Statistical Analysis Plan U=U Trial

Project title: Improving HIV Testing, Linkage, and Retention in Care for Men Through U=U Messaging

Unique Protocol ID: R01MH129223
NCT number: NCT05602376

Document date: 25 October 2024

Statistical Analysis Plan U=U Trial
(Reviewed by Mary Putt, ScD, Elizabeth Bair, MS, Alison Buttenheim, PhD, Andrew Medina-Marino, PhD)

Note: SAP created prior to statistical team accessing information about treatment group.

Created on 2024/09/03 16:34:00
Current Revisions: 2024/10/25

A schema of the study design appears on page 2.

Primary outcome: The primary outcome is verified ART initiation or reinitiation within 30 days of testing. Verification is based on a record of treatment (re)-initiation in the NHLS LabTrac system or in the clinic records.

Secondary outcomes: The secondary outcomes (dark yellow and dark blue boxes) include:

1. ART Initiation /Reinitiation sensitivity analysis: Either verified OR self-reported ART initiation or reinitiation with 30 days of testing – this will be a more "generous" definition of initiation for sensitivity/robustness check.
2. Tested Positive for HIV as part of our study AND no evidence of being on ART in past 28 days. NOT COUNTED IN THIS OUTCOME: Tested positive with us but evidence of ART in past 28 days.
3. Presented, consented, and tested AKA "tested"

Descriptive Analysis

Baseline characteristics: Site-level and participant-level data will be summarized overall and by three strata (Eastern Cape Rural, Eastern Cape Urban and Western Cape Urban) as well as by intervention-arms. Site-level variables include the total number of days visited, the total number of days cancelled and, if recorded, the reasons for cancelling. The number of cards distributed by site and time will be described.

Participant-level characteristics can only be described for those who present for testing and consent to the study (Conditional Analysis: Men who Test in the Schema). Participants are self-selected and their characteristics may differ across study arms. We will estimate absolute (means, differences in rates) and standardized mean differences between arms along with 95% confidence intervals for the differences. Relevant control variables to include in the propensity score estimation are:

Variable:	Source in Aim 1 baseline instrument:
Age	Constructed from details_birthdate
Unemployed	Recode demo_employment, 1=0 2=1 3=1 4=0 5=0 6=0 7=0 8=0
Matric or above	Recode demo_education 1=0 2=0 3=0 4=1 5=1
Ever tested before	Recode survey_hiv_test 1=1 97=0 99=0
Most recent test result =pos	=1 if survey_hiv_result =1 & survey_hiv_test =1, all else 0
Partner status unknown	Recode survey_hivstatus_partner 1=0 2=1 97=0 *99=1
Partner status = positive	=1 if survey_hiv_partner=1 & survey_hivstatus_partner=1, all else 0
Logged # of partners	If survey_sexpartners= 0, the 0, else log(survey_sexpartners)
Any transactional sex	survey_sexformoney = yes
Regular Clinic not close to home	cae_q1 = no
Regular clinic not close to work	cae_q2 = no

Regular clinic more friendly to men cae_q4 = 1

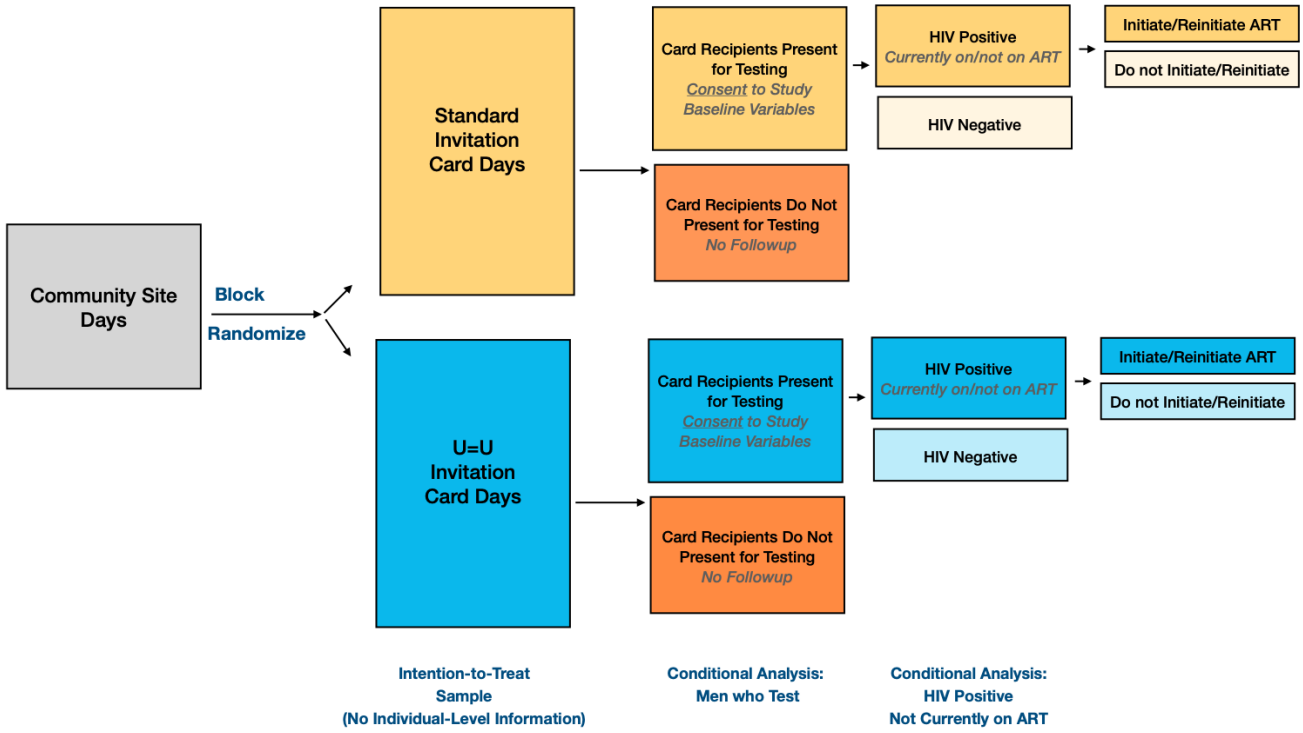
Design Integrity: Each site is visited on multiple days and receives one of the two interventions (standard of care, SOC, or U=U) each day. The study was designed to achieve approximate balance of the intervention frequency both within and across sites, as well as over time. We will tabulate and plot the number of cards distributed at approximately semi-monthly intervals over time by site and intervention.

We will report the number of intervention cards distributed using site, by the stratification variable (Eastern Rural, Eastern Urbans, Western Urbans) and intervention type (U=U versus SOC).

Outcome variables: Using graphical and tabular approaches, rates of each of the primary and secondary outcomes will be summarized as aggregates across site-interventions (cluster-level summaries by site) and as aggregates across interventions (individual-level summaries).

Incomplete data: Rates of incomplete data will be tabulated overall and by arm.

Schema for study design.



Primary analysis. The study structure is outlined above. The primary analysis is intention-to-treat based on randomization of the sample of participants at the site-day level. The analysis will test the hypothesis that men who are invited to test for HIV using U=U versus standard messaging cards will have increased rates of ART initiation within 30 days of testing. The primary analysis will use generalized estimating equations (GEE) with a log link, an independence working correlation matrix and robust standard errors. The model will include terms for the stratification variable (Eastern Rural, Eastern Urban or Western Urban). In the absence of any detrimental effects of U=U in promoting HIV testing in previous work, the statistical test will be one-sided with a Type I error rate of 0.05. An over-dispersion parameter will be

reported. Fay and Graubard adjustments for small sample cluster numbers will be used to adjust the standard errors.

The primary analysis includes the sample of participants who are randomized prior to consent. Information on individual-level baseline characteristics is not available for those who do not present for testing (Orange boxes). Hence no covariate adjustment for imbalances in individual-level baseline characteristics can be made for the primary analysis.

Secondary analysis .
Intention to Treat Analysis: The primary analysis using the intention-to-treat sample will be repeated using the secondary outcomes.

Analyses Conditional on Presenting, Consenting and Testing: Here we consider the population of men who present for testing, are consented and enrolled in the study, and complete testing (hereafter, "test"). Two questions are of interest:

1. Among men who test, do rates of HIV positivity or initiation/reinitiation of care differ based on whether they received a U=U versus SoC card? (dark yellow and dark blue boxes)
2. Men who present for testing with different card types (U=U versus SoC) may differ in their baseline characteristics (e.g. the same variables listed above for inclusion in the propensity score). The second part of this analysis asks whether conditional on baseline characteristics at the time of testing, is receipt of a U=U card associated with increased ART initiation/re-initiation.

The analysis will follow the strategy outlined for the primary analysis but will use the "Conditional on Testing" Sample (See Schema). Unadjusted and covariate-adjusted analyses will be presented. The number of events is anticipated to be small limiting the number of covariates in any one model. Using a logistic regression model, we will create a propensity score model for the card type (U=U versus SoC) using a logistic regression model with all covariates. Independent variables included in the propensity score will include those listed above. The propensity score will be used in adjusted analysis, specifically using inverse-probability of treatment weighting (IPTW).

Analyses Conditional on Testing Positive
Of primary interest are men who test positive who are also eligible to initiate or reinitiate care.

1. Do rates initiation/reinitiation of ART increase among men who received a U=U versus SoC card? (dark yellow and dark blue boxes)
    a. Repeat this analysis separately for initiation among those not previously known to be HIV positive and reinitiation among those previously known to be HIV positive and not currently on ART
2. We anticipate that the number of HIV positive participants will be small and hence no adjustment for covariates will be attempted.

Subgroup Analyses
The analyses of the intention-to-treat sample and the 'conditional on testing' sample will be repeated for each of the 3 geographic strata (Eastern rural, Eastern and Western urbans) presented using forest plots and for age (<25 years, 25-34 years, >34 years).

Missing Data: Rates of missing data will be tabulated by arm, both for the outcome and baseline covariate data. If necessary, we will describe the missing data by arm and by site at each step after testing. (Rates of missingness appear to be very very low in the provisional dataset).

Westgate PM, Cheng DM, Feaster DJ, Fernández S, Shoben AB, Vandergrift N. Marginal modeling in community randomized trials with rare events: Utilization of the negative binomial regression model. Clin Trials. 2022 Apr;19(2):162-171. doi: 10.1177/17407745211063479. Epub 2022 Jan 6. PMID: 34991359; PMCID: PMC9038610.

Sample Size Calculations and Statistical power.

A sample size 28,880 invitation cards distributed (14,440 in each arm) allows 84% power to detect a risk difference in our primary outcome (yield, or ART initiations as a percent of invitations distributed) of 0.16%. We assume a yield of 0.1% (1 ART initiation per 1000 invitation cards) for SoC versus 0.257% (2-3 ART initiations per 1000 invitation cards) for the U=U intervention. Our sample size calculations are based on pilot data, experience in the field, and recent HIV incidence. For the U=U intervention arm, we assumed that 15% would present for testing, 3.5% of those who test would test HIV-positive, and 49% of those receiving a positive test result would initiate ART, for a yield of 0.257%. For the SoC arm, we estimated that 10% would present for testing, 2.5% of those who test would test HIV-positive and 40% of those testing positive would initiate ART, for an overall yield of ART initiation among those invited to test of 0.10%. The sample size calculation assumed a within-site-day intraclass correlation (ICC) of 0.01 for subjects sampled on the same day at a single site, and a between-site-day ICC of 0.009 for subjects sampled on two consecutive sampling days at a site. A discrete-time decay correlation structure using an exponential decay function of the time ‘j’ between site-days allowed the correlation between subjects sampled at longer intervals apart to decrease over time. Power calculations were conducted by uploading a design matrix generated in R(V4.2.2) to Shiny CRT (60). To achieve this sample size, we will distribute 90 invitation cards per site day for a total of 320 site days (160 for U=U intervention and 160 for SoC. We will visit 8 sites from each of Western and Eastern Cape for 20 days each. Within Eastern Cape 3? Sites are rural and 5 are Urban; all 8 Western Cape sites are Urban.