TITLE PAGE

SCT-MD-60

A Randomized, Multicenter, Double-Blind, Flexibly-dosed, Efficacy and Safety Study of Escitalopram in the Treatment of Children and Adolescents With Generalized Anxiety Disorder

STATISTICAL ANALYSIS PLAN

Version 4.0: 22 JUL 2021

1.0 OVERALL TABLE OF CONETENTS

1.1 LIST OF IN-TEXT TABLES

1.2 LIST OF IN-TEXT FIGURES

2.0 LIST OF ABBREVIATIONS

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
ATC4	anatomic therapeutic class 4
BMI	body mass index
CDC	Center for Disease Control
CGAS	Children's Global Assessment Scale
CGI-S	Clinical Global Impression - Severity
CRF	case report form
C-SSRS	Columbia-Suicide Severity Rating Scale
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ECG	electrocardiogram, electrocardiographic
eCRF	electronic case report form
ET	end of treatment
GAD	generalized anxiety disorder
LOCF	last observation carried forward
MCH	mean cell hemoglobin
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MMRM	mixed-effects model for repeated measures
MMRM-CAT	categorical mixed effects model for repeated measures
NEAE	newly emergent adverse event
OC	observed cases

PARS	pediatric anxiety rating scale
PCS	potentially clinically significant
PK	pharmacokinetic
PT	preferred term
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett formula
QTcF	QT interval corrected for heart rate using the Fridericia formula
SAE	serious adverse event
SAP	statistical analysis plan
SE	Standard Error
SI	Le Système International d'Unités (International System of Units)
SID	subject identification
SOC	System Organ Class
TBL	total bilirubin
TEAE	treatment-emergent adverse event
ULN	upper limit of normal laboratory reference range

3.0 INTRODUCTION

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of the efficacy and safety data as outlined and/or specified in the protocol of Study SCT-MD-60. Specifications of tables, figures, and data listings are contained in a separate document.

This is a randomized, multicenter, double-blind, flexibly-dosed, placebo-controlled parallel group study of escitalopram in the treatment of children (aged 7 through 11 years) and adolescents (aged 12 through 17 years) with generalized anxiety disorder (GAD).

The study design incorporates an up to 3 week screening period, an 8-week double-blind, acute treatment period, and a 1-week double-blind down taper period with a follow-up telephone contact 1 week after the end of the double-blind down taper period.

Approximately 430 subjects will be screened and a total of approximately 256 subjects are planned to be randomized in a 1:1 ratio to either escitalopram or placebo. The randomization schedule will be stratified by age at randomization (grouped as 7-11 years vs. 12–17 years) and sex, and blocked. The sample size may be increased based on the outcome of a blinded interim analysis but the number of subjects in each treatment arm will be capped at 160.

Subjects will begin with a 10 mg/day dose of escitalopram (taken orally) or matching placebo for the first 2 weeks of double-blind treatment. At the end of Week 2, subjects who can tolerate the current dose of 10 mg/day may have a dose escalation to 20 mg/day of escitalopram or matching placebo. Dose escalation will be at the investigator's discretion taking into account the Clinical Global Impression of Severity (CGI-S) score. Subjects who remain on the 10 mg/day dose of escitalopram or matching placebo will be evaluated again at Week 4 for a possible dose escalation to 20 mg/day at the investigator's discretion. Dose escalations cannot be made at any other time during the study. The dose may be decreased from 20 mg/day to 10 mg/day following discussion with the medical director if the subject does not tolerate the 20 mg/day dose. At the end of the treatment period or at early withdrawal, subjects will enter a 1-week double-blind down taper period.

For subjects receiving escitalopram 10 mg/day, the doses given during the double-blind down taper period will be as follows:

- Week 9: Subjects will receive placebo.

For subjects receiving escitalopram 20 mg/day, the doses given during the double-blind down taper period will be as follows:

- Week 9: Subjects will receive 10 mg/day.

Subjects randomized to the placebo arm will receive matching placebo during the double-blind acute treatment period. Subjects evaluated at the end of Week 2 and Week 4 and tolerating 10 mg/day matching placebo may be indicated for dose escalation by the investigator but in a blinded fashion will continue to receive matching placebo. Subjects receiving placebo will follow a similar down tapering schedule at the end of the double-blind acute treatment period.

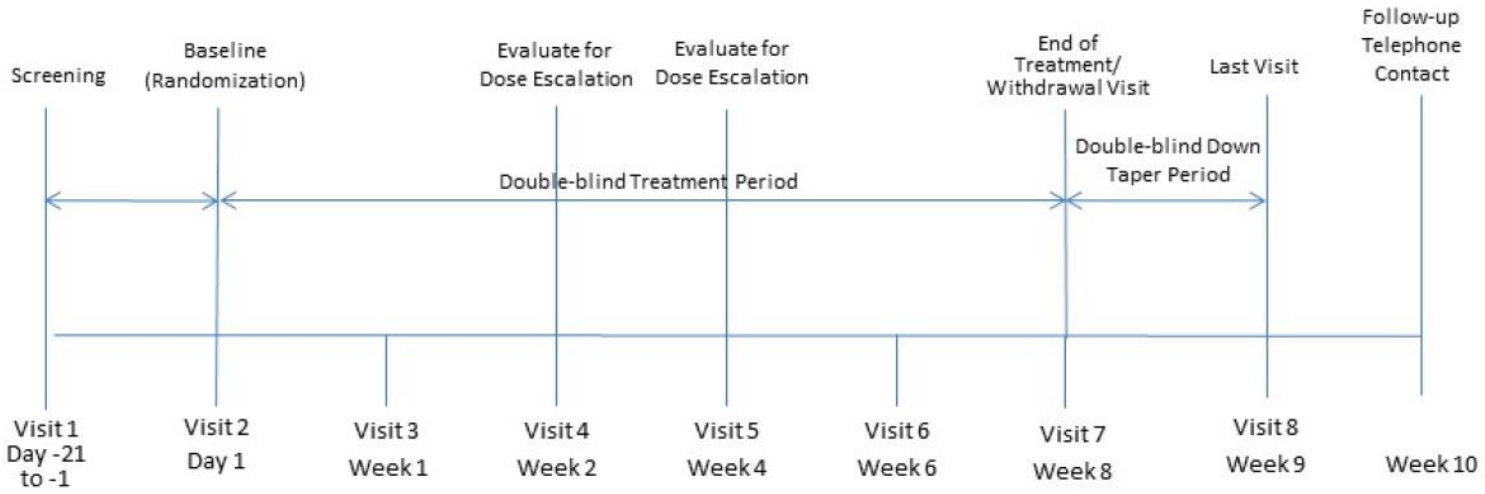
A follow-up telephone contact will take place 1 week after the end of the double-blind down taper period.

A blinded interim analysis will be conducted when 75% of randomized subjects have either completed or discontinued the study in order to check the assumptions in the sample size calculation. Based on this, the sample size may be increased to ensure an adequate power. However, due to the difficulties in recruiting pediatric subjects with GAD, the number of subjects in each treatment arm will be capped at 160.

The study design is shown in Figure 1.

3.1 STUDY DESIGN

Figure 1. Flow Chart of the Study



3.2 SCHEDULE OF ASSESSMENTS

Period	Screening Period	Double-blind Acute Treatment Period						Double-blind Down Taper Period	Follow-up
Visit	1	2	3	4	5	6	7/ET[a]	8[b]	Telephone Contact
Day/Week	Day -21 to -1	Baseline Day 1	Week 1[c]	Week 2[c]	Week 4[c]	Week 6[c]	Week 8/ET[c]	Week 9[c]	Week 10[c]
Subject Informed Consent/Assent	X								
Parent Informed Consent	X								
Inclusion/Exclusion Criteria	X	X							
Demographics	X								
Medical History	X								
Psychiatric History	X								
Medication History	X								
MINI Kid	X								
Physical Examination	X						X		
Body Mass Index (BMI)	X				X		X		
Randomization		X							
Vital Signs[d]	X	X	X	X	X	X	X	X	
Electrocardiogram (ECG)	X						X		
Urine Drug Screen	X						X		
Screen for HBsAg, HCVAb, HIV	X								
Clinical Chemistry including Thyroid	X[e]						X		
Hematology	X						X		
Urinalysis	X						X		
Pregnancy Test[f]	X	X					X		
C-SSRS	X	X	X	X	X	X	X	X	
CGAS	X	X	X	X	X	X	X		
PARS	X	X		X			X		
CGI-S	X	X		X			X		
Global COVID-19 Impact Assessment[g]	X	X	X	X	X	X	X		

Evaluate for Dose Escalation[h]				X	X				
PK Blood Samples[i]					X		X		
Dispense Study Drug		X[j]	X	X	X	X	X		
Study Drug Accountability			X	X	X	X	X	X	
Adverse Events		◄----------							----------►
Concomitant Medications	◄----------								----------►

Abbreviations: BMI = body mass index; CGAS = Children's Global Assessment Scale; CGI-S = Clinical Global Assessment of Severity; COVID-19 = coronavirus disease 2019; C-SSRS = Columbia-Suicide Severity Rating Scale; ECG = electrocardiogram; ET = End of Treatment; HBsAg = hepatitis B surface antigen; HCVAb = hepatitis C virus antibody; HIV = human immunodeficiency virus; MINI Kid = Mini-International Neuropsychiatric Interview for Children and Adolescents; PARS = Pediatric Anxiety Rating Scale; PK = pharmacokinetic.

[a] End of Treatment or Early Withdrawal.

[b] If a site becomes aware of any adverse events after study completion (occurring outside of a defined study visit, including any contact up to 30 days after study completion), all adverse events reported by the patient or patient representative or observed or otherwise identified by the investigator or other study personnel must be documented.

[c] Visit window ± 3 days. Refer to Section 8.2.2.1, Table 1 of the protocol for visit windows for safety assessments (physical examination, BMI, vital signs, ECG, urine drug screen, clinical chemistry including thyroid, hematology, urinalysis, pregnancy test) and/or PK assessments missed due to remote visits.

[d] Vital signs (blood pressure, pulse rate, temperature, and respiration rate) will be recorded at every visit. Height and weight will be recorded at Visit 1 (Screening), Visit 5 (Week 4), and Visit 7 (Week 8/ET) only.

[e] Fasting blood sample.

[f] A serum pregnancy test will be performed at Visit 1 (Screening); a urine pregnancy test will be performed at Visit 2 (Baseline) and Visit 7 (Week 8/ET) for females of childbearing potential.

[g] To be completed last after all other study assessments.

[h] Subjects who can tolerate the current dose of 10 mg/day may have a dose escalation to 20 mg/day at the investigator's discretion taking into account the CGI-S score.

[i] For PK blood sampling for assented/consented subjects at Visit 5 (Week 4) and Visit 7 (Week 8/ET), 2 blood samples per subject will be collected at each of these visits for escitalopram and S-demethylcitalopram analysis at the following times: Visit 5 (Week 4): for subjects receiving dose in the morning, at 0 h (predose) and ≥ 1 h postdose and for subjects receiving dose in the evening, 2 samples at random times during the visit (i.e., postdose from the dose received the previous evening), collected at least 2 hours apart from each other; and Visit 7 (Week 8/ET): 2 samples at random times during the visit (i.e., postdose from the dose received in the morning or the previous evening), collected at least 2 hours apart from each other.

[j] Subjects should take their first dose of study drug in clinic.

3.3 SCHEDULE OF ASSESSMENTS THAT CAN BE DONE REMOTELY

Period	Double-blind Acute Treatment Period						Double-blind Down Taper Period	Follow-up
Visit	2	3	4	5	6	7/ET[a]	8[b]	Telephone Contact
Day/Week	Baseline Day 1	Week 1[c]	Week 2[c]	Week 4[c]	Week 6[c]	Week 8/ET[c]	Week 9[c]	Week 10[c]
Inclusion/Exclusion Criteria	X							
Randomization	X							
C-SSRS	X	X	X	X	X	X	X	
CGAS	X	X	X	X	X	X		
PARS	X		X			X		
CGI-S	X		X			X		
Global COVID-19 Impact Assessment[d]	X	X	X	X	X	X	X	
Evaluate for Dose Escalation[e]			X	X				
Dispense Study Drug[f]	X[g]	X	X	X	X	X		
Study Drug Accountability		X	X	X	X	X	X	
Adverse Events	◄----------							----------►
Concomitant Medications								

Abbreviations: CGAS = Children's Global Assessment Scale; CGI-S = Clinical Global Assessment of Severity; COVID-19 = coronavirus disease 2019; C-SSRS = Columbia-Suicide Severity Rating Scale; ET = End of Treatment; PARS = Pediatric Anxiety Rating Scale; PK = pharmacokinetic.

Notes for Remote Study Visits:

• The Baseline Visit can be remote but it is not preferred. For all female subjects of childbearing potential, the Baseline Visit must take place as an in-clinic visit with all assessments outlined in Section 8.1.2 of the protocol.

• Subject data are collected via telephone or telehealth video calls using a technology provider. Sites should remain consistent with either method (video or telephone) for each subject.

• In-clinic safety and PK assessments missed due to remote visits may be collected per Section 8.2.2.1, Table 1 of the protocol.

[a] End of Treatment or Early Withdrawal.

[b] If a site becomes aware of any adverse events after study completion (occurring outside of a defined study visit, including any contact up to 30 days after study completion), all adverse events reported by the patient or patient representative or observed or otherwise identified by the investigator or other study personnel must be documented.

[c] Visit window ± 3 days.

[d] To be completed last after all other study assessments.

[e] Subjects who can tolerate the current dose of 10 mg/day may have a dose escalation to 20 mg/day at the investigator's discretion taking into account the CGI-S score.

[f] For remote visits study drug will be sent to subjects via curbside (preferred) or courier.

[g] In case of a remote Baseline Visit, subjects will receive and take their first dose of study drug via curbside (preferred) or they will receive it by courier and take it at home.

4.0 STUDY OBJECTIVES

4.1 PRIMARY OBJECTIVE

The primary objective of this study is to assess the safety and efficacy of escitalopram relative to placebo in the acute treatment of children (aged 7 through 11 years) and adolescents (aged 12 through 17 years) who meet criteria for GAD as defined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).

4.2 SECONDARY OBJECTIVE

The secondary objective of this study is to characterize the pharmacokinetic (PK) profile of escitalopram in the pediatric population (aged 7 through 17 years).

4.3 EXPLORATORY OBJECTIVE

The exploratory objective of this study is to assess the exposure/dose-response relationship.

5.0 SUBJECT POPULATIONS

Five populations will be considered in the statistical analysis of the study.

5.1 SCREENED POPULATION

The Screened Population will consist of all subjects who underwent a Screening Visit and received a subject number, and signed the informed consent form.

5.2 RANDOMIZED POPULATION

The Randomized Population will consist of all subjects in the Screened Population who were randomized to a treatment group in the study. Randomized treatment will be utilized for the summaries of Randomized Population by treatment.

5.3 SAFETY POPULATION

The Safety Population will consist of all subjects in the Randomized Population who took at least one dose of study medication. Actual treatment received based on the treatment received the most during the double-blind acute treatment period (>50% of doses) will be utilized for the summaries of Safety Population by treatment. If placebo and escitalopram are received an equal number of times, escitalopram will be considered the actual treatment received.

5.4 MODIFIED INTENT-TO-TREAT (MITT) POPULATION

The mITT Population is defined as all subjects who were randomized and received at least one dose of study medication, and have both baseline and at least one post-baseline primary efficacy measure [ie, Pediatric Anxiety Rating Scale (PARS) severity score]. The mITT population will be used for all of the efficacy analyses. Randomized treatment will be utilized for the summaries of mITT Population by treatment.

5.5 PHARMACOKINETIC POPULATION

The Pharmacokinetic Population is defined as all subjects in the Safety Population with at least one evaluable plasma concentration of escitalopram or S-demethylcitalopram.

6.0 SUBJECT DISPOSITION

The number of subjects in 3 of the study populations (Randomized, Safety, and mITT) will be summarized by treatment group and study center; the Screened Population will be summarized overall only by study center.

Screen-failure subjects (i.e., subjects screened but not randomized) and the associated reasons for failure to randomize will be tabulated overall for the Screened Population. Subjects who are randomized but not in the Safety Population will be listed.

Subjects completing the Week 8 assessment of double-blind acute treatment period will be considered completers. The number and percentage of subjects who complete the double-blind acute treatment period and of subjects who prematurely discontinue during the same period will be presented for each treatment group and overall for the Safety Population.

The reasons for premature discontinuation from the double-blind acute treatment period as recorded on the disposition pages of the electronic case report forms will be summarized (number and percentage) by treatment group for the Safety Population. All subjects who prematurely discontinue during the double-blind acute treatment period will be listed by discontinuation reason for the Safety Population.

The number and percentage of subjects impacted by the COVID-19 pandemic will be summarized by treatment group and overall for the Screened Population. The summary will also include the extent of the impact, including:

• discontinuations related to COVID-19 during screening, double-blind acute treatment period, and down-taper period

• visits missed due to COVID-19

• video/phone visits conducted in place of on-site visits due to COVID-19

• assessments missed due to COVID-19 (this will only include assessments missed that were not permitted to be missed during off-site visits)

• significant protocol deviations related to COVID-19

• study drug disruption due to COVID-19

• TEAEs of coronavirus infection or positive coronavirus test

The number and percentage of subjects with missing or remote visits and with missing assessments due to COVID-19 will be summarized by treatment group and overall for each visit:

The number and percentage of subjects with significant protocol deviations will be summarized overall and by treatment group for the Randomized Population. Deviations related to the following categories will be included:

• inclusion or exclusion criteria

• withdrawal criteria

• treatment or dose

• concomitant medications

These and any additional significant protocol deviations will be reviewed and documented before database lock and unblinding of treatment codes.

Significant protocol deviations that occur due to COVID-19 will be summarized separately. All protocol deviations that occur due to COVID-19 will be listed.

7.0 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

7.1 DEMOGRAPHICS

Demographic parameters (age, age group [7-11 vs. 12-17 years], sex, race, ethnicity, weight, height, BMI, BMI z-score group [≥ 2, < 2]) and other baseline characteristics will be summarized descriptively by treatment group and overall for the Safety and mITT populations, respectively.

Continuous variables will be summarized by number of subjects and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of subjects.

7.2 MEDICAL HISTORY

Medical history will be classified by system organ class and preferred term using the Medical Dictionary for Regulatory Activities (MedDRA), version 22.0 or newer. The number and percentage of subjects in each system organ class (SOC) and preferred term (PT) will be summarized by treatment group for the Safety Population.

Psychiatric history and result from MINI Kid will also be summarized descriptively by treatment group for the Safety Population.

Listings of medical history, psychiatric history and results from MINI Kid will be presented

7.3 CONCOMITANT MEDICATION

The World Health Organization Drug Dictionary, version Global B3, Mar 2019 or newer, will be used to classify prior and concomitant medications by therapeutic class and drug name.

Prior medication is defined as any recorded medication taken before the date of the first dose of double-blind investigational product. Concomitant medication is defined as any medication taken on or after the date of the first dose of double-blind investigational product through the last study visit, including the follow-up call. The imputation of incomplete and missing start/stop date of medication are specified in the section 17.

Both prior and concomitant medications use will be coded by PT and Anatomic Therapeutic Class 4 (ATC4). The use of prior and concomitant medications will be summarized by the number and percentage of subjects in each treatment group and overall for the Safety Population. Summaries will be ordered in alphabetical order of ATC4 and then, within an ATC4, in decreasing frequencies by PT in total column. If a

subject took a specific medication multiple times or took multiple medications within a specific therapeutic class, that subject would be counted only once for the coded drug name or therapeutic class.

Summaries for concomitant medication use will be presented for the double-blind acute treatment period and the down-taper period, separately. Any concomitant medications which started after the date of the last dose of double-blind investigational product in the study will not be presented in the summary tables but will be included in the subject data listings.

8.0 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

8.1 EXTENT OF EXPOSURE

Total exposure to double-blind investigational product for the Safety Population during the double-blind acute treatment period will be summarized, calculated as the number of days from the date of the first dose of double-blind investigational product taken to the date of the last dose taken through the Week 8 End of Treatment or Early Withdrawal visit, inclusive. Descriptive statistics (n, mean, standard deviation, minimum, median, and maximum) will be presented by treatment group. Similarly, the treatment duration in the down taper period, defined as the first dose in the down taper period to the last dose in the down taper period, inclusive, will be summarized. The total duration of double-blind investigational product exposure, defined as the sum of the treatment duration through Week 8 and the duration in the down taper period will be summarized.

The number and percentage of subjects who had their dosage escalated at Week 2 from 10 mg/day to 20 mg/day will be summarized by treatment group. The denominator will be the number of subjects who had study medication dispensed at Week 2. Similarly, the number and percentage of subjects who had their dosage escalated at Week 4 from 10 mg/day to 20 mg/day will be summarized by treatment group. The number of subjects who required a dose reduction from 20 mg/day to 10 mg/day will be summarized by treatment group. The number of days from first date of escalation to 20 mg/day to the dose reduction will be summarized using descriptive statistics.

The mean daily dose of investigational product between visits and overall during double-blinded acute treatment period will be summarized by treatment group for the Safety Population. The calculation is as below:

Mean daily dose = (Total dose taken during the interval) / (Total number of days in the interval).

Total number of days = Stop date of study drug – start date of study drug + 1.

The total dose, start date, and stop date will be based on the eCRF Study Drug Administration page.

8.2 MEASUREMENT OF TREATMENT COMPLIANCE

Dosing compliance for a specified period is defined as the total number of capsules actually taken by a subject during that period divided by the number of capsules prescribed to be taken for the same period multiplied by 100.The total number of capsules actually taken during a specific time period is calculated based on the study medication record. The number of capsules prescribed to be taken for a specific treatment period will be calculated by multiplying the number of days in that period by the number of capsules to be taken per day.

Descriptive statistics for investigational product compliance will be presented by treatment group for each period between 2 consecutive visits, as well as for the entire double-blind acute treatment period, the period from first dose to the Week 8 End of Treatment or Early Withdrawal visit, and the down taper period.

9.0 EFFICACY ANALYSES

The efficacy analyses will be based on the mITT Population. Baseline for efficacy is defined as the last non-missing efficacy assessment before the first dose of double-blind investigational product. All statistical tests will be 2-sided hypothesis tests performed at the 5% level of significance for main effects. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise.

For efficacy analyses in which study center is a factor, a small center will be defined as a center with fewer than 2 subjects in any treatment group in the mITT Population. Small centers will be pooled to form pseudo-centers so that each treatment group includes at least 2 mITT subjects within the center. Pooling will be done (within each country first) using the following algorithm:

Based on the number of mITT subjects, small centers will be ordered from the largest to the smallest, and centers of the same size will be ordered from the largest center code to the smallest center code. The pooling process starts with the largest small center from the top, which will be pooled with the smallest from the bottom until a non-small pseudo center is formed. The process will be repeated using the small centers left out after the first pass. If any centers are left out at the end of the process, they will be pooled with the smallest pseudo center. If there is more than 1 smallest pseudo center, the pseudo center with the smallest center code will be selected. In case that the pseudo center formed by pooling all small centers is still a small center, it will be pooled with the smallest non-small center. If there is more than 1 smallest non-small center, the one with the smallest center code will be selected.

These pseudo-centers will be used for all efficacy analyses when the model is adjusted for study center.

By-visit analyses based on the mixed-effects model for repeated measures (MMRM) using the observed case (OC) approach will be performed for all continuous efficacy parameters with multiple post-baseline measurements.

9.1 PRIMARY EFFICACY PARAMETER

9.1.1 Primary Objective Estimand

The estimand used to address the primary objective “To determine efficacy of escitalopram relative to placebo in the acute treatment of children (aged 7 through 11 years) and adolescents (aged 12 through 17 years) who meet criteria for GAD as defined in the Diagnostic and Statistical Manual of Mental Disorders” is defined by the following:

- Population:

- The target population is children (aged 7 through 11 years) and adolescents (aged 12 through 17 years) with GAD and who satisfy the inclusion and exclusion criteria as specified in the protocol.
- The analysis population is mITT population defined in section 5.4.

- Variable:
  - Change from baseline to Week 8 in PARS severity. The PARS severity score is derived by summing 5 of 7 severity/impairment/interference items (2, 3, 5, 6, and 7).

- Intercurrent events:
  - To evaluate the efficacy at Week 8 in the mITT population, participants are assumed to adhere to the assigned treatment for the duration of the study. As a result, data after the discontinuation from the study treatment due to all reasons will not be included in the primary analysis and they will be assumed as missing at random.
  - For the sensitivity analyses of primary parameter, the LOCF approach and pattern-mixture model approach will be utilized to impute the missing value. The details are provided in Section 10.1.2.

- Population-level summary:
  - The change from baseline of PARS severity score at Week 8 in subjects treated with escitalopram compared to the change from baseline of PARS severity score at Week 8 in subjects treated with placebo

9.1.2 Analysis of Primary Parameter

The primary efficacy parameter is the change from baseline to Week 8 in PARS severity. The PARS severity score for GAD is assessed for all symptoms identified in the generalized anxiety section of the PARS symptom checklist. The PARS severity score for GAD will be derived by summing 5 of 7 severity/impairment/interference items (2, 3, 5, 6, and 7). If any of the items is missing, the severity score will be considered missing.

The primary analysis will be performed using an MMRM with treatment group, age group strata (7-11 vs. 12-17 years), sex, pooled study center, visit, and treatment group-by-visit interaction as the fixed effects and the baseline value and baseline value-by-visit interaction as the covariates. An unstructured covariance matrix will be used to model the covariance of within-subject scores. The Kenward-Roger approximation (Kenward and Roger, 1997) will be used to estimate denominator degrees of freedom. If the model convergence for the unstructured covariance matrix fails, then the Fisher scoring

algorithm will be used to provide better initial values of the covariance parameters. In the event that the model still does not converge after using those initial values, simplified covariance structures will be used to fit the model in the following order until the model converges: (1) ante-dependence, (2) heterogeneous autoregressive, (3) Toeplitz, and (4) compound symmetry. If a structured covariance is used, then a robust sandwich estimator will be used for estimating the variance of the treatment effect estimate.

This analysis will only use the observed cases of post-baseline scores without imputation of missing values. The significance test for the difference in means at Week 8 will be based on least squares (LS) means using a 2-sided α of 0.05. Differences between treatments at the other post-baseline visits will also be estimated based on LS means.

The primary analysis will treat measurements after study withdrawal as missing in the MMRM modeling, leading to an estimate of the treatment effect while on treatment. Two sensitivity analyses, last observation carried forward (LOCF) and pattern-mixture model, will be performed on the primary efficacy parameter to assess the impact of the missing at random assumption of MMRM.

For the LOCF approach, only the postbaseline PARS severity score for GAD will be imputed; individual item scores will not be carried forward to derive the PARS severity score. Baseline total score will be carried forward only for the missing scores immediately after baseline. If all the postbaseline values are missing, the baseline value will not be carried forward. The LOCF approach is based on an analysis-of-covariance (ANCOVA) model including treatment group and pooled study center as factors and baseline PARS severity score as a covariate

Another sensitivity analysis using a pattern-mixture model approach based on non-future dependent missing value restrictions (Kenward et al., 2003) will be performed to assess the robustness of the primary MMRM results to the possible violation of the missing-at-random assumption. The details of this sensitivity analysis are as follows:

The pattern for the pattern-mixture model will be defined by the subject's last visit with an observed value. The observed PARS severity score at a visit is assumed to have a linear relationship with the subject's prior measurements. The missing values will be imputed under the assumption that the distribution of the missing observations differs from that of the observed only by a shift parameter value Δ. The dataset with missing values imputed will be analyzed using an analysis of covariance (ANCOVA) model with treatment group and pooled study center as factors and baseline PARS severity score as a covariate for between–treatment group comparisons at Week 8. The imputation of missing values and the analysis will be performed multiple times and the inference of this sensitivity analysis will be based on the combined estimates using the standard multiple imputation technique. The shift parameter values Δ for the multiple imputation will start at 0 and increase by 1 to 6 or until a tipping point is reached (ie, the p-value switches from <= 0.05 to > 0.05), whichever is higher.

More details of the proposed pattern-mixture model approach (e.g., the models for the pattern-specific identifiable densities and the unidentified conditional distributions, the shift parameter Δ, and the multiple imputation algorithm) are provided in Appendix I to this SAP.

To examine the consistency of the primary efficacy results across study centers, the treatment differences versus placebo treatment groups (mean ± SE) in the change from baseline to Week 8 in PARS severity score, both OC and LOCF, will be graphed by pooled study center.

The comparison of escitalopram vs. placebo will be performed to assess study assay sensitivity.

To assess the impact of COVID-19 on the primary efficacy parameter:

- The primary analysis will be repeated for subjects that were not significantly impacted by COVID-19 (Global COVID-19 Impact Assessment score of 4 or higher, regardless of direction of impact; see section 9.2.5).

- A sensitivity analysis will be conducted by treating remote visits as missing, and analyzing the remaining data using the primary MMRM model.

- A sensitivity analysis will be conducted by repeating the primary analysis and adding the following to the model: an indicator of whether data was collected before or on and after the start of the pandemic (17 March 2020) and the indicator by treatment interaction.

- A subgroup analysis will be conducted by repeating the primary analysis for randomized patients before or on and after the start of the COVID-19 pandemic (17 March 2020).

The subject-level PARS scores will be listed

9.2 SECONDARY EFFICACY PARAMETERS

The secondary efficacy endpoints are as follows:

- Response rate on the PARS at acute treatment endpoint (Week 8)
- Remission rate on the PARS at acute treatment endpoint (Week 8)
- Change on the CGI-S from baseline to acute treatment endpoint (Week 8)
- Remission rate on CGI-S at acute treatment endpoint (Week 8)
- Change on the CGAS from baseline to acute treatment endpoint (Week 8)
- Remission rate on the CGAS at acute treatment endpoint (Week 8)
- COVID-19 impact assessment score at each visit.

Inferential testing for the secondary endpoints will be reported with nominal p-values. No adjustments for multiplicity will be made. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise

9.2.1 Analysis of Response Rate on PARS

The response rate on the PARS at acute treatment endpoint (Week 8) is defined as 50% reduction on PARS severity score for GAD (using 5 PARS items: 2, 3, 5, 6, and 7) from baseline. The response rate of the PARS will be analyzed using a repeated measures model for categorical outcomes. Modeling will be completed using a categorical MMRM (MMRM-CAT) approach, in which a pseudo-likelihood-based repeated measures analysis will be used and implemented in SAS procedure GLIMMIX. A binary distribution with logit link will be used for the dependent variable. Classification variables for treatment, age group, sex, visit (as a categorical variable), treatment-by-age group interaction, treatment-by-sex interaction, treatment by visit interaction, treatment-by-age group-by-sex interaction, and baseline score as a covariate will be considered in the modeling. If the GLMM does not converge, a logistic regression model with treatment group and baseline score as explanatory variables will be used. For the logistic regression analysis, postbaseline missing data will be imputed using the LOCF approach.

9.2.2 Analysis of Remission Rate on the PARS

Remission rate on the PARS at acute treatment endpoint (Week 8) is defined as PARS severity score for GAD <8 (using 6 PARS items: 2, 3, 4, 5, 6, and 7). The analysis of remission rate will be conducted using the same methodology as the response rate on the PARS.

9.2.3 Summary of PARS Severity Score at each Visit

Observed values and change from baseline in the PARS severity score will be summarized by treatment at each visit.

9.2.4 Analysis of Change on the CGI-S from Baseline and Remission Rate of CGI-S

The CGI-S is a 7-point scale developed for use in clinical trials to provide an assessment of the clinician's view of the patient's global functioning. The change from baseline in CGI-S will be analyzed using the same methodology described for the primary efficacy parameter. The subject-level CGI-S score will be listed.

The remission rate of CGI-S at acute treatment endpoint (Week 8) is defined as percentage of subjects having a CGI-S score ≤2 at endpoint. The remission rate will be analyzed using the same methodology described for the analysis of the response rate on the PARS.

9.2.5 Summary of CGI-S Score at each Visit

Observed values and change from baseline in the CGI-S score will be summarized by treatment at each visit.

9.2.6 Analysis of Change on the CGAS from Baseline and Remission Rate of CGAS

The CGAS is a 100-point scale range from 1 to 100 used as a measure of global functioning and impairment. The change in the CGAS will be analyzed using the same methodology described for the primary efficacy parameter.

Remission rate on the CGAS at acute treatment endpoint (Week 8) is defined as CGAS >70. The analysis of the remission rate for the CGAS will use the same methodology described for the analysis of the response rate on the PARS.

9.2.7 Summary of CGAS Score at each Visit

Observed values and change from baseline in the CGAS score will be summarized by treatment at each visit.

9.2.8 Analysis of COVID-19 impact assessment score

The global COVID-19 impact assessment will be used to evaluate the impact of the coronavirus disease pandemic on the severity of the subject's GAD during the last 1 week. The rating will include scores from 0 (not applicable, subject not enrolled in the trial during the pandemic) to 7 (extreme impact). Raters will also evaluate whether the impact was an improvement or worsening of the subject's GAD.

Subjects with scores from 4 (moderate impact) to 7 (extreme impact) at any visit other than the Screening Visit will be considered to have been significantly impacted, regardless of the direction of impact.

Number and percentage of subjects with each score on the global COVID-19 impact assessment will be summarized descriptively at each visit by treatment group for the mITT population, including subjects that were significantly impacted. For each score from 2 to 7, the number and percentage of subjects who improved or worsened will also be summarized.

SAP SCT-MD-60
Allergan

10.0 SAFETY ANALYSES

The safety analysis will be performed for the double-blind acute treatment period and the double-blind down-taper period separately, using the Safety Population unless stated otherwise.

The safety parameters include adverse events (AEs), clinical laboratory parameters, vital signs, electrocardiographic (ECG) parameters, and suicide risk assessment using the Columbia-Suicide Severity Rating Scale (C-SSRS). For each safety parameter, the last non-missing safety assessment before the first dose of double-blind investigational product will be used as baseline. Continuous variables will be summarized by number of subjects and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of subjects.

Safety data collected outside of planned visit windows but within allowance limits (up to 1 week late for Visits 7 and 8) will be included in the summary tables. Safety data collected outside of planned visit windows and outside of the allowance limits (>1 week to 6 weeks late for Visits 7 and 8) will be listed, but will not be included in the summary tables. Subgroup analyses may be performed to assess the impact of COVID-19 on the safety analyses; details are listed in each safety section.

10.1 ADVERSE EVENTS

Adverse events will be coded by SOC and PT using the MedDRA version 22.0 or newer.

An AE (classified by PT) that occurs during the double-blind acute treatment period, during the down-taper period, or within 30 days after the date of the last dose of double-blind investigational product will be considered a treatment-emergent adverse event (TEAE) if it was not present before the date of the first dose of double-blind investigational product, or was present before the date of the first dose of double-blind investigational product and increased in severity during the double-blind acute treatment period or during the down-taper period, respectively.

If more than 1 AE was reported before the date of the first dose of double-blind investigational product and coded to the same preferred term, the AE with the greatest severity will be used as the benchmark for comparison with the AEs occurring during the double-blind acute treatment period or during the down-taper period that were also coded to that preferred term. An AE that occurs more than 30 days after the date of the last dose of double-blind investigational product in the study will not be counted as a TEAE but will be included in the listings.

An AE occurring during the down-taper period or safety follow-up period will be considered a newly emergent AE (NEAE) if it is not present before the start of the down-taper period or was present before the start of the down-taper period but increased in severity during the down-taper period. All NEAEs will be summarized by SOC, PT, and treatment group.

The number and percentage of subjects in each treatment group reporting TEAEs during the double-blind acute treatment period will be tabulated by SOC and PT; by SOC, PT, and severity; and by SOC, PT, and causal relationship to the investigational product. If more than one AE is coded to the same PT for the same subject, the subject will be counted only once for that PT using the most severe and most related occurrence for the summarization by severity and by causal relationship to the investigational product, respectively. Study-treatment-related TEAEs are defined as TEAEs with relationship to treatment categorized as Possible, Probable, Very, Like/Certain or missing.

The distribution of TEAEs by severity and causal relationship to the investigational product will be summarized by treatment group. The incidence of common (≥ 2% of subjects in any treatment group) TEAEs during the double-blind acute treatment period will be summarized by PT and treatment group.

A serious adverse event (SAE) that occurred between the date of the first dose of double-blind investigational product and 30 days after the date of the last dose of double-blind investigational product in the study, inclusive, will be considered an on-therapy SAE. The incidence of SAEs and AEs leading to premature discontinuation of the study will also be summarized by study period (separately for double-blind acute treatment period and double-blind down taper period), SOC, PT, and treatment group.

The above summaries will be sorted by decreasing frequency of SOC and by the decreasing frequency of PT within SOC for the escitalopram treatment group. If incidence for more than 1 term is identical, they will then be sorted alphabetically. In addition, the above summaries will be presented as all AEs, AEs that are associated with the COVID-19 pandemic (as determined by the investigator), and AEs that are not associated with the COVID-19 pandemic. AEs associated with the COVID-19 pandemic will be listed separately.

Listings will be presented for subjects with SAEs, subjects with AEs leading to premature discontinuation, and subjects who died for the Screened Population. All subjects with SAEs, including those reported during the screening period or more than 30 days after the date of the last dose of the double-blind investigational product, and subjects discontinuing due to AEs before the start of double-blind investigational product will be included in these listings

10.2 CLINICAL LABORATORY PARAMETERS

Descriptive statistics for clinical laboratory values (in SI units) and changes from the baseline values at each assessment time point (including the end of the double-blind acute treatment period) will be presented by treatment group for the following laboratory parameters:

Hematology: Hemoglobin, hematocrit, white blood cell count (total and differential), red blood cell count, platelet count, mean cell volume, mean cell hemoglobin (MCH), and MCH concentration.

Chemistry: Creatinine, urea (or blood urea nitrogen), aspartate transferase, alanine transferase, gamma glutamyltransferase, alkaline phosphatase, lactate dehydrogenase, total bilirubin, albumin, total protein, sodium, potassium, chloride, glucose, uric acid, calcium, phosphorus, thyroid-stimulating hormone, T3, and T4.

Urinalysis: pH, glucose, ketones, blood, protein, and microscopy.

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criteria listed in Table 10.2–1. The number and percentage of subjects who have PCS post-baseline clinical laboratory values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values and at least 1 postbaseline assessment. The numerator will be the total number of subjects with available non-PCS baseline values and at least 1 PCS post-baseline value. A supportive listing of subjects with PCS post-baseline values will be provided, including the subject identification (SID) number, study center number, and baseline and all post-baseline (including non-PCS) values.

In addition, a listing of all AEs that occurred in subjects who had PCS post-baseline clinical laboratory values will be provided.

Table 10.2–1. Criteria for Potentially Clinically Significant Laboratory Values

Parameter	SI Unit	Lower Limit	Higher Limit
HEMATOLOGY			
Basophils	%	—	> 6
Eosinophils	%	—	> 10
Hematocrit	1.0	< 0.9 × LLN	—
Hemoglobin	g/L	< 0.9 × LLN	—
Lymphocytes	%	< 10	> 60
Monocytes	%	—	> 20
Neutrophils	%	< 30	> 90
Platelet count	× 10^9/L	≤ 75	≥ 700

White blood cell count	× 10^9/L	≤ 2.5	≥ 15
CHEMISTRY			
Albumin	g/L	< 0.9 × LLN	> 1.1 × ULN
Alkaline phosphatase	U/L	—	≥ 3 × ULN
Alanine aminotransferase	U/L	—	≥ 3 × ULN
Aspartate aminotransferase	U/L	—	≥ 3 × ULN
Bicarbonate	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Blood urea nitrogen	mmol/L	—	> 1.2 × ULN
Calcium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Chloride	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Cholesterol, total	mmol/L	—	> 200mg/dL
Creatinine	µmol/L	—	> 1.3 × ULN
Glucose	mmol/L	< 0.8 × LLN	> 1.4 × ULN
high-density lipoprotein	mmol/L	< 40 mg/dL	—
Low-density lipoprotein	mmol/L	—	> 130 mg/dL
Magnesium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Phosphate	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Potassium	mmol/L	< 1.0 × LLN	> 1.0 × ULN
Sodium	mmol/L	< 1.0 × LLN	> 1.0 × ULN
Total bilirubin	µmol/L	—	> 1.5 × ULN
Total protein	g/L	< 0.9 × LLN	> 1.1 × ULN
Triglyceride	mmol/L	—	> 100 mg/dL (ages 7-10), >130 mg/dL (ages 11-17)
URINALYSIS			
Protein	g/L	—	at least two +
Glucose	mmol/L	—	at least two +

LLN = lower limit of normal value provided by the laboratory; SI = Le Système International d'Unités (International System of Units); ULN = upper limit of normal value provided by the laboratory.

Shift tables from baseline to the end of double-blind acute treatment period for clinical laboratory parameters will be presented by treatment group for the following categories: low, normal, and high.

Potential Hy's Law criteria is defined by a postbaseline elevation of alanine aminotransferase (ALT) or aspartate aminotransferase (AST) ≥ 3× ULN (ULN = upper limit of normal laboratory reference range), along with total bilirubin (TBL) ≥ 2× ULN and a non-elevated alkaline phosphatase (ALP) < 2× ULN, all based on blood draws collected within a 24-hour period. The number and percentage of subjects who meet the following criteria of clinical interest during the double-blind acute treatment period will be tabulated by treatment group. A supportive list will also be provided.

- ALT: ≥ 3× ULN, ≥ 5× ULN, ≥ 10× ULN, ≥ 20× ULN

- AST: ≥ 3× ULN, ≥ 5× ULN, ≥ 10× ULN, ≥ 20× ULN
- ALT or AST: ≥ 3× ULN, ≥ 5× ULN, ≥ 10× ULN, ≥ 20× ULN
- TBL: ≥ 1.5 x ULN, 2 × ULN
- ALP: < 2 × ULN
- Concurrent elevations:
  - ALT or AST ≥ 3× ULN with TBL ≥ 1.5× ULN;
  - ALT or AST ≥ 3× ULN with TBL ≥ 2× ULN;
  - ALT or AST ≥ 3× ULN with TBL ≥ 2× ULN and ALP < 2× ULN.

10.3 VITAL SIGNS

Descriptive statistics for vital signs (sitting systolic and diastolic blood pressure, pulse rate, respiration rate, temperature, height, weight, BMI, height percentile, weight percentile, and BMI percentile) and their changes from baseline values at each visit and at the end of study will be presented by treatment group. Height, weight and BMI z-scores will be calculated based on Centers for Disease Control (CDC) growth charts. Only subjects with available baseline and at least 1 postbaseline assessment will be included in the summary.

Vital sign values will be considered PCS if they meet both the observed-value criteria and the change-from-baseline criteria listed in Table 10.3–1. The number and percentage of subjects with PCS postbaseline values will be tabulated by treatment group separately. The percentages will be calculated relative to the number of subjects with available baseline and at least 1 postbaseline assessment. The numerator will be the total number of subjects with available baseline values and at least 1 PCS postbaseline value. A supportive listing of subjects with PCS postbaseline values will be provided, including the SID number, study center number, and baseline and all postbaseline (including non-PCS) values.

In addition, a listing of all AEs that occurred in subjects who had PCS postbaseline vital sign values will be provided.

Table 10.3–1. Criteria for Potentially Clinically Significant Vital Signs

Parameter	***Flag***	***Criteria[a]***		
		Observed Value		***Change From Baseline***
		Age 7-11	***Age 12-17***	
	High	≥ 125	≥ 140	Increase of ≥ 20

Sitting systolic blood pressure, mm Hg	Low	≤ 85	≤ 90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥ 90	≥ 95	Increase of ≥ 15
	Low	≤ 40	≤ 50	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 130	≥ 110	Increase of ≥ 15
	Low	≤ 55	≤ 45	Decrease of ≥ 15
Weight, kg	High	—	—	> 2 deviations above the age-gender mean weight
	Low	—	—	> 2 deviations below the age-gender mean weight

a A post-baseline value is considered potentially clinically significant if it meets both the observed-value and the change-from-baseline criteria.

bpm = beats per minute kg = kilogram, mmHg = millimeters of mercury hemoglobin.

10.4 ELECTROCARDIOGRAM

Descriptive statistics for ECG parameters (heart rate, PR interval, QRS interval, RR interval, QT interval, and QTc interval) at baseline and changes from baseline values at each assessment time point will be presented by treatment group. The QTc interval is calculated using both the Bazett (QTcB = QT/(RR)½) and Fridericia (QTcF = QT/(RR)⅓) corrections. Only subjects with available baseline and at least 1 postbaseline assessment will be included in the summary.

Electrocardiographic parameter values are considered PCS if they meet or exceed the higher-limit PCS criteria listed in Table 10.4–1. The number and percentage of subjects with PCS postbaseline ECG values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with available non-PCS baseline values and at least 1 postbaseline assessment. The numerator is the total number of subjects with available non-PCS baseline values and at least 1 PCS postbaseline value. A supportive listing of subjects with PCS postbaseline values will be provided, including the SID number, study center number, baseline, and all postbaseline (including non-PCS) values.

Table 10.4–1. Criteria for Potentially Clinically Significant Electrocardiographic Parameters

Parameter	***Unit***	***Higher Limit***
QRS interval	msec	≥ 150
PR interval	msec	≥ 250
QTcB	msec	> 460
QTcF	msec	> 450

QTcB = QT interval corrected for heart rate using the Bazett formula ($QTcB = QT/(RR)^{1/2}$); QTcF = QT interval corrected for heart rate using the Fridericia formula ($QTcF = QT/(RR)^{1/3}$).

In addition, a listing of all AEs that occurred in subjects who had postbaseline PCS ECG values will be provided.

A listing of subjects with postbaseline clinically significant ECG abnormalities, as reported by the investigator or by the central cardiologist, will also be provided.

The number and percentage of subjects with an increase > 30 msec but ≤ 60 msec, and with an increase > 60 msec in QTcB or QTcF will be tabulated. A supportive listing of subjects with postbaseline QTcB or QTcF increases > 30 msec will be provided, including the SID number, study center number, and all QTcB and QTcF values (including changes from baseline). A listing of all AEs for subjects with postbaseline QTcB or QTcF increases > 30 msec will also be provided.

Shift tables from baseline to the end of double-blind acute treatment period in the Investigator's overall interpretation of the ECG and the central reader's overall interpretation will be presented separately by treatment group for the following categories: normal; abnormal, not clinically significant; abnormal, clinically significant.

10.5 COLUMBIA-SUICIDE SEVERITY RATING SCALE

For the C-SSRS, the number and percentage of subjects with any suicidal ideation or suicidal behavior will be presented by treatment group, separately for the lifetime history, the double-blind acute treatment period and the double-blind down-taper period. The distribution of responses for the most severe suicidal ideation and suicidal behavior will be summarized by treatment group, separately for the lifetime history, the double-blind acute treatment period, and the double-blind down-taper period. Supportive listings will be provided and will include the SID number, treatment group, visit number, lifetime history, and postbaseline values for each subject. Intensity of ideation, suicidal behavior type, and lethality of suicidal behavior will also be included in these listings.

11.0 PHARMACOKINETIC ANALYSIS

PK analysis of Escitalopram will be conducted using a non-linear mixed-effect modeling approach. Population PK parameters, inter-subject, and residual variability will be determined. The effect of covariates on the escitalopram PK will also be investigated. From the developed population PK model (Pop-PK), PK parameters like volume of distribution and clearance will be estimated. Additional PK parameters including, but not limited to, area under the curve at steady-state (AUCSS) will be derived from the individual Bayes estimates from the pop-PK model.

The relationship between efficacy parameters and plasma exposure using a nonlinear mixed effects approach will be explored. The relationship between safety parameters and plasma exposure may also be explored.

A separate analysis plan will be created to detail this approach.

12.0 HEALTH OUTCOMES ANALYSES

Not applicable.

13.0 BLINDED INTERIM ANALYSIS

A blinded interim analysis will be conducted when approximately 75% of randomized subjects have either completed the study or discontinued from the study. The blinded interim analysis is to obtain an estimate of the pooled standard deviation of the change from baseline in the PARS score to Week 8 of the double-blind acute treatment period. The sample size will be re-estimated using simulation in the same manner as the original sample size as described in Section 14 using the estimated pooled standard deviation and dropout rate from the blinded interim analysis. If the estimated pooled standard deviation is larger than the assumed pooled standard deviation due to assumption deviation as well as COVID-19 remote visits or dropout rate is increased, the sample size in placebo and escitalopram treatment groups may be increased to ensure an adequate power. However, due to the difficulties in recruiting pediatric subjects with GAD, the number of subjects in each treatment arm will be capped at 160.

As the sample size re-estimation requires only the pooled standard deviation and dropout rate to be updated, all data and personnel will remain blinded for the interim analysis. Randomization code will not be released for the sample size re-estimation. There are no potential unblinding issues in the sample size re-estimation. As this is a blinded re-estimation, there is no impact on the overall Type I error rate of the study.

14.0 DETERMINATION OF SAMPLE SIZE

The sample size estimate was constructed using a simulation approach as described in Lu K. (2012).

The primary efficacy parameter is the change from baseline to Week 8 of the double-blind acute treatment period in PARS total score. To detect an effect size (treatment group difference of 2.30 units relative to pooled standard deviation of 5.79) of 0.39, a sample size of 128 subjects in the escitalopram treatment group and 128 subjects in the placebo group will provide 85% power based on an MMRM model using simulation method. The simulation assumed a correlation of 0.7 between the repeated measures, and a common dropout rate of 14%, based on historical data of escitalopram in pediatric subjects.

15.0 STATISTICAL SOFTWARE

Statistical analyses will be performed using version 9.4 (or newer) of SAS.

16.0 DATA HANDLING CONVENTIONS

16.1 VISIT TIME WINDOWS

Table 16.1–1 and Table 16.1–2 present the visits assigned for efficacy and safety analyses and the corresponding ranges of treatment days (window) during which an actual visit may occur.

Table 16.1–1. Visit Time Windows

Derived Visit	*Scheduled Visit (Day)*[a]	*Window*
Baseline[b]	Week 0 (Day 1)	Days ≤ 1
Week 1	Week 1 (Day 8)	Days [2, 11]
Week 2	Week 2 (Day 15)	Days [12, 21]
Week 4	Week 4 (Day 29)	Days [22, 35]
Week 6	Week 6 (Day 43)	Days [36, 49]
Week 8	Week 8 (Day 57)	Days ≥ 50 days and within double-blind acute treatment period
End of double-blind acute treatment period[c]	Final or termination visit during the double-blind acute treatment period	

a Relative to the date of the first dose of double-blind investigational product. Day 1 = the date of the first dose of double-blind investigational product.
b The window for selecting baseline values is from 30 days before the first dose of double-blind investigational product to Day 1.
c Presented in analysis tables for safety parameters, including but not limited to electrocardiograms, clinical laboratory values, vital signs, and C–SSRS. If repeated assessments occur at the end of double-blind acute treatment period visit or during any visit window, the algorithm in Section 16.3 will be used to determine which values are used in the analysis.

Table 16.1–2. Visit Time Windows for the Double-Blind Down-Taper Period

Derived Visit	*Scheduled Visit*	*Window*
Week 9/Down-Taper Period	Week 9	Within the double-blind down-taper period (from 1 day after the end of double-blind acute treatment period to the end of the double-blind down-taper period)
End of double-blind down-taper period[a]	Final or termination visit during the double-blind down-taper period	

a Presented in analysis tables for safety parameters collected during the double-blind down-taper period. If any assessments are repeated at the end-of-double-blind-down-taper-period visit or during any visit window, the algorithm in Section 16.3 will be used to determine which values are used in the analysis.

If the visit date is on or after the date of the first dose of double-blind investigational product, the study day is calculated by visit date – date of the first dose of double-blind investigational product + 1. If the visit date is before the date of the first dose of double-blind investigational product, the study day is calculated by visit date – date of the first dose of double-blind investigational product. Therefore, a negative day indicates a day before the start of the double-blind investigational product.

If a subject has 2 or more visits within the same window, the last visit with a nonmissing value will be used for analysis.

16.2 AGE-AND-GENDER-CORRELATED VALUES FOR WEIGHT AND HEIGHT

To adjust weight (kg) and height (cm) for sex and age, one needs to compare them to standard reference values for the same sex and age group, which are available in the United States Growth Charts and can be downloaded from:
http://www.cdc.gov/growthcharts/percentile_data_files.htm

The z-score is calculated as below

$$z = \frac{(X/M)^L - 1}{SL}, \text{ if } L \neq 0 \text{ and}$$

$$z = \frac{\ln(X/M)}{S}, \text{ if } L = 0,$$

where X is the physical measurement (e.g. weight and height) and L, M and S are the values from the appropriate table corresponding to the age in months (or length/stature) and sex (1 = male; 2 = female). X must be in metric measurements (kilograms or meters). This is called LMS method (Cole TJ, 1990), and parameters L, M, and S are the Box-Cox transformation power, median, and standard deviation, respectively, in the reference data, which again are provided in the reference data tables.

16.3 REPEATED OR UNSCHEDULED ASSESSMENTS OF SAFETY PARAMETERS

If a subject has repeated assessments before the date of first dose of double-blind investigational product, the results from the final nonmissing assessment made before the date of the first dose of double-blind investigational product will be used as baseline. If end-of-double-blind-treatment period assessments are repeated or if unscheduled visits occur, the last nonmissing post-baseline assessment during the double-blind acute treatment period will be used as the end-of-double-blind-acute-treatment-period assessment for generating summary statistics. Likewise, if end of double-blind down-taper period assessments are repeated or if unscheduled visits occur, the last non-missing assessment during the double-blind down-taper period will be used as the end-of-double-blind-down-taper-period assessment for generating summary statistics.

Similarly, if a subject has repeated assessments within a given visit window, the last non-missing post-baseline assessment during the window will be used as the assessment for that visit window when generating summary statistics.

However, all post-baseline assessments will be used to determine PCS values for laboratory parameters, vital signs and ECG parameters, and to determine most severe suicidal ideation and most severe suicidal behavior from C-SSRS. All assessments will be presented in the data listings.

16.4 MISSING DATE OF THE LAST DOSE OF INVESTIGATIONAL PRODUCT

When the date of the last dose of investigational product in the study taken during the double-blind acute treatment period is missing for a subject in the Safety Population, all efforts should be made to obtain the date from the Investigator. If it is still missing after all efforts, then the last available dosing record date will be used as the last dose date.

16.5 MISSING SEVERITY ASSESSMENT FOR ADVERSE EVENTS

If severity is missing for an AE that started before the date of the first dose of double-blind investigational product, an intensity of mild will be assigned. If severity is missing for an AE that started on or after the date of the first dose of double-blind investigational product, an intensity of severe will be assigned. The imputed values for severity assessment will be used for the incidence summary. The number of imputed severity value will be presented in the footnote of the incidence summary. The imputed values will be shown as missing in the data listings.

16.6 MISSING CAUSAL RELATIONSHIP TO INVESTIGATIONAL PRODUCT FOR ADVERSE EVENTS

If the causal relationship to the investigational product is missing for an AE that started on or after the date of the first dose of double-blind investigational product, a causality of Likely/Certain will be assigned. The imputed values for causal relationship to double-blind treatment will be used for the incidence summary. The number of imputed severity value will be presented in the footnote of the incidence summary. The imputed values will be shown as missing in the data listings.

16.7 MISSING DATE INFORMATION FOR ADVERSE EVENTS

The following imputation rules only apply to cases in which the start date for AEs is incomplete (i.e., partly missing).

Missing month and day

- If the year of the incomplete start date is the same as the year of the first dose of double-blind investigational product, the month and day of the first dose of double-blind investigational product will be assigned to the missing fields

- If the year of the incomplete start date is before the year of the first dose of double-blind investigational product, *December 31* will be assigned to the missing fields

- If the year of the incomplete start date is after the year of the first dose of double-blind investigational product, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the first dose of double-blind investigational product, the day of the first dose of double-blind investigational product will be assigned to the missing day

- If either the year of the incomplete start date is before the year of the date of the first dose of double-blind investigational product or if both years are the same but the month of the incomplete start date is before the month of the date of the first dose of double-blind investigational product, the last day of the month will be assigned to the missing day

- If either the year of the incomplete start date is after the year of the date of the first dose of double-blind investigational product or if both years are the same but the month of the incomplete start date is after the month of the date of the first dose of double-blind investigational product, the first day of the month will be assigned to the missing day

If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

If the start date is completely missing and the stop date is complete, the following algorithm will be used to impute the start date:

- If the stop date is after the date of the first dose of double-blind investigational product, the date of the first dose of double-blind investigational product will be assigned to the missing start date

- If the stop date is before the date of the first dose of double-blind investigational product, the stop date will be assigned to the missing start date

16.8 MISSING DATE INFORMATION FOR PRIOR OR CONCOMITANT MEDICATIONS

For prior or concomitant medications, including rescue medications, incomplete (i.e., partly missing) start dates and/or stop dates will be imputed. When the start date and the stop date are both incomplete for a subject, the start date will be imputed first.

16.8.1 Incomplete Start Date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication start date. If the stop date is complete (or imputed) and the imputed start date is after the stop date, the start date will be imputed using the stop date.

Missing month and day

- If the year of the incomplete start date is the same as the year of the first dose of double-blind investigational product, the month and day of the first dose of double-blind investigational product will be assigned to the missing fields

- If the year of the incomplete start date is before the year of the first dose of double-blind investigational product, *December 31* will be assigned to the missing fields

- If the year of the incomplete start date is after the year of the first dose of double-blind investigational product, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the first dose of double-blind investigational product, the day of the first dose of double-blind investigational product will be assigned to the missing day

- If either the year of the incomplete start date is before the year of the date of the first dose of double-blind investigational product or if both years are the same but the month of the incomplete start date is before the month of the date of the first dose of double-blind investigational product, the last day of the month will be assigned to the missing day

- If either the year of the incomplete start date is after the year of the date of the first dose of double-blind investigational product or if both years are the same but the month of the incomplete start date is after the month of the date of the first dose of double-blind investigational product, the first day of the month will be assigned to the missing day

16.8.2 Incomplete Stop Date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication stop date. If the date of the last dose of double-blind investigational product is missing, impute it as described in Section 16.4. If the imputed stop date is before the start date (imputed or nonimputed start date), the imputed stop date will be equal to the start date.

Missing month and day

- If the year of the incomplete stop date is the same as the year of the last dose of double-blind investigational product, the month and day of the last dose of double-blind investigational product in the study will be assigned to the missing fields

- If the year of the incomplete stop date is before the year of the last dose of double-blind investigational product, *December 31* will be assigned to the missing fields

- If the year of the incomplete stop date is after the year of the last dose of double-blind investigational product, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete stop date are the same as the month and year of the last dose of double-blind investigational product, the day of the last dose of double-blind investigational product will be assigned to the missing day

- If either the year of the incomplete stop date is before the year of the date of the last dose of double-blind investigational product or if both years are the same but the month of the incomplete stop date is before the month of the date of the last dose of double-blind investigational product, the last day of the month will be assigned to the missing day

- If either the year of the incomplete stop date is after the year of the date of the last dose of double-blind investigational product or if both years are the same but the month of the incomplete stop date is after the month of the date of the last dose of double-blind investigational product, the first day of the month will be assigned to the missing day

16.9 CHARACTER VALUES OF CLINICAL LABORATORY PARAMETERS

If the reported value of a clinical laboratory parameter cannot be used in a statistical summary table because, for example, a character string is reported for a parameter of the numeric type, a coded value must be appropriately determined for use in the statistical analyses. The actual values, as reported in the database will be presented in the data listings.

Table 16.9–1 shows examples of how some possible laboratory results should be coded for analysis.

Table 16.9–1. Examples of Coding Special Character Values for Clinical Laboratory Parameters

Laboratory Test, SI Unit	***Possible Laboratory Results***	***Coded Value for Analysis***
CHEMISTRY		
ALT, U/L	< 5	5
AST, U/L	< 5	5
Bilirubin, total, µmol/L	< 2	2
URINALYSIS		
Glucose, mmol/L	< 2.775, Negative	0
	[2.775, 5.55)	1+
	[5.55, 13.875)	2+
	[13.875, 27.75)	3+
	[27.75, 55.5)	4+
	≥ 55.5	5+
ketones, mmol/L	> 0.444, ≥ 0.444, >0, Trace	Positive
	≤ 0, Negative	Negative
pH	> 8.0, ≥ 8.0	8.0
	≥ 8.5	8.5
Protein, g/L	< 0, Negative	Negative
	[0, 0.15), Trace	0
	[0.15, 0.3)	1+
	[0.30, 1.0)	2+
	[1.0, 5.0)	3+
	>= 5.0	4+

ALT = alanine aminotransferase; AST = aspartate aminotransferase; SI = *Le Système International d'Unités* (International System of Units).

17.0 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

Only the following changes have been made to the protocol version 3.0: 30 Jun 2020.

Per protocol section 11.3, demographic parameters to be summarized include weight z-score and height z-score. However, these will be listed but not summarized.

No other changes have been made to the analyses specified in the protocol.

18.0 REFERENCES

Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr.,* 44: 45-60, 1990.

Kenward MG, Molenberghs G, Thijs H. Pattern-mixture models with proper time dependence. *Biometrika,* 2003;90:53-71.

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics,* 1997;53:983-997.

Lu K. Sample size calculations with multiplicity adjustment for longitudinal clinical trials with missing data. Stat Med 2012;31:19-28.

19.0 APPENDICES

APPENDIX I PATTERN-MIXTURE MODEL DETAILS

For repeated measures with the monotone missing mechanism, the pattern-mixture model with non-future dependent missing assumption, proposed by Kenward et al. (2003), provides a feasible solution to accommodate certain missing not at random (MNAR) mechanism. The methodology relies on constructing unidentifiable conditional densities using identifiable densities and borrows techniques from standard multiple imputation.

1. Non-Future Dependent Missing Assumption

Assume there are *T* designed visits in a longitudinal study and let $y_i (i = 1,2,...,T)$ represent subject's measurement at Visit *i*. When the missing mechanism is monotone, the pattern of missing data can be defined by the number of measurements (L) actually observed from the subject. Let $f(y_i,...,y_j \mid L = t)$ denote the conditional density of y_i ,... y_j, given that the last observed measurement is at Visit *t*. Then the overall density function for Pattern *t* can be written as

$$f(y_1,..., y_T \mid L = t) = f(y_1,..., y_t \mid L = t) f(y_{t+1} \mid y_1,..., y_t, L = t) \times \prod_{s = t+2}^{T} f(y_s \mid y_1,..., y_{s-1}, L = t) \quad (1)$$

Note on the right hand side of (1) the first factor is clearly identifiable from the observed data, while the second and the beyond are not, due to lack of available data. The second factor $f(y_{t+1} \mid y_1,...,y_t, L = t)$ could be identifiable based on an assumed relationship between $f(y_{t+1} \mid y_1,...,y_t, L = t)$ and $f(y_{t+1} \mid y_1,...y_t, L \geq t + 1)$. The third and beyond factors $f(y_s \mid y_1,...,y_{s-1}, L = t)$ (with all $s \geq t + 2$) could be identifiable with the help of non-future dependent missing assumption.

For longitudinal data with dropouts, non-future dependent missing (NFD) mechanism (Kenward et al., 2003) assumes that the unidentifiable conditional distributions of y_s ($s \geq t + 2$), given earlier measurements, in Pattern t, is equal to the corresponding distribution in patterns $L \geq s - 1$:

$$f(y_s \mid y_1,..., y_{s-1}, L = t) = f(y_s \mid y_1,..., y_{s-1}, L \geq s - 1) \quad (2)$$

The right hand side of (2) can further be partitioned into

$$f(y_s \mid y_1,\ldots, y_{s-1}, L \geq s-1) = \sum_{j=s-1}^{T} \omega_{s-1,j} \cdot f(y_s \mid y_1,\ldots, y_{s-1}, L = j) \qquad (3)$$

Where mixture probabilities $\omega_{s-1,j}$ are:

$$\omega_{s-1,j} = \frac{\alpha_j f(y_1,\ldots,y_{s-1} \mid L = j)}{\sum_{t=s-1}^{T} \alpha_t f(y_1,\ldots,y_{s-1} \mid L = t)}, \text{ and } \alpha_j \text{ represents the fraction of} \qquad (4)$$

subjects from Pattern j.

Each factor of the unidentifiable conditional distribution of y_s ($s \geq t+2$) on the right side of (1) can be expressed using the following:

- $f(y_s \mid y_1,\ldots,y_{s-1}, L = s-1)$, the unidentifiable conditional distribution of the first missing in pattern $s-1$,

- $f(y_s \mid y_1,\ldots,y_{s-1}, L = j)$, the identifiable conditional distributions of y_s given $y_1,\ldots,y_{s-1}$ of pattern j ($j \geq s$), and

- α_j, the fraction of subjects from pattern j ($j \geq s-1$).

So under NFD, all the unidentifiable conditional distribution on the right side of (1) can be estimated and missing value could be therefore imputed based on the assumption for unidentifiable conditional distribution of the first missing.

We re-formulate the partition in (3), for $s \geq t+2$, as the following:

$$f(y_s \mid y_1,\ldots, y_{s-1}, L = t) = \delta_{s-1} f(y_s \mid y_1,\ldots, y_{s-1}, L = s-1) + (1-\delta_{s-1}) f(y_s \mid y_1,\ldots, y_{s-1}, L \geq s) \qquad (5)$$

for $s \geq t+2$ with $\delta_{s-1} = \omega_{s-1,s-1}$.

Therefore, under monotone missing and NFD assumption, the unidentifiable conditional densities for Visit s in Pattern t ($s \geq t+2$) can be expressed as a mixture distribution of $f(y_s \mid y_1,\ldots,y_{s-1}, L = s-1)$ - the unidentifiable conditional distribution of the first missing measurement y_s in Pattern $s-1$, and $f(y_s \mid y_1,\ldots,y_{s-1}, L \geq s)$ - the identifiable conditional distribution of y_s from all the patterns with observed data at Visit s or beyond:

$$f(y_s \mid y_1,\ldots, y_{s-1}, L \geq s) = \sum_{j=s}^{T} \lambda_{s-1,j} f(y_s \mid y_1,\ldots, y_{s-1}, L = j) \qquad (6)$$

where the mixture probability

$$\lambda_{s-1,j} = \omega_{s-1,j} / (1 - \omega_{s-1,\,s-1}) = \frac{\alpha_j f(y_1,\ldots,y_{s-1} \mid L = j)}{\sum_{t=s}^{T} \alpha_t f(y_1,\ldots,y_{s-1} \mid L = t)} \text{ for } j \geq s, \text{ where } \alpha_j \text{ is the fraction of} \qquad (7)$$

subjects from Pattern j.

The conditional densities for the first missing are selected as:

$$f(y_s \mid y_1,\ldots, y_{s-1}, L = s-1) = f(y_s - \Delta \mid y_1,\ldots, y_{s-1}, L \geq s) \text{ for } s = 2, \ldots, T, \qquad (8)$$

Note that the two distributions only differ by a shift (Δ) parameter. When $\Delta = 0$, the missing value y_s in Pattern $s-1$ is imputed based on the distribution of all observed data up to Visit s, as a result, leading to missing at random (MAR) missingness. When $\Delta \neq 0$, (8) will introduce a scenario of MNAR. The similar idea was also presented in the recent publication "The Prevention and Treatment of Missing Data in Clinical Trials" by the National Academies Press. The selection of the plausible values for the shift parameter (Δ) is discussed in Section 3 of this appendix.

Note that per recommendation in Wang and Daniels (2011), only the observed data within pattern is assumed to be multivariate normal. The observed data distribution can be expressed in terms of the marginal distribution of baseline measurement and the conditional distributions of postbaseline measurements given earlier measurements. Assuming that these distributions are normal, the linear regression of each observation on prior observations will yield least-squares estimates of model parameters that can be utilized for independent posterior draws of model parameters for observed data. Multiple imputation approach will be used to estimate the overall mean at the final time point.

2. Imputation Procedure

All the missing data will be imputed to create complete datasets, then statistical analysis can be performed using appropriate techniques such as ANCOVA. The imputation can accommodate MNAR missing data mechanisms, based on the theory discussed in the previous section.

The model parameters for each dropout pattern, i.e., the mean, variance and proportions of observations in each pattern, are drawn from their posterior distributions prior to the imputation of missing data for a single imputation.

The details of imputation within a pattern, say Pattern t, are as the following:

Step 1. Impute the first missing value y_{t+1} for each subject in Pattern t $(t = 1,..., T - 1)$:

a. Compute estimates of mixture probabilities $\lambda_{s-1,j}$ in (7) with $s = t+1$ given the posterior draw of proportions of observations in each pattern and the posterior draw of regression parameters for the observed data.

b. Draw a random integer from $\{s, ..T\}$ to index a component distribution on the right hand side of (6), using mixture probabilities obtained in a). Draw y^*_{t+1} from the identified component normal distribution. Impute the missing y_{t+1} as $\tilde{y}_{t+1} = y^*_{t+1} + \Delta$.

Step 2. Impute the rest of the missing values of $y_{t+2}, y_{t+3}, ..., y_T$ for subjects in Pattern t:

Starting with imputation for y_{t+2}, first, similar to Step 1, draw y^*_{t+2} from the normal mixture (6) based on the observed $y_1,...,y_t$ and the already imputed $\tilde{y}_{t+1}$ for the subject. Then the missing y_{t+2} is imputed as $\tilde{y}_{t+2} = y^*_{t+2} + \Delta$ with probability δ_{t+1} and as $\tilde{y}_{t+2} = y^*_{t+2}$ with probability $1 - \delta_{t+1}$, where the mixture probability $\delta_{t+1} = \omega_{t+1,t+1}$ is obtained from (4) given the posterior draw of proportions of observations in each pattern and the posterior draw of regression parameters for the observed data.

Missing values of y_{t+3} through y_T can be imputed similarly as y_{t+2}.

To summarize, the imputations of y_{t+1} through y_T is done recursively within each Pattern t (for all $t = 1,...,T - 1$) to create a complete dataset after imputation is done for all patterns with missing values.

The above imputation procedure is applied to all subjects in each missing data pattern to create a single imputed data set. Repeating the process of drawing parameters from the posterior distribution and imputing missing data given the posterior draw m times will yield m imputed data sets. The observed or imputed values at the final data point are averaged to obtain the overall mean estimate for each imputed data set, and the multiple imputation estimate is obtained by averaging the estimates across m imputations.

In this sensitivity analysis, m is set to equal to 20. The value of m is discussed in the context of imputation efficiency in standard multiple imputation theory (Rubin, 1987, p. 114), and m = 20 would provide at least 96% of relative efficiency as compared with using an large number of imputations (SAS/STAT User's Guide, p. 3796).

3. Determination of the Shift Parameter Values

The common shift parameter Δ is the difference between the mean of y_{t+1} among those who drop out at Visit t and those who remain beyond Visit t ($1 \leq t \leq T - 1$). The exact value of Δ is unknown and can't be estimated from data because of missingness. The magnitude of Δ depends on the medical aspects of the trial. Using relevant historical data, one may select Δ as a proportion of the sample standard deviation or a proportion of observed treatment efficacy.

References:

Kenward MG, Molenberghs G, Thijs H. Pattern-mixture models with proper time dependence. *Biometrika,* 2003;90:53-71.

Rubin DB. Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York; 1987.

SAS Institute Inc. SAS/STAT 9.2 User's Guide, Second Edition. Cary, NC: SAS Institute Inc; 2009.

Wang C, Daniels MJ. A note on MAR, identifying restrictions, model comparison, and sensitivity analysis in pattern mixture models with and without covariates for incomplete data. *Biometrics,* 2011; 67(3):810-818.