Statistical Analysis Plan (SAP)

Minocycline as adjunctive treatment for treatment-resistant depression: a double blind, placebo-controlled, randomized trial (MINDEP2)

Principal Investigator:	Dr. Ishrat Husain MBBS, MD(Res.), FRCPsych(U.K.), FRCP(C)
Sponsor:	Centre for Addiction and Mental Health (CAMH) Toronto, ON, Canada

Protocol version:	13.1 (2023-05-25)
Research Ethics Board number:	135-2018
ClinicalTrials.gov identifier:	NCT03947827

SAP version:	2.0
Version date:	2026-03-08

Author:	Marcos Sanches, MSc Trial statistician, Biostatistics Core, CAMH

Table of contents

1. Document Information

SAP Version: 2.0
REB Number: 135-2018
Principal Investigator: Dr. Ishrat Hussain

2. Study Overview

2.1 Study Title

Minocycline as adjunctive treatment for treatment-resistant depression (TRD): a double blind, placebo-controlled, randomized trial (MINDEP2).

2.2 Study Design

- **Design:** Individually randomized parallel group placebo-controlled trial
- **Duration:** 48 months
- **Sample Size:** 100 participants (50 per arm)
- **Allocation Ratio:** 1:1
- **Blinding:** Double-blind (participant and investigator)
- **Participating Sites:** CAMH

- **Study Intervention:** 200mg of minocycline added to standard antidepressants compared to placebo added to standard antidepressants
- **Reference Therapy/Comparator:** Placebo
- **Duration of the Intervention:** 12 weeks (3 months)

2.3 Schedule of Assessments

Assessment	Screening	Baseline	Week 2	Week 4	Week 6	Week 8	Week 10	Week 12	Ad hoc
Demography	✓								
TAU/Concomitant medication	✓	✓	✓		✓			✓	✓
ATHF	✓								
SCID-5 (Medical Confirmation of Eligibility)	✓								
MoCA	✓								
Clinical Bloodwork	✓								
HRSD-17	✓	✓	✓		✓			✓	
Randomization		✓							
Physical health measures		✓							
CGI (Clinical Follow-Up)		✓	✓		✓			✓	
GAD-7		✓	✓		✓			✓	
WHOQOL-BREF		✓	✓		✓			✓	
Minocycline dispensed		✓	✓		✓				
Biomarkers (CRP, cytokines)		✓			✓			✓	
Adverse Event Checks			✓		✓			✓	✓
Telephone contact				✓		✓	✓		✓

2.4 Study Objectives

2.4.1 Primary Objective

To determine whether minocycline added to standard antidepressants for 3 months in patients with TRD leads to a reduction in depressive symptoms as measured by the 17-item Hamilton Rating Scale for Depression (HRSD-17) compared with placebo added to antidepressants.

2.4.2 Secondary Objective

To determine the safety and acceptability of minocycline added to antidepressants, compared to placebo added to antidepressants.

2.4.3 Exploratory Objectives

1. To determine whether TRD patients with evidence of inflammation at baseline are more likely to respond to minocycline
2. To determine whether changes in inflammatory cytokines and CRP predict a response to minocycline

2.5 Study Hypotheses

Primary Hypothesis (H1): Patients receiving adjunctive minocycline will experience a greater reduction in depressive symptoms as measured by Total HSRD-17 total score compared to patients receiving placebo added to antidepressants, at the end of 12 weeks as compared to baseline (week 0).

Secondary Hypothesis (H1a): Patients receiving adjunctive minocycline will experience a greater reduction in secondary outcomes: GAD, CGI and WHOQOL total score compared to patients receiving placebo added to antidepressants, at the end of 12 weeks as compared to baseline (week 0).

Secondary Hypothesis (H1b): Subgroup analysis. Exploratory analysis of subgroups difference in treatment effect at the end of the trial will be conducted.

Secondary Hypothesis (H1c): Response and Remission. Exploratory analysis of response (HRSD-17 change of -50% or more at 12 weeks) and remission (HRSD-17 score 7 or lower at week 12) will be reported.

Secondary Hypothesis (H1d): Adherence analysis. Descriptive analysis of adherence to the Investigational Product (IP) and comparison of adherence between groups.

Secondary Hypothesis (H2): Adverse effects (AE) will be quantified in minocycline and placebo groups as well as group differences in total AE and categories (mild, moderate, severe).

Exploratory Hypothesis (H3): Higher baseline plasma pro-inflammatory cytokines (e.g., CCL-2, TNF-α, IL-1, IL-6, IL-8, IL-12, IL-23 levels) and C-Reactive Protein (CRP) will predict response to minocycline as defined by the change in HRSD-17 total scores from week 0 to week 12.

Exploratory Hypothesis (H4): Changes in plasma pro-inflammatory cytokines levels and CRP will mediate treatment response to minocycline.

3. Study Endpoints

3.1 Primary Endpoint

Change in HRSD-17 total score from baseline to week 12

- HRSD-17 is assessed at: Screening, Baseline (Week 0), Week 2, Week 6, and Week 12
- Range: 0-52 (higher scores indicate more severe depression)
- The baseline value is defined as the assessment at Week 0 (randomization visit), not screening

3.2 Secondary Endpoints

The secondary endpoints are measured at baseline, weeks 2, 6 and 12. They are:

- GAD (Generalized Anxiety Disorder-7)
- CGI (Clinical Global Impression)
- WHOQOL (World Health Organization Quality of Life)

3.3 Safety Outcomes

Adverse events (AE): Frequency of AE in total and by categories (mild, moderate, severe).

3.4 Exploratory Endpoints

1. **Baseline inflammatory markers as moderators:** Baseline CCL-2, TNF-α, IL-1, IL-6, IL-8, IL-12, IL-23 and CRP
2. **Change in inflammatory markers as mediators:** Change from baseline CCL-2, TNF-α, IL-1, IL-6, IL-8, IL-12, IL-23 and CRP

4. Sample Size and Power

A sample of 100 participants (50 per group) provides 80% power to detect a standardized effect size Cohen's f = 0.13, which represents a small effect size. This calculation assumes 20% dropout rate (n = 40 completers per group), uses two-tailed tests with α = 0.05 and is based on a repeated measures ANOVA model and testing the group × time interaction effect. Power calculation w conducted using G*Power 3.1.9.2, prior to data collection.

5. Analysis Populations

- **Randomized Analysis Set (RAS):** The RAS includes all randomized participants in the groups they were assigned to and will be used for sample characterization and possibly sensitivity analysis.
- **Full Analysis Set (FAS):** The FAS includes all randomized participants who received at least one dose and with at least one post-baseline assessment, following ITT principle (i.e., analyzed according to their randomized treatment assignment, regardless of treatment received or protocol adherence). The FAS will be used for all efficacy analyses, unless otherwise specified.

- **Complete Response Set (CRS):** The CPS includes participants who have outcome response data at all time points. The CRS will be used for sensitivity analyses and for the analysis of Response and Remission rates.
- **Safety Analysis Set (SAS):** The SAS includes all participants who received at least one dose of study medication (minocycline or placebo), analyzed according to treatment actually received. This is the primary population for safety analyses of AE.

6. General Statistical Principles

6.1 Statistical Software

All analyses will be conducted using R (version ≥4.5.1) and Mplus version 9. The following are the main packages used in R:

- `lme4` and `nlme` for mixed effects models
- `emmeans` for marginal means and contrasts
- `lmerTest` for p-values in linear mixed models
- `tidyverse` for data cleaning and graphs.
- Report of statistical analysis results will include `sessionInfo()` listing all packages and version used.

6.2 Statistical Inference

All statistical tests will use a two-sided significance level of $\alpha = 0.05$ unless otherwise specified. 95% confidence intervals (CIs) will be reported for all estimates unless otherwise specified.

For the primary hypothesis test (group difference in change from baseline to Week 12), no adjustment for multiplicity is required as this is a single pre-specified comparison.

For secondary analyses p-values will be reported but not used for decisions. If adjustment is used for multiple tests, they will be clearly stated.

Exploratory analyses will not be adjusted for multiplicity, and results will be interpreted as hypothesis-generating.

6.3 Missing Data

The primary and most of the other analyses use a likelihood-based mixed effects model, which provides valid inference under the Missing at Random (MAR) assumption without requiring imputation. Detailed approach for missing values handling for primary analysis are described in the respective sections.

7. Descriptive Analysis

Descriptive analysis will be conducted and reported for the variables to be used in the analysis prior to the analyzes involving statistical inference. This phase will be used to inspect variable contents, flag outliers, atypical or suspicions data points. No modification of the original data that involves changing records will be made without approval of the Principal Investigator, and those changes made will be fully documented and justified. Outliers and atypical data points will be handled via model diagnostic and sensitivity analysis which to the extent of possible are pre-specified. Changes related to variable recoding and new variables generation from existing information in the data will be conducted and documented via reproducible analysis code.

7.1 Summary statistics

Baseline characteristics will be summarized by treatment group of interest using descriptive statistics and distribution graphs.

- Group of interest
    - Study groups (minocycline and placebo)
    - Study completion groups (Completers and non-completers)
- Continuous variables:
    - Mean
    - Standard Deviation (SD)
    - Median
    - First and third quartiles
    - Minimum
    - Maximum
    - Histogram
    - Sample Size
    - Count of missing values
- Categorical Variables:
    - Sample Size
    - Percent
    - Count of missing values
- Statistical Tests

In the context of baseline variables, the RAS will be used and statistical tests are provided as descriptive measure of effect or measure of evidence within the context of the test, but will not be used for decisions.

- **Continuous variables:** Mann-Whitney U test (non-parametric)
- **Categorical variables:** Fisher's exact test

- **All Variables:** Standardized mean difference between groups (SMD)

Definition SMD: For continuous variables, the usual Cohen's d:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p}$$

with pooled standard deviation

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where:

- $\bar{x}_1, \bar{x}_2$ = group means
- s_1, s_2 = group standard deviations
- n_1, n_2 = group sample sizes
- s_p = pooled standard deviation

For categorical variables:

$$\text{SMD} = \sqrt{\sum_{k=1}^{K} \frac{(p_{1k} - p_{2k})^2}{p_k}}$$

where:

- $k = 1, \ldots, K$ indexes category levels
- p_{1k} and p_{2k} are proportions in treatment groups 1 and 2
- $p_k = (p_{1k} + p_{2k})/2$ is the pooled proportion

For binary variables, this reduces to the familiar form:

$$\text{SMD} = \frac{p_1 - p_2}{\sqrt{p(1-p)}}$$

8. Primary Analysis (H1)

8.1 Outcome and Estimand

The primary estimand is the treatment policy estimand of the effect of assignment to minocycline versus placebo on depressive symptom severity in the target population.

- **Population:** Patients who meet the trial's inclusion/exclusion criteria (i.e., the population represented by randomized participants).

- **Treatment conditions:** Assignment to minocycline augmentation versus placebo augmentation (as randomized).
- **Variable (endpoint):** Change from baseline in HRSD-17 total score at Week 12.
- **Population-level summary:** Difference in means (minocycline minus placebo) in the change from baseline to Week 12.

- **Strategies for intercurrent events:** Intercurrent events are treatment discontinuation, non-adherence, use of prohibited rescue medication, switching to alternative therapy. The value of the variable is used regardless of whether these events occur.

The estimand reflects the treatment effect under the intention-to-treat principle: the difference that would be observed if all randomized participants were followed and assessed as randomized, regardless of adherence, dropout, or protocol deviations (where feasible). Missing data (e.g., due to loss to follow-up or withdrawal of consent) will be handled under the assumption that data are missing at random (MAR) conditional on observed data and relevant covariates, using mixed model estimated via full information maximum likelihood.

8.2 Statistical Model

8.2.1 Model Specification

A **Linear Mixed Effect Model (LMM)** will be fitted to the repeated HRSD-17 measurements with the following specification:

Model formula:

$$Y_{ij} = \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{Time}_j + \beta_3 \left(\text{Group}_i \times \text{Time}_j\right) + \beta_4 \text{BaselineHRSD}_i + b_{0i} + \epsilon_{ij}$$

Where:

- Y_{ij} = HRSD-17 score for participant i at time point j
- Group_i = Treatment group (0 = Placebo, 1 = Minocycline)
- Time_j = Time point coded in three levels:
  - 2 = Week 2
  - 6 = Week 6
  - 12 = Week 12
    - The time point for the primary analysis is defined by the planned visit time.
- $\text{Group}_i \times \text{Time}_j$ = Interaction between group and time
- BaselineHRSD_i = Baseline HRSD-17 score (continuous, centered at sample mean)
- $b_{0i} \sim N(0, \sigma_b^2)$ = Random intercept for participant i
- $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ = Residual error

Fixed effects:
- Intercept (β_0)
- Treatment group (β_1)
- Time (β_2, categorical)
- Group × Time interaction (β_3)
- Baseline HRSD-17 (β_4)

Random effects:
- Random intercept for each participant

Residual structure: Independent and identically distributed errors (conditional on random effects)

8.2.2 Rationale for Model Specification

1. **Baseline as covariate:** Baseline HRSD-17 is included as a covariate to improve precision and account for baseline imbalance (ANCOVA-type approach within mixed model framework, where baseline HRSD-17 is not part of the outcome).

2. **Time as categorical:** Time is coded as a categorical variable (2, 6, 12 weeks) because the hypothesis targets a specific time point: 12 weeks. This avoids the assumption of linear time effect. The categorical specification of time assumes that every participant were measured at the same time point, or equivalently, that differences in time of measurement don't impact the outcome which is expected to be reasonable for this trial.

3. **Random intercept:** A random intercept accounts for correlation of repeated measures within participants.

4. **No additional correlation structure:** Conditional on the random intercept, residuals are assumed to be independent. This assumption is expected to be reasonable due to the use of only 3 time points.

8.2.3 Model Fitting

The model will be fitted using **Restricted Maximum Likelihood (REML)** estimation via the `lmer()` function from the `lme4` package in R:

```
library(lme4)
library(lmerTest)
library(emmeans)

analysis_data <- long.data %>%
   mutate(Time = factor(Time)) %>%
   arrange(id,Time) %>%
   group_by(id) %>%
   mutate(BaselineHRSD = first(HRSD17)) %>%
```

```
   ungroup() %>%
   filter(Time != "0")

primary_model <- lmer( HRSD17 ~ Group +
                 Time +
                 Group:Time +
                 BaselineHRSD +
                 (1|id),
                 data = analysis_data)
```

Note: Lack of convergence or boundary problems related to singularities are not uncommon for mixed models, but not expected here. If happening, they will be handled first by using different estimation algorithm and the `nlme`. If that fails, the robustness of the model will be tested by comparing it with a simpler model that solves the estimation problem fixing the offending parameter to zero (e.g., removing it).

8.3 Primary Hypothesis Test

8.3.1 Hypotheses

- **Null Hypothesis:** H_0: There is no difference between minocycline and placebo in the mean change in HRSD-17 score from baseline to Week 12.
- **Alternative Hypothesis:** H_A: There is a difference between minocycline and placebo in the mean change in HRSD-17 score from baseline to Week 12.

8.3.2 Test Statistic

The primary hypothesis will be tested using a **linear contrast** of the estimated marginal means at Week 12:

$$\Delta_{12} = \left[\text{EMM}_{\text{Minocycline, Week 12}} - \text{EMM}_{\text{Placebo, Week 12}}\right]$$

Note: This contrast will be computed using the `emmeans` package. Kenward-Roger method for degrees of freedom is the default in `emmeans` and will be used for degree of freedom of the t-statistic.

```
# Estimate marginal means at Week 12.
emm_week12 <- emmeans(primary_model,
                      specs = ~ Group | Time,
                      at = list(Time = "12"))

# Contrast: Minocycline - Placebo at Week 12
primary_contrast <- contrast(emm_week12,
                             method = "pairwise",
                             adjust = "none")

# Results
summary(primary_contrast, infer = TRUE)
```

8.3.3 Decision Rule

The null hypothesis will be rejected if the two-sided p-value for the contrast is less then 0.05.

Additionally, we will report:
- Estimated marginal means (EMMs) for each group at Week 12 with 95% confidence intervals, and conditional on average baseline HRSD-17 - Mean change from baseline to Week 12 for each group with 95% confidence intervals, conditional on mean baseline HRSD-17 - Effect size: Cohen's d calculated as the estimated difference in change (from model) divided by the pooled standard deviation of the raw change score, which scales the treatment effect relative to the natural change from baseline to week 12.

8.4 Model Diagnostics

The following diagnostic checks will be performed on the fitted primary model, their results reported as well as results of sensitivity analyses, when conducted.

8.4.1 Normality

Check: Visual inspection of Q-Q plot and histogram of standardized residuals.

- **Criteria:**
  - If severe departure: We will investigate the probably source of model mis-specification, refit the model and report results as sensitivity analysis.
  - HRSD-17 can have truncated distribution in some populations, which makes the normal model not appropriate. In this case truncation is not caused by censoring. We will use mixed ordinal regression models as primary sensitivity analysis for the HESD-17 score. Alternative models, like beta distribution on re-scaled HRSD-17 score may also be considered.

8.4.2 Homoscedasticity

Check: Visual inspection of plot of absolute standardized residuals vs. fitted values; Levene's test for equality of variances across groups and time points will be conducted manually.

- **Criteria:**
  - If evidence of non-constant variance, fit a model with heterogeneous residual variance and report results and report the results as sensitivity analysis.

R code

```
analysis_data$resids <- residuals(primary_model, type = "pearson")
analysis_data$abs_resids <- abs(analysis_data$resids)
```

```
levene_mixed <- lmer(abs_resids ~ Group + (1|id),
                     data = analysis_data)
summary(levene_mixed)

ggplot(analysis_data, aes(x = fitted(primary_model), y = sqrt(abs(resids))))
+
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", color = "red") +
  labs(title = "Scale-Location Plot",
       x = "Fitted Values",
       y = "Square Root of |Pearson Residuals|")
```

8.4.3 Influential Observations and individuals

Check: Cook's distance for data points and individuals calculated using `influenceME` package. DFBETAS for group × time interaction.

- **Criteria:**
  - Histogram to visualize the distribution of Cook's distance and DFBETAS. Points standing out are flagged as influential
  - Document influential cases (table with id and values for variables in the model)
  - Verify data accuracy
  - Re-fit model excluding influential cases as sensitivity analysis and report results.

8.4.4 Outliers

Check: Standardized residuals

- **Criteria:**
  - Check histogram for cases that stand out with special attention for values larger than 3
  - Document outliers (table with id and value for variables in the model)
  - Verify data accuracy
  - Conduct sensitivity analysis excluding outliers and report results.

8.4.5 Random Effects Distribution

Check: Visual inspection of histogram of random intercepts (BLUPs)

- **Criteria:**
  - Search for source of model mis-specification, fix the model and report results with alternative model as sensitivity analysis
  - Mixed models are robust to this violation so we will only consider severe departure from normality.

8.4.6 Linearity of Baseline HRSD-17

Check: scatter plot of residuals against baseline HRSD-17 with smoothing Loess line

- **Criteria:**
  - If departure from linearity, adjust alternative model with spline function of baseline HRSD-17 and report results.

8.5 Missing Data Handling

8.5.1 Primary Approach

The primary analysis uses a **likelihood-based mixed effects model**, which provides valid inference under the Missing at Random (MAR) assumption.

8.5.2 Missing Data Pattern Analysis

Characterization of missing data in the outcome HRSD-17:

1. **Frequency:** Tabulate number and percentage of missing HRSD-17 assessments by treatment group and time point
2. **Patterns:** Identify monotone vs. non-monotone missingness patterns
3. **Predictors:** Compare baseline characteristics between participants with and without missing data
4. **Informative dropout:** Compare treatment groups on dropout rates using logistic regression. Report proportion of dropout in each group, Odds Ratio and Risk Difference.

8.5.3 MAR Assumption Assessment

While MAR cannot be proven, it will be made more plausible by:

1. Including baseline HRSD-17 as a covariate (MAR conditional on baseline severity)
2. Examining whether missingness is associated with previous HRSD-17 values
3. Testing whether missingness rates differ by treatment group (if significant imbalance, MAR may not hold)

8.5.4 Sensitivity Analyses for Missing Data

If >10% of participants have missing week 12 outcome data, the following sensitivity analyses will be conducted:

8.5.4.1 Multiple Imputation

As a sensitivity analysis to assess the robustness of the primary analysis to the handling of missing outcome data, the primary model will be re-estimated using multiple imputation (MI) for missing HRSD-17 values.

The MI analysis will be conducted using the Randomized Analysis Set (RAS), including all randomized participants regardless of treatment exposure or availability of post-baseline assessments. Missing post-baseline HRSD-17 scores will be imputed under the Missing at Random (MAR) assumption using a multiple imputation by chained equations (MICE) procedure.

The imputation model will include variables expected to be associated with missingness and/or outcome, including:

- Treatment group (minocycline vs placebo)
- Baseline HRSD-17 score
- Observed HRSD-17 scores at Weeks 2, 6, and 12
- Time indicators
- Additional baseline covariates associated with dropout indicator at SMD > 0.2.

Continuous HRSD-17 scores will be imputed using predictive mean matching (PMM) to preserve the distribution of observed values.

A total of 50 imputed datasets will be generated. Each imputed dataset will be analyzed using the same Mixed Model for Repeated Measures (MMRM) specified for the primary analysis (including treatment group, time, group × time interaction, and baseline HRSD-17 as a covariate). Estimates from the imputed datasets will be combined using Rubin's rules to obtain pooled parameter estimates, standard errors, confidence intervals, and p-values.

The treatment effect estimate at Week 12 from the MI analysis will be compared with the primary MMRM result to evaluate the robustness of conclusions to the handling of missing data.

8.5.4.2 Pattern Mixture Model (MNAR)

A pattern mixture model sensitivity analysis will be conducted under Missing Not at Random (MNAR) assumptions if the primary analysis demonstrates a statistically significant treatment effect ($\alpha = 0.05$). Missing HRSD-17 outcomes will be assumed to differ from MAR-based predictions by a shift parameter (δ), representing worse outcomes among participants with missing data. Initially, δ will be applied equally across treatment arms and varied from 0 to +10 points on the HRSD-17 scale. The tipping point will be defined as the smallest value of δ at which the primary treatment effect loses statistical significance. If statistical significance is maintained across this range, an additional sensitivity analysis will apply δ selectively to the intervention arm to evaluate robustness to differential MNAR mechanisms.

8.5.4.3 Complete Case Analysis

Primary analysis will be repeated including only participants with complete HRSD-17 data at all time points (Baseline,Weeks 2, 6, 12), which is the CAS.

8.5.4.4 Covariates associated with missingness

The primary analysis will be repeated using maximum likelihood estimation and adjusting for all variables found to be associated with missingness at standardized mean difference (SMD) larger than 0.2. This make the MAR assumption given covariates more plausible. The analysis use the FAS.

8.6 Sensitivity Analyses for Primary Endpoint

8.6.1 Covariate Adjustment for Baseline Imbalance

If any baseline characteristic differs between groups with SMD > 0.20, they will be adjusted for in the primary model. This analysis use the FAS.

8.7 Protocol Deviations

Protocol deviations are departures from the study protocol occurring during the conduct of the trial. Deviations will be identified and documented by the study team prior to database lock and classified as major or minor according to predefined criteria established by the study investigators and clinical monitoring team.

Examples of protocol deviations may include, but are not limited to:

- Violation of key inclusion or exclusion criteria
- Use of prohibited concomitant medications
- Incorrect study treatment administration
- Failure to receive study medication
- Substantial deviations from the scheduled visit window
- Other departures from protocol procedures judged relevant by the study team

All protocol deviations will be listed and summarized by treatment group. Major protocol deviations will be summarized separately from minor deviations.

The primary efficacy analysis will follow the intention-to-treat principle using the Full Analysis Set (FAS) and will therefore include participants regardless of protocol deviations.

9. Secondary Analyses

These analyzes are secondary and focus will be on estimation. P-values are reported as a measure of evidence within the context of the model and its assumptions, possibly used to rank evidence, it is not intended to be dichotomized as significant/not significant. This analysis uses the FAS.

9.1 Secondary Outcomes (H1a)

Three scales are collected as secondary outcomes: GAD-7, CGI and WHOQOL. These outcomes are collected at the same time points as the primary outcome HRSD-17 and their analysis will use the same ANCOVA-style mixed effect models. As part of the descriptive analysis, trajectory plots will be used as well as histograms, separated by group and time point. Descriptive statistics including means, standard deviation, maximum and minimum will be reported by group and time point. The primary effect of interest will be the group difference at week 12, estimated using estimated marginal means with point estimates and 95% confidence intervals reported. Graphs of average estimated trajectories by group will also be reported regardless of p-values. Alpha = 0.10 will be used due to the exploratory nature of this analysis.

9.1.1 GAD-7

Model: Same model as the primary analysis. The primary model will be linear mixed effect model that assumes normality for total score. Model residuals will be inspected for normality, heterocedasticity and outliers and alternative model specifications may be used. If the default normal model assumption shows severe violation, beta mixed model will be used after normalizing total GAD-7 scores to the [0-1] interval.

9.1.2 CGI

Model: The primary model will be ordinal linear mixed effect model that assumes ordinal outcome and proportional odds. It will otherwise follow the same specifications for the primary model.

9.1.3 WHOQOL

Model: Same model as the primary analysis. The primary model will be linear mixed effect model that assumes normality for total score. Model residuals will be inspected for normality, heterocedasticity and outliers and alternative model specifications may be used. Transformations (log and square root) will be considered if normal assumption is severely violated.

9.2 Subgroup Analysis (H1b)

We will explore the moderation of the treatment effect by demographics and clinical variables of interest, by adding these variables as predictors in the primary model interacting with the Group by Time interaction. The so called subgroup variables will enter the model in their natural form, that is, continuous or categorical. In the case of continuous variables, model diagnostic will be used to inspect for evidence of non-linear moderation. Estimated marginal means will be used to compare subgroups at week 12 regarding their treatment effect. Graphs for interpretation of the moderation effect using marginal means will be created if the three-way interaction is significant at alpha = 0.10 level (exploratory).

- Subgroups
    - Age in years
    - Gender
    - Biological Sex
    - Elevated baseline CRP, i.e. +/- 3mg/L
    - Baseline HRSD-17
    - Baseline BMI

Each moderation analysis will be conducted using a separate model and all results will be reported regardless of p-values.

9.3 Response and Remission Rates (H1c)

To support clinical interpretation and decision, some additional outcomes related to HRSD-17 are reported. These analyses only include participants with HRSD-17 scores at week 12. A sensitivity analysis will be conducted coding dropouts as non-responders and non-remitters to account for possible imbalances in dropout rate. This analysis use the CAS.

9.3.1 Response Rate

Definition: Proportion of participants achieving ≥50% reduction in HRSD-17 score from baseline to Week 12.

Analysis:

1. Calculate binary response indicator for each participant
2. Compare groups using logistic regression, adjusting for baseline HRSD-17
3. Estimate response rate in each group using marginal means on response scale.

Model:

$$\text{logit}\left(P\left(\text{Response}_i = 1\right)\right) = \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{BaselineHRSD}_i$$

Test: Wald test for β_1, two-sided α = 0.05

- **Report:**
    - Response rates (% and n/N) by group
    - Odds ratio with 95% CI
    - Risk difference with 95% CI
    - Number needed to treat (NNT) with 95% CI

R code:

```
# Calculate response
response_data <- long.data %>%
  arrange(id,Time) %>%
```

```
  group_by(id) %>%
  mutate(BaselineHRSD = first(HRSD17)) %>%
  filter(Time %in% c(0, 12)) %>%
  pivot_wider(id_cols = id:BaselineHRSD,
              names_from = Time,
              values_from = HRSD17,
              names_prefix = "Week") %>%
  mutate(
    Change = Week12 - Week0,
    Response = if_else(Change / Week0 <= -0.5, 1, 0)
  )

# Logistic regression
response_model <- glm(Response ~ Group + BaselineHRSD,
                      data = response_data,
                      family = binomial)
summary(response_model)

# Odds Ratio and CI
exp(coef(response_model)["GroupMinocycline"])
exp(confint(response_model)["GroupMinocycline", ])

# Probability of responders in each group
emm_probs <- emmeans(response_model, ~Group,
          type = "response",infer = c(T,T))

# Risk Difference
risk_dif <- contrast(regrid(emm_probs, transform = "unlink") ,
                         method = "pairwise",
                         infer = TRUE  )

# NNT is 1/Risk Difference
```

9.3.2 Remission Rate

Definition: Proportion of participants achieving HRSD-17 score ≤7 at Week 12

Analysis: Same approach as response rate

Model:

$$\text{logit}\big(P(\text{Remission}_i = 1)\big) = \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{BaselineHRSD}_i$$

9.4 Longitudinal Treatment Effect (Descriptive)

Explicit additional reporting from primary model.

- The primary model will be used for:

- Report time × group interaction p-value using type 3 F-test from default anova function in `lmerTest` package
- Report marginal HRSD-17 means and 95% confidence intervals for both groups and time points after baseline
- Report group differences in HRSD-17 for every time point after baseline
- Create graph with average trajectory for each group as well as average change from baseline estimated with baseline at mean.

This is exploratory analysis and p-values will not be adjusted.

9.5 Adherence Analysis (H1d)

Adherence to the investigational product (IP) will be summarized descriptively overall, by treatment group, and by time point. The primary adherence analysis will use the same mixed model framework as the primary efficacy analysis, with adherence (%) as the outcome, and fixed effects for treatment group, time, and treatment × time interaction. Estimates will be reported as model-adjusted means (least-squares means) with 95% confidence intervals. The treatment × time interaction will allow assessment of whether adherence differs over time between groups.

10. Safety Analyses (H2)

10.1 Adverse Effects (AE)

Outcome: Frequency of adverse effects by category (mild, moderate, severe) and in total.

Analysis: Separate analysis will be conducted for each AE category and for the total number of AE. Generalized linear Model with negative binomial distribution, log link and log-number of weeks as offset. The analysis primarily aims to estimate the rate ratio of AE between the two groups in the post-treatment period, with reporting of within group estimates of frequency of AEs. Exact 95% Clopper-Pearson confidence interval for proportion of individuals with at last one AE in each group will also be reported, with Fisher's Exact test for group comparison and reporting of Odds Ratio as association measure. All estimated results are reported with 95% confidence intervals.

Model:

$$\log(\mu_i) = \beta_0 + \beta_1 \text{Group}_i + \log(\text{exposure}_i)$$

- Where:
  - μ_i = Expected count of AEs for participant i
  - Exposure offset accounts for variable follow-up time
  - β_1 is the coefficient of interest

R code:

```r
# Data for model
data.ae <- long.data %>%
         mutate(nweeks = as.numeric(as.character(Time))) %>%
         mutate(nweeks = nweeks - lag(nweeks)) %>%
         filter(Time != "0") %>%
         mutate(Time = droplevels(Time))

summary_data <- data.ae %>%
  group_by(id, Group) %>%
  summarise(
    total_AE = sum(AE_count, na.rm = T),
    total_weeks = sum(nweeks, na.rm + T),
    .groups = "drop"
  )

nb_model <- glm.nb(
  total_AE ~ Group + offset(log(total_weeks)),
  data = summary_data
)

# Fisher's test pvalue
fisher.test(summary_data$Group, summary_data$any_AE)

# Exact 95% Confidence Interval (do for both group)
binom.test(n_AE_Placebo, N_Total_Placebo, conf.level = 0.95)
```

11. Exploratory Analyses (H3 and H4)

11.1 Exploratory Hypothesis 3: Baseline Inflammation as Moderator of the treatment effect on HRSD-17.

11.1.1 Variables

- **Baseline inflammatory markers:**
  - C-Reactive Protein (CRP, mg/L)
  - CCL-2
  - TNF-α
  - IL-1
  - IL-6
  - IL-8
  - IL-12
  - IL-23

11.1.2 Analysis Plan

For CRP and each inflammatory marker separately the interaction between the inflammatory marker, time and group will be added to the primary model. The inflammatory marker will be entered to the model as a continuous variable, after descriptive analysis of the marker's distribution. Sensitivity analysis with transformation will be conducted if the marker if found to have distribution that is lightly skewed, or concentrated at zero. Log-transformation will be used for slewed to the right markers, after 0.1 * X, transformation for the marker where X is the minimum value of the marker that is larger than 0, if the marker has 0 (to avoid log(0)). In the case of concentration of values at zero, hurdle transformation will be applied. The transformation implies creating an indicator for zeros, and adding it to the model to estimate separate effects of zero and continuous non-zero values.

Moderation Model

Model with three-way interaction:

$$\begin{aligned} HRSD17_{ij} = \beta_0 \quad & +\beta_1 \text{Group}_i + \beta_2 \text{Time}_j + \beta_3 \text{Inflammation}_i \\ & +\beta_4 (\text{Group}_i \times \text{Time}_j) + \beta_5 (\text{Group}_i \times \text{Inflammation}_i) \\ & +\beta_6 (\text{Time}_j \times \text{Inflammation}_i) \\ & +\beta_7 (\text{Group}_i \times \text{Time}_j \times \text{Inflammation}_i) \\ & +\beta_8 \text{BaselineHRSD}_i + b_{0i} + \epsilon_{ij} \end{aligned}$$

Test: F-test for three-way interaction (β_7) will be the primary test of interest.

Significance level: α = 0.10 (exploratory) will be considered for definition of exploratory evidence and interpretation of the results. In this case the test is not expected to indicate strong evidence without independent replication.

Reporting

Estimate treatment effect (Minocycline vs. Placebo trajectories) will be reported using estimated marginal means and 95% confidence intervals with interaction plots. High, medium and low values for the marker is defined as mean + SD, mean and mean - SD, and will be used in the graphs. Besides predicted trajectories, treatment effect at 12 weeks for different levels of the inflammatory marker will be reported.

Diagnostic Analysis

The linearity of the inflammatory marker and time will be investigated visually using model residuals plotted against the value of the marker and time; a smoothing loess line will be fitted to it. Sensitivity analysis will be conducted using spline functions for the inflammatory marker and time if evidence of non-linearity is strong.

R code:

```
# Example: IL-6 as moderator

moderation_model <- lmer(
  fixed = HRSD17 ~ Group * Time * IL6 + BaselineHRSD +
  (1 | id),
  data = analysis_data %>% filter(Time > 0) )

# Test three-way interaction
anova(moderation_model)

# If significant, estimate simple effects
# using list of IL6 values (mean - sd, mean, mean + sd)
# possibly more value sif effect is non-linear.
emm <- emmeans( moderation_model,
    ~ Time | Group * IL6,
    at = list( IL6 = il6_values,
    Time = 1:12 ) )

# time slope (change in HRSD per week)
emtrends( moderation_model,
    ~ Group | IL6,
       var = "Time",
       at = list(IL6 = il6_value)
)
```

11.2 Exploratory Hypothesis 4: Change in Inflammatory markers will mediate the treatment effect on HRSD-17.

- **Baseline inflammatory markers:**
    - C-Reactive Protein (CRP, mg/L)
    - CCL-2
    - TNF-α
    - IL-1
    - IL-6
    - IL-8
    - IL-12
    - IL-23

Mediation models will be fitted in Mplus Version 9 with separated model for each marker.

11.2.1 Longitudinal mediation

IL6 is used as example, the same analysis will be done for all 4 inflammatory markers.

The mediator (IL6 change from baseline to Week 6) temporally precedes the outcome (depression change from baseline to Week 12), strengthening causal interpretability of the indirect pathway.

Measurement Model: Latent Change Scores

A latent change factor (ΔIL6_0–6) will be defined using IL6 at baseline and Week 6:

- IL6_6 regressed on IL6_0 (autoregressive parameter fixed to 1)
- Latent change factor defined to capture residual change from baseline to Week 6

A latent change factor (ΔHRSD_0–12) will be defined using HRSD at baseline and Week 12:

- HRSD_12 regressed on HRSD_0 (autoregressive parameter fixed to 1)
- Latent change factor defined to represent change from baseline to Week 12

Structural Model

The mediation model will include the following paths:

ΔIL6_0–6 regressed on:

- Treatment group
- Baseline HRSD
- Baseline covariates associated with missingness at SMD > 0.2

ΔHRSD_0–12 regressed on:

- ΔIL6_0–6
- Treatment group
- Baseline HRSD
- Baseline covariates associated with missingness at SMD > 0.2

The 0.2 threshold is chosen to avoid including all baseline variables in the model. The assumption is that the threshold 0.2 will capture all baseline variables with meaningful impact on MAR assumption.

Statistical significance of the indirect effect of ΔHRSD_0–12 on treatment group via ΔIL6_0–6 will (estimated as the product of coefficients) be assessed using Bias corrected bootstrap confidence intervals with 5000 re sampling replications and 90% and 95% confidence intervals.

The model as specified in Mplus below will account for missingness under MAR assumption and will include HRSD at weeks 2 and 6 in the model to make MAR assumption for dropouts more plausible. We chose to define change from baseline to week 12 as outcome of interest in order to capture the full treatment effect since the alternative, change from week 6 to 12, is expected by the less affected by the randomized treatment

groups. We will, however, conduct a secondary analysis exploring the mediation effect of the inflammatory marker on HRSD change from week 6 to 12.

MPLUS example code

```
TITLE: Longitudinal Mediation Model
       Latent Change IL6 (0-6) mediating Treatment effect on
       Latent Change HRSD (0-12)

DATA:
  FILE = analysis_wide_data.dat;

! X1 is a placeholder for baseline variables associated
! with missingness at SMD >= 0.2..
! Group is coded as (0 Placebo, 1 Minocycline)

VARIABLE:
  NAMES = ID GROUP IL6_0 IL6_6 HRSD_0 HRSD_2 HRSD_6 HRSD_12 X1;
  USEVARIABLES = GROUP IL6_0 IL6_6 HRSD_0 HRSD_12 X1;
  MISSING = ALL (-999);
  AUXILIARY = HRSD_2 HRSD_6; !to include time 2 and 6 in FIML.

ANALYSIS:
  ESTIMATOR = ML;
  BOOTSTRAP = 5000;

MODEL:

  !--------------------------------------
  ! Latent Change in IL6 (0 to 6 weeks)
  !--------------------------------------

  dIL6 BY IL6_6@1;
  IL6_6 ON IL6_0@1;
  IL6_6@0;


  !--------------------------------------
  ! Latent Change in HRSD (0 to 12 weeks)
  !--------------------------------------

  dHRSD BY HRSD_12@1;
  HRSD_12 ON HRSD_0@1;
  HRSD_12@0;
  dHRSD@0;

  ! IL6_0 WITH HRSD_0; correlation between exogenous variable
```

```
  ! are included by default in Mplus.

  !--------------------------------------
  ! Structural Paths (Mediation)
  !--------------------------------------

  ! a-path
  dIL6 ON GROUP (a1)
          X1
          X2;

  ! b-path and direct effect
  dHRSD ON dIL6 (b1)
           GROUP (cprime)
           HRSD_0
           X1
           X2;

MODEL INDIRECT:
  dHRSD IND GROUP;

OUTPUT:
  STDYX CINTERVAL(BCBOOTSTRAP);
```

11.2.2 Parallel Process Mediation

IL6 is used as example; identical models will be estimated for the other inflammatory markers.

The mediator (within-person linear slope of IL6 over time) is hypothesized to mediate the effect of treatment on the trajectory of depressive symptoms across time (within-person linear slope of HRSD over time).

This approach uses a **parallel process latent growth model**, allowing simultaneous estimation of inflammatory and symptom trajectories and their within subject association as part of the mediation mechanism. In this model we don't explicitly use the longitudinal lag to establish time precedence and instead time precedence is assumed to happen at the same time that by modelling the full trajectory and their association, this model can be seen as more mechanistic.

Measurement Model: Latent Growth Factors

- IL6 Growth Model
    - IL6 measurements at baseline, weeks 6 and 12
    - Latent IL6 intercept (initial level of the marker)
    - Latent IL6 slope (weekly linear change - assumption of linearity is made due to only 3 time points being available and for simpler interpretation of mediation effect)

- HRSD Growth Model
    - HRSD measurements at baseline, weeks 2, 6 and 12
    - Latent HRSD intercept (initial level of symptoms)
    - Latent HRSD slope (weekly linear change will be made for simpler interpretation of mediation effect)

Structural Model

The path model for the indirect mediation effect is represented by:

$$Treatment \rightarrow IL6\ slope \rightarrow HRSD\ slope$$

The IL6 and HRSD slope will be regressed on:

- Treatment group
- Baseline HRSD
- Baseline covariates associated with missingness (SMD > 0.2)
- With HRSD slope regressed on IL6 Slope

This specification estimates whether changes in inflammation are associated with changes in depressive symptoms, above and beyond treatment assignment.

Statistical significance for the indirect effect will be evaluated using **bias-corrected bootstrap confidence intervals with 5000 re sampling replications** (90% and 95% CIs).

Missing Data Handling

The model will be estimated using **maximum likelihood (ML)** under the assumption of missing at random (MAR). Baseline HRSD will be included as a predictor of HRSD slope.

Mplus Example Code

```
TITLE: Parallel Process Longitudinal Mediation Model
       IL6 Trajectory mediating Treatment effect on HRSD Trajectory

DATA:
  FILE = analysis_wide_data.dat;

  ! X1 is a placeholder for baseline variables ot be adjusted for.

VARIABLE:
  NAMES = ID GROUP IL6_0 IL6_6 HRSD_0 HRSD_2 HRSD_6 HRSD_12 X1;
  USEVARIABLES = GROUP IL6_0 IL6_6 HRSD_0 HRSD_6 HRSD_12 X1;
  MISSING = ALL (-999);

ANALYSIS:
  ESTIMATOR = ML;
```

```
  BOOTSTRAP = 5000;

MODEL:

  !----------------------------
  ! IL6 Growth Model
  !----------------------------

  iIL6 sIL6 | IL6_0@0 IL6_6@1;

  !----------------------------
  ! HRSD Growth Model
  !----------------------------

  iHRSD sHRSD | HRSD_0@0 HRSD_6@0.5 HRSD_12@1;

  !----------------------------
  ! Structural Mediation
  !----------------------------

  sIL6 ON GROUP (a1)
          X1
          iHRSD;

  sHRSD ON sIL6 (b1)
           GROUP (cprime)
           X1
           iIL6;

MODEL INDIRECT:
  sHRSD IND GROUP;

OUTPUT:
  STDYX CINTERVAL(BCBOOTSTRAP);
```

12. Data Monitoring and Interim Analyses

No formal interim analyses are planned for efficacy.

13. Documentation and Reporting

13.1 Analysis Outputs

All analyses will be documented with:

1. **R and Mplus scripts:** Annotated, version-controlled scripts for all analyses
2. **Analysis log:** Record of all analyses conducted, including date, analyst, and results
3. **Output:** All statistical results in annotated HTML format for analysis in R and possibly in Excel format for analysis in Mplus.

13.2 Reproducibility

To ensure reproducibility:

- Set random seed for all procedures involving randomness: `set.seed(20260102)`
- Document R version and package versions: `sessionInfo()`
- Archive analysis data set with date stamp
- Complete analysis code available in the report.

13.3 Deviations from SAP

Any deviations from this SAP must be:

- Documented with rationale
- Approved by principal investigator
- Noted in final analysis report
- Reported in publication methods