

DISCLOSURE

REDACTED STATISTICAL ANALYSIS PLAN

ACE-536-B-THAL-001

A Phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to beta (β)-thalassemia.

The “BELIEVE” Trial

The information contained in the attached report is the property of Celgene and should not be shared or used for any purpose other than that for which it was provided.

Celgene is committed to providing information about its clinical trials to researchers and patients with the goal of furthering science and enhancing healthcare worldwide. Laws and regulations require, however, that Celgene protect patient privacy. The company may further have legal or contractual obligations not to disclose commercial or technical information provided by or related to certain partner companies or vendors.

The attached report is presented in its original format, but certain information has been redacted in order to comply with the aforementioned obligations or to protect Celgene’s confidential commercial information. The redactions are based on the following principles:

- Redacted information has been replaced by grey space, maintaining original spacing and pagination.
- Any information that might allow the identification of individuals has been redacted for anonymization.
- Attachments to this report that contain confidential information are not made available. Such attachments include those that contain identifiable patient information, such as subject listings, narratives, and profiles. They also may contain confidential commercial information such as methodologies, and hypothesis generating and exploratory analyses.
- Cross-references to these attachments (such as links to subject listings in Section 16.2) are not redacted from the body of the report. However, the hyperlinks in the electronic document are no longer functional.
- Information about Celgene vendors and their services are redacted because many contracts prohibit disclosure of that information. Further, laws and regulations prevent us from disclosing certain information about our vendors or their services because it is protected by copyright.

Information about Celgene’s redaction policies and the availability of additional data from this report may be found at <http://www.celgene.com/research-development/clinical-trials/clinical-trials-data-sharing/>.

STATISTICAL ANALYSIS PLAN

A Phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to beta (β)-thalassemia.

The "BELIEVE" Trial

STUDY DRUG: Luspatercept (ACE-536)
PROTOCOL NUMBER: ACE-536-B-THAL-001
DATE FINAL: 19 Jun 2018

Prepared by:

CCI

On behalf of

Celgene Corporation

86 Morris Avenue

Summit, NJ 07901

CONFIDENTIAL

The information contained in this document is regarded as confidential and, except to the extent necessary to obtain informed consent, may not be disclosed to another party unless such disclosure is required by law or regulations. Persons to whom the information is disclosed must be informed that the information is confidential and may not be further disclosed by them.

TABLE OF CONTENTS

1.	LIST OF ABBREVIATIONS.....	7
2.	INTRODUCTION	10
3.	STUDY OBJECTIVES	11
3.1.	Primary Objective.....	11
3.2.	Secondary Objectives	11
CCI	[REDACTED]	
4.	INVESTIGATIONAL PLAN.....	13
4.1.	Overall Study Design and Plan.....	13
4.2.	Study Endpoints.....	15
4.2.1.	Primary Efficacy Endpoint	15
4.2.2.	Secondary Efficacy Endpoints.....	15
CCI	[REDACTED]	
4.2.4.	Safety Endpoints	16
4.3.	Stratification, Randomization, and Blinding	16
4.4.	Sample Size Determination	17
5.	GENERAL STATISTICAL CONSIDERATIONS.....	18
5.1.	Reporting Conventions	18
5.2.	Analysis Populations	20
5.2.1.	Intent-to-Treat Population	20
5.2.2.	Safety Population.....	20
5.2.3.	Health-related QoL Evaluable Population.....	20
6.	SUBJECT DISPOSITION.....	21
7.	PROTOCOL DEVIATIONS	23
8.	DEMOGRAPHICS AND BASELINE CHARACTERISTICS	24
8.1.	Demographics	24
8.2.	Baseline Characteristics.....	24
8.3.	Beta-thalassemia comorbidities/Medical History.....	26
8.4.	Prior Beta-Thalassemia Treatment	26
8.5.	Prior, Concomitant and Post Treatment Medications.....	26
8.5.1.	Prior Medications.....	26
8.5.2.	Concomitant Medications	26

8.5.3.	Post Treatment Medications	27
8.6.	Concomitant Procedures/Surgeries.....	27
8.7.	Prior/Concomitant/Post Iron Chelation Therapies.....	27
9.	STUDY TREATMENTS AND EXTENT OF EXPOSURE.....	28
9.1.	Treatment Duration.....	28
9.2.	Number of Doses Received per Subject	28
9.3.	Average Number of Days Between Doses	28
9.4.	Dose Delay/Adjustment.....	28
9.5.	Investigational Drug Overdose	29
10.	EFFICACY ANALYSIS	30
10.1.	Multiplicity	30
10.2.	Analysis of Primary Efficacy Endpoint.....	30
10.3.	Analyses of Key Secondary Efficacy Endpoints	31
10.3.1.	The Transfusion Reduction based on the Rolling Method	32
10.4.	Other Efficacy Analyses	33
10.4.1.	Mean Change in Liver Iron Concentration.....	33
10.4.2.	Mean Change in Mean Daily Dose of Iron Chelation Therapy.....	35
10.4.3.	Mean Change in Serum Ferritin Level	35
10.4.4.	Bone Mineral Density Assessed by DXA Scan.....	35
10.4.5.	Mean Change in Myocardial Iron by T2* MRI.....	36
10.4.6.	Change in Quality of Life assessed by TranQol and SF-36	36
10.4.7.	Duration of Transfusion Burden Reduction	39
10.4.8.	Time from First Dosing Date to the First Erythroid Response.....	40
10.4.9.	Transfusion Independence	40
10.4.10.	Post-baseline Transfusion Events Frequency	41
10.4.11.	Pre-transfusion Hemoglobin Change from Baseline	41
10.4.12.	Healthcare Resource Utilization	41
10.5.	Subgroup Analysis.....	42
10.6.	Missing Data Imputation	43
11.	SAFETY ANALYSIS	44
11.1.	Adverse Events	44
11.2.	Adverse Events of Special Interest (AESI)	45
11.3.	Other Adverse Events That Require Safety Analysis.....	45

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE

SAP TITLE ACE-536-B-THAL-001 Statistical Analysis Plan

SAP VERSION, DATE Final Version 1.0 19JUN2018

SAP AUTHOR

PPD

PPD

Printed Name and Title

Signature and Date

PROTOCOL TITLE

A Phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to beta (β)-thalassemia

INVESTIGATIONAL PRODUCT

luspatercept (ACE-536)

PROTOCOL NUMBER

ACE-536-B-THAL-001

PROTOCOL VERSION, DATE

21-APR-2017

SIGNATURE STATEMENT

By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable.

Statistical Therapeutic Area Head

Signature

Printed Name

Date

Lead Clinical Research Physician / Clinical Research Physician

Signature

Printed Name

Date

Lead Product Safety Physician

Signature

Printed Name

Date

Lead Statistician

Signature

Printed Name

Date

CELGENE PROPRIETARY INFORMATION

ICT	Iron chelation therapy
IP	Investigational product
ITT	Intent-to-treat
IVRS	Integrated voice response system
IWRS	Integrated Web Response System
LDH	Lactic dehydrogenase
LIC	Liver iron concentration
LLN	Lower limit of normal
LVEF	Left ventricular ejection fraction
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic resonance imaging
MUGA	Multi Gated Acquisition Scan
NA	Not applicable
NCI	National cancer institute
PK	Pharmacokinetic
QoL	Quality of life
RBC	Red blood cell
RDW	Red blood cell distribution width
SAP	Statistical analysis plan
SBP	Systolic blood pressure
SC	Subcutaneous
SD	Standard deviation
SE	Standard error
SGOT	Serum glutamic oxaloacetic transaminase (AST)
SGPT	Serum glutamic pyruvic transaminase (ALT)
SOC	System and organ class

PT	Preferred term
TEAE	Treatment emergente adverse event
ULN	Upper limit of normal
WHO-DD	World Health Organization Drug Dictionary

CELGENE PROPRIETARY INFORMATION

2. INTRODUCTION

This statistical analysis plan (SAP) describes the analyses and data presentations for Celgene's protocol ACE-536-B-THAL-001 "A Phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of luspatercept (ACE-536) versus placebo in adults who require regular red blood cell (RBC) transfusions due to beta (β)-thalassemia." which was issued on 25AUG2015, with an amendment version issued on 21APR2017. It contains definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety.

This analysis plan includes one primary analysis. There is no interim analysis planned for this study. Throughout this SAP, the treatment groups will be referred to as luspatercept group, which is ACE-536 plus best supportive care (BSC), and placebo group, which is placebo plus BSC. The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to database lock and any data analysis for the final analyses. This SAP will be finalized and signed prior to the clinical database lock for the final analysis. All statistical analyses detailed in this SAP will be conducted using SAS® Version 9.3 or higher.

Operational details for the Data Monitoring Committee (DMC) during the course of the study will be described in a separate DMC charter, therefore, it will not be included in this SAP.

3. STUDY OBJECTIVES

3.1. Primary Objective

The primary objective is to determine the proportion of subjects treated with luspatercept plus BSC versus placebo plus BSC who achieve erythroid response, defined as $\geq 33\%$ reduction from baseline in transfusion burden (units RBCs / time) with a reduction of at least 2 units, from Week 13 to Week 24.

3.2. Secondary Objectives

The secondary objectives are:

- To evaluate the proportion of subjects who achieve $\geq 33\%$ reduction from baseline in transfusion burden from Week 37 to Week 48 versus placebo
- To evaluate the proportion of subjects who achieve $\geq 50\%$ reduction from baseline in transfusion burden from Week 13 to Week 24 versus placebo
- To evaluate the proportion of subjects who achieve $\geq 50\%$ reduction from baseline in transfusion burden from Week 37 to Week 48 versus placebo
- To evaluate the mean change from baseline in transfusion burden from Week 13 to Week 24
- To evaluate the mean change from baseline in liver iron concentration (LIC) versus placebo
- To evaluate the mean change from baseline in mean daily dose of iron chelation therapy (ICT) used versus placebo
- To evaluate the mean change from baseline in serum ferritin versus placebo
- To evaluate the effect of luspatercept on osteoporosis/osteopenia, total hip and lumbar spine measured by bone mineral density (BMD) versus placebo
- To evaluate mean change from baseline in myocardial iron versus placebo
- To evaluate mean change from baseline in QoL as assessed by TranQoL and SF-36, versus placebo
- To evaluate the effect of luspatercept on healthcare resource utilization versus placebo
- To evaluate the proportion of subjects who are transfusion independent for ≥ 8 weeks versus placebo
- To evaluate the duration of reduction in transfusion burden or transfusion independence
- To evaluate the time to erythroid response
- To evaluate the post-baseline transfusion events frequency versus placebo

- To evaluate the population pharmacokinetics (PK) of luspatercept in subjects with β -thalassemia
- To evaluate the safety and immunogenicity of luspatercept versus placebo

CCI

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

CELGENE PROPRIETARY INFORMATION

4. INVESTIGATIONAL PLAN

4.1. Overall Study Design and Plan

This is a Phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of luspatercept (ACE-536) plus BSC versus placebo plus BSC in adults who require regular red blood cell transfusions due to β -thalassemia. Approximately 300 subjects diagnosed with transfusion-dependent β -thalassemia (including Hemoglobin E/ β -thalassemia, excluding Hemoglobin S/ β -thalassemia and Hemoglobin H) requiring regular transfusions will be randomized worldwide at a 2:1 ratio of luspatercept plus BSC versus placebo plus BSC.

Study participation for each subject includes Screening/Run-in Period, a 48-week placebo-controlled double-blind Treatment Period, followed by double-blind Long-term Treatment Period, an Open-label Phase and a Post-treatment Follow-up Period. At the end of the double-blinded Long-term treatment period, unblinding will occur to assess individual subject's eligibility to enter the Open-label Phase. Subjects initially receiving luspatercept and not discontinuing the double-blind phase are eligible to enter the Open-label Phase; subjects initially receiving placebo and meeting the screening criteria for Open-label Phase are eligible to enter even if they might have discontinued the double-blinded phase. This analysis plan will only address data summary up to unblinding.

Subject's eligibility will be determined during the Screening/Run-in period of at least 12 weeks. The qualified subjects will be randomized to luspatercept group or placebo group with 2:1 ratio based on subjects' geographical region as a stratification factor (refer to section 4.3).

During the double-blind Treatment Period, subjects will receive their first subcutaneous (SC) dose of luspatercept or placebo (1 mg/kg) on Day 1 of each dosing cycle. BSC will be available to all study subjects, which will include RBC transfusions, iron-chelating agents, use of antibiotic therapy, antiviral and antifungal therapy, and/or nutritional support as needed, to minimize the safety risk to subjects. The double-blind Treatment Period will last up to 48 weeks from study day 1, regardless of dose delays. Upon completion of the 48-week treatment period, subjects can enter the long-term treatment period and continue receiving the investigational product (IP) that they are originally randomized to at the investigator's discretion. The Long-term Treatment Period will continue until all subjects have completed or discontinued their 48-week treatment period and last maximally up to 48 weeks post last subject's first dose or discontinued early, whichever occurs first. Treatment is administered every 21 days (3 weeks) during the Treatment Period and during the Long-term Treatment Period, unless dose delay or treatment discontinuation is indicated. Subjects randomized to the luspatercept group will start luspatercept at 1 mg/kg dose level and can be dose titrated up to a maximum of 1.25 mg/kg.

Upon last subject completing 48 weeks after first dose date, the study will be unblinded. Placebo subjects who are still ongoing or in follow-up have to fulfill eligibility criteria and opt to receive luspatercept in the Open-label Phase or to discontinue treatment and enter the Post-treatment Follow-up Period. In the Open-label Phase, subjects may receive luspatercept until all subjects initially assigned to luspatercept in the double-blind Treatment Period complete the total treatment duration of 5 years from subjects' Dose 1 Day 1 or discontinue early. The Open-label Phase will continue to be monitored by an independent external DMC.

All subjects who discontinue treatment will undergo a 156-week Post-treatment Follow-up Period, following the last dose of IP (monitored at week 9, week 24, and every 24 weeks until week 144) and then the end of study visit at week 156. Specifically, discontinued placebo subjects may have several follow up visits before unblinding and re-enter the treatment in Open-label Phase. Subjects who stay on treatment until end of the Open-label Phase will enter the end of study visit directly.

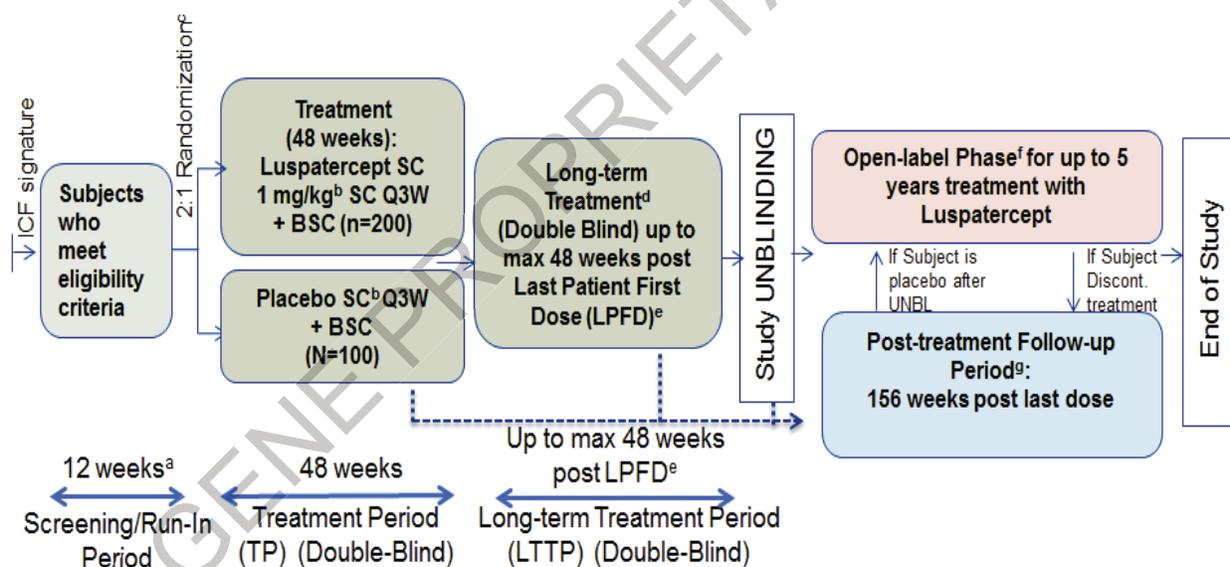
The end of treatment is defined as the last visit during the Treatment Period, the Long-term Treatment Period, or the Open-label Phase, whichever comes later. For data summary in this analysis plan, the end of treatment is defined as the last visit during the Treatment Period, or the Long-term Treatment Period, whichever comes later.

The end of study is defined as the time upon completion of the Post-treatment Follow-up visit or end of the Open-label Phase or end of trial as described as below.

The End of Trial is defined as when all subjects initially assigned to luspatercept in the double-blind Treatment Period, reach the maximum treatment duration of 5 years from subjects' Dose 1 Day 1 or discontinue earlier and complete the 156 weeks of the Post-treatment Follow-up Period, whichever occurs later; or the date of receipt of the last data point from the last subject that is required for primary, secondary, CCI analysis, as pre-specified in the protocol, whichever is the later date.

The study schematic is presented in [Figure 1](#):

Figure 1: Overall Study Design



BSC = Best Supportive Care; DMC = Data Monitoring Committee; ICF = Informed Consent Form; Q3W = every 3 weeks; SC = subcutaneous; UNBL = unblinding; LPFD = Last Patient First Dose.

^a The historical documentation of transfusion dependence for β -thalassemia subjects (including units transfused and hemoglobin (Hb) levels measured prior to each transfusion) for 24 weeks prior to subject randomization, should be available.

^b Dose may be titrated up to a maximum of 1.25 mg/kg.

^c Randomization will be 2:1, luspatercept plus BSC versus placebo plus BSC.

^d All subjects, who complete 48 weeks of the double-blind Treatment Period of this study will have the opportunity to continue to a double-blind Long-term Treatment Period at the Investigator's discretion. Subjects who do not enroll in the double-blind Long-term Treatment Period or who discontinue early will proceed to the Post-treatment Follow-up Period.

^e Maximum duration of 48 weeks after LPFD or when all subjects completed 48 weeks of double-blind treatment or discontinued before reaching 48 weeks double-blind treatment, or in the event the study is unblinded per DMC recommendation.

^f Open-label Phase: Subjects who were compliant with the protocol 48 weeks post Dose 1 Day 1 can enter in the Open-label Phase, unless medically contraindicated and as described in protocol Section 3.1.4..

^g Early discontinued subjects, i.e., subjects who discontinue before completing the double-blind treatment period (48 weeks), will continue to be monitored on week 9, followed by 24, 48, 72, 120, 144 after the last dose up to Week 156, ie 3 years (refer to protocol Section 3.1.5).

4.2. Study Endpoints

4.2.1. Primary Efficacy Endpoint

The primary endpoint is the proportion of subjects with hematological improvement, which is defined as RBC transfusion burden reduction from baseline $\geq 33\%$ with a reduction of at least 2 units during Week 13 - 24 compared to the 12-week interval on or prior to Dose 1 Day 1 for luspatercept plus BSC versus placebo plus BSC.

4.2.2. Secondary Efficacy Endpoints

4.2.2.1. Key Secondary Efficacy Endpoints

The key secondary endpoints will be measured at Week 24 and Week 48, and will be statistically tested in a sequential order at $\alpha = 0.05$ level, details related to multiplicity adjustment can be found in Section 10.1. The key secondary efficacy endpoints include:

- Proportion of subjects with hematological improvement, defined as $\geq 33\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 37 to Week 48 compared to the 12-week interval on or prior to Dose 1 Day 1 for luspatercept plus BSC versus placebo plus BSC.
- Proportion of subjects $\geq 50\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 13 to Week 24 compared to the 12-week interval on or prior to Dose 1 Day 1 for luspatercept plus BSC versus placebo plus BSC.
- Proportion of subjects $\geq 50\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 37 to Week 48 compared to the 12-week interval on or prior to Dose 1 Day 1 for luspatercept plus BSC versus placebo plus BSC.
- Mean change from baseline in transfusion burden (RBC units) from Week 13 to Week 24.

4.2.2.2. Other Efficacy Endpoints

Other efficacy endpoints include:

- Mean change from baseline in liver iron concentration (LIC, mg/g dw) by MRI
- Mean change from baseline in mean daily dose of ICT
- Mean change from baseline in serum ferritin
- Mean change from baseline in total hip and lumbar spine BMD by DXA
- Mean change from baseline in myocardial iron by MRI
- Mean change from baseline (screening) in Quality of Life assessed by Transfusion-dependent QoL questionnaire (TranQoL) and 36-item Short Form Health Survey (SF-36) at Week 24 and Week 48
- Healthcare resource utilization
- Proportion of subjects who are transfusion independent for ≥ 8 weeks during treatment
- Duration of reduction in transfusion burden
- Duration of transfusion independence
- Time to erythroid response
- Post-baseline transfusion events frequency versus placebo
- PK analysis: CCI [REDACTED]

4.2.4. Safety Endpoints

Safety endpoints include:

- Type, frequency and severity of adverse events and relationship to luspatercept (per NCI CTCAE version 4.0)
- Frequency of antidrug antibodies CCI [REDACTED]

4.3. Stratification, Randomization, and Blinding

Subjects will be randomized to receive luspatercept or placebo at a 2:1 ratio. Randomization will be accomplished by an IVRS/IWRS to ensure timely registration and randomization. A stratified

randomization schedule will be implemented. Randomization will be stratified by the following geographical regions:

- North America and Europe (including Bulgaria, Canada, France, Greece, Italy, United Kingdom and United States)
- Middle East and North Africa (including Israel, Lebanon, Tunisia and Turkey)
- Asia-Pacific (including Australia, Malaysia, Taiwan and Thailand)

4.4. Sample Size Determination

Based on data in the luspatercept Phase 2 (A536-04/A536-06) studies, the assumed targeted response rate for the primary endpoint is 40% in the luspatercept group and 20% for the placebo group. A total sample size of 300 subjects (200 in the luspatercept group, 100 in the placebo group) will have 90% power to detect the difference between the luspatercept group and the placebo group with a 2-sided alpha of 0.05 and assumed 10% drop-out rate for each treatment group.

5. GENERAL STATISTICAL CONSIDERATIONS

5.1. Reporting Conventions

Summary tables, listings, figures and any supportive SAS outputs will include a “footer” of explanatory notes that will indicate, at a minimum, the following:

- Program source (e.g., SAS program name, including the path, that generates the output) and
- Data extraction date (e.g., the data cutoff date, database lock date, run date)

The purpose of the data extraction date is to link the output to a final database, either active or archived, that is write-protected for replication and future reference. An output date will also appear on each output page and will indicate the date the output was generated by the analysis program.

The following reporting conventions apply generally to tables, listings, and figures:

- Data from all study centers will be combined for analysis;
- All stratified efficacy analyses will use randomization factor as stratum;
- All statistical tests of the treatment effect will preserve a significance level of 0.050 for 2-sided tests. Testing of interactions will be performed at the 0.100 significance level, unless specified otherwise;
- P-values will be rounded to 4 decimal places. P-values that round to 0.0000 will be presented as '<0.0001' and p-values that round to 1.000 will be presented as '>0.9999';
- Confidence intervals (CIs) will be presented as 2-sided 95% CIs unless specified differently in specific analysis;
- Summary statistics will consist of the number and percentage of subjects (or cycles, if appropriate) in each category for discrete variables, and the sample size, mean, median, Standard Deviation (SD), Q1, Q3, minimum, and maximum for continuous variables;
- All mean, median, Q1, and Q3 values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value; Minimum and maximum values will be presented to the same number of decimal places as the measured value.
- All percentages will be rounded to one decimal place. The number and percentage of responses will be presented in the form XX (XX.X%), where the percentage is in the parentheses; when the number of a response is zero, percentage will not be presented for that response;
- All listings will be sorted for presentation in order of treatment group, study center, subject, and date of procedure or event if not otherwise specified;

- All listings will display original collected values, cases with special marks (i.e., <500) will be listed as it is. The special mark will be removed if the value is used for calculation in tables.
- All analysis and summary tables will have the analysis population sample size (i.e., number of subjects) if not otherwise specified;
- All summary tables will be displayed by treatment (“Luspatercept + BSC” and “Placebo + BSC”), the “Total” group will be added for sections if specified;
- The day of the first dose of IP will be defined as Day 1; for erythroid response related endpoint, if a subject is not treated, the randomization date will be used as Day 1.
- In general, if not otherwise specified, baseline value will be defined as the last value (including “unscheduled”) on or before the date of the first dose of IP (if collecting time is available, date/time will be used to compare with first dosing date/time to identify baseline record, if there is no time available, only date will be used); if multiple values are present for the same date/time, the average of these values will be used as the baseline. For subjects who were not treated, the value on or prior to randomization date will be used. Specifically, for the laboratory hematology parameter ‘Leukocytes’, the baseline is defined as the highest value between screening visit and dose 1 day 1 visit.
- For data handling in change from baseline and shift tables (except for MRI and DXA parameters), “unscheduled” visits will be grouped with the closest scheduled visit based on assessment date. The average will be used as value for that scheduled visit in change from baseline tables; the worst category will be used in shift tables. If an unscheduled visit has equal distance to two scheduled visits, it will be grouped with the later visit. Specifically, for ADA titer summary, the titer value won’t be averaged if an “unscheduled” visit is mapped to the closest scheduled visit. The titer for “unscheduled” visit will only be used for summary if the original scheduled visit has no titer result.
- For RBC transfusion related efficacy endpoints summary, the 12-week interval is defined as:
 - Baseline 12-week interval: from Day -83 to Day 1;
 - Week 1 - 12 interval: from Day 2 to Day 85;
 - Week 13 - 24 interval: from Day 86 to Day 169;
 - Week 25 - 37 interval: from Day 170 to Day 253;
 - Week 37 - 48 interval: from Day 254 to Day 337.

This SAP addresses efficacy and safety endpoints during the double-blinded treatment period (including the initial 48-week treatment period for all subjects and the long-term treatment period for early enrolled subjects). The database will be locked upon all subjects completed 48 weeks of the double-blind treatment period or discontinued before reaching 48 weeks. All the data collected up to 11MAY2018 will be used for summary. Data selection rules will be applied in each summary panel as needed, please refer to individual part for details.

5.2. Analysis Populations

5.2.1. Intent-to-Treat Population

The intent-to-treat (ITT) population will consist of all randomized subjects regardless of whether or not the subject received IP. All efficacy analyses will be conducted for the ITT population and will be analyzed based on randomization group.

5.2.2. Safety Population

The safety population will consist of all subjects who were randomized and received at least one dose of IP. Subjects will be included in the treatment group corresponding to the IP they actually received.

5.2.3. Health-related QoL Evaluable Population

The Health-related QoL (HRQoL) evaluable population consists of all subjects in the ITT population who completed the health-related QoL assessment at baseline (screening) and at least one postbaseline assessment visit. The completion of a health-related QoL assessment is defined for each health-related measure:

- TranQoL: Completion at a given visit is defined as $\geq 75\%$ of all items that were answered (i.e., ≥ 27 items of the 36 items or a nonmissing total score).
- SF-36: Completion at a given visit is defined as $\geq 50\%$ of all items that were answered (i.e., ≥ 18 items of the 36 items or a nonmissing total score).

6. SUBJECT DISPOSITION

The total number of subjects screened and total number of subjects with screen failure will be summarized. Reasons subjects did not qualify for the study will be displayed by category. A corresponding listing will be provided.

A summary of analysis populations will be presented by treatment group and total, including ITT population and safety population.

Subject disposition summary will present the number and percentage of subjects for the following categories: subjects who were randomized, subjects who received treatment, subjects who discontinued study treatment, subjects whose treatment were ongoing, subjects who completed 24 weeks of treatment, subjects who completed 48 weeks of treatment, and subjects who discontinued from the study by treatment group and total. The reasons for discontinuation of study treatment and the reasons for discontinuation of study participation will also be summarized in the table. All percentages will be based on the number of subjects randomized using the ITT population.

The reasons for treatment discontinuation will be collected on the electronic case report form (eCRF) and summarized for all treated subjects based on the following categories:

- Death
- Adverse event/Adverse event: Other
- Pregnancy
- Progressive disease
- Lack of efficacy
- Recovery
- Withdrawal by subject
- Non-compliance with study drug
- Lost to follow up
- Study terminated by sponsor
- Transition to commercially available treatment
- Physician decision
- Disease relapse
- Symptomatic deterioration
- Protocol violation
- Adverse event: Leukocytosis Grade 3
- Adverse event: Hematological malignancy
- Other

The reasons for study discontinuation will be collected on the eCRF (only when FUP period is not completed) and will be summarized for all randomized subjects based on the following categories:

- Death
- Adverse event
- Pregnancy
- Lack of efficacy
- Recovery
- Withdrawal by subject
- Non-compliance with study drug
- Lost to follow up
- Study terminated by sponsor
- Transition to commercially available treatment
- Physician decision
- Disease relapse
- Symptomatic deterioration
- Protocol violation
- Other

A summary of subjects enrolled by geographic region, country and site will be provided in a separate table by treatment group and total.

A subject disposition listing will be provided.

7. PROTOCOL DEVIATIONS

The protocol deviations will be identified and assessed by clinical research physician or designee following company standard operational procedure. A violation occurs when there is any departure from the approved protocol that: impacts the safety, rights, and/or welfare of the subject; or negatively impacts the quality or completeness of the data; or makes the informed consent document/form inaccurate. Protocol violations are identified based on blinded data reviews of deviation log throughout the study and are finalized prior to database lock.

The number and percentage of the subjects with any protocol deviation or protocol violation will be provided for the ITT population respectively by treatment group and total. For protocol violations, the number and percentage of subjects within each subcategory will be summarized as well.

A listing for protocol deviation will be provided.

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

The demographics and baseline characteristics will be summarized for the ITT population. Individual subject listings will be provided to support the summary tables.

8.1. Demographics

Summary statistics will be provided descriptively by treatment group and total for the following continuous variables:

- Age
- Weight (kg)
- Height (cm)
- Body mass index (BMI; kg/m²)

Age or date of birth will be recorded on the eCRF. Where age is not recorded, age will be calculated as described in section 17.1.1.

Body mass index (BMI) will be calculated as follows:

BMI (kg/m²) = baseline weight in kg / (height in m)².

A frequency summary (number and percentage) will be provided by treatment group for the following categorical variables:

- Age category (≤ 32 years, $> 32 - \leq 50$ years and > 50 years)
- Sex (Male, Female with or without childbearing potential)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Not Reported, Other)
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Region (North America and Europe, Middle East and North Africa, Asia-Pacific)
- BMI category (< 20 , ≥ 20 to < 25 , ≥ 25 to < 30 , ≥ 30 kg/m²)

8.2. Baseline Characteristics

The following baseline characteristics will be summarized. Baseline characteristics will also be summarized by subgroups stated in section 10.5:

- Beta-thalassemia diagnosis;
- Age when subject started regular transfusions (in years) ;
- Baseline transfusion burden in units/12 weeks based on 12 weeks historical data, i.e., transfusion data between day -167 and day -84;
- Baseline transfusion burden in units/12 weeks based on 12 weeks run-in data, i.e., transfusion data between day -83 and day 1 (descriptive

- and categorized level: ≤ 6 and > 6 ; as well as categorized level: low transfusion burden (≤ 5), medium transfusion burden ($> 5 - \leq 7$) and high transfusion burden (> 7));
- Baseline transfusion burden in units/24 weeks based on 12 weeks historical data and 12 weeks run-in data (descriptive and categorized level: low transfusion burden (≤ 10), medium transfusion burden ($> 10 - \leq 15$) and high transfusion burden (> 15));
 - 24 week pre-transfusion hemoglobin threshold, defined as mean of all documented pre-transfusion hemoglobin values during the 24 weeks prior to Dose 1 Day 1 (descriptive and categorized level: < 9 g/dL and ≥ 9 g/dL);
 - 12 week pre-transfusion hemoglobin threshold, defined as mean of all documented pre-transfusion hemoglobin values during the 12 weeks prior to Dose 1 Day 1;
 - Beta-thalassemia gene mutation grouping: B0/B0 and Non-B0/B0. All reported beta gene mutations were validated by a trained molecular biologist [REDACTED]. [REDACTED]. Mutations were homogenized to HGVS and legacy nomenclature and the beta severity (beta0 or beta+) described. Hemoglobin E variants were considered a beta+ mutation. Co-inheritance of alpha thalassemia (single or double gene deletion) or alpha gene triplication of quadruplication were also documented.
 - Eastern Cooperative Oncology Group (ECOG) performance status (0 or 1) at screening visit;
 - Splenectomy;
 - Hepatitis B and C results;
 - MRI liver iron content (LIC) (descriptive and categorized level: 0-3, >3 (including subgroups $> 3 - \leq 7$, $> 7 - \leq 15$ and > 15). The value of LIC will be either the value collected from eCRF or the value derived from T2*, R2* or R2 parameter depending on which techniques and software was used for MRI LIC acquisition. Please refer to section 10.4.1 for more imputation details.
 - MRI myocardial T2* and Iron. If myocardial iron content is missing, it will be derived from non-missing myocardial T2* value: $45/(T2*)^{1.22}$
 - Bone mineral density DXA scan (BMD scores and T-scores);

Specifically, bar plot will be provided for baseline transfusion burden in units/12 weeks and in units/24 weeks. The baseline transfusion burden in units/12 weeks is grouped to low transfusion burden (≤ 5); medium transfusion burden ($> 5 - \leq 7$); and high transfusion burden (> 7). The baseline transfusion burden in units/24 weeks is grouped to low transfusion burden ($> 6 - \leq 10$); medium transfusion burden ($> 10 - \leq 15$); and high transfusion burden (> 15). Percent of subjects within each category will be displayed by treatment group.

8.3. Beta-thalassemia comorbidities/Medical History

The Beta-thalassemia comorbidities and medical history will be coded by Medical Dictionary for Regulatory Activities (MedDRA; Version 20.0), and summarized by system organ class (SOC) and preferred term (PT) by treatment group and total. The SOCs and PTs will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each SOC/PT.

A separate table will be provided to summarize Beta-thalassemia comorbidities by SOC and PT for each treatment group and total.

Corresponding listing will be provided.

8.4. Prior Beta-Thalassemia Treatment

Prior β -thalassemia treatment will be coded by the Anatomical Therapeutic Chemical (ATC) coding scheme of the World Health Organization Drug Dictionary (WHO-DD; Version March 2017) and summarized together with prior medications.

Details of prior treatment will be provided in a listing.

8.5. Prior, Concomitant and Post Treatment Medications

Prior, concomitant, and post treatment medication collected in the eCRF will be coded by the ATC coding scheme of WHO-DD (Version March 2017), same as prior β -thalassemia treatment. Details of prior, concomitant and post treatment medications will be provided in a listing.

8.5.1. Prior Medications

Prior medications are defined as medications that were started before the start of the study treatment and either ended before the start of the study treatment or continued after study treatment. A summary showing the number and percentage of subjects who took prior medications or prior β -thalassemia treatment will be presented by ATC4 level and PT by treatment group and total. ATC4 level and preferred terms (PTs) will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each ATC4/PT.

8.5.2. Concomitant Medications

Concomitant medications are defined as non-study medications that are started on or after the start but on or before the end of the study treatment, or started before the start of the study treatment and ended or remain ongoing during the study treatment.

A summary showing the number and percentage of subjects who took concomitant medications will be presented by ATC4 level and PT by treatment group and total. ATC4 level and PTs will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each ATC4/PT.

8.5.3. Post Treatment Medications

Post treatment medications are defined as medications that were initiated after the last dose of the study treatment. A summary showing the number and percentage of subjects who took post treatment medications will be presented by ATC4 level and PT by treatment group and total. ATC4 level and PTs will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each ATC4/PT.

8.6. Concomitant Procedures/Surgeries

Procedures/surgeries will be coded by MedDRA (Version 20.0). A summary showing the number and percentage of subjects who had concomitant procedures will be presented by SOC and PT by treatment group and total. The SOCs and PTs will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each SOC/PT.

Corresponding listing will be provided.

8.7. Prior/Concomitant/Post Iron Chelation Therapies

Iron Chelation therapies are coded by the ATC coding scheme of WHO-DD (Version March 2017). Details of therapies will be provided in a listing.

Prior iron chelation therapies are defined as therapies that were started before the start of the study treatment and either ended before the start of the study treatment or continued after study treatment. Concomitant iron chelation therapies are defined as therapies that are started on or after the start but on or before the end of the study treatment, or started before the start of the study treatment and ended or remain ongoing during the study treatment.

Post treatment iron chelation therapies are defined as therapies that are initiated after the last dose of the study treatment.

The number and percentage of subjects who had prior/concomitant/post iron chelation therapies will be presented by ATC4 level and PT in separate summary tables by treatment group and total. ATC4 level and PTs will be listed in descending frequency within the luspatercept group. A subject will be counted only once for multiple events within each ATC4/PT.

9. STUDY TREATMENTS AND EXTENT OF EXPOSURE

Subjects will be assigned to one of following regimens during the treatment phase:

- Luspatercept starting dose level 1 mg/kg SC once every 21 days
- Placebo SC once every 21 days

Study treatment and extent of exposure summaries will be provided based on the safety population. Descriptive statistics will be provided for treatment duration, number of doses received per subject/treatment and average number of days between doses by treatment group and total. The number and percentage of subjects will be summarized for maximum dose level received, maximum dose received within 24 weeks, maximum dose received beyond 24 weeks, and the reduced dose level by treatment group and total. Corresponding listing will be provided.

9.1. Treatment Duration

Treatment duration (weeks) is defined as:

$$[(\text{The treatment end date}) - (\text{The treatment start date}) + 1]/7,$$

where the treatment start date is the date of the first dose of study drug. The treatment end date is min [(date of last dose + 20), death date].

Descriptive statistics for treatment duration will be summarized by treatment group and total.

9.2. Number of Doses Received per Subject

Total number of doses received per subject is defined as the total number of doses the subject received (i.e., total number of non-zero doses). It will be summarized descriptively and categorically (1, 2, 3, 4, 5, 6, 7, 8, 9 - 16, 17 - 24, 25 - 32, >32) by treatment group and total. The total number of doses received will be calculated from all subjects within each treatment group. The total number and percentage of doses received for each planned dose level (1.25 mg/kg, 1.0 mg/kg, 0.80 mg/kg, 0.60 mg/kg, 0.45 mg/kg) will be calculated within treatment group, with the total number of doses received as the denominator.

9.3. Average Number of Days Between Doses

Average number of days between doses is defined as the number of days on treatment (treatment duration) divided by the number of doses (where a subject received a non-zero dose) for each subject. Descriptive summary statistics will be provided for average number of days between doses by treatment group and total.

9.4. Dose Delay/Adjustment

Dose delay is defined as delay of planned dose schedule due to increased hemoglobin ($\geq 11.5\text{g/dL}$) or adverse events (any related events \geq grade 2) or WBC count $\geq 2\text{x}$ baseline in the absence of an associated condition (e.g., Infection or concomitant corticosteroid use) or WBC count $\geq 3\text{x}$ baseline or Grade 3 Leukocytosis or other reasons. If dose delay exceeds 15 weeks

from last dose administration date, the treatment should be discontinued. Dose adjustment includes dose reduction and dose titration (increase). Dose reduction can be caused by increased or high level of hemoglobin or adverse events. Titration is based on erythroid response during previous two dose cycles. It only occurs when transfusion reduction is obtained at specific level and reviewed by sponsor.

The dose delay, dose reduction and dose titration will be summarized in separate tables by treatment group and total. The number of subjects with at least one dose delay/reduction/titration, number of dose delays/reduction/titration per subject, reason for dose delay/reduction/titration, time to first dose delay/reduction/titration (days), and time to first dose delay/reduction due to AE (days) will be summarized by treatment group and total. Corresponding listing will be provided.

9.5. Investigational Drug Overdose

Overdose refers to luspatercept only. It is defined as SC 10% over the protocol-specified dose level assigned to a given subject, regardless of adverse events or sequelae. A listing will be provided for any overdose, which occurs accidentally or intentionally as collected in the eCRF.

10. EFFICACY ANALYSIS

All efficacy evaluations will be conducted using the ITT population, with the exception of Quality of Life analyses that will be conducted on the HRQoL evaluable population. Statistical comparisons will be made between luspatercept plus BSC vs. placebo plus BSC. Key secondary efficacy results will be considered statistically significant after consideration of the strategy for controlling the family-wise Type 1 error rate, as described in Section 10.1, Multiplicity. All statistical tests will be 2-sided at the significance level of $\alpha = 0.05$, and the corresponding p-values and 2-sided CIs for point estimates will be reported, unless specified otherwise.

For the early treatment discontinued subjects, i.e., patients who did not complete 24 weeks or 48 weeks of double-blinded treatment period, the transfusion records will still be collected up to 48 weeks or 9 weeks post last dose, whichever is the later date. All the transfusion records collected throughout the entire study period up to the efficacy cutoff date will be used in the RBC related efficacy analyses. The efficacy cutoff date is defined as the minimum date among death date, study discontinuation date, last dose date + 20, and 11MAY2018. For the primary and key secondary endpoints, if at the time of data summary, a subject's efficacy cutoff date is before the end of the 12-week interval or a subject has any invalid transfusion records (i.e., transfusion unit not available) during the specified 12-week interval, this subject will be included in the analysis as a non-responder.

10.1. Multiplicity

Gate-keeping methods will be used to control the overall Type 1 error rate for the key secondary endpoints. After the result from the primary efficacy analysis in the ITT population shows statistical significance, the key secondary endpoint 1 will be tested next. The key secondary endpoint 2 will be tested only if the test results for both primary endpoint and the key secondary endpoint 1 are significant. The key secondary endpoint 3 will be tested only if the test results for primary endpoint and the key secondary endpoints 1 and 2 are all significant. The testing procedure above will be implemented strictly in order to control the overall Type 1 error rate of 0.05 due to multiplicity. Details regarding the gate-keeping methods are described in section 10.3.

10.2. Analysis of Primary Efficacy Endpoint

The primary endpoint response rate is defined as the number of responders (subjects who achieve an erythroid response during the 12-week interval from Week 13 to Week 24 compared to baseline) divided by the number of subjects in the ITT population within each treatment group. The erythroid response is defined as subjects with $\geq 33\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units, where the 12-week interval on or prior to Dose 1 Day 1 is used as baseline value. Specifically, if a subject is not treated, the 12-week interval on or prior to randomization date will be used as baseline value.

The 12-week RBC transfusion burden (units/12 weeks) is calculated as

Number of RBC units transfused during the 12 - week interval

The following statistical hypothesis will be tested:

$H_0: P_1$ (response rate in the luspatercept group) = P_2 (response rate in the placebo group)

$H_a: P_1 \neq P_2$

The treatment comparison (luspatercept plus BSC versus placebo plus BSC) will be conducted by the Cochran Mantel-Haenszel (CMH) test stratified by the geographical regions defined at randomization as stratification factor. The odds ratio (OR) (luspatercept versus placebo) with corresponding 2-sided (at 0.05 alpha level) 95% CI and p-value will be provided.

The number and percentage of responders will be summarized by each treatment group and the difference in proportions (luspatercept – placebo) and corresponding 95% CI will also be calculated by unconditional test.

A forest plot showing the ORs, 95% CI and p-value for the overall result and the results in each subgroup will be constructed.

Listing of individual RBC transfusion data will be provided.

10.3. Analyses of Key Secondary Efficacy Endpoints

The key secondary endpoints are:

1. Proportion of subjects with hematological improvement, defined as $\geq 33\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 37 to Week 48.
2. Proportion of subjects with hematological improvement, defined as $\geq 50\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 13 to Week 24.
3. Proportion of subjects with hematological improvement, defined as $\geq 50\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units from Week 37 to Week 48.
4. Mean change in transfusion burden (RBC units/12 weeks) from Week 13 to Week 24.

To control the overall Type 1 error rate for the endpoints 1~3, the testing procedure will be implemented strictly in order: the test for 33% hematological improvement from week 37 to week 48 (endpoint 1) will only be conducted when there is evidence showing that erythroid response is achieved in the luspatercept group from week 13 to week 24 (primary endpoint); the test for 50% hematological improvement from week 13 to week 24 (endpoint 2) will only be conducted when there is evidence showing that erythroid response is achieved in the luspatercept group from week 13 to week 24 (primary endpoint) and 33% hematological improvement is achieved in the luspatercept group from week 37 to week 48 (endpoint 1); the test for 50% hematological improvement from week 37 to week 48 (endpoint 3) will only be conducted when there is evidence showing that erythroid response is achieved in the luspatercept group from week 13 to week 24 (primary endpoint), the 33% hematological improvement is achieved in the luspatercept group from week 37 to week 48 (endpoint 1) and the 50% hematological improvement is achieved in the luspatercept group from week 13 to week 24 (endpoint 2).

For the first three secondary endpoints, the number and percentage of responders in ITT population will be summarized by treatment group and the treatment comparison will be analyzed analogous to the primary efficacy endpoint, using the CMH model stratified by the geographical regions defined at randomization.

The OR (luspatercept versus placebo) with corresponding 2-sided (at 0.05 alpha level) 95% CI and p-value will be provided. The difference in proportions (luspatercept – placebo) will also be calculated. A forest plot showing the ORs, 95% CI and p-value for the overall result and the results in each subgroup will be constructed for each endpoint.

The fourth secondary endpoint, mean change in RBC transfusion burden at the 12-week interval of Week 13 to Week 24 from the baseline 12-week interval, will be analyzed using analysis of covariance (ANCOVA) with baseline values and geographical regions defined at randomization taken as covariates for the ITT population. Treatment effect will be evaluated as a contrast of luspatercept versus placebo. Least squares (LS) means with corresponding standard errors (SE) for each treatment group, along with LS mean of treatment difference with corresponding 95% CI and p-value will be presented. Also, summary statistics for RBC transfusion burden, change from baseline and percent change from baseline in RBC transfusion burden will be provided by treatment group for the following 12-week intervals: Week 1 to Week 12, Week 13 to Week 24, Week 25 to Week 36, and Week 37 to Week 48.

The totality of transfusion burden reduction will be evaluated using 24 weeks baseline (sum of 12 weeks historical data and 12 weeks run in data). Baseline of 48 weeks transfusion burden will be defined as 2 times 24 weeks baseline transfusion burden. Descriptive statistics (n, mean, median, SD, range) for totality of transfusion burden reduction along with the change from baseline will be summarized for each treatment group by 24 weeks and 48 weeks. Only subjects whose efficacy cutoff date is on or beyond end of the week 24/48 interval are included in the analysis.

A waterfall plot will be provided for the 24 week post-baseline transfusion burden percent change from baseline by treatment group and time point. Individual subject's transfusion burden percent change from baseline will be displayed in a single bar.

10.3.1. The Transfusion Reduction based on the Rolling Method

To measure the duration of transfusion burden reduction, time to the first erythroid response, transfusion independence and duration of transfusion independence, rolling method will be applied. The summary based on rolling method intends to reflect consistency with clinical practice. Please refer to section 10.4.7, 10.4.8, and 10.4.9 for more details.

The transfusion reduction will be measured using the consecutive “rolling” 12-week (or 24-week) time interval within the entire study period up to the efficacy cutoff, i.e., Days 2 to 85, Day 3 to 86 (or Day 2 to 169, Day 3 to 170 for 24-week) and so on. Note that, day 1 transfusion belongs to baseline. The transfusion reduction by 12-week is defined as subjects with $\geq 33\%$ (or 50%) reduction from baseline in RBC transfusion burden with a reduction of at least 2 units. The transfusion reduction by 24-week is defined as subjects with $\geq 33\%$ (or 50%) reduction from baseline in RBC transfusion burden only (i.e., without considering the absolute unit reduction).

The treatment comparison with “rolling” method (luspatercept plus BSC versus placebo plus BSC) will be conducted by the CMH test stratified by the geographical regions defined at

randomization as stratification factor. The OR (luspatercept versus placebo) with corresponding 2-sided (at 0.05 alpha level) 95% CI and p-value will be provided. The number and percentage of responders will be summarized by each treatment group and the difference in proportions (luspatercept – placebo) and corresponding 95% CI will also be calculated by exact unconditional test with “rolling” method as well.

A forest plot showing the ORs, 95% CI and p-value for the erythroid response (33% and 50%) in each subgroup will be provided with “rolling” method (by 12-week or 24-week respectively) as well.

A waterfall plot will be provided for the transfusion burden percent change from baseline during any 12-week or 24-week interval by treatment group respectively. Individual subject’s transfusion burden percent change from baseline will be displayed in a single bar. The displayed transfusion burden percent change is each subject’s largest percent decrease in transfusion burden during any post-baseline 12-week or 24-week interval.

10.4. Other Efficacy Analyses

In general, descriptive statistics will be provided and statistical tests will be applied if appropriate. For continuous variables, LS means with corresponding SEs for each treatment group, along with LS mean of treatment difference (luspatercept versus placebo) with corresponding 95% CI and p-value will be presented for ANCOVA method. Kaplan-Meier methods will be used to analyze time to event variables. Counts and percentages will be used to describe categorical variables. If an ANCOVA method is used, the statistical assumption will be validated first, log transformation will be applied as needed.

10.4.1. Mean Change in Liver Iron Concentration

Quality Control for LIC measurements: An imaging charter (MRI Manual) was distributed to all participating sites to specify appropriate MRI imaging parameters and scanning techniques. All sites were required to submit one dummy MRI LIC assessment file for review of the imaging parameters prior to enrolling subjects into the study. Dummy scans were reviewed by an independent expert. Feedback was provided back to the sites if dummy scan review highlighted non-compliance with MRI manual. Following the baseline and 24 week visits, random and for-cause spot checks of source DICOM images were performed. Independent analysis of the LIC values was performed using techniques validated against MRI LIC acquisition (Wood et al, 2005). The sponsor was informed of the independent audit results.

The value of LIC will be either the value collected from eCRF or the value derived from T2*, R2* or R2 parameter depending on which techniques and software were used for MRI LIC acquisition and post-processing. The LIC value will be derived as below:

Technique	Site number	Derivation source	Alternative derivation if still missing
Ferriscan R2	PPD [REDACTED] [REDACTED]	Reported LIC from CRF	$(29.75 - \text{SQRT}(900.7 - 2.283 * \text{EXP}((-0.19043 + 1.016385 * \ln(R2))) / 0.983615))$

Technique	Site number	Derivation source	Alternative derivation if still missing
	PPD) ^{1.4265} .
T2*/R2*	PPD	$31.94 * (T2*)^{-1.014}$	$0.029*(R2*)^{1.014}$ If both T2* and R2* are missing, LIC from CRF will be used
	PPD	$25.4/T2* + 0.2$	$0.0254 * (R2*) + 0.2$ if site is not PPD. If both T2* and R2* are missing, LIC from CRF will be used. LIC from CRF will be used for site PPD
	PPD	$0.0254 * (R2*) + 0.2$	$25.4/T2* + 0.2$ If both R2* and T2* are missing, LIC from CRF will be used

The derived LIC value will be used for analysis. If a subject has any LIC value > 43, the subject's LIC value will be excluded from analysis. Note that, subjects with LIC value > 43 are not excluded from the LIC baseline summary and efficacy subgroup analysis by LIC categories.

Descriptive statistics for LIC measurements and change from baseline will be summarized at week 24/48. The 24/48-week LIC change from baseline will be analyzed using ANCOVA model with geographical regions defined at randomization and baseline LIC as covariates for the ITT population. Additionally, a shift table representing the shift from the baseline to week 24/48 category (≤ 3 ; $> 3 \leq 7$; $> 7 \leq 15$ and > 15) will be provided for LIC. A subject will have maximum two post-baseline LIC assessments (including "unscheduled") during the 48 week double-blinded treatment period per protocol. If a subject has only one assessment, it will be counted as "Week 48" visit; if a subject has two assessments, the first one will be counted as "Week 24" visit, and the later one as "Week 48" visit regardless of the collected nominal visit name. This logic will be used in model based summary, change from baseline summary and the shift table summary.

Additionally, bar plot will be provided for baseline, week 24 and week 48 LIC categories (≤ 3 ; $> 3 \leq 7$; $> 7 \leq 15$; > 15). Percent of subjects within each category will be displayed by treatment group. All the LIC data will be presented in a listing.

10.4.2. Mean Change in Mean Daily Dose of Iron Chelation Therapy

The ICT mean daily dose summary will be provided for subjects who did not change ICT drug from baseline to post-baseline and only one ICT drug has been used. Descriptive statistics for mean daily dose will be summarized at baseline and the post-baseline visit for each ICT drug. The baseline mean daily dose will be calculated using the ICT dosage during the 12 weeks on or prior to first study drug treatment and the post-baseline mean daily dose is calculated during the last 12 weeks of the 48-week double-blind treatment period or the last 12 weeks of the study treatment for early discontinued subjects.

The same descriptive statistic summary for baseline and post-baseline ICT drug mean daily dose will be provided for subjects in each baseline liver iron content category (≤ 3 mg/gr dry weight and > 3 mg/gr dry weight).

The change from baseline in mean daily dose at post-baseline will be analyzed using an ANCOVA model with the geographical regions defined at randomization and baseline ICT mean daily dose as covariates for the ITT population for subjects who did not change ICT drug from baseline to post-baseline and only one ICT drug has been used.

A summary showing the number and percentage of subjects who took monotherapy (i.e., only one ICT drug) vs. combo therapy (i.e., more than one ICT drug) at the 12 week baseline period and post-baseline period will be provided.

Bar plot will be provided for percent of subjects who took each ICT drug at baseline and post-baseline for subjects who did not change ICT drug from baseline to post-baseline and only one ICT drug has been used. A similar bar plot will be provided for percent of subjects who took monotherapy or combo therapy at baseline and post-baseline.

10.4.3. Mean Change in Serum Ferritin Level

Descriptive statistics for serum ferritin level will be summarized at baseline and the post-baseline visit, where the baseline mean serum ferritin is measured during the 12 weeks prior to the first dose and the post-baseline mean serum ferritin is calculated during the last 12 weeks of the 48-week double-blind treatment period or the last 12 weeks of the study treatment for early discontinued subjects. Change from baseline will be summarized at the post-baseline visit.

The change from baseline in serum ferritin at the post-baseline will be analyzed using an ANCOVA model with the geographical regions defined at randomization and baseline serum ferritin value as covariates for the ITT population.

All the serum ferritin data will be presented in a listing.

10.4.4. Bone Mineral Density Assessed by DXA Scan

A subject will have only one post-baseline DXA assessment during the 48 week double-blinded treatment period per protocol amendment 1. The only post-baseline assessment will be counted as “week 48” visit regardless of the collected nominal visit name. For patients enrolled before protocol amendment 1, there would be maximum 2 DXA assessments during the 48 week double-blinded treatment period. In this case, if a subject has only one assessment, it will be counted as “Week 48” visit; if a subject has two assessments, the first one will be counted as “Week 24” visit, and the later one as “Week 48” visit regardless of the collected nominal visit

name. The analysis will be done on subjects that have at least two measurements (one baseline and one post-baseline measurement). Descriptive statistics for BMD and T-score baseline, post-baseline, and change from baseline will be summarized for week 48 visit. Change from baseline at Week 48 will be analyzed using ANCOVA model with the geographical regions defined at randomization and baseline measurement as covariates for the ITT population. All the DXA data will be presented in a listing.

10.4.5. Mean Change in Myocardial Iron by T2* MRI

A subject will have one post-baseline myocardial assessment during the 48 week double-blinded treatment period per protocol. The only post-baseline assessment will be counted as “week 48” visit regardless of the collected nominal visit name. Descriptive statistics for myocardial iron measurements and change from baseline will be summarized for week 48 visit. The 48-week myocardial iron change from baseline will be analyzed using an ANCOVA model with the geographical regions defined at randomization and baseline myocardial iron as covariates for the ITT population.

A shift table representing the shift from the baseline to week 48 category (≤ 10 ms and > 10 ms) will be provided for Myocardial Iron T2* value.

A histogram plot of myocardial iron by T2* will be provided for baseline and week 48.

All the myocardial iron data will be presented in a listing.

10.4.6. Change in Quality of Life assessed by TranQol and SF-36

TranQol:

The TranQol is a disease-specific, self-administered, well-validated health-related quality of life tool developed for beta-thalassemia patients (Klaassen, 2014; Klonizakis, 2017). The adult self-report version, used in this study, includes 36 questions assessed on a 5-point response, that are grouped into 5 domains (Physical Health, Emotional Health, Sexual Health, Family Functioning, School/Career Functioning). Scores are calculated according to author’s guidelines and scoring rules (please refer to appendix 18.3.1 for calculation algorithms). Both the total score and the domain scores range from 0 (worst) to 100 (best).

To interpret the difference in change score from baseline (screening), between treatment groups and change score at the individual level, 2 threshold values are usually used: 1) the minimally important differences (MIDs), used as a benchmark to interpret mean score difference between groups as clinically meaningful in a clinical trial (FDA, 2006); and 2) the responder definitions (RDs), defined as the individual patient HRQoL score change over a predetermined time period that should be interpreted as a treatment benefit (FDA, 2009). Both thresholds have not yet been well established. However, data reported in literature seemed to suggest a change of 4–6 points in the total TranQoL score can be considered as the RD. However, as MIDs for the domains of the TranQoL are still not available from literature, $0.3 \times \text{SD}$ of the domain scores at baseline (screening) from the pooled data will be used as proxies for MIDs. The $0.5 \times \text{SD}$, which is generally considered as an approximation for RD (Norman, 2003; Norman, 2004) will not be used for MIDs because it is intended for differences in individual respondent scores rather than in group respondent scores, and there is an emerging consensus that the criterion of $0.5 \times \text{SD}$ is considered too high for MID (Maruish, 2011).

SF-36:

The SF-36 is a generic, self-administered instrument consisting of 8 multi-item scales that assess 8 health domains (Maruish ME, 2011; McHorney, 1994; Ware, 1992): Physical functioning (PF), Role-Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social functioning (SF), Role-Emotional (RE), Mental Health (MH). Two summary scales, Physical Component Summary (PCS) and Mental Component Summary (MCS), will be calculated using norm-based scores from the 8 health domains. The primary interests of the SF-36 are the 8 health domain scores and the PCS and MCS scores. Scores are calculated according to author's guidelines and scoring rules (please refer to appendix 18.3.2 for calculation algorithms). The raw score for each health domain can be transformed into a 0 (worst) to 100 (best) domain score, which can be transformed into norm-based T-scores, with a mean of 50 and a standard deviation (SD) of 10. Higher norm-based T-scores indicate better health/QoL based on data from a nationally representative sample of adults from the US

In order for one health domain scale to be meaningfully compared with the other scales and for domain scores to have a direct interpretation in relation to the distribution of scores in the US general population, the 0–100 scale score for each health domain can be converted to norm-based scores using a T-score transformation, with a mean of 50 and a standard deviation (SD) of 10. Higher norm-based T-scores indicate better health/QoL, based on data from a nationally representative sample of adults from the US. Two summary scales, the Physical Component Summary (PCS) and Mental Component Summary (MCS), in norm-based metric, can also be calculated from these eight health domains. The SF-36 has been thoroughly assessed, showing good psychometric properties.

Table 1 describes the range of possible T-scores, minimally important difference and responder definition for all SF-36 scale scores.

Table 1: Composition and Interpretation of T-Scores for SF-36 (Version 2) Component Summary Measures and Health Domain Scales

Scale/Measure	Composition	Range of Possible T-Scores		MID	RD
Physical functioning (PF)	Items 3a–3j	19.26	57.54	3.0	4.3
Role-physical (RP)	Items 4a–4d	21.23	57.16	3.0	4.0
Bodily pain (BP)	Items 7, 8	21.68	62.00	3.0	5.5
General health (GH)	Items 1, 11a–11d	18.95	66.50	2.0	7.0
Vitality (VT)	Items 9a, 9e, 9g, 9i	22.89	70.42	2.0	6.7
Social functioning (SF)	Items 6, 10	17.23	57.34	3.0	6.2
Role-emotional (RE)	Items 5a–5c	14.39	56.17	4.0	4.6
Mental health (MH)	Items 5a–5c	11.63	63.95	3.0	6.7
PCS	All scales	5.02	79.78	2.0	3.8
MCS	All scales	-3.33	80.09	3.0	4.6

MCS = Mental Component Summary; MID = minimally important difference; PCS = Physical Component Summary; RD = responder definition

*Highest and lowest observed T-scores in 2009 general population normative sample

Statistical Analyses

For both QOL endpoints, records beyond week 48 will be used.

The preselected primary domains of interest for the assessment are:

- TranQoL
 - Total score
 - Physical Health domain
- SF-36
 - Physical Functioning domain
 - General Health domain
 - Physical Component Summary

For the assessment of changes from baseline (screening), the analysis of scores at Week 24 will be considered as primary; CCI [REDACTED]

To assess the extent of missing data at each assessment visit by treatment group, compliance rates for the TranQoL and SF-36 will be estimated on the ITT population separately based on the number of subjects included in the ITT population per treatment group who are eligible for assessment at a given scheduled visit.

Subjects will be considered compliant with completion of the TranQoL if at least 75% of the items are non-missing (ie, ≥ 27 items of the 36 items completed or non-missing total score) for a given assessment visit, and compliant with completion of the SF-36 assessment if at least half of the 36 items (ie, ≥ 18 items) are completed.

To assess the effect of luspatercept + BSC versus placebo + BSC on health-related QoL, the key analysis below will be performed based on the HRQoL evaluable population (Section 5.2.3).

To determine whether the findings of the primary analyses are generalizable to the entire ITT population, the comparability of the HRQoL evaluable and non-evaluable populations will be assessed at baseline (screening). The HRQoL non-evaluable population in this particular analysis will be defined as those subjects in the ITT population who are not included in both HRQoL-evaluable populations.

A cross-sectional analysis of change from baseline (screening) will be performed to compare the scores at Week 24 and Week 48 between treatment groups using ANCOVA models adjusted for baseline (screening) domain scores and randomization stratification factors. The least squares (LS) mean (95% CI and p-value) for changes from baseline (screening) at each postbaseline/screening visit in all domain scores within each treatment group, and the difference in the LS means (95% CI, p-value) between treatment groups at each postbaseline/screening visit will be estimated.

To avoid bias when interpreting differences between groups in HRQoL score changes from baseline (screening) to last QoL assessment, the corresponding treatment duration will be described per treatment groups for patients analyzed.

A number of sensitivity analyses, including analysis with imputation of missing data, are planned to assess the robustness of HRQoL findings from the main analysis and the impact of missing data.

Additional details of the highlighted analyses as well as sensitivity analyses, and additional analyses to assess subgroups and treatment effects on HRQoL are provided in a separate HRQoL statistical analysis plan, which will be finalized prior to database lock. The HRQoL statistical analysis plan will be appended to the separate specific HRQoL report.

The raw scores for each individual question and the calculated domain and component scores will be presented in the listings for TranQoL and SF-36 respectively.

10.4.7. Duration of Transfusion Burden Reduction

The duration of the longest continuous 12-week based erythroid response (based on 33% and 50% criteria) during the entire study period up to the efficacy cutoff will be summarized by the Kaplan–Meier method (LOGLOG transformation will be used). The 33% response is defined as subjects with $\geq 33\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units. The 50% response is defined as subjects with $\geq 50\%$ reduction from baseline in RBC transfusion burden with a reduction of at least 2 units. The median duration of response, 25th and 75th quartiles with the associated 2-sided (at 0.05 alpha level) 95% CIs will be presented for each treatment group. The min and max of duration will be provided for all responders. The Kaplan-Meier plots of the response duration, defined as time from the start of the longest response to end of the response (see below) will also be provided. Only subjects who achieve a response will be included in the analysis.

The duration of the individual continuous response is defined as Last Day of Response – First Day of Response + 1,

where

First Day of Response = the first day of the first 12-week interval when the subject meets response.

Last Day of Response = the last day of the last 12-week interval when the subject meets response.

The subject has to meet response on all the days within the above duration.

For subjects who have one response and continue to respond at the efficacy cutoff date, the end day of the response will be censored at the date of efficacy cutoff and the duration of response will be calculated as date of efficacy cutoff – first day of response + 1,

where date of efficacy cutoff is defined in section 10.

For subjects who have multiple responses and the last one continue to respond at the efficacy cutoff date, the longest response will be the last one if the duration from response start to censoring is longer than all the previously occurred response durations. If the continuing response duration is not the longest compared with the previously occurred responses, the

response with longest duration will be selected. Summary statistics will be provided for total duration of transfusion burden reduction (33% and 50% criteria) within the entire study period and the ratio of total response duration versus entire study duration. The entire study duration is defined as period from day 2 to date of efficacy cutoff.

10.4.8. Time from First Dosing Date to the First Erythroid Response

The descriptive statistics for the time from first dosing date to the first erythroid response (for both 33% and 50% criteria) will be provided by treatment group, where time from first dosing date to the first erythroid response is defined as First Day of Response – Date of First Study Drug +1. The difference in time from first dosing date to the first erythroid response (luspatercept – placebo), corresponding 95% CI and p-value will be calculated by t-test.

Only subjects who have a response will be included.

10.4.9. Transfusion Independence

The number and percent of subjects who achieve transfusion independence will be summarized by treatment group and the treatment comparison will be analyzed using the CMH model stratified by the geographical regions defined at randomization. Transfusion independence is defined as the absence of any transfusion during any consecutive “rolling” 6-week, or 8-week or 12-week time interval within the entire study period up to the efficacy cutoff date, i.e., Day 2 to 43, Day 3 to 44, ..., Day x to efficacy cutoff date for 6-week interval (or Day 2 to 57, Day 3 to 58, ..., Day x to efficacy cutoff date for 8-week interval, or Day 2 to 85, Day 3 to 86, ..., Day x to efficacy cutoff date for 12-week interval). Subjects whose efficacy cutoff date that is before day 43 (for 6-week based), or day 57 (for 8-week based) or day 85 (for 12-week based) will be counted as non-responders.

The duration of transfusion independence will be summarized by the Kaplan–Meier method. The median duration of response, 25th and 75th quartiles with the associated 2-sided (at 0.05 alpha level) 95% CIs will be presented for each treatment group. The min and max of duration will be provided for all responders. The Kaplan-Meier plots of the transfusion independence duration, defined as time from first response to end of response (see below) will also be provided. Only subjects who achieve a response will be included in the analysis.

The duration of transfusion independence will be calculated similar to the duration of erythroid response, which is defined as Last Day of Response – First Day of Response +1,

where

First Day of Response = the first day of the first 6-week (or 8-week or 12-week) interval when the subject meets response.

Last Day of Response = the last day of the last 6-week (or 8-week or 12-week) interval when the subject meets response.

For subjects who continue to respond at the efficacy cutoff date, the end day of the response will be censored at the date of efficacy cutoff and the duration of response will be calculated as date of efficacy cutoff – first day of response +1,

where date of efficacy cutoff is defined in section 10.

10.4.10. Post-baseline Transfusion Events Frequency

The post-baseline transfusion event frequency during the individual 12-week interval (week 1-12, week 13-24, week 25-36, week 37-48, week 49-60, etc...until the maximum available 12-week interval) and the overall post-baseline transfusion frequency will be analyzed using negative binomial regression with the geographical regions defined at randomization and baseline transfusion frequency (12 weeks on or prior to Dose 1 Day 1) in the regression model. The p-value will be provided from the model. For the definition of transfusion events, if multiple transfusions happen on the same date, they are counted as one event; if multiple transfusions happen on two consecutive dates, they are counted as one event; if multiple transfusions happen on three consecutive dates, they are counted as two events.

The same summary statistics for transfusion event frequency will be repeated for erythroid responders (33% and 50% during any 12-week interval respectively).

10.4.11. Pre-transfusion Hemoglobin Change from Baseline

To estimate the change of pre-transfusion hemoglobin value after dosing within each 12-week interval (week 1-12, week 13-24, week 25-36, week 37-48), summary statistics for the baseline, post-baseline and change from baseline pre-transfusion hemoglobin values will be provided by treatment group. The number and percentage of subjects meeting selected change category (Decrease ≥ 2 g/dL, Decrease ≥ 1.5 -<2 g/dL, Decrease ≥ 1 -<1.5 g/dL, Decrease ≥ 0.5 -<1 g/dL, Decrease >0 -<0.5 g/dL, Increase ≥ 0 -<0.5 g/dL, Increase ≥ 0.5 -<1 g/dL, Increase ≥ 1 -<1.5 g/dL, Increase ≥ 1.5 -<2 g/dL and Increase ≥ 2 g/dL) will be provided separately. Baseline pre-transfusion hemoglobin is defined as mean of all documented pre-transfusion hemoglobin values collected during the 24 weeks prior to Dose 1 Day 1. Post-baseline pre-transfusion hemoglobin is defined as mean of all documented pre-transfusion hemoglobin values collected during each 12-week interval. The same summary statistics and categorized pre-transfusion hemoglobin change from baseline analyses will be repeated for erythroid responders (33% and 50% at week 13-24), non-responders (33% and 50% at week 13-24), and responders (33% and 50% during any 12-week interval) respectively.

In addition, hemoglobin values from central lab will be summarized in the same way by 12-week interval. Only hemoglobin measurements on the same day as transfusion or 14 days after transfusion date will be included, i.e., if a hemoglobin measurement occurs within 14 days after any transfusion, the hemoglobin value will be excluded from analysis.

10.4.12. Healthcare Resource Utilization

Summary statistics of the number of subjects who had a doctor office visit (non-study scheduled) or emergency room visit, or a hospitalization after signing informed consent, and the number of subjects hospitalized in a higher level of care unit by type (e.g., Intensive care unit, Coronary care unit, Other, Missing), the number of days in higher care units during the study will be presented by treatment group. Number of days of hospitalization will be defined as (hospitalization end date – hospitalization start date) + 1. If hospitalization has unknown start or/and end date, it will be counted in 'Missing' category.

Reasons for inpatient hospitalization will be summarized categorically using number and percentage of subjects with the following categories:

- Adverse Events
- Protocol-driven procedure (e.g., bone marrow aspiration)
- Non-protocol driven assessments or procedure (e.g., ultrasound)
- Transfusion (e.g., RBC, platelets)
- Procedure planned prior to signing informed consent (e.g., coronary arteriogram)
- Elective procedure for a pre-existing condition (e.g., hernia repair)
- Social, technical or practical reason in the absence of an Adverse Event (e.g., travel distance from the clinic prohibitive for study participation).

A listing will be provided for healthcare utilization.

10.5. Subgroup Analysis

The primary, key secondary efficacy endpoints, and the 33%, 50% erythroid response endpoint by 'rolling' base interval will also be summarized for the following subgroups:

1. Geographic region:
 - North America and Europe
 - Middle East and North Africa
 - Asia-Pacific
2. Age:
 - ≤ 32 years
 - > 32 years
3. Splenectomy:
 - Yes
 - No
4. Sex :
 - Male
 - Female
5. Beta-thalassemia gene mutation grouping:
 - B0/B0
 - Non-B0/B0

6. Baseline transfusion burden
 - ≤ 6 units/12 weeks
 - > 6 units/12 weeks
7. Mean pre-transfusion hemoglobin at baseline
 - < 9 g/dL
 - ≥ 9 g/dL
8. Baseline liver iron content
 - ≤ 3 mg/g dry weight
 - $> 3-\leq 7$ mg/g dry weight
 - $> 7-\leq 15$ mg/g dry weight
 - > 15 mg/g dry weight
9. Baseline transfusion burden (units/24 weeks): for endpoints based on 24-week rolling method only
 - Low transfusion burden (≤ 10 units/24 weeks)
 - Medium transfusion burden ($> 10-\leq 15$ units/24 weeks)
 - High transfusion burden (> 15 units/24 weeks);
10. Baseline transfusion burden (units/12 weeks): for endpoints based on 12-week rolling method only
 - Low transfusion burden (≤ 5 units/12 weeks)
 - Medium transfusion burden ($> 5-\leq 7$ units/12 weeks)
 - High transfusion burden (> 7 units/12 weeks);

10.6. Missing Data Imputation

In case of any missing data for RBC transfusion units records and MRI liver iron content, imputation will be applied for each section.

The imputation for RBC transfusion units is stated in section 10: if at the time of data summary, a subject's efficacy cutoff date is before the end of the 12-week interval or a subject has any invalid transfusion records (i.e., transfusion unit not available) during the specified 12 week interval, this subject will be included in the analysis as a non-responder.

The imputation logic for missing LIC value is stated in section 10.4.1: the value of LIC will be either the value collected from eCRF or the value derived from T2*, R2* or R2 parameters depending on which techniques and software were used for MRI LIC data acquisition.

11. SAFETY ANALYSIS

The purpose of this section is to define the safety parameters for the study. All summaries of safety data will be conducted using the safety population. The safety analysis includes adverse events (AEs), clinical laboratory tests, vital signs, electrocardiogram (ECG), cardiac Doppler or Multi Gated Acquisition Scan (MUGA), and antidrug antibody (ADA) testing. In addition, pregnancy test and menstrual status assessments will be provided for female subjects.

If not otherwise specified (for example, AEs), safety summaries will use all collected records.

11.1. Adverse Events

Adverse events will be analyzed in terms of treatment-emergent adverse events (TEAEs) which are defined as any AEs that begin or worsen on or after the start of study drug through 63 days after the last dose of IP (i.e., AE start date on or after the first dose date and within last dose date + 63). In addition, an AE that occurs beyond this timeframe and that is assessed by the investigator as possibly related (suspected) to study drug will be considered to be treatment-emergent.

All AEs will be coded using the MedDRA (Version 20.0). The incidence of TEAEs will be summarized by MedDRA SOC and PT. The AE tables will be sorted by SOC and PT (within SOC) in descending frequency within the luspatercept group. If a subject experiences multiple AEs under the same PT (or SOC), then the subject will be counted only once for that PT (or SOC).

The intensity of AEs will be graded 1 to 5 according to the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. If a subject experiences the same AE more than once with different toxicity grades, then the event with the highest grade will be tabulated in “by grade” tables. In addition, AEs with a missing intensity will be presented in the summary table as an intensity category of “Missing” only if the same event category has no other valid grades.

Tables summarizing the incidence of TEAEs will be generated for each of the following by treatment group (if not otherwise specified, the summary is by SOC and PT):

- TEAEs;
- TEAEs by SOC only;
- Treatment-related TEAEs;
- Serious TEAE;
- Treatment-related serious TEAEs;
- TEAEs by CTCAE maximum severity;
- Treatment-related TEAEs by CTCAE maximum severity;
- TEAE with CTCAE Grade ≥ 3 ;
- Treatment-related TEAE with CTCAE Grade ≥ 3 ;

- TEAEs leading to study drug discontinuation;
- Treatment-related TEAEs leading to study drug discontinuation;
- TEAEs leading to study drug dose reduction;
- Treatment-related TEAEs leading to study drug dose reduction;
- TEAEs leading to study drug dose delay;
- Treatment-related TEAEs leading to study drug dose delay;
- TEAEs leading to death;
- Treatment-related TEAEs leading to death;
- Most frequent TEAEs by PT ($\geq 5\%$ in PT frequency of all subjects or subjects from any treatment group); Most frequent TEAEs by SOC ($\geq 5\%$ in SOC frequency of all subjects); Most frequent TEAEs by SOC and high level term (HLT) ($\geq 5\%$ in HLT frequency of all subjects)
- TEAEs by Age group (≤ 32 years, > 32 years);
- TEAEs by Age group (≤ 32 years, $32-\leq 50$ years, > 50 years);
- TEAEs by Gender (Male and Female);
- TEAEs by Splenectomy status (Yes/No);
- [REDACTED]
- TEAE by preferred term for QTC Prolongation and Atrial Fibrillation events;
- All death by cause of death;

Listings for AEs, SAE, and AEs leading to discontinuation will be presented separately. Treatment-emergent AEs will be flagged in the listings. A death listing will be provided for all death events.

11.2. Adverse Events of Special Interest (AESI)

The following adverse events are of special interest:

- Malignancy
- Premalignant Conditions

An AESI summary by PT terms will be provided for TEAEs by treatment group.

A listing for AESI will be provided as support.

11.3. Other Adverse Events That Require Safety Analysis

Other adverse events that require safety analysis include AEs that fall under the “Embolic and Thrombotic Events” SMQ category and AEs with PT “Bone pain”.

Similar to AESI, summary will be provided for AEs that fall under the “Embolic and Thrombotic Events” SMQ category by preferred term. For “Bone pain” events, number of subjects with bone

pain events by worst CTCAE grade will be provided from grade 1 to grade 5. The number and percentage of subjects with bone pain occurred during the first 24 weeks and after 24 weeks will be provided. Summary statistics will also be provided for time to first bone pain, total duration of bone pain (in days), total duration of bone pain that occurred during the first 24 weeks (in days), and total duration of bone pain that occurred after 24 weeks (in days). Total duration of bone pain is defined as sum of all bone pain duration within a subject, excluding overlapped period.

The “Embolic and Thrombotic Events” SMQ category analysis will include subgroup analysis by splenectomy status, platelet above ULN, concomitant medications, and comorbidities as needed.

11.4. Clinical Laboratory Evaluations

Clinical laboratory data is collected by central lab and local lab (if relevant to dose administration, modification and AE, or when no central lab results are obtained). Central laboratory assessments include hematology, chemistry, immunology, and pregnancy test. Local laboratory assessments include hematology, chemistry and urinalysis. Lab data will be collected over time during the study. All summaries will be based on the SI units and missing values will not be imputed. Clinical laboratory values will be graded (grade 0-4) according to NCI-CTCAE version 4.0 for applicable tests. Normal ranges will be used to determine the “High”, “Low”, and “Normal” categories for all numeric laboratory tests. Only central lab results and local lab “Reticulocyte (Blood)” parameter will be used for table summaries.

11.4.1. Hematology/Chemistry/Immunology

The laboratory results and change from baseline will be summarized by visit by treatment group for central lab hematology and chemistry panels separately.

A shift table representing the shift from the baseline grade to maximum NCI-CTC Grades (high or low) will be provided for selected hematology and chemistry parameters having toxicity grade by visit by treatment group. The shift summary for high category will be done for hepatic function parameter (ALT, AST, ALP, and total bilirubin) and renal function parameter (serum creatinine). The shift summary for low category will be done for hematology parameter (platelets, leukocytes and absolute neutrophil counts).

To estimate the incidence of subjects who have passed the predefined threshold for selected parameters, a summary table representing the number and percentage of subjects with lab assessments satisfying the threshold criteria will be provided by treatment group. A subject with post-baseline result (including “unscheduled” visits) meeting the criteria will be counted. The threshold criteria includes below:

LIVER FUNCTION	
Post-baseline Alanine Aminotransferase (ALT)	$\geq 3x$ upper limit of normal (ULN)
Post-baseline Aspartate Aminotransferase (AST)	$\geq 3x$ ULN
Post-baseline Direct Bilirubin (BILDIR)	$\geq 2x$ ULN
Post-baseline ALT/AST and BILDIR	(ALT $\geq 3x$ ULN or AST $\geq 3x$ ULN) and BILDIR $\geq 2x$ ULN

RENAL FUNCTION	
Post-baseline Creatinine Clearance (CREATCLR)	< 0.5x baseline
Post-baseline Serum Creatinine (CREAT)	> 2x baseline
Albuminuria Category (ACR: mg/g)	<30
	≥30- <300
	>300-≤1000
	>1000-≤3500
	>3500
HEMATOLOGY	
Post-baseline Leukocytes (WBC)	≥ 2x baseline and > ULN
	≥ 3x baseline and > ULN
	≥ 2x baseline and > ULN and lasts for at least 42 days
	≥ 3x baseline and > ULN and lasts for at least 42 days
Maximum Post-baseline Platelets (PLAT)	≥ 1.5x baseline and > ULN
	≥600 - <1000x10 ⁹ /L
	≥ 1000 x 10 ⁹ /L

Specifically, the threshold summary for platelets will be based on the maximum post-baseline value. The summary will be provided by baseline category based on normal range (i.e., within normal limit at baseline, > upper limit at baseline), and by splenectomy status (Yes, No) in separate tables. Furthermore, the number and percentage of subjects with maximum post-baseline WBC exceeding 3x baseline value and >ULN, maximum post-baseline WBC exceeding 3x baseline value and > ULN and lasts for at least 42 days, and subjects with maximum post-baseline platelets ≥600 - <1000x10⁹/L and maximum post-baseline platelets exceeding 1000 x 10⁹/L will be provided separately by splenectomy status (Yes, No).

For some key lab parameters (ALT, AST, WBC), plots will be presented to show the pattern of the lab test values over time by treatment group. Mean and SE will be presented in the plot.

Listings of clinical laboratory data will be provided for central lab and local lab respectively for each panel (excluding serum erythropoietin and serum ferritin). Abnormal observations will be noted. Specifically, subjects with any WBC differential count exceeding 2x baseline value will be listed in a separate listing. All WBC differential count records of qualified subjects will be presented.

11.4.2. Serum Erythropoietin and Serum Ferritin

Serum erythropoietin and serum ferritin are collected from central laboratory. The summary of serum erythropoietin test results and change from baseline will be provided by visit by treatment

group. The summary of serum Ferritin is described in section 10.4.3 as a study endpoint. A plot will be presented to show the pattern of the serum erythropoietin test results over time by treatment group. Mean and SE will be presented in the plot.

A listing will be provided for both serum erythropoietin and serum ferritin.

11.4.3. Local lab “Reticulocyte(Blood)” parameter

The “Reticulocyte (Blood)” parameter is only collected at local lab. The summary of absolute reticulocyte count and change from baseline will be provided by visit and treatment group in the same way as other lab parameters.

A line plot will be presented to show the pattern of the reticulocyte test results over time by treatment group. Mean and SE will be presented in the plot.

11.5. Vital Sign Measurements

Vital sign is collected over time during the study. Vital sign parameters include weight, temperature, pulse rate, seated blood pressure (diastolic blood pressure (DBP) and systolic blood pressure (SBP)). The DBP and SBP are collected twice at each visit with 10 minutes apart. The average of the two assessments will be used. Summary statistics of observed values and change from baseline values will be presented for each parameter by visit by treatment group.

To further estimate the incidence of subjects whose maximum post-baseline blood pressure have passed selected criteria, summary tables representing the number and percentage of subjects with post-baseline (including ‘unscheduled’ visits) SBP/DBP assessments satisfying each criteria will be provided by treatment group. The selected criteria includes below:

Maximum post-baseline SBP	No increase
	Increased < 20 mmHg
	Increased \geq 20 mmHg
	Increased \geq 20 mmHg and SBP \geq 140 mmHg
	Increased \geq 20 mmHg and SBP \geq 150 mmHg
	Subjects only with baseline values
Maximum post-baseline DBP	No increase
	Increased < 20 mmHg
	Increased \geq 20 mmHg
	Increased \geq 20 mmHg and DBP \geq 100 mmHg

A plot will be presented to show the pattern of the SBP and DBP test results over time by treatment group. Mean and SE will be presented in the plot. Additionally, a spaghetti plot for SBP and DBP values overtime for individual subjects with maximum post-baseline SBP increased \geq 20 mmHg and SBP \geq 150 mmHg or maximum post-baseline DBP increased \geq 20 mmHg and DBP \geq 100 mmHg will be provided.

Corresponding listing will be provided for vital sign data.

11.6. Electrocardiograms

The 12-lead electrocardiogram (ECG) is collected over time at selected visits. ECG parameters include heart rate, PR interval, QRS duration, RR interval and QT. The RR interval value will be derived per formula: $RR \text{ interval (msec)} = 60000 \text{ (msec)} / \text{heart rate (bpm)}$.

The corrected value for QT interval will be derived based on Fridericia's formula as below:

Fridericia's formula: $QTcF = QT / (RR)^{1/3}$

where RR is the calculated RR interval as above.

The calculated RR interval value, recorded values of other ECG parameters and change from baseline values will be summarized by visit by treatment group.

To further estimate the incidence of subjects whose baseline or post-baseline QTcF values have passed the selected ICH E14 Criteria, summary tables representing the number and percentage of subjects with ECG assessments satisfying the CPMP (Committee for Proprietary Medicinal Products) criteria will be provided for QTcF by treatment group and visit (for baseline and post-baseline respectively). A subject with baseline or any post-baseline (including 'unscheduled' visits) result meeting individual criteria will be counted. The selected CPMP criteria includes below:

Baseline/Post-baseline QTcF Interval	> 450 msec
	> 480 msec
	> 500 msec
QTcF Interval Increase from Baseline	≥ 30 msec
Post-baseline QTcF Interval and Increase from Baseline	Post-baseline Interval > 480 msec and Increase from Baseline ≥ 60 msec

Corresponding listing will be provided for ECG data.

11.7. Cardiac Doppler or Multi Gated Acquisition Scan

The Left ventricular ejection fraction (LVEF) is collected over time at selected visits. It will be measured by either echocardiography (ECHO), Multi Gated Acquisition Scan (MUGA) or MRI. Recorded values of LVEF and change from baseline values will be summarized by treatment group and by visit.

Corresponding listing will be provided for LVEF data.

11.8. ECOG Performance Status

The eastern cooperative oncology group (ECOG) scale is used to assess how the disease affects subjects' daily activities. ECOG is classified into 6 categories:

0 = fully active, able to carry on all pre-disease performance without restriction;

1 = restricted in physically strenuous activity but ambulatory and able to carry out work of a light

or sedentary nature, e.g., light housework, office work;

2 = ambulatory and capable of all self-care but unable to carry out any work activities, up and about more than 50% of waking hours;

3 = capable of only limited self-care, confined to bed or chair more than 50% of waking hours;

4 = completely disabled, cannot carry on any self-care, totally confined to bed or chair;

5 = dead.

The ECOG status at screening visit is summarized in the demographic table. A listing will be provided for ECOG data.

11.9. Antidrug Antibody Testing

The anti-luspatercept antibody test will be conducted over time. There are 4 ADA parameters:

CCI Titer information will be collected CCI. Specificity test, Nab and titer data are only available for subjects who are positive CCI.

To evaluate the treatment-emergent ADA level, the number and percentage of positive ADA result will be provided by parameter. The summary will be broken down by ADA positive categories (CCI) within each treatment: “Preexisting”, “Treatment-Emergent” and “Positive total”. The “Positive total” category is the sum of “Preexisting” and “Treatment-Emergent”. A subject is counted as ‘Treatment-Emergent’ if there is a positive post-baseline sample while the baseline sample is ADA negative, or there is a positive post-baseline sample with a titer \geq 4-fold of the baseline titer while the baseline sample is ADA positive. A subject is counted as ‘Preexisting’ if the baseline sample is ADA positive and the subject is not qualified for ‘Treatment-Emergent’.

A separate table will summarize the ADA titer information. For placebo group, only subjects with ADA sample collected are included; for luspatercept group, only subjects who are positive CCI will be included. The median, min and max value will be provided for ADA titer by treatment group and visit. Specifically, the titer summary for luspatercept subjects will be split to: “Preexisting”, “Treatment-Emergent” and “positive total” groups as defined above. Corresponding listing will be provided to support the table.

Additionally, a bar plot will be provided for subject’s ADA status (“Preexisting”, “Treatment-Emergent” and “Negative”). Percent of subjects within each category will be displayed by treatment group.

11.10. Pregnancy Test and Menstrual Status for Female Subjects

The number and percentage of the subjects for each pregnancy test result category (i.e., positive, negative) will be presented by treatment group. A subject is counted as ‘positive’ if there is any positive result captured after first dose date, a subject is counted as ‘negative’ if there is no positive result captured after first dose date.

The pregnancy test along with menstrual status will be provided in a listing.

CCI

[REDACTED]

CEL GENE PROPRIETARY INFORMATION

13. PK ANALYSIS

This SAP does not provide the details of statistical methods for PK analysis, which will be developed in a separate analysis plan. PK data will be provided in a listing for ITT set, including sample date/time and luspatercept concentration.

CELGENE PROPRIETARY INFORMATION

14. QUALITY OF LIFE ANALYSIS

The QoL analyses are addressed in section 10.4.6. This SAP only provides description of main QoL analysis. A detailed statistical analysis of QoL data will be provided in a separate HRQoL SAP, which will be finalized prior to database lock. The HRQoL SAP will be appended to the separate specific HRQoL report.

15. GENERAL INFORMATION

There is no interim analysis planned for this study. There are two clinical study reports (CSR) planned: one CSR report for marketing authorization application (MAA), and one final CSR report. This SAP will only address the MAA CSR report.

15.1. CSR for MAA

The MAA CSR will include safety and efficacy parameters at the time of the final analysis when all subjects completed 48 weeks of a double-blind Treatment Period or discontinued before reaching 48 weeks. Upon which, data base will be locked and the study will be unblinded. The CSR for MAA will include primary/secondary efficacy endpoints and safety endpoints.

15.2. Final CSR

The final CSR will include efficacy and safety data at the time of the End of the Trial (i.e., when all subjects initially assigned to luspatercept in the double-blind Treatment Period have completed the total treatment duration of 5 years from the subject's Dose 1 Day 1 or have discontinued earlier and completed the 156-week Post-treatment Follow-up Period, whichever occurred later.). The final analyses will be conducted on primary/secondary efficacy endpoints and safety endpoints. The analysis plan for final CSR will be addressed in a separate SAP.

15.3. DMC

An independent DMC will review the unblinded safety data. If the p-value for primary endpoint comparison crosses the O'Brien-Fleming boundary, with supportive other efficacy and safety data, the DMC may recommend stopping of further randomization. The Sponsor will then discuss the results in the context of the totality of the data package with the Health Authority for next steps. Meanwhile, the study will continue as planned until the Sponsor and the Health Authority have reviewed the information and reached an agreement on the path forward.

Operational details for the DMC and the algorithm and its validation by an expert panel will be detailed in the DMC charter.

16. REFERENCES

- Meloni A, Rienhoff HY, Jr., Jones A, Pepe A, Lombardi M, Wood JC. The use of appropriate calibration curves corrects for systematic differences in liver R2* values measured using different software packages. *Br J Haematol.* 2013;161(6):888-891.
- Wood JC, Enriquez C, Ghugre N, et al. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood.* 2005;106(4):1460-1465.
- Hankins JS, McCarville MB, Loeffler RB, et al. R2* magnetic resonance imaging of the liver in patients with iron overload. *Blood.* 2009;113(20):4853-4855.
- Garbowski MW, Carpenter JP, Smith G, et al. Biopsy-based calibration of T2* magnetic resonance for estimation of liver iron concentration and comparison with R2 Ferriscan. *J Cardiovasc Magn Reson.* 2014;16:40.
- St Pierre TG, Clark PR, Chua-anusorn W, et al. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. *Blood.* 2005;105(2):855-861.
- St Pierre TG, El-Beshlawy A, Elalfy M, et al. Multicenter validation of spin-density projection-assisted R2-MRI for the noninvasive measurement of liver iron concentration. *Magn Reson Med.* 2014;71(6):2215-2223.
- Ramazzotti A, Pepe A, Positano V, et al. Multicenter validation of the magnetic resonance T2* technique for segmental and global quantification of myocardial iron. *J Magn Reson Imaging.* 2009;30(1):62-68.
- Bacigalupo L, Paparo F, Zefiro D, et al. Comparison between different software programs and post-processing techniques for the MRI quantification of liver iron concentration in thalassemia patients. *Radiol Med.* 2016;121(10):751-762.
- Caocci G, La Nasa G, Efficace F. Health-related quality of life and symptom assessment in patients with myelodysplastic syndromes. *Esp Rev Hematol.* 2009;2:69-80
- Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH), U.S. Department of Health and Human Services. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Draft Guidance. Feb, 2006; <https://www.fda.gov/ohrms/dockets/98fr/06d-0044-gdl0001.pdf>. Accessed August 12, 2017a.
- Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH), U.S. Department of Health and Human Services. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Dec, 2009;

<https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf>. Accessed August 12, 2017.

Klaassen RJ, Barrowman N, Merelles-Pulcini M, et al. Validation and reliability of a disease-specific quality of life measure (the TranQol) in adults and children with thalassaemia major. *Br J Haematol*. 2014;164(3):431-7.

Klonizakis P, Klaassen R, Sousos N, Liakos A, Tsapas A, Vlachaki E. Evaluation of the Greek TranQol: a novel questionnaire for measuring quality of life in transfusion-dependent thalassemia patients. *Ann Hematol*. 2017;96(11):1937-44.

Maruish ME. User's manual for the SF-36v2 Health Survey (3rd ed.). Quality Metric Incorporated. 2011

McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32(1):40-66.

Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-92.

Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4(5):581-5.

Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-83.

17. APPENDICES

17.1. Handling of Dates

Dates will be stored as numeric variables in the SAS analysis files and reported in DDMMYYYY format (i.e., the Date9. datetime format in SAS). Dates in the clinical database are classified into the categories of procedure dates, log dates, milestone dates, outcome dates, and special dates.

- **Procedure Dates** are the dates on which given protocol-specified procedure are performed. They include the dates of laboratory testing, physical examinations, tumor scans, etc. They should be present whenever data for a protocol-specified procedure are present and should only be missing when a procedure are marked as NOT DONE in the database. Procedure dates will not be imputed.
- **Log Dates** are dates recorded in eCRF data logs. Specifically, they are the start and end dates for adverse events and concomitant medications/procedures. They should not be missing unless an event or medication is marked as *ongoing* in the database. Otherwise, incomplete log dates will be imputed according to the rules in Appendix 17.2 (e.g., for duration or cycle assignment, etc.). However, in listings, log dates will be shown as recorded without imputation.
- **Milestone Dates** are dates of protocol milestones such as randomization, study drug start date, study drug termination date, study closure date, etc. They should not be missing if the milestone occurs for a subject. They will not be imputed.
- **Special Dates** cannot be classified in any of the above categories and they include the date of birth. They may be subject to variable-specific censoring and imputation rules.

Dates recorded in comment fields will not be imputed or reported in any specific format.

17.1.1. Calculation Using Dates

Calculations using dates (e.g., subject's age or relative day after the first dose of study drug) will adhere to the following conventions:

- Study days after the start day of study drug will be calculated as the difference between the date of interest and the first date of dosing of study drug plus 1 day. The generalized calculation algorithm for relative day is the following:
 - If TARGET DATE \geq DSTART then STUDY DAY = (TARGET DATE – DSTART) + 1;
 - Else use STUDY DAY = TARGET DATE – DSTART.

Note that Study Day 1 is the first day of treatment of study drug. Negative study days are reflective of observations obtained during the baseline/screening period. Note: Partial dates for the first study drug are not imputed in general. All effort should be made to avoid incomplete study drug start dates.

- Age (expressed in years) is calculated as the number of months between birth date and informed consent date divided by 12 (if both dates are not missing), the integer part will be kept. If the month of birth date is the same as informed consent date and the day of birth date is greater than informed consent date, then the age calculated by above will minus 1. If any date is missing, AGE will be set to the age collected from CRF.
 - Partial birth date: impute missing day as 15th of the month; impute missing month as July; set missing age for missing year
- Intervals that are presented in weeks will be transformed from days to weeks by using (without truncation) the following conversion formula:

$$\text{WEEKS} = \text{DAYS} / 7$$

- Intervals that are presented in months will be transformed from days to months by using (without truncation) the following conversion formula:

$$\text{MONTHS} = \text{DAYS} / 30.4167$$

17.2. Date Imputation Guideline

17.2.1. Impute Missing Adverse Events/ Prior or Concomitant Medications, Procedures/Surgeries

Incomplete Start Date:

Missing day and month

- If the year is the **same** as the year of the first dosing date, then the day and month of the first doing date will be assigned to the missing fields.
- If the year is **prior to** the year of first dosing date, then December 31 will be assigned to the missing fields.
- If the year is **after** the year of first dosing, then January 1 will be assigned to the missing fields.

Missing day only

- If the **month** and year are the **same** as the year and month of first dosing date, then the first doing date will be assigned to the missing day.
- If either the year of the partial date is **before** the year of the first dosing date or the years of the partial date and the first dosing date are the same but the month of partial date is **before** the month of the first dosing date, then the last day of the month will be assigned to the missing day.
- If either the year of the partial date is **after** the year of the first dosing date or the years of the partial date and the first dose date are the same but the month of partial date is **after** the month of the first dosing date, then the first day of the month will be assigned to the missing day.

- If the stop date is not missing, and the imputed start date is after the stop date, the start date will be imputed by the stop date.

Missing day, month, and year

- No imputation is needed, the corresponding AE will be included as TEAE.

Incomplete Stop Date: If the imputed stop date is before the start date, then the imputed stop date will be equal to the start date.

Missing day and month

- If the year of the incomplete stop date is the **same** as the year of the last dosing date, then the day and month of the last dosing date will be assigned to the missing fields.
- If the year of the incomplete stop date is **prior to** the year of the last dosing date or prior to the year of the first dosing date, then December 31 will be assigned to the missing fields.
- If the year of the incomplete stop date is **prior to** the year of the last dosing date but is the same as the year of the first dosing date, then the first dosing date will be assigned to the missing date.
- If the year of the incomplete stop date is **after** the year of the last dosing date, then January 1 will be assigned to the missing fields.

Missing day only

- If the month and year of the incomplete stop date are the **same** as the month and year of the last dosing date, then the day of the last dosing date will be assigned to the missing day.
- If either the year of the partial date is **not equal to** the year of the last dosing date or the years of the partial date and the last dosing date are the same but the month of partial date is **not equal to** the month of the last dosing date, then the last day of the month will be assigned to the missing day.



Celgene Signing Page

This is a representation of an electronic record that was signed electronically in Livelink.
This page is the manifestation of the electronic signature(s) used in compliance with
the organizations electronic signature policies and procedures.

UserName: PPD
Title: PPD
Date: Wednesday, 20 June 2018, 12:32 PM Eastern Daylight Time
Meaning: Approved, no changes necessary.
=====

UserName: PPD
Title: PPD
Date: Wednesday, 20 June 2018, 12:51 PM Eastern Daylight Time
Meaning: Approved, no changes necessary.
=====

UserName: PPD
Title: PPD
Date: Wednesday, 20 June 2018, 01:12 PM Eastern Daylight Time
Meaning: Approved, no changes necessary.
=====

UserName: PPD
Title: PPD
Date: Wednesday, 20 June 2018, 01:50 PM Eastern Daylight Time
Meaning: Approved, no changes necessary.
=====