Published in final edited form as:

Arch Sex Behav. 2018 January ; 47(1): 133–142. doi:10.1007/s10508-016-0932-y.

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

Matthew R Beymer, PhD, MPH[1,2], **Ian W. Holloway, PhD, MPH, MSW**[3], and **Christian Grov, PhD, MPH**[4]

[1]Division of Infectious Diseases, Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, United States

[2]Los Angeles LGBT Center, Los Angeles, California, United States

[3]Department of Social Welfare, Luskin School of Public Affairs, University of California, Los Angeles, Los Angeles, California, United States

[4]City University of New York (CUNY) Graduate School of Public Health and Health Policy, New York City, New York, United States

Abstract

Recruitment for HIV research among gay, bisexual, and other men who have sex with men (MSM) has increasingly moved to the online sphere. However, there are limited data comparing the characteristics of clinic-based respondents versus those recruited via online survey platforms. MSM were recruited from three sampling sites (STI clinic, MTurk, and Qualtrics) to participate in a survey from March 2015 to April 2016. Respondents were compared between each of the sampling sites on demographics, sexual history, substance use, and attention filter passage. Attention filter passage was high for the online sampling sites (MTurk = 93%; Qualtrics = 86%), but significantly lower for the clinic-based sampling site (72%). Clinic-based respondents were significantly more racially/ethnically diverse, reported lower income, and more unemployment than online respondents. Clinic-based respondents reported significantly more male sexual partners in the previous three months (mean clinic-based = 6; MTurk = 3.6; Qualtrics = 4.5), a higher proportion of gonorrhea, chlamydia, and/or syphilis in the last year, and a greater proportion of methamphetamine use (clinic-based = 21%; MTurk = 5%), and inhaled nitrates use (clinic-based = 41%; MTurk = 11%). The clinic-based sample demonstrated more demographic diversity and a greater proportion of HIV risk behaviors when compared to the online samples, but

Research involving Human Participants and/or Animals

This study was approved by the University of California, Los Angeles North General Institutional Review Board (IRB00004474; Project No. 14-000750).

Informed consent

Informed consent was provided for each participant completing the survey. An example of the informed consent can be found here: https://uscsocialwork.qualtrics.com/SE/?SID=SV_beXsWwODysDi2od

Disclosure of potential conflicts of interest

The authors disclose no conflicts of interest.

also a relatively low attention filter passage rate. We recommend the use of attention filters across all modalities to assess response validity and urge caution with online survey engines as samples may differ demographically and behaviorally when compared to clinic-based respondents.

Keywords

Survey methods; Mechanical Turk; MTurk; Qualtrics; Instructional Manipulation Check; Attention filters

INTRODUCTION

Gay, bisexual, and other men who have sex with men (MSM) are at high risk for a range of health issues including substance abuse and psychiatric disorders which are associated with an elevated incidence of HIV (Paul, Catania, Pollack, & Stall, 2001; Parsons, Grov, & Golub, 2012). Periodic surveys are important in order to understand emerging and changing HIV risk factors among MSM. These data can then be used to design HIV prevention interventions that address the diversity and specific circumstances of MSM environments. Although MSM are a priority population in HIV prevention research, conducting extensive research has been particularly difficult due to length of time and cost of in-person and telephone surveys (Grov, Bux Jr., Parsons, & Morgenstern, 2009; Parsons, Vial, Starks, & Golub, 2013; Vial, Starks, & Parsons, 2014; Vial, Starks, & Parsons, 2015).

In the past ten years, traditional telephone and face-to-face survey methods have increasingly been replaced by online research methods (Baker et al., 2010; Grov et al., 2016). Previous studies have shown that online surveys can produce as reliable data as both telephone (Braunsberger, Wybenga, & Gates, 2007; Rankin et al., 2008; Simons & Chabris, 2012) and in-person (Birnbaum, 2000; Touvier et al., 2011) surveys. However, while these new methods are promising, they have a number of challenges for behavioral research including repeat participation, higher rates of dropout, underrepresentation of African-American and Latino respondents, and reduced control when compared to laboratory conditions (Birnbaum, 2004; McKee, Picciano, Roffman, Swanson, & Kalichman, 2006; Sullivan et al., 2011; Grov et al., 2016). These issues can lead to biased estimates and a lack of data for sub-populations of MSM facing the highest rates of HIV infection.

Previous studies have attempted to assess the comparability of online and in-person samples for MSM. A 2009 study assessed the comparability and random digit dialing internet-based samples and found that internet samples yielded more accurate responses (Chang & Krosnick, 2009). A European study found that MSM recruited in the European MSM Internet Survey had a lower median age when compared to surveillance data (Marcus, Hickson, Weatherburn, Schmidt, & Network, 2013). A study among MSM in the United States found that individuals recruited from internet venue-based sampling differed significantly from those in the National HIV Behavioral Surveillance and Web-Based HIV Behavioral Surveillance by self-reported factors such as age, income, substance use, and HIV serostatus (Raymond et al., 2010). A study based in New York City compared MSM recruited from bathhouses, bars/clubs and Craigslist.org and found substantial demographic and behavioral differences (Grov, 2012). Lastly, a study in Atlanta found that MSM recruited

via Facebook yielded similar results to those recruited via venue-based, time-space sampling (Hernandez-Romieu et al., 2014). Although previous studies have utilized incentives, to our knowledge, no studies have assessed the comparability of in-person MSM samples to those drawn from paid survey engines like Qualtrics and Amazon's Mechanical Turk (MTurk).

Qualtrics and MTurk have existed since 2002 and 2005, respectively, and allow researchers to conduct surveys in communities that are traditionally hard to reach (Rosser et al., 2009). The Qualtrics model allows researchers to develop surveys using Qualtrics software and then request participant pools, or panels, where subjects are recruited by Qualtrics using the researcher's specified criteria. Researchers are quoted a price per subject based on the specificity of criteria, and the results are returned to the researcher at the completion of recruitment. In contrast, MTurk allows researchers to choose the compensation amount and survey modality through the creation of human intelligence tasks (HITs). Researchers directly post advertisements to their surveys on the MTurk worker space, and MTurk users can find surveys that pertain to them by searching by keywords and compensation amount. Previous studies have assessed the validity of MTurk for behavioral studies (Buhrmester, Kwang, & Gosling, 2011; Mason & Suri, 2012), but no studies to date have assessed the comparability of MSM respondents to clinic-based surveys.

The primary objective of this study was to compare the demographic and HIV risk differences between online samples of MSM recruited on MTurk and Qualtrics with a convenience sample of MSM recruited in community-based clinic in Los Angeles. A secondary objective was to compare the costs and time of participant recruitment and incentives across the MTurk, Qualtrics, and an STI clinic (subsequently referred to as "clinic-based") sampling sites to inform researchers and providers about the advantages and disadvantages of each modality for HIV behavioral research.

METHOD

Recruitment

The survey was programmed for all three sampling sites with Qualtrics software with a target recruitment of 200 respondents per sampling site. The survey was completed directly on the Qualtrics portal for the Qualtrics sampling site, and a link was provided for the clinic-based and MTurk sampling sites (Link to Survey: https://uscsocialwork.qualtrics.com/SE/?SID=SV_3HLElR6IWN2QTw9).

MTurk respondents were recruited from April 1st, 2015 to May 11th, 2015. Respondents were allowed to access the survey to determine eligibility if they had an address located in the United States and had an HIT approval rate (proportion of tasks that have been submitted and approved) of 95% or above (Peer, Vosgerau, & Acquisti, 2014). Respondents were directed to the survey if they used one or more of the following key words: gay/bisexual men, men who have sex with men, research, Los Angeles, Chicago, Miami, New York. These cities were chosen due to their size as well as having large defined, gay-centered areas. The recruitment script can be found in Appendix 1. Following survey completion, respondents were provided with a unique code to redeem their incentive.

Qualtrics respondents were recruited from June 24th to June 30th, 2015 by a requested Qualtrics panel that consisted of the following criteria: residents of Los Angeles, Chicago, Miami, or New York City metro areas; ages 18+; must be assigned a male sex at birth and currently identify as male; sex with a man in the last three months. The inclusion/exclusion criteria were the same for all three sampling sites.

Clinic-based respondents were recruited from July 28th, 2015 to March 15th, 2016 by HIV/STI testing counselors of a community-based clinic in Los Angeles. All clients receiving HIV/STI testing interface with an HIV/STI testing counselor and therefore were eligible for recruitment during this period. For recruitment script, see Appendix 1.

Survey

Upon clicking on the survey link, individuals were provided informed consent and asked to complete a screening tool to assess study eligibility. Individuals who met the inclusion criteria were asked a series of questions about their demographics, sexual behaviors in the past three months, history of sexually transmitted infections (STIs), substance use history, and knowledge about pre-exposure prophylaxis (PrEP). Questions for the survey were adapted from the full risk assessment used during HIV/STI screening by the community-based clinic for which the in-person sample was drawn. Qualtrics policy did not allow for the inclusion of the substance use questions, but these questions were asked in the MTurk and clinic-based sampling sites. Otherwise, the three questionnaires were identical.

Qualtrics and MTurk participants completed the survey from their own computers. Clinic-based respondents completed the survey on one of two tablets that were located in the waiting room of the community-based clinic in Los Angeles. Both tablets were equipped with privacy screens to protect the confidentiality of respondents.

Compensation

MTurk respondents were initially paid $0.50 for their participation. After one month of recruitment (April 1st, 2015 – May 1st, 2015), only 3 respondents had been recruited and the compensation level was increased to $1 for participation (May 8th – May 11th, 2015). Upon survey completion, a unique code was used to redeem their incentive. Qualtrics was paid at a rate of $6 per subject, but the actual payment amount from Qualtrics to respondents was between $2 and $3. Individuals who did not meet the inclusion criteria in MTurk were still monetarily compensated for their participation following the screening tool. Clinic-based respondents were provided with a code at the end of the survey and instructed to inform the HIV/STI testing counselor of this code to confirm survey completion and redeem a $10 gift card to either Target or Trader Joe’s.

Statistical analyses

The primary goal of this study was to compare the two online venues (MTurk and Qualtrics sampling sites) to a sample of MSM recruited in clinic. Categorical predictors were compared between these three groups under study via χ^2 tests or Fisher Exact Tests when one or more cells had counts less than 5. Continuous predictors had non-normal distributions and were compared Kruskal-Wallis tests. Four multivariable logistic regression models were

run to determine differences between sampling sites on demographics, sexual risk behaviors, substance use, and PrEP knowledge. All statistical tests used an alpha level of 0.05 and were calculated in SAS version 9.4 (Cary, NC). All protocols for the study were approved and overseen by the institutional review board of the researchers' academic institution.

RESULTS

Recruitment and Screening

A total of 511 respondents clicked on the survey for the MTurk sampling site, and 267 met the inclusion criteria for analysis. Of the 244 who were not qualified, 53 were cisgender female (22%); 21 reported a non-male gender identity (9%); 77 had not had sex with another man in the past three months (32%); 84 did not live in Los Angeles, Chicago, Miami, or New York City (34%); 9 quit before eligibility could be determined (4%).

A total of 366 respondents clicked on the survey for the Qualtrics sampling site, and 211 met the inclusion criteria for analysis. Of the 155 who were not qualified, six were cisgender female (4%); 13 reported a non-male gender identity (8%); 104 had not had sex with another man in the past three months (67%); 25 did not live in Los Angeles, Chicago, Miami, or New York City (16%); three were less than 18 years of age (2%); four quit before eligibility could be determined (3%).

A total of 317 respondents clicked on the survey for the clinic-based sampling site, and 231 met the inclusion criteria for analysis. Of the 86 who were not qualified, 11 were cisgender female (13%); 19 reported a non-male gender identity (22%); 24 had not had sex with another man in the past three months (28%); 11 did not live in Los Angeles, Chicago, Miami, or New York City (13%); 21 quit before eligibility could be determined (24%).

Attention Filter

A question was inserted three quarters through the survey that asked participants to answer "Disagree" for the question to determine if they were paying attention (known as an "attention filter" or "instructional manipulation check") and thus assess survey validity (Oppenheimer, Meyvis, & Davidenko, 2009). Approximately 86% of Qualtrics users, 93% of MTurk users, and 72% of clinic-based users passed this attention filter (Chi-square = 41; $p < 0.0001$). Participants were allowed to complete the survey and compensated regardless of their passage of the attention filter. However, demographic and behavioral results are presented only for those who passed the attention filter.

Demographic Differences

Qualtrics users were significantly older than the other two sampling sites with a reported mean age of 41 (median = 41; SD = 14) compared to a mean age of 29 for both MTurk and clinic-based respondents (median = 28; SD = 7) (Table 1). Qualtrics respondents and MTurk respondents were more likely to report a White race, 66% and 63% respectively, when compared to clinic-based respondents (25%). The highest proportion of Hispanic participants was recruited for the clinic-based sampling site (43%) compared to 15% for MTurk and 20% for Qualtrics. Approximately 22% of individuals in the Qualtrics sampling

site reported a graduate degree whereas only 12% reported this in the MTurk sampling site and 7% in the clinic-based sampling site. Lastly, Qualtrics users reported a much higher income on average with 45% reporting more than $50,000 per year compared to only 27% and 11% with this proportion for MTurk and clinic-based sampling sites, respectively. However, a substantial proportion of individuals in all three sampling sites refused to answer this question (Range: 15–34%).

Sexual History Differences

MTurk respondents reported the highest mean number of female sexual partners in the past three months (1.9) followed by Qualtrics respondents (1.5) and clinic-based respondents (1.1) (Table 2). Conversely, clinic-based respondents reported the highest mean number of male sexual partners in the past three months (6) followed by Qualtrics respondents (4.5) and MTurk respondents (3.6). Clinic-based respondents reported most frequently meeting partners on geosocial networking apps like Grindr and Scruff (44%) whereas MTurk respondents met most commonly in bars/clubs (53%) and Qualtrics users met most commonly on gay-centered websites like Adam4Adam and Manhunt (25%).

All three sampling sites differed on reported history of gonorrhea: 28% of clinic-based respondents, 2% of MTurk respondents, and 6% of Qualtrics respondents reported testing positive for gonorrhea within the past year. Similar trends were also observed for chlamydia and syphilis.

Substance Use Differences

As stated previously, Qualtrics users were not asked about substance use due to a company policy that prohibited questions about illegal drug use from panel participants. A similar proportion of clinic-based (18%) and MTurk respondents (12%) reported using cocaine or crack in the past year (Table 3). The same trend was observed for ecstasy use among clinic-based and MTurk respondents at 16% and 17%, respectively. However, 21% of clinic-based respondents reported methamphetamine use in the last year compared with only 5% of MTurk respondents. Similarly, 41% of clinic-based respondents reported nitrates/poppers use in the last year compared to only 11% of MTurk respondents.

Differences in Pre-Exposure Prophylaxis (PrEP) Knowledge/Attitudes

Clinic-based respondents and Qualtrics respondents were most likely to report that they knew a fair amount or a lot about PrEP, and MTurk respondents were most likely to report that they either knew only a little or nothing at all about PrEP before taking the survey (Table 4). In total 5% of Qualtrics respondents, 3% of MTurk respondents, and 10% of clinic-based respondents indicated they were currently taking PrEP. After being informed that PrEP was at least 90% effective in preventing HIV if taken every day, significantly more participants from the clinic-based sampling site said that they would definitely want to take PrEP when compared to the MTurk and Qualtrics sampling sites. However, when asked about their likelihood to use PrEP, the greatest proportion from each sampling site choose the option that stated, “I’m not sure, I might begin taking PrEP.”

DISCUSSION

Our study found that online samples of MSM recruited from Qualtrics and MTurk differed on demographics, sexual history, substance use, and PrEP knowledge when compared to participants surveyed in a clinic-based sample in Los Angeles, California. The advantages of the clinic-based sampling site included the racial/ethnic diversity and the greatest HIV risk, but this could be the artifact of the clinic's location in Los Angeles, an economically and racially/ethnically diverse urban environment. The disadvantages of the clinic-based sampling site included the low passage of the attention filter, long recruitment time (230 days), high cost per subject ($10 per subject), and high indirect costs (staff time for recruitment, purchase of iPads for survey completion, and hardware for securing the iPads). The advantages of the Qualtrics sampling site were the ease and speed of recruitment (7 days) and lower cost per subject when compared to the in-person sample ($6 per subject), but the disadvantages included relatively lower demographic diversity, a substantially different HIV risk profile when compared to other modalities, and the inability to ask questions about substance use. The MTurk sampling site advantages included the lowest cost per subject ($1 per subject), the quickest recruitment (only 3 days after increasing the incentive from $0.50 to $1), and the highest attention filter passage rate (93%). Although the MTurk sampling site was more diverse and reported higher HIV risk behaviors than the Qualtrics sampling site, there was lower demographic diversity and condomless anal sex (CAS) when compared to the clinic-based sampling site.

There are numerous limitations to this analysis. While MTurk and Qualtrics respondents were recruited from four major metropolitan areas across the United States (Los Angeles, Chicago, Miami, or New York City), the clinic-based sampling site only recruited participants from Los Angeles. Furthermore, individuals who visited the community-based clinic in Los Angeles for HIV/STI testing, treatment, or other biomedical prevention services and agreed to participate in the survey, may be different from other MSM in Los Angeles who either receive testing services at another facility or did not agree to participate in the survey. A second limitation was that the survey time could not be accurately recorded for all clinic-based respondents due to frequent network errors that prompted reloading of the survey home page. A third limitation was the high proportion of missing data for certain questions (Range: 0% – 34%). A fourth limitation was that while the clinic-based sampling site reported both higher levels of substance use and previous STIs, the polling of individuals from an HIV/STI testing clinic presents a potential for selection bias in recent CAS events may have prompted the visit to the HIV/STI clinic. Other methods for identifying potential participants, such as time-space sampling (Parsons, Grov, & Kelly, 2008) have the ability to reduce some bias in sample selection, but are also not without limitations including high cost and high refusal rates. Studies looking to replicate our findings should consider the strengths and weaknesses of various recruitment and sampling approaches to obtain the most representative sample of MSM. Lastly, we note that there are many venues in which we did not recruit from (e.g., bars/clubs, bathhouses) as well as emerging popular digital venues (i.e., geosocial networking apps like Grindr and Scruff) that may capture yet another segment of MSM (Holloway et al., 2014; Grov, Rendina, Jimenez,

& Parsons, 2016). Future studies looking at sampling representativeness should also incorporate these survey modalities into their study designs.

Despite these limitations, there are numerous strengths to our study. Few studies have compared demographics between traditional and online recruitment methods. In a study comparing samples from a telephone interview and MTurk, researchers showed that the samples produced similar demographic profiles (Simons & Chabris, 2012). In a study comparing MTurk users to 1) a sample recruited on Twitter, Facebook, and Reddit and 2) an in-person sample of college students, researchers found that the demographics of the samples differed, but the MTurk participants were more economically and racially diverse (Casler, Bickel, & Hackett, 2013). To the best of our knowledge, there are no studies to date that have compared the representativeness of Qualtrics users to more traditional recruitment methods. However, the older age and lack of racial diversity in the Qualtrics panel highlights the potential need to specify age, race, and ethnicity quotas when requesting Qualtrics panels. Future studies seeking to replicate similar population comparisons should also consider recruiting across multiple clinics as our results may be biased due to the location and population served.

A second strength is that we compared the validity of survey responses based on the use of an attention filter. A previous study showed that attention filters are a powerful tool in assessing accuracy of self-assessments (Hauser & Schwarz, 2016). Given these findings, we strongly recommend the use of attention filters to gauge response validity from both online and clinic-based respondents. In addition, future studies may consider comparing the validity of results with multiple attention filters compared to only a single attention filter to further understand how to maximize respondent engagement.

As the use of online assessments continues to grow, evaluations of new and existing survey modalities are important. Online survey tools like Qualtrics and MTurk have many strengths when compared to clinic-based samples including lower relative cost, faster recruitment due to decreased barriers for study participation, increased confidentiality, and the possibility to include a more diverse sample. Despite these benefits, these methods are not without limits. Our study found sampling biases that lead us to suggest caution when using these methods for behavioral research. It's important for researchers planning to use these methods to carefully develop quotas (in the case of Qualtrics) and/or detailed inclusion/exclusion criteria (in the case of MTurk) to obtain demographically representative samples. The use of one or more attention filters should be implemented to maximize respondent validity. Lastly, future research should focus on how to best obtain samples that best approximate population demographics and trends. These methods will never be perfect, but the potential benefits to both researchers and participants warrant further development of best practices for the field.

Acknowledgments

The authors would like to acknowledge the work of the Los Angeles LGBT Center's HIV/STI testing counselors for recruiting participants for the clinic-based arm of the study

The funding is as follows:

- This work was funded by the UCLA School of Public Affairs and the Los Angeles LGBT Center.

- MRB was supported by the UCLA Postdoctoral Fellowship Training Program in Global HIV Prevention Research (Currier and Gorbach, PIs); T32MH080634
- IWH was supported by California HIV/AIDS Research Program (CHRP; grant RP15-LA-007); the Center for HIV Identification, Prevention, and Treatment (CHIPTS; NIMH grant MH58107); the UCLA Center for AIDS Research (CFAR; NIAID grant AI028697); and the National Center for Advancing Translational Sciences through UCLA (CSTI; grant UL1TR000124). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH.

APPENDIX 1

MTurk Recruitment Script

"The survey is to assess the feasibility of conducting online research among gay and bisexual men. The survey will take approximately 15 minutes and ask various questions about your health and demographics."

Clinic-Based Recruitment Script

" XYZ $Clinic$ is conducting a study in collaboration with UCLA to determine if online samples of gay, bisexual, and other men who have sex with men are comparable to individuals who come to the XYZ $Clinic$ for HIV/STD screening on sexual health behaviors. The survey will last between 10 and 20 minutes, and you will be given a $10 gift card to either Target or Trader Joe's for your participation. You are not required to participate, and your agreement or rejection to participate will NOT affect either current or future services at the XYZ $Clinic$. Are you interested in participating?"

$Highlighted$ $Areas$ $de-identified$ for $manuscript$ $review$ $purposes$

References

Baker R, Blumberg SJ, Brick JM, Couper MP, Courtright M, Dennis JM, Zahs D. Research Synthesis. Public Opinion Quarterly. 2010; 74:711–781. DOI: 10.1093/poq/nfq048.

Birnbaum, MH. Psychological experiments on the internet. San Diego: Academic; 2000.

Birnbaum MH. Human research and data collection via the internet. Annual Review of Psychology. 2004; 55:803–832. DOI: 10.1146/annurev.psych.55.090902.141601

Braunsberger K, Wybenga H, Gates R. A comparison of reliability between telephone and web-based surveys. Journal of Business Research. 2007; 60:758–764. DOI: 10.1016/j.jbusres.2007.02.015

Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? Perspectives on Psychological Science. 2011; 6:3–5. DOI: 10.1177/1745691610393980 [PubMed: 26162106]

Casler K, Bickel L, Hackett E. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. Computers in human behavior. 2013; 29:2156–2160. DOI: 10.1016/j.chb.2013.05.009

Chang L, Krosnick JA. National surveys via RDD telephone interviewing versus the internet. Public Opinion Quarterly. 2009; 73:641–678. DOI: 10.1093/poq/nfp075

Grov C. HIV risk and substance use in men who have sex with men surveyed in bathhouses, bars/clubs, and on Craigslist.org: Venue of recruitment matters. AIDS and Behavior. 2012; 16:807–817. DOI: 10.1007/s10461-011-9999-6 [PubMed: 21748276]

Grov C, Bux D Jr, Parsons JT, Morgenstern J. Recruiting hard-to-reach drug-using men who have sex with men into an intervention study: Lessons learned and implications for applied research.

Substance Use & Misuse. 2009; 44:1855–1871. DOI: 10.3109/10826080802501570 [PubMed: 20001284]

Grov C, Cain D, Whitfield THF, Rendina HJ, Pawson M, Ventuneac A, Parsons JT. Recruiting a US national sample of HIV-negative gay and bisexual men to complete at-home self-administered HIV/STI testing and surveys: Challenges and opportunities. Sexuality Research and Social Policy. 2016; 13:1–21. DOI: 10.1007/s13178-015-0212-y [PubMed: 26858776]

Grov C, Rendina HJ, Jimenez R, Parsons JT. Using online settings to identify gay and bisexual men willing to take or with experience taking PrEP: Implications for researchers and providers. AIDS Education and Prevention. 2016; 28:378–392. [PubMed: 27710085]

Hauser DJ, Schwarz N. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. Behavior Research Methods. 2016; 48:400–407. DOI: 10.3758/s13428-015-0578-z [PubMed: 25761395]

Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del Rio C, Salazar LF, Frew PM, Rosenberg ES. The comparability of men who have sex with men recruited from venue-time-space sampling and Facebook: A cohort study. JMIR Research Protocols. 2014; 3:e37.doi: 10.2196/resprot.3342 [PubMed: 25048694]

Holloway IW, Dunlap S, Del Pino HE, Hermanstyne K, Pulsipher C, Landovitz RJ. Online social networking, sexual risk and protective behaviors: Considerations for clinicians and researchers. Current Addiction Reports. 2014; 1:220–228. DOI: 10.1007/s40429-014-0029-4 [PubMed: 25642408]

Marcus U, Hickson F, Weatherburn P, Schmidt AJ, Network, E. Age biases in a large HIV and sexual behaviour-related internet survey among MSM. BMC Public Health. 2013; 13doi: 10.1186/1471-2458-13-826

Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. Behavior Research Methods. 2012; 44:1–23. DOI: 10.3758/s13428-011-0124-6 [PubMed: 21717266]

McKee MB, Picciano JF, Roffman RA, Swanson F, Kalichman SC. Marketing the "sex check": Evaluating recruitment strategies for a telephone-based HIV prevention project for gay and bisexual men. AIDS Education and Prevention. 2006; 18:116–131. DOI: 10.1521/aeap.2006.18.2.116 [PubMed: 16649957]

Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of experimental social psychology. 2009; 45:867–872. DOI: 10.1016/j.jesp.2009.03.009

Parsons JT, Grov C, Golub SA. Sexual compulsivity, co-occurring psychosocial health problems, and HIV risk among gay and bisexual men: Further evidence of a syndemic. American Journal of Public Health. 2012; 102:156–162. DOI: 10.2105/ajph.2011.300284 [PubMed: 22095358]

Parsons JT, Grov C, Kelly BC. Comparing the effectiveness of two forms of time-space sampling to identify club drug-using young adults. Journal of Drug Issues. 2008; 38:1061–1081. [PubMed: 20686625]

Parsons JT, Vial AC, Starks TJ, Golub SA. Recruiting drug using men who have sex with men in behavioral intervention trials: A comparison of internet and field-based strategies. AIDS and Behavior. 2013; 17:688–699. DOI: 10.1007/s10461-012-0231-0 [PubMed: 22684657]

Paul JP, Catania J, Pollack L, Stall R. Understanding childhood sexual abuse as a predictor of sexual risk-taking among men who have sex with men: The Urban Men's Health Study. Child Abuse & Neglect. 2001; 25:557–584. DOI: 10.1016/s0145-2134(01)00226-5 [PubMed: 11370726]

Peer E, Vosgerau J, Acquisti A. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. Behavior Research Methods. 2014; 46:1023–1031. DOI: 10.3758/s13428-013-0434-y [PubMed: 24356996]

Rankin KM, Rauscher GH, McCarthy B, Erdal S, Lada P, Il'yasova D, Davis F. Comparing the reliability of responses to telephone-administered versus self-administered web-based surveys in a case-control study of adult malignant brain cancer. Cancer Epidemiology Biomarkers & Prevention. 2008; 17:2639–2646. DOI: 10.1158/1055-9965.epi-09-0304

Raymond HF, Rebchook G, Curotto A, Vaudrey J, Amsden M, Levine D, McFarland W. Comparing internet-based and venue-based methods to sample MSM in the San Francisco Bay Area. AIDS and Behavior. 2010; 14:218–224. DOI: 10.1007/s10461-009-9521-6 [PubMed: 19160034]

Rosser BRS, Miner MH, Bockting WO, Ross MW, Konstan J, Gurak L, Coleman E. HIV risk and the internet: Results of the Men's INTernet Sex (MINTS) Study. AIDS and Behavior. 2009; 13:746–756. DOI: 10.1007/s10461-008-9399-8 [PubMed: 18512143]

Simons DJ, Chabris CF. Common (mis)beliefs about memory: A replication and comparison of telephone and Mechanical Turk survey methods. Plos One. 2012; 7doi: 10.1371/journal.pone.0051876

Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, DiClemente RJ. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. Journal of Medical Internet Research. 2011; 13doi: 10.2196/jmir.1797

Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K, Hercberg S. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. British Journal of Nutrition. 2011; 105:1055–1064. DOI: 10.1017/s0007114510004617 [PubMed: 21080983]

Vial AC, Starks TJ, Parsons JT. Finding and recruiting the highest risk HIV-negative men who have sex with men. AIDS Education and Prevention. 2014; 26:56–67. DOI: 10.1521/aeap.2014.26.1.56 [PubMed: 24450278]

Vial AC, Starks TJ, Parsons JT. Relative efficiency of field and online strategies in the recruitment of HIV-positive men who have sex with men. AIDS Education and Prevention. 2015; 27:103–111. DOI: 10.1521/aeap.2015.27.2.103 [PubMed: 25915696]

Table 1

Bivariate Chi-Square/Fisher Exact Tests and Multivariable Logistic Regression of Demographics Differences between Clinic-based, Mechanical Turk (MTurk), and Qualtrics Survey Respondents who Identify as Gay/Bisexual Men and Passed the Attention Filter, April 2015 – March 2016.

	Clinic-based Arm* (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		2010 US Census***	Unadjusted p-value	Adjusted p-value
Demographics	***n***	**%**	***n***	**%**	***n***	**%**			
Age Group									
18–24	45	27%	66	27%	23	13%	NA		
25–29	53	32%	84	34%	20	11%	7%		
30–39	42	25%	72	29%	32	18%	13%	<0.0001	<0.0001
40–49	17	10%	17	7%	38	21%	14%		
50+	1	1%	5	2%	59	33%	32%		
Missing	8	5%	4	2%	9	5%	NA		
Race/Ethnicity									
White	41	25%	157	63%	119	66%	64%		
Hispanic	71	43%	38	15%	36	20%	16%		
Black/African-American	19	11%	25	10%	11	6%	13%	<0.0001	0.006
Asian/PI	7	4%	13	5%	8	4%	5%		
Other	22	13%	11	4%	5	3%	2%		
Missing	6	4%	4	2%	2	1%	NA		
Orientation									
Gay/Homosexual	117	70%	137	55%	138	76%			
Bisexual	24	14%	90	36%	27	15%			
Hetero/Straight	4	2%	7	3%	4	2%	NA	<0.0001	0.06
Other	16	10%	13	5%	11	6%			
Declined	4	2%	0	0%	0	0%			
Missing	1	1%	1	0%	1	1%			
Education Level									
Less than High School	5	3%	2	1%	1	1%	18%		
High School Grad	32	19%	25	10%	11	6%	30%		
Some college	68	41%	88	35%	60	33%	27%	<0.0001	0.06

	Clinic-based Arm* (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		2010 US Census***	Unadjusted p-value	Adjusted p-value
Demographics	***n***	**%**	***n***	**%**	***n***	**%**			
Bachelor's degree	47	28%	100	40%	69	38%	17%		
Graduate degree	11	7%	30	12%	39	22%	9%		
Prefer not to disclose	2	1%	2	1%	0	0%	NA		
Missing	1	1%	1	0%	1	1%	NA		
Individual Yearly Income									
Less than $15,000	28	17%	19	8%	5	3%	6%		
Between $15,000 and $24,999	30	18%	27	11%	11	6%	15%		
Between $25,000 and $34,999	15	9%	40	16%	7	4%	18%	<0.0001	<0.0001
Between $35,000 and $49,999	12	7%	54	22%	18	10%	22%		
More than $50,000	18	11%	66	27%	81	45%	40%		
Prefer not to disclose	6	4%	4	2%	7	4%	NA		
Missing	57	34%	38	15%	52	29%	NA		
Employment Status									
Employed	109	66%	210	85%	129	71%			
Not Employed	46	28%	23	9%	51	28%	NA	<0.0001	NA
Declined	10	6%	14	6%	1	1%			
Missing	1	1%	1	0%	0	0%			
City of Residence†									
Los Angeles	166	100%	70	28%	59	33%			
Chicago	0	0%	50	20%	24	13%	NA	0.24	NA
New York City	0	0%	87	35%	71	39%			
Miami	0	0%	41	17%	27	15%			
HIV Status									
Positive	9	5%	9	4%	33	18%	NA	<0.0001	0.009
Negative	140	84%	211	85%	140	77%			
Not Sure	17	10%	28	11%	8	4%			
Time from Start to Completion**									
Mean Time	Not Reported		8 minutes		9 minutes				
Standard Deviation	Not Reported		5 minutes		5 minutes				

	Clinic-based Arm* (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		2010 US Census***	Unadjusted p-value	Adjusted p-value
Demographics	***n***	**%**	***n***	**%**	***n***	**%**			
Total	166	100%	248	100%	181	100%	10000%		

*Clinic-based respondents were recruited for the survey by testing counselors at a community-based clinic in Los Angeles

**Periodically, the iPad for the clinic-based arm needed to be checked for internet connectivity. This was done by navigating from the consent screen to the first screening question and back. Although this was not counted as a response, the time counter was continuous during these checks and therefore the data for survey completion time for clinic-based respondents overestimate the actual time spent on the survey. Therefore, data are not presented.

***Calculation of each p-value did not include the 2010 US Census Data or Missing Data

†Calculation of the p-value did not include the clinic-based arm

Table 2

Bivariate Chi-Square, Fisher Exact/Kruskal-Wallis Tests and Multivariable Logistic Regression of Sexual History Differences between Clinic-based, Mechanical Turk (MTurk), and Qualtrics Survey Respondents who Identify as Gay/Bisexual Men and Passed the Attention Filter, April 2015 – March 2016.

	Clinic-based Arm* (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		Unadjusted p-value*	Adjusted p-value*
Sexual Risk Behaviors	***n***	**%**	***n***	**%**	***n***	**%**		
Number of Female Sex Partners in the Past 3 Months							<0.0001	0.36
Mean	1.15		1.94		1.54			
Standard Deviation	0.56		2.3		4.5			
Number of Male Sex Partners in the Past 3 Months							<0.0001	0.24
Mean	6.01		3.56		4.5			
Standard Deviation	6.83		3.26		8.63			
Most Common Place to Meet Sex Partners in the Past 3 Months								
Bars/Clubs	20	12.0%	131	52.8%	29	16.0%		
Bathhouses	7	4.2%	4	1.6%	15	8.3%	<0.0001	<0.0001
Websites	29	17.5%	39	15.7%	45	24.9%		
Mobile phone apps	73	44.0%	42	16.9%	35	19.3%		
Other	32	19.3%	23	9.3%	49	27.1%		
History of Gonorrhea Diagnosis								
Missing	4	2.4%	0	0.0%	0	0.0%		
No	82	49.4%	216	87.1%	142	78.5%	<0.0001	<0.0001
Yes, but over a Year Ago	33	19.9%	26	10.5%	28	15.5%		
Yes, within the Past Year	47	28.3%	6	2.4%	11	6.1%		
History of Chlamydia Diagnosis								
Missing	7	4.2%	2	0.8%	1	0.6%		
No	104	62.7%	223	89.9%	146	80.7%	<0.0001	0.31
Yes, but over a Year Ago	28	16.9%	14	5.6%	21	11.6%		
Yes, within the Past Year	27	16.3%	9	3.6%	13	7.2%		
History of Syphilis Diagnosis								
Missing	7	4.2%	1	0.4%	0	0.0%		

	Clinic-based Arm* (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		Unadjusted p-value*	Adjusted p-value*
Sexual Risk Behaviors	***n***	**%**	***n***	**%**	***n***	**%**		
No	124	74.7%	228	91.9%	147	81.2%	<0.0001	0.003
Yes, but over a Year Ago	12	7.2%	11	4.4%	22	12.2%		
Yes, within the Past Year	23	13.9%	8	3.2%	12	6.6%		
Total	166	100.0%	248	100.0%	181	100.0%	181	100.0%

*Calculation of each p-value did not include Missing Data

Table 3

Bivariate Chi-Square Tests and Multivariable Logistic Regression of Substance Use Differences between Clinic-based, Mechanical Turk (MTurk), and Qualtrics Survey Respondents who Identify as Gay/Bisexual Men and Passed the Attention Filter, April 2015 – March 2016.

	Clinic-based Arm (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics* (*n* = 181)		Unadjusted p-value**	Adjusted p-value**
Substance Use	***n***	**%**	***n***	**%**	***n***	**%**		
Cocaine/Crack Use								
Not Asked	0	0.0%	0	0.0%	181	100.0%	0.13	0.49
Missing	10	6.0%	0	0.0%	–	–		
Never Used	104	62.7%	190	76.6%	–	–		
Used Before, but over a Year Ago	23	13.9%	29	11.7%	–	–		
Used within the Past Year	29	17.5%	29	11.7%	–	–		
Crystal Methamphetamine Use								
Not Asked	0	0.0%	0	0.0%	181	100.0%		
Missing	5	3.0%	3	1.2%	–	–	<0.0001	0.05
Never Used	113	68.1%	206	83.1%	–	–		
Used Before, but over a Year Ago	14	8.4%	27	10.9%	–	–		
Used within the Past Year	34	20.5%	12	4.8%	–	–		
Ecstasy Use								
Not Asked	0	0.0%	0	0.0%	181	100.0%		
Missing	12	7.2%	3	1.2%	–	–	0.6	0.01
Never Used	90	54.2%	153	61.7%	–	–		
Used Before, but over a Year Ago	38	22.9%	50	20.2%	–	–		
Used within the Past Year	26	15.7%	42	16.9%	–	–		
GHB Use								
Not Asked	0	0.0%	0	0.0%	181	100.0%		
Missing	11	6.6%	4	1.6%	–	–	0.01	0.23
Never Used	124	74.7%	217	87.5%	–	–		
Used Before, but over a Year Ago	10	6.0%	15	6.0%	–	–		
Used within the Past Year	21	12.7%	12	4.8%	–	–		
Nitrates/Poppers Use								

	Clinic-based Arm (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics* (*n* = 181)		Unadjusted p-value**	Adjusted p-value**
Substance Use	***n***	**%**	***n***	**%**	***n***	**%**		
Not Asked	0	0.0%	0	0.0%	181	100.0%		
Missing	8	4.8%	1	0.4%	–	–	<0.0001	<0.0001
Never Used	77	46.4%	199	80.2%	–	–		
Used Before, but over a Year Ago	13	7.8%	22	8.9%	–	–		
Used within the Past Year	68	41.0%	26	10.5%	–	–		
Erectile Dysfunction Drug (without a prescription) Use								
Not Asked	0	0.0%	0	0.0%	181	100.0%		
Missing	10	6.0%	2	0.8%	–	–	0.23	0.0002
Never Used	123	74.1%	177	71.4%	–	–		
Used Before, but over a Year Ago	11	6.6%	25	10.1%	–	–		
Used within the Past Year	22	13.3%	44	17.7%	–	–		
Total	166	100.0%	248	100.0%	181	100.0%		

* p-values for substance use variables are only for the clinic-based and Mturk arms since substance use questions were not allowed on the Qualtrics panel

** Calculation of each p-value did not include Missing Data

Table 4

Bivariate Fisher Exact Tests and Multivariable Logistic Regression of PrEP and PEP Knowledge and Attitudes between Clinic-based, Mechanical Turk (MTurk), and Qualtrics Survey Respondents who Identify as Gay/Bisexual Men and Passed the Attention Filter, April 2015 – March 2016.

	Clinic-based Arm (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		Unadjusted p-value*	Adjusted p-value*
PrEP Knowledge and Attitudes	***n***	**%**	***n***	**%**	***n***	**%**		
Are you currently taking PrEP?								
Yes, I am currently on PrEP	17	10.2%	7	2.8%	9	5.0%		
No, I've never taken PrEP	124	74.7%	223	89.9%	135	74.6%	0.0002	NA
Yes, but I am no longer taking PrEP	16	9.6%	9	3.6%	4	2.2%		
Missing	9	5.4%	9	3.6%	33	18.2%		
How familiar would you say you are with PrEP?								
Know a fair amount or a lot about it	51	30.7%	18	7.3%	35	19.3%	<0.0001	0.46
Does not know about it	115	69.3%	230	92.7%	146	80.7%		
Based on what you've heard, how effective is PrEP in preventing HIV infection?**								
More than 90% effective	41	24.7%	2	0.8%	20	11.0%	<0.0001	<0.0001
Less than 90% effective	8	4.8%	13	5.2%	11	6.1%		
I don't know	2	1.2%	3	1.2%	4	2.2%		
Suppose that PrEP is at least 90% effective in preventing HIV when taken daily. How likely to use? (Willingness to Use)								
I would definitely take it	70	42.2%	71	28.6%	39	21.5%	<0.0001	0.08
I would probably take it	18	10.8%	54	21.8%	28	15.5%		
I might take it	25	15.1%	68	27.4%	35	19.3%		
I would probably not take it	10	6.0%	15	6.0%	21	11.6%		
I would definitely not take it	1	0.6%	15	6.0%	12	6.6%		
Missing	42	25.3%	25	10.1%	46	25.4%		
PrEP is currently available with a prescription from your doctor, will you start taking it? (Likelihood to Use)								
Yes, I definitely will begin taking PrEP	29	17.5%	36	14.5%	11	6.1%		
Yes, I will probably begin taking PrEP	28	16.9%	54	21.8%	20	11.0%	0.0001	0.4
I'm not sure, I might begin taking PrEP	43	25.9%	95	38.3%	57	31.5%		

	Clinic-based Arm (*n* = 166)		MTurk Arm (*n* = 267)		Qualtrics (*n* = 181)		Unadjusted p-value*	Adjusted p-value*
PrEP Knowledge and Attitudes	***n***	**%**	***n***	**%**	***n***	**%**		
No, I probably will not begin taking PrEP	19	11.4%	21	8.5%	32	17.7%		
No, I definitely will not begin taking PrEP	4	2.4%	17	6.9%	15	8.3%		
Missing	43	25.9%	25	10.1%	46	25.4%		
Total	166	100.0%	248	100.0%	181	100.0%		

*Calculation of each p-value did not include Missing Data

**Only analyzed for individuals who stated that they “Know a fair amount or a lot about it”