DARAJA Clinical Trial

Statistical Analysis Plan
Version 1.0
12 May 2023

Reducing post-hospital mortality in HIV-infected adults in Tanzania (DARAJA)

Sponsored by:
National Institutes of Health

Funding Mechanism
R01

Grant Number:
MH118107-01

Trial protocol:
version 2.2, July 8, 2019

Trial registration:
https://clinicaltrials.gov/ct2/show/NCT03858998

Protocol paper:
Kisigo GA, Issarow B, Abel K, et al. A social worker intervention to reduce post-hospital mortality in HIV-infected adults in Tanzania (Daraja): Study protocol for a randomized controlled trial. Contemp Clin Trials. 2022;113:106680. doi:10.1016/j.cct.2022.106680

INTRODUCTION

This statistical analysis plan (SAP) describes the statistical procedures that address the study objectives specified in the DARAJA Study Protocol, Protocol Paper and Trial Registration (referenced above).

The SAP is applicable to the final analyses following end line survey. The end line survey was conducted about 12 months after the final enrollment.

Study design

This is a randomized controlled trial to evaluate the efficacy of a social worker intervention in 500 HIV-infected adults who have been hospitalized in Mwanza, Tanzania.

Statistical hypothesis for primary objective

The **primary hypothesis** is that the social worker intervention will significantly increase survival at 12 months compared to controls. The study is powered to have 98% power to detect a difference in survival from 75% in the control group to 90% in the intervention group.

Analysis sets

The full analysis set is all participants in the end line survey (500 HIV adults).

Statistical analysis

Descriptive analysis

Summary statistics that are appropriate to specific variable types will be used to describe socio-demographic information of the participants in the full analysis set according to trial arm (see **shell Table 1**). Socio-demographic information were collected at baseline interview. For the descriptive analysis of the baseline characteristics, the decision to report mean (SD) or median (ranges) and the categorization of continuous variables will depend on the final distribution of these variables. These categories may be modified prior to final report to allow for optimal characterization of the respective variables. Separately, frequency, and percentages will be used to summarize participants' characteristics that were collected using categorical variables. The levels of categorical variables with few observations will be merged provided that this does not meaningfully impact on the intent of the analysis.

NB: the terms enrollment date and randomization date are used interchangeably in this document since these dates are the same for all participants.

Study endpoints

Primary endpoint: Survival at 12 months

The number of participants who die in the first 12 months post-hospitalization were recorded. Death was determined by phone calls to relatives and was confirmed by verbal autopsies, hospital records, or death certificates. Survival at 12 month was confirmed during the 12-month face to face questionnaire interview or by phone call at 12 months. We will compare 12-month survival between the intervention and control arms using the log-rank test following the intention-to-treat principle. Survival times are measured in days from the date of randomization. This is an individual randomized trial, and we anticipate the two arms will have similar baseline characteristics. Therefore, the primary analysis will be done without adjusting for baseline variables. Since we have one primary outcome, we will test the primary hypothesis without multiple testing adjustments. Statistical tests will be two-tailed, with a significance level of 0.05. If some variables at baseline are notably imbalanced, we will conduct adjusted analyses as secondary or sensitivity analysis using multivariable-adjusted Cox regression and Cochran–Mantel–Haenszel statistics. The results will be summarized in terms of survival rates as well as hazards ratios, along with confidence interval and statistical significance.

We will also conduct a secondary “per-protocol’ analysis. In this analysis, we will exclude all participants who died before 90 days since these participants could not have completed all 5 sessions according to the protocol. Among participant who survived to 90 days in both arms, we will will use proportional hazard cox regression to determine whether intervention participants who completing all 5 sessions experienced a lower hazards of mortality from 90 days to 1 year compared to control participants.

Secondary endpoints

The secondary endpoints considered in this analysis will be **linkage and other steps in the HIV Care Continuum.** Most secondary outcome variables will be determined by medical record extraction from the HIV clinic electronic medical record (except ART adherence). Medical record extraction was performed every 3 months using a standardized CRF ("Follow-up Medical Records Data Extraction Form").

Linkage and ART initiation will be defined as the first visit to the outpatient HIV clinic after enrollment and the time of ART initiation, respectively. Twelve-month retention will be defined as being alive and on ART between 10 and 12 months. HIV RNA levels are retrieved from the EMR. HIV viral suppression will be defined as being alive and on ART with a plasma HIV RNA level <1,000 copies/μl according to the World Health Organization definition. Participants who have ≥1,000 copies/μl will be considered unsuppressed. Adherence will be measured with a 4-day recall questionnaire during the 12 month interview.

Time to linkage and ART initiation will be measured in days from the date of randomization to the first HIV clinic (CTC) visit or ART initiation, respectively. For participants without linkage, follow up time will be censored at the date of lost-to-follow up (defined as the date that the research team last had contact with the participant by phone call or face-to-face visit), death, or study end, whichever happens first. Time to linkage and ART initiation will be compared between groups using the log-rank test. We will conduct adjusted analyses using Cox regression. HIV clinic retention, ART adherence and viral suppression will be treated as binary outcomes at 12 months. We will compare these outcomes between study arms using the Fisher exact test. We will conduct adjusted analyses using multivariable adjusted logistic regression. Cochran–Mantel–Haenszel statistics will be used for subgroup analyses.

1) **HIV clinic attendance (linkage)** will be defined as the first visit to the outpatient HIV clinic after enrollment as recorded on the 3 monthly medical record extraction of HIV clinic data. Medical record extraction was performed every 3 months using a standardized CRF ("Follow-up Medical Records Data Extraction Form"). The Fine-Grey approach will be used to adjust for the competing risk of death, if necessary (*Austin et al, Stat Med 2017*).

2) Antiretroviral therapy (ART) initiation will be defined as the date that the participant first started (or re-started) ART after enrollment. ART initiation was determined using medical record extraction of CTC clinic data. Medical record extraction was performed every 3 months using a standardized CRF ("Follow-up Medical Records Data Extraction Form"). We also defined in-hospital initiation of ART as initiation of ART on or before the day of hospital admission. The Fine-Grey approach will be used to adjust for competing risks, if necessary.

3) **HIV clinic retention** at 12 month will be defined as being alive and on ART between 10 and 12 months. HIV clinic retention was determined using medical record extraction of CTC clinic data. Medical record extraction was performed every 3 months using a standardized CRF ("Follow-up Medical Records Data Extraction Form"). Participant who died or withdrew before 12 months will be excluded from the analysis. Stratification will be used to adjust for competing risks, if necessary.

4) **Suppressed viral load** will be defined as a binary outcome based upon the WHO definition of viral suppression as a plasma HIV-1 RNA level <1000 copies/μl. HIV viral load suppression was determined using medical record extraction of CTC clinic data. Medical record extraction was performed every 3 months using a standardized CRF ("Follow-up Medical Records Data Extraction Form"). HIV RNA levels are measured routinely at the HIV clinics at 6, 12 and 24 months after ART initiation. For this analysis we will use the most current viral load available for each participant. Participant who died or withdrew before 12 months will be excluded from the

analysis. Missing values are assumed to be not suppressed since viral load is obtained on all participants who have been on ART for at least 6 months. Stratification will be used to adjust for competing risks, if necessary.

5) **ART adherence** will be assessed using the ACTG 4-Day ART Recall Questionnaire based on the 12 month research visit or phone call. Poor adherence will be defined as missing one or more pills in the past 4 days. Participant who died or withdrew before 12 months will be excluded from the analysis. Stratification will be used to adjust for competing risks, if necessary.

Intervention fidelity and completion

An experienced social scientist independent of the conduct unannounced field visit to observe approximately 5% of the intervention sessions of each social worker. The social scientist observed and rated the fidelity of the intervention session using a structured form including the four core elements of DARAJA intervention. A score was provided on a scale of zero to four for each observed intervention session. We will report the average fidelity scores. We will also report the number of intervention participants who completed all 5 prescribed sessions and reasons for participants who did not complete all 5 sessions including died vs. unavailable vs. moved before completing all 5 sessions.

Subgroup analyses

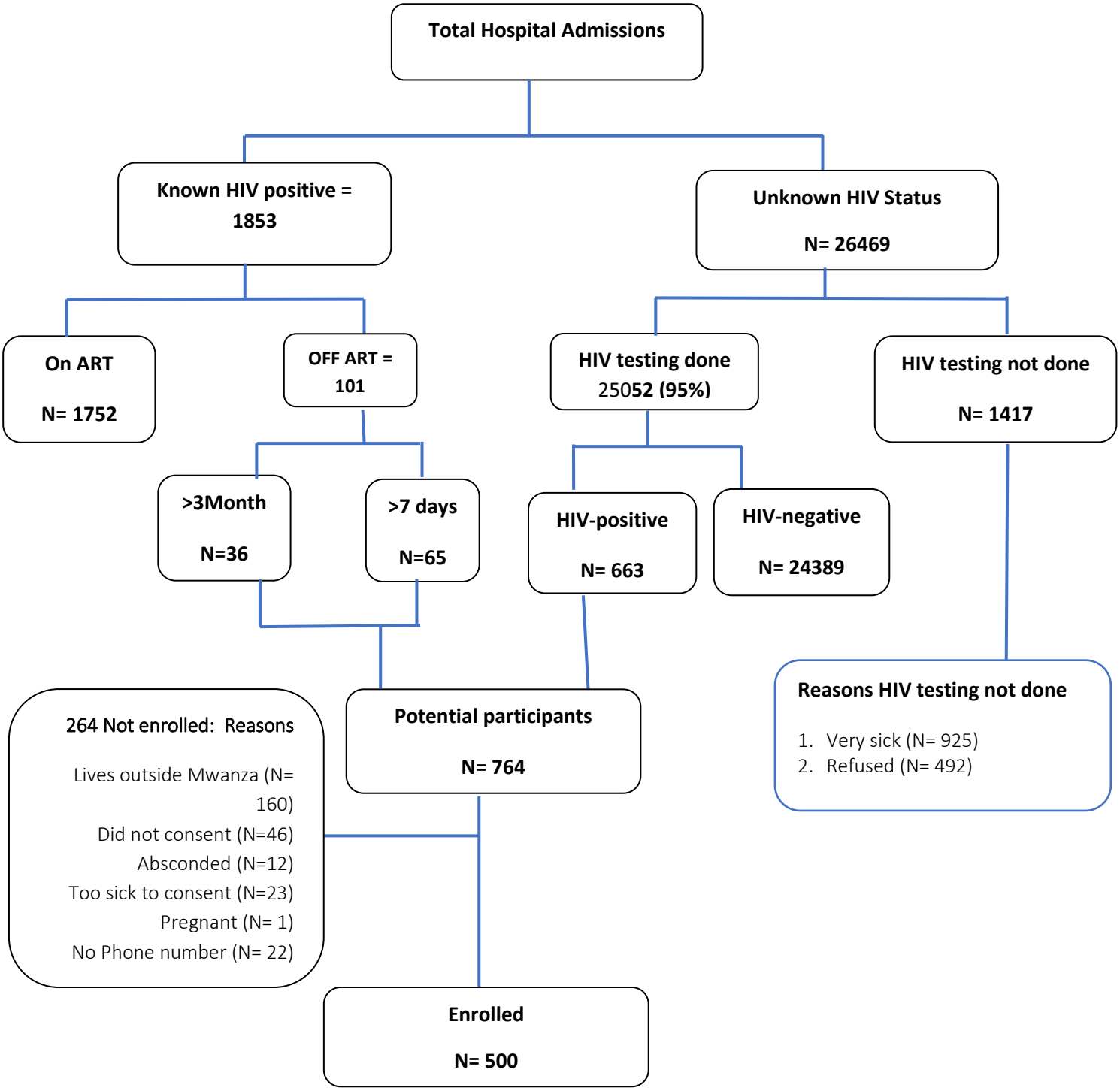
For any primary or secondary outcomes that are statistically significant, we will conduct subgroup analyses to see if the effect differs by sex. For time-to-event analyses, stratification and interaction p-values will be used to assess for effect modification. For categorical outcomes, stratification with Cochran–Mantel–Haenszel statistics will be used determine if the benefit/harm of the intervention differed significantly by age group (<> median age), sex or any other variables of particular interest.

Proposed tables and figures

Proposed Table 1: Baseline characteristics among participants in DARAJA clinical trial (with question numbers from the baseline survey)

Variable		**Intervention (N = 250)**	**Control (N = 250)**
		n(%)	**n(%)**
	Age in years (Q105 or Q106)		
sex	Sex (Q104)		
marst	Marital Status (Q108)		
head	Head of household?		
edlev	Education level (Q113)		
Insorc_recat	Occupation		
earn	Income level		
alcohol	Alcohol use (current / ever)		
antiretroviralsp	ART use (new diagnosis)		
healthinsu	Health Insurance		
	Yes		
	No		
bmi	BMI		
sbp	SBP		
cdcount	CD4		
haemoglobin	Hemoglobin		
admin_hosp	Admission Hospital		
depression	Depression (PHQ 9 < 10 vs. >=10)		
covid	COVID time period (pre COVID vs. COVID time period)		

Proposed Figure 1: Flow diagram for screening to enrollment from 20 study sites



Proposed Figure 2: Kaplan-Meier curves for survival by study arms with log rank test (primary outcome)

Proposed Figures 3A/3B: Nelson-Aalen curves for time to linkage and ART initiation in study arms with log rank test (secondary outcomes)

Proposed Table 2: Secondary outcome table for viral suppression, HIV clinic retention, and ART adherence

Secondary Outcome	**Intervention (N = ??*)**	**Control (N = ??*)**	**p-value** (Fisher's exact test)
	n(%)	**n(%)**	
Viral load suppression ("hvl_status")			
Suppressed			
Not suppressed			
HIV clinic retention at 12 months ("retention")			
Retained on ART			
Not retained on ART			
ART Adherence at 12 months ("adherence")			
100% Adherent			
Not adherent			

*** N= 250 minus participants who died / withdrew**