NCT02431806

Study ID: LVM-MD-11

Title: A Double-blind, Placebo- and Active-Controlled Evaluation of the Safety and Efficacy of Levomilnacipran ER in Adolescent Patients With Major Depressive Disorder

Statistical Analysis Plan Amendment 2 Date: 3 Oct 2019

1.0 TITLE PAGE

Forest Research Institute, Inc.
Harborside Financial Center, Plaza V
Jersey City, NJ 07311

LVM-MD-11

A Double-blind, Placebo- and Active-Controlled Evaluation of the Safety and Efficacy of Levomilnacipran ER in Adolescent Patients with Major Depressive Disorder

STATISTICAL ANALYSIS PLAN

Original SAP Date: 06 Mar 2015

Amendment #1: 16 July 2015

Amendment #2: 3 October 2019

2.0 TABLE OF CONTENTS

3.0 LIST OF ABBREVIATIONS

AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ANCOVA	analysis of covariance
BP	blood pressure
CDRS-R	Children's Depression Rating Scale-Revised
CGI-I	Clinical Global Impression - Improvement
CGI-S	Clinical Global Impression - Severity
[redacted]	[redacted]
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision
ECG	electrocardiogram, electrocardiographic
eCRF	electronic case report form
ET	early termination
GLMM	generalized linear mixed model
IP	investigational product
ITT	intent to treat
LAR	legally authorized representative
LOCF	last observation carried forward
MDD	major depressive disorder
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed-effects model for repeated measures
MNAR	missing not at random

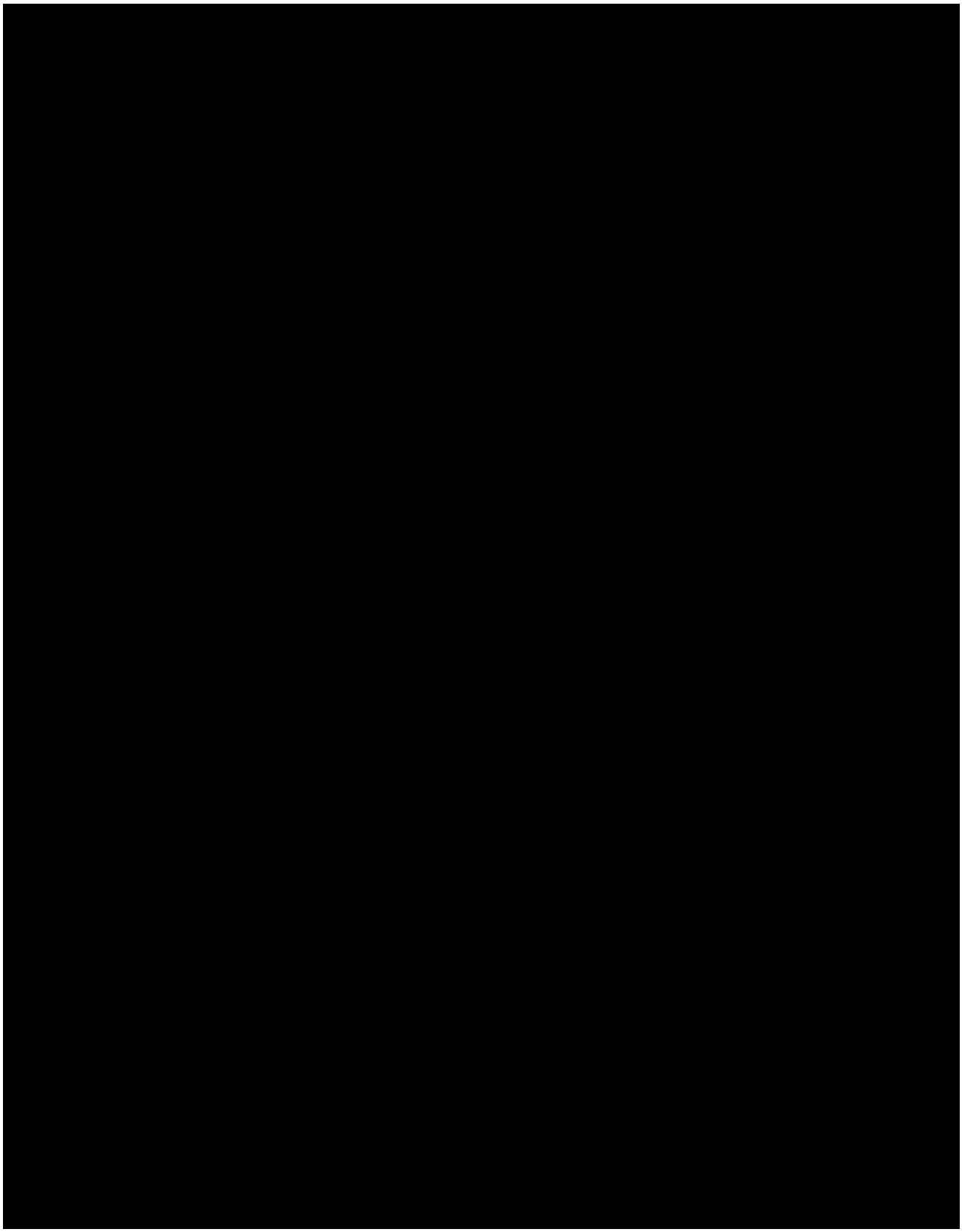
NEAE	newly emergent adverse event
NFD	non-future dependence
OC	observed cases
PCS	potentially clinically significant
PID	patient identification
PK	pharmacokinetic
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett formula
QTcF	QT interval corrected for heart rate using the Fridericia formula
SAE	serious adverse event
SAP	statistical analysis plan
SI	*Le Système International d'Unités* (International System of Units)
TBL	total bilirubin
TEAE	treatment-emergent adverse event
ULN	upper limit of normal laboratory reference range

4.0 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 LVM-MD-11. Specifications of tables, figures, and data listings are contained in a separate document. The statistical analysis for pharmacokinetic parameters will be specified in a separate document.

Study LVM-MD-11 is a multicenter, randomized, double-blind, placebo- and active-controlled, parallel-group, fixed-dose study in adolescent patients, ages 12 - 17 years, who have been diagnosed with major depressive disorder (MDD) using Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) diagnostic criteria. This study includes 4 treatment groups: placebo, levomilnacipran 40 mg/day, levomilnacipran 80 mg/day, and fluoxetine 20 mg/day. A total of ***544*** patients (***136*** per treatment group) are planned to be randomized.

The length of this study will be 10 weeks, starting with a 1-week screening period, followed by 8-week of double-blind treatment period, and 1-week of double-blind down-taper period. Patients must provide assent to participation and their legally authorized representative (LAR) and caregiver must provide written informed consent prior to the conduct of any study-specific procedures. At the end of Visit 2 (Baseline), patients meeting eligibility criteria will be randomized in 1:1:1:1 ratio to treatment groups of placebo, levomilnacipran 40 mg, levomilnacipran 80 mg, and fluoxetine 20 mg, respectively. Patients randomized to the levomilnacipran 40 mg/day group will receive 10 mg/day for Days 1 to 2 (2 days), 20 mg/day for Days 3 to 7 (5 days), and 40 mg/day for Weeks 2 through Week 8; patients randomized to the levomilnacipran 80 mg/day group will receive 10 mg/day for Days 1 to 2 (2 days), 20 mg/day for Days 3 to 4 (2 days), 40 mg/day for Days 5 to 7, and 80 mg/day for Weeks 2 through Week 8; Patients randomized to fluoxetine 20 mg/day group will receive 10 mg/day for Week 1, and 20 mg/day for Weeks 2 through 8. (Table 4–1).



5.0 OBJECTIVES

The objective of this study is to evaluate the efficacy, safety, and tolerability of levomilnacipran relative to placebo in adolescent outpatients (12 - 17 years) with MDD.

In addition, the study is designed to obtain pharmacokinetic (PK) data to guide the dose selection for future levomilnacipran pediatric studies.

6.0 PATIENT POPULATIONS

Four populations will be considered in the statistical analysis of the study as specified below.

6.1 SCREENED POPULATION

The Screened Population will consist of all patients who underwent a Screening Visit, received a screening number, and for whom informed consent was obtained.

6.2 RANDOMIZED POPULATION

The Randomized Population will consist of all patients in the Screened Population who were randomized to a treatment group in the study.

6.3 SAFETY POPULATION

The Safety Population will consist of all patients in the Randomized Population who took at least 1 dose of double-blind investigational product.

6.4 INTENT-TO-TREAT POPULATION

The Intent-to-Treat (ITT) Population will consist of all patients in the Safety Population who had the baseline and at least 1 postbaseline assessment of the CDRS-R total score.

7.0 PATIENT DISPOSITION

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

Screen-failure patients (i.e., patients screened but not randomized) and the associated reasons for failure to randomize will be tabulated overall for the Screened Population. The number and percentage of patients who complete the double-blind treatment period and of patients who prematurely discontinue during the same period will be presented for each treatment group and pooled across treatment groups for the Safety Population. The number and percentage of patients who enter the double-blind down-taper period and of patients who complete the double-blind down-taper period will be presented for each treatment group and pooled across treatment groups for the Safety Population.

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

8.0 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographic parameters (age, sex, race, and ethnicity) and other baseline characteristics (weight, height, body mass index) will be summarized descriptively by treatment group for the Safety and ITT populations, respectively. Baseline efficacy variables will be summarized by treatment group for the ITT Population.

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

Medical and surgical history/physical findings will be classified by system organ class and preferred term using the *Medical Dictionary for Regulatory Activities*, version 16.1 or newer. The number and percentage of patients with abnormalities in medical and surgical histories in each system organ class and preferred term will be summarized by treatment group for the Safety Population. Psychiatric history of MDD and nondrug psychiatric treatment history will also be summarized by treatment group for the Safety Population.

The *World Health Organization Drug Dictionary*, version March 2013 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.

Both prior and concomitant medications use will be coded by drug name and therapeutic class. The use of prior and concomitant medications will be summarized by the number and percentage of patients receiving each drug within each therapeutic class in each treatment group for the Safety Population. If a patient took a specific medication multiple times or took multiple medications within a specific therapeutic class, that patient 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 treatment period and the double-blind down-taper period, separately. Any concomitant medications 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 patient data listings.

9.0 EXTENT OF EXPOSURE AND TREATMENT COMPLIANCE

9.1 EXTENT OF EXPOSURE

Exposure to the investigational product for the Safety Population during the double-blind treatment period will be summarized in terms of treatment duration, calculated as the number of days from the date of the first dose of double-blind investigational product to the date of the last dose of double-blind investigational product during the double-blind treatment period, inclusive. Descriptive statistics (number of patients, mean, SD, median, minimum, and maximum) will be presented by treatment group. The number and percentage of patients will be presented for specific intervals of exposure by treatment group.

Patient-years, defined as total exposure to double-blind investigational product in years (excluding the double-blind down-taper period), will be summarized by treatment group for the Safety Population.

In addition, weekly and overall mean daily dose of investigational product during the double-blind treatment period will be summarized by treatment group for the Safety Population.

9.2 MEASUREMENT OF TREATMENT COMPLIANCE

Dosing compliance for a specified period is defined as the number of capsules actually taken by a patient during that period divided by the number of capsules prescribed for the same period multiplied by 100. The total number of capsules actually taken during a specific time period will be calculated from the study medication records. The number of capsules expected 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 prescribed to be taken per day.

Descriptive statistics for investigational product compliance during the double-blind treatment period will be presented by treatment group for each period between 2 consecutive visits, as well as for the whole double-blind treatment period, for the Safety Population.

10.0 EFFICACY ANALYSES

The efficacy analyses will be based on the ITT Population. *Baseline* for each efficacy parameter is defined as the last nonmissing 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 patients in any treatment group in the ITT Population. Small centers will be pooled to form pseudo-centers so that each treatment group includes at least 2 ITT patients within the center. Pooling will be done using the following algorithm:

Based on the number of ITT patients, 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.

In addition, by-visit analyses using the last-observation-carried-forward (LOCF) approach will be presented for all efficacy parameters. For the LOCF approach, only the postbaseline total score of a parameter will be imputed; individual item scores will not be carried forward to derive the total score. Baseline total score will be carried forward only for the intermittent missing scores immediately after baseline. If all the postbaseline values are missing, the baseline value will not be carried forward.

10.1 PRIMARY EFFICACY PARAMETER

The primary efficacy parameter is the change from baseline to Week 8 in CDRS-R total score. The primary analysis for comparing levomilnacipran 40 mg vs. placebo, and levomilnacipran 80 mg vs. placebo for the primary efficacy parameter will be performed using an MMRM with treatment group, pooled study center, visit, and treatment group-by-visit interaction as fixed effects and the baseline value and baseline value-by-visit interaction as covariates. An unstructured covariance matrix will be used to model the covariance of within-patient scores. The Kenward-Roger approximation (Kenward and Roger, 1997) will be used to estimate denominator degrees of freedom. This analysis will only use the observed cases of postbaseline scores without imputation of missing values. To control the overall type I error rate for the multiple comparisons across the primary and the secondary hypotheses, the matched parallel gatekeeping procedure (Chen et al.,2005) will be applied. Details are provided in Section 10.2.

Graphical display of treatment difference vs. placebo (mean difference ± SE) in primary efficacy parameter will be provided by study center.

In addition, two sensitivity analyses, LOCF and pattern-mixture model approaches, will be performed on the primary efficacy parameter.

The LOCF approach is based on an analysis-of-covariance (ANCOVA) model including treatment group and pooled study center as factors and baseline CDRS-R total score as a covariate. The pattern-mixture model approach is based on the non-future dependent missing value restrictions (Kenward et al., 2003) and is 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 patient's last visit with an observed value. The observed CDRS-R total score at a visit is assumed to have a linear relationship with the patient'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 all missing values imputed will be analyzed using an ANCOVA model with treatment group and pooled study center as factors and baseline CDRS-R total 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 range of values for the shift parameter Δ is selected as 0 to 6 based on experience with historical data. Technical details of the proposed pattern-mixture model approach (eg, 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.

Since the investigator of Site 054 didn't conduct the study in accordance with signed statement, one more sensitivity analysis excluding all data from Site 054 will be performed on the primary efficacy parameter. The same MMRM would be applied on all OC of change from baseline to Week 8 in CDRS-R total score except all values of Site 054.

10.2 SECONDARY EFFICACY PARAMETER

The secondary efficacy parameter is the change from baseline to Week 8 in CGI-S score, which will be analyzed using the MMRM approach similar to the one used for the primary efficacy parameter. A sensitivity analysis will also be performed using the LOCF approach as described in Section 10.1.

To control the overall type I error rate for multiple comparisons across the primary and the secondary efficacy parameters and multiple doses, the matched parallel gatekeeping procedure (Chen et al., 2005) will be implemented. The primary hypotheses family, F_1, consists of 2 null hypotheses H_{11} and H_{12} which are for comparisons of levomilnacipran 40 mg and 80 mg, respectively, with placebo in regard to the primary efficacy parameter, change from baseline to Week 8 in CDRS-R total score. Similarly for the secondary efficacy parameter, change from baseline to Week 8 in CGI-S score, we have the corresponding hypotheses family $F_2 = \{H_{21}, H_{22}\}$. Family F_1 will serve as the parallel gatekeeper for F_2. Then the matched gatekeeper procedure utilizes the special logical relationship between the primary and the secondary parameters to enhance the power of statistical testing, i.e., primary and secondary hypotheses are matched on dose. The secondary efficacy parameter will be tested at a specific dose only if the corresponding primary efficacy parameter is statistically significant. Weighted Simes test will be performed to derive the local p-values for the intersection hypotheses. Assignment of weight for all $2^4 - 1 = 15$ intersection hypotheses is shown in Table 10.2–1.

Table 10.2–1. Weights for Intersection Tests in the Matched Parallel Gatekeeping Procedure

Intersection hypothesis	***Weight***			
	H_{11}	H_{12}	H_{21}	H_{22}
$H_{11} \cap H_{12} \cap H_{21} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{21}$	0.5	0.5	0	0
$H_{11} \cap H_{12} \cap H_{22}$	0.5	0.5	0	0
$H_{11} \cap H_{12}$	0.5	0.5	0	0
$H_{11} \cap H_{21} \cap H_{22}$	0.5	0	0	0.5
$H_{11} \cap H_{21}$	1	0	0	0
$H_{11} \cap H_{22}$	0.5	0	0	0.5
H_{11}	1	0	0	0
$H_{12} \cap H_{21} \cap H_{22}$	0	0.5	0.5	0

$H_{12} \cap H_{21}$	0	0.5	0.5	0
$H_{12} \cap H_{22}$	0	1	0	0
H_{12}	0	1	0	0
$H_{21} \cap H_{22}$	0	0	0.5	0.5
H_{21}	0	0	1	0
H_{22}	0	0	0	1

The adjusted p-values for the four elementary hypotheses will be calculated from the local p-values based on the closed testing principle. Statistical significance is determined by comparing the adjusted p-values to $\alpha = 0.05$. The multiplicity method controls the family-wise type I error rate since the Simes inequality holds for multivariate normal distributions with nonnegative correlations (Sarkar, 2008). The large sample size of this study ensures asymptotic normality while nonnegative correlations are well-known among different active versus placebo comparisons and are expected to exist across different time points and between the CDRS-R and CGI-S endpoints based on historical data.

11.1 ADVERSE EVENTS

Adverse events will be coded by system organ class and preferred term using the Medical Dictionary for Regulatory Activities, version 16.1 or newer.

An AE (classified by preferred term) that occurs during the double-blind treatment period or during the double-blind down-taper period 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 treatment period or during the double-blind 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 treatment period or during the double-blind 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 double-blind down-taper period will be considered a newly emergent AE (NEAE) if it is not present before the start of the double-blind down-taper period or was present before the start of the double-blind down-taper period but increased in severity during the double-blind down-taper period. The NEAEs during the double-blind down-taper period will be summarized by body system, preferred term, and treatment group for all patients who enter the double-blind down-taper period.

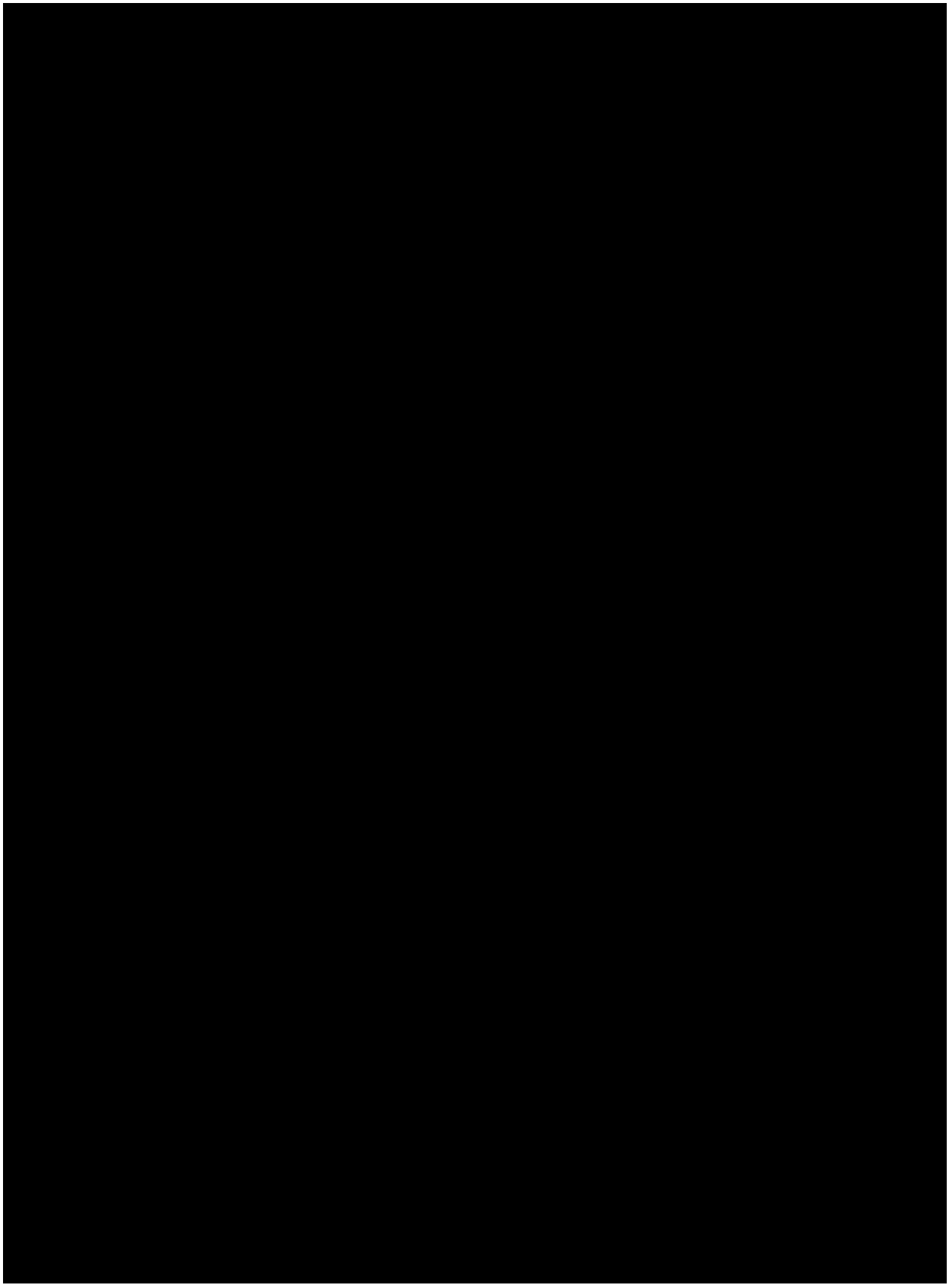
The number and percentage of patients reporting TEAEs in each treatment group will be tabulated by system organ class and preferred term and further categorized by severity and causal relationship to the investigational product, for both the double-blind treatment period and the double-blind down-taper period, separately. If more than 1 AE is coded to the same preferred term for the same patient, the patient will be counted only once for that preferred term using the greatest severity and strictest causality for the summarization by severity and by causal relationship to the investigational product, respectively.

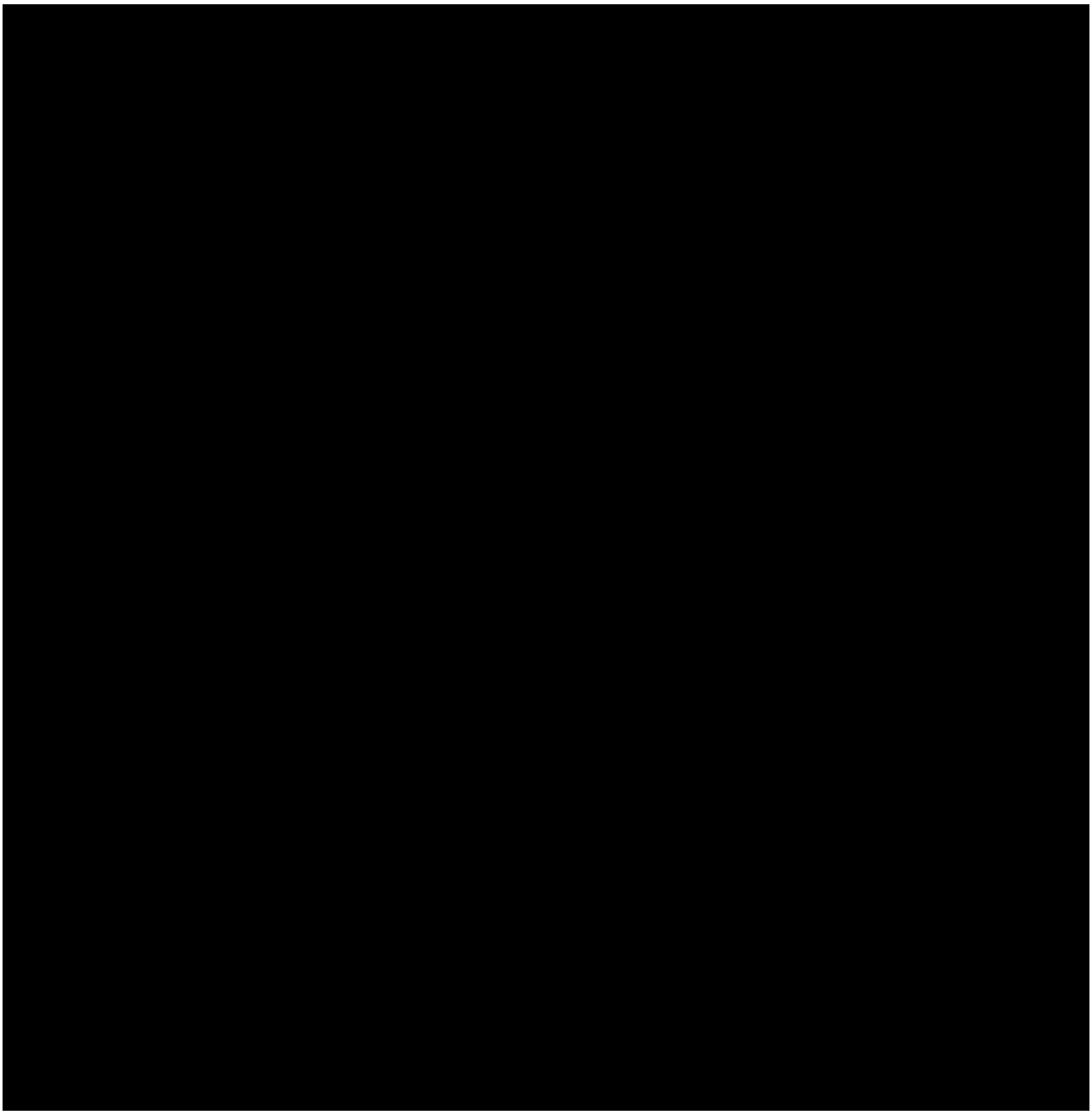
The distribution of TEAEs by severity and causal relationship to the investigational product will be summarized by treatment group separately for the double blind treatment period and the double-blind down-taper period. Common TEAE during the double-blind treatment period is defined in 2 ways: $\geq 1\%$ of patients in any treatment group and $\geq 2\%$ of patients in any treatment group. The incidence of common TEAEs will be summarized separately by preferred term, and treatment group and will be sorted by decreasing frequency for the levomilnacipran treatment group (first levomilnacipran 80 mg/day, then levomilnacipran 40 mg/day).

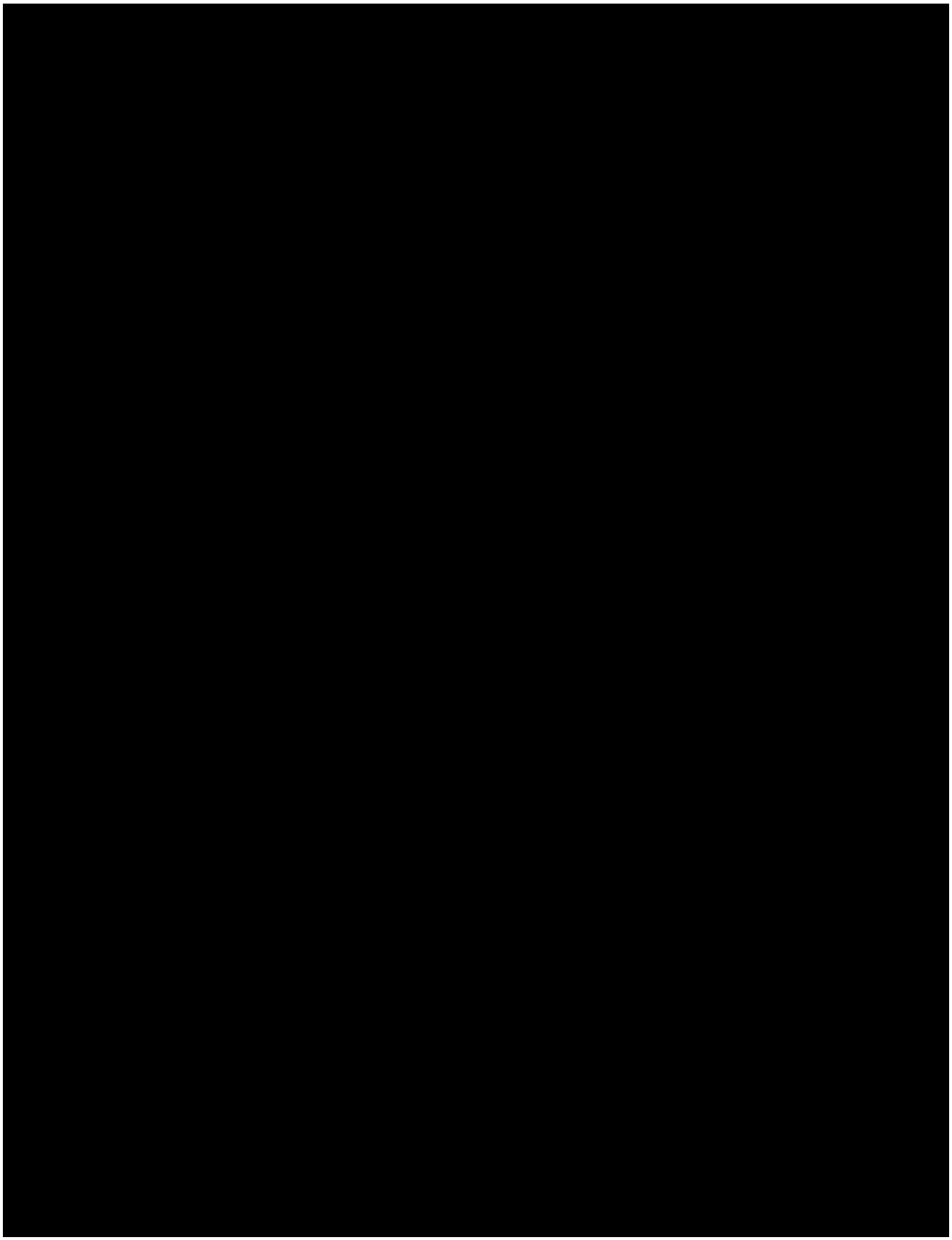
A serious adverse event (SAE) that occurred between the date of the first dose of double-blind investigational product and 30 days after the date of the last dose of double-blind investigational product, inclusive, will be considered an on-therapy SAE.

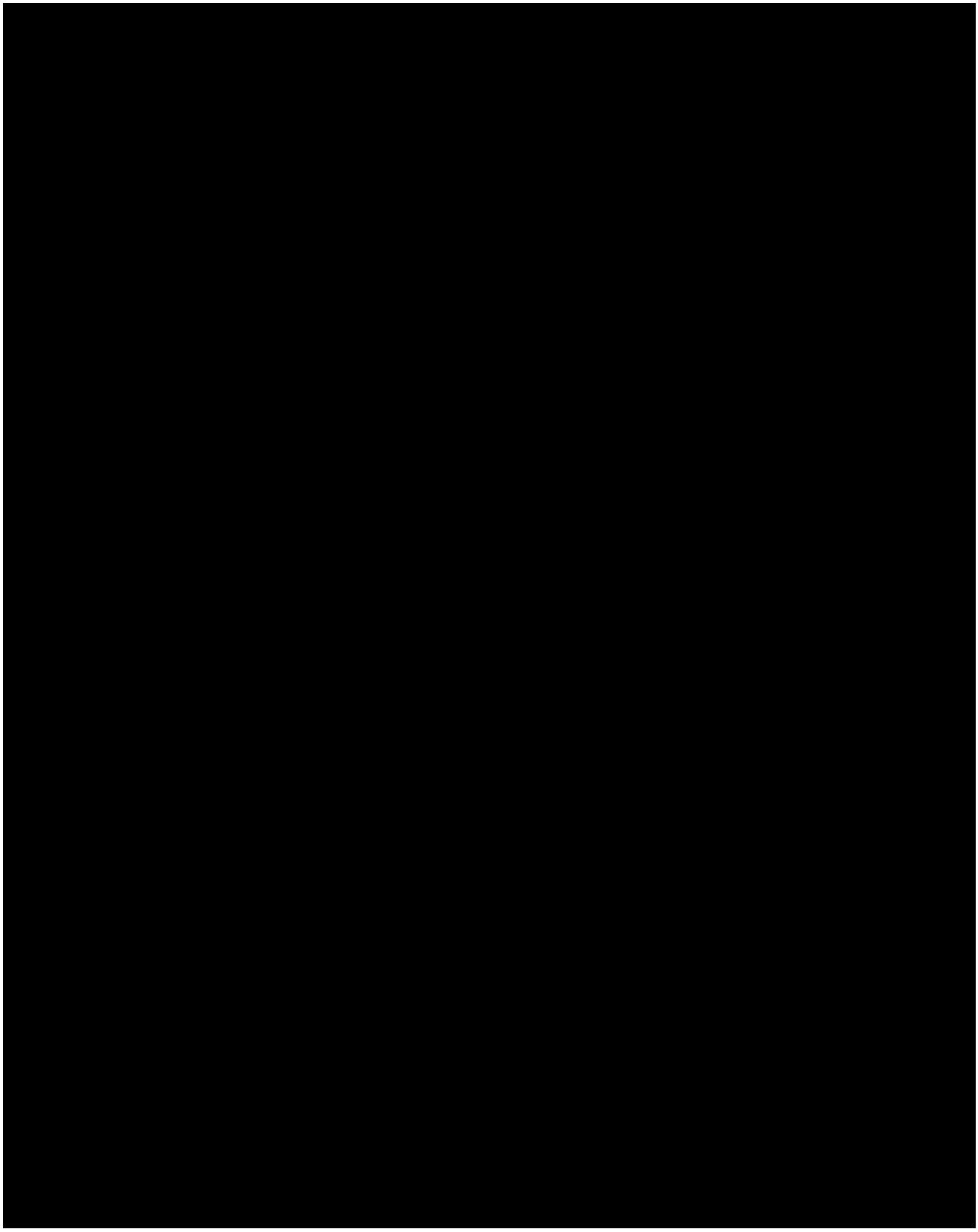
The number and percentage of patients who have on-therapy SAEs will be summarized by preferred term and treatment group and will be sorted by decreasing frequency for the levomilnacipran treatment groups. In addition, the incidence of on-therapy SAEs that led to death will be summarized separately by treatment group and preferred term. The number and percentage of patients in the Safety Population who have AEs leading to premature discontinuation of the investigational product will be summarized by preferred term and treatment group and will be sorted by decreasing frequency for the levomilnacipran treatment group.

Listings will be presented for patients with SAEs, patients with AEs leading to discontinuation, and patients who died (if any) for the Screened Population. All patients 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 patients discontinuing because of AEs before taking double-blind investigational product will be included in these listings.









12.0 HEALTH OUTCOMES ANALYSES

Not applicable.

13.0 BLINDED INTERIM ANALYSIS

A blinded interim analysis will be conducted when approximately 75% of randomized patients 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 to Week 8 in the CDRS-R total score. If the estimated pooled standard deviation is larger than the assumed pooled standard deviation specified in Determination of Sample Size (Section 14.0), the sample size may be increased to ensure an adequate power. However, due to the difficulties in recruiting pediatric patients with MDD, the total number of patients of the study will be capped at ***800*** (***200*** per treatment group). Detailed specification of the blinded interim analysis will be provided in the blinded interim analysis plan, a separate document.

14.0 DETERMINATION OF SAMPLE SIZE

Original sample size calculation: the effect size (treatment group difference relative to pooled standard deviation) of ***0.36*** for both levomilnacipran and fluoxetine is based on a treatment difference of ***4 units*** with a common pooled standard deviation of ***11.1*** for the primary efficacy parameter, change from baseline to Week 8 in CDRS-R total score. Adjusting for multiple comparisons of two levomilnacipran groups with placebo across the primary and secondary endpoints by using the matched parallel gatekeeping procedure, a sample size of ***660*** patients (***165*** per treatment group) will provide 85% power for the primary analysis (levomilnacipran vs. placebo) and for assay sensitivity analysis (fluoxetine vs. placebo) based on an MMRM model using simulation method (Lu, 2012). The simulation assumed a correlation of 0.7 between the repeated measures, and a dropout rate of 17%, based on historical data in pediatric patients.

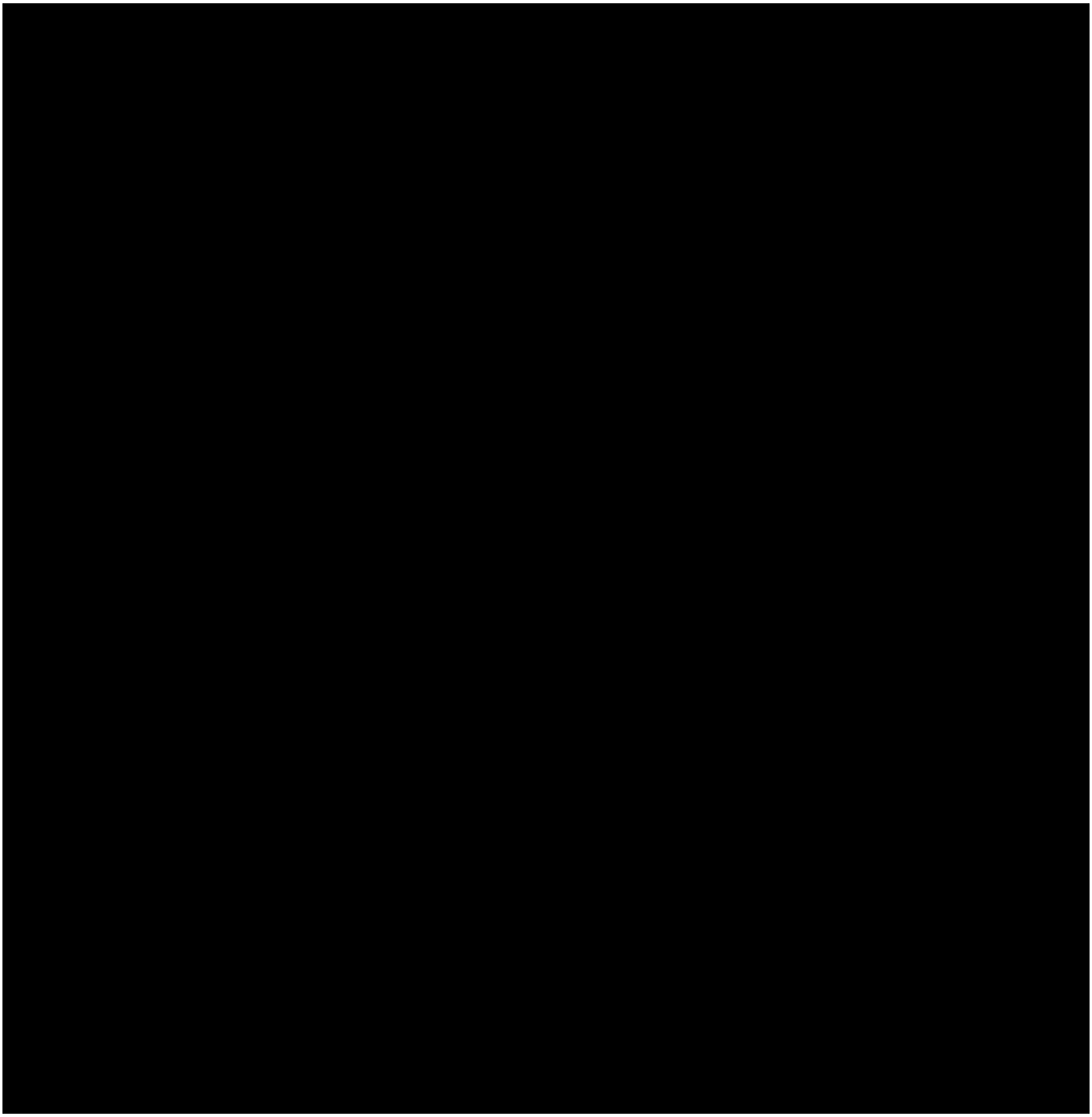
Sample size recalculation based on discussion with FDA: following the discussion with FDA, the study will be deemed as successful if statistical significance is achieved in at least one of two levomilnacipran doses (40 mg/day and 80 mg/day) versus placebo based on the primary efficacy endpoint. Therefore, resultant sample size is ***520*** patients (***130*** per treatment group) as required by 85% power based upon statistical significance in at least one of the two dose levels with the same effect size assumption.

Sample size recalculation based on interim analysis: following a blinded interim analysis, the sample size has been re-estimated. The MMRM model to estimate the pooled variance for change from baseline in CDRS-R total score at Week 8 includes pooled study site and week as factors and baseline CDRS-R total score and baseline value-by-week interaction as covariates, where week is included as a class variable. A total of 421 patients were included in the ITT analysis. Of these 421 patients, 347 (82.4%) patients have observed data at Week 8, so the estimated drop-out rate at Week 8 is 17.6%. Based on the MMRM model, the variance estimate for the change from baseline in CDRS-R total score at Week 8 is 129.53, and the estimated pooled standard deviation is 11.38. Thus the estimated effect size is 0.35 based on a treatment difference of 4 units and a pooled standard deviation of 11.38. Assuming an effect size of ***0.35***, a common dropout rate of ***17.6%***, and a correlation of 0.7 between the repeated measures, an estimated total sample size of ***544*** patients (***136*** per treatment group) will be needed to maintain the 85% power to detect the treatment difference of 4 units for at least one levomilnacipran group versus placebo and for fluoxetine versus placebo.

15.0 STATISTICAL SOFTWARE

Statistical analyses will be performed using [redacted]

16.0 DATA HANDLING CONVENTIONS



[redacted]

16.2 DERIVED EFFICACY VARIABLES

The efficacy variables are derived as follows:

- *CDRS-R responders* are defined as patients with a ≥ 40% reduction from baseline in CDRS-R total score. The value is 1 for CDRS-R responders and 0 otherwise

- *CDRS-R remitters* are defined as patients with CDRS-R ≤ 28. The value is 1 for CDRS-R remitters and 0 otherwise

The total score at a particular visit will be calculated using (sum of nonmissing items) × (total number of items) / (number of nonmissing items) only if the number of missing items is ≤ 2 for the CDRS-R score.

If a patient misses a postbaseline visit or if his/her postbaseline visit is outside of the visit time window, a record for the scheduled visit will be imputed using the last observed nonmissing value immediately before the missing value. If the missing value occurs at Week 1, the baseline value will be carried forward for Week 1, provided that at least 1 subsequent postbaseline assessment is available. For a composite scale such as CDRS-R total score, individual items of the rating scale will not be carried forward. Only total scores will be carried forward using the LOCF approach.

16.3 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 (eg 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.

16.4 REPEATED OR UNSCHEDULED ASSESSMENTS OF SAFETY PARAMETERS

If a patient 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 postbaseline assessment during the double-blind treatment period will be used as the end-of-double-blind-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. However, all postbaseline 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.

16.5 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 treatment period is missing for a patient 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 visit date will be used as the last dose date.

16.6 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 values will be shown as missing in the data listings.

16.7 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 yes will be assigned. The imputed values for causal relationship to double-blind treatment will be used for the incidence summary; the values will be shown as missing in the data listings.

16.8 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

16.9 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 patient, the start date will be imputed first.

16.9.1 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

16.9.2 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, replace it with the last visit date in the imputations described below. If the imputed stop date is before the start date (imputed or non-imputed 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

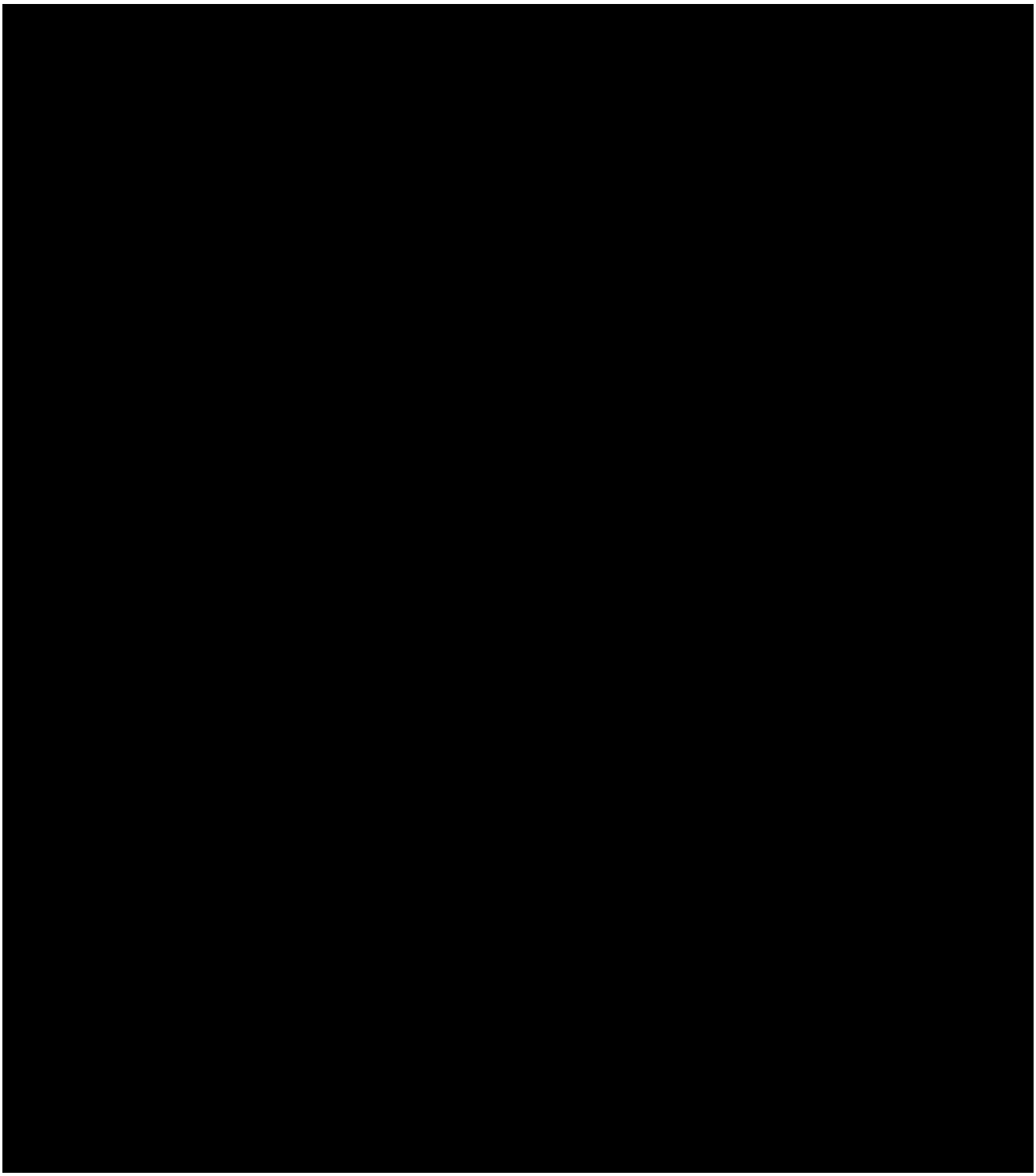
- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

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



17.0 CHANGES TO ANALYSES SPECIFIED IN PROTOCOL

The major change to the analyses specified in the Protocol Amendment #3, dated 25 March 2019, is as follows:

- Delete the Fisher exact test to compare the percentage of premature discontinuation of levomilnacipran with placebo (specified in Section 9.7.2 of the protocol amendment #3) to reflect the current sponsor standards.

- Delete the comparability analysis with placebo for continuous variables and categorical variables (specified in Section 9.7.3 of the protocol amendment #3) to reflect the current sponsor standards.

18.0 REFERENCES

Chen X, Luo X, Capizzi, T. The application of enhanced parallel gatekeeping strategies. *Statistics in Medicine,* 2005;24:1385-1397.

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

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.

Sarkar SK. On the Simes inequality and its generalization. IMS Collections, Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen, 2008; 1: 231-242.

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 patient'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 patient. 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 (NFD) missing 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:

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

patients 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 patients 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 $f(y_s \mid y_1,...,y_{s-1}, L = s-1)$ - the unidentifiable conditional distribution of the first missing measurement y_s in Pattern $s-1$, and $f(y_s \mid y_1,...,y_{s-1}, L \geq s)$ - the identifiable conditional distribution of y_s from all the patterns with observed data at Visit s or beyond:

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

where the mixture probability

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

patients from Pattern j.

The conditional densities for the first missing are selected as:

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

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.

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 patient 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 patients 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 patient. 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 ($1 \le t \le T-1$). 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.

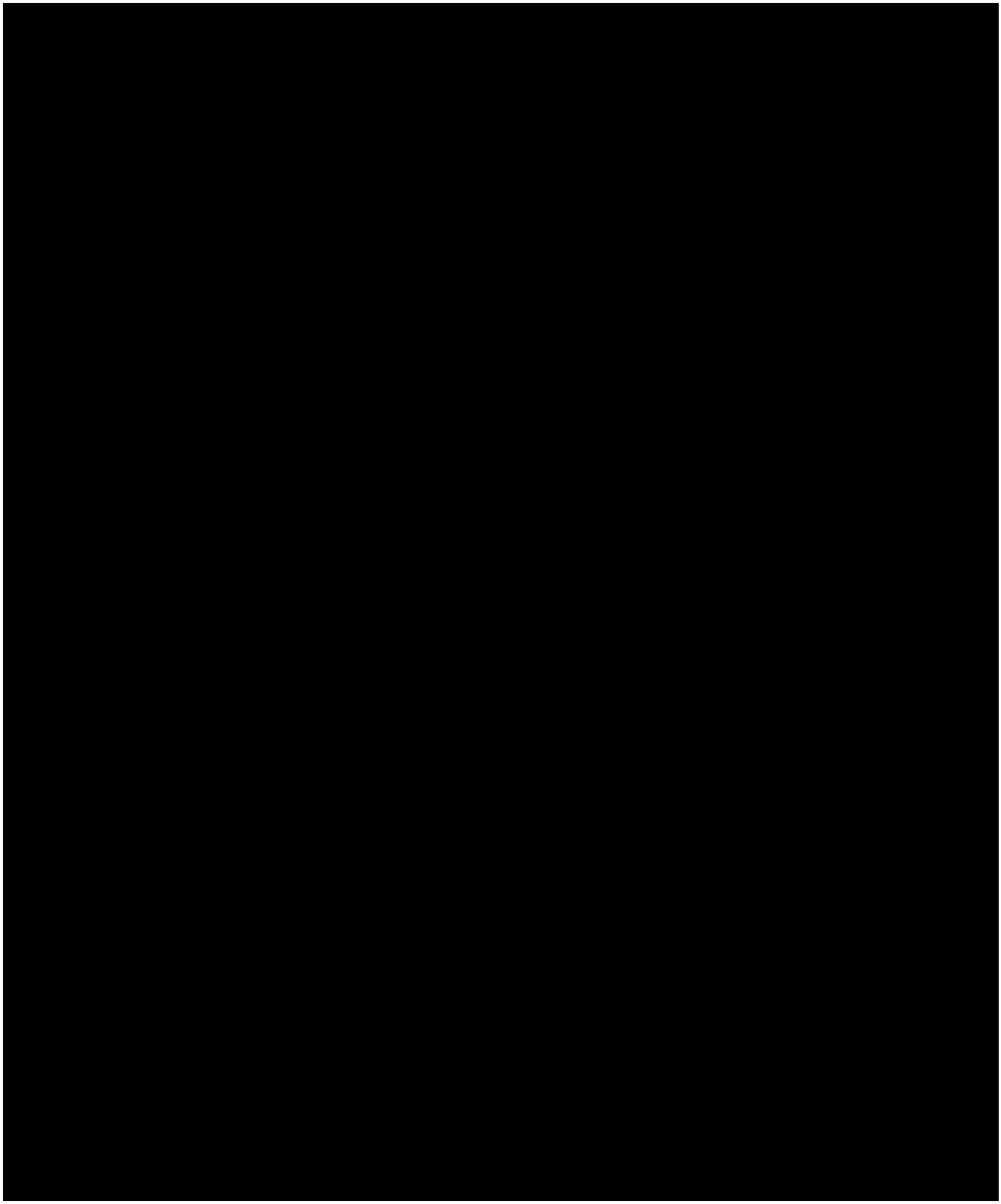
References:

Kenward MG, Molenberghs G, Thijs H. Pattern-mixure 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.



APPENDIX III SAP AMENDMENT #2 SUMMARY

Title: A Double-blind, Placebo- and Active-Controlled Evaluation of the Safety and Efficacy of Levomilnacipran ER in Adolescent Patients with Major Depressive Disorder

Study LVM-MD-11 SAP Amendment #2

Date of Amendment #2: 3 October 2019

Amendment Summary

This summary includes changes made to SAP LVM-MD-11 Amendment #1 (16 July 2015).

Minor editorial and document formatting revisions have not been summarized. The following is a summary of the major changes that were made to each section of the SAP, and a brief rational for these changes.

Section	Major Revision	Rationale
Section 7.0	*Delete the Fisher exact test to compare the percentage of premature discontinuation of levomilnacipran with placebo.*	To reflect current sponsor standards.
Section 8.0	*Delete the comparability analysis with placebo for continuous variables and categorical variables.*	To reflect current sponsor standards.
Section 10.1	*Add a sensitivity analysis on primary efficacy parameter to exclude the data from Site 054.*	Investigator of Site 054 didn't conduct in accordance with signed statement. Operational concerns to add this sensitivity analysis.
[redacted]	[redacted]	[redacted]
Section 14.0	*Update the determination of sample size.*	To reflect the discussion with FDA and blinded interim analysis.