STATISTICAL ANALYSIS PLAN

A 104-WEEK, FLEXIBLE-DOSE, OPEN-LABEL, MULTICENTER, EXTENSION STUDY TO EVALUATE THE LONG-TERM SAFETY AND EFFECTIVENESS OF LURASIDONE IN PEDIATRIC SUBJECTS

Lurasidone, SM-13496, Latuda

Study No. D1050302

Version Number: Final 2.0

Version Date: Nov 19, 2018

Authorization Signature Page

A 104-WEEK, FLEXIBLE-DOSE, OPEN-LABEL, MULTICENTER, EXTENSION STUDY TO EVALUATE THE LONG-TERM SAFETY AND EFFECTIVENESS OF LURASIDONE IN PEDIATRIC SUBJECTS

Author:

_____ _____

Name: Date
Position: Senior Director, Biostatistics
Company: Sunovion Pharmaceuticals Inc.

Approved by:

_____

Name: Date
Position: Senior Director, Clinical Development and Medical Affairs, CNS
Company: Sunovion Pharmaceuticals Inc.

_____ _____

Name: Date
Position: Executive Director, Biostatistics
Company: Sunovion Pharmaceuticals Inc.

Table of contents

Table of tables

Table 1 List of Abbreviations and Definition of Terms

ABC	Aberrant Behavior Checklist
ADHD-RS	Attention-Deficit/ Hyperactivity Disorder Rating Scale
ADI-R	Autism Diagnostic Interview - Revised
AE	Adverse Event
AIMS	Abnormal Involuntary Movement Scale
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical
BARS	Barnes Akathisia Rating Scale
β-hCG	human chorionic gonadotropin, beta chain
BMI	body mass index
CDC	Centers for Disease Control and Prevention
CDRS-R	Children's Depression Rating Scale, Revised
CGI-I	Clinical Global Impression – Improvement
CGI-S	Clinical Global Impression – Severity
CGI-BP-S	Clinical Global Impression Bipolar Version – Severity
CGAS	Clinician-rated Children's Global Assessment Scale
CGSQ	Caregiver Strain Questionnaire
CI	confidence interval
CRO	contract research organization
CSR	Clinical Study Report
CY-BOCS	Children's Yale-Brown Obsessive Compulsive Scales
DSMB	Data Safety Monitoring Board
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision
ECG	electrocardiogram
eCRF	Electronic Case Report Form
EMA	European Medicine Agency
EPS	extrapyramidal symptoms
FDA	Food and Drug Administration

Table 1 List of Abbreviations and Definition of Terms (Continued)

FSH	follicle stimulating hormone
GCP	Good Clinical Practices
HbA1c	Hemoglobin A1c
HDL	high-density lipoprotein
HR	Heart Rate
ISAC	Independent Statistical Analysis Center
ITT	intent-to-treat
IXRS	Interactive Voice Response/Web Response System
LDL	low-density lipoprotein
LH	luteinizing hormone
LLOQ	lower limit of quantification
LOCF	last observation carried forward
LPO	last patient out (LPO)
LSD	Lab Specification Document
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
NA	Not Applicable
PANSS	Positive and Negative Syndrome Scale
PARS	Pediatric Anxiety Rating Scale
PMM	Pattern mixture model
PP	per protocol
PQ-LES-Q	Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire
QTc	QT interval corrected
SAE	serious adverse event
SAS	Simpson-Angus Scale
SAP	Statistical Analysis Plan
SCHARP	Statistical Center for HIV/AIDS Research & Prevention
SD	standard deviation
SI	International System of Units

Table 1 List of Abbreviations and Definition of Terms (Continued)

SOC	System Organ Class
sNDA	Supplemental New Drug Application
TEAE	treatment-emergent adverse event
ULOQ	upper limit of quantification
WHO	World Health Organization
YMRS	Young Mania Rating Scale

1. INTRODUCTION

This statistical analysis plan (SAP) contains the definitions of analysis populations, derived variables, and statistical methods for the analyses of safety and effectiveness data from study D1050302 for subjects who continued from core studies D1050301, D1050325, and D1050326.

Study D1050302 enrolls pediatric or adolescent subjects with following 3 disease indications:

- Adolescent subjects with schizophrenia (13-17 years) from the United States and Worldwide (Study D1050301)
- Children and adolescent subjects (6-17 years) with irritability associated with autistic disorder from the United States (Study D1050325)
- Children and adolescent subjects with bipolar I depression (10-17 years) from the United States and Worldwide (Study D1050326)

Study completion status of above lurasidone pediatric core studies is given below: study D1050325 was completed and database was locked on Dec 16, 2014; study D1050301 was completed and database was locked on Feb 10, 2016; study D1050326 was completed by Oct 12, 2016 and database was locked on Nov 22, 2016.

Two interim analyses were planned and implemented. The first interim analysis (data cut date of Mar 01, 2016) was done to support the United States Supplemental New Drug Application (sNDA) for adolescent subjects with schizophrenia. The second interim analysis (data cut date of Oct 27, 2016) was conducted to support the United States sNDA for children and adolescent subjects with bipolar I depression disorder. For each of the interim analyses, separate and independent database lock was planned and separate interim SAPs were developed for study D1050302. Considering that there exists a significant difference in study completion dates for subjects with schizophrenia and subjects with bipolar depression in study D1050302, to fulfill regulatory requests from both United States Food and Drug Administration (FDA) and European Medicine Agency (EMA),final database lock are conducted in 2-batches as described below: final database lock for adolescents with schizophrenia (hereafter, referred it as final database lock 1) occurred on June 15, 2018; final database lock for subjects with autistic disorder and subjects with bipolar depression (hereafter, referred it as final database lock 2) will occur when every subject from study D1050326 completes the study D1050302, which is targeted to happen on November 21, 2018. Clinical study report (CSR) for subjects with schizophrenia has been completed based on the results per final database lock 1 for EMA submission. Full final clinical study report (CSR) for all subjects in study D1050302 will be completed after the final database lock 2 occurs.

This SAP is developed based on the study protocol with Protocol Version 4.0 (Amendment 3.0) dated 30 June, 2017 for final analysis. It outlines analyses to be conducted per final database lock 1 and final database lock 2. SAPs for the interim analyses were prepared separately.

1.1. Study Objectives

Primary Objective

The primary objective of this study is to evaluate the long-term safety, tolerability, and effectiveness of lurasidone (20, 40, 60 or 80 mg/day, flexibly dosed) in pediatric subjects who have completed a prior lurasidone study.

Secondary Objective

For all subjects:

- Proportions of subjects with adverse events (AEs), discontinuations due to AEs, and serious AEs (SAEs).

For subjects continued from Study D1050301, the following will be assessed:

- Change in Positive and Negative Syndrome Scale (PANSS) total, positive, negative, general psychopathology, and excitability subscale scores;
- Change in the Clinical Global Impression – Severity (CGI-S);
- Change in the Clinician-rated Children's Global Assessment Scale (CGAS);
- Change in the Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q).

For subjects continued from Study D1050325, the following will be assessed:

- Change in Aberrant Behavior Checklist (ABC) irritability subscale, and the following subscale scores (hyperactivity, stereotypy, inappropriate speech, and lethargy/social withdrawal);
- Change in the Clinical Global Impression – Severity (CGI-S);
- Change in Children's Yale-Brown Obsessive Compulsive Scales (CY-BOCS) modified for pervasive developmental disorders (PDDs);
- Change in the Caregiver Strain Questionnaire (CGSQ).

For subjects continued from Study D1050326, the following will be assessed:

- Change in the Children's Depression Rating Scale, Revised (CDRS-R);
- Change in the Clinical Global Impression Bipolar Version – Severity (CGI-BP-S);
- Change in the Clinician-rated Children's Global Assessment Scale (CGAS);
- Change in the Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q);
- Change in anxiety symptoms as measured by the Pediatric Anxiety Rating Scale (PARS);
- Change in attention-deficit/hyperactivity symptoms as measured by the Attention-Deficit/ Hyperactivity Disorder Rating Scale (ADHD-RS) total score.

1.2. Study Design

This is an open-label, 104-week, multicenter, extension study designed to evaluate the long-term safety, tolerability and effectiveness of flexibly dosed lurasidone (20, 40, 60 or 80 mg/day) in

pediatric subjects who have completed a 6-week treatment period in one of the preceding studies, D1050301, D1050325, or D1050326.

Subjects who complete participation in one of the preceding studies, D1050301, D1050325 or D1050326, are eligible for enrollment in this study. Informed consent/assent is obtained from all subjects (where developmentally appropriate) before any study procedures are performed. Subjects who meet entry criteria transition to this study directly from Study D1050301, D1050325 or D1050326. All subjects are treated with flexibly dosed lurasidone during the trial.

All eligible subjects are treated with lurasidone 40 mg/day for Days 1-7. Beginning with Day 8, dose adjustments (20, 40, 60 or 80 mg/day) are permitted, based on investigator judgment, to optimize tolerability and effectiveness.

Dose adjustment of study drug should occur at the regularly scheduled visits and in increments or decrements of 1 dose level. However, dose reductions for tolerability or safety purposes may occur beginning on Day 2, based on investigator judgment. These dose reductions may be between study visits and at more than 1 dose level at a time (maximum of 2 dose levels at a time). If dose reductions are required between regular study visits, the subject must return to the study site for an unscheduled visit to receive new medication kits and return all used/unused medication kits at the time of dose adjustment.

Safety and effectiveness assessments are conducted at scheduled visits during the study. A follow-up visit occurs 1 week post last dose of study drug.

A Data and Safety Monitoring Board (DSMB) reviews safety and clinical outcome data including data on AEs and SAEs at regular intervals until the D1050301, D1050325, and D1050326 studies are complete and as long as necessary for the D1050302 study, as determined by the Sponsor. The DSMB is independent of the Sponsor, contract research organization (CRO), and the investigators and is empowered to recommend stopping the study due to safety concerns, but not for efficacy or futility.

1.2.1. Determination of Sample Size

Subjects who complete the respective double-blind studies D1050301, D1050325 or D1050326, sign the consent, and meet all entry criteria are included in this study. Studies D1050301, D1050325, and D1050326 have finished, with a total of 271 subjects from D1050301, 125 subjects from D1050325, and 306 subjects from D1050326 enrolling into D1050302. Based on an estimated attrition rate of approximately 30% of subjects over six months, it is expected that at least 100 subjects are to be exposed to lurasidone for a minimum of 6 months for subjects continued from studies D1050301 and D1050325 in study D1050302.

1.2.2. Randomization and Blinding

This is an open-label study, therefore randomization and blinding is not employed.

2. ANALYSES PLANNED

2.1. General Analysis Definition

2.1.1. Logic of Final Database Lock for Adolescent Subjects with Schizophrenia

Final database lock 1 is expected to occur on June 15, 2018 for subjects who continued from study D1050301 (ie, adolescent subjects with schizophrenia). All data for these subjects will be cleaned and 'frozen' prior to the planned database lock (DBL) date and all related parties, including site personnel, clinical research organization (CRO), non-data management (DM) team of the sponsor will be removed access authority of the clinical database for adolescent subjects with schizophrenia. The CSR for adolescent subjects with schizophrenia will be based on the final database lock 1. In case there is a data error identified after the final database lock 1 for adolescent subjects with schizophrenia, depending on the nature of the issue, DM of the sponsor will re-grant the site an access permission to fix the data issue. Errata will be filed as needed and the error will be noted in the final full CSR after the final database lock 2, which will occur after all subjects in study D1050302 complete the study.

2.1.2. Definition of Open-Label Treatment Phase in Extension Study

The day of the 1st dose date in the extension study will be denoted as study Day 1. All measurements at each visit will be assigned a study day relative to this date. All data collected on/after Day 1 for subjects previously randomized to placebo group in core study or on/after the extension baseline visit for subjects previously randomized to lurasidone groups up to Week 104, including the End-of-treatment/Early Withdrawal visit, will be considered within the open-label treatment phase. That is to say, the start of the open-label treatment period is defined as Day 1 for subjects previously randomized to placebo group and the OL Baseline visit date for subjects previously randomized to lurasidone groups; the end of the open-label treatment phase is defined as the maximum of the last date of administration of study medication and the withdrawal/ completion date for subjects who have completed the study.

2.1.3. Visit Windows

While the Time and Events Schedule of the protocol indicates the visit timing and procedures for each visit, the timing of a subject's actual visit timing may not be exactly as per the protocol indicated target day/visit window (note that the protocol does not indicate required visit windows, but rather provides clear guidance for visit timing). Consequently, for the purpose of analysis, the conventions outlined below will be used to allocate the data collected at each actual visit to a planned protocol visit by defining analysis visit windows.

The reference day is study Day 1. Definition of baseline and endpoint is provided in Sections 2.1.5 - 2.1.6.

If a subject has 2 or more visits (scheduled or unscheduled) in one analysis visit window, the visit closest to the target day will be used as the study visit for that analysis visit window. The other additional visit(s) in the same analysis visit window will not be included in the by-visit

descriptive summary analyses. However, measurements from this/these other additional visit(s) will be included the subject listings and in the other safety analyses (eg, Potentially Markedly Abnormal Laboratory Values or the worst case analysis) to assure all on-treatment measurements are included. If 2 study visits occur the same number of days from the target day within the same visit window, the later one will be considered as the study visit for that target day.

Although all (scheduled and unscheduled) visits will be allocated to an analysis visit window, only planned protocol visits for each measurement will be included in the by-visit analyses. All tables and figures presenting data by visit will present only those timepoints where the applicable assessment was scheduled to be collected, unless otherwise noted below. Unscheduled and early termination data will be included for definition of endpoint or overall assessments. Data listings will present all data, regardless of visit.

Table 2 summarizes the analysis visit windows for all questionnaires, laboratory, weight, height, waist circumference, vital signs, ECG, cogstate, tanner stage, menstrual cyclicity, and other safety variables assessed at scheduled visits.

Table 2 Time Intervals for Analysis Visit Windows

Scheduled Visit (label on output in Extension study)	Target Day [(a)]	Time Interval (Day)
Day 1	1	<=1
Week 2	15	2 – 21
Week 4	29	22 – 35
Week 6	43	36 – 49
Week 8	57	50 - 70
Week 12	85	71 - 98
Week 16	113	99 - 126
Week 20	142	127 - 154
Week 24	169	155 - 182
Week 28	197	183 - 210

Table 2 Time Intervals for Analysis Visit Windows (Continued)

Scheduled Visit (label on output in Extension study)	Target Day [(a)]	Time Interval (Day)
Week 32	225	211 - 238
Week 36	253	239 - 266
Week 40	281	267 - 294
Week 44	309	295 - 322

Week 48	337	323 - 350
Week 52	365	351 - 378
Week 56	393	379 - 406
Week 60	421	407 - 434
Week 64	449	435 - 462
Week 68	477	463 - 490
Week 72	505	491 - 518
Week 76	533	519 - 546
Week 80	561	547 - 574
Week 84	589	575 - 602
Week 88	617	603 - 630
Week 92	645	631 - 658
Week 96	673	659 - 686
Week 100	701	687 - 714
Week 104	729	715 - EOT
(a): Relative to the day of the 1st dose date		

2.1.4. Pooling Strategy for Analysis of Centers

There will be no pooling of centers in this study.

2.1.5. Definition of Baseline Assessments

The following definitions will be used for effectiveness and safety assessments unless otherwise specified:

- **DB Baseline** (ie, Baseline assessment of the double-blind studies [D1050301, D1050325, or D1050326]): the last assessment made on or before the 1st dose of double-blind study medication as described in the SAPs of the core studies;
- **OL Baseline** (ie, Endpoint assessment of the double-blind studies [D1050301, D1050325, or D1050326] or Baseline assessment of the open-label study [D1050302]): the last assessment made on or before Day 1 of the extension study.

According to the protocol, measurement collected at OL Baseline is same as the value at visit 9 (ie, Week 6) (Study 301/325) or visit 8 (ie, Week 6) (study 326). For operation convenience, the protocol specified that data collected at the EOT visit of the core studies were to be recorded in core studies only and were not to be saved in the clinical database RAVE of the extension study. OL Baseline data in the Study Data Tabulation Model (SDTM) datasets of the study D1050302 will be obtained from corresponding core studies directly.

Unless otherwise specified, at each study visit of study D1050302, data derivation rule and missing imputation method for aggregated parameter that involves multiple individual items (eg, total score, subscale score, or composite score), are same as the corresponding ones in individual core study SAP. DB Baseline is defined same as the Baseline in individual core study per core study SAPs. DB Baseline used in the analysis of study D1050302 will be extracted from corresponding Analysis Data Set (ADS) of the core studies directly unless otherwise specified.

2.1.6. Definition of Endpoint Assessments for Final Analysis

- **Endpoint**: the last post-OL Baseline assessment during the open-label treatment period.
- **Week 52 LOCF/Week 52 Endpoint:** the assessment at Week 52 visit or the last post-OL baseline assessment before the Week 52 visit if the value at the Week 52 visit is missing

Week 52 LOCF will be used to refer effectiveness data only and week 52 Endpoint will be used to summarize safety data. Endpoint will be derived for both effectiveness and safety data.

2.1.7. Definition of Aggregated Analysis Timepoint

Unless otherwise specified, aggregated analysis timepoints will be derived for limited safety assessments (eg, the worst case analysis) as below.

- Any visit in open-label treatment period (extension study): defined as any assessment collected in open-label treatment period.
- Any visit in the first 52 week treatment period (extension study): defined as any assessment collected in the first 52-weeks open-label treatment period (ie, data up to Day 365).

Definition of 'Any visit in DB treatment period (core study)' (ie, as any assessment collected in DB treatment period) is same as what were defined in the core studies.

Specifically, above 3 aggregated analysis timepoints in extension study will be generated for the following worst-case analysis:

- Potentially Markedly Abnormal Laboratory Values (PMALV) (Section 2.6.5, Table 3.1)
- Potentially Markedly Abnormal Vital Signs (PMAVS) (Section 2.6.7, Table 7)
- OTc prolongation (Section 2.6.6, Table 4)
- ECG abnormality (Section 2.6.6, Table 5).

See Sections 2.6.3 (Treatment-Emergent Occurrences in OL Period**)** and 2.6.10 (C-SSRS), and 2.6.11 UKU for further details about the application of 'Any visit in DB treatment period (core study)' and 'Any visit in open-label treatment period (extension study)'.

2.1.8. Analysis Population

The analysis population for the extension study analysis will be the Safety population. The Safety population will consist of all subjects who receive at least one dose of study drug in this study. If a subject receives dispensed drug but has not returned the blister card (eg, lost follow-up), the subject will be included in the Safety population (ie, treated as taking at least one dose).

2.1.9. Analysis Group and Analysis Subgroup

A total of three analysis groups will be formed based on a subject's previous participation of double-blind core studies. Column group "301/325/326" that combines these three analysis groups will be also presented. Summary tables will be labeled as follows:

- 301-Lur: Subjects continued from Study D1050301
- 325-Lur: Subjects continued from Study D1050325
- 326-Lur: Subjects continued from Study D1050326
- 301/325/326-Lur: All subjects combined

Subjects continued from Study D1050301 will be further grouped into 3 analysis subgroups (hereafter, referred as 'Analysis Subgroup A'). Depending on the type of safety analysis in summary table they will be labeled as follows (ie, with 2 additional pooled groups: AllLur-Lur and 301-Lur):

- Pbo-Lur: previously randomized to placebo
- Lur40-Lur: previously randomized to lurasidone 40 mg/day
- Lur80-Lur: previously randomized to lurasidone 80 mg/day
- AllLur-Lur: previously randomized to lurasidone (40 mg/day or 80 mg/day)
- 301-Lur: subjects continued from study D1050301

Subjects continued from Study D1050325 will be further grouped into 3 analysis subgroups (hereafter, referred as 'Analysis Subgroup B**'**). Depending on the type of safety analysis in summary table they will be labeled as follows (ie, with 2 additional pooled groups: AllLur-Lur and 325-Lur):

- Pbo-Lur: previously randomized to placebo
- Lur20-Lur: previously randomized to lurasidone 20 mg/day
- Lur60-Lur: previously randomized to lurasidone 60 mg/day
- AllLur-Lur: previously randomized to lurasidone (20 mg/day or 60 mg/day)
- 325-Lur: subjects from study D1050325

Subjects continued from Study D1050326 will be further grouped into 2 analysis subgroups (hereafter, referred as 'Analysis Subgroup C**'**). Depending on the type of safety analysis in summary table they will be labeled as follows (ie, with one additional pooled groups: 326-Lur):

- Pbo-Lur: previously randomized to placebo
- Lur-Lur: previously randomized to lurasidone 20-80 mg/day
- 326-Lur: subjects from study D1050326

Subjects may be further grouped into 4 analysis subgroups based on the age at OL baseline (years) as needed (hereafter, referred as 'Analysis Subgroup D'). Depending on the type of safety analysis in summary table they will be presented as follows follow (ie, with 2 additional pooled groups: 6-12 years old and ≥ 13 years old):

- 6-9 years old
- 10-12 years old
- 6-12 years old (ie, combined group of '6-9 years old',' 10-12 years old')
- 13-15 years old
- ≥ 16 years old
- ≥ 13 years old (ie, combined group of '13-15 years old',' ≥ 16 years old')

To fulfill EMA's request, for subjects continued from study D1050301, selected safety summary will be also presented by following age subgroup as needed:

- 13 to 14 years old
- ≥ 15 years old

Summary tables, wherever applicable, will be presented for all subjects in the study by analysis group and by analysis subgroup as needed.

2.1.10. Average Dose and Predominant Dose

The predominant dose (ie, modal dose, used interchangeably hereafter in this document) in the open-label treatment period is defined as the lurasidone dose to which the subject was exposed for the greatest. If a subject has 2 or more doses with the great duration, the one with higher dosage will be selected as predominant treatment.

The average total daily dose during the open-label treatment phase will be calculated as the sum of the total daily dose divided the treatment duration (see Section 2.2.2.4). The average total daily dose during a visit period is defined in a similar way.

The predominant dose and average total daily dose in the first 52-weeks period are defined in a similar way; specifically, for a subject, the treatment duration in the first 52-weeks period (in days) is defined to be same as the treatment duration in open-label treatment period if the subject was early terminated by Week 52 visit; otherwise, it is defined as the amount of elapsed time between the first day that study medication was taken and the Week 52 visit date (inclusive).

For any study visit , if medication kit usage information is incomplete (ie, meet one of following conditions: any dispensed medication kit at the visit has missing number of tables returned, or

has missing dosing start or ending date), average total daily dose at the visit will not be derived and will be set as missing. For an assessment period that involves multiple study visits (eg, the open-label treatment period), if at some visit(s) the drug usage information is incomplete, then the other visits with complete information of kit usage will be used to derive the average daily dose and predominant dose accordingly.

2.1.11. Presenting 'Endpoint' in Outputs for Final Analysis

For safety summary, data will be summarized by visit, Week 52 Endpoint, and 'Endpoint' . For effectiveness summary, data will be presented by study visit, 'Week 52 LOCF', and 'Endpoint'.

2.1.12. Age over Time

The following ages will be derived: age (years) at DB Screening, age (years) at OL Baseline, age (years) at Endpoint.

2.1.13. Subgroup Analysis

For a limited number of descriptive analyses, data will be presented for analysis group and analysis subgroup C by subgroup as below as needed:

- Age group at OL Baseline (years): "6-9 years", "10-12 years", "6-12 years", "13-15 years", "≥ 16 years", "≥ 13 years".
- Country: US, Non-US.

For analysis subgroup C, age subgroups to be displayed in output will be "10-12 years", "13-15 years", "≥ 16 years", "≥ 13 years".

2.1.14. Data Imputation Rules

2.1.14.1. Composite Score

Imputation rule of each assessment in this extension study will be consistent with what was specified in core studies. That is to say, unless otherwise specified, any individual missing item in any scale will not be imputed. When calculating a total score, subscale score, composite score, or any assessment with more than one item (e.g., CDRS-R total score, ABC subscale scores, CY-BOCS, CGSQ , PANSS total score, PQ-LES-Q, BARS, SAS, AIMS, etc.), if one or more items are missing at a visit, then the associated score (ie, total score or subscale score) will be set to missing. Same as the core studies, the only exceptions are the standardized composite score of the CogState Computerized Cognitive Test Battery, and the category scores of UKU. See

Section 2.6.9 for missing imputation rule of Cogstate standardized composite score and Section 2.6.11 for missing imputation rule of the UKU category scores and UKU total score.

Endpoint

In the event of missing data at Endpoint, the last post-OL baseline observation carried forward approach prior to Endpoint will be used to impute missing values to obtain Endpoint;

Week 52 LOCF/Week 52 Endpoint

In the event of missing data at Week 52 (per visit window in Section 2.1.3), the last post-OL baseline observation carried forward approach prior to Week 52 will be used to impute missing values.

OL Baseline Assessment

If a value at OL Baseline visit (ie, Week 6 visit in core study) is missing, no imputation will be made.

2.1.14.2. Incomplete/Missing Dates

2.1.14.2.1. Start/End Date for Adverse Event

Treatment-emergent adverse events (AEs) are those events defined in Section 2.6.4. A conservative approach will be used to handle the missing dates for AE onset date and end date to identify the treatment-emergent AE (ie, if the available incomplete date cannot determine treatment emergent status, the AE will be considered as a treatment-emergent AE). No imputation will be done to calculate AE duration.

2.1.14.2.2. Start/End Date for Prior/Concomitant Medication

Partial dates of non-study medication start/end date need to be imputed in order to classify the medications as prior or concomitant medications. A conservative approach will be used to handle the incomplete date(s) for medication start/end date (ie, if the available incomplete date(s) cannot determine the concomitant relationship relative to the open-label treatment period, the medication will be considered as a concomitant medication).

2.1.14.2.3. Start/End Date of Psychiatric Hospitalization

The duration (days) of the hospitalization during the open-label treatment period is calculated as discharge date – hospitalization admission date +1. If the admission date and the discharge date are completely or partially missing, no imputation will be done. If admission date is prior to the start of open-label treatment period, it will be imputed by the start date of open-label treatment period; if discharge date is after the last day of OL treatment phase, it will be imputed by the last day of the OL treatment period. See Section 2.1.2 for definition of start/end of open-label treatment period.

2.1.14.2.4. First Menses Date

A conservative approach will be used to handle the incomplete date of the first menses date. The partial first menses date will be imputed as follows: (1) If year and month are known, and if the year or month is previous to Screening, use the 15th of the month; otherwise, if the month is the month of Screening, use the 1st of the month. (2) If only year is known, and it is previous to the year of Screening, use June 30th of that year; otherwise, if the year is the year of Screening, use the Jan 01 of the year. No imputation will be done if the first menses date is completely missing.

2.1.14.2.5. Adverse Event Grade and Relationship to Study Medication

The Adverse Event with missing severity will be classified as 'unspecified'; similarly, the AE with missing causality to study medications will be classified as 'unspecified' as well.

2.1.14.2.6. Lab Data

Where applicable, the upper limit of quantification (ULOQ) will be substituted for chemistry, hematology, urinalysis, and urine microscopic values that are greater than the ULOQ. Similarly, if the value is below the lower limit of quantification (LLOQ), then the LLOQ will be substituted.

2.1.14.3. BMI

At each visit, BMI will be derived only when both height and weight are available at the visit and no imputation will be made if one or two of them are missing.

2.1.14.4. Z-score and Percentile of Height, Weight, and BMI per WHO and CDC Growth Reference

Age-and-sex specific z-scores and percentiles in height and BMI will be derived at each assessed visit using WHO 2007 growth reference for each subject. Considering that no weight related growth reference is available per WHO standard, for each subject (US and non-US), CDC 2000 growth chart for weight will be used to derive age-and-sex specific z-scores and percentiles in weight. In addition, for US subjects, age-and-sex specific z-scores and percentiles in height and BMI will be also derived using CDC 2000 growth chart.

For height and BMI, for summary of z-score and percentile of analysis group or analysis subgroups A and C (ie, studies D1050301 and D1050326), WHO based ones will be reported; for summary of analysis subgroup B (ie, study D1050325) or US only population, the ones per CDC growth chart will be displayed.

For weight, for summary of z-score and percentile of analysis group or analysis subgroup, CDC based ones will be reported consistently.

2.2. Methods of Analysis

2.2.1. Statistical Hypotheses for Trial Objectives

Because of the nature of an open-labeled study, no hypotheses are planned and to be tested.

2.2.2. Subject Information

Unless otherwise specified, subject information will be presented by analysis group and analysis subgroup as needed (see Section 2.1.9) for the safety population.

2.2.2.1. Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized by analysis group and analysis subgroup for the safety population.

Baseline descriptive statistics (N, mean, SD, median, and range) of demographic characteristics will be provided by analysis group and analysis subgroup for age at OL baseline (in years), body weight, height, body mass index (BMI, calculated as weight(kg)/ [height(m)^2]), and waist circumference. At OL Baseline, age-and-sex specific z-scores and percentiles in height and BMI will be obtained using WHO 2007 growth reference (see Attachment 4.1); corresponding z-score and percentile in weight will be obtained using CDC 2000 growth chart. In addition, age-and-sex specific z-scores and percentiles in height weight and BMI at OL Baseline will be reported using CDC 2000 growth chart for US subjects as well.

The number and percentage of subjects in each of the following categories (and sub-categories) at OL baseline will also be summarized by analysis group and by analysis subgroup:

- Gender: Male, Female;
- Race: white, black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Other;
- Ethnicity: Hispanic or Latino, Not Hispanic or Latino;
- Age at OL Baseline group 1: ≤12 years, 13-17, 18-20 years;
- Age at OL Baseline group 2: ≤9 years, 10-12 years, 13-15 years, 16-17, 18-20 years;
- OL Baseline category for height (WHO): <5th, 5th to <95th, >95th percentile;
- OL Baseline category for weight (CDC): <5th, 5th to <95th, >95th percentile;
- OL Baseline BMI category (WHO): <3th, 3th to 85th, >85th to 97th , >97th percentile;
- Country: US, non-US;
- Region: North America, South America, Europe, and Asia;

The number and percentage of subjects in each of the following categories (and sub-categories) at OL baseline will also be summarized by analysis subgroup for subjects from D1050325 or overall US subjects from D1050301, D1050325, and D1050326 as well:

- OL Baseline category for height (CDC): <5th, 5th to <95th, >95th percentile;
- OL Baseline category for weight (CDC): <5th, 5th to <95th, >95th percentile;
- OL Baseline BMI category (CDC): <5th, 5th to 85th, >85th to 95th , >95th percentile;

Baseline summary statistics (N, mean, SD, median, and range) of disease characteristics will be provided by analysis subgroup for subjects continued from Study D1050301

- DB Baseline and OL Baseline of PANSS total score
- DB Baseline and OL Baseline of CGI-S score

Baseline summary statistics (N, mean, SD, median, and range) of disease characteristics will be provided by analysis subgroup for subjects continued from Study D1050326

- DB Baseline and OL Baseline of CDRS-R total score
- DB Baseline and OL Baseline of CGI-BP-S score

2.2.2.2. Disposition Information

Subject disposition for all enrolled subjects will be summarized by analysis group and analysis subgroup using the following categories:

- Subjects who entered the extension study;
- Subjects who entered the extension study, but not dosed;
- Subjects in the Safety population;
- Subjects in the Safety who completed the study;
- Subjects who discontinued.

Subject disposition for one year study (ie, at Week 52) will be also presented using a similar way for the above categories with one more category for number and percentage of subjects who were still ongoing at 1-year cut.

For subjects who are consented and discontinue from the study, the corresponding reasons for discontinuation will be summarized. Kaplan-Meier plot of the time to early discontinuation will be presented by analysis group.

2.2.2.3. Treatment Compliance

At each post OL-baseline (Weeks 2 – 104) visit, prior to dispensing study medication, previously dispensed study medication will be retrieved and assessed by tablet count.

Compliance in the open-label treatment phase will be calculated for each post-OL baseline visit, the first 52-weeks treatment period, and overall as follows:

- Percent compliance = (number of tablets taken / number of tablets should have taken) x 100.
- Number of tablets taken = number of tablets dispensed – number of tablets returned.
- Number of tablets should have taken = (number of tablets supposed to take in a day) x (number of exposure days).
- Number of exposure days = last dose date – first dose date + 1.

For subjects assigned to the 20 mg per day or 40 mg per day dose, number of tablets supposed to be taken in a day is 1 tablet; for subjects assigned to the 60 mg per day or 80 mg per day, number of tablets supposed to be taken in a day is 2 tablets. If the number of tablets returned is missing, compliance will be set as missing.

For any study visit, if medication kit usage information is incomplete (ie, meet one of following conditions: any dispensed medication kit at the visit has missing number of tables returned, or has missing dosing start or ending date), compliance rate at the visit will not be derived and will be set as missing. For an assessment period that involves multiple study visits (eg, the open-label treatment period), if at some visit(s) the drug usage information is incomplete, then the other visits with complete information of kit usage will be used to derive the compliance rate.

Non-compliance is defined as less than 75% or more than 125% non-missing compliance with the study medication. Subjects with missing compliance are not classified as non-compliant.

Descriptive summary statistics (N, mean, median, and range) will be provided by analysis group and by analysis subgroup, respectively, for percent compliance by study visit. In addition, the number and percentage of subjects in the categories of the following variables will be also summarized by study visit:

- Compliant;
- Non-compliant (<75% or >125%);
- Non-compliant: <75%;
- Non-complaint: >125%;
- Missing.

Incomplete medication kit usage information will be summarized for overall OL treatment period by analysis group as follows: incomplete kit usage for 1 visit period, 2 visit periods, 3 visit periods, ≥ 4 visit periods.

2.2.2.4. Extent of Exposure

Treatment duration is defined as the amount of elapsed time between the first and the last day that study medication was taken (inclusive). It will be calculated (in days) in terms of the

difference in relative study days between the last and first dose of study medication, plus one day. If the end date of the medication intake is unknown for subjects who either completed or early terminated the study (eg, subject is lost to follow-up), it will be imputed as the earlier of the disposition date (ie, date of the last contact with subject) and 14 days from the date that the last medication kit is dispensed if the last contact visit is either V1E (OL baseline), V2E (W2), V3E (W4) , or V4E (W6), and 28 days from the date that the last medication kit is dispensed is V5E (W8) or after.

Descriptive statistics (N, mean, standard deviation, median, and range) will be presented by analysis group and analysis subgroup for treatment duration. Number of subjects with duration in each of following categories (days): ≥ 14, ≥ 28, ≥ 56, ≥ 84, ≥ 112, ≥ 140, ≥ 168, ≥ 196, ≥ 252, ≥ 308, and ≥ 364; 1-13, 14-27, 28-55, 56-83, 84-111, 112-139,140-167,168-195, 196-251, 252-307, 308-363, ≥ 364 days will also be presented. Return of the blister card at any given visit is not a prerequisite for inclusion in exposure summary statistics if corresponding dose start and end dates are reported.

2.2.2.5. Study Medication and Prior and Concomitant Medications

2.2.2.5.1. Study Medication

Descriptive summary statistics (N, mean, median) for the mean daily dose will be provided by analysis group and analysis subgroup, respectively, at each post-baseline visit, in the first 52-weeks treatment period, and overall treatment period. In addition, frequency distribution of the following dose information will be summarized:

- dispensed dose by study visit.
- predominant dose in the first 52-weeks treatment period, and during the open-label treatment period, respectively.

Return of the blister card at any given visit is not a prerequisite for inclusion in average daily dose summary statistics or predominant dose summary statistics if corresponding dose start and end dates and amount are reported.

2.2.2.5.2. Concomitant Medications

Concomitant medications during the open-label treatment period (see Section 2.1.2) (ie, medications taken during the OL treatment period) are coded using World Health Organization (WHO) Drug Dictionary and Anatomical Therapeutic Chemical (ATC) codes.

Number and percentage of subjects taking concomitant medications will be provided by level 3 ATC classification and preferred name by analysis group. Since medications are coded to ATC classification by indication, preferred names may appear under multiple ATC classifications. Medications started and stopped before the open-label treatment period as well as the medications started after last dose of study drug will be included in the prior and concomitant medication listing only.

For subjects continued from Study D1050326

The number and percentage of subjects taking the following categories of concomitant medications will be provided:

- With Concomitant Mood Stabilizer;
- With Concomitant Antidepressant;
- With Concomitant Mood Stabilizer or Antidepressant;
- With Concomitant treatment of ADHD;
- With Concomitant treatment of ADHD continuing from core study

Mood stabilizers, antidepressants, and treatment of ADHD will be determined by medical review prior to sign-off of the SAP.

2.2.2.5.3. Concomitant Use of Benzodiazepine

The number and percentage of subjects, who received benzodiazepine medications during the open-label treatment phase, will be summarized by preferred term and analysis group, respectively.

For subjects continued from Studies D1050301 and D1050326

A conversion chart for standardizing selected benzodiazepines to an equivalent dosage of Lorazepam is provided in Attachment 4.4. Specifically, for each subject, a weight value corresponding to the number of days of benzodiazepine use in a given study visit will be used for calculating the mean daily dose for each study visit; the average total daily dose during the open-label treatment phase will be calculated as the sum of the total daily dose divided by the total days of the benzodiazepine medication administration. The usages of each benzodiazepine medication, including average total daily dose and equivalent dosage to Lorazepam over time, will be provided in a listing.

2.2.2.6. Psychiatric History and Medical History Prior to DB Phase of Core Study

Psychiatric history and medical history prior to DB phase of core study will be summarized in the same way as in the core study.

2.3. Efficacy

2.3.1. Efficacy Analysis Specifications

2.3.1.1. Definition

The effectiveness endpoints in this study are:

For subjects continued from Study D1050301

- Change in PANSS total, positive, negative, general psychopathology, and excitability subscale scores
- Change in the CGI-S
- Change in the PQ-LES-Q percentage maximum possible score
- Change in the CGAS total score

For subjects continued from Study D1050325

- Change in the ABC irritability subscale score and the other subscale scores (hyperactivity, stereotypy, inappropriate speech, and lethargy/social withdrawal)
- Change in the CGI-S
- Change in the CY-BOCS modified for pervasive developmental disorders (PDDs)
- Change in the CGSQ

For subjects continued from Study D1050326

- Change in the CDRS-R total score
- Change in CGI-BP-S score (depression, mania, overall bipolar illness)
- Proportion of subjects who meet symptom remission, defined as CDRS-R Total Score ≤ 28 and YMRS total score ≤ 8 and CGI-BP-S depression score ≤ 3 at study visit
- Change in the PARS total score
- Change in the PQ-LES-Q score
- Change in the CGAS score
- Change in the ADHD-RS total score

2.3.1.1.1. Positive and Negative Syndrome Scale PANSS

The PANSS is only applied to subjects who continued from Study D1050301.

The PANSS is comprised of 30 items and three scales:

- the Positive scale contains seven questions to assess delusions, conceptual disorganization, hallucinations behavior, excitement, grandiosity, suspiciousness /persecution, and hostility; Positive subscale (range 7-49) is calculated as sum of Items P1 to P7 in the positive subscale;
- the Negative scale contains seven questions to assess blunted effect, emotional withdrawal, poor rapport, passive/apathetic social withdrawal, lack of motivation, and

similar symptoms; Negative subscale (range 7-49) is calculated as sum of Items N1 to N7 in the negative subscale;

- the General Psychopathology subscale addresses other symptoms such as anxiety, somatic concern, and disorientation; General psychopathology subscale (range 16-112) is calculated as sum of Items G1 to G16 in the general psychopathology subscale.

An anchored Likert scale from 1-7, where values of 2 and above indicate the presence of progressively more severe symptoms, is used to score each item. Individual items are then summed to determine scores for the three scales, as well as a total score. As a result, the total score may range from 30 to 210, positive and negative sub-scores may range from 7 to 49, and the general psychopathology sub-scores may range from 16 to 112. In addition, a subscale of excitability will be derived, which consists of the following four items from the PANSS: excitement, hostility, uncooperativeness, and poor impulse control. The sum of the four items ranges from 4 to 28.

In general, higher values of PANSS scores represent greater severity of illness. If one or more items are missing, no imputation will be performed and the total score and the scores of the subscales that include these items will be left missing.

2.3.1.1.2. Clinical Global Impression – Severity

The CGI-S is applied to subjects who continued from Study D1050301 and Study D1050325. The CGI-S is a single value, clinician-rated assessment of illness severity, and 7-point scale with range from 1='Normal, not at all ill' to 7='Among the most extremely ill patients'. A higher score is associated with greater illness severity.

2.3.1.1.3. Quality of Life Enjoyment and Satisfaction Questionnaire (Pediatric Q-LES-Q)

The Pediatric Q-LES-Q is applied to subjects who continued from Study D1050301 and Study D1050326.

The Pediatric Q-LES-Q is a 15-item self-report measure of the degree of enjoyment and satisfaction in various areas of daily living, based on the content of the Short From of the Q-LES-Q. Each item is rated on a 5-point scale, ranging from 1 (very poor) to 5 (very good). The first 14 items are the same as the General Activities section of the regular Q-LES-Q form and are used to compute the raw score. The PQ-LES-Q-SF percentage maximum possible score is calculated as follows:

% Max = 100 × (Raw Score – Minimum Score) / (Maximum Score – Minimum Score),

where the Minimum Score equals 14 and the Maximum Score equals 70, and the % maximum possible score can range from 0% to 100%. Higher scores indicate better quality of life.

2.3.1.1.4. Clinician- rated Children's Global Assessment Scale (CGAS)

The CGAS is applied to subjects who continued from Study D1050301and Study D1050326.

The Children's Global Assessment Scale (CGAS) is a numeric scale (1 through 100) used by mental health clinicians to rate the general functioning of children under the age of 18, where 1 represents the most impaired functioning and 100, superior functioning. Each decile (e.g., 1–10, 11–20) has a descriptive header (e.g., "Moderate impairment in functioning in most domains") and examples of behaviors and types of environmental accommodations that might be seen at that level of functioning. Scores above 70 on the CGAS indicate functioning within the range of typically developing children of the same age as the child being rated while scores below 60 indicate a definite clinical case.

2.3.1.1.5. Children's Depression Rating Scale, Revised (CDRS-R)

The CDRS-R is only applied to subjects who continued from Study D1050326.

The CDRS-R is a semi-structured, clinician-rated instrument designed for use with children and adolescents between the ages of 6-17 years. It contains 17 ordinally-scaled items that evaluate the presence and severity of symptoms commonly associated with depression in childhood. The CDRS-R is administered separately to the patient and to the caregiver; among 17 items, 14 items are based on separate interviews with child and parent, 3 items are based solely on the rater's observation of child (ie, no questions). The 14 items (see below) are rated on a 1 (no psychopathology) to 7 (most psychopathology) scale, where a rating of 3 represents mild psychopathology. The 3 items (sleep disturbance, appetite disturbance, listless speech) are rated on a 1 (no pathology) to 5 (most pathology) scale.

1) Impaired Schoolwork
2) Difficulty Having Fun
3) Social Withdrawal
4) Sleep Disturbance
5) Appetite Disturbance
6) Excessive Fatigue
7) Physical Complaints
8) Irritability
9) Excessive Guilt
10) Low Self-Esteem
11) Depressed Feelings
12) Morbid Ideation
13) Suicidal Ideation
14) Excessive Weeping

The CDRS-R total score ranges from 17-113. In general, higher values of CDRS-R total score represent greater severity of illness

2.3.1.1.6. Clinical Global Impression-Bipolar Version, Severity of Illness (CGI-BP-S)

The CGI-BP-S is only applied to subjects who continued from Study D1050326.

The CGI-BP-S is a three-question clinician-rated assessment of the subject's current illness state (depression, mania, and overall) using a 7-point scale (1(normal, not ill) to 7 (very severely ill)) for each question, where a higher score is associated with greater illness severity.

2.3.1.1.7. Pediatric Anxiety Rating Scale (PARS)

The PARS is only applied to subjects who continued from Study D1050326.

The PARS is a clinician-rated instrument for assessing over time the severity of anxiety symptoms associated with common DSM-IV anxiety disorders (generalized anxiety disorder, separation anxiety, and social phobia) in children ages 6-17 years. The PARS is administered separately to the subject and to the caregiver. The instrument has 2 sections. The first section includes a 50-item symptom checklist, which the clinician rates as present or absent during the past week. The second section is comprised of 7 severity impairment items reflecting the severity/impairment of all symptoms endorsed in Section 1 of the PARS (during the past week). Each question is answered on a 0-5 Likert scale (0 for none, and 1-5 for minimal to extreme) with alternative responses of 8=Not Applicable and 9=Does Not Know. Scores of 8 or 9 are not counted in the summation as per the PARS instructions. The PARS total score over all 7 questions ranges in value from 0 to 35.

2.3.1.1.8. Attention-Deficit/ Hyperactivity Disorder Rating Scale (ADHD-RS)

The ADHD-RS IV is only applied to subjects who continued from Study D1050326.

The ADHD-RS IV is a validated scale that measures the behaviors of children with ADHD. The ADHD-RS IV consists of 18 items reflecting current symptomatology of ADHD based on DSM-IV-TR criteria. Each item is scored from a range of 0 (no symptoms) to 3 (severe symptoms) with total scores ranging from 0 to 54. The 18 items may be grouped into two sub-scales: hyperactivity/impulsivity (even number items 2 through 18) and inattentiveness (odd number items 1 through 17), ranging from 0 to 27.

2.3.1.2. Analysis Methods

No inferential statistics on effectiveness will be presented. For all of the continuous effectiveness assessments, descriptive summary statistics (N, mean, standard deviation, median, range, and a 95% confidence interval) will be presented by analysis subgroup, at DB baseline, OL baseline, each of post-OL visit, Week 52 LOCF, and the endpoint in the open-label study. All analyses for the scheduled assessments will be based on observed cases (OC). In addition, for all effectiveness evaluations, changes from DB baseline and OL baseline will be presented in a similar way using summary statistics, as described above.

For subjects continued from study D1050301

At a visit in extension study, responders relative to DB Baseline are defined as those who show a 20% or more reduction (i.e., improvement) from baseline in the PANSS total score from DB Baseline. The percent change of PANSS total score relative to DB Baseline is defined as [100*Change from DB Baseline/(DB Baseline – 30)] since 30 is the lowest possible value for PANSS. At a post-OL Baseline visit, responders relative to OL Baseline are defined as those

who show a 20% or more reduction (i.e., improvement) from OL Baseline. The percent change of PANSS total score relative to OL Baseline is defined as [100*Change from OL Baseline/(OL Baseline – 30)].

For subjects continued from study D1050325

At a visit in extension study, responders relative to DB Baseline are defined as those who show a 25% or more reduction (i.e., improvement) from baseline in the ABC irritability subscale score from DB Baseline. The percent change of ABC irritability subscale score relative to DB Baseline is defined as [100*Change from DB Baseline/(DB Baseline)]. At a post-OL Baseline visit, responders relative to OL Baseline are defined as those who show a 25% or more reduction (i.e., improvement) from OL Baseline. The percent change of ABC irritability score relative to OL Baseline is defined as [100*Change from OL Baseline/(OL Baseline)].

For subjects continued from study D1050326

At a visit in extension study, responders relative to DB Baseline are defined as those who show a 50% or more reduction (i.e., improvement) from DB Baseline in CDRS-R total score. The percent change of CDRS-R total score relative to DB Baseline is defined as [100*Change from DB Baseline/(DB Baseline – 17)] since 17 is the lowest possible value for CDRS-R total score; At a post-OL Baseline visit, responders relative to OL Baseline are defined as those who show a 50% or more reduction (i.e., improvement) from OL Baseline. The percent change of CDRS-R total score relative to OL Baseline is defined as [100*Change from OL Baseline/(OL Baseline – 17)].

Frequencies and percentages of CDRS-R responders relative to both DB Baseline and OL Baseline over time will be reported by analysis subgroup, respectively. Frequencies and proportions of remitters as well as the corresponding 95% Clopper-Pearson (ie, exact) CI of the proportion at DB baseline, OL baseline, each of post-OL visit, Week 52 LOCF, and endpoint will be reported by analysis subgroup.

Subgroup analyses for CDRS-R total score and CGI-BP-S depression score will be conducted on observed values and changes from DB baseline and OL baseline by study visit, W52 LOCF, and endpoint by analysis subgroup, to examine the effects of geographic region (US, non-US) in a similar approach as described above.

2.4. Pharmacokinetics (PK) Analysis

Not applicable.

2.5. Multiplicity Adjustment

Not applicable.

2.6. Safety

2.6.1. Analysis Specifications

All safety analyses and summaries will be based on the safety analysis population. Subgroup analysis per age group may be conducted as needed.

There will be no imputation of missing values for clinical laboratory test results, vital sign measurements, and ECG evaluations in the by-visit analyses.

2.6.2. Age Over Time

Descriptive statistics (N, mean, SD, median, and range) of age will be provided by analysis group at DB screening, OL baseline, and endpoint by analysis group. Shifts from DB screening to endpoint and from OL Baseline to endpoint in age group (years) below will be produced by analysis group to show the percentage of subjects within the age categories:

- Age group 2: 6-9 years, 10-12 years, 13-15 years, 16-17, 18-20 years.

2.6.3. Definition of Treatment-Emergent Occurrences in Open-Label Extension Study

Definition of treatment-emergent adverse event is provided in Section 2.6.4. Treatment-emergent occurrences of adverse reactions, including Potentially Markedly Abnormal Laboratory Values (PMALV), Potentially Markedly Abnormal Vital Signs (PMAVS), and abnormal ECG values, are defined as the assessment values that meet specific safety criteria within the open-label treatment phase (see Section 2.1.2).

2.6.4. Adverse Events

A treatment-emergent adverse event (TEAE) is defined as an AE onset on or after the start of open-label treatment period through 7 days after study drug discontinuation (14 days for serious adverse events and deaths). For sake of simplicity, hereafter TEAE is referred as AE in the section.

The overall incidence (ie, number and percent of subjects with one or more AE in each category) of AEs, serious AEs, deaths, AEs leading to discontinuation, drug-related AEs, serious drug-related AEs, drug-related AEs leading to discontinuation, serious AEs leading to discontinuation, and serious drug-related AEs will be summarized by analysis group and by analysis subgroup, separately.

AEs by system organ class (SOC) (number and proportion of subjects with one or more AEs within a system organ class) will be summarized by analysis group. For each AE, the percentage of subjects who experienced at least one occurrence of the given event will be provided by preferred term, grouped by system organ class, and presented by analysis group and by analysis subgroup. AEs for age subgroup summary will be also presented by SOC and preferred term and by analysis group and analysis subgroup.

Furthermore, additional summaries will be created based on the following combined preferred terms:

- Dystonia - preferred terms: dystonia, trismus, oculogyric crisis, oromandibular dystonia, tongue spasm, torticollis
- Parkinsonism - preferred terms: bradykinesia, drooling, extrapyramidal disorder, glabellar reflex abnormal, hypokinesia, parkinsonism, psychomotor retardation
- Somnolence - preferred terms: hypersomnia, sedation, somnolence, hypersomnolence

According to MedDRA coding guidelines for Lurasidone pediatric program,

- EPS-related reported terms for Tremor, Rigidity or Muscle Rigidity will be queried for further specificity. Sites will qualify if the reported term is Parkinsonian or non-Parkinsonian in nature; if confirmed, the reported term will be coded to lowest level terms of Parkinsonism; if not, the term will be coded to specific symptom.
- For reported term Restlessness, sites will qualify if the reported term is Akathisia or non-Akathisia in nature; if confirmed, the reported term will be coded to lowest level terms of Akathisia; if not, the term will be coded to specific symptom.

In clinical database of this study, AEs with preferred term of Tremor, Muscle Rigidity, and Cogwheel Rigidity should be read as non-Parkinsonian; AEs with preferred term of Restlessness should be read as non-Akathisia. Due to above reasons, preferred terms Tremor, Muscle Rigidity, and Cogwheel Rigidity are not included in the list for combined preferred terms of 'Parkinsonism'. In addition, these three preferred terms plus term Restlessness are not included in the list for extra-pyramidal symptoms (EPS)-related AEs in Attachment 4.5.1.

Summaries of AEs by preferred term only by analysis group for both combined and uncombined terms will be generated as well, respectively.

In addition, the incidence of AEs (by preferred term, grouped by SOC, and presented by analysis group and by analysis subgroup) will also be summarized by severity, by the relationship to study medication, by the action taken regarding the study medication, as well as by the outcome. The relationship to study medication will be presented by 2 categories: related (which includes possibly related, probably related, and related, as determined by investigators), and not related (which includes unrelated and unlikely to be related, as determined by investigators). Listings will also be generated for deaths, SAEs, and discontinuations due to AEs. All adverse events regardless of treatment emergent status will be provided in the AE data listing.

AEs with Special Interest

A list of preferred terms that are to be combined for the assessment of each of the pre-specified adverse events (including Extrapyramidal Symptoms (EPS)-related AEs, metabolic-related AEs, hypersensitivity-related AEs, and suicidality and self-injury related AEs) is provided in Attachments 4.5.1 - 4.5.4. These AE preferred terms are determined and verified by medical review of coded AE terms during data review meeting and finalized prior to database lock. The incidence and exposure adjusted incidence rates of EPS-related AEs, metabolic-related AEs,

hypersensitivity-related AEs, and suicidality and self-injury related AEs will be summarized by analysis group and by analysis subgroup, respectively. In addition, EPS-related AEs will be summarized for akathisia and non-akathisia event as well.

AEs by Exposure

The TEAEs will be summarized by exposure to study drug using time intervals beginning with the earliest concurrent exposure to study drug for following AE summary as needed. Denominators for exposure intervals will be based on the number of subjects who were exposed as of the first day of the interval; for cases where events meet the treatment-emergent definition but start after the last dose, the exposure interval for the event will be considered the interval of the last day of exposure to study medication.

- AEs by system organ class (SOC) and preferred term, Serious AEs by SOC and preferred term, and AEs leading to discontinuation by SOC and preferred term (sorted alphabetically by SOC and preferred term within SOC).
- AEs by preferred term, AEs with special interest by preferred term (sorted by descending order per frequency count)

Exposure category to be assessed is given below:

- Exposure categories : 1-13 days, 14-27 days, 28-55 days, 56-83 days, 84-167 days, 168-363, and ≥ 364 days.

Summary of AEs Onset after the Last Study Dose

For subjects who either complete or are early terminated the study, incidence of hyperactivity-related AEs will be summarized by analysis group. Attachment 4.5.5 provides the definition of hyperactivity-related AEs per preferred term and lower level term. These terms were originally defined in Appendix 2 of SN-0434 and were communicated with the Psychiatry Division of FDA on Oct 24, 2014.

Subjects continued from Study D1050326

Treatment-Emergent Mania:

Treatment-emergent mania is defined as a YMRS score of ≥ 20 on any 2 consecutive OL visits (ie, including OL baseline visit for subjects who were randomized to lurasidone group in the core study or excluding the visit(s) with study day ≤ Day 1 for subjects randomized to placebo group) or at the final assessment in OL treatment period, or any TEAE of mania or hypomania.

Frequency and percentage of subjects experiencing treatment-emergent mania in the first 52-week treatment period and during the overall open-label treatment period will be presented by analysis subgroup, respectively.

In addition, frequency and percentage of subjects of following categories will be presented for the first 52-week treatment period and the overall open-label treatment period, accordingly.

- YMRS score of ≥ 20 on any 2 consecutive OL visits or at the final assessment
- YMRS score of ≥ 16 on any 2 consecutive OL visits or at the final assessment
- YMRS score of ≥ 15 on any 2 consecutive OL visits or at the final assessment
- Any TEAE of mania or hypomania
- Any TEAE of mania or hypomania or YMRS score of ≥ 16 on any 2 consecutive OL visits or at the final assessment
- Any TEAE of mania or hypomania or YMRS score of ≥ 15 on any 2 consecutive OL visits or at the final assessment

2.6.5. Laboratory Measurements

Laboratory data will be summarized for each laboratory test listed in Attachment 4.3.

Descriptive statistics (N, mean, standard deviation, median, and range) will be reported by analysis group and analysis subgroup for each continuous laboratory analyte at the DB baseline, OL baseline, each post-OL baseline visit, W52 endpoint, and endpoint for absolute value as well as changes from DB baseline and OL baseline, respectively. Categorical results (eg, urinalysis tests) will be summarized at DB baseline, OL baseline, each post-OL baseline visit, W52 endpoint, and endpoint by analysis group using frequency and percentage. All lab values will be presented in the data listings.

Per a notice from lab vendor of the study, the test agent or assay for Estradiol was switched from Estradiol II to Estradiol III during the mid-course of the study. Lab Specification Document (LSD) was updated on Sep 01, 2015 to add normal reference range for Estradiol III. Due to this change, Estradiol data are presented under 2 test names (ie,'Estradiol II' and 'Estradiol III') in the raw lab dataset. It should be noted that for three completed core studies (D1050325, D1050301, D1050326), only Estradiol II was used and reported in the raw lab dataset; 'Estradiol II' in raw data was mapped as 'Estradiol' in the lab analysis dataset of the core studies. In the lab analysis dataset of study D1050302, all data under test name of 'Estradiol', 'Estradiol II', or 'Estradiol III' will be mapped to 'Estradiol (overall)'; Estradiol results under different test agents will be kept as is (ie, mapped as 'Estradiol II', and 'Estradiol III', respectively). In summary output, results for 'Estradiol (overall)', 'Estradiol II', and 'Estradiol III' will be presented, separately.

Prolactin values will be summarized by analysis group and gender (male, female, and overall). Results for glucose, insulin, and lipid tests (HDL cholesterol, LDL cholesterol, total cholesterol, and triglycerides), along with a derived variable for homeostasis model assessment of insulin resistance (HOMA-IR), will be presented separately by fasting status, which includes fasting only, and overall (fasting, non-fasting, or unknown). HOMA-IR will be derived based on glucose and insulin results as follows:

HOMA-IR = Glucose (mg/dL) x Insulin (mU/L) / 405 for conventional unit;

HOMA-IR = Glucose (mmol/L) x Insulin (μU/L) /22.5) for SI unit.

The normal reference ranges for laboratory tests will be used to determine whether the laboratory test value is below, within, or above the normal range. Shifts from DB baseline and shifts from OL baseline over time will be produced by analysis group to show the percentage of subjects with laboratory test values below, within, and above the normal range.

Potentially Markedly Abnormal Laboratory Values (PMALV) per conventional unit and the equivalent International System of Units (SI) for selected laboratory parameters can be found in Table 3.1. The criterion of PMALV is satisfied if a value falls into the potentially markedly abnormal range. The number and percentage of subjects with treatment emergent PMALV in the first 52-weeks treatment period and during overall treatment period, and the number and percentage of subjects who meet the threshold in Table 3.2 per conventional unit and the equivalent International System of Units (SI) by study visit will be presented by analysis group and by analysis subgroup as needed.

Table 1.1 Criteria for Potentially Markedly Abnormal Laboratory Values for Pediatric and Adolescent Subjects

Conventional Unit		SI Unit	
Parameter (Unit)	**Potentially Markedly Abnormal Range**	**Parameter (Unit)**	**Potentially Markedly Abnormal Range**
Hematology			
Hemoglobin (g/dL)	Female: ≤9.5, Male: ≤ 11.5 Female or Male: ≥ 17.2	Hemoglobin (g/L)	Female: ≤95, Male: ≤ 115 Female or Male: ≥ 172
Haematocrit (fraction, %)	≤30, ≥ 50	Haematocrit (V/V)	≤0.3, ≥ 0.5
WBC (10^3/μL)	≤2.8, ≥16	WBC (10^9/L)	≤2.8, ≥16
RBC (10^6/μL)	≤3.0, ≥6.0	RBC (10^{12}/L)	≤3.0, ≥6.0
Platelet Count (10^3/μL)	≤100, ≥ 500	Platelet Count (10^9)/L)	≤100, ≥ 500
Eosinophils (%)	≥ 10	Eosinophils (%)	≥ 10
Neutrophils (%)	≤ 15	Neutrophils (%)	≤ 15
Clinical Chemistry			
ALP (U/L)	≥ 3 x ULN	ALP (U/L)	≥ 3 x ULN
ALT (U/L)	≥ 2 x ULN	ALT (U/L)	≥ 2 x ULN
AST (U/L)	≥ 2 x ULN	AST (U/L)	≥ 2 x ULN
Total Bilirubin (mg/dL)	≥ 2.0	Total Bilirubin (umol/L)	≥ 34.2
Albumin (g/dL)	< 50% LLN	Albumin (g/L)	< 50% LLN
Creatinine (mg/dL)	≥ 2.0	Creatinine (umol/L)	≥ 176.8
Creatine Phosphokinase (CPK) (U/L)	≥ 450	Creatine Phosphokinase (CPK) (U/L)	≥ 450
LDH (U/L)	≥ 3 x ULN	LDH (U/L)	≥ 3 x ULN
GGT (U/L)	≥ 150	GGT (U/L)	≥ 150
Sodium (mEq/L)	≤130, ≥150	Sodium (mmol/L)	≤130, ≥150
Potassium (mEq/L)	≤3, ≥5.5	Potassium (mmol/L)	≤3, ≥5.5
Bicarbonate (mEq/L)	<18, >30	Bicarbonate (mmol/L)	<18, >30
Calcium (mg/dL)	< 8.4, >11.5	Calcium (mmol/L)	< 2.1, > 2.875
Chloride (mEq/L)	< 90, > 115	Chloride (mmol/L)	< 90, > 115
Blood Urea Nitrogen (mg/dL)	≥ 30	Blood Urea Nitrogen (mmol/L)	≥ 11.25

Glucose (fasting) (mg/dL) (random) (mg/dL)	 ≤45, ≥126 ≤45, >200	Glucose (fasting) (mmol/L) (random) (mmol/L)	 ≤2.4975, ≥ 6.993 ≤2.4975, > 11.1
Prolactin (ng/mL)	≥ 1 x ULN	Prolactin (mIU/L)	≥ 1 x ULN
Lipid Panel (Fasting)			
Total Cholesterol (mg/dL)	≥ 240	Total Cholesterol (mmol/L)	≥ 6.216
Triglycerides (mg/dL)	Female: ≥170, Male: ≥ 200	Triglycerides (mmol/L)	Female: ≥ 1.921 , Male: ≥ 2.26
HDL-C (mg/dL)	≤40	HDL-C (mmol/L)	≤ 1.036
LDL-C (mg/dL)	≥160	LDL-C (mmol/L)	≥ 4.144
Urinalysis			
RBC (hpf)	>25	RBC (hpf)	>25
WBC (hpf)	>25	WBC (hpf)	>25

ULN = Upper Limit of the Normal range. LLN = Lower Limit of the Normal range.

Table 3.2 Thresholds for Selected Laboratory Parameters

Conventional Unit		**SI Unit**	
Parameter (Unit)	**Threshold**	**Parameter (Unit)**	**Threshold**
Prolactin (ng/mL)	≥ 1 x ULN	Prolactin (mIU/L)	≥ 1 x ULN
	≥ 2 x ULN		≥ 2 x ULN
	≥ 3 x ULN		≥ 3 x ULN
	≥ 5 x ULN		≥ 5 x ULN
HDL-C (mg/dL)	≤30	HDL-C (mmol/L)	≤ 0.777
	≤40		≤ 1.036
LDL-C (mg/dL)	≥ 160	LDL-C (mmol/L)	≥ 4.144
	≥ 200		≥ 5.180
Total Cholesterol (mg/dL)	≥ 200	Total Cholesterol (mmol/L)	≥ 5.180
	≥ 240		≥ 6.216

A listing of subject data for those who have experienced at least once PMALV in the open-label study will be generated.

Urine drug screen results will be presented for each parameter at DB baseline, OL baseline, post-OL baseline visit, W52 endpoint, and endpoint. Number and percentage of positive and negative results will be displayed by analysis group.

2.6.6. ECG Evaluations

As a part of each ECG, heart rate (HR), axis, and the following intervals will be measured: RR, PR, QRS, and QT. Corrected QT intervals (QT_cB and QT_cF) values are based on the following formula:

- Bazett: QT_cB (msec) = QT (msec) / $(RR/1000)^{1/2}$
- Fridericia: QT_cF (msec) = QT (msec) / $(RR/1000)^{1/3}$

where RR is reported in milliseconds (msec).

Descriptive statistics (eg, mean, SD, median, range, 95% CI) for observed value at the DB baseline, OL baseline, each post-OL baseline visit, W52 endpoint, and endpoint and changes from DB baseline and OL baseline, respectively will be displayed for continuous variables (ie, HR, QRS, QTcF, etc) by analysis group.

The overall ECG assessment and ECG tracing results (T wave, U wave, ST segment etc.) are conducted by central reading and a determination is made whether the reading is normal (or absent if applicable) or abnormal. Shifts from OL baseline and Shifts from DB baseline over time will be produced by analysis group.

Bazett and Fridericia corrected QT_c values will be classified as having QT_c prolongation according to Table 4.

Table 4 QTc Prolongation for Pediatric and Adolescent Subjects

QTc Prolongation
QTc >460 msec QTc >480 msec QTc >500 msec
Increase from OL baseline QTc ≥ 30 msec Increase from DB baseline QTc ≥ 30 msec Increase from OL baseline QTc ≥ 60 msec Increase from DB baseline QTc ≥ 60 msec

The number and percentage of subjects having a QT_c prolongation using Fridericia correction method for DB baseline, OL baseline, post-OL baseline visit, W52 endpoint, endpoint, and any treatment emergent occurrence in the first 52-weeks treatment period, and during the open-label treatment period (the worst case), will be summarized by analysis group. A listing of ECG data for subjects with at least one prolonged QT_c will also be produced.

For other ECG parameters, subjects will be classified as normal or abnormal using the limits specified in Table 5. For abnormal ECG parameter reporting, counts and percentages of subjects will be presented for DB baseline, OL baseline, each post-OL baseline visit, W52 endpoint, endpoint, and any treatment emergent occurrence in the first 52-weeks treatment period, and during the open-label treatment phase by analysis group.

Table 2 Definition of Abnormal ECG Values by Parameter for Pediatric and Adolescent Subjects

ECG parameter (unit)	Age (years old)	Abnormally Low	Abnormally High
HR (bpm)	6 to <8	<65	> 115
	8 to <12	<55	>110
	12 to <16	<50	>105
	≥16	<50	>100
PR interval (msec)	6 to <8	--	>160
	8 to <12	--	>175
	12 to <16	--	>180
	≥16	--	>200
QRS interval (msec)	6 to <8	--	>100

	8 to <12	--	>105
	12 to <16	--	>110
	≥16	--	>120

2.6.7. Vital Signs, Weight, Waist Circumference, and BMI

Descriptive statistics (mean, SD etc.) will be provided for observed at the DB baseline, OL baseline, post-OL baseline visit, W52 endpoint, and endpoint, and change from DB baseline and OL baseline values, respectively of each vital sign parameters, weight, height, BMI, waist circumference listed below:

- Systolic blood pressure (SBP) in supine and standing positions
- Diastolic blood pressure (DBP) in supine and standing positions
- Heart rate in supine and standing positions
- Orthostatic change in SBP
- Orthostatic change in DBP
- Orthostatic change in heart rate
- Respiratory rate
- Temperature
- Body weight, height, BMI, and waist circumference,
- Age-and-sex specific z-scores of height and BMI (WHO)
- Age-and-sex specific z-score of weight (CDC)

WHO 2007 growth reference provides the reference information up to age of 19 years and 0 month while CDC 2000 growth chart provides the reference for United State subjects up to age of 20 years and 0 month. A small proportion of subjects in study 302 will reach 19 years old and a few subjects may reach 20 years old during in the extension study. For each subject, WHO 2007 growth reference at 19 years and 0 month will be used to derive the z-score and percentile of height and BMI for data collected after 19 years old; CDC 2000 growth chart at 20 years and 0 month will be used to derive the z-score and percentile of weight for data collected after 20 years old. For US subjects, CDC 2000 growth chart at 20 years and 0 month will be used to derive the z-score and percentile of height and BMI for data collected after 20 years old.

For analysis group and analysis subgroups A and C, WHO based z-scores and percentiles will be reported for height and BMI and CDC based z-score and percentile will be reported for weight. For analysis subgroup B and the output specific for US subjects, CDC based z-scores and percentiles of height, weight, and BMI will be summarized accordingly.

To better interpret these growth indexes, the expected value of height and BMI relative to DB baseline per WHO and the expected value of height, weight, and BMI relative to DB baseline per CDC at each open-label visits will be derived; corresponding changes in expected value from DB baseline will be generated as well. The number and percentage of subjects with weight status categories, defined as a value falling into the categories in Table 6.1 and Table 6.2, will be

summarized by analysis group and by analysis subgroup, respectively. This includes subjects having experienced at least one occurrence during the open-label extension study, up to and including endpoint.

Table 6.1 Weight Status Categories for Pediatric and Adolescent Subjects per WHO Standard

BMI (kg/m2)	Weight Status
< 3th percentile	Underweight
3th percentile ≤ BMI ≤ 85th percentile	Normal
85th percentile < BMI ≤ 97th percentile	Overweight
> 97th percentile	Obese

Table 6.2 Weight Status Categories for Pediatric and Adolescent Subjects per CDC Standard

BMI (kg/m2)	Weight Status
< 5th percentile	Underweight
5th percentile ≤ BMI ≤ 85th percentile	Normal
85th percentile < BMI ≤ 95th percentile	Overweight
>95th percentile	Obese

Shifts from OL baseline and shifts from DB baseline over time for weight status categories (Tables 6.1) will be summarized by analysis group to show the number and percentage of subjects that fall into above categories. A similar shift analysis will be done for US subjects (Table 6.2) by analysis subgroup D.

The number and percentage of subjects with Potentially Markedly Abnormal Vital Signs (PMAVS), defined in Table 7, will be presented by analysis group at DB baseline and OL baseline (when applicable), post-OL baseline visits, W52 endpoint, endpoint, and in the first 52-weeks treatment period, and any treatment-emergent occurrence during the open-label treatment phase (the worst case).

Table 7 Criteria for Potentially Markedly Abnormal Vital Signs for Pediatric and Adolescent Subjects

Parameter (unit)	Age (years old)	Markedly Low	Markedly High
SBP (supine, standing) (mmHg)	6-12	Value ≤ 70 and ≥ 20 decrease from baseline	Value ≥ 120 and ≥ 20 increase from baseline
	13-18	Value ≤ 90 and ≥ 20 decrease from baseline	Value ≥ 135 and ≥ 20 increase from baseline
	≥ 19	Value ≤ 90 and	Value ≥ 180 and

		≥ 20 decrease from baseline	≥ 20 increase from baseline
DBP (supine, standing) (mmHg)	6-12	Value ≤ 40 and ≥ 15 decrease from baseline	Value ≥ 80 and ≥ 15 increase from baseline
	13-18	Value ≤ 50 and ≥ 15 decrease from baseline	Value ≥ 90 and ≥ 15 increase from baseline
	≥ 19	Value ≤ 50 and ≥ 15 decrease from baseline	Value ≥ 105 and ≥ 15 increase from baseline
Pulse rate (supine, standing) (bpm)	6-10	Value ≤ 60 and ≥ 15 decrease from baseline	Value ≥ 135 and ≥ 15 increase from baseline
	11-18	Value ≤ 50 and ≥ 15 decrease from baseline	Value ≥ 120 and ≥ 15 increase from baseline
	≥ 19	Value ≤ 50 and ≥ 15 decrease from baseline	Value ≥ 120 and ≥ 15 increase from baseline
SBP orthostatic criteria (mmHg)	~	≥ 20 decrease from supine to standing position	NA
DBP orthostatic criteria (mmHg)	~	≥ 10 decrease from supine to standing position	NA
Pulse rate orthostatic criteria (bpm)	~	NA	≥ 20 increase from supine to standing position
Temperature (˚C)	~	NA	Value ≥ 38.3°C and ≥0.8°C increase from baseline
Note: ~ means that the abnormal range is applicable for all subjects; Baseline represents 'DB baseline' and 'OL baseline' respectively.			

A listing of subject data for those who satisfy MAPVS during the open-label study will be presented.

2.6.8. Young Mania Rating Scale (YMRS)

The YMRS is only applied to subjects who continued from Study D1050326.

The YMRS is an 11-item instrument used to assess the severity of mania in patients with a diagnosis of bipolar disorder. Ratings are based on patient self-reporting, combined with clinician observation. Seven items are rated on a 5-point scale, ranging from 0 to 4, and four items are rated on a 9-point scale, ranging from 0 to 8.

- Seven Items that are scored 0 to 4 are:
  - Elevated mood
  - Increased Motor Activity – Energy
  - Sexual Interest
  - Sleep
  - Language/Thought Disorder
  - Appearance
  - Insight
- Four items that are scored 0 to 8 are
  - Irritability

- Speech (Rate and Amount)
- Content
- Disruptive/Aggressive Behavior

The YMRS total score is the sum of the 11 individual items and ranges from 0 to 60. A higher score is associated with a greater severity of mania.

Analysis Methods

Subjects continued from Study D1050326

Summary statistics for the observed at the DB baseline, OL baseline, each post-OL baseline time point, W52 endpoint, and endpoint, change from DB baseline, and change from OL baseline in YMRS total score will be presented by analysis subgroup.

2.6.9. Extrapyramidal Symptoms Assessment Scales

2.6.9.1. Barnes Akathisia Rating Scale

The Barnes Akathisia Rating Scale (BARS) consists of four items: an objective rating (ranging 0 to 3), 2 subjective ratings of symptoms of akathisia (awareness of restlessness and reported distress related to restlessness: ranging 0 to 3), and a global clinical rating of akathisia, ranging from 0 (absent) to 5 (severe). The global rating score (item 4), that is scored separately, is the most relevant measure of severity of akathisia. Higher scores denote worsening akathisia. The subjective and objective items (items 1 through 3) are summed to yield a BARS total score with a range of 0 to 9. Higher values of the BARS total score indicate higher severity of akathisia.

For subjects continued from Studies D1050301 and D1050326

Two items for subjective ratings of symptoms of akathisia (awareness of restlessness and reported distress related to restlessness) range from 0 to 3.

For subjects continued from Study D1050325

Two items for subjective ratings of symptoms of akathisia (awareness of restlessness and reported distress related to restlessness) range from 0 to 3, and Not Applicable (NA). When any subjective rating is recorded as NA, the BARS total score will not be calculated and will be set as NA as well.

2.6.9.2. Abnormal Involuntary Movement Scale

The Abnormal Involuntary Movement Scale (AIMS) contains items related to: facial and oral movements; extremity movements; trunk movements; global judgments and dental status. The individual items on the AIMS range from 0= 'None' to 4= 'Severe'. Three other items assess the subject at a global level, and two items assess dental status. The individual items on the AIMS are rated from 0= 'None' to 4= 'Severe'. The (non-global) total AIMS score is the sum of items 1 through 7 (ranging from 0 to 28). Items 8 through 12 will not be calculated in total AIMS

score. Higher values of the total AIMS score indicate increased severity in abnormal movement. The global severity score (Item 8) is the response to, "Severity of abnormal movements", found within the global judgments section (0 - None, normal, 1 – Minimal, 2 – Mild, 3 – Moderate and 4 – Severe) and will be summarized separately.

2.6.9.3. Simpson-Angus Scale

The Simpson-Angus Scale (SAS) is a clinician-rated assessment of neuroleptic-induced Parkinsonism consisting of 10 items. Items are anchor-based, rated on a 5-point scale (0-4), and address rigidity, gait (bradykinesia), tremor, glabellar tap, and salivation. Each individual item on the SAS ranges from 0= 'Normal' to 4= 'Extreme Symptoms'. The SAS mean score is defined as the average of all 10 items and ranges between 0 and 4. Higher value of the SAS mean score indicates more severe condition of EPS symptom

SAS version 2 was to be used in studies D1050301, D1050325, D1050326, and D1050302. It was noticed on December 9, 2016 that all Hungarian sites in studies D1050301, D1050326, and D1050302 used SAS version 1 instead. There is no plan to update the SAS version for Hungarian sites; ie, Hungarian subjects will keep use SAS version 1 till the end of study D1050302.

The difference between SAS version 1 and version 2 is highlighted below:

Version 2	**Version 1 (Hungarian Sites Only)**
1. Gait	1. Gait
2. Arm dropping	2. Arm dropping
3. Shoulder shaking	3. Shoulder shaking
4. Elbow rigidity	4. Elbow rigidity
5. Wrist rigidity or fixation of position	5. Wrist rigidity or fixation of position
6. Head rotation	6. Leg Pendulousness
7. Glabella Tap	7. Head rotation
8. Tremor	8. Glabella Tap
9. Salivation	9. Tremor
10. Akathisia	10. Salivation

The SAS total score per V1 and V2 will be derived, separately. Since the SAS version related issue was found after the DBL of studies D1050301 and D1050326, DB baseline and OL baseline associated with these two SAS versions was derived and will be used in final analysis.

Analysis Methods

Summary statistics for the observed at the DB baseline, OL baseline, each post-OL baseline time point, W52 endpoint, and endpoint, change from DB baseline, and change from OL baseline in BARS total score, BARS global clinical assessment of akathisia score, AIMS total score, SAS 10-item mean score (all versions combined) will be presented by analysis group, and by analysis subgroup.

The post-OL baseline BARS global clinical assessment of Akathisia responses and the post-OL baseline AIMS global severity scores will be classified as ‘worsened’, ‘unchanged’, or ‘improved’, relative to a subject’s corresponding DB baseline response and OL baseline response, respectively. A higher post-OL baseline response than that of DB baseline or OL baseline would be classified as ‘worsened’ relative to DB baseline and OL baseline, respectively. Conversely, a lower value than that of DB baseline or OL baseline would be classified as ‘improved, respectively. Frequency distribution over time of these variables will be summarized at each post-OL visit and endpoint by analysis group.

The AIMS total score at DB baseline, OL baseline, each post-OL visit, W52 endpoint, and endpoint will also be classified as ‘abnormal’ if either: at least two items have a response of ‘mild’ or higher; or at least one item has a response of ‘moderate’ or higher; otherwise, non-missing total scores will be classified as ‘normal’. This is a modification of the Schooler-Kane Criteria for Tardive Dyskinesia. The SAS 10-item mean score (all versions combined) at DB baseline, OL baseline, each post-OL visit, W52 endpoint, and endpoint will be classified as ‘abnormal’ if it exceeds 0.3; otherwise, non-missing mean scores will be classified as ‘normal.’ Shifts from OL baseline and shifts from DB baseline will be summarized for these variables (ie, abnormality status of the AIMS total score and the SAS 10-item mean score) at each assessed visit, endpoint, and overall for the open-label study up to and including endpoint (based on maximum severity) by analysis group.

2.6.10. CogState Computerized Cognitive Test Battery

Cogstate is only applied to the subjects continued from Studies D1050301 and D1050326.

The CogState Computerized Cognitive Test Battery consists of 4 primary, individually administered tests in 4 distinct MATRICS cognitive domains, as detailed in Table 8. Components of the CogState Computerized Cognitive Test Battery:

- Detection Task

In this task, subjects must respond to a card being displayed face up on the screen as quickly as possible. The score will be calculated as the mean of log10 transformed reaction times for correct responses. Lower scores indicate better performance.

- Identification Task

Subject must identify the color of a card as quickly as possible in this task. The score will be calculated as the mean of log10 transformed reaction times for correct responses. Lower scores indicate better performance.

- One Card Learning Task

In this task, subjects must remember which cards have been shown before in the task. The arcsine transformation of the square root of the proportion of correct responses will be recorded. Higher scores indicate better performance.

- One Back Task

In this task, subjects must

(1) decide if each card is exactly the same as the previous card presented. The score will be calculated as the arcsine transformation of the square root of the proportion of correct responses. Higher scores indicate better performance.

(2) identify the color of a card as quickly as possible in this task. The score will be calculated as the mean of log10 transformed reaction times for correct responses. Lower scores indicate better performance.

Table 8 Cognitive Domains Evaluated in CogState Tests

MATRICS Cognitive Domain	CogState Test	Data Type	Transformed Data from CogState
Psychomotor speed	Detection Task	Speed	Log10(S), S= reaction times for correct responses
Attention	Identification Task	Speed	Log10(S), S= reaction times for correct responses
Learning/Memory	One Card Learning Task	Accuracy	arcsine $\sqrt{R}$, R=proportion of correct responses
Working Memory	One Back Task	Speed	Log10(S), S= reaction times for correct responses
		Accuracy	arcsine $\sqrt{R}$, R=proportion of correct responses

A normative table of CogState Computerized Cognitive Test Battery for children and adolescent subjects is provided in Attachment 4.2. For an individual domain, the age-specific standardized score for each task and data type is derived as below based on the subject's age (in years) at the study visit. The standardized scores are calculated by CogStat vendor (COGSTATE CT).

$$(\text{Score} - \mu_{age,domain}) \times \text{Multiplicand} / SD_{age,domain},$$

Where

$\mu_{age,domain}$: mean of the norm group at the age (in years)

$SD_{age,domain}$: standard deviation of the norm group at the age (in years)

Score : individual subject performance

Multiplicand: an indicator,

- Multiplicand = –1 when a lower score indicates improved performance (e.g. Detection task (speed), Identification task (speed), One Back Task (speed));
- Multiplicand = +1 when a higher score indicates improved performance (e.g. One Card Learning Task (accuracy), One Back Task (accuracy)).

Based on the standardized scores of individual task and data type (accuracy, speed) provided by the vendor, the sponsor will further derive the following composite scores for final analyses:

(1) The standardized composite score is calculated as the mean of the standardized scores of Detection Task (speed), Identification Task (speed), One Back Task (speed), and One

Card Learning Task (accuracy). If two or more components are missing at a visit, the composite score will be set to missing.

(2) For each subject, change from baseline in standardized score for each of the 4 cognitive domains, and the standardized composite score will be derived at each assessed visit.

Definition of the standardized composite score in this SAP is same as the one in SAPs of interim analysis 1 and study D1050326 and is also same as the standardized composite score 2 in clinical study report of study D1050301.

Analysis Methods

For subjects continued from Studies D1050301 and D1050326

For CogState computerized cognitive composite standardized score and individual standardized domains scores, descriptive summaries (N, mean, standard deviation, median, and range) of observed and change from DB Baseline and OL Baseline will be provided by visit, W52 endpoint, and endpoint and analysis group and analysis subgroup, respectively.

2.6.11. Columbia Suicide Severity Rating Scale (C-SSRS)

C-SSRS is applied to the subjects continued from Study D1050301 and C-SSRS children version is applied to the subjects continued from Study D1050326.The C-SSRS Childrens' Version is distinguished from the adult version by the absence of questions regarding the duration, controllability, deterrents and reasons for ideation.

Severity of suicidal ideation is rated on a 6-point scale from 0='No ideation present' to 5='Active ideation with plan and intent'. A score of 4 or 5 on this scale indicates serious suicidal ideation. Suicidal ideation score is defined as the most severe suicidal ideation score (1-5 on the CSSRS) present at the assessment; assign a score of 0 if no ideation is present.

For subjects continued from study D 1050301, the ideation intensity total score is calculated as summation of the five items from the Ideation Intensity scale: frequency, duration, controllability, deterrents, and reasons for ideation. If one or more of these 5 items are missing at a visit, no imputation will be made and the corresponding ideation intensity total score will be set as missing. Items corresponding to frequency and duration are each rated on a scale from 1 to 5 and items corresponding to controllability, deterrents, and reasons for ideation are each rated on a scale from 0 to 5. If the patient did not endorse any suicidal ideation, set the intensity rating to 0. Thus, the possible range for the intensity total score is 0 to 25. For subjects continued from study D 1050326, the ideation intensity total score is not applicable and therefore will not be derived.

Suicidal behavior is collected as presence/absence of actual attempts, non-suicidal self-injurious behavior, interrupted attempts, aborted attempts, preparatory acts or behavior, and any suicidal behavior. In addition, the number of actual attempts, interrupted attempts, and aborted attempts are captured. Any attempt will be defined as suicidal behavior.

The lethality associated with actual attempts is rated on a 6-point scale from 0='No physical damage or very minor physical damage' to 5='Death'. Potential lethality of attempts is rated on a

3-point scale from 0='Behavior not likely to result in injury' to 2='Behavior likely to result in death despite available medical care'.

According to FDA draft guidance for suicidal ideation and behavior ([5], Appendix A), the following C-SSRS outcomes, which have binary responses (yes/no), are re-ordered and numbered to facilitate the further definitions of the composite endpoint and comparative endpoint of C-SSRS as below.

Category 1 – Wish to be Dead
Category 2 – Non-specific Active Suicidal Thoughts
Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent

Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt
Category 10 – Completed Suicide
Category 11– Self-injurious behavior without suicidal intent

Unless otherwise specified, C-SSRS table outputs will be presented by analysis subgroup for one analysis time point 'Any extension', whose definition is given in Table 9.

Table 9 Analysis TimePoint for C-SSRS

Analysis Timepoint	**eCRF Visit(s) to be Involved**	**C-SSRS Version /Assessment Period**	**Derivation Rule**
Any extension visit*	Up to EOT	Since last visit	Most severe value
*: including all visits in open-label treatment period (see Section 2.1.2).			

Comparative endpoints for 'Emergence' (outcomes that include events that first emerge) and 'Worsen' (outcomes that include events that worsen) will be derived below).

Table 10.1 Comparative Endpoint Relative to DB Baseline

Comparative Endpoints	**Analysis Time Points Involved**	**Derivation Rule**
Emergence of Suicidal Ideation	DB Baseline(a) Any extension visit(b)	No suicidal ideation at DB Baseline, and any type of suicidal ideation in 'any extension visit'
Emergence of Serious Suicidal Ideation		No suicidal ideation at DB Baseline, and any serious suicidal ideation [ideation score of 4 or 5] in 'any extension visit'

Emergence of Suicidal Behavior		No suicidal behavior at DB Baseline, and any type of suicidal behavior in 'any extension visit'
Emergence of Suicidality		No suicidality at DB Baseline, and any suicidality in 'any extension visit'
Worsening of Suicidal Ideation		Most severe suicidal ideation in 'any extension visit' is more severe than the most severe value at DB Baseline
(a): Definition of DB Baseline: see Table 9 in D1050301 SAP, and Table 9 in D1050326 SAP ; (b): see Table 9.		

Table 10.2 Comparative Endpoint Relative to DB Treatment Period

Comparative Endpoints	**Analysis Time Points Involved**	**Derivation Rule**
Emergence of Suicidal Ideation	Any DB treatment period [(a)] Any extension visit[(b)]	No suicidal ideation in DB treatment period, and any type of suicidal ideation in 'any extension visit'
Emergence of Serious Suicidal Ideation		No suicidal ideation in DB treatment period, and any serious suicidal ideation [ideation score of 4 or 5] in 'any extension visit'
Emergence of Suicidal Behavior		No suicidal behavior in DB treatment period, and any type of suicidal behavior in 'any extension visit'
Emergence of Suicidality		No suicidality in DB treatment period, and any suicidality in 'any extension visit'
Worsening of Suicidal Ideation		Most severe suicidal ideation in 'any extension visit' is more severe than the most severe value in any DB treatment period
[(a)]: Definition of 'any DB treatment period': see 'any post-baseline' in Table 9 of D1050301 SAP, and 'any post-baseline' in Table 9 of D1050326 SAP ; [(b)]: see Table 9.		

Analysis Methods

For subjects continued from Studies D1050301 and D1050326

The number and percentage of subjects with following categories will be summarized by analysis group and analysis subgroup for any extension visit.

- each type of suicidal ideation,
- any suicidal ideation,
- each type of suicidal behavior,
- any suicidal behavior,

- any having self-injurious behavior without suicidal intent,
- any suicidality

In addition, for any extension visit, the number and percentage of subjects with following categories will be also summarized by analysis group and analysis subgroup.

- emergence of suicidal ideation,
- emergence of serious suicidal ideation,
- worsening of suicidal ideation,
- emergence of suicidal behavior,
- emergence of suicidality

Shifts from DB baseline and OL baseline to any post-OL baseline visit (ie, the worst case) in suicidal ideation score will be presented by analysis group and analysis subgroup, to show the percentage of subjects with values 0 ('No ideation present') to 5 ('Active ideation with plan and intent'). For subjects continued from study D1050301, ideation intensity total score will be provided in data listing. Data collected on actual suicide attempts (lethality of actual attempts and potential lethality of attempts) will be presented in a data listing.

2.6.12. Udvalg for Kliniske Undersogelser Side Effect Rating Scale (UKU)

UKU is only applied to the subjects continued from Studies D1050301 and D1050326.

The UKU consists of 48 adverse effects items, divided into 4 categories: psychic (10 items), neurologic (8 items), autonomic (11 items), and other (19 items). Each item is rated on a 0-3 point scale: 0– no side effects, 1– mild, 2– moderate, 3– severe with causal relationship ('improbable', 'possible', 'probable') also assessed for scores of 1-3. Score of 9 is given if symptom is not assessed or cannot be assessed if it does not apply to a subject (gender, age, sexual maturation); for individual items that are not able to be assessed the rater would line out the item and enter N/A and not assessed would be checked in the symptom checklist. Specifically, male subjects and female subjects who have not achieved menarche will be rated with a 9 ('Not Assessed) for items 4.7, 4.8, and 4.16, since these items are designed only for female subjects who have achieved menarche; likewise, female subjects will be rated with a 9 ('Not Assessed) for items 4.10, 4.13, and 4.14, since these items are designed for male subjects only.

Unlike other UKU adverse effect item, which is composed of a single assessment, UKU items 4.1 (rash, sub-items: 4.1a~4.1e) and 4.17 (headache, sub-items: 4.17a~4.17c) contains assessments for multiple sub-categories. For these two items, eCRF collects data in the following way: when an adverse event exists, a rating score is to be collected at a specific sub-item level only; when none of these sub-items show adverse event, only item 4.1 (or 4.17) will have value 0 (or 9 when applicable). The rating score related to items 4.1 and 4.17 will be further derived as below:

- If the rating score for at least one of its sub-items is 1 and higher (ie, 1~3), the rating score of item 4.1 (or 4.17) will be set as the most severe value among its sub-items;
- If item 4.1 (or 4.17) has value 0 or 9, its corresponding sub-items will be set as 0 or 9.

In addition to the 48 questions for symptom, two global assessment of effect on functioning are collected in UKU (one by clinician and one by patient); clinical "consequence" with respect to management will be rated by clinician only.

The following four UKU category scores will be calculated as the summation of the assessed individual scores within each category:

- psychic side effects score (range from 0 up to 30),
- neurologic side effects score (range from 0 up to 24),
- and autonomic side effects score (range from 0 up to 33)
- other side effects score (range from 0 up to 48)

Specifically, item with value 9 ('Not Assessed') will not be included in calculation of these category scores. Particularly, UKU other side effects score is gender-specific, ie, items 4.10, 4.13 and 4.14 will not be included in UKU other side effects score calculation for female subjects while items 4.7, 4.8, and 4.16 will not be included in UKU other side effects score calculation for male subjects.

If two or more components are missing (not 'Not Assessed') at a visit, the corresponding UKU category score will be set as missing; when only one component item is missing, the UKU category scores will be calculated based on the available items at a visit. The total UKU score is calculated as the sum of four UKU category scores (ranging from 0 up to 135). If one UKU category score is missing, the total UKU score will be set as missing. Higher values of the total UKU score and four category scores indicate more severe side effects.

Unless otherwise specified, UKU table outputs for following binary effect and the corresponding drug-related effect will be presented by analysis subgroup for analysis timepoint 'any extension visit'.

- Any psychic side effect (ie, Items 1.1 – 1.10)
- Any neurologic side effect (ie, Items 2.1 – 2.8)
- Any autonomic side effect (ie, Items 3.1 – 3.11)
- Any other side effect (ie, Items 4.1 – 4.17)

For UKU assessed in the open-label treatment period, causal relationship to study medication will be grouped into 2 categories (determined by investigators): related (which includes 'possible', 'probable'), and not related ('improbable'). For any UKU assessment prior to the start of the open-label treatment period, causal relationship for UKU items will be 'not related' for previously placebo subjects. During the any open-label treatment, drug-related effect will be defined similarly as above based on the worst case related to the study medication.

A score of 1, 2 or 3 on any UKU item that first occurred or worsened during the open-label treatment period indicated 'cases' of adverse events. Adverse events were conceptualized as a

dichotomous variable, either 'present' (1 (mild) or higher) or 'not present' (0, 9). The severe order is defined as 3>2> 1> 0, 9. Both DB baseline and 'the worst case of DB treatment period' will be used to derive the 'Emergence' and 'Worsen' for 'Any extension visit' respectively for the parameters below.

- Any Individual side effect (by individual item)
- Any psychic side effect,
- Any neurologic side effect,
- Any autonomic side effect,
- Any other side effect

If one or more of the two required values are missing for 'emergence of side effect' or 'worsen of side effect', corresponding side effect value will be set as missing.

Table 11.1 Comparative Endpoint of UKU Relative to DB Baseline

Comparative Endpoints	**Derivation Rule**
Emergence of side effect Relative to **DB Baseline**	'not present' at DB Baseline and any 'present' in the open-label treatment period
Emergence of drug-related side effect Relative to **DB Baseline**	'not present' at DB Baseline and any drug-related 'present' in the open-label treatment period
Worsening of side effect Relative to **DB Baseline**	the most severe rating in the open-label treatment period is more severe than it was at DB Baseline
Worsening of drug-related side effect Relative to **DB Baseline**	the most severe drug-related rating in the open-label treatment period is more severe than it was at DB Baseline
*: the value of the open-label treatment period is defined as the worst case in the open-label treatment period.	

Table 11.2 Comparative Endpoint of UKU Relative to the Core DB Treatment Period

Comparative Endpoints	**Derivation Rule**
Emergence of side effect Relative to Core DB Treatment Period	'not present' in any core DB treatment period* and any 'present' at any open-label treatment period
Emergence of drug-related side effect Relative to Core DB	'no drug-related present' in any core DB treatment period* and any drug-related 'present' in open-label treatment period

Treatment Period	
Worsening of side effect Relative to Core DB Treatment Period	the most severe rating at any open-label treatment period is more severe than the worst case of any core DB treatment period*
Worsening of drug-related side effect Relative to Core DB Treatment Period	the most severe drug-related rating in open-label treatment period is more severe than any drug-related rating in core DB treatment period*
Note: *: is referred to any post-DB baseline visit in DB treatment period of the core study (see Table 11 of D1050301 SAP or Table 11 of D1050326 SAP).	

Analysis Methods

For subjects continued from Studies D1050301 and D1050326

For individual item and four side effect categories, numbers and percentages of subjects who experience any side effect and any drug-related side effect will be summarized by analysis subgroup at DB baseline, any core DB treatment period, and any extension treatment period (when applicable), respectively.

In addition, numbers and percentages of subjects with emergence of side effect and worsening of side effect relative to DB baseline and to the core DB treatment period will be summarized by analysis group and analysis subgroup for each individual item and each side effect category, respectively.

For each side effect category score and UKU total score, descriptive summaries (N, mean, SD, median etc.) of observed and change from DB Baseline and change from OL baseline in four UKU four category side effects scores and UKU total score will be presented by analysis group and analysis subgroup at each assessment point, W52 endpoint, and end point.

Frequency distribution of clinical global assessment of effect, patient global assessment of effect, consequence assessed by clinician will be also tabulated by visit, W52 endpoint, and endpoint and by analysis group and analysis subgroup.

2.6.13. Tanner Staging and Menstrual Cyclicity

The tanner staging is a scale of physical development in children, adolescents and adults. The scale defines physical measurements of development based on external primary and secondary sex characteristics, such as the size of the breasts, genitalia, testicular volume and development of pubic and axillary hair. Tanner staging for each measurement ranges from Stage I (Preado-lescent) (numeric value 1) to Stage V (numeric value 5).

Analysis Methods

All outputs for tanner staging will be tabulated separately by gender. The frequency distribution of each category of the tanner staging will be tabulated at DB baseline, OL Baseline, post-OL Baseline visit, W52 endpoint, and endpoint by analysis group. In addition, frequency distribution of changes from DB baseline and OL baseline for each category of the tanner staging will be summarized for each post-OL visit, W52 endpoint, and endpoint by analysis group. Shifts from DB baseline and OL baseline to Endpoint in tanner stage will be presented by analysis group. The frequency distribution of quality and frequency of menstrual cyclicity will be presented by analysis group.

3. REFERENCES

1. FDA (2012), FDA written request for Latuda (lurasidone) pediatric program: NDA 200603, Reference ID: 3118727
2. WHO Growth reference 5-19 years, 2007, http://www.who.int/growthref/en/
3. WHO AnthroPlus software: http://www.who.int/growthref/tools/en/
4. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: Methods and development. *National Center for Health Statistics.* Vital Health Stat 11(246). 2002 (http://www.cdc.gov/growthcharts/2000growthchart-us.pdf).
5. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), 2012. Guidance for Industry Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials

4. ATTACHMENTS

4.1. Computation of Centiles and Z-scores for Height-For-Age, Weight-For-Age (CDC only), and BMI-For-Age

4.1.1. WHO 2007 Reference for Growth Charts

World Health Organization

COMPUTATION OF CENTILES AND Z-SCORES FOR HEIGHT-FOR-AGE, WEIGHT-FOR-AGE AND BMI-FOR-AGE

The method used to construct the 2007 WHO references relied on GAMLSS with the Box-Cox power exponential distribution (Rigby and Stasinopoulos, 2004). However, the final selected models simplified to the LMS model (Cole and Green, 1992) since none of the references required adjustment for kurtosis. As a result, the computation of percentiles and z-scores for all three indicators uses formulae based on the LMS method. However, a restriction was imposed on all indicators to enable the derivation of percentiles only within the interval corresponding to z-scores between -3 and 3. The underlying reasoning is that percentiles beyond ±3 SD are invariant to changes in equivalent z-scores. The loss accruing to this restriction is small since the inclusion range corresponds to the 0.135th to 99.865th percentiles.

For all indicators, the tabulated fitted values of Box-Cox power, median and coefficient of variation corresponding to age (or height) *t* are denoted by *L(t)*, *M(t)* and *S(t)*, respectively.

Centiles and z-scores for height-for-age

For this indicator, *L(t)* is equal to 1, simplifying the Box-Cox normal distribution used in the LMS method to the normal distribution. Therefore, differences between adjacent standard deviations (e.g. between 2 SD and 3 SD) were constant for a specific age but varied at different ages.

In this case, the centiles at age *t* can be estimated from:

$$C_{100\alpha}(t) = M(t)\left[1 + L(t)S(t)Z_\alpha\right]^{1/L(t)} = M(t)\left[1 + S(t)Z_\alpha\right]$$
$$= M(t) + StDev(t)Z_\alpha, \qquad -3 \le Z_\alpha \le 3$$

where Z_α is the normal equivalent deviate for tail area α, C100α(t) is the 100α-th centile, and $StDev(t)$ is the standard deviation at age *t* (derived from multiplying $S(t)$ by $M(t)$).

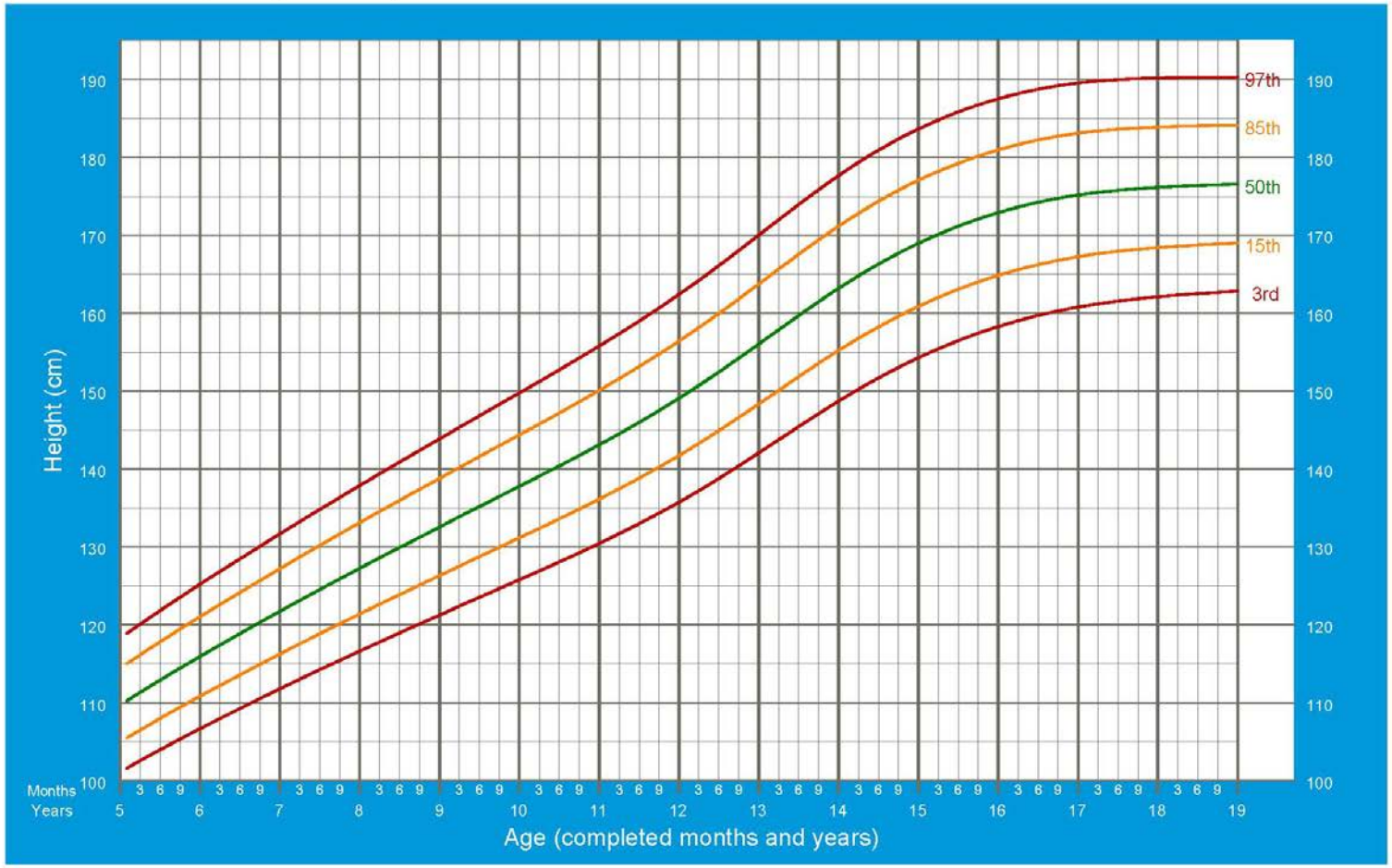
The individual z-score for a measurement *y* at age *t* was computed as:

$$z_{ind} = \frac{\left[y/M(t)\right]^{L(t)} - 1}{S(t)L(t)} = \frac{y - M(t)}{StDev(t)}$$

Height-for-age BOYS

5 to 19 years (percentiles)

World Health Organization

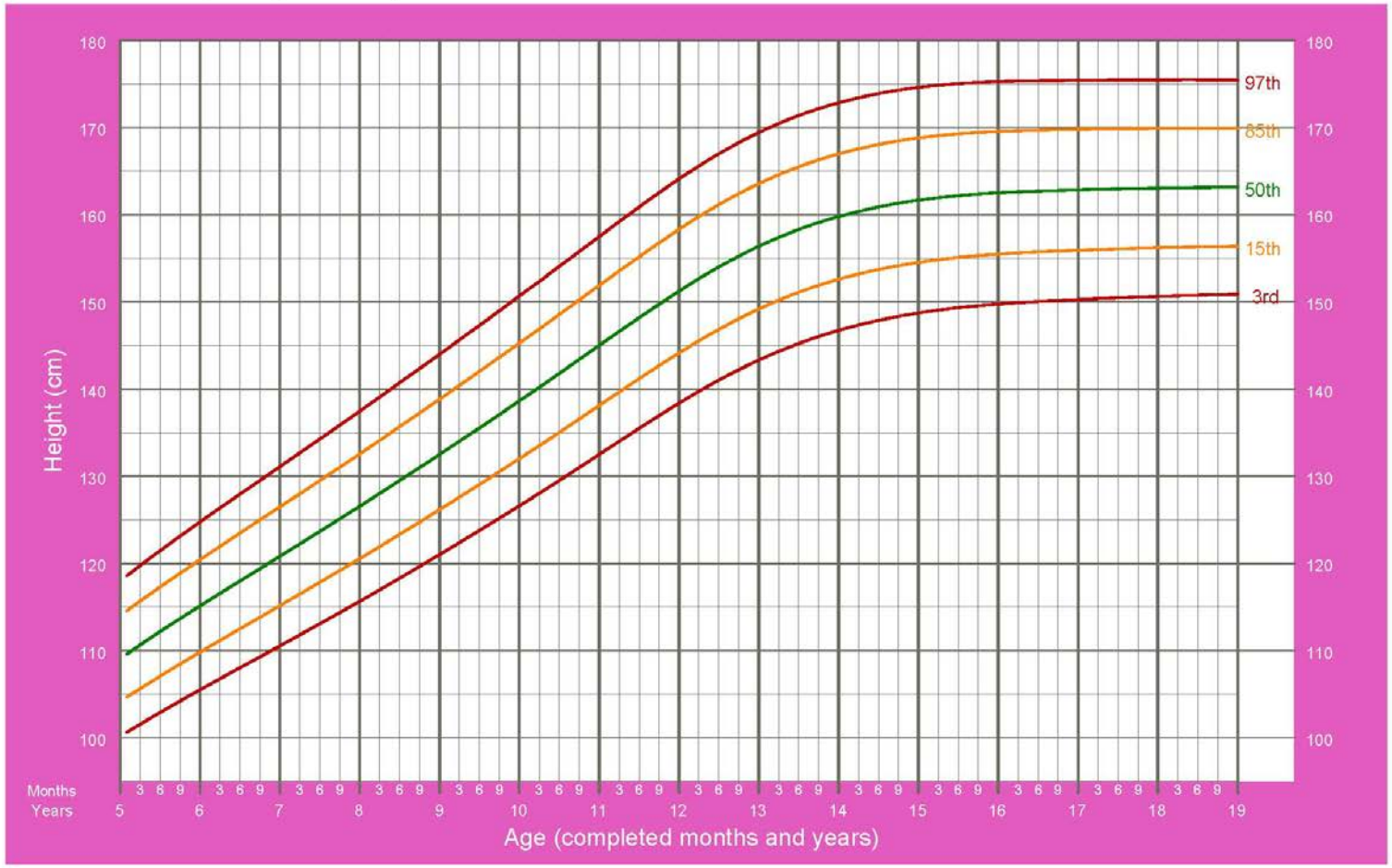


2007 WHO Reference

Height-for-age GIRLS

5 to 19 years (percentiles)

World Health Organization

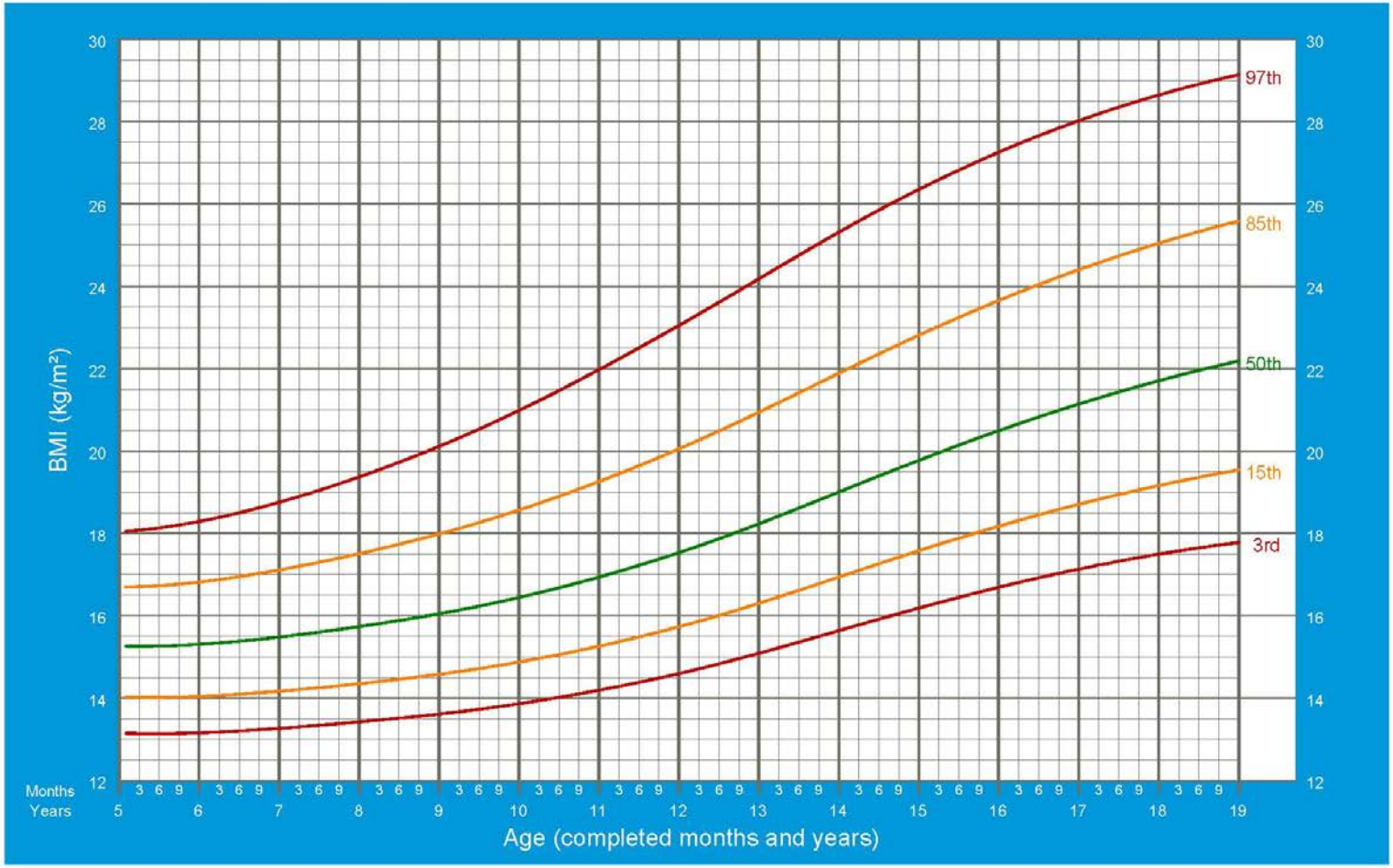


2007 WHO Reference

BMI-for-age BOYS

5 to 19 years (percentiles)

World Health Organization

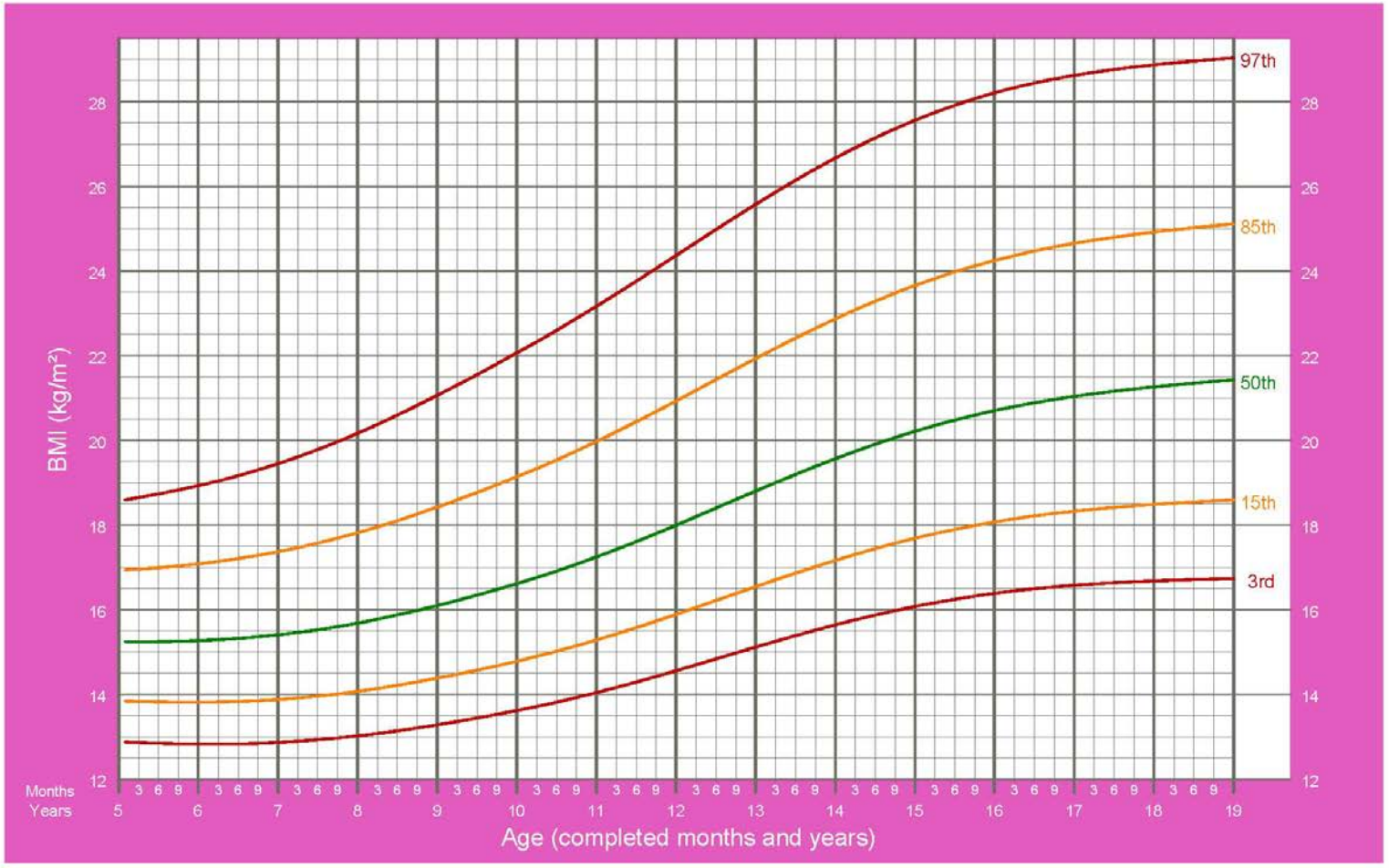


2007 WHO Reference

BMI-for-age GIRLS

5 to 19 years (percentiles)

World Health Organization



2007 WHO Reference

4.1.2. 2000 CDC Growth Charts for the United States

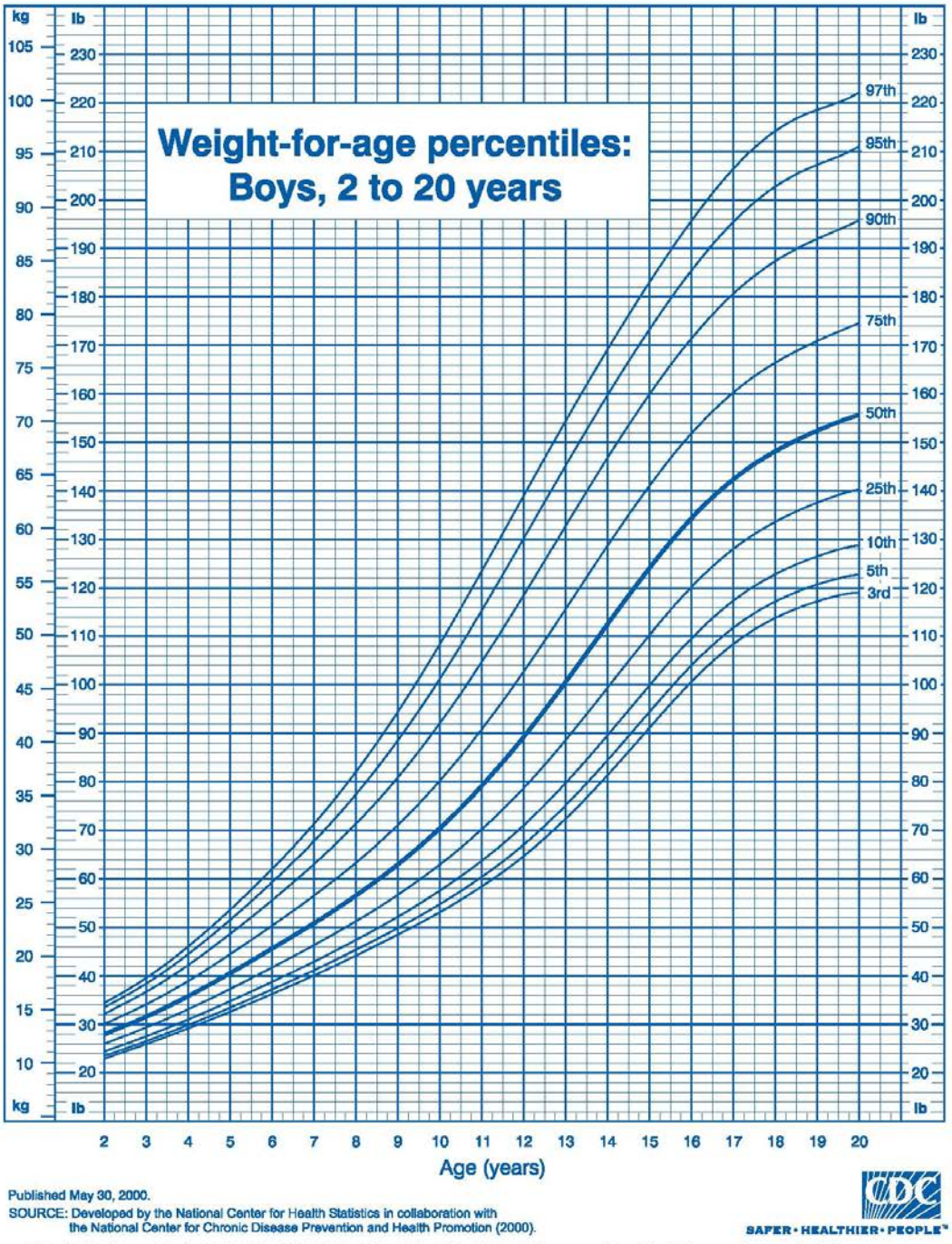


Figure 9. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, 2 to 20 years: Boys weight-for-age

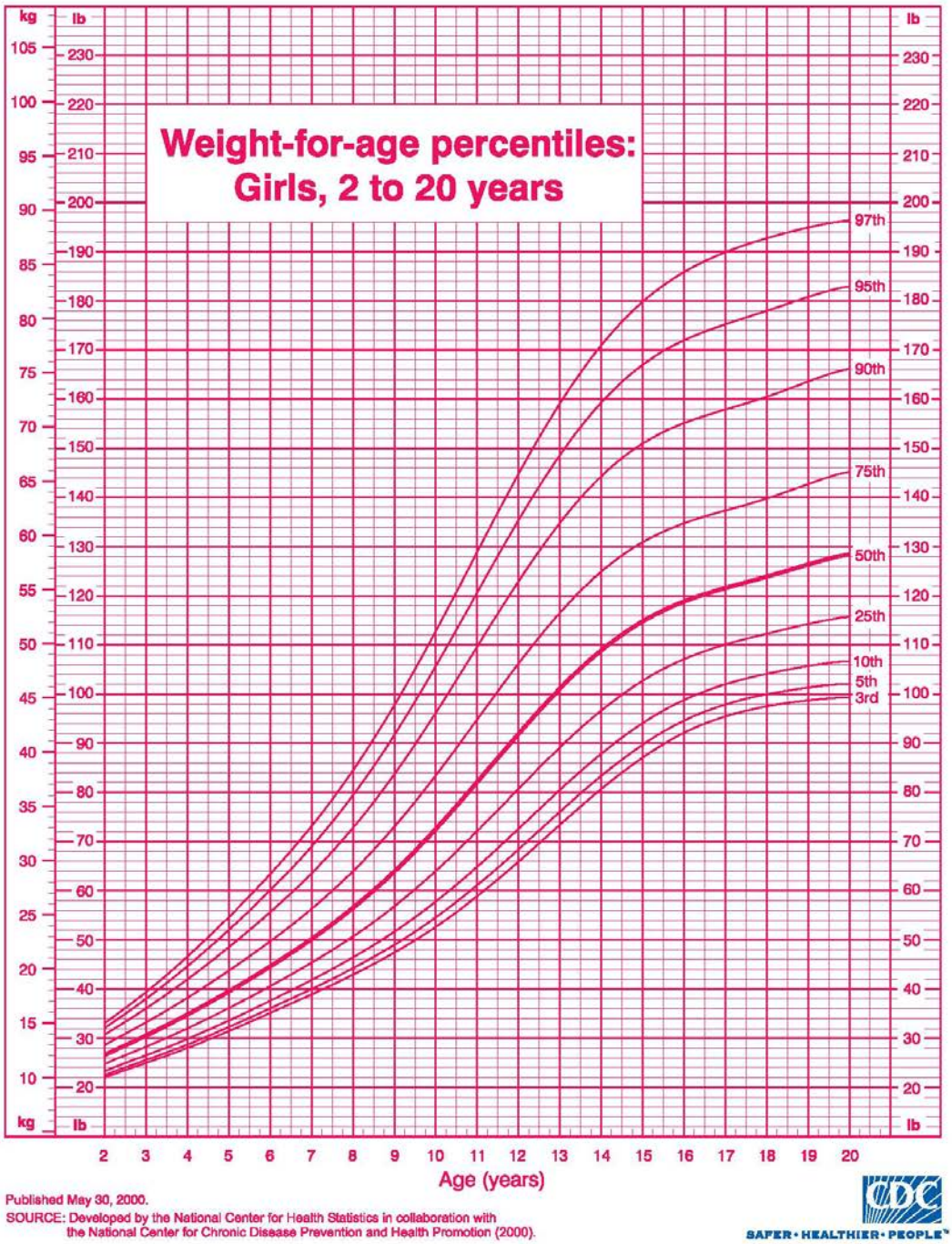


Figure 10. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, 2 to 20 years: Girls weight-for-age

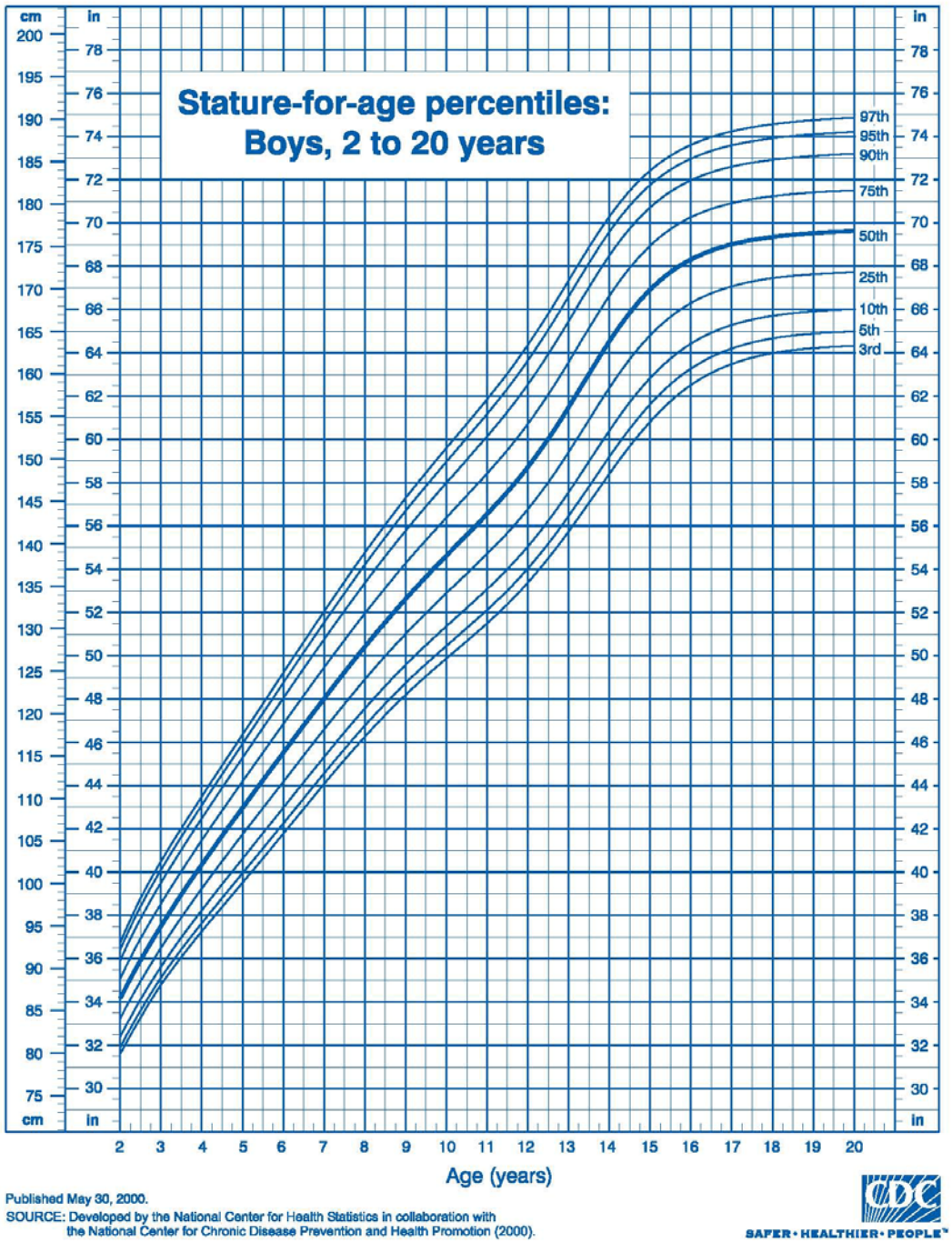


Figure 11. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, 2 to 20 years: Boys stature-for-age

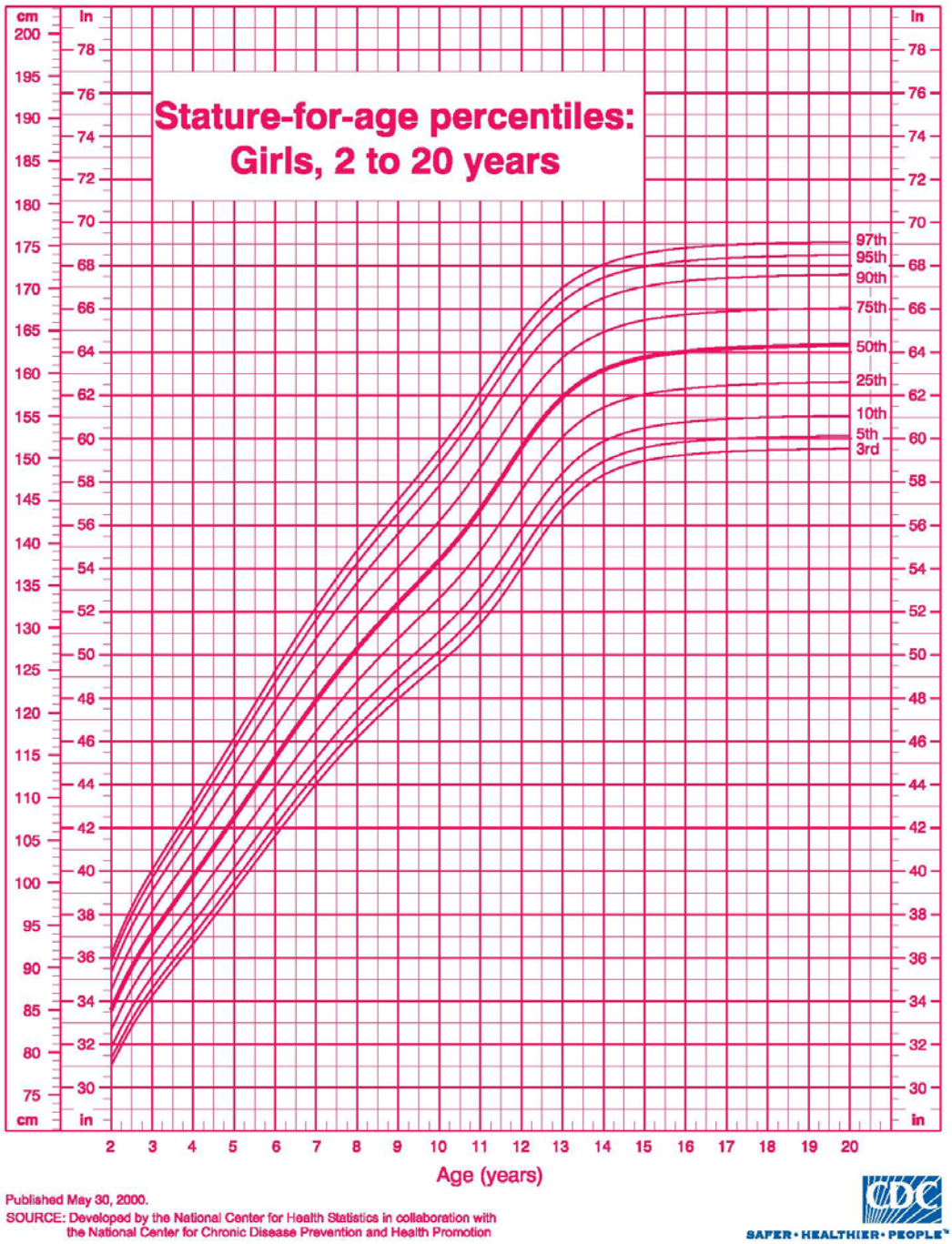
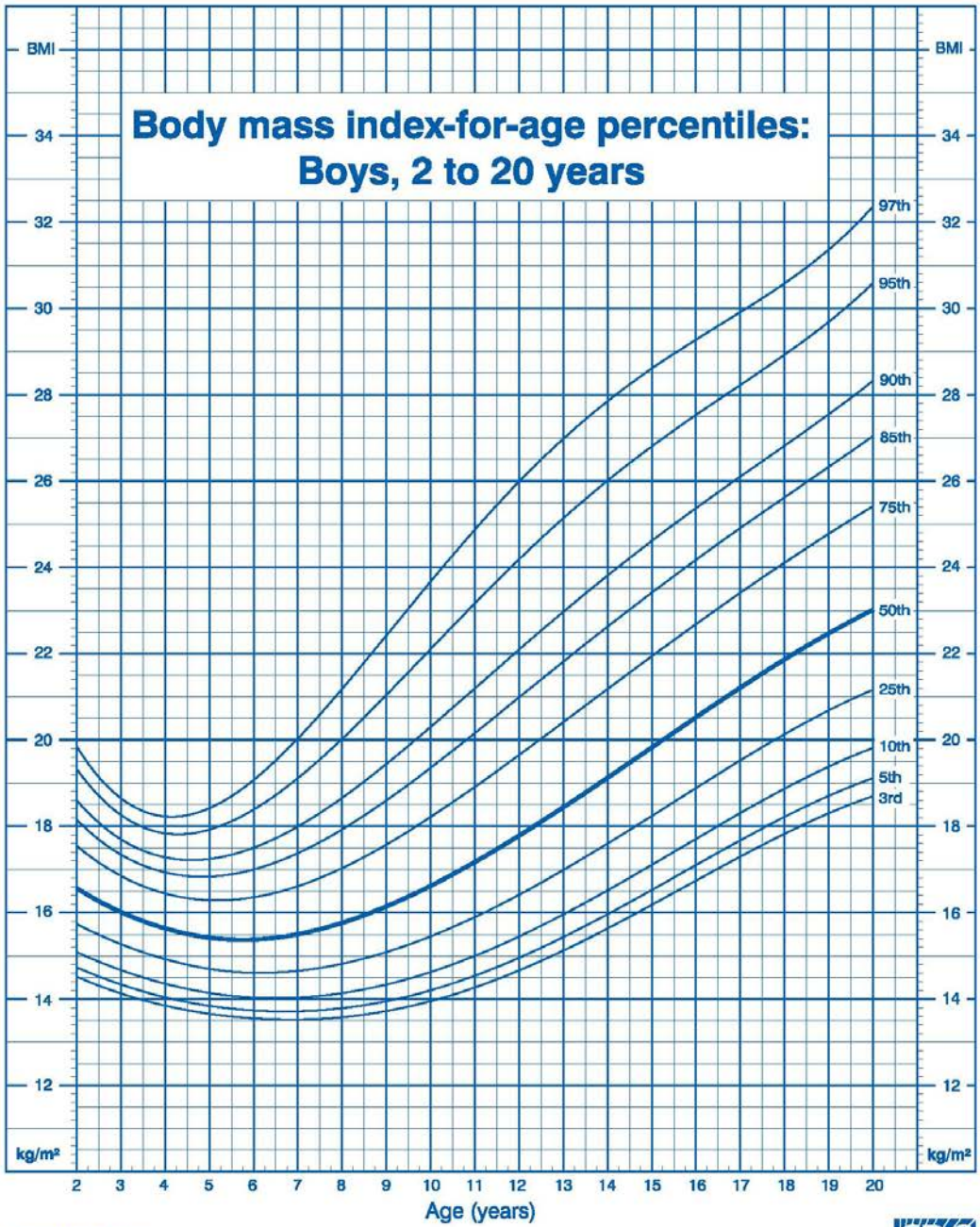


Figure 12. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, 2 to 20 years: Girls stature-for-age



Published May 30, 2000.
SOURCE: Developed by the National Center for Health Statistics in collaboration with the National Center for Chronic Disease Prevention and Health Promotion

SAFER • HEALTHIER • PEOPLE™

Figure 13. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, 2 to 20 years: Boys body mass index-for-age

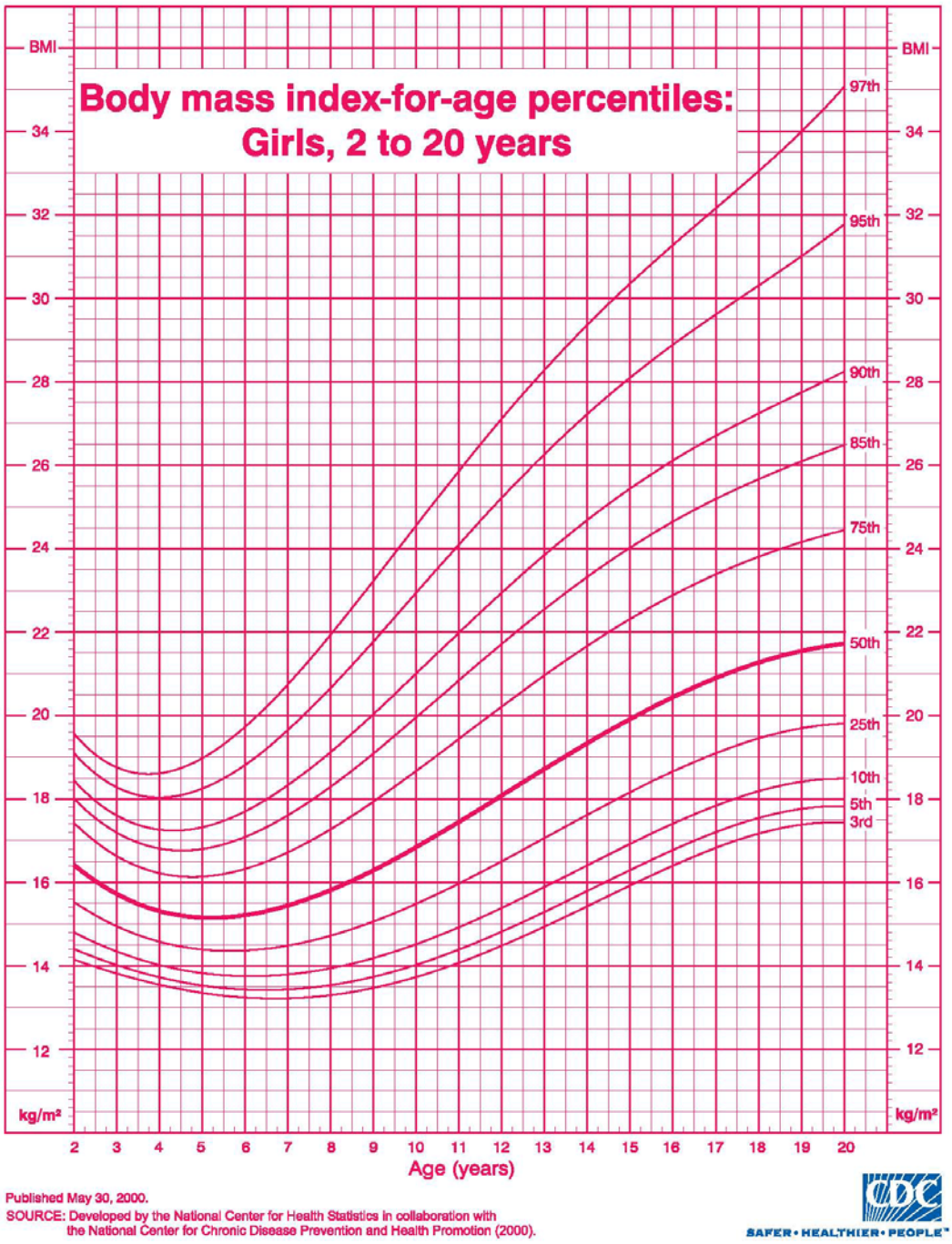


Figure 14. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 85th, 90th, 95th, 97th percentiles, 2 to 20 years: Girls body mass index-for-age

4.2. Cogstate Child and Adolescent Normative Data

The Cogstate normative data from healthy children and adolescents aged between 10 and 20 years are provide by Cogstate and displayed below.

Cogstate Normative Data for Children and Adolescent Subjects

Age (years)	Detection – Speed		Identification – Speed		One Back – Speed	
	Mean	SD	Mean	SD	Mean	SD
10	2.57	0.084	2.77	0.075	2.95	0.095
11	2.54	0.081	2.74	0.074	2.92	0.089
12	2.54	0.081	2.72	0.075	2.90	0.091
13	2.52	0.078	2.70	0.071	2.88	0.088
14	2.52	0.081	2.69	0.074	2.87	0.090
15	2.51	0.078	2.68	0.071	2.86	0.089
16	2.51	0.077	2.68	0.069	2.85	0.090
17	2.50	0.075	2.67	0.066	2.84	0.090
18	2.51	0.072	2.67	0.066	2.84	0.089
19	2.50	0.066	2.67	0.060	2.83	0.085
20	2.50	0.069	2.66	0.060	2.82	0.089

Age (years)	One Card Learning – Accuracy		One Back − Accuracy	
	Mean	SD	Mean	SD
10	0.98	0.098	1.28	0.18
11	0.98	0.096	1.30	0.17
12	0.98	0.092	1.31	0.16
13	0.98	0.089	1.32	0.15
14	0.98	0.088	1.32	0.16
15	0.98	0.088	1.33	0.16
16	0.99	0.088	1.34	0.15
17	0.99	0.090	1.34	0.15
18	1.00	0.089	1.35	0.15
19	1.01	0.086	1.37	0.14
20	1.02	0.090	1.38	0.13

4.3. Clinical Laboratory Tests

Blood Chemistry Tests
aspartate aminotransferase
alanine aminotransferase
albumin
alkaline phosphatase
bicarbonate
blood urea nitrogen
calcium
chloride
creatinine
gamma-glutamyl transferase
phosphorous

potassium
sodium
total bilirubin[a]
total protein
lactate dehydrogenase
fasting (non-random) triglycerides
fasting (non-random) serum cholesterol
fasting (non-random) serum high-density lipoprotein cholesterol
fasting (non-random) serum low-density lipoprotein cholesterol
fasting glucose
prolactin
whole blood hemoglobin A1c
creatine phosphokinase
Endocrine Tests
free thyroxine test
thyroid-stimulating hormone test
Hormonal Parameters
follicle stimulating hormone (female subjects only)
luteinizing hormone (female subjects only)
testosterone (male subjects only)
estradiol (female subjects only)
serum human chorionic gonadotropin (β-hCG) (female subjects ≥ 11 years of age only)
urine human chorionic gonadotropin (β-hCG) (female subjects ≥ 11 years of age only)
Hematology Tests
white blood cell count
white blood cell differential
eosinophilic leukocyte count
basophilic leukocyte count
neutrophil count
lymphocyte count
monocyte count
platelet count
hemoglobin
blood hematocrit
red blood cell count
red cell distribution width
red blood cell indices:
mean corpuscular volume
mean corpuscular hemoglobin concentration
mean corpuscular hemoglobin
Urinalysis Tests
color
appearance
total ketones
urobilinogen
bilirubin
red blood cells
leukocyte esterase
nitrite
pH
protein
specific gravity

glucose	
microscopic evaluation[b]	
Urine Drug Screen	
amphetamines	
benzodiazepines	
barbiturates	
cocaine	
tetrahydrocannabinol	
ethanol	
methadone	
methamphetamine	
opiates	
phencyclidine	

4.4. Benzodiazepine Equivalence Table Conversion Factors to Convert Benzodiazepine Dose to Oral Equivalent Doses Standardized to Lorazepam

Benzodiazepine	Comparative dose	Benzodiazepine	Comparative dose
Alprazolam	0.5	Ketazolam	7.5
Amobarbital *	50	Loflazepate *	1.67
Barbital *	75	Lorazepam	1
Bromazepam	3	Lormetazepam *	1
Bromvalerylurea *	500	Medazepam *	10
Brotizolam *	0.5	Mexazolam *	1.67
Butoctamide *	500	Midazolam **	5 – 7.5
Chloral hydrate *	250	Nimetazepam *	5
Clobazam *	10	Nitrazepam	2.5
Clonazepam	0.25	Nordiazepam	5
Clorazepate	10	Oxazepam	15
Chlordiazepoxide	25	Oxazolam *	20
Cinolazepam **	40	Passiflora *	100
Clotiazepam *	10	Pentobarbital *	50
Cloxazolam *	1.5	Phenazepam	0.5
Delorazepam	1	Phenobarbital *	15
Diazepam	5	Prazepam	10
Estazolam	1	Quazepam	7.5
Etizolam *	1	Rilmazafone *	2
Gidazepam	50		

Fludiazepam *	0.5	Secobarbital *	50
Flunitrazepam *	1	Tandospirone *	25
Flurazepam	15	Tofisopam *	125
Flutazolam *	15	Temazepam	10
Flutoprazepam *	1.67	Triazolam	0.25
Halazepam	40	Zolpidem *	10
Haloxazolam *	5	Zopiclone *	7.5

From: Virani AS, Bezchlibnyk-Butler KZ, Jeffries JJ. *Clinical handbook of psychotropic drugs.* 18th revised version. Ashland, OH: Hogrefe & Huber Publishers; 2009.
* *Japanese Journal of Clinical Psychophamacology 9: 1446, 2006.*
** *Sostmann HJ, Sostmann H, Crevoisier C, Bircher J (1989). "Dose equivalence of midazolam and triazolam. A psychometric study based on flicker sensitivity, reaction time and digit symbol substitution test". Eur. J. Clin. Pharmacol. 36 (2): 181–7. PMID 2721543.*
Benzodiazepine equivalence table; V1.6 24-Feb-2012

4.5. Selected Adverse Events with Special Interest

4.5.1. Extrapyramidal Symptoms (EPS)-Related AEs

BRADYKINESIA	PARKINSONISM
DROOLING	PSYCHOMOTOR RETARDATION
DYSTONIA	TORTICOLLIS
GLABELLAR REFLEX ABNORMAL	TRISMUS
HYPOKINESIA	PSYCHOMOTOR HYPERACTIVITY
OCULOGYRIC CRISIS	AKATHISIA
OROMANDIBULAR DYSTONIA	DYSKINESIA
EXTRAPYRAMIDAL DISORDER	TARDIVE DYSKINESIA
SALIVARY HYPERSECRETION	

4.5.2. Metabolic-Related AEs

BLOOD CHOLESTEROL INCREASED	HYPERCHOLESTEROLAEMIA
BLOOD GLUCOSE INCREASED	HYPERGLYCAEMIA
BLOOD INSULIN INCREASED	HYPERINSULINAEMIA
BLOOD TRIGLYCERIDES INCREASED	HYPERLIPIDAEMIA
DIABETES MELLITUS	HYPERTRIGLYCERIDAEMIA
DYSLIPIDAEMIA	IMPAIRED FASTING GLUCOSE
GLUCOSE TOLERANCE IMPAIRED	LOW DENSITY LIPOPROTEIN INCREASED
GLYCOSURIA	METABOLIC SYNDROME
GLUCOSE URINE PRESENT	TYPE 2 DIABETES MELLITUS

GLYCOSYLATED HAEMOGLOBIN INCREASED	TYPE 1 DIABETES MELLITUS
HIGH DENSITY LIPOPROTEIN DECREASED	WEIGHT INCREASED
WEIGHT DECREASED	LOW BLOOD SUGAR

4.5.3. Hypersensitivity Related AEs

ANAPHYLACTIC SHOCK	LIP SWELLING
ANGIOEDEMA	LIP OEDEMA
BUTTERFLY RASH	OEDEMA PERIPHERAL
CHEST DISCOMFORT	PHOTOSENSITIVITY ALLERGIC REACTION
CONJUNCTIVITIS	PRURITUS
DERMATITIS ALLERGIC	PRURITUS GENERALISED
DYSPNOEA	RASH
EYE PRURITUS	RASH GENERALIZED
FACE OEDEMA	RASH PRURITIC
GLOSSODYNIA	RASH PUSTULAR
HYPERSENSITIVITY	URTICARIA
MOUTH ULCERATION	WHEEZING
PRURITUS GENITAL	SWOLLEN TONGUE
	PHOTOSENSITIVITY REACTION

4.5.4. Suicidality and Self-Injury Related AEs

• COMPLETED SUICIDE	
COMPLETED SUICIDE	
• SUICIDE ATTEMPT	
INTENTIONAL SELF-INJURY	INTENTIONAL OVERDOSE
SELF INJURIOUS BEHAVIOUR	MULTIPLE DRUG OVERDOSE INTENTIONAL
SUICIDAL BEHAVIOUR	POISONING DELIBERATE
SUICIDE THREAT	INTENTIONAL DRUG MISUSE
SUICIDE ATTEMPT	
• SUICIDAL IDEATION	
DEPRESSION SUICIDAL	SELF-INJURIOUS IDEATION
SUICIDAL IDEATION	

4.5.5. Hyperactivity Related AEs

High-level Term		
Increased Physical Activity Levels	Increased Physical Activity Levels	Attention Deficit and Disruptive Behavior Disorders
Preferred Term		
Psychomotor Hyperactivity	Restlessness	Attention Deficit/Hyperactivity Disorder
Lower Level Term		
Activity motor exaggerated	Activity motor exaggerated	Marked restlessness
Behavior hyperactive	ADHD, combined type	Muscular unrest
Behaviour hyperactive	ADHD, predominantly hyperactive-impulsive type	Physically restless
Hyper	Attention deficit disorder of childhood with hyperactivity	Psychomotor restlessness
Hyperactive	Hyperactive child syndrome	Restless
Hyperactivity	Hyperactivity syndrome aggravated	Restlessness
Increased activity	Hyperkinesis of childhood with developmental delay	Restlessness (drug related)
Irritable hyperkinesis	Hyperkinetic conduct disorder of childhood	Restlessness aggravated
Motor activity exaggerated	Hyperkinetic syndrome	Restlessness marked
Muscular hyperactivity	Hyperkinetic syndrome of childhood	Unrest
Overactive	Other specified manifestations of hyperkinetic syndrome of childhood	
Overactivity	Overactive child	
Pedaling	Syndrome hyperkinetic	
Psychomotor agitation	Unspecified hyperkinetic syndrome of childhood	
Psychomotor excitability		
Psychomotor hyperactivity		
Source: Lurasidone IND 61292 SN0434, Communicated with FDA on 10/24/2014		