

## 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.*

---

<b>Title</b>	Relative Effectiveness of Social Media, Dating Apps, and Information Search Sites in Promoting HIV Self-testing: Observational Cohort Study
<b>CRU/Department/Division/Center</b>	National Institute on Drug Abuse
<b>IRB Number</b>	18-001580
<b>Investigators:</b>	
<b>Primary Investigator</b>	Jeffrey D. Klausner
<b>Collaborative Lead</b>	
<b>Co-authors (if known)</b>	Chrysovalantis Stafylis, Gabriella Vavala, Qiao Wang, Bethany McLeman, Shea M Lemley, Sean D Young, Haiyi Xie, Abigail G Matthews, Neal Oden, Leslie Revoredo, Dikla Shmueli-Blumberg, Emily G Hichborn, Erin McKelle, Landhing M Moran, Petra Jacobs, Lisa A Marsch
<b>Analysis Biostatistician(s)</b>	Haiyi Xie
<b>Biostatistics Supervisor</b>	
<b>Lead Biostatistician</b>	
<b>Subject Matter Expert</b>	Jeffrey D. Klausner
<b>Original Creation Date</b>	01/16/2026
<b>Version Date</b>	01/22/2026
<b>Project Folder Location</b>	
<b>Project Goal(s)</b>	Manuscript, abstract
<b>Submission Deadline(s)</b>	01/28/2026
<b>Effort Estimate (optional)</b>	

- 
- |                               |   |
|-------------------------------|---|
| <b>Investigator Agreement</b> | <input type="checkbox"/> 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).<br><input type="checkbox"/> 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

The initial plan specified a Poisson regression model to compare self-test kit ordering rates by platform type (social media, information search site, and dating app). During analysis, significant platform-by-wave interactions and differences in ordering rates across sites within the same platform were observed, making pooling across platform types inappropriate. Thus, the analysis approach was revised to compare ordering rates between specific sites within the same wave. Pairwise comparisons across all six sites from the two waves were conducted with multiple testing adjustment using the Hochberg method.

---

## Acronyms

MSM	Men who have sex with men
CDC	Centers for Disease Control and Prevention
PrEP	Pre-Exposure Prophylaxis
TAPS	Tobacco, Alcohol, Prescription medication, and other Substance
SD	Standard deviation
IQR	Interquartile range
HIV	Human Immunodeficiency Virus
AIDS	Acquired Immunodeficiency Syndrome
NIDA	National Institute on Drug Abuse

---

## 1 Study Overview

**Background/Introduction:** This study is a longitudinal observational study designed to evaluate the effectiveness of different web platforms in promoting HIV self-test kit ordering among young Black and Latinx MSM at increased risk for HIV infection. Participants were recruited through advertisements placed on social media sites, dating apps, and web browsers during two recruitment waves, with each wave simultaneously including one site from each platform type. The primary outcome was the number of HIV self-test kits ordered per day from each platform.

## **1.1 Study Aims**

Primary aim:

- Compare the relative effectiveness of social media sites, informational search sites, and dating apps in promoting HIV self-testing

Secondary aims:

- Evaluate the association between substance use and HIV self-test kit ordering
- Evaluate whether stage of change for HIV testing is associated with HIV self-test kit ordering
- Evaluate associations between attitudes toward HIV testing and HIV treatment and HIV self-test kit ordering
- Evaluate the relationship between HIV stigma and HIV self-test kit ordering
- Evaluate whether medical mistrust is associated with HIV self-test kit ordering
- Evaluate whether participant opinions toward PrEP are associated with HIV self-test kit ordering

Exploratory aim:

- Compare advertisement reach and cost metrics across web-based platforms used to promote HIV self-testing

## **1.2 Study Hypotheses**

### **1.2.1 Primary Hypothesis**

- H0: Within each recruitment wave, HIV self-test kit ordering rates do not differ across the recruitment sites active in that wave vs. H1: Within each recruitment wave, HIV self-test kit ordering rates differ for at least one site compared with another site in the same wave

### **1.2.2 Secondary Hypotheses**

- H0: HIV self-test kit ordering does not differ by substance use vs. H1: HIV self-test kit ordering differs by baseline substance use
- H0: HIV self-test kit ordering does not differ by stage of change for HIV testing vs. H1: HIV self-test kit ordering differs by stage of change for HIV testing
- H0: HIV self-test kit ordering does not differ by attitudes toward HIV testing and HIV treatment vs. H1: HIV self-test kit ordering differs by attitudes toward HIV testing and HIV treatment
- H0: HIV self-test kit ordering does not differ by HIV stigma vs. H1: HIV self-test kit ordering differs by HIV stigma
- H0: HIV self-test kit ordering does not differ by medical mistrust vs. H1: HIV self-test kit ordering differs by medical mistrust
- H0: HIV self-test kit ordering does not differ by opinions toward PrEP vs. H1: HIV self-test kit ordering differs by opinions toward PrEP

## **2 Study Population**

### **2.1 Inclusion Criteria**

- Men who have sex with men (MSM)
- Self-identified as Black/African American or Latinx, including multiracial and multiethnic individuals of these groups
- Reported condomless anal sex in the past 90 days or more than one male sex partner in the past 90 days
- Be between the ages of 18-30 years old

### **2.2 Exclusion Criteria**

- HIV-positive status
- HIV testing within the 90 days prior to enrollment
- Current use of PrEP or use of PrEP at any time during the 6 months prior to enrollment

## 2.3 Data Acquisition

Fill in all relevant information:

Study design	Longitudinal observational cohort study
--------------	---

Notes: Participants enrolled from Google and Facebook during periods when Grindr was inactive were excluded from analysis to ensure comparability across the different platforms. The two advertisement periods of Wave 1 were combined prior to analysis. Wave 3 data were excluded entirely due to zero enrollment during the COVID-19 emergency, making statistical models inestimable.

## 3 Outcomes, Exposures, and Additional Variables of Interest

### 3.1 Primary Outcome(s)

Outcome	Description	Variables and Source	Specifications
HIV self-test kit ordering	Whether a participant ordered an HIV self-test kit	Variable: ora_redeemed Source: Orasure redemption records	Categorical variable indicating HIV kit ordering status: <ul style="list-style-type: none"><li>• Yes = participant ordered an HIV kit</li><li>• No = participant did not redeem voucher</li><li>• Over 60 days = voucher redeemed outside analytic window</li></ul>

### 3.2 Secondary Outcome(s)

Outcome	Description	Variables and Source	Specifications
Substance use	Participant-reported substance use at baseline	Variables: Q12_2, Q12_3, Q12_4 (past 12 months frequency); Q13_1–Q13_23 (past 3 months yes/no); Q13_24 (“other drug”) Source: Baseline questionnaire	Q12_2–Q12_4: categorical frequency, 1=Daily or almost daily; 4=Weekly; 5=Monthly; ... Q13_1–Q13_23: 1=Yes; 2=No. Q13_24: text
Stage of change for HIV testing	Participant readiness for HIV testing at baseline	Q15_1 Source: Baseline questionnaire	Ordinal categorical 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 and treatment	Participant attitudes and beliefs regarding HIV testing and HIV treatment	Variables: HIV testing attitudes: Q15_3, Q15_4, Q15_5, Q15_6, Q15_7 HIV treatment perceptions/optimism: Q94_1, Q94_5, Q94_6, Q94_7, Q94_8, Q94_9, Q94_10, Q94_12, Q94_13 Source: Baseline questionnaire	Q15_3–Q15_7: 1=Agree; 2=Disagree. Q94_*: numeric responses
HIV-related stigma	Participant-perceived HIV stigma	Variables: Q14_2, Q14_3, Q14_4, Q14_5 Source: Baseline questionnaire	Likert categorical responses 1=Strongly agree; 2=Agree; 3=Somewhat agree; 4=Disagree; 5=Strongly disagree
Medical mistrust	Participant mistrust of medical organizations/providers	Variables: Q16_1–Q16_7 Source: Baseline questionnaire	Likert categorical responses. Most items: 1=Strongly agree; 2=Agree; 6=Disagree; 7=Strongly disagree
Opinions about PrEP	Participant opinions, barriers, and facilitators related to PrEP use	Variables: Q7_5_fu1, Q7_8_fu1, Q7_9_fu1, Q7_10_fu1, Q7_16_fu1, Q8_1_6_fu1, Q8_1_12_fu1, Q8_1_13_fu1 Source: Follow-up questionnaires	Q7_*_fu1: Likert 1=Strongly agree ... 5=Strongly disagree) Q8_1_*_fu1: 1=Not at all important; 2=Slightly important; 3=Moderately important; 4=Very important; 5=Extremely important

### 3.3 Additional Variables of Interest

Variable	Description	Variables and Source	Specifications
Age	Participant age at enrollment	Q3_1 Source: Baseline questionnaire	Continuous (years)
Sex at birth	Sex assigned at birth	Q4_1 Source: Baseline questionnaire	1 = Male, 2 = Female

Ethnicity	Hispanic/Latinx identity	Q5_1 Source: Baseline questionnaire	1 = Yes, 2 = No
Race	Self-identified race	Q5_3, Q5_3_28_TEXT Source: Baseline questionnaire	Categorical; multiple race categories; text for “Other”
PrEP history	Prior PrEP use	Q6_2 Source: Baseline questionnaire	Categorical
Sexual risk behavior	Number of male sex partners (90 days); condom use frequency; condomless receptive anal sex (90 days)	Q11_2, Q11_3, Q11_4 Source: Baseline questionnaire	Q11_2: numeric; Q11_3: ordinal (Never → Always); Q11_4: 1 = Yes, 2 = No
HIV testing history	Ever tested for HIV and time since last test	Q11_5, Q11_6, Last_hiv_test_months, Last_hiv_test_interval Source: Baseline questionnaire	Binary for ever tested; date field; interval categories
Reason for not testing	Main reason participant had not tested for HIV	Q11_7 Source: Baseline questionnaire	Categorical
Recruitment platform	Platform type and specific site of recruitment	Platform, site Source: Baseline questionnaire	Categorical (Social media, Dating app, Informational site; site-level categories)
Recruitment wave	Study recruitment wave	Wave Source: Baseline questionnaire	Categorical (1, 2, 3, 4)
Geographic location	Participant state of residence	State_of_use Source: Baseline questionnaire	Categorical (US states)

## 4 Statistical Analysis Plan

### 4.1 Demographic and Clinical Characteristics (“Table 1”)

Table 1 summarizes demographic and behavioral characteristics of participants included in the study. Participant age is summarized using median/interquartile range (IQR). Race and ethnicity are presented as frequencies and percentages. Behavioral characteristics include history of PrEP uptake, number of male sex partners in the past 90 days, condom use frequency, and engagement in condomless receptive anal sex during the past 90 days. HIV testing history is defined by lifetime HIV testing status and time since last HIV test for those who have been tested before. For participants who reported never testing for HIV, the primary reasons for not testing are summarized using frequencies and percentages.

### 4.2 Analyses Plan for Aim 1

The primary objective is to compare HIV self-test kit ordering rates across online recruitment sites within each recruitment wave. Analyses are restricted to periods during which all sites within a wave

were simultaneously active. Participants enrolled during Wave 1 from Google and Facebook while Grindr was inactive are excluded. The two advertisement periods within Wave 1 are combined prior to analysis. Participants who did not order an HIV self-test kit within 60 days of receiving a test code are classified as not having ordered a kit.

The original analysis plan had a Poisson regression evaluate differences in HIV self-test kit ordering rates by platform type (social media, dating apps, and information search sites), accounting for varying recruitment durations with a log time offset. However, significant platform-by-wave interactions and substantial differences in ordering rates between sites within the same platform were observed, so pooling sites by platform type is not appropriate for statistical inference.

Instead, differences in HIV self-test kit ordering rates are evaluated at the site level within each recruitment wave. Comparisons are done to sites within the same wave, allowing for direct assessment of relative ordering rates without pooling by platform types. Pairwise comparisons are done for all six recruitment sites across the two waves, and statistical significance is assessed using Hochberg-adjusted P values to account for multiple testing.

### **4.3 Analyses Plan for Aim 2**

Secondary analyses evaluate the associations between participant-level characteristics and HIV kit ordering. This is intended to identify potential relationships of HIV kit ordering with individual factors, including substance use, stage of change for HIV testing, attitudes toward HIV testing and treatment, HIV stigma, medical mistrust, and perceptions of PrEP.

These characteristics are compared between participants who ordered a test kit and those who did not. Continuous variables are summarized using measures of central tendency and variability (median, 25<sup>th</sup>, 75<sup>th</sup> percentiles) and compared between groups using Student's t tests. Ordinal variables (like Likert-scale variables) are compared using Wilcoxon rank-sum tests. Categorical variables are summarized using frequencies and percentages and compared between groups using Fisher's exact tests.

### **4.4 Descriptive Analysis**

Descriptive analyses are done to summarize study recruitment, performance of the chosen platforms, and HIV self-test outcomes. The recruitment process is outlined using counts, including the number of individuals screened, the number of eligible participants, and the number included in the final primary analysis population. HIV self-test outcomes among that final sample are summarized using counts and include HIV self-test kit ordering, the number who reported results, and the actual test results. Among participants who were HIV-positive, the number that visited a doctor and then subsequently initiated HIV treatment are counted. Among participants reporting a negative test result, those who sought PrEP services and then potentially initiated PrEP are summarized with counts.

Recruitment characteristics are summarized by platform, site, and recruitment wave, including the duration of advertising, the number of HIV self-test kits ordered, and observed rate of kits per day. Advertisement performance metrics are summarized descriptively by platform, including number of impressions, number of clicks, click-through rates (in %), users screened, enrolled participants, and total funds spend per platform (in US dollars).

### **4.5 Sensitivity Analysis**

Three sensitivity analyses are conducted to assess how robust the primary findings are, using the statistical approach, Poisson regression, and posthoc contrast. The primary sensitivity analysis includes all HIV self-test kits ordered at any time during the study, including orders occurring outside the 60-day window used in the primary analysis, and includes all participants in the validated participant population. The second sensitivity analysis addresses the two-phase structure of Wave 1

resulting from the temporary suspension of advertising on Grindr. The final sensitivity analysis evaluates the impact of the COVID pandemic on recruitment and ordering rates.

#### **4.6 Missing Data**

The analysis of the primary outcome does not include missing data, as HIV self-test kit ordering counts and recruitment duration are received from automated records. For the individual-level survey, participants were permitted to skip questions. Skipped responses are classified as missing and are excluded from the calculation of summary statistics and from analyses involving these variables. No imputation methods are applied in the study.

#### **4.7 Software**

Statistical analyses were conducted using Statistical Analysis Software (SAS), version 9.4.

### **5 Limitations**

Several limitations should be considered when interpreting analyses conducted in this study. The study was conducted in selected geographic areas with high HIV incidence, so findings may not be generalizable to populations outside these regions. Enrollment was varied across recruitment waves, which limited the ability to conduct broad comparisons between platforms and made pooling of data across waves impossible.

In addition, recruitment was conducted on a selected set of commonly used apps and websites, which were grouped into platform categories based on similar characteristics. Analyses are therefore specific to the platforms and sites that were chosen to display the ads and may not extend to other platforms or sites not included in this study.

### **6 References**

- **UCLA Statistical Consulting Group.**  
*Poisson Regression*. Institute for Digital Research and Education, University of California, Los Angeles.  
<https://stats.oarc.ucla.edu/r/dae/poisson-regression/> (accessed January 21, 2026).
- **Statology.**  
*Benjamini–Hochberg Procedure*.  
<https://www.statology.org/benjamini-hochberg-procedure/> (accessed January 21, 2026).



## Appendix

Original model specification for primary analysis:

$$\log(o_{ij}) = \log(t_i) + \alpha + \beta_i + \gamma_j + \beta\gamma_{ij}$$

where

$o_{ij}$  is the number of kits ordered by the site in Wave  $i$  (i.e. time period  $i$ ), platform type  $j$ .

$t_i$  is the time that the Wave platforms were recruiting.

$\beta_i$  is the main effect of wave.

$\gamma_j$  is the main effect of platform type.

$\beta\gamma_{ij}$  is the interaction term.

Under this model, the rate for any site  $ij$  is given by:

$$rate_{ij} = \exp(\alpha + \beta_i + \gamma_j + \beta\gamma_{ij})$$

Population sociodemographic and behavioral characteristics:

Characteristic	Value
Age in years, median (IQR)	25 (23-27)
<b>Ethnicity, n (%)</b>	
Hispanic/Latinx	66 (26)
<b>Race, n (%)</b>	
American Indian or Alaskan Native	1 (0.4)
Black or African American	196 (78.4)
White	28 (11.2)
Other	14 (5.6)
Multiracial	11 (4.4)
<b>History of PrEP<sup>a</sup> uptake, n (%)</b>	
Never taken PrEP	232 (91.3)
In the past 6 months	22 (8.9)
Number of male sex partners in the past 90 days, median (IQR)	4 (3-6)
<b>Condom use, n (%)</b>	
Never	36 (14.2)
Sometimes	108 (42.5)
About half the time	37 (14.5)
Most of the time	68 (26.8)
Always	5 (2)
Condomless receptive anal sex in the past 90 days, n (%)	210 (82.7)
Ever tested for HIV during lifetime, n (%)	191 (75.2)
<b>If tested for HIV, median (IQR)</b>	
Months since last HIV test	11 (6-21)
If not tested for HIV, n (%)	63 (24.8%)
<b>Main reasons cited by the 63 participants for not getting tested, n (%)</b>	
Unlikely to be exposed to HIV	8 (12.7)
Afraid of testing HIV-positive	26 (41.3)
Did not want to think about HIV/HIV-positive	8 (12.7)
Worried about names being reported if positive	3 (4.8)
Dislike for needles	5 (8)
Unable to trust that the results will be confidential	3 (4.8)
Unaware of where to get tested	7 (11.1)
Other reasons	3 (4.8)

Number and rate of HIV home self-test kits ordered through promotional platforms by wave per protocol sample:

Type of platform	Wave	Number of days for each wave	Number of test kits ordered	Order rate (ordered test kits/day)
<b>Social media site</b>				
Facebook	1 <sup>a</sup>	70	13	0.19
Instagram	2	38	13	0.34
Subtotal	N/A	108	26	0.24
<b>Dating app</b>				
Grindr	1 <sup>a</sup>	70	9	0.13
Jack'D	2	38	125	3.29
Subtotal	N/A	108	134	1.24
<b>Information search site</b>				
Google	1 <sup>a</sup>	70	17	0.24
Bing	2	38	0	0.00
Subtotal	N/A	108	17	0.16
Total	N/A	108	177	1.64

Performance of advertisements by platform throughout the advertisement campaign:

Platform	Impressions <sup>a</sup>	Clicks <sup>b</sup>	Click-through rate		Enrolled participants	Total funds spent (US \$)
			(%) <sup>c</sup>	Users screened <sup>d</sup>		
Social media	3,864,778	21,399	0.6	2679	33	16,221.52
Dating apps	1,331,200	53,067	4	4390	202	17,939.40
Information search sites	708,770	10,869	1.5	2562	19	15,978.86