A MULTICENTER, OPEN-LABEL STUDY OF SEBELIPASE ALFA IN PATIENTS WITH LYSOSOMAL ACID LIPASE DEFICIENCY

Unique Protocol ID:	**LAL-CL06**
NCT Number	**NCT02112994**
EudraCT Number	**2011-004287-30**
Date of Statistical Analysis Plan	**17 November 2016**

Alexion Pharmaceuticals, Inc.

STATISTICAL ANALYSIS PLAN

PROTOCOL NUMBER: **LAL-CL06**

A MULTICENTER, OPEN LABEL STUDY OF SEBELIPASE ALFA IN PATIENTS WITH LYSOSOMAL ACID LIPASE DEFICIENCY

Author: PPD

Date: November 17, 2016

Version: Final 2.0

1. APPROVAL SIGNATURES

PPD

17 NOV 2016
Date dd mmm yyyy

17NOV 2016
Date dd mmm yyyy

17 NOV 2016
Date dd mmm yyyy

2. TABLE OF CONTENTS, LIST OF TABLES, AND LIST OF FIGURES

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

3. LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Table 1: Abbreviations and acronyms

Abbreviation or acronym	Explanation
ADA	Antidrug Antibodies
AE	Adverse Event
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
ATC	Anatomic Therapeutic Class
AUC	The Area Under the Concentration curve
BLQ	Below the Limit of Quantification
BMI	Body Mass Index
CI	Confidence Interval
CL	Apparent clearance
CLDQ	Chronic Liver Disease Questionnaire
Cmax	The maximum serum level observed over the dose interval
CRF	Case Report Form
CS	Clinically Significant
CSR	Clinical Study Report
DDST	Denver II Developmental Screening Test
DNA	Deoxyribonucleic Acid
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EMA	European Medicines Agency
ERT	Enzyme Replacement Therapy
FACIT	Functional Assessment of Chronic Illness Therapy
FDA	Food and Drug Administration
GGT	Gamma glutamyltransferase
HDL-C	High Density Lipoprotein Cholesterol
HS-CRP	High-Sensitivity C-Reactive Protein
ICH	International Conference on Harmonization
IMP	Investigational Medicinal Product
IAR	Infusion Associated Reaction
INR	International Normalized Ratio
IV	Intravenous
LAL	Lysosomal acid lipase
LDL-C	Low Density Lipoprotein Cholesterol
LLOQ	Low Limit of Quantification
MCS	Mental Component Summary (SF-36)

Abbreviation or acronym	Explanation
MedDRA®	Medical Dictionary for Regulatory Activities
MEGE	Multi-Echo Gradient Echo
MRI	Magnetic Resonance Imaging
NCS	Not Clinically Significant
PedsQL	Pediatric Quality of Life Inventory
PK	Pharmacokinetic
PT	Preferred Term
PT (INR)	Prothrombin Time
PPT (INR)	Partial Thromboplastin Time
QOW	Every other week dosing
QTc	QT-interval for ECG corrected for heart rate
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Safety Committee
SFA	Stature-for-Age
SOC	System Organ Class
SOP	Standard Operating Procedure
TEAE	Treatment-Emergent Adverse Event
Tmax	The time to reach the Cmax
UK-MELD	United Kingdom Model for End-Stage Liver Disease
V	Volume of distribution after IV infusion
WFA	Weight-for-Age
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

4. Description of the Protocol

LAL-CL06 is an open-label, multicenter study to evaluate the safety and efficacy of SBC-102 (sebelipase alfa) in a more broad population of subjects with Lysosomal Acid Lipase Deficiency (LAL-D) than have been previously studied.

Following a screening period of up to 45 days, each eligible subject will receive an intravenous (IV) infusion of SBC-102 at an initial dose of 1mg/kg every other week (qow). A dose escalation to 3 mg/kg qow - and subsequently to 3 mg/kg every week (qw) - may be considered for a subject who satisfies the criteria outlined in Section 6.1 of the study protocol. Each subject will receive up to 96 weeks of treatment, then proceed to an expanded treatment period for a maximum of 48 weeks, after which a follow-up phone call will be conducted to assess adverse events (AEs) and concomitant medications approximately 30 days later.

The primary objective of the study is to evaluate the safety and tolerability of SBC-102 in LAL-D subjects who have not been previously studied.

The secondary objectives are:

- to evaluate effects of SBC-102 relative to Baseline assessment of lipid metabolism and liver abnormalities (including histopathology);
- to evaluate the effects of SBC-102 on additional clinical parameters of LAL-D including those not previously well characterized in the literature;
- to evaluate the effects of SBC-102 on growth parameters in pediatric subjects presenting with evidence of growth delay;
- and to evaluate the immunogenicity of SBC-102.

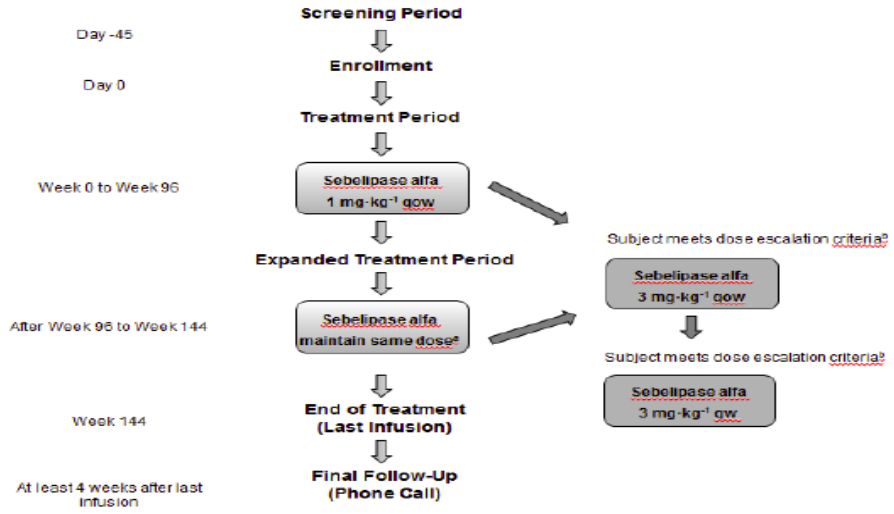
The exploratory objectives are:

- to further characterize the pharmacokinetics (PK) of SBC-102 in pediatric subjects, subjects with severe hepatic dysfunction, and those who have had a previous liver or hematopoietic stem cell transplant;
- and to evaluate the effect of SBC-102 on health-related quality of life (HRQOL) assessments.

An independent program-level safety committee (SC) appointed by the Sponsor will provide additional oversight of subject safety in this study through periodic and ad-hoc reviews of safety data.

A schematic of the study is provided in Figure 1. Further details on target population, inclusion/exclusion criteria, schedule of assessments and study procedures are available in the study protocol.

Figure 1: Study Flow Diagram



[a] Each subject will initiate treatment in the open-label extension period (Week 100) at the same dose of sebelipase alpha they received during the last infusion of the treatment period (Week 96), unless criteria for dose modification are met (see Section 6.1 of Study Protocol).
[b] Refer to Section 6.1 of Study Protocol for dose escalation criteria.

4.1. Changes from Analyses Specified in the Protocol

The following items are specified in the protocol and are not part of this analysis plan:

- The urinalysis summary of the change from baseline will not be summarized. However urinalysis results will be provided in the subject listings.
- The summary of the change in vital signs during and after infusion relative to pre-infusion values will not be summarized; instead, the change from baseline (defined below) to each post-baseline visit will be summarized. Given that multiple vital sign measurements are taken at each post-baseline visit, the average of the measurements within a visit will be summarized.
- At the time of this Statistical Analysis Plan (SAP) amendment, only 2 subjects have been anti-drug antibody (ADA) positive. As a result, a clarification was added that many of the the ADA analyses will only be performed in the event that at least 5 subjects are ADA positive.
- The effect of sebelipase alpha on growth parameters in pediatric subjects will be evaluated among all subjects younger than 18 years old. The protocol specifies analyzing these data conditional on subjects having had manifestations of impaired growth. However, the analysis will not be limited to subjects showing signs of impaired growth.
- The AE table showing the incidence of AEs across time intervals will not be presented.
- The protocol indicates that percentages will be calculated including subjects with missing values. However, percentages will be based on subjects with non-missing values unless otherwise specified.

4.2. Changes from Analyses Specified in the Previous Version of the SAP

In addition to items described in Section 4.1, Changes from Analyses Specified in the Protocol, the additions, clarifications and a few deletions have been made to the previous version of the SAP.

The following additional summaries have been described in this version of the SAP:

- Additional baseline characteristics are summarized including, among others, a summary of the LAL enzyme activity and DNA results at baseline.
- The number and percentage of subjects on a modified diet at baseline and with changes in diet over the course of the treatment period are summarized.
- The number of weeks on specific doses of sebelipase alpha is summarized.
- A listing of infusion associated reactions (IARs) for those subjects with a change in sebelipase alpha dosing is added.
- Summaries of select laboratory parameters will be presented relative to the upper limit of normal (ULN).
- The summary of the Child-Pugh Status is described.
- Details of the physical examination summaries are provided.
- A shift from baseline in the interpretation of the 12-lead ECGs has been described.
- The percent change from baseline is added as a summary of each endpoint for which a change from baseline was planned.
- The United Kingdom Model for End-Stage Liver Disease (UK-MELD) summary is described.
- A summary of the post-baseline LAL enzyme activity is described.

The following clarifications have been provided to this version of the SAP:

- Medical histories are summarized separately for those identified as prior and concomitant disease.
- The Denver II assessment will be presented using the overall score and not the domain scores.
- Exploratory analyses of the effect of ADAs on the efficacy of SBC-102 will not be explored using stratified analyses. This may be added if at least 5 subjects are ADA positive at the time of the analysis.
- Laboratory abnormalities will not be specified as CS (Clinically Significant) or NCS (Not-Clinically Significant) by the Investigator or designee; instead, clinically significant changes from a subject's baseline value or previous values are recorded as AEs.
- In the appendices, the scoring of the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue was changed to correctly reflect that Items 7 and 8 are not reverse scored, and that the Pediatric Quality of Life Inventory (PedsQL) Total Score and Psychosocial Health Summary Score are not computed if at least 1 scale score is missing.

The following summaries have been removed from this version of the SAP:

- Spaghetti plots of laboratory parameters are replaced with waterfall plots of the percent change from baseline to Week 96. Spaghetti plots may be provided.
- For anthropometric measures in which results are compared to population norms using z-scores and percentiles, only the percentiles will be summarized.

5. DEFINITIONS

5.1. Safety

5.1.1. Primary Endpoint(s)

This study is designed primarily to evaluate the safety of IV infusion of SBC-102. Safety will be assessed based on: incidence and severity of TEAEs, serious adverse events (SAEs), and IARs; changes from baseline through trial completion in vital signs, Electrocardiogram (ECG), and routine clinical laboratory tests (chemistry, hematology, urinalysis); physical examination findings; use of concomitant medications/treatments; functional and overall development measured using the Denver II Development Screening Test; and incidence/titer of antibodies to SBC-102.

5.1.1.1. Adverse Events (AEs), Serious Adverse Events (SAEs) and Infusion-Associated Reactions (IARs)

See Section 7.1 of study protocol for definitions.

5.1.1.2. Vital Signs, Height and Weight

Vital signs – including pulse rate, respiratory rate, systolic and diastolic blood pressure, and body temperature – will be obtained at the time points specified in the schedule of assessments (Appendix B of study protocol). On dosing days, vital signs will be recorded pre-infusion, every 30 minutes (±10 min) during the infusion and every 30 minutes (±10 min) from 0 to 2 hours after the end of the infusion.

Body weight and height will be measured, prior to SBC-102 infusion at study at the time points specified in the schedule of assessments (Appendix B of the study protocol). Height is measured in inches or centimeters, and weight is recorded in pounds or kilograms. Both will be reported in metric units.

5.1.1.3. Physical Examination

A complete, age-appropriate physical examination will be performed by the Investigator or qualified designee, at the time points specified in the schedule of assessments. The examination will include an assessment of the subject's general appearance, skin, head (including head circumference for children up to 3 years of age), eyes, ears, nose, throat, heart, lungs, abdomen, extremities/joints, and neurological status. Abnormal findings will be recorded in the electronic Case Report Form (eCRF).

Every physical examination will also include the following:

- Child-Pugh Score (see Appendix A of Study Protocol);
- Abdominal photograph for subjects < 4 years of age at Screening;
- Liver size: A clinical assessment of liver size (palpable/non palpable and centimeters below costal margin), regularity (smooth/nodular), and sensitivity (tender/nontender);
- Spleen size: A clinical assessment of spleen size (palpable/nonpalpable and centimeters below costal margin), regularity (smooth/nodular), and sensitivity (tender/nontender);
- Lymphadenopathy: An assessment of the size, location, and character of any palpable lymph nodes. Areas to be examined include cephalic (occipital, preauricular, postauricular, submental, submandibular), cervical, clavicular, axillary, and inguinal. Any enlarged nodes will be characterized as tender or nontender;

- Arterial disease: A clinical assessment of the right and left posterior tibialis and dorsalis pedis pulses and carotid bruits;
- Skin manifestations: A clinical assessment of signs of liver disease and portosystemic anastomoses such as periumbilical venous engorgement (caput medusae), spider nevia, or gynecomastia, dyslipidemia such as xanthomas (tendinous, tuberous) and xanthelasma, and allergies including assessment of allergic skin rashes.

5.1.1.4. Laboratory Assessments

Blood and urine samples for clinical laboratory tests will be collected at the time points indicated in the schedule of assessments. Table 2 lists the laboratory panels and associated tests to be performed on blood (or urine) samples. Note that not all tests are performed at the same time points due to limitations on blood volume collection that is considered acceptable in young children with very small total circulating blood volumes (see Section 5.9 of study protocol for further details).

A central laboratory will be responsible for analysis of all laboratory tests, with the exception of prothrombin time (international normalized ratio) [PT(INR)], activated partial thromboplastin time (aPTT), urinalysis, and urine pregnancy.

Clinically significant changes from a subject's baseline value or previous values will be recorded as AEs.

Table 2:Clinical Laboratory Tests

Laboratory Panel	Tests
Hematology	White blood cell count; red blood cell count; hemoglobin; hematocrit; mean corpuscular volume, hemoglobin, and hemoglobin concentration (MCV, MCH, MCHC); platelet count; neutrophil; lymphocytes; monocytes; eosinophils; basophils; peripheral smear for examination of cell morphology
Liver Function Tests	ALT, AST, alkaline phosphatase, GGT, albumin, bilirubin (direct, indirect, total)
Serum Lipids	LDL-C, total cholesterol non-HDL-C, triglyceride, HDL-C
Other Chemistries	Serum electrolytes (sodium, potassium, chloride, calcium, magnesium, phosphorus), glucose, creatinine, bicarbonate, total protein, blood urea nitrogen, high-sensitivity C-reactive protein
	HbA1c
Macrophage Activation Markers	serum chitotriosidase, serum ferritin, hs-CRP , CK-18
Urinalysis[1]	pH, clarity, color, specific gravity, glucose, ketones, blood, protein, nitrite, and leukocytes (microscopic examination will only be done if urinalysis is positive for blood, nitrite, or leukocytes, or if protein is >1+)
Viral Hepatitis Screen	HBsAg and HCV serology
Alcohol Use	CDT
Coagulation Parameters[1]	PT (INR), aPTT
Pregnancy[2]	Serum hCG at Screening and urine hCG1 at all other visits
Anti-drug Antibody	Anti-sebelipase alfa antibody
Exploratory Biomarkers	See Section 5.9.5 of Study Protocol
LAL Enzyme Activity	Dried blood spot. See Section 5.9.6 of Study Protocol
DNA Sample	See Section 5.9.7 of Study Protocol
PK Samples	See Section 5.9.4 of Study Protocol
Malnutrition Markers and Fat Soluble Vitamins	See Laboratory Manual for details
1 Test performed by a local laboratory. 2 Performed for female subjects of childbearing potential only.	

5.1.1.5. Concomitant Medications and Treatments

Prior medications are defined as medications taken or treatments received by subjects prior to the first dose of SBC-102. Concomitant medications are defined as medications taken or treatments received by subjects during the study after first dose of SBC-102.

Concomitant medications include prescription and over-the-counter medications, herbal medications, prophylactic and therapeutic vaccines, vitamins, and dietary supplements. Concomitant treatments include diagnostic, palliative, or interventional procedures (e.g., lipid-lowering diet, surgery, physical therapy). Information on all concomitant medications and treatments will be recorded in the eCRF and will include the name of the medication (brand or generic) or therapy, reason for use, start date, stop date, dose, route of administration (if applicable), and frequency of administration.

5.1.1.6. Electrocardiogram (ECG)

Age-appropriate 12-lead ECG measurements will be obtained at the time points specified in the schedule of assessments. ECGs will be reviewed by a qualified clinician, and any abnormalities will be specified as CS or NCS.

5.1.1.7. Anti-drug Antibodies (ADAs)

Blood samples will be collected to test for antibodies to SBC-102 in plasma just prior to dosing, at Baseline and every 4 weeks until Week 12. After Week 12, ADA samples will be collected every 12 weeks and at end of study (or early withdrawal) visit, as well as after an IAR.

5.1.1.8. Denver II Developmental Screening Test (DDST)

The Denver II Developmental Screening Test (DDST) will be administered at the time points specified in the schedule of assessments.

The test is a standardized measure to assess development in children from 1 month to 6 years of age (Frankenberg et al., 1992). It includes performance-based and parent-reported items in 4 functional areas: fine motor-adaptive, gross motor, personal-social, and language skills. The test was normed on a diverse sample of children who were full term and had no obvious developmental disabilities; the norms indicate when 25%, 50%, 75%, and 90% of children passed each item. The instrument has good inter-rater and test-retest reliability (correlations ≥ 0.90 for most tests).

The test must be administered by a trained clinician and takes an average of 10 to 20 minutes to complete, and up to an hour depending on the age of the child and number of assessments completed. Administration and scoring of DDST is based upon the child's age.

Refer to the Study Operations Manual (SOM) for further information on administration and scoring of DDST.

5.2. Efficacy

5.2.1. Secondary Endpoint(s)

Efficacy endpoints will include absolute changes or percent changes over time from Baseline as follows: Decrease in Alanine Aminotransferase (ALT); Decrease in Aspartate Aminotransferase (AST); Decrease in Low Density Lipoprotein Cholesterol (LDL-C); Increase in High Density Lipoprotein Cholesterol (HDL-C); Decrease in non-HDL-C; Decrease in triglyceride; Decrease in Child-Pugh status for subjects with Child-Pugh class C or B at Baseline; Decrease in United Kingdom Model for End-Stage Liver Disease (UK-MELD) score; Improvement in liver histopathology; Decrease in liver and spleen volume by magnetic resonance imaging (MRI); Decrease in liver fat fraction by MRI.

Other efficacy endpoints will be determined as follows:

- For Subjects ≤24 months of age, z-scores and percentile scores based on World Health Organization (WHO) growth charts (WHO Multicentre Growth Reference Study Group, 2006 and 2007) will be determined for the following parameters:
  - Weight-for-age (WFA);
  - Weight-for-length;
  - Length-for-age;
  - Body Mass Index (BMI)-for-age;
  - Head circumference-for-age.
- For subjects >24 months of age to 18 years, z-scores and percentile scores based on Centers for Disease Control and Prevention (CDC) growth charts (Kuczmarski et al., 2000) will be determined for the following parameters:
  - WFA;
  - Weight-for-stature (weight-for-height);
  - Stature-for-age (SFA);
  - BMI-for-age.
- For all subjects <18 years, the proportion of subjects who meet criteria for under nutrition (underweight, wasting, and stunting) based on WFA, weight-for-length/weight-for-stature, and length-for-age/SFA, respectively (UNICEF, 2009), and combinations of these indicators, will be determined.
- Additional clinical, biochemical and hematological abnormalities including: total and conjugated bilirubin; Gamma glutamyltransferase (GGT); hemoglobin level; platelet count; and markers of macrophage activation, including absolute reductions in serum ferritin, serum chitotriosidase and high-sensitivity C-reactive protein (hs-CRP).
- Changes from baseline in scores from different HRQOL instruments briefly described in Section 5.2.1.3 below.

5.2.1.1. ALT, AST, LDL-C, non HDL-C, Triglycerides and HDL-C

Blood samples to determine ALT, AST, non HDL-C, triglycerides and HDL-C are collected as part of clinical laboratory assessments, as specified in Table 2, and at timepoints specified in the schedule of assessments.

5.2.1.2. Child-Pugh Score

See Appendix A of the study protocol.

5.2.1.3. Health-Related Quality of Life (HRQOL)

HRQOL questionnaires will be completed at the time points specified in the schedule of assessments for subjects who are ≥5 years of age. The specific HRQOL questionnaires completed by each subject will be based on the respondent's age on the date that informed consent is obtained as indicated in Table 3. The questionnaires will be administered prior to any other study procedures being conducted at that visit.

The 13-item Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue scale was developed to measure levels of fatigue in people living with a chronic disease. In this study, the FACIT-Fatigue scale version 4 will be self-administered by all subjects who are ≥17 years of age at informed consent.

The Chronic Liver Disease Questionnaire (CLDQ) is a disease-specific instrument designed to assess HRQOL in subjects with chronic liver disease (Younossi et al., 1999). The CLDQ will be self-administered by all subjects who are ≥17 years of age at informed consent.

The Pediatric Quality of Life Inventory TM (PedsQL™) is composed of generic core scales and disease-specific modules. The 23-item PedsQL™ 4.0 Generic Core Scales was designed to measure the core dimensions of health, as delineated by the World Health Organization, as well as role (school) functioning in healthy children and those with acute or chronic health conditions. The PedsQL™ Generic Core Scales includes 4 multidimensional scales of physical functioning (8 items), emotional functioning (5 items), social functioning (5 items) and school functioning (5 items). In addition to the total scale score (all 23 items), two summary scores, the Physical Health Summary (8 items) and Psychosocial Health Summary (15 items), are also reported. In this study, the PedsQL™ 4.0 Generic Core Scales will be self-administered by subjects who are 5 to <18 years of age on the date of informed consent, using one of the three self-report forms (ages 5-7, 8-12, or 13-18), as appropriate to the subject's age (Varni et al., 2009). Parent proxy reports will not be used in this study.

Sample questionnaires and detailed instructions on administration of HRQOL questionnaires are provided in the SOM. Also see Section 9.4.9 for details on derivation of summary scores from the HRQOL instruments.

Table 3: Health-Related Quality of Life Assessments

Patient Age	FACIT-Fatigue	CLDQ	PedsQL™ Generic Core Scales
5 to <17 years[1]			X
17 to 18 years[1]	X	X	X
>18 years[1]	X	X	

[1] Based on the patient's age on the date that informed consent is obtained.

5.3. Pharmacokinetic Endpoints

Blood samples for determination of SBC-102 serum concentrations will be collected from subjects <18 years of age, from subjects with severely impaired liver functioning (e.g. based on imaging, biopsy or Child-Pugh Class C status at Screening), and for subjects who have previously undergone liver or hematopoietic stem-cell transplantation. Sampling will be completed - as blood volume permits - at Baseline, Week 24 and Week 48, and on the day of a dose increase, as applicable. Details on sampling time and sampling strategy are described in the study protocol (Section 5.9.8).

PK parameter estimates to be reported may include: serum clearance and apparent volume of distribution; Area Under the Concentration Curve (AUC), maximum observed concentration (Cmax), time to maximum observed concentration (Tmax), and terminal elimination half-life (t1/2). Details of pharmacokinetic analyses will be described in a separate analysis plan.

Further details will be provided in a separate analysis plan.

5.4. Exploratory Biomarker Assessment

Blood samples (for serum isolation) will be obtained at the time points specified in the schedule of assessments, where local regulations and blood volume limitations permit, for exploratory analyses to identify and evaluate disease-related biomarkers. The intent is to investigate baseline disease and dynamic markers that will enable the Sponsor to better understand the pathogenesis of LAL-D and related comorbidities and response to SBC-102 therapy. Given the rarity of LAL-D and the paucity of information on disease characteristics, the definitive list of analytes remains to be determined.

No biomarkers are planned for analysis at this time.

6. DATA SETS ANALYZED (STUDY POPULATIONS)

Safety will be examined for the Full Analysis Set overall and, as sample size permits, adverse events will besummarized by age. Efficacy will be examined for the Full Analysis Set and, as sample size permits, summarized by age, gender, baseline use of lipid-lowering medication, genetic mutation category, baseline fibrosis or cirrhosis status, baseline ALT (>1.5 x ULN and <=1.5 x ULN), and other potential subgroups of interest at the time of analysis. Summaries of HRQOL may be summarized by subgroups.

6.1. Full Analysis Set (FAS)

This analysis set will include all subjects who received at least 1 infusion of SBC-102. This analysis set will be used for analysis of safety, efficacy, and subject health outcomes data.

7. STATISTICAL ANALYSIS

Alexion or designee will be responsible for data collection, reviewing and validating all the information in the eCRFs, statistical analysis, and generation of the clinical report. The final statistical analysis will not begin until the entire database is locked and signed off.

The Alexion Biostatistics Department or designee will perform the statistical analysis of the data derived from this trial. The analysis will be performed using the SAS® statistical software system Version 9.2 or higher.

No formal inferential statistical testing will be performed. As appropriate, 95% two-sided confidence intervals (CIs) will be calculated around the estimates based on the exact binomial distribution (Clopper-Pearson method) for binomial endpoints and the t-distribution for continuous endpoints. P-values and 95% confidence intervals, where presented, will be considered descriptive and will be provided to facilitate clinical review and interpretation.

All data collected in this study will be provided in subject data listings sorted by subject number; the dose of SBC-102 temporally associated with the data will be included on the listing. Summary tables and/or graphs will be presented for each endpoint, as appropriate to the data, by evaluation time point. Visits will be summarized as collected, and data collected from unscheduled visits will be included in subject listings.Unless otherwise noted, the following standard conventions will be used for creating descriptive summaries:

- Continuous numeric endpoints will be summarized by providing the number of subjects with nonmissing data, the mean and standard deviation (SD) of the data, and the minimum, first quartile (Q1), median, third quartile (Q3), and maximum value;
- For categorical endpoints, the number and percentage of subjects with each possible outcome will be displayed. The denominator for percentages will not include subjects with missing data;
- Mean, median values, Q1 and Q3 will be presented with one decimal point more than the original value, SD will be displayed with two decimal points more than the original value and minimum/maximum values will have the same number of decimal points as the original value;
- In frequency tables the percentage will be displayed with no decimal place.

7.1. Study Subjects

7.1.1. Disposition of Subjects

Subject disposition will be presented in a listing, and will include age, gender, date of consent, date and dose of first infusion of study drug, date and dose of last infusion of study drug, date of completion or premature discontinuation from the study, and reason for discontinuation if the subject discontinues prematurely. Data from all screened subjects will be included in the summary of subject disposition. The frequency and percentage of subjects who are treated and not-treated in the study, continuing treatment at the time of the analysis (for Week 96 analysis only), completed the study, and discontinued from the study, along with reasons for discontinuation, will be summarized. In addition, the number of subjects with at least one infusion during the treatment period (Week 0 to Week 96), and the number of subjects with at least one infusion in the extended treatment period (Week 97 to Week 144) will be presented.

7.1.2. Protocol Deviations

Protocol deviations will be listed by subject, as applicable.

7.1.3. Demographics and Medical History

All demographic and medical history information will be summarized using the FAS set.

7.1.3.1. Demographics

The following demographic variables will be summarized:

- Age (Years) reported to 1 decimal place (i.e. xx.x years), <12, ≥12 years, and 2-4 years
- Sex
- Race
- Ethnicity
- If Asian, Japanese (yes/no), and generation (first, second, third)
- Birth weight (kg), for subjects <4 years old at enrollment

By-subject listings of demographic information will be produced.

7.1.3.2. Baseline Characteristics

Summaries of baseline height (cm), weight (kg), and BMI (kg/m^2) will be presented, as well as disease history. Disease history will include time since diagnosis (years), method of diagnosis, and whether any family members have been diagnosed with LAL-D.

7.1.3.3. Medical History

Baseline medical history information, i.e., number (%) of subjects who have a medical history will be summarized for the FAS. Summaries will be presented as prior disease and concomitant disease, with the prior and concomitant status determined by whether the condition was active (yes/no) at the time of screening. By subject listings will be created for medical history which will include medical history condition, start and end date, and if ongoing.

7.1.3.4. Baseline LAL Enzyme Activity and DNA Results

The number and percentage of subjects with LIPA gene sequencing results consistent with LAL-D and with fibroblast results consistent with LAL-D will be presented. In addition, the number and percentage of subjects homozygous for common mutation, heterozygous for common mutation, with another mutation, or no variants found will be presented. LAL Enzyme Activity at baseline will be categorized as affected (<=0.016 nmol/punch/hr), carrier (0.017 to 0.026 nmol/punch/hr); and unaffected (0.027 to 0.0152 nmol/punch/hr) are summarized.

7.1.4. Prior and Concomitant Medications/Treatments

Medications will be coded using the World Health Organization Drug Dictionary version in use by Alexion at the time of analysis. Summaries will be performed on the FAS.

The number (%) of subjects using prior and concomitant medications will be summarized for the World Health Organization Anatomical Therapeutic Chemical (WHO ATC) Class code and generic name. Subjects with multiple medications in the same class code will only contribute once to the number and percentage of subjects with a medication in that class; likewise, subjects with multiple entries for the same medication will only contribute once to the summary of the generic name.

Similar tables will be provided for lipid-lowering medications. However, in this case, prior lipid-lowering medications will be limited to medications administered in the period 6 weeks prior to the first study drug infusion.

In addition, the number of subjects using any lipid-lowering medication at baseline and post-baseline will be summarized along with a summary of whether post-baseline use of LLM was accompanied with a change in dosing (yes/no).

All prior and concomitant medications and treatments (pharmacological and non-pharmacological) will be provided in a data listing. The data listing will include start and end dates (or indication of ongoing), dose, unit, frequency, route, and an indication of whether the medication usage is prior, concomitant or both.

7.1.5. Modified Diet

The number and percentage of subjects on any low cholesterol or low saturated fat diet at baseline will be summarized along with the number and percentage of subjects on reduced consumption of fatty dairy products, non-dairy animal-derived fatty foods, and/or increased consumption of cholesterol lowering foods.

7.2. Safety Analyses

No formal hypothesis testing is planned. Baseline is defined as the last available assessment prior to the first infusion of SBC-102. In case of multiple pre-treatment measurements, Baseline is the average of the last (up to 3) measurements.

Listings of all safety data will be produced. Results from pregnancy tests will be presented in subject listings.

The analysis of safety and tolerability will include the following data:

- SBC-102 exposure.
- Adverse events.
- Clinical laboratory investigations.
- Vital signs (body temperature, pulse rate, respiratory rate, blood pressure).
- Physical exam.
- 12-Lead ECG.
- Concomitant medications and therapies.
- Antidrug antibody formation.

As sample size permits, safety analyses stratified by ADA status, and for other subgroups of interest might be explored.

7.2.1. SBC-102 Exposure

Number of weeks in the study and number of study drug infusions received will be summarized overall and; any changes in dose will be described in listings, and may also be provided in tabular summaries.

Exposure to study drug (SBC-102) will be summarized overall doses for the following parameters: the total number of infusions attempted, the number of infusions completed both without and with a rate change or infusion interruption and the number of infusions where the entire infusion volume was not administered will be presented. In addition, a summary will be provided showing the number of weeks on specific doses.

A data listing will be presented, sorted by subject within initial dose of SBC-102, providing the dates, total planned and actual volumes of drug infused, infusion rate and duration of the each infusion along with any modification to the infusion.

7.2.2. Adverse Events and Infusion-Associated Reactions

Adverse Events (AE) will be coded by primary system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) (version 15.1 or higher). AEs will be determined as occurring from the signing of informed consent and prior to the initiation of SBC-102 (pre-treatment AEs), or as on/after first dose of SBC-102 (treatment-emergent) as described in Section 9.4.7. Pre-Treatment Adverse Events (PTAEs) will be presented in listings only, and Treatment-Emergent Adverse Events (TEAEs) will be presented in summary tables and listings.

Study displays for TEAEs are described below.

7.2.2.1. Overall Summary of Adverse Events

All TEAEs will be presented using summary statistics (n and n, %), showing the number of events as well as the number and percentage of subjects with the event.

Table 4: Subcategories For Treatment-Emergent Adverse Events (TEAEs)

Relationship
Related AEs
Not Related AEs
Severity
Mild
Moderate
Severe

The relationship between an AE and the study drug will be assessed as: not related; unlikely related; possibly related; or related. Subjects with at least one related AE and subjects with at least one not related AE will be summarized. Related AEs will include those AEs considered 'possibly related' or 'related' by the investigator; and not related AEs will include those AEs considered 'not related' or 'unlikely related'. The severity or intensity of the severity will be assessed and summarized as mild, moderate, or severe.

The number and percentage of subjects with an event meeting these criteria will be summarized according to the most severe occurrence of an AE for summaries of severity or intensity, and the most extreme relationship of the AE for summaries of relationship to study drug. . The number of events will be summarized as reported, with no accounting for the most extreme severity or relationship. Additionally, the number and percentage of subjects who withdrew from the study due to an AE or who died on study will be presented along with the number of events meeting each of these criteria. These statistics will be prepared for all TEAEs and, separately, for SAEs and IARs.

7.2.2.2. AEs, SAEs and IARs by System Organ Class (SOC) and Preferred Term (PT)

The frequencies of TEAEs, SAEs, IARs and percentage of subjects with events will be tabulated by PT within SOC. Subjects are counted once in each SOC and PT. Percentages will be based on the total number of subjects in the specific analysis population. SOCs will be listed in order of frequency of occurrence.

Frequencies will also be presented by the classifications of severity and causality (i.e., relationship to study drug). The incidence of AEs leading to study discontinuation and SAEs leading to discontinuation will also be summarized.

A listing of IARs will be presented for those subjects who had a change in dose. This listing will include the subject number, and the prior dose administered, in addition to the standard adverse event listing elements.

Detailed listings of AEs, SAEs and IARs will be presented. These listings will include severity and relationship to treatment, seriousness, as well as action taken regarding study treatment, other action taken, and subject outcome. A separate listing of subjects who discontinued from the study due to a treatment-emergent AE will also be provided.

7.2.2.3. Deaths

A listing of subject deaths and cause of death will be produced, if applicable.

7.2.3. Laboratory Tests

Observed measurements and changes from Baseline to each study time point in clinical laboratory data will be summarized. Frequencies of results relative to the laboratory normal range as well as shift tables showing the change from baseline to post-baseline relative to the normal range will be summarized, if sufficient data are available for analysis. Summaries will be presented for liver function tests, serum lipids, hematology, serum chemistry, HbA1c, and coagulation parameters as identified in Table 2, Clinical Laboratory Tests, above.

Plots of mean observed values, absolute change from baseline and mean percent change from baseline over time with standard error bars may be prepared for the following:

- Liver function tests: ALT/AST, GGT, total bilirubin, direct bilirubin, indirect bilirubin, albumin, alkaline phosphatase
- Lipids: total cholesterol, LDL-C, non-HDL-C, HDL-C, triglycerides
- Hematology: hemoglobin, platelets

Also see Section 7.3.2 Clinical Laboratory Tests as described in the Efficacy Analyses section.

7.2.4. Child-Pugh Status

Using the Total Child-Pugh score, subjects will be classified into categories A (score 5 to 6), B (score 7 to 9), and C (score 10 to 15). Higher scores and higher categories represent a worse outcome.

Total Child-Pugh scores will be summarized as a continuous endpoint showing the observed values at baseline and each post-baseline visit, as well as the change from baseline and the percent change from baseline to each post-baseline visit. This summary will be presented for those subjects with a Class B or C at baseline as well as for all subjects in the FAS.

The number and percentage of subjects in each of the categories (A, B, and C) will be summarized for each visit at which the test was performed. In addition a shift from baseline will be presented showing improvement, no change, and decline. Generally, percentages will be based on the number of subjects with baseline and post-baseline results. However, percentages for shifts of improvement (i.e., B to A, C to B, and C to A) will be based on subjects who had scores of B or C at baseline. Similarly, percentages for shifts of worsening (i.e., A to B, A to C, and B to C) will be based on subjects who had scores of A or B at baseline.

7.2.5. Physical Examination

A summary of findings including liver size, spleen size, lymphadenopathy and arterial disease at Screening and follow-up will be presented for baseline and each post-baseline visit.

Summaries of liver and spleen size will include the percentage of subjects with palpable and non-palpable organ, the regularity (smooth or nodular), and the sensitivity (tender or non-tender).

Summaries of lymphadenopathy will include whether any lymphadenopathy is present (yes/no) and the presence/absence and tender/non-tender status of each of five specific regions: cephalic, cervical, clavicular, axillary and inguinal.

Summaries of arterial disease will include the number and percentages of subjects with a pulse result (not palpable, faint, diminished, normal, bounding) in each of four categories (left posterior tibialis, right posterior tibialis, left dorsalis pedis, and right dorsalis pedis. Results will be presented for baseline and each post-baseline visit.

7.2.6. Vital Signs

Observed values, change from baseline, and percent change from baseline will be presented for the following vital sign parameters:

- Body temperature.
- Pulse rate.
- Blood pressure.
- Respiratory rate.

Baseline values will be the average of at most three measurements taken prior to the first infusion, and post-dose values at each visit will be the average the measurements taken as part of that visit including pre-infusion, during infusion, and post-infusion measurements. A listing of the vital signs measurements will be also provided.

7.2.7. Electrocardiograms (ECG)

Categorical summaries (e.g. normal, abnormal, abnormal clinically significant) and a shift from baseline to post-baseline visit will be presented. The shift from baseline will be categorized as no change and worsening. The number of subjects with a worsening ECG will include shifts from Normal to Abnormal – NCS, Normal to Abnormal – CS, and Abnormal – NCS to Abnormal – CS. The percentage of subjects worsening will be based on the number of subjects with ECGs at baseline of Normal or Abnormal-NCS and a post-baseline result. All other percentages in the shift table will be based on the number of subjects with both a baseline and a post-baseline evaluation. A listing of ECG results will be provided.

7.2.8. Denver II Developmental Screening Test (DDST)

The instrument screens children's performance in four developmental domains: personal-social, fine motor-adaptive, language, and gross motor with respect to the age-matched population. The overall results from DDST are expressed as Normal, Abnormal, Questionable, and Untestable.

DDST results will be listed. The results will also be summarized with the number of subjects who had the test done and tested 'normal', 'suspect', or 'untestable' by visit. A shift of the change from baseline to post-baseline result will also be summarized; percentages in the shift table will be based on the number of subjects in the FAS with baseline and post-baseline results.

7.2.9. Antidrug Antibody (ADA) Titer

A subject will be considered as having been ADA positive if the test was reactive and specific.

ADA data will be reported in the subject listings. The number and percentage of subjects who became ADA positive, and the number and percentage of subjects who were positive to neutralizing antibodies will be presented. Should at least 5 subjects be ADA-positive, then additional analyses will be performed to include: time to peak titer; ADA titer by time point; and median and peak ADA titer. Time to peak ADA titer will be analyzed by Kaplan-Meier method; if no titer is detected, time to peak titer will be censored at the date of the last visit during treatment with SBC-102.

Other exploratory analyses of the effect of ADAs on the safety of SBC-102 may also be performed as suggested by the type and amount of data available.

7.3. Efficacy Analyses

Efficacy will be examined for the FAS and, as subject numbers permit, stratified by ADA status, and for other subgroups of interest. Summaries will be provided in the event that at least 5 subjects are included in each subgroup defined by a characteristic (eg. at least 5 subjects with ADA positive and ADA negative).

Efficacy of SBC-102 is expected to be reflected through changes in liver and spleen fat content and organ volume, clinical laboratory data and, where applicable, growth parameters. Evaluation of change in parameters will be made from key assessment time points to the pre-dosing baseline and, as appropriate, to dose-specific baseline assessments (i.e., the assessments taken just prior to escalation to a higher dose level). Additional comparisons of change between sequential assessment time points may also be made.

For continuous endpoints, the observed values, change from baseline and percent change from baseline at each time of assessment will be summarized, and will include 95% confidence intervals.

7.3.1. Handling of Dropouts or Missing Data

All data will be analyzed as they were collected in the database. Missing data in general will not be imputed, limited imputation techniques if applied are described in Section 9.

7.3.2. Clinical Laboratory Tests

Both the change and the percent change from Baseline in serum liver biochemical parameters (liver function tests), serum lipids, hemoglobin levels, and platelet count will be tabulated for each evaluation timepoint. Data will be summarized for the endpoints as continuous variables. In addition, summaries of categories relative to the upper limit of normal (≤ 1 x ULN, 1 to ≤1.2 x ULN, 1.2 to ≤1.5x ULN, 1.5 to ≤2 x ULN, >2 x ULN) will be provided with shifts from baseline to post-baseline in these categories. Waterfall plots showing the percent change from baseline to Week 96 will be presented.

- Liver function tests: ALT/AST, GGT, total bilirubin, direct bilirubin, indirect bilirubin, albumin, alkaline phosphatase
- Lipids: total cholesterol, LDL-C, non-HDL-C, HDL-C, triglycerides
- Hematology: hemoglobin, platelets

The UK-MELD assessment, which is computed from clinical laboratory values, will be summarized similarly to laboratory data, although these scores will not be compared against standard laboratory reference ranges for individual laboratory assessments.

Also see Section 7.2.3, Laboratory Tests, as described in the Safety Analysis section.

7.3.3. Abdominal Magnetic Resonance

Change and percent change in liver and spleen volume and fat content will be calculated from Baseline and tabulated for each evaluation timepoint by multiples of normal. Data will be summarized for the endpoints as continuous variables. Spaghetti plots (1 line per subject) and plots of mean change in measurements over time may also be created for selected endpoints of interest. Ultrasound results will be summarized for subjects with no MRI.

7.3.4. Anthropometric Parameters

Anthropometric indicators of growth status (WFA, plus additional parameters if sufficient data are available) will be evaluated for subjects ≤ 18 years of age. Weight-for-length/weight-for-stature and BMI, if computed, will be derived from data on weight and length/stature. Anthropometric parameters will be plotted on standard growth curves. Z-scores and percentiles based on the age-gender standardized norms will be calculated in accordance with the methodology described by the WHO (subjects ≤ 24 months) or CDC (subjects > 24 months to 18 years) and using the growth charts relevant to the respective methodology. When possible, historical data on growth parameters will also be incorporated into the analyses. The growth curve will be generated for each subject. In addition, the percentages of subjects who meet criteria for under nutrition (underweight, stunting, and wasting) will be tabulated at each time point.

For subjects who are >18 years of age, BMI will be derived and summarized over time.

7.3.5. Exploratory Efficacy Analyses

Evaluation of the relationship between non-invasive measurements of liver fat content and liver histopathology may also be conducted using scatter plots and Spearman's rank correlation analysis.

7.4. Pharmacokinetic Analyses

Details of pharmacokinetic analyses will be described in a separate analysis plan.

7.5. Liver Biopsies

Data from Liver biopsies conducted will be summarized using tables and graphical displays, where applicable.

7.5.1. Ishak Fibrosis Score, Microvesicular Steatosis, Macrovesicular Steatosis, Lobular Inflammation and Portal Inflammation

These scores are reported on an ordinal scale; the Ishak Fibrosis Score ranges from 0 (No fibrosis/Normal) to 6 (Cirrhosis, probable or definite), while the other scores range from 0 to 4.

Each of the scores will be provided in a frequency table with number and percentage within each ordinal level at the timepoints specified in the Schedule of Assessments.

For each of the scores, shift tables will be used to summarize changes in categories (or levels) from Screening/Baseline to each subsequent timepoint at which biopsy data are collected.

7.5.2. Other Histopathology Variables

Percent fat, percent collagen, percent SMA and percent CD68 will be summarized as continuous variables for each timepoint as specified in the Schedule of Assessments. Observed results, the change from baseline and the percent change from baseline will be presented.

7.6. Endoscopic Biopsies

Per study protocol, endoscopic biopsy was conducted for subjects with evidence of growth abnormalities or malabsorption (Criterion 5d on List of Inclusion/Exclusion Criteria) to measure lipid substrates. Results may be summarized.

7.7. Esophagogastroduodenoscopy

Per study protocol, esophagogastroduodenoscopy (EGD) was conducted for subjects with evidence of advanced liver disease (Criterion 5a on List of Inclusion/Exclusion Criteria) to document the presence, size and quality of esophageal varices. Data may be summarized.

7.8. Subject Health-Related Quality of Life (HRQOL) Outcomes Analyses

For the subset of subjects ≥ 5 years of age, change in HRQOL measures will be calculated from Baseline and tabulated for each evaluation timepoint as an exploratory analysis. In addition to the evaluation of changes in the overall scores for each HRQOL measure, changes in subscales and summary scores, as applicable to the HRQOL instrument, will be summarized.

For the PedsQL™, in addition to the Generic Core Scale, the 4 subscales (physical functioning, emotional functioning, social functioning, and school functioning) may also be summarized. Observed measurements and changes and/or percent changes from baseline to study time-points will be summarized including 95% confidence intervals.

Summaries may be presented by subgroup.

7.9. LAL Enzyme Activity

LAL enzyme activity will be summarized using the results from the dried blood spot, if available, or alternatively the results from the whole blood. Results will be categorized as affected (<=0.016 nmol/punch/hr), carrier (0.017 to 0.026 nmol/punch/hr); and unaffected (0.027 to 0.0152 nmol/punch/hr). The number and percentage of subjects in each of these categories will be summarized for each visit at which the test was performed.

In addition, a shift from baseline will be presented showing improvement, no change, and worsening. Generally, the percentages of the shifts will be based on the number of subjects with baseline and post-baseline data. However, summaries of subjects showing improvement (i.e., affected to carrier, affected to unaffected, and carrier to unaffected) will be the number of subjects who were classified as carrier or affected at baseline and had a post-baseline result. Similarly, the summaries of subjects showing worsening (i.e., unaffected to carrier, unaffected to affected, and carrier to affected) will be based on the number of subjects who were in classified as unaffected or carrier at baseline and had a post-baseline result.

Summaries may be presented by subgroup.

7.10. Dietary Change

A summary of changes in diet will be presented for the FAS, as well as the status whether the subject was on a low fat or low cholesterol diet at baseline (yes/no). Any change in diet (yes/no) will be summarized. Changes in diet will reflect any post-treatment change in fat content, cholesterol content, or cholesterol lowering foods; each will be summarized as to whether there was ever any reduction, ever any increase, or no change.

In addition a summary of whether there was ever a change in dosing for subjects who were on lipid lowering medication will be summarized.

7.11. Exploratory Biomarker Analyses

Exploratory disease-related biomarkers - which may be identified, based on emerging information from the SBC-102 development program and scientific literature - will be listed and observed values and changes and/or percent changes from baseline will be summarized at available timepoints.

7.12. Subgroup Analyses

Demographic and baseline disease characteristics and efficacy outcomes will be summarized by: age at start of infusion (<12 Years *versus* >=12 Years), gender (Male *versus* Female), baseline use of lipid-lowering medications (Yes *versus* No), genetic mutation category (Homozygous for common mutation *versus* Heterozygous for common mutation *versus* Other mutation), baseline fibrosis or cirrhosis status, and baseline ALT level (<1.5 x ULN vs >=1.5 x ULN). Efficacy outcomes and selected demographic or baseline disease characteristics will also be summarized for subjects 2-4 years old. HRQOL outcomes will not be summarized by subgroup.Adverse event summaries will be presented by age at start of infusion.

Summaries will be provided in the event that at least 5 subjects are included in each subgroup defined by a characteristic (eg. At least 5 subjects with ADA positive and ADA negative).

7.13. Additional Exploratory Analyses

Post-hoc subgroup analyses may be performed as needed.

7.14. Interim Analyses

There will be two Clinical Study Reports (CSRs) generated for study LAL-CL06:

(1.) CSR#1 will present analyses described in the SAP, but with data accumulated for all subjects up to- and including – the Week 96 Visit; and

(2.) CSR#2 will present all analyses described in the SAP with data accumulated for all subjects over the entire study duration. This will be the final study CSR.

CSR#1 is being planned to enable the study team meet regulatory requirements. The analysis and subsequent release of CSR#1 will neither impact the operational characteristics of the overall study, nor affect the conduct of the remaining portion of LAL-CL06.

To facilitate the successful implementation of this strategy, two Data Base Locks (DBLs) are being planned: one for CSR#1 (DBL#1) and the other for CSR#2 (DBL#2). Efforts will be made to ensure that all queries related to data accrued up to – and including – Week 96 are resolved, and all eCRF pages up to – and including – Week 96 Visit are frozen except for AEs and concomitant medications reported as “Ongoing” as of DBL#1.

The proposed analyses are descriptive and no adjustments will be made for multiple comparisons or level of statistical significance due to multiple data locks/analyses.

7.15. Data Monitoring Committee

An independent Safety Committee (SC) will perform periodic reviews of aggregated safety data for study LAL-CL06 on an at least biannual basis (i.e., every 6 months) from the date of enrollment of the first subject until completion of dosing for all subjects in the study. Details of SC procedures, processes and analyses for the study are provided in an SC Charter.

8 REFERENCES

Frankenberg, WK, Dodds, J, Archer P, et al. The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics*. 1992;89:91-97.

Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: Methods and development. Centers for Disease Control and Prevention, National Center for Health Statistics, US Dept. of Health and Human Services. *Vital Health Stat*. 2002;(246):1-190.

Varni JW. The PedsQL™ 4.0 Generic Core Scales Young Adult Version: Feasibility, reliability, and validity in a university student population. *J Health Psychol*. 2009;14:611-622.

WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: *World Health Organization*, 2006 (312 pages).

WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: Methods and development. Geneva: *World Health Organization*, 2007 (217 pages).

Younossi ZM, Guyatt G, Kiwi M, Boparai N, King D. Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. *Gut*. 1999;45:295-300.

9 APPENDICES

9.1 Protocol Schedule of events

Refer to the approved study protocol for a schedule of events.

9.2 Changes from Analyses Specified in the Previous Version of the SAP

Not applicable

9.3 Sample Size, Power, and Randomization

No formal sample size calculations were performed for this study; the projected enrollment is based on feasibility. Given the rarity of LAL-D, it is not expected that more than 20 subjects will be treated.

9.4 Technical Specifications for Derived Variables

9.4.1 Age

Age will be presented as the number of years between date of birth and the reference date. The following age will be computed, with reference dates indicated:

Table 5: Age and reference date

AGE	REFERENCE DATE
• Age at Enrollment	• Date of Signing ICF
• Age at First Dose of SBC-102	• Date of First Dose of SBC-102

In cases where only the month and year are provided for a date, the day will be imputed as 15. Missing month will be imputed as June. In cases where the day and the month are missing, the date will be imputed as July 1.

The 2-4 year age group will be defined based on the date of first dose of SBC-102. Subjects whose date of first dose was on or after the subject's second birthday but prior to the subject's fifth birthday will be included in this subgroup.

9.4.2 Time to Peak ADA Titer

Time to Peak Titer (Weeks) = (Date of Peak ADA Titer - Date of First SBC-102 Infusion + 1)/7;

This derived variable will be used in the calculation of Kaplan-Meier estimate of median time to peak ADA titer (if calculable), and probability of "no titer detected" at selected timepoints.

9.4.3 Body Mass Index (BMI)

BMI is calculated as:

$$BMI = \text{body weight (Kg)} / [\text{Height (m)}]^2$$

BMI is expressed in Kg/m^2

9.4.4 Derivation of Study Day

The date and time of first study drug infusion is the reference day for deriving study day, and will be calculated as follows:

- for measurements on or after first study drug infusion, study day is:
  - collection date – first study infusion date + 1.
- for measurements prior to first study drug infusion, study day is:
  - collection date – first infusion date -1.
- for measurements on the day of first study drug infusion starting at the same time or after first study infusion, study day is:
  - "Day 1".
- for measurement on the day of first study drug infusion starting prior to the first study infusion, study day is:
  - "Day 0".

9.4.5 Definition of Baseline Values and Change from Baseline

9.4.5.1 Baseline

Baseline is defined as the last available pre-first study dose assessment for all subjects. In case of multiple pre-treatment measurements, Baseline is the average of the last (up to 3) measurements.

For laboratory parameters if any Baseline test is done at local and central laboratory, then the central laboratory results will be used for analysis.

9.4.5.2 Change from Baseline

Change from baseline will be calculated as

Change of Baseline = Assessment Value – Baseline Assessment Value.

9.4.5.3 Percent Change from Baseline

Percent change from baseline will be calculated as

Percent Change of Baseline = [(Assessment Value – Baseline Assessment Value)/(Baseline Assessment Value)]*100

and rounded to 1 more decimal place than observed in the raw data.

9.4.6 United Kingdom Model for End-Stage Liver Disease

The UK-MELD is calculated as (5.395 x lnINR) + (1.485 x ln serum creatinine) + (3.13 x ln serum bilirubin) – (81.565 x ln serum sodium) + 435.

9.4.7 Adverse Events

The analysis of Adverse Events is described in detail in Section 7.2.2.

9.4.7.1 Treatment-emergent Adverse Events (TEAEs)

Treatment-emergent AEs (TEAEs) are events with start dates and start times on or after the date and time of the first study drug dose (SBC-102). All other AEs are considered Pre-Treatment Adverse Events (PTAEs).

Subject percentages are based on the total number of treated subjects.

The relationship between an AE and the study drug will be classified as: not related; unlikely related; possibly related; or related. AEs assessed as “possibly related” or “related” are further categorized as “Related AEs”. AEs assessed as “not related” or “not related” are further categorized as “Unrelated AEs”.

9.4.7.2 Missing AE Dates

If the start date of an AE is partially or completely missing and the end (stop) date and time of the AE does not indicate that it occurred prior to first dose, then the determination of treatment-emergent status will be based on the following:

- If the start year is after the year of the first study drug dose, then the AE is treatment-emergent; else,
- If the start year is the same as the year of the first study drug dose and
  - the start month is missing, then the AE is treatment emergent; else if
  - the start month is present and is the same or after the month of the first study drug dose, then the AE is treatment-emergent; else.
- If the start date is completely missing, then the AE is treatment-emergent.

9.4.7.3 AE Duration (Days)

AE Duration (days) = Date of stop of AE – Date of start of AE + 1;

Duration will be set to ‘missing’ if dates are incomplete.

9.4.8 Calculation of Multiples of Normal for Liver and Spleen Volume

Multiples of Normal (MN) are derived by determining the expected “normal” weight of a specific organ based on body weight.

- Liver: 2.5% of body weight in kg gives the expected "normal" liver weight in liters;
- Spleen: 0.2% of body weight in kg gives the expected "normal" spleen volume in liters

For example: a 50 kg subject would have an expected "normal" liver volume of 1.25 liters (0.025*50=1.25) and an expected "normal" spleen volume of 0.10 liters (0.002*50=0.1). The patient's observed organ volume is then divided by the expected "normal" volume. Values of MN >1.0 indicate an organ volume which is larger than the expected normal volume.

9.4.9 Health-Related Quality of Life Questionnaires

9.4.9.1 The Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue scale

The FACIT-Fatigue instrument consists of 13 items, with each item rated on a 4-point ordinal scale ranging from 0 (Not at all) to 4 (Very Much).

		Not at all	A little bit	Somewhat	Quite a bit	Very Much
1	I feel fatigue	0	1	2	3	4
2	I feel weak all over	0	1	2	3	4
3	I feel listless ("washed out")	0	1	2	3	4
4	I feel tired	0	1	2	3	4
5	I have trouble starting things because I am tired	0	1	2	3	4
6	I have trouble finishing things because I am tired	0	1	2	3	4
7	I have energy	0	1	2	3	4
8	I am able to do my usual activities	0	1	2	3	4
9	I need to sleep during the day	0	1	2	3	4
10	I am too tired to eat	0	1	2	3	4
11	I need help doing my usual activities	0	1	2	3	4
12	I am frustrated by being too tired to do things I want to do	0	1	2	3	4
13	I have to limit my social activity because I am tired	0	1	2	3	4

Scoring of FACIT-Fatigue: The total score ranges from 0-52. Higher scores suggest better quality of life for the subject. All items, except Items #7 and #8, are reversed scored.

	Reverse Item?		Item Response	Item Score
1	4	-		=
2	4	-		=
3	4	-		=
4	4	-		=
5	4	-		=
6	4	-		=
7	0	+		=
8	0	+		=
9	4	-		=
10	4	-		=
11	4	-		=
12	4	-		=
13	4	-		=

Total Score: (sum of Item Scores) * 13 and divide by the number of items answered.

Missing Data: The total FACIT-Fatigue Scale can be calculated as long as more than 50% of the items were answered (e.g. a minimum of 7 of 13 items).

9.4.9.2 The Chronic Liver Disease Questionnaire (CLDQ)

CLDQ is a 29-item instrument to measure quality of life in 6 domains: abdominal symptoms; fatigue; systemic symptoms; activity; emotional function; and worry. Each item is rated on a 7-point ordinal scale ranging from 1 (All of the time) to 7 (none of the time).

Domains of the CLDQ:

Abdominal Symptoms (AB):	Items 1, 5, 17
Fatigue (FA):	Items 2, 4, 8, 11, 13
Systemic Symptoms (SY):	Items 3, 6, 21, 23, 27
Activity (AC):	Items 7, 9, 14
Emotional Function (EM):	Items 10, 12, 15, 16, 19, 20, 24, 26
Worry (WO):	Items 18, 22, 25, 28, 29;

Scoring of CLDQ:
Single domains and overall score range from 1-7. Higher scores implicate a better quality of life.

Single domain = (Sum individual item scores) x (number of scheduled items per domain) and divide by number of items answered.

Total score = (Sum single domains) divide by 6.

Total score can only be calculated if each single domain score is available.

Missing data:
Single domain scores can be calculated as long as more than 50% of the items were answered (e.g. a minimum of 2 of 3 items, 3 of 5 items, 4 of 8 items).

9.4.9.3 The Pediatric Quality of Life Inventory TM (PedsQL ™)

The PedsQL TM is a 23-item instrument, with each item rated on either a 5-point or 3-point Likert Scale. The 23 items are divided into four dimensions:

Physical Functioning	8 items
Emotional Functioning	5 items
Social Functioning	5 items
School Functioning	5 items

Scoring of the PedsQL ™:

Note -

- To aid interpretation, items on the PedsQL ™ instrument are reversed scored and linearly transformed to a 0-100 scale, so that higher scores indicate better quality of life.
- To reverse score, transform the 0-4 scale items to 0-100 as follows: 0=100, 1=75, 2=50, 3=25, 4=0.

Response Choices	Never	Almost Never	Some-times	Often	Almost Always
Raw Scores	0	1	2	3	4
0-100 Scale Scores	100	75	50	25	0

The following summary scores are produced –

1. **Scale Score (each domain):** the scale score – for each domain listed above – is the sum of all nonmissing item responses divided by the number of responses (to account for missing data). Items where no responses were provided should not be included in the denominator for calculating mean. ***If more than 50% of the items in the scale are missing, the Scale Score should not be computed.***
2. **Total Score:** It is calculated as the sum of all nonmissing item responses divided by the number of responses across all scales) ***if and only if all 4 scale scores (Physical, Emotional, Social, and School Functioning) are nonmissing. If ≥ 1 scale score is missing then Total Score is not computed.***
3. **Psychosocial Health Summary Score:** It is calculated as the sum of all nonmissing item responses divided by the number of responses in the: Emotional; Social; and School Functioning Scales ***if and only if all 3 scale scores (Emotional, Social and School Functioning Scales) are non-missing. If ≥1 scale score is missing then Psychosocial Health Summary Score is not computed.***
4. **Physical Health Summary Score:** The Physical Health Summary Score is the same as the Physical Functioning Scale Score.

9.4.10 Geometric Mean of ADA Titer

$$Geometric\ Mean\ (at\ Time\ i) = (v_1 \times v_2 \times \cdots \times v_n)^{\frac{1}{n}}$$

where v_n is the titer value for the n^{th} subject at timepoint *i* (e.g. baseline), and *n* is the number of subjects with non-missing titer values.

9.4.11 The Z-Score (or SD Score) and Percentiles

9.4.11.1 Z-Score

Is the deviation of an individual's observed value* from the median value of a reference population, divided by the standard deviation (SD) of the reference population.

$$\text{Z-score (or SD-score)} = \frac{\text{observed value - median value of the reference population}}{\text{standard deviation value of reference population}}$$

* Weight; Height; BMI; Head circumference etc.

9.4.11.2 Z-Scores in Growth Curves

Alternatively, the formula for the Z-score can also be expressed as:

$$\boldsymbol{Z = \frac{\left[\frac{observed\ value}{M}\right]^{L} - 1}{L * S}; \qquad whenever\ L \neq 0}$$

$$\boldsymbol{Z = \frac{\log_e\left(\frac{observed\ value}{M}\right)}{S}; \qquad whenever\ L = 0}$$

where L (lambda: power in Box-Cox transformation, reflecting degree of skewness), M (mu: the median) and S (sigma: the generalized coefficient of variation) are curve parameters – which vary according to the child's sex and age (e.g. weight-for-age among boys) or according to the child's sex and height (e.g. weight-for-height among boys) – determined from smoothed percentile curves using a series of linear and non-linear regression models and a modified LMS procedure *(Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med. 1998;17:407–429).*

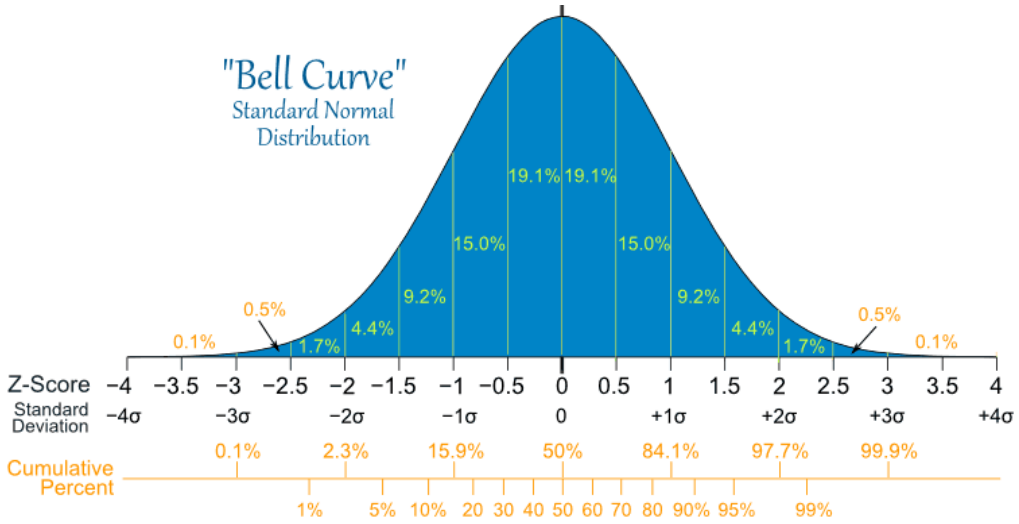
The estimates of L, S and M have been tabulated - for different anthropometric parameters - for a series of ages by WHO and CDC using datasets from surveys conducted in the United States. The estimates for these parameters are available from publicly available WHO Growth Charts (for subjects ≤24 months of age) and CDC Growth Charts (subjects >24 months of age to 18 years) and can be used to calculate Z-scores and percentiles for observed anthropometric parameters.

9.4.11.3 Classification of Z-Scores for malnutrition

1. Underweight:
    - Moderate: Weight-for-age < -2 SD from the international reference median value
    - Severe: Weight-for-age < -3 SD from the international reference median value
2. Stunting:
    - Height-for-age/Length-for-age < -2 SD from the international reference median value
3. Wasting:
    - Weight-for-height/ Weight-for-length < -2 SD from the international reference median value

9.4.11.4 Percentiles

Percentiles are cumulative probabilities associated with z-scores using Standard Normal distribution theory. Z-scores, by definition, have a Standard Normal Distribution.



Source: Pierce, Rod. (11 Oct 2014). "Normal Distribution". Math Is Fun. Retrieved 14 Jan 2016 from http://www.mathsisfun.com/data/standard-normal-distribution.html

For example, a z-score of +2 would be equal to the 97.7 percentile on the standard normal curve, meaning the subject's Z-score is in the top 2.3% in comparison to his/her reference population.

9.4.11.5 Note on Z Scores and Percentiles

1. Will only be calculated for subjects ≤ 18 years at informed consent
2. For Subjects ≤24 months of age, z-scores and percentile scores based on World Health Organization (WHO) growth charts (WHO Multicentre Growth Reference Study Group, 2006 and 2007)

3. For subjects >24 months of age to 18 years, z-scores and percentile scores based on Centers for Disease Control and Prevention (CDC) growth charts (Kuczmarski et al., 2000).