Statistical Analysis Plan: ATN *[TMI: ATN 146]*

PROTOCOL:	ATN *[146]*
PROTOCOL DATE:	
PROTOCOL VERSION:	
CLINICALTRIALS.GOV ID:	NCT03681912
PRINCIPAL INVESTIGATOR:	Sylvie Naar
SENIOR STATISTICIAN:	Jason E. Chapman
SPONSOR:	The Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) with co-funding from: National Institute on Drug Abuse (NIDA), National Institute of Mental Health (NIMH), and National Institute on Minority Health and Health Disparities (NIMHD)
SAP DATE:	6/30/2020
SAP VERSION:	1.0
SAP PREPARED BY:	SIU Analytic Core
AUTHORS:	

[The page numbers in the table of contents link to each section of the SAP. Once the SAP is complete, the table can be updated by right clicking the first page number in the table and choosing "Update field" then "Update entire table".]

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS 5
2. INTRODUCTION 6
2.1. BACKGROUND 6
2.2. OBJECTIVES/RESEARCH HYPOTHESES 6
2.2.1. Primary Objective 6
2.2.2. Secondary Objective 6
2.2.3. Other Objectives 6
2.3. OUTCOMES 6
2.3.1. Primary Outcomes 7
2.3.2. Secondary Outcomes 7
2.4. KEY UPDATES 7
3. STUDY METHODS 7
3.1. STUDY DESIGN 7
3.2. RANDOMIZATION 8
3.2.1. Primary Objective ..... **Error! Bookmark not defined.**
3.2.2. Secondary Objective *(as needed)* ..... **Error! Bookmark not defined.**
3.2.3. Other Objectives *(as needed)* ..... **Error! Bookmark not defined.**
3.3. SAMPLE SIZE AND POWER CALCULATION 8
3.3.1. Primary Objective 9
3.3.2. Secondary Objective *(as needed)* 9
3.3.3. Other Objectives *(as needed)* 9
3.4. STATISTICAL INTERIM ANALYSES 9
3.5. STOPPING GUIDANCE 9
3.6. TIMING OF FINAL ANALYSIS 10
3.7. TIMING OF OUTCOME ASSESSMENTS 10
4. STUDY POPULATION 11
4.1. STUDY POPULATION CHARACTERISTICS 11
4.1.1. Primary Objective 11
4.1.2. Secondary Objective *(as needed)* 11
4.1.3. Other Objectives *(as needed)* 11

4.2. SCREENING DATA 11
4.3. ELIGIBILITY CRITERIA 11
4.3.1. Primary Objective ..... **Error! Bookmark not defined.**
4.3.2. Secondary Objective *(as needed)* ..... **Error! Bookmark not defined.**
4.3.3. Other Objectives *(as needed)* ..... **Error! Bookmark not defined.**
4.4. RECRUITMENT 12
4.5. WITHDRAWAL AND LOST TO FOLLOW-UP 12
4.5.1. Withdrawal 12
4.5.2. Lost to Follow-up 12
5. STATISTICAL ANALYSIS 12
5.1. STATISTICAL PRINCIPLES 12
5.1.1. Confidence Intervals and P-values 12
5.1.2. Descriptive Statistics 12
5.1.3. Adherence to Intervention 13
5.1.4. Protocol Deviation 13
5.1.5. Analysis Population 13
5.2. SPECIFICATION OF KEY VARIABLES 13
5.2.1. Outcomes 13
5.2.2. Exposures 14
5.2.3. Covariates 14
5.3. EXPLORATORY ANALYSES 15
5.4. OUTCOME ANALYSES 16
5.4.1. Primary Objective 16
5.4.2. Secondary Objective *(as needed)* 17
5.4.3. Other Objectives *(as needed)* 17
5.5. MISSING DATA PROCEDURES 17
5.6. REPRODUCIBILITY 17
5.7. SOFTWARE 17
6. EFFICACY AND SAFETY ANALYSES 18
6.1. EFFICACY ANALYSIS 18
6.1.1. Primary Efficacy Outcome 18

6.1.2. Secondary and Other Efficacy Outcomes 18
6.1.3. Quality of Life Outcomes 18
6.1.4. Analytic Methods 18
6.2. SAFETY ANALYSIS 18
6.2.1. Adverse Events 18
6.2.2. Analytic methods 18
7. REFERENCES 19

1. LIST OF ABBREVIATIONS

AC	(SIU) Analytic Core
ART	Antiretroviral Treatment
ATN	Adolescent Medicine Trials Network for HIV/AIDS Interventions
CoP	Community of Practice
CT	Change Talk
DAP	Dynamic Adaptation Process
DWLD	Dynamic Wait-listed Design
EBP	Evidence-Based Practice
EPIS	Exploration, Preparation, Implementation, Sustainment Model
IRT	Item Response Theory
IS	Implementation Science
ISC	(SIU) Implementation Science Core
MC	(SIU) Management Core
MI	Motivational Interviewing
MOO	Manual of Operations
MPIs	Main Principle Investigators
PI	Principle Investigator
REC	(SIU) Recruitment and Enrollment Center
SC	Site Coordinator
sIRB	Single Institutional Review Board
SIU	Scale It Up
TMI	Tailored MI Implementation Intervention
TMI Team	Coordination Team, excluding Study Sites
YLH	Youth Living with HIV

2. INTRODUCTION

2.1. Background

The goal of this study is to test a multi-faceted Tailored Motivational Interviewing Implementation Intervention (TMI) in a hybrid type-3 implementation-effectiveness trial using a dynamic wait-listed design. The TMI intervention is based on the Dynamic Adaptation Process (DAP—which aims to balance flexibility and fidelity—to scale up an Evidence-Based Practice (EBP) in multidisciplinary adolescent HIV care settings.

2.2. Objectives/Research Hypotheses

2.2.1. Primary Objective

The objective of this study is to determine the effect of TMI on provider competency.

Specific Aims:

1) To determine the effect of TMI on provider MI competency, and secondarily HIV care cascade outcomes, in 10 multidisciplinary adolescent HIV care teams.
2) To compare internal facilitation plus Communities of Practice (CoPs) to CoPs alone in sustaining fidelity.
3) To explore the role of the barriers and facilitators to implementation identified in the EPIS model as these impact on provider MI competency.
4) To determine the cost effectiveness of TMI with or without internal facilitation sustainment by combining fidelity and cascade outcomes with monies spent on implementation strategies.

Primary Hypotheses:

1) MI competency will increase during the TMI phase relative to the baseline phase.
2) During the sustainment phase, the level of MI competence form the Implementation phase will be maintained. (Exploratory)
3) During the sustainment phase, CoP with IF will have better maintenance of MI competence relative to CoP only. (Exploratory)
4) Cascade-related outcomes—including viral load detectability, number of clinic visits among YLH, number of youth tested and linked to care—will improve during the implementation and sustainment phases relative to the baseline phase.

2.2.2. Secondary Objective

2.2.3. Other Objectives

2.3. Outcomes

[To list and briefly describe the outcomes of interest]

2.3.1. **Primary Outcomes**

The primary Outcome is MI Competency, which includes a raw average score and a criterion score (Beginner, Novice, Intermediate, Advanced). The MI Coach Rating Scale was developed and evaluated as part of Dr. Naar's NHLBI funded center grant and in our preliminary studies (NIMH R34 for TMI, ATN 128 Peacoc). It was developed and evaluated using methods based in Item Response Theory (IRT), specifically, Rasch models, Many-Facet Rasch Models (MFRMs), and item bifactor measurement models.. The MI-CRS is comprised of 12 items that are rated on a 4-point ordered categorical scale (*Poor*, *Fair*, *Good*, *Excellent*) by trained observational coders, supervisors, or other TMI team members. The ratings are based on observation of a provider-patient interactions, whether a recording of an actual interaction or a standardized patient interaction. Psychometric performance of the MI-CRS is detailed in Naar et al. (in press).

2.3.2. **Secondary Outcomes**

The secondary outcome is HIV Cascade Variables.

- Record/chart abstraction will be conducted for all patients in care at the site for
    - the 12 months prior to the start of implementation;
    - the 12 months during implementation; and
    - 6 months after the end of the implementation intervention.
- These data will include the date of diagnosis and care entry, antiretroviral adherence prescriptions, CD4 counts, viral load, and the CD4 and viral load testing dates.
- Retention in care variables will include: missed visits (using an absolute count or a minimum number of missed visits); Appointment adherence (proportion of scheduled visits that are kept); Persistence/constancy (a minimum standard of visits/time period, attending at least 1 visit every 90 days); and gaps in care (no more than 4 months without a visit). In addition, number of youth receiving HIV C&T services from the site as well as number of new diagnoses identified and linked to care will also be captured.

2.4. Key Updates

[To document any adjustments to the SAP as the manuscript progresses]

SAP Version	Changes Made	Rationale	Effective Date

3. STUDY METHODS

3.1. Study Design

The hybrid type-3 implementation-effectiveness trial used a Dynamic Wait-Listed Design (DWLD) to evaluate the impact of TMI and CoP on provider MI competency and HIV care cascade outcomes. Ten adolescent HIV sites in the United States were randomly assigned to one of five clusters, which determined the timing of the TMI implementation intervention. There

were three study phases, baseline, implementation, and sustainment, each of which included repeated measurements of provider MI competency. Across the 10 sites, there were 139 participating providers and cascade outcomes for [*N*] patients.

At the time of each randomization, the two sites were randomly selected as a cluster to enter the TMI implementation phase, with the other clusters remaining in the baseline phase. This continued until all blocks were randomized to the TMI implementation phase. After each cluster received one year of TMI implementation, a second randomization assigned the two sites in each cluster to the sustainment phase condition, either CoPs or CoPs plus internal facilitation monitoring and coaching. For each cluster, the duration of this phase was variable (e.g., 17 months for the first cluster, 6 months for the last cluster), with each ending at the end of study data collection (May 2020). MI competency was assessed on a quarterly basis across the baseline, implementation, and sustainment phases.

3.2. Randomization

With the DWLD, the 10 sites, in groups of two, were randomly assigned to one of five clusters. To allow sufficient time for scheduling and planning the initial workshop component, each wave of randomization occurred six months prior to the workshop month, which was three months prior to the first pre-intervention fidelity assessment. At each wave of randomization, two of the remaining (i.e., untrained) sites were randomly selected for the next wave of implementation. The specific randomization method involved several steps, each of which was implemented using SPSS software: (1) The full list of sites was entered, (2) each site had a dichotomous Intervention Status indicator (0 = No, 1 = Yes) to reflect whether it had been randomized to the implementation intervention. There were also fields to record the randomization date and initial workshop date. (3) For each wave of randomization, the data were sorted in ascending order by training status and site name. (4) A new random number variable was created for each wave. The assigned random value was obtained using the SPSS random number generator, set using the Mersenne Twister algorithm with a random starting point, and computed using the random uniform function with minimum and maximum values specified as 100000, 999999. (5) After a new random number was assigned to each site, the data were be resorted in ascending order by intervention status and the random number variable. (6) The first two sites (those with the lowest randomly assigned numbers) were selected for the next wave of training. This process was repeated for each subsequent wave of randomization. Drs. MacDonell and Chapman were responsible for implementing the randomization procedure, and TMI study staff were responsible for communicating the randomization results to sites.

3.3. Sample Size and Power Calculation

The final sample size was 10 sites, and 151 providers. For the HIV care cascade outcomes, there were 1462 individuals.

3.3.1. Primary Objective

For Aim 1 (i.e., the effect of TMI on provider MI competence), estimation of power for mixed-effects regression models is complex because of the number of parameters involved, and in the present case, this is further complicated by the DWLD. Traditionally, and for conventional mixed-effects power analysis software, intervention condition is modeled at the highest level, and the intervention effect is the cross-level interaction between condition and a level-1 growth term. In the present case, because each provider has measurements in each of three research phases—that is, the intervention condition changes over time for each provider—the terms for testing the intervention effect are modeled at level-1 (i.e., time and phase terms). As such, a special method was required to estimate power, which was based on Hox (2018). There were three steps: (1) Compute the *actual* sample size of observations. For the primary outcome of provider MI competence, with 10 sites, an average of 18.9 providers each, and an average of 6.5 measurements per provider, there were 1,237 observations (i.e., non-independent, nested observations). (2) Penalize the sample of observations for nesting effects using a reorganization of the design effect formula (i.e., $neff = n / \{1+nclus - 1\} \rho$), the result of which is the *effective* sample size: the sample size of truly independent observations after adjusting the observed sample size for nesting effects that decrease statistical power. This step was performed twice, first to adjust for repeated measurements within providers and then to adjust for providers within sites. The supplied values for ρ were based on estimates from a fully unconditional, three-level mixed-effects regression model for the MI competence outcome. Note that the final models will not estimate a random effect for sites (due to the small number), and instead, they will control site effects using fixed effect indicators. For the purpose of power analysis, the site penalty was applied to provide a more conservative estimate. For the provider MI competence outcome, with a nesting effect of $\rho = .35$ for repeated measurements within providers, the sample of 1,237 non-independent observations provides the statistical power of 420 independent observations. The sample of 420 observations, with a nesting effect of $\rho = .09$ for providers in sites, provides statistical power equivalent to 90 truly independent observations. (3) Power was then estimated for a conventional, single-level regression model with the effective sample size of independent observations, which indicated that the effective sample of 90 independent observations was sufficient to detect a small-to-medium effect of $R^2 = .09$. Thus, power is sufficient to detect the effects of interest, though it is important to note that, due to the specific features of the design, this method did not directly account for the number of parameters in the model or the reliability of estimated slopes.

3.3.2. Secondary Objective *(as needed)*

3.3.3. Other Objectives *(as needed)*

3.4. Statistical Interim Analyses

[To outline interim analyses to be conducted and to detail planned adjustment to significance level due to interim analyses]

3.5. Stopping Guidance

[To outline guidelines for stopping the trial early]

3.6. Timing of Final Analysis

The final analyses will be performed following completion of all data collection and processing of all data.

3.7. Timing of Outcome Assessments

The figure below illustrates the timing of outcome assessments across the course of the study, with shading to differentiate measurements in the baseline (light gray), implementation (medium gray), and sustainment (dark gray) phases. Outcome assessments began in the first month of the baseline phase for all sites and clusters. The assessments began, continued on a quarterly basis, and ended at the same points in study time for all sites and clusters. What differed across clusters was the duration of the baseline and sustainment phases, whereas the implementation phase was fixed at 12 months for all sites. Additionally, for all sites and clusters, the first month of the implementation phase included two measurements of MI competence, though by definition, the timing of this varied by cluster.

STANDARD PATIENT INTERACTION SCHEDULE

MONTH	Sep '17 0	Oct '17 1	Nov '17 2	Dec '17 3	Jan '18 4	Feb '18 5	Mar '18 6	Apr '18 7	May '18 8	Jun '18 9	Jul '18 10	Aug '18 11	Sep '18 12	Oct '18 13	Nov '18 14	Dec '18 15	Jan '19 16	Feb '19 17	Mar '19 18
Block 1 (TN & PA)	X			X	X=2		X			X			X			X			X
Block 1 (MI & SU)	X			X			X	X=2		X			X			X			X
Block 3 (2)	X			X			X			X	X=2		X			X			X
Block 4 (2)	X			X			X			X			X	X=2		X			X
Block 5 (2)	X			X			X			X			X			X	X=2		X
Block 6 (1)	X			X			X			X			X			X			X

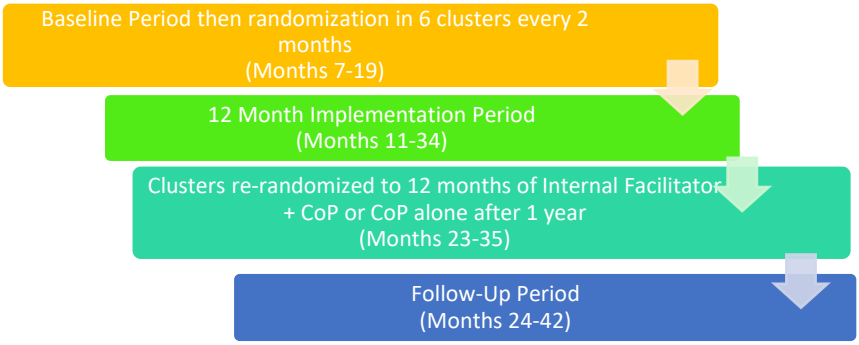
MONTH	Apr '19 19	May '19 20	Jun '19 21	Jul '19 22	Aug '19 23	Sep '19 24	Oct '19 25	Nov '19 26	Dec '19 27	Jan '20 28	Feb '20 29	Mar '20 30	Apr '20 31	May '20 32	Jun '20 33	Jul '20 34	Aug '20 35	Sep '20 36
Block 1 (TN & ?)			X			X			X			X			X			X
Block 1 (MI & SU)			X			X			X			X			X			X
Block 3			X			X			X			X			X			X
Block 4			X			X			X			X			X			X
Block 5			X			X			X			X			X			X
Block 6	X=2		X			X			X			X			X			X

Color Key

- Light Grey: Baseline Period until Initial Workshop based on randomization
- Medium Grey: Implementation Period (1 year post-first randomization)

- Dark Grey: Sustainment Period (1 year post-second randomization; except Block 5, which ends at 9 months, and Block 6, which ends at 6 months)

Study Timeline:



4. STUDY POPULATION

4.1. Study Population Characteristics

[Organized by objective (if different by objective)]
[To briefly outline study population characteristics; to include but not to be limited to:

- *Demographic characteristics*
- *Medical history*
- *Other background characteristics]*

4.1.1. Primary Objective

4.1.2. Secondary Objective *(as needed)*

4.1.3. Other Objectives *(as needed)*

4.2. Screening Data

NA

4.3. Eligibility Criteria

Inclusion Criteria:

All youth HIV providers (prevention and care) at our target sites that meet a minimum of four hours per week of patient contact are eligible to participate.

Exclusion Criteria:

Clinic personnel that have limited to no patient contact are excluded.

4.4. Recruitment

[To describe information to be included in the CONSORT flow diagram]

TMI study staff will work with the SC at each site to introduce the project and recruit participants by scheduling and conducting introductory meetings (via phone) with the SC and site PI and clinic staff. After the introductory meetings, the SC or PI from each site will send contact information (email, phone number) to the TMI study staff. The TMI study staff will contact participants via email and/or phone call to provide the information sheet and schedule of assessments. TMI study staff will also work with the SC and site PI to schedule the MI training session(s).

4.5. Withdrawal and Lost to Follow-Up

[To outline withdrawal/lost to follow-up details including but not limited to:

- *How withdrawal/lost to follow-up will be defined*
- *Level of withdrawal (e.g. from intervention and/or follow-up)*
- *Timing of withdrawal/lost to follow-up data*
- *Reasons and details of how withdrawal/lost to follow-up data will be presented]* [1]

4.5.1. Withdrawal

4.5.2. Lost to Follow-up

TMI Study staff who no longer wish to participate in follow-ups are lost to follow-up.

5. STATISTICAL ANALYSIS

5.1. Statistical Principles

5.1.1. Confidence Intervals and P-values

The test statistical for estimated effects will be the Wald statistic (i.e., β / *SE*), evaluated relative to a threshold of $p < .05$. To reflect the magnitude and precision of the unstandardized effect, 95% confidence intervals will be computed and reported for all estimated effects. To address multiple testing, planned contrasts will be specified to obtain all comparisons of interest from a single model for each outcome.

5.1.2. Descriptive Statistics

Standard descriptive statistics will be computed in preparation for final data analyses, and of these, the most relevant descriptive statistics will be included in the main outcome manuscript. Continuous variables will be described based on the mean, median, standard deviation,

minimum, and maximum; and categorical variables will be described based on observed frequencies across the range of possible categories.

5.1.3. Adherence to Intervention

[To specify details related to adherence to intervention including but not limited to:

- *Definition of adherence to the intervention and how this will be assessed including extent of exposure*
- *Duration of intervention exposure*
- *Description of how adherence to the intervention will be presented]*[2]

5.1.4. Protocol Deviation

[To specify details related to protocol deviation including but not limited to:

- *Definition of protocol deviations*
- *Description of which deviations will be summarized]*[2]

5.1.5. Analysis Population

Intention-to-treat analyses will be performed, with each provider retained in the original site, cluster, and intervention phase corresponding to the measurement occasion.

5.2. Specification of Key Variables

5.2.1. Outcomes

	Outcome Type *[e.g. primary, secondary, etc.]*	**Variable Name** *[To include Calculation/ Transformation Specifications, if derived]*	**Definition** *[to include reference to definition or name of author/coauthor that provided definition.* *To specify timings, specific measurements, units of measurement.]*	**Primary/Key Secondary Endpoints** *[if applicable; listed in order of importance]*	**Algorithm for Value Calculation in Case of Partially Complete Data**
MI Competency, Raw Average Score	Primary	MICRS_AV	MI Competency, Raw Average Score on the 12-Item MI Coach Rating Scale		Mean(MIRating1 to MIRating 12)

			(MI-CRS), Ranging from 1-4		
MI Competency, Criterion Score	Primary	MICRS_CR	MI Competency Level, Criterion Score (Beginner, Novice, Intermediate, Advanced), Ordered Categorical Criterion Score Based on Defined Thresholds for Competency Raw Scores		
HIV Care Cascade Related Outcomes, ***	Secondary	Viral Load			
…					

5.2.2. Exposures

	Exposure Type *[e.g. primary, secondary, etc.]*	**Variable Name** *[to include calculation/ transformation specifications, if derived]*	**Definition** *[to include timings, specific measurements, units of measurement]*	**Additional Details** *(as needed)*
[Exposure A]				
[Exposure B (as needed)]				
[Exposure C (as needed)]				
…				

5.2.3. Covariates

	Variable Name (or Calculation/	**Definition**	**Additional Details**

	Transformation Specifications if derived)	*[to include timings, specific measurements, units of measurement]*.	*(as needed)*
Implementation Phase Indicator	PH_IMP	Dummy-Coded Indicator for Measurements in the Implementation Phase	
Sustainment Phase Indicator	PH_SUS	Dummy-Coded Indicator for Measurements in the Sustainment Phase	
Linear Time	L_MO	Linear number of months from the start of the Baseline Phase	
Linear Time, Implementation	L_MO_IM	Linear number of months from the start of the Implementation Phase	
Linear Time, Sustainment	L_MO_SU	Linear number of months from the start of the Sustainment Phase	
Sustainment Phase Condition	COND_IF	Indicator for sites randomized to the CoP with IF condition in the Sustainment Phase	

5.3. Exploratory Analyses

Six types of exploratory and preliminary analyses will be performed: (1) There are too few sites and clusters to support estimation of random effects for those levels. To address that, the general strategy will be to include fixed effect indicators to control for systematic differences across clusters. However, preliminary tests will be performed to evaluate the extent of differences across both sites and clusters and to determine the most effective statistical controls for such differences. (2) Preliminary analyses will be performed to identify provider control variables for the final models. (3) As part of standard model-building steps, variance component estimates

from the mixed-effects regression models will be used to compute intraclass correlation coefficients to characterize the proportion of the total outcome variance that is attributable to each level of the model. (4) With three main intervention phases (i.e., Baseline, Implementation, Sustainment), alternative configurations of phase indicators are possible, and preliminary analyses will be performed to determine the final modeling strategy. (5) With each provider having repeated outcome measurements across multiple intervention phases, but also within phases, preliminary analyses will be performed to inform the appropriate modeling strategies for these measurements (e.g., estimating linear change within a phase versus the average level of the outcome for the phase). (6) The criterion score for the primary MI competency outcome will be evaluated both as a continuous outcome and an ordered categorical outcome, which will inform the final modeling strategy.

5.4. Outcome Analyses

5.4.1. Primary Objective

The primary MI competency outcome data are structured with repeated measurements (level-1) nested within providers (level-2). As noted previously, providers are also nested within sites, and for the dynamic wait-listed design, sites are nested within clusters. Systematic differences across clusters, and site-specific differences if indicated, will be controlled using dummy-coded fixed effect indicators. There are two other critical features of the design: (1) There are three intervention phases (i.e., Baseline, Implementation, Sustainment). (2) The intervention "condition" varies over time for each provider, site, and cluster. Also relevant is that there are repeated measurements of MI competency within all phases. Thus, the two-level mixed-effects regression model will be formulated with a series of indicators to differentiate the measurements occurring in each phase. For the primary models, these will be main effects only, and follow-up models may be performed with interactions to test for differential change over time in competency across the five clusters of sites. The primary parameters of interest will include the phase indicators and time polynomials. This formulation will test for an overall shift in the level of MI competence from the baseline to implementation and sustainment phases, and if supported, it will test for a shift in the rate of change in MI competence during each phase. Additional models will be performed to test for differences in MI Competence between the CoP and CoP with IF conditions in the Sustainment phase. The key model parameters, including regression coefficients, standard errors, confidence intervals, test statistics, degrees of freedom, probability values, variance components, and planned contrasts will be reported in tables and text. As noted previously, the analyses will follow and intention-to-treat approach. Subgroup analyses are not anticipated.

5.4.2. Secondary Objective *(as needed)*

5.4.3. Other Objectives *(as needed)*

5.5. Missing Data Procedures

There are two aspects of missing data procedures. The first relates to missing responses to individual items at a given measurement occasion, and the second relates to measurement occasions that are missed entirely by some participants. For TMI, the former is less of a concern, with low rates of observed missingness that can be remedied by standard scoring procedures (e.g., use of average scores rather than total scores). However, given the real-world implementation focus of TMI, there is substantial missingness of measurement occasions across providers. To address this, key considerations include the reasons for missingness, the amount of missingness, the nature of the missingness mechanism, and the resulting statistical remedies.

The primary outcome of MI competency was based on observer ratings across 12 items, and there were no missing item responses; thus, the primary missing data consideration pertains to missed measurement occasions. Across five blocks and ten sites, there were 189 providers. Of these providers, 99% had at least one MI competency measurement in the baseline phase (i.e., prior to the workshop), 75% had at least one measurement during the implementation phase, and 49% had at least one measurement during the sustainment phase. Given the decrease in completed assessments across the three phases, analyses will be performed to determine whether the likelihood of missingness was associated with baseline MI competency or other provider demographic characteristics. A pattern-mixture approach will also be used to determine whether the amount of change between baseline and implementation differed for providers who were missing sustainment data. Combined, the analyses will identify any potential evidence of a non-random missing data pattern that, if identified, can be controlled in subsequent analyses. Assuming there is no evidence of a non-random missingness mechanism, the analytic strategy will be to rely on maximum likelihood estimation, utilization of all available data, and estimation of individual slopes for each provider.

It is also important to note that the preceding description pertained only to the presence of *at least one* measurement in each of the three phases. However, based on the research protocol, providers were expected to have repeated measurements in each of the three phases. Compared to the possible number of measurements in each phase, the actual numbers revealed more considerable missing data. This missingness will be evaluated as previously described, with the addition of tests for differences in missingness *within* each phase.

5.6. Reproducibility

[To detail where data and/or code will be made available to the public (if applicable)][1,2]

5.7. Software

Data management, processing, and scoring for analysis will be performed using SAS, R, and SPSS software. Descriptive statistics will be computed using SPSS. Mixed-effects regression

models will be performed primarily using HLM software, with SuperMix software used to validate estimates for the ordinal MI competency outcome.

6. EFFICACY AND SAFETY ANALYSES

6.1. Efficacy Analysis

[To specify details of efficacy analyses to be conducted]

6.1.1. Primary Efficacy Outcome

[To specify the primary efficacy outcome and how it will be measured]

6.1.2. Secondary and Other Efficacy Outcomes

[To specify secondary and other efficacy outcomes and how they will be measured]

6.1.3. Quality of Life Outcomes

[To specify quality of life outcomes and how they will be measured]

6.1.4. Analytic Methods

[To specify methods to be used for efficacy analyses; may include the following (as needed):

- *Missing data procedures*
- *Sensitivity analysis (if applicable)]*

6.2. Safety Analysis

[To specify safety summaries/analyses to be conducted.]

6.2.1. Adverse Events

[To specify details of adverse events (AEs) including but not limited to:

- *Laboratory abnormalities*
- *Vital sign abnormalities*
- *Concomitant Medications*
- *Death*
- *Discontinuations for AEs*
- *Severity, expectedness, and causality]*

6.2.2. Analytic methods

[To specify methodological details of safety analyses including but not limited to:

- *How AEs will be coded and categorized*
- *How AE data will be analyzed/summarized*
- *What analyses/summaries will be included in manuscript]*

7. REFERENCES

[To list relevant references including but not be limited to:

- *Data Management Plan*
- *Trial Master File*
- *Statistical Master File*
- *Other Standard Operating Procedures or documents to be adhered to*
- *Sources regarding any nonstandard statistical methods]*[2]

1. Adolescent Medicine Trials Network for HIV/AIDS Interventions. ATN manual of policies and procedures. Version 1.0. March, 2018. https://sites.cscc.unc.edu/atn/policies-procedures-grp. Accessed October 1, 2017.
2. Hox, J. J., Moerbeek, M., & van de Schoot, R. (2018). Multilevel analysis: Techniques and applications (3rd ed.). New York: Routledge.
3. Naar, S., Chapman, J. E., Cunningham, P., Ellis, D., MacDonnel, K., & Todd, L. (in press). Development of the Motivational Interviewing Coach Rating Scale (MI-CRS) for health equity implementation contexts. *Health Psychology*.