

# Statistical Analysis Plan: Randomized Trial for Youth Health Africa

## **Part 1: Assess cumulative impact by difference-in-difference analysis**

We will use linear models to measure the difference in outcomes for the intervention group as compared to the control group from a baseline period to study period.

$$Y = \beta_0 + \beta_1 \text{Time} + \beta_2 \text{Treat} + \beta_3 \text{DID}$$

### ***Parameters***

- Time represents the time period (baseline or study)
- 'Treat' represents the treatment group (control or intervention)
- 'DID' is calculated as Time\*Treat

### ***Parameter of Interest***

We are interested in  $\beta_3$  (difference in change between baseline and study period for the intervention versus the control group), i.e., the difference-in-difference.

### ***Hypothesis***

For each outcome:

- Null hypothesis:  $\beta_3=0$
- Alternative hypothesis:  $\beta_3 \neq 0$

This will be assessed at the 0.05 significance level using the standard "stats" package in R.

### ***Time Period:***

- The baseline period should mirror the study period in its calendar time and number of months.
- The first three months of this intervention will be considered a "run-in" period and will be excluded from the analysis.

### ***Measures:***

- Testing and treatment outcomes will be aggregated across all time periods (e.g., % tested for HIV = total tested for HIV over 9 months / total headcount for 9 months).
- Retention outcomes will not be aggregated across all time periods, as the denominator cannot be aggregated across months. Monthly means will be calculated for these indicators.

## **Part 2: Assess monthly variation using controlled, interrupted time series analyses**

We will use controlled, interrupted time series analyses for HIV service outcomes to assess differences in monthly proportions and trends between the intervention and control groups over time.

### ***Segmented Linear Model:***

$$Y_{ijt} = \beta_0 + \beta_1 \text{Time}_{ijt} + \beta_2 \text{Intervention}_{ijt} + \beta_3 \text{TimeAfterIntervention}_{ijt} + \beta_4 \text{Treat}_{ijt} + \beta_5 \text{Treat*Time}_{ijt} + \beta_6 \text{Treat*Intervention}_{ijt} + \beta_7 \text{Treat*TimeAfterIntervention}_{ijt} + E_{ij}$$

- Time will be measured in calendar months
- We will use restricted maximum likelihood estimates (REML)

- We will adjust for autocorrelation (AR1)
- We will adjust for clustering at the facility level
- We fit all models with random slopes and intercepts, where possible
- We will run these analyses using the “nlme” package in R

### **Model Parameters:**

- ‘Time’ represents the number of months since the start of the baseline period
- ‘Intervention’ represents the period (pre-intervention or intervention)
- ‘TimeAfterIntervention’ represents the number of months since the placement of the first interns
- ‘Treat’ represents the treatment group (control or intervention)

### **Parameters of Interest**

- $\beta_6$ : the difference-in-difference for immediate change, i.e., the immediate difference observed after intern placement in the intervention group minus the immediate difference observed in the control groups)
- $\beta_7$ : the difference-in-difference for slope change, i.e., the change in slope after the intervention was implemented for the intervention group minus the change in slope observed in the control group.

**Hypothesis:** For every outcome, the null hypotheses are as follows:

- $\beta_6=0$  (no difference in immediate change between groups)
- $\beta_7=0$  (or difference in trends between groups)

This will be assessed at the 0.05 significance level.

**Time Period:** We will use at least one-year pre-intervention and the full intervention period

### **For all analyses, outcomes are defined as follows**

Outcome	Calculation	Numerator	Denominator
<b>Testing</b>			
<b>% Tested*</b>	HTS_TST / Headcount	Number tested for HIV	Total headcount
<b>% Positive</b>	HTS_POS/ HTS_TST	Number testing positive for HIV	Number tested for HIV
<b>Treatment</b>			
<b>% Initiated in 14 days</b>	INITIATED_14DAYS / HTS_POS	Number initiating treatment within 14 days of diagnosis	Number testing positive for HIV
<b>Retention</b>			
<b>% Early Default</b>	ART_DEFAULT_EARLY / TX_CURR90	Number who did not return for treatment within 28 days of appointment	Number on treatment
<b>% Late Default</b>	ART_DEFAULT_LATE / TX_CURR90	Number who did not return for treatment within 89 days of appointment	Number on treatment
<b>% Loss to Follow-Up</b>	ART_DEFAULT_ULTF / TX_CURR90	Number of patients out of care for $\geq 90$ days with no outcome (i.e., on the unconfirmed lost to follow-up list)	Number on treatment