 10.1007/s11524-005-9008-3 [PubMed: 16736356]
- Curtis BL (2014). Social networking and online recruiting for HIV research: Ethical challenges. *Journal of Empirical Research on Human Research Ethics*, 9(1), 58–70. 10.1525/jer.2014.9.1.58 [PubMed: 24572084]
- Dawson P (2014). Our anonymous online research participants are not always anonymous: Is this a problem? *British Journal of Educational Technology*, 45(3), 428–437. 10.1111/bjet.12144
- Dedoose Version 5.0.11, web application for managing, analyzing, and presenting qualitative and mixed method research data. (n.d.). Retrieved 7 10, 2019, from <https://www.dedoose.com/>
- DiCicco-Bloom B, & Crabtree BF (2006). The qualitative research interview. *Medical Education*, 40(4), 314–321. 10.1111/j.1365-2929.2006.02418.x [PubMed: 16573666]
- Du Bois SN, Johnson SE, & Mustanski B (2012). Examining racial and ethnic minority differences among YMSM during recruitment for an online HIV prevention intervention study. *AIDS and Behavior*, 16(6), 1430–1435. 10.1007/s10461-011-0058-0 [PubMed: 21986869]
- Fain DC, & Pedersen JO (2006). Sponsored search: A brief history. *Bulletin of the American Society for Information Science and Technology*, 32(2), 12–13. 10.1002/bult.1720320206
- Fisher CB, Arbeit MR, Dumont MS, Macapagal K, & Mustanski B (2016). Self-consent for HIV prevention research involving sexual and gender minority youth: Reducing barriers through evidence-based ethics. *Journal of Empirical Research on Human Research Ethics*, 11(1), 3–14. 10.1177/1556264616633963 [PubMed: 26956988]
- Fisher CB, & Mustanski B (2014). Reducing health disparities and enhancing the responsible conduct of research involving LGBT youth. *Hastings Center Report*, 44(s4), S28–S31. 10.1002/hast.367
- Franks J, Mannheimer SB, Hirsch-Moverman Y, Hayes-Larson E, Colson PW, Ortega H, & El-Sadr WM (2018). Multiple strategies to identify HIV-positive black men who have sex with men and

- transgender women in New York City: a cross-sectional analysis of recruitment results. *Journal of the International AIDS Society*, 21(3), e25091. 10.1002/jia2.25091
- Galbraith KL (2017). Terms and conditions may apply (but have little to do with ethics). *The American Journal of Bioethics: AJOB*, 17(3), 21–22. 10.1080/15265161.2016.1274796
- Gelinas L, Pierce R, & Cohen IG (& others). (2017). Using social media as a research recruitment tool: Ethical issues and recommendations. *American Journal of Bioethics*, 17(3), 3–14.
- Grey JA, Konstan J, Iantaffi A, Wilkerson JM, Galos D, & Rosser BRS (2015). An updated protocol to detect invalid entries in an online survey of men who have sex with men (MSM): How do valid and invalid submissions compare? *AIDS and Behavior*, 19(10), 1928–1937. 10.1007/s10461-015-1033-y [PubMed: 25805443]
- Grov C, Breslow AS, Newcomb ME, Rosenberger JG, & Bauermeister JA (2014). Gay and bisexual men's use of the Internet: research from the 1990s through 2013. *Journal of Sex Research*, 51(4), 390–409. 10.1080/00224499.2013.871626 [PubMed: 24754360]
- Guidance regarding social media tools. (2016, 2 16). Retrieved July 10, 2019, from National Institutes of Health (NIH) website: <https://www.nih.gov/health-information/nih-clinical-research-trials-you/guidance-regarding-social-media-tools>
- Gupta S (2017). Ethical issues in designing internet-based research: Recommendations for good practice. *Journal of Research Practice*, 13(2). Retrieved from <http://eric.ed.gov/?id=EJ1174008>
- Hills SR (2013). Considerations and Recommendations Concerning Internet Research and Human Subjects Research Regulations, with Revisions. 18.
- Hirshfield S, Grov C, Parsons JT, Anderson I, & Chiasson MA (2015). Social media use and HIV transmission risk behavior among ethnically diverse HIV-positive gay men: Results of an online study in three U.S. states. *Archives of Sexual Behavior*, 44(7), 1969–1978. 10.1007/s10508-015-0513-5 [PubMed: 26179596]
- HIV by Group | HIV/AIDS | CDC. (2018, 11 13). Retrieved June 21, 2019, from <https://www.cdc.gov/hiv/group/index.html>
- Iribarren SJ, Ghazzawi A, Sheinfil AZ, Frasca T, Brown W, Lopez-Rios J, ... Carballo-Diéguez A (2018). Mixed-Method Evaluation of Social Media-Based Tools and Traditional Strategies to Recruit High-Risk and Hard-to-Reach Populations into an HIV prevention intervention study. *AIDS and Behavior*, 22(1), 347–357. 10.1007/s10461-017-1956-6 [PubMed: 29124420]
- Jones J, & Salazar LF (2016). A review of HIV prevention studies that use social networking sites: Implications for recruitment, health promotion campaigns, and efficacy trials. *AIDS and Behavior*, 20(11), 2772–2781. 10.1007/s10461-016-1342-9 [PubMed: 26902295]
- Kosinski M, Matz SC, Gosling SD, Popov V, & Stillwell D (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, 70(6), 543–556. 10.1037/a0039210 [PubMed: 26348336]
- Kraut R, Olson J, Banaji M, Bruckman A, Cohen J, & Couper M (2004). Psychological research online: report of Board of Scientific Affairs' advisory group on the conduct of research on the internet. *American Psychologist*, 59(2), 105–117. 10.1037/0003-066X.59.2.105 [PubMed: 14992637]
- Law E, Gajos KZ, Wiggins A, Gray ML, & Williams A (2017). Crowdsourcing as a tool for research: Implications of uncertainty. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1544–1561. 10.1145/2998181.2998197
- Liau A, Millett G, & Marks G (2006). Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. *Sexually Transmitted Diseases*, 33(9), 576. 10.1097/01.olq.0000204710.35332.c5 [PubMed: 16540884]
- Lorence DP, Park H, & Fox S (2006). Racial disparities in health information access: resilience of the digital divide. *Journal of Medical Systems*, 30(4), 241–249. 10.1007/s10916-005-9003-y [PubMed: 16978003]
- Luque AE, Van Keken A, Winters P, Keefer MC, Sanders M, & Fiscella K (2013). Barriers and facilitators of online patient portals to personal health records among persons living with HIV: Formative research. *JMIR Research Protocols*, 2(1), e8. 10.2196/resprot.2302 [PubMed: 23612564]
- Macapagal K, Coventry R, Arbeit MR, Fisher CB, & Mustanski B (2017). "I won't out myself just to do a survey": Sexual and gender minority adolescents' perspectives on the risks and benefits of sex

- research. *Archives of Sexual Behavior*, 46(5), 1393–1409. 10.1007/s10508-016-0784-5 [PubMed: 27469352]
- Madkins K, Greene GJ, Hall E, Jimenez R, Parsons JT, Sullivan PS, & Mustanski B (2018). Attrition and HIV risk behaviors: A comparison of young men who have sex with men recruited from online and offline venues for an online HIV prevention program. *Archives of Sexual Behavior*, 47(7), 2135–2148. 10.1007/s10508-018-1253-0 [PubMed: 30069614]
- Markham A, & Buchanan E (2012). Recommendations from the AoIR Ethics Working Committee (Version 2.0). 19.
- Martín-Santana JD, & Beerli-Palacio A (2012). The effectiveness of web ads: Rectangle vs contextual banners *Online Information Review*; Bradford, 36(3), 420–441. <http://dx.doi.org.avoserv2.library.fordham.edu/10.1108/14684521211241431>
- Mustanski B, Coventry R, Macapagal K, Arbeit MR, & Fisher CB (2017). Sexual and gender minority adolescents' views on HIV research participation and parental permission: A mixed-methods study. *Perspectives on Sexual and Reproductive Health*, 49(2), 111–121. 10.1363/psrh.12027 [PubMed: 28445608]
- Mustanski B, & Fisher CB (2016). HIV rates are increasing in gay/bisexual teens. *American Journal of Preventive Medicine*, 51(2), 249–252. 10.1016/j.amepre.2016.02.026 [PubMed: 27061893]
- Mustanski BS (2001). Getting wired: Exploiting the internet for the collection of valid sexuality data. *The Journal of Sex Research*, 38(4), 292–301. 10.1080/00224490109552100
- Pequegnat W, Rosser BRS, Bowen AM, Bull SS, DiClemente RJ, Bockting WO, ... Zimmerman R (2007). Conducting Internet-based HIV/STD prevention survey research: Considerations in design and evaluation. *AIDS and Behavior*, 11(4), 505–521. 10.1007/s10461-006-9172-9 [PubMed: 17053853]
- Raymond HF, Rebchook G, Curotto A, Vaudrey J, Amsden M, Levine D, & McFarland W (2010). Comparing Internet-based and venue-based methods to sample MSM in the San Francisco Bay Area. *AIDS and Behavior*, 14(1), 218–224. 10.1007/s10461-009-9521-6 [PubMed: 19160034]
- Rendina HJ, & Mustanski B (2018). Privacy, trust, and data sharing in web-based and mobile research: Participant perspectives in a large nationwide sample of men who have sex with men in the United States. *Journal of Medical Internet Research*, 20(7), e233 10.2196/jmir.9019 [PubMed: 29973332]
- Sanchez TH, Zlotorzynska M, Sineath RC, Kahle E, Tregear S, & Sullivan PS (2018). National trends in sexual behavior, substance use and HIV Testing among United States men who have sex with men recruited online, 2013 Through 2017. *AIDS and Behavior*, 22(8), 2413–2425. 10.1007/s10461-018-2168-4 [PubMed: 29948340]
- Saxton P, Dickson N, & Hughes A (2013). Who is omitted from repeated offline HIV behavioural surveillance among MSM? Implications for interpreting trends. *AIDS and Behavior*, 17(9), 3133–3144. 10.1007/s10461-013-0485-1 [PubMed: 23605157]
- Shilton K, & Sayles S (2016). “We aren’t all going to be on the same page about ethics”: Ethical practices and challenges in research on digital and social media. 2016 49th Hawaii International Conference on System Sciences (HICSS), 1909–1918. 10.1109/HICSS.2016.242
- Simon Rosser BR, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, & Peterson JL (2009). HIV sexual risk behavior by men who use the internet to seek sex with men: Results of the Men’s INternet Sex Study-II (MINTS-II). *AIDS and Behavior*, 13(3), 488–498. 10.1007/s10461-009-9524-3 [PubMed: 19205866]
- Teitcher JEF, Bockting WO, Bauermeister JA, Hoefler CJ, Miner MH, & Klitzman RL (2015). Detecting, preventing, and responding to “fraudsters” in Internet research: Ethics and tradeoffs. *The Journal of Law, Medicine & Ethics*, 43(1), 116–133. 10.1111/jlme.12200
- Ur B, Leon PG, Cranor LF, Shay R, & Wang Y (2012). Smart, useful, scary, creepy: Perceptions of online behavioral advertising. *Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS '12*, 1 10.1145/2335356.2335362
- Young SD (2012). Recommended guidelines on using social networking technologies for HIV prevention research. *AIDS and Behavior*, 16(7), 1743–1745. 10.1007/s10461-012-0251-9 [PubMed: 22821067]

- Yuan P, Bare MG, Johnson MO, & Saberi P (2014). Using online social media for recruitment of human immunodeficiency virus-positive participants: A cross-sectional survey. *Journal of Medical Internet Research*, 16(5), e117 10.2196/jmir.3229 [PubMed: 24784982]
- Zou H, & Fan S (2017). Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: A systematic review and meta-analysis. *Archives of Sexual Behavior*, 46(4), 885–894. 10.1007/s10508-016-0709-3 [PubMed: 27043837]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Demographic information of principal investigators interviewed for this study

Characteristic		N = 14
		n (%)
Gender		
	Male	8 (57.1)
	Female	5 (35.7)
	Transgender	1 (7.1)
Ethnicity		
	Hispanic	1 (7.1)
	Not Hispanic	13 (92.9)
Race		
	Black/ African American	2 (14.3)
	White	12 (85.7)
Geographic Region		
	Northeast	5 (35.7)
	Midwest	6 (42.9)
	South	2 (14.3)
	West	1 (7.1)
Age Category		
	30s	6 (42.9)
	40s	5 (35.7)
	50s	2 (14.3)
	60s	1 (7.1)
Job Title		
	Researcher	2 (14.3)
	Asst. Professor	7 (50.0)
	Assoc. Professor	1 (7.1)
	Full Professor	4 (28.6)
Mean years at current university		8.46 (range, 1-23)
University has a doctoral program		14 (100)
Discipline		
	Epidemiology	4 (28.6)
	Public Health	4 (28.6)
	Psychology	3 (21.4)
	Interdisciplinary	2 (14.3)
	Psychiatry	1 (7.1)

Table 2.

Themes and representative quotes reflecting the experiences and perspectives of principle investigators using online recruitment for HIV research.

THEME 1: Sampling advantages and disadvantages: Online recruitment facilitates sampling of hidden populations at relatively low cost, but advertising restrictions and oversampling cause difficulties ($k = .95, p < .05$).	
Low cost access to hidden and geographically dispersed populations.	
•	“You can really efficiently target your ads to these groups that otherwise either may be hidden or very difficult to reach in large numbers. I think to me that is probably the greatest strength at least in terms of the recruitment team.”
•	“Many of the minorities within minorities communities who are geographically disparate that I work with do have online communities. So the internet has been a place where people have come together, where before they felt isolated and found support and validation for their identity and experiences”
•	“We stratified by race ethnicity because we wanted to oversample MSM of color. And that was pretty successful because one of the issues is often the samples are overly white online.”
Cost effectiveness, oversampling and Advertising Restrictions	
•	“The benefit of using the internet is I think greater access to the source or target population and efficiency, money, time, because you can do these surveys in person, it would just take a huge budget to go visit people and take the time to do it. But the computer can do it fast.”
•	“We just blew through the budget really fast—our recruitment dollars. Within two weeks we’d spent \$3000 on recruitment... Most of the people that clicked on the ads were not eligible because we weren’t really going to our target demographic”
•	“We’ve encountered images that Facebook didn’t approve... “We don’t want this on our website,” ...It wasn’t even that graphic. It was just two men without shirts like hugging each other.”
THEME 2: Data Integrity: The anonymity of online recruitment may prevent social desirability bias, but incentives lead to fraudulent or duplicate responses ($k = .93, p < .05$)	
Anonymity encourages honest responses	
•	“Some people may be more truthful in their responses because there is not someone sitting in front of them potentially judging them. But I am not sure as a methodologist that is true. I am not convinced yet.”
•	“So with an internet study, people can choose where they feel safe and where they are doing the interview from or taking the survey from. So there is a level of confidentiality and safety that was just simply not possible in a lot of settings kind of offline.”
Incentives and fraudulent responses	
•	“Incentives matter, but I think also there are bots and... search tools that people use to scavenge the web and enter junk. But, if you are not mindful, you will think that it is a true responder when in fact it is not... we have also seen sometimes it is something between 15 to 20% of all entries are suspicious if not fraudulent.”
Strategies for data integrity	
•	“We look at people’s IP address, operating system, browser, time... to complete the survey, number of entries from the same IP address. We verify their state of residence to their IP region... crosscheck their email with if they have a Facebook account... we compare their Facebook information for their demographic information.”
THEME 3: Control of Privacy Protections: Researchers have limited control of participant privacy when using social media platforms and online survey companies for recruitment and data collection ($k = .88, p < .05$).	
Limited control over privacy and data protection	
•	“We can do all we can and want but there’s going to be a limit to how much protection we can provide... participants and researchers are in it, they’re aware of that and they still go for it. And I think in most cases, it’s fine. But you’re right, there’s always risk for breach.”
Participant understanding of privacy protections	

THEME 1: Sampling advantages and disadvantages: Online recruitment facilitates sampling of hidden populations at relatively low cost, but advertising restrictions and oversampling cause difficulties ($k = .95, p < .05$).
<ul style="list-style-type: none"> • “And I mean, we’ve conducted focus groups before...and they were like, “Okay, well, why can’t you just develop your own platform and do the, just some screening to the actual study?” Don’t use Facebook, don’t use Google... just develop your own platform. And that would be ideal, I think. But the funds are usually not there.”
Strategies for control of privacy protections
<ul style="list-style-type: none"> • “All the data is immediately housed and stored on our server and there’s no intermediary. So I did not want to use Facebook, Google Plus, Skype...I didn’t want to use any of those where the data could be housed externally to us... So, even though that might have been more convenient to have a Facebook chat with someone, it wasn’t secure, right?”
THEME 4: Competence and Responsibility: Researcher and participant responsibility to understand online privacy protections and implement appropriate safeguards ($k = .92, p < .05$)
Researcher responsibility to understand, implement and train staff on data protection
<ul style="list-style-type: none"> • “I would read their policies and be sure that the person has consented to say what they are collecting. I would not advertise on a social network that I was not comfortable using myself. I would have to be able to see what the person could see.”
Researcher responsibility to consider participant understanding and communicate data protection limits to participants
<ul style="list-style-type: none"> • “We do a lot of housekeeping at the beginning of every session... it’s filled with reassurance and asking them not to change their password and asking them to clear their history and to not friend anyone.”
Participant responsibility to understand privacy protections on websites they use
<ul style="list-style-type: none"> • “There is a presumption of confidentiality but I also think if you think about the sort of person who is comfortable putting their HIV status and then requesting for a Starbucks gift card... I just don’t think that is the sort of person that would truly be harmed in any way by the connection between the status and the name but I could be wrong.”
<ul style="list-style-type: none"> • “We use the reasonable person standard. Does the reasonable person on Facebook appreciate that Facebook is using all these games and ads and click-throughs to collect data to make money? And I think our team has concluded that for an adult on Facebook we assume they know that Facebook is collecting all that information.”
THEME 5: Resources: Different experiences with sources of support, guidance, and accountability for online recruitment ($k < 1.00, p < .05$).
Institutional Review Boards
<ul style="list-style-type: none"> • “The IRB people I don’t feel like were very well informed about technology stuff and seem to be... oppositional about like any time you might suggest that some of their information might be out of date... they really, firmly believe that with an IP address you can get someone’s home address. Just by going to one of those lookup services. I kept saying to them no. It is the address of where the domain is hosted”
National or regulatory guidelines
<ul style="list-style-type: none"> • “We need a real protocol. But we do not have that yet. I think that type of best practices would be so helpful.”
<ul style="list-style-type: none"> • “I am more interested in changing the culture and practice than the rules. And I do not know how to change practice better. But to me... the current rules are fine. People just need to follow them better.”
IT experts and other researchers
<ul style="list-style-type: none"> • “When it goes to full board, it requires that there is an IT person looking specifically, looking exactly into protocol, what is the research, what are the concerns and what are the technological pieces that need to be taken into account.”
<ul style="list-style-type: none"> • “I went to a meeting at Columbia recently they were talking about problems with people faking participating in online research. We all went through and kind of shared our best practices. It was really helpful because I learned about the way other people were using IP addresses and using other things that we hadn’t been doing.”
Community stakeholders

THEME 1: Sampling advantages and disadvantages: Online recruitment facilitates sampling of hidden populations at relatively low cost, but advertising restrictions and oversampling cause difficulties ($k = .95, p < .05$).

- “Expert advisory panels of stakeholders who are often men who might be filling out our surveys. And we ask them, hey...does this explain it to you? Do you understand the risk? Can we say it this way? Would you assume this risk if you knew this piece of information?... And by expert I mean not just sort of ivory tower scholars but regular people.”
- “Because there is so much clutter on Facebook we were much more marketing savvy and we usually get young MSM themselves to give us, as part of our quality of work to give us feedback on our materials.”

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript