Statistical Analysis Plan (SAP)

The creation of this template for the Duke BERD Methods Core was made possible by Grant Number UL1TR002553 from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research.

Title	**Relative Effectiveness of Social Media, Dating Apps, and Information Search Sites in Promoting HIV Self-testing: Observational Cohort Study**
CRU/Department/Division/Center	Department of Medicine, UCLA/CTN
IRB Number	18-001580
Investigators:	
Primary Investigator	Jeffrey D. Klausner, MD, MPH
Collaborative Lead	Lisa A. Marsch
Co-authors (if known)	Chrysovalantis Stafylis, MD, MPH; Gabriella Vavala, BA; Qiao Wang, MPH; Bethany McLeman, BA; Shea M Lemley, PhD; Sean D Young, PhD; Haiyi Xie, PhD; Abigail G Matthews, PhD; Neal Oden, PhD; Leslie Revoredo, PhD; Dikla Shmueli-Blumberg, PhD; Emily G Hichborn, BSc; Erin McKelle, MA; Landhing M Moran, PhD; Petra Jacobs, MD, MHS
Analysis Biostatistician(s)	
Biostatistics Supervisor	
Lead Biostatistician	Haiyi Xie, PhD
Subject Matter Expert	Jeffrey D. Klausner, MD, MPH; Lisa A. Marsch, PhD
Original Creation Date	11.05.2019
Version Date	05.26.2020
Project Folder Location	
Project Goal(s)	Comparing the effectiveness of HIV self-testing promotion in social media sites versus informational sites versus dating sites
Submission Deadline(s)	09.25.2020
Effort Estimate (optional)	

Investigator Agreement ☐ All statistical analyses included in an abstract or manuscript should reflect the work of the biostatistician(s) listed on this SAP.

No changes or additional analyses should be made to the results or findings without discussing with the project biostatistician(s).

☐ All biostatisticians on this SAP should be given sufficient time to review the full presentation, abstract, manuscript, or grant and be included as co-authors on any abstract or manuscript resulting from the analyses.

☐ If substantial additional analysis is necessary or the aims of the project change, a new SAP will need to be developed.

☐ Publications resulting from this SAP are supported in part by the Duke CTSA and must cite grant number UL1TR002553 and be submitted to PubMed Central.

☐I have reviewed the SAP and understand that any changes must be documented.

Acknowledged by: Click or tap here to enter text.
Date: Click or tap to enter a date.

Activity Log

Acronyms

AE	Adverse Effect
AHF	AIDS Healthcare Foundation
AIDS	Acquired Immunodeficiency Syndrome
CAPA	Corrective and Preventative Action
CCC	Clinical Coordinating Center
CCTN	Center for the Clinical Trials Network
CDC	Centers for Disease Control and Prevention
CoC	Certificate of Confidentiality
CRF	Case Report Form
CTN	Clinical Trials Network
DSC	Data and Statistics Center
DSMB	Data and Safety Monitoring Board
eICF	Electronic Informed Consent Form
ETR	Education, Training and Research
FDA	Food and Drug Administration
GBMMS	Group-Based Medical Mistrust Scale
GCP	Good Clinical Practice
HAART	Highly Active Antiretroviral Therapy
HHS	Department of Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act
HIV	Human Immunodeficiency Virus
HOPE	Harnessing Online Peer Education
IRB	Institutional Review Board
LN	Lead Node
MOP	Manual of Operating Procedures
MSM	Men who have sex with men
NIH	National Institutes of Health
NIDA	National Institute on Drug Abuse
OHRP	Office for Human Research Protections

PI	Principal Investigator
PRB	Protocol Review Board
PrEP	Pre-Exposure Prophylaxis
QA	Quality Assurance
RHBA	Rapid HIV Behavioral Assessment
SAE	Serious Adverse Event
SAMHSA	Substance Abuse and Mental Health Services Administration
SOP	Standard Operating Procedure
SSO	Single Sign On
STD	Sexually Transmitted Disease
STI	Sexually Transmitted Infection
TAPS	Tobacco, Alcohol, Prescription medication, and other Substance use
TTM	Transtheoretical Model
TDF-FTC	Emtricitabine - tenofovir disoproxil fumarate
UCLA	University of California, Los Angeles
WHO	World Health Organization
YTH	Youth Technology and Health

1 Study Overview

Background/Introduction:
Despite accounting for a small fraction of the U.S. population (6%), men who have sex with men (MSM) account for a majority of those infected with HIV and experience the largest burden of new infections1. Those 18-30 years of age account for approximately two-thirds of all new HIV infections, and a vast majority of these infections result from male-to-male sexual contact. Nearly half of those individuals do not know that they are infected. It is essential that prevention efforts directly focus on MSM as a targeted high-risk group to prevent new infections by facilitating HIV testing and Pre-Exposure Prophylaxis (PrEP) uptake.

HIV testing and use of PrEP medication are two of the most promising methods for reducing HIV transmission rates. HIV testing leading to treatment can reduce the transmission of infection by 95%. Aside from the 'traditional' point-of-care laboratory or rapid tests, the World Health Organization (WHO) has endorsed the use of self-testing to increase awareness of HIV infection status. The OraQuick® Home HIV test kit [OraSure Technologies, Inc., PA] is an FDA-approved self-test kit that detects HIV antibodies in oral fluid. The kit is commercially available in pharmacies and online, and it has been proven to be acceptable and user-friendly. The research team has utilized this kit in previous screening programs.

Transmission rates can further be reduced with PrEP. PrEP is the combination of two antiviral drugs (tenofovir–emtricitabine; TDF–FTC) in the form of a daily pill that prevents new infections. Randomized controlled trials in high-risk populations have shown that infection rates with PrEP use drop dramatically, as low as 0%. The Centers for Disease Control and Prevention (CDC) recommends prescribing PrEP to all adults with substantial risk for HIV infection. Nevertheless, PrEP awareness and most importantly uptake has been slow, especially among young (ages 18- 24) Black-and-Latino MSM.

Substance use is prevalent among young MSM and is associated with condomless anal sex, a key risk factor for new HIV infection. Further, Patel et al. reported that some young Black and Latino MSM used social media to exchange sex for money or drugs, behavior that increases risk of HIV transmission. Substance use is also associated with this high-risk population's likelihood of HIV testing and PrEP uptake. Among young, Black MSM, binge drinking is associated with decreased odds of recent HIV testing, while heavy marijuana use is associated with increased likelihood of unknown HIV infection. Substance use is prevalent among MSM users of social media. In a study among 118 Latino and Black MSM Facebook users, we reported high prevalence of marijuana, methamphetamine and cocaine use15. Among internet-using MSM, only 85% had ever tested for HIV and only 58% had tested in the last year, and both younger age and substance use are associated with lower likelihood of testing, indicating an unmet need for HIV testing promotion among these subpopulations.

The exponential popularity of social networking sites has drawn scholarly attention in recent years, and rapidly growing efforts have been made in applying these websites to behavioral health interventions. Nearly 7 out of 10 Americans use social media to connect with one another, engage with news content, share information and entertain themselves. Compared to all other age groups, regardless of sexual preferences, those between 18-30 years are the age group that is most likely to be active and engaged in daily social media use. Facebook is the most popular social media platform among other platforms, including Instagram and Twitter. Social media is also a new way to meet sex partners. Young MSM frequently use social media, including dating- centered applications ("apps") like Grindr, to connect with new sex partners. In a recent study by Goedel et. al. among 92 young gay and bi-sexual men, participants reported that they used Grindr to find sexual partners, had accounts in multiple other dating apps, and spent nearly one hour a day using these apps.

To our knowledge, it is not yet known whether leveraging social media sites or dating apps—the current online methods for recruiting and engaging HIV participants—is the more effective method for health promotion among participants at risk for HIV and related substance use. Because users may be in a more action-oriented stage of change when seeking HIV information on informational sites such as Google, they might be more ready to engage in health services. It is possible that these individuals could be more likely to click on an HIV self-testing-focused ad as well as actually seek testing and PrEP services. Although substance use is associated with lower rates of HIV testing among MSM, how substance use impacts likelihood of HIV home self-testing when MSM are recruited online is underexplored. This study seeks to address these questions.
(Protocol, Section 4.0 Introduction)

A. Study Aims

This study seeks to evaluate the effectiveness of an online campaign promoting HIV self-testing on different web platforms: social media sites (Facebook, Instagram, Twitter); dating apps (Grindr/ alternative, Hornet, Jack'd); and informational sites (Google, Bing, Yahoo). (Protocol, Section 4.0 Introduction, Study Rationale)

B. Study Hypotheses

The study will evaluate the relative effectiveness of using social media sites (i.e., Facebook, Instagram, Twitter), informational sites (i.e., Google, Bing, Yahoo), and dating apps (i.e., Grindr/alternative, Hornet, Jack'd) to promote self-testing of HIV infection among MSM who are at increased risk of HIV exposure and/or infection.

C. Primary Hypotheses

H0: The rate of HIV home self-test kits orders per day across the platforms (social media, dating apps, and information search sites) is equal.
H1: The rate of HIV home self-test kits orders per day differs at least by one platform.

D. Secondary Hypotheses

H0: The PrEP uptake across all three platforms is equal.
H1: The PrEP uptake differs at least by one platform.

H0: The ad conversion rate is equal across all platforms.
H1: The ad conversion rate differs at least by one platform.

2 Study Population

2.1 Inclusion Criteria

- Have clicked on one of the study-specific advertisements posted on the platforms/ websites described in the protocol;
- Have been biologically born male (cis-gender man), per participant self-report;
- Report condomless anal intercourse and more than one male sex partner in the 90 days prior to the date of the screening questionnaire;
- Be between the ages of 18-30 years old, inclusive;
- Self-identify as Latino and/or Black/African American;
- Not currently on PrEP and haven't taken PrEP in the last six months prior to the date of the screening questionnaire (per participant self-report);
- Have not tested for HIV in the last 3 months prior to the date of the screening questionnaire (per participant self-report);
- Have a Facebook account (for identity validation to reduce duplicate attempts at enrollment); and
- Be willing to provide contact information (phone number, email) to the study team.

E. Exclusion Criteria

- Are unwilling or unable to provide informed consent.
- Are unwilling to provide contact information (phone number, email address).
- Report having a preliminary positive or positive HIV result in a test completed less than 30 days prior to the date of screening or report being currently under treatment for HIV infection.

2.2 Data Acquisition

Fill in all relevant information:

Study design	Observational model, non-randomized cohort study, time perspective
Data source/how the data were collected	All study data will be captured electronically from participants via the Qualtrics survey site managed by the study team at UCLA. Advertising performance information will be recorded automatically by the websites (ETR) and reported on an approximately bi-weekly basis.
Contact information for team member responsible for data collection/acquisition	N/A

Date or version (if downloaded, provide date)	NA
Data transfer method and date	NA
Where dataset is stored	UCLA servers

3 Outcomes, Exposures, and Additional Variables of Interest

3.1 Primary Outcome(s)

Outcome	**Description**	**Variables and Source**	**Specifications**
number of HIV home self-test kits ordered per day by promotional platform	The primary outcome will be measured by the number of home HIV test kits ordered through the electronic study system by promotional platform. These data will allow the study team to link individual kit orders to study participants, and to promotional platform.		e.g., how is it coded?

3.2 Secondary Outcome(s)

Outcome	**Description**	**Variables and Source**	**Specifications**
number of participants who used the study-provided HIV home self-test kit by 60- day follow-up	The number of participants who used the study-provided HIV home self-test kit by 60-day follow-up will be collected via self-report at		e.g., how is it coded?

	the time of follow-up (either 14-day or 60-day)		
number of participants who tested positive for HIV using the study-provided home test kit by 60-day follow-up	The number of participants who tested positive for HIV using the study-provided home test kit will be collected via self-report and/or sent as an image to the study team, at the time of follow-up (either 14-day or 60-day).		
number of participants who reported a high-risk substance abuse profile at baseline who ordered an HIV home self-test kit	The number of participants who reported a high-risk substance abuse profile at baseline who ordered an HIV home self-test kit. Participants' substance abuse profiles will be collected via the TAPS Tool at baseline. These data will be used to characterize the sample, examine differences across platform types, and determine how substance abuse modifies the primary outcome of HIV testing.		
number of participants who reported a high-risk substance	The number of participants who reported a high-risk substance abuse		

abuse profile at baseline who reported seeking PrEP services or started PrEP at either follow-up	profile at baseline who reported seeking PrEP services or started PrEP at either follow-up. Similarly, participants' substance abuse profiles will be collected via the TAPS Tool at baseline. These data will be used to characterize the sample, examine differences across platform types, and determine how substance abuse modifies the secondary outcome of PrEP uptake		
number of participants in the "Determination" stage of change at baseline who ordered an HIV home self-test kit	The number of participants in the "Determination" stage of change at baseline who ordered an HIV home self-test kit. Participants' stages of change or readiness for HIV testing will be collected via the Transtheoretical Model (Stage of Change) tool at baseline. These data will be used to characterize the sample, examine differences across platform		

	types and determine how stage of change modifies the primary outcome of HIV testing		
number of participants in the "Determination" stage of change at baseline who reported seeking PrEP services or started PrEP at either follow-up	The number of participants in the "Determination" stage of change at baseline who reported seeking PrEP services or started PrEP at either follow-up. Participants' stages of change or readiness for HIV testing will be collected via the Transtheoretical Model (Stage of Change) tool at baseline. These data will be used to characterize the sample, examine differences across platform types and determine how stage of change modifies the secondary outcome of PrEP uptake.		
number of participants who tested positive for HIV using the study-provided home test kit and were linked to care within 30 days after the self-reported test date.	The number of participants who tested positive for HIV using the study-provided home test kit and were linked to care within 30 days after self-testing		

	will be collected via self-report in a study-designed questionnaire at the time of follow-up (14-day or 60-day).		
number of participants who tested positive for HIV using the study-provided home test kit, were linked to care and started treatment within 30 days after self-testing.	The number of participants who tested positive for HIV using the study-provided home test kit and started treatment within 30 days after self-testing will be collected via self-report in a study-designed questionnaire at the time of follow-up (14-day or 60-day).		
number of participants who tested negative for HIV using the study-provided home test kit and sought PrEP services within 30 days after self-testing.	The number of participants who tested negative for HIV using the study-provided home test kit and sought PrEP services within 30 days after self-testing will be collected via self- report in a study-designed questionnaire at the time of follow-up (14-day or 60-day).		
The number of participants who tested negative for HIV using the study-provided home test kit,	The number of participants who tested negative for HIV using the study-provided home test		

sought PrEP services and started PrEP within 30 days after self-testing.	kit and started PrEP in 30 days after self-testing will be collected via self-report in a study-designed questionnaire at the time of follow-up (14-day or 60-day).		
amount of money spent per test kit ordered per promotion type, including all the costs of the intervention (advertisement, test kit)	The amount of money spent per test kit ordered per platform type, including all the cost of the intervention (advertisement, test kit), will be collected via website metrics (i.e., conversion), data from the follow up (14-day or 60-day), and internal tracking of the expenses related to the implementation of the intervention.		
proportion of people from each website and the proportion of people from each type of website who clicked on advertisements	The proportion of people from each website and the proportion of people from each type of website who clicked on advertisements, will be collected via website metrics from each website.		
Kit ordering rates expressed as kits	Kit ordering rates expressed		

ordered per unique participants visiting the study informational page on Qualtrics	as kits ordered per unique participants visiting the study informational page on Qualtrics. This can be calculated using the same approach as in the primary analysis, except that we will use number of visiting participants rather than time as the offset. As in the primary analysis, we assume that it will be rare for participants to order multiple copies of kits, although visits to multiple websites may be more common. If there are one or more orders within a wave, we will use only the first order, together with its website. If there are multiple websites with no orders, we will use the first website.		
The number of participants with high rate of sexual delay discounting at	The number of participants with high rate of sexual delay discounting at		

baseline who ordered an HIV home self-test kit.	baseline who ordered an HIV home self-test kit will be collected via the Sexual Discounting Task tool at baseline. These data will be used to characterize the sample, examine differences across platform types, and in moderator analyses to determine how substance abuse modifies the primary outcome of HIV testing and secondary outcome of PrEP uptake.		

3.3 Additional Variables of Interest

Variable	**Description**	**Variables and Source**	**Specifications**
			e.g., how is it coded?

4 Statistical Analysis Plan

4.1 Demographic and Clinical Characteristics ("Table 1")

Participant's baseline demographics and characteristics will be presented using summary statistics. Descriptive summaries of the distribution of continuous variables will be presented with percentiles (median, 25th and 75th percentiles), and with mean and standard deviation. Categorical variables will be summarized in terms of frequencies/counts and percentages

4.2 Analyses Plan for Aim 1

The analysis model will be a Poisson regression model that will use time as an offset.

4.3 Analyses Plan for Aim 2

Secondary and exploratory analyses will be conducted using univariate tests and descriptive statistics such as counts, percentages, and 95% confidence intervals.
ETC

5 Limitations

- The study was conducted in 9 areas with high HIV incidence; thus, the conclusions may not be generalizable to the whole country;
- Low enrollment and participation in waves affected the capacity to make broader comparisons between platforms and potentially between sites;
- Only the most popular apps were selected; thus, the findings are specific to the sites included in the campaigns.

6 Addendum for Additional Analyses

[insert if applicable]
Secondary and exploratory analyses will be conducted using univariate tests and descriptive statistics such as counts, percentages, and 95% confidence intervals.

7 Appendix

Table 1. Summary of National Institute on Drug Abuse Clinical Trials Network Social Media Pre-exposure Prophylaxis Study, 2020, population

Characteristic	Value
Age in years, median (IQR)	
Ethnicity, n (%)	
Hispanic/Latinx	
Race, n (%)	
American Indian or Alaskan Native	
Black or African American	
White	
Other	
Multiracial	
History of PrEP[a] uptake, n (%)	
Never taken PrEP	
In the past 6 months	
Number of male sex partners in the past 90 days, median (IQR)	
Condom use, n (%)	
Never	
Sometimes	
About half the time	
Most of the time	
Always	
Condomless receptive anal sex in the past 90 days, n (%)	
Ever tested for HIV during lifetime, n (%)	
If tested for HIV, median (IQR)	
Months since last HIV test	
If not tested for HIV, n (%)	
Main reasons cited by 63 participants for not getting tested, n (%)	
Unlikely to be exposed to HIV	
Afraid of testing HIV-positive	
Did not want to think about HIV/HIV-positive	
Worried about names being reported if positive	

Dislike for needles	
Unable to trust that the results will be confidential	
Unaware of where to get started	
Other reasons	

Table 2. Number and rate of HIV home self-test kits ordered through promotional platforms by wave per protocol sample in the National Institute on

Type of platform	Wave	Number of days for each wave	Number of test kits ordered	Order rate (ordered test kits/day)
Social media site				
Facebook	N/A	N/A	N/A	N/A
Instagram	N/A	N/A	N/A	N/A
Subtotal	N/A	N/A	N/A	N/A
Dating app	N/A	N/A	N/A	N/A
Grindr	N/A	N/A	N/A	N/A
Jack'D	N/A	N/A	N/A	N/A
Subtotal	N/A	N/A	N/A	N/A
Information search site				
Google	N/A	N/A	N/A	N/A
Bing	N/A	N/A	N/A	N/A
Subtotal	N/A	N/A	N/A	N/A
Total	N/A	N/A	N/A	N/A

Table 3. Performance of advertisements by platform throughout the advertisement campaign in the National Institute on Drug Abuse Clinical Trials Network Social Media PrEP Study, 2020

			Click-through rate			
Platform	Impressions	Clicks	(%)[c]	Users screened	Enrolled Participants	Total funds spent (US$)
Social Media	N/A	N/A	N/A	N/A	N/A	N/A
Dating Apps	N/A	N/A	N/A	N/A	N/A	N/A
Information search sites	N/A	N/A	N/A	N/A	N/A	N/A

Table B1: Primary outcome sensitivity analysis including Validated participants (n=271). Wave 3 was not included in the analytic model

Platform	Site	Wave**	Days	Test Kits Ordered	Order rate
Informational					
	Google	1	N/A	N/A	N/A
	Bing	2	N/A	N/A	N/A

Social media					
	Facebook	1	N/A	N/A	N/A
	Instagram	2	N/A	N/A	N/A
Dating					
	Grindr	1	N/A	N/A	N/A
	Jack'd	2	N/A	N/A	N/A
Total					

Table B3: Primary outcome sensitivity analysis including participants enrolled prior March 13th, 2020 to assess the impact of the COVID19 pandemic.

Platform	Site	Wave	Days	Test Kits Ordered	Test Kit Ordered
Informational					
	Google	1	N/A	N/A	N/A
	Bing	2	N/A	N/A	N/A
Social media					
	Facebook	1	N/A	N/A	N/A
	Instagram	2	N/A	N/A	N/A
Dating					
	Grindr	1	N/A	N/A	N/A
	Jack'd	2	N/A	N/A	N/A
Total					

a. **Tobacco, alcohol, prescription medications, and other substance use**						
		Ordered test kit (n=177) n/N	**Percent**	**Did not order test kit (n=77) n/N**	**Percent**	**P-value[1]**
Alcohol	None	NA	NA	NA	NA	NA
	Problem use[2]	NA	NA	NA	NA	NA
	High Risk Substance Use	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
Cannabis	None	NA	NA	NA	NA	NA
	Problem use	NA	NA	NA	NA	NA
	High Risk Substance Use	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
Stimulants	None	NA	NA	NA	NA	NA
	Problem Use/High Risk Substance Use[3]	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
Opioid	None	NA	NA	NA	NA	NA
	Problem Use/ High Risk Substance Use[3]	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
Sedative	None	NA	NA	NA	NA	NA
	Problem use/High Risk Substance Use[3]	NA	NA	NA	NA	NA
	Missing	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
Prescribed stimulant	None	NA	NA	NA	NA	NA
	Problem use/High Risk Substance Use[3]	NA	NA	NA	NA	NA
	Missing	NA	NA	NA	NA	NA

b. Stage of Health Behavior Change						
		Ordered test kit (n=177) n/N	**Percent**	**Did not order test kit (n=77) n/N**	**Percent**	**P-value[1]**
Precontemplation	*I do not see any need to regularly test for HIV*	NA	NA	NA	NA	NA
Contemplation	*I think I should get tested for HIV regularly, but I am not sure*	NA	NA	NA	NA	
Determination	*I am ready to start getting regularly tested for HIV*	NA	NA	NA	NA	
Action	*I am trying to get tested regularly for HIV*	NA	NA	NA	NA	
Maintenance	*I have been getting tested for HIV regularly over the past few years*	NA	NA	NA	NA	

c. Attitudes toward human immunodeficiency virus (HIV) testing						
		Ordered test kit (n=177) n/N	**Percent**	**Did not order test kit (n=77) n/N**	**Percent**	**P-value[1]**
Getting tested for HIV helps people feel better	Agree	NA	NA	NA	NA	NA

	Disagree	NA	NA	NA	NA	NA
Getting tested for HIV helps people from getting HIV	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
People in my life would leave if I had HIV	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
People who tested positive for HIV should hide it from others	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
I would rather not know if I have HIV	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	

d. Attitudes toward human immunodeficiency virus (HIV) treatment (continuous scale from 1 [strongly disagree] to 7 [strongly agree])

Statements	Ordered test kit (n=177) Mean (SD)	Did not order test kit (n=77) Mean (SD)	P-value (*Wilcoxon rank test*)
I am less threatened by the idea of being HIV positive than I used to be	NA	NA	NA
I am less worried about HIV infection than I used to be	NA	NA	NA
I think HIV/AIDS is less of a problem than it used to be	NA	NA	NA
I think HIV/AIDS is a less serious threat than it used to be because of new HIV/AIDS treatments	NA	NA	NA
I am much less concerned about becoming HIV positive myself because of new HIV/AIDS treatments	NA	NA	NA
I think that condom use during sex is less necessary now that new HIV/AIDS treatments are available	NA	NA	NA
I think that someone who is HIV positive now needs to care less about condom use	NA	NA	NA
I think that the need for condom use is less than it used to be, because you can always start new treatments	NA	NA	NA
I think that someone who is HIV positive and uses new HIV/AIDS treatments can be cured	NA	NA	NA
I think that new HIV/AIDS treatments can eradicate the virus from your body	NA	NA	NA

e. Human immunodeficiency virus (HIV)-related stigma among study participants

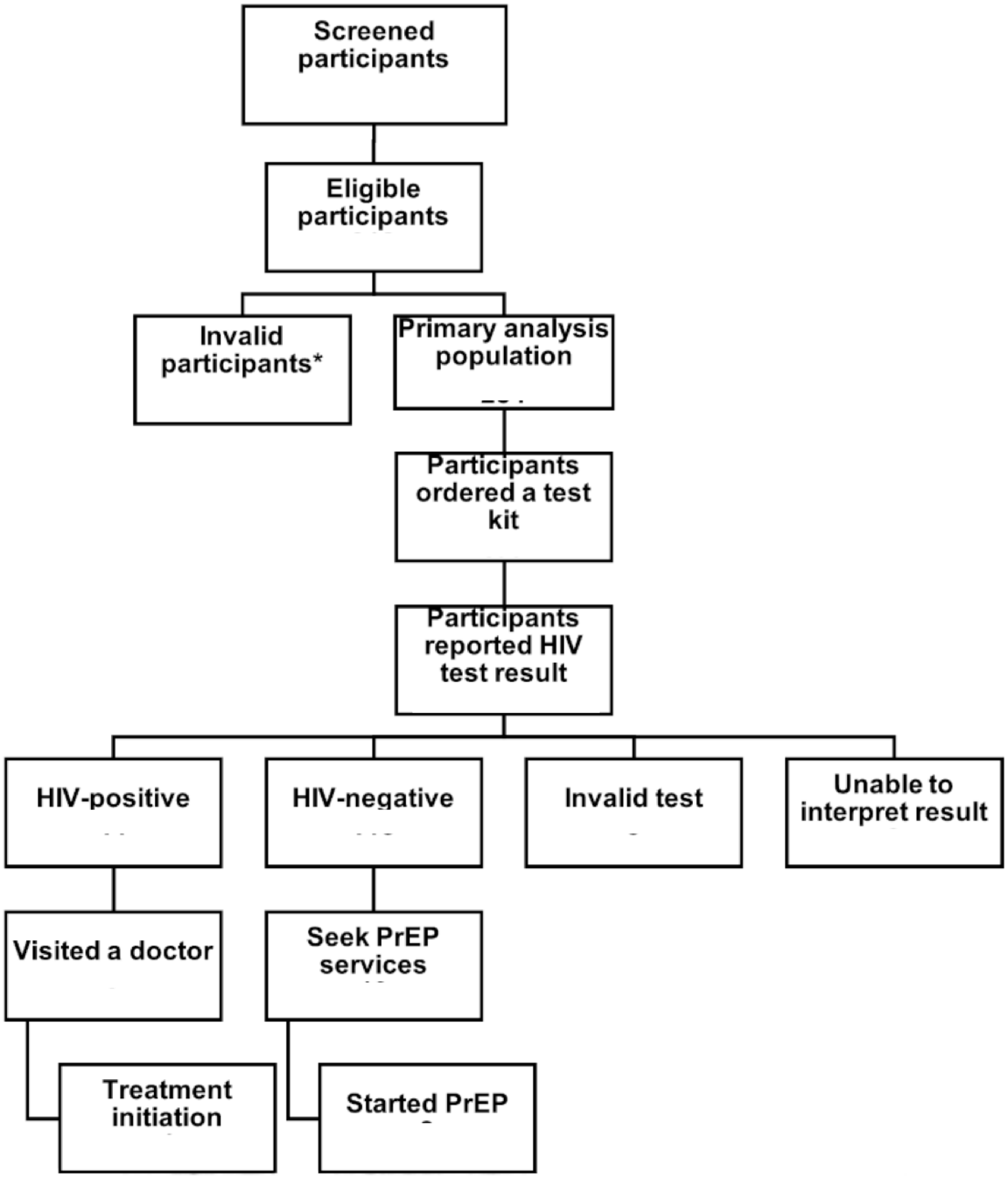
		Ordered test kit (n=177) n/N	Percent	Did not order test kit (n=77) n/N	Percent	P-value[1]
I feel afraid of people living with HIV/AIDS	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Somewhat agree	NA	NA	NA	NA	NA
	Neither agree nor disagree	NA	NA	NA	NA	NA
	Somewhat disagree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
	Strongly agree	NA	NA	NA	NA	NA

	Agree	NA	NA	NA	NA	NA
	Somewhat agree	NA	NA	NA	NA	NA
I could not be friends with someone who has HIV/AIDS	Neither agree nor disagree	NA	NA	NA	NA	NA
	Somewhat disagree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
People who get HIV/AIDS through sex or drug use got what they deserve	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Somewhat agree	NA	NA	NA	NA	NA
	Neither agree nor disagree	NA	NA	NA	NA	NA
	Somewhat disagree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
I feel anger toward people with HIV/AIDS	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Somewhat agree	NA	NA	NA	NA	NA
	Neither agree nor disagree	NA	NA	NA	NA	NA
	Somewhat disagree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA

f. Medical mistrust						
		Ordered test kit (n=177) n/N	**Percent**	**Did not order test kit (n=77) n/N**	**Percent**	**P-value[1]**
You'd better be cautious when dealing with health care organizations	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
Patients have sometimes been deceived or misled by health care organizations	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
When health care organizations make mistakes they usually cover it up	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
Health care organizations have sometimes done harmful experiments on patients without their knowledge	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
Health care organizations don't always keep your information totally private	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA

Sometimes I wonder if health care organizations really know what they are doing	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA
Mistakes are common in health care organizations	Strongly agree	NA	NA	NA	NA	NA
	Agree	NA	NA	NA	NA	NA
	Disagree	NA	NA	NA	NA	NA
	Strongly disagree	NA	NA	NA	NA	NA

Figure 1. Enrollment, HIV home test kit use, pre-exposure prophylaxis uptake, and linkage to care among participants of the National Institute on Drug Abuse Clinical Trials Network Social Media PrEP Study, 2020. PrEP: pre-exposure prophylaxis. *Invalid participants include duplicate entries, fake accounts, and participants outside the country.



References

- Statistical Analysis Software (version 9.4, SAS Institute)
- Chiu YF. Multiple comparisons and multiple tests. Using the SAS system. Peter H. Westfall, Randall D.Tobias, Dror Rom, Russell D. Wolfinger and Yosef Hochberg, SAS Institute, Cary, U.S.A. 2000. No. of pages: xiv + 397. Price: DKK 412.00. ISBN 1-58025-397-0. Statist Med 2002 May;21(10):1499-1500. [doi: 10.1002/sim.1168]

Statistical Analysis Plan Checklist

Below you will find a checklist of recommended items to include in a statistical analysis plan. Some of these are specific to clinical trials (based on this JAMA paper) and some are other are specific to observational studies (based on STROBE/RECORD guidelines), so every item will not be necessary for every project. The biostatistician should start with the SAP template above and add in necessary information from the checklist. Item numbers that are starred (*) are not explicitly included in the SAP template and should be added by the author if relevant to the project. This checklist was developed using the CONSORT 2010 Checklist.

Section/Topic	Item #	Description	Included (Yes/No/NA)
Administrative Information			
Study Information	1a	Descriptive title that matches the protocol, with SAP either as a forerunner or subtitle	Yes
	1b	Trial registration number, protocol version number, and/or IRB number.	Yes
	1c	CRU/Department/Division/Center/other collaborative unit that the study falls under	Yes
Roles and	2a	Listing of principal investigators, clinical leads, and co-authors (if known)	Yes
responsibility	2b	Name and affiliation of SAP author(s)	NA
	2c	Names, affiliations, and roles of other SAP contributors (e.g. senior statistician)	NA
SAP Information	3	SAP version number, with date of current version and original creation date	Yes
Project Information	4a	Project folder location	NA
	4b	Project goals (e.g. manuscript, abstract, presentation, etc.)	Yes
	4c	Project deadlines (of listed goals)	NA
	4d	Effort estimate	NA
Investigator Agreement			
Investigator Agreement	5	Confirmation that BERD Method Core's collaborative process has been reviewed, that all statistical analyses included in an abstract or manuscript should reflect the SAP, no changes should be made to the SAP without discussing with the SAP author, all biostatisticians on the SAP are co-authors on the manuscript, and that publications resulting from the SAP must cite grant number UL1TR002553 and be submitted to PubMed Central	No
Signatures	6	Signatures of SAP author, senior statistician, and principal investigator(s)	No
Activity Log			
SAP revisions	7a	SAP revision history with dates	NA
	7b	Justification for each SAP revision	NA
	7c*	Timing of SAP revision in relation to any interim analyses or submissions	NA

[1]Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev*. 2012;11:MR000030. Published 2012 Nov 14. doi:10.1002/14651858.MR000030.pub2

Study Overview			
Background and introduction	8	Synopsis of scientific background and rationale for the study	Yes
Aims and Hypotheses	9a	List of all scientific aims/objectives of the study, with specifications of primary, secondary, etc.	Yes
	9b	List of all statistical hypotheses (corresponding to the scientific aims), with specifications of primary, secondary, etc.	Yes
Variables of Interest	10a	List of all outcome/endpoint variables, with a description of their coding/units, timing, and source, corresponding to the statistical hypotheses. If any variables are defined using ICD or CPT codes, list them out.	NA
	10b	List of all exposure variables, with a description of their coding/units, timing, and source, corresponding to the statistical hypotheses. If any variables are defined using ICD or CPT codes, list them out.	NA
	10c	List of any additional variables of interest (e.g. covariates, potential confounders, effect modifiers, etc.) in the analysis	NA
	10d*	Location of data dictionary (or provided as an appendix)	NA
	10e*	Report category boundaries if continuous variables are collapsed into categories, and describe any other relevant data transformations	No
Causal Graph	11*	May be helpful to include a DAG or other graph/diagram that describes the way the variables of interest are presumed to relate to each other	NA
Study Methods			
Study Plan and Design	12a	Description of the study design (e.g. parallel group randomized trial, case-control study, cohort study, etc.)	Yes
	12b*	Study setting, location, and relevant dates (e.g. periods of enrolment, exposure, follow-up, and collection)	Yes
	12c*	Description of intervention or exposure groups, with allocation ratios, and details of any matching criteria	No
	12d*	Details on randomization (e.g. stratification factors) and blinding procedures	No
	12e	List of eligibility and/or inclusion/exclusion criteria	Yes
	12f*	Description of screening/enrolment/recruitment processes	Yes
	12g*	Description of patient flow (e.g. CONSORT diagram)	Yes
	12h*	Description of analysis population (e.g. intention to treat, per protocol, etc.)	No
	12i*	Definitions of adherence/compliance, protocol deviations, loss-to-follow-up, adverse events, etc.	No
	12j*	Time points at which outcomes are measured	Yes

	12k*	Timing of final analyses (are all outcomes analysed collectively, or will short-term outcomes be analysed separately from long-term outcomes, etc.)	Yes
Sample Size	13a*	Sample size calculation or justification (either provided in full or summarized, with link to original source)	No
	13b*	Description of pre-planned subgroup analyses, power for these analyses, and planned multiple comparison adjustment procedures	No
Interim Analyses	14a*	Description of what interim analyses will be conducted at which time points, and what methods used to adjust significance levels due to the interim analysis	No
	14b*	Details of any guidelines (e.g. safety, futility) for stopping the study early	No
	14c*	Details of any changes to trial design due to interim analyses (e.g. enrolling more patients)	No
Data	15a	Description of data collection/acquisition process, with contact information for team member responsible	Yes
	15b	Description of data flow/transfer from primary data collection through to creation of final analysis dataset	No
	15c	Data transfer method and date	NA
	15d	Folder location where datasets are stored	NA
	15e*	Description of any additional data management, quality control, or processing undertaken	Yes
	15f*	If any data are extracted from a database, a description of the database and the query used for the extraction, and whether/how it was merged with any data from outside that database. If the study involved linkage of databases, consider use of a flow diagram to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	No
	15f*	Description of any other data sources incorporated in the analysis	No
Missing Data	16a*	Description of sources and magnitudes of missing data	No
	16b*	Description of how missing data patterns will be presented/summarized (may be helpful to have a table shell or draft CONSORT-style diagram)	No
	16c*	Description of contingency plans for handling missing data in analysis	No
Simulations	17a*	If conducting a simulation, a description of the purpose of the simulation and its design (e.g. fully factorial, partially factorial, grid search, etc.)	No
	17b*	Define the fixed and variable factors or parameters in the simulation, the estimands/targets of the simulation, and the performance measures to be estimated (with justifications of their relevance to the estimands/targets)	No

	17c*	Description of the tabular and graphical presentations of simulation results and their interpretation	No
Statistical Analysis Plan			
Statistical Significance	18a*	Hypothesis testing framework (e.g. superiority, equivalence, non-inferiority), or description of alternative analytic framework (e.g. evaluation of a posterior in a Bayesian analysis, etc.)	Yes
	18b*	Level of significance for primary hypotheses, including a description and rationale for any multiple comparisons adjustment or Type I error control procedures	NA
	18c*	Description of any decision-making rules based on confidence intervals, credible intervals, prediction intervals, Bayes' factors, or other alternative inferential methods	Yes
	18d*	Description of how the results of any hypothesis tests (or alternative inferential methods) will be interpreted with respect to both the statistical hypotheses and scientific aims/objectives of the study	
Descriptive Statistics	19a*	List of characteristics (e.g. demographic, clinical) to be summarized descriptively (e.g. "Table 1")	Yes
	19b*	Description of how these characteristics will be summarized descriptively (e.g. means/medians vs. N (%), tabular displays, graphical displays, etc.)	Yes
	19c*	Summarize follow-up time (e.g. average and total amount) and number of events	Yes
Analysis Methods	20a	For each aim/hypothesis (see items 9a/9b), a description of what analysis method will be used and how the results from this method will be reported and interpreted	Yes
	20b*	Description of any transformations, standardizations, covariate or confounder adjustments, weighting, or stratification methods to be used and why.	No
	20c*	For each analytic method proposed, a description of the assumptions of that method and what processes will be used to evaluate whether or not those assumptions hold	No
	20d*	Details of contingency plans/alternative methods to be used if the assumptions are found not to hold	NA
	20e*	In the case of non-standard test statistics, formulas provided for the test statistic with a description of the mathematical null hypothesis, how significance is determined, and how the test statistic is interpreted	No
	20f*	In the case of regression models, formulas provided for the full model with a description of which parameters are to be used, how they will be interpreted, how confidence intervals will be constructed, etc.	No
	20g*	In the case of survey, hierarchical/nested, or clustered data, a description of what methods will be used to adjust for the data structure and why (e.g. if using a GEE, describing which correlation structure and why it was chosen, etc.)	No
	20h*	For non-continuous outcomes, clearly explain the effect used (e.g. risk difference, risk ratio, odds ratio, etc.), whether it is relative or absolute, and justify why that was chosen as the effect measure of interest	No

	20i*	Documentation of any non-standard methods used (e.g. using alternative degree of freedom calculation methods, using a non-canonical link function, etc.)	Yes
	20j*	Description of any limitations, sources of bias, internal/external validity, and other relevant discussions concerning the interpretation and generalizability of the design or methods used	Yes
Additional Analysis Methods	21a*	Description of any pre-planned sensitivity analyses and how they will be interpreted	No
	21b*	Description of pre-planned subgroup analyses, power for these analyses, and planned multiple comparison adjustment procedures	No
	21c*	Description of any additional post-hoc calculations or analyses (e.g. evaluating interaction/modification effects, calculating mediation or local average treatment effects, evaluation of AUROC curves, etc.)	Yes
	21d*	If conducting any bootstrap analyses, a description of the sampling algorithm and number of iterations used	No
	21e*	If conducting any cross-validation procedures, a description of how the cross-validation is conducted (e.g. leave-one-out, train/validation/test, etc.)	No
Exploratory Analyses	22a*	Description and justification for any pre-planned exploratory analyses and what methods will be used to conduct them	Yes
	22b*	Framework for conducting any unplanned exploratory analyses and how they will be integrated into the planned analysis	No
Software	23*	List of statistical software (along with version numbers) to be used for each phase of the analysis; in the case of R or Stata, additionally list any requisite installed packages and their version numbers	Yes
Other	24*	Description of any additional planned analyses of the data (e.g. a safety analysis looking at adverse event rates for a Data Safety Monitoring Board, etc.)	No
Tables and Figures			
Table Shells	25*	Example tables related to any of the conducted analyses; if possible including any available preliminary data	Yes
Example Figures	26*	Example figures related to any of the conducted analyses; if possible including any available preliminary data.	Yes
References			
References	27a	References for any non-standard statistical methods used	Yes
	27b	References (and locations) for any relevant protocols, standard operating procedures, or other documents cited in the SAP	No
Additional Information			
Appendices	28*	If necessary, appendices may be included (e.g. a full data dictionary, a copy of a Case Report Form, etc.)	No
Addendums	29*	Any additional analyses conducted that were not included in the SAP should be documented in an addendum, describing the purpose of the additional analysis, when it was conducted, and by whom	Yes