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	*04.06.2020 - 05.06.2020 - Wave 3 - due to low enrollment, excluded from the analyses*

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.

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**
kit_ordered	Orasure ordered within 60 days?	ora_within60_yesno	1=Yes; 2= No; Missing (note: Missing refers to those who did not redeem the code)

3.2. Secondary Outcome(s)

Outcome	Description	Variables and Source	Specifications
Substance use	In the PAST 3 MONTHS, did you have a drink containing alcohol? ... In the PAST 3 MONTHS, what were the other drug(s) you used? Please specify.	Q13_1 - Q13_24	1=Yes;2=No
Stage of health behavior change	Which of these statements is most true for you?	Q15_1	1=I don't see any need to regularly test for HIV; 2=I think I should get tested for HIV regularly, but I am not sure; 3= I'm ready to start getting regularly tested for HIV; 4=I'm trying to get tested regularly for HIV; 5=I've been getting testing for HIV regularly over the past few years.
Attitudes toward HIV testing	Getting tested for HIV helps people feel better. Getting tested for HIV helps people from getting HIV. People in my life would leave if I had HIV. People who tested positive for HIV should hide it from others. I would rather not know if I have HIV.	Q15_3 - Q15_7	1=Agree; 2=Disagree Binary(1,0)

Attitudes toward HIV treatment	I am less threatened by the idea of being HIV positive than I used to be. ... I think that new HIV/AIDS treatments can eradicate the virus from your body	Q94_1 - Q94_13	Numeric (1-7)
HIV-related stigma	I feel afraid of people living with HIV/AIDS ... I feel anger toward people with HIV/AIDS.	Q14_2 - Q14_5	1=Strongly agree; 2=Agree; 3=Somewhat agree; 4=Neither agree nor disagree; 5=Somewhat disagree; 6=Disagree; 7=Strongly disagree
Medical mistrust	You'd better be cautious when dealing with healthcare organizations ... Mistakes are common in health care organizations.	Q16_1 - Q16_7	1=Strongly agree; 2=Agree; 6=Disagree; 7=Strongly disagree

3.3. Additional Variables of Interest

Variable	**Description**	**Variables and Source**	**Specifications**
Q6_1	What is your HIV status?		1=Negative; 2=Positive; 4=I don't know; 5=Refuse to answer
Q6_2	Have you ever taken PrEP		1=Yes, but haven't taken in past 6 months; 2=Yes, currently taking; 3=No, never taken
Q3_1	How old are you?		Numeric
Site	Advertising site/platform		Facebook; Google; Instagram; Jack'd; Grindr; Bing; Yahoo; Hornet; Twitter

4. Statistical Analysis Plan

4.1. Demographic and Clinical Characteristics ("Table 1")

Table 1 is listed next page as a table shell.
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 to compare kit ordering rates across sites (social media, info search sites, and dating apps) that will use time (number of advertising days) as an offset. The model was specified as:
$\log(\text{orders}) = \beta_0 + \beta_1(\text{site}) + \log(\text{advertising\_days})$

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.

Secondary analyses are focused on participants who order a test kit versus those who did not.
Wilcoxon rank test will be used for continuous variables (Attitudes toward HIV treatment, HIV related stigma, medical mistrust).
Fisher's exact test will be used for categorical variables (substance use, stage of health behavior change, attitudes toward HIV testing).

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

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					
	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
I could not be friends with someone who has 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
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
	Strongly agree	NA	NA	NA	NA	NA

When health care organizations make mistakes they usually cover it up	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

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/
Administrative Information			
Study Information	1a	Descriptive title that matches the protocol, with SAP either as a forerunner	Yes
	1b	Trial registration number, protocol version number, and/or IRB number.	Yes
	1c	CRU/Department/Division/Center/other collaborative unit that the study	Yes
Roles and responsibility	2a	Listing of principal investigators, clinical leads, and co-authors (if known)	Yes
	2b	Name and affiliation of SAP author(s)	NA
	2c	Names, affiliations, and roles of other SAP contributors (e.g. senior	NA
SAP	3	SAP version number, with date of current version and original creation	Yes
Project	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-	No

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

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
Study Overview			
Background and	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	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	NA
	10b	List of all exposure variables, with a description of their coding/units, timing, and source, corresponding to the statistical hypotheses. If any	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-	Yes
	12b	Study setting, location, and relevant dates (e.g. periods of enrolment,	Yes
	12c	Description of intervention or exposure groups, with allocation ratios, and	No
	12d	Details on randomization (e.g. stratification factors) and blinding	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-	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,	Yes
Sample Size	13a	Sample size calculation or justification (either provided in full or	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	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	No
Data	15a	Description of data collection/acquisition process, with contact	Yes

	15b	Description of data flow/transfer from primary data collection through to	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	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	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	No
	17c	Description of the tabular and graphical presentations of simulation results	No
Statistical Analysis Plan			
Statistical Significance	18a *	Hypothesis testing framework (e.g. superiority, equivalence, non-inferiority), or description of alternative analytic framework (e.g.	Yes
	18b *	Level of significance for primary hypotheses, including a description and rationale for any multiple comparisons adjustment or Type I error control	NA
	18c *	Description of any decision-making rules based on confidence intervals, credible intervals, prediction intervals, Bayes' factors, or other alternative	Yes
	18d *	Description of how the results of any hypothesis tests (or alternative inferential methods) will be interpreted with respect to both the statistical	
Descriptive	19a	List of characteristics (e.g. demographic, clinical) to be summarized	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	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	Yes
	20b *	Description of any transformations, standardizations, covariate or confounder adjustments, weighting, or stratification methods to be used	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	No
	20d	Details of contingency plans/alternative methods to be used if the	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	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	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	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	No
	20i*	Documentation of any non-standard methods used (e.g. using alternative degree of freedom calculation methods, using a non-canonical link	Yes
	20j*	Description of any limitations, sources of bias, internal/external validity, and other relevant discussions concerning the interpretation and	Yes
Additional Analysis Methods	21a	Description of any pre-planned sensitivity analyses and how they will be	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	Yes
	21d	If conducting any bootstrap analyses, a description of the sampling	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,	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	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,	No
Tables and Figures			
Table Shells	25*	Example tables related to any of the conducted analyses; if possible	Yes
Example	26*	Example figures related to any of the conducted analyses; if possible	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	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	Yes