NCT02453334 1VIT14039

Can Iron Lessen Anemia Due to Cancer and Chemotherapy: A Study to Investigate the Efficacy and Safety of Injectafer® (IRON CLAD)

Statistical Analysis Plan Protocol No. 1VIT14037, Version 1.0 - 22 November 2016

16.1.9 Documentation of Statistical Methods

Table of Contents

16.1.9 Documentation of Statistical Methods 1
Table of Contents 2
Statistical Analysis Plan Protocol No. 1VIT14037, Version 1.0 - 22 November 2016 3

Statistical Analysis Plan

Protocol No.: 1VIT14037

A Double-Blind, Multi-Center, Randomized, Placebo-Controlled Study to Investigate the Efficacy and Safety of Injectafer® (Ferric Carboxymaltose) in the Treatment of Restless Legs Syndrome (RLS)

Sponsor:	Luitpold Pharmaceuticals, Inc.
Version:	Version 1.0
Issue Date:	22 November, 2016

STATISTICAL ANALYSIS PLAN APPROVAL

Author:

Lina Yin, PhD
Biostatistician, Biostatistics, FMD K&L Inc.

Date: 12, 15, 2016

Reviewed and approved by:

Kevin Chen, PhD
Sr. Director, Biostatistics, FMD K&L Inc.

Date: 12, 15, 2016

Sponsor Review:

Linda M. Mundy, MD, PhD VP, CMO
Sr. Medical Director, Clinical Development
Luitpold Pharmaceuticals, Inc.

Date: 15 Dec 2016

David Morris, PhD
Sr. Director, Statistics
WebbWrites, LLC

Date: 16 Dec 2016

TABLE OF CONTENTS

ABBREVIATIONS 5

1. INTRODUCTION 7

1.1 Study Objectives 7
1.1.1 Primary Objective(s) 7
1.2 Study Endpoints 7
1.2.1 Primary Efficacy Endpoint(s) 7
1.2.2 Secondary Efficacy Endpoint(s) 7
1.2.3 Other Efficacy Endpoint(s) 7
1.2.4 Safety Endpoints 8
1.3 Summary of the Study Design 8
1.3.1 General Study Design and Plan 8
1.3.2 Randomization and Blinding 10
1.3.3 Sample Size and Statistical Power Considerations 10

2. STATISTICAL METHODS 11

2.1 General Considerations 11
2.1.1 Reporting Precision 11
2.2 Definitions of Analysis Populations (Analysis Sets) 12
2.2.1 Safety Population 12
2.2.1 Full Analysis Set Population 12
2.3 Time Windows for Analysis 12
2.4 Pooling of Centers 14
2.5 Handling of missing data 14
2.6 Analysis software 15

3. STUDY SUBJECTS 15

3.1 Disposition of Subjects 15
3.2 Protocol Deviations 15

4. DEMOGRAPHICS AND BASELINE CHARACTERISTICS 15

4.1 Demographics and Baseline Characteristics 15
4.2 Medical History 16

5. STUDY DRUG AND EXPOSURE 16

5.1 Treatment Compliance and Extent of Exposure 16
5.2 Prior and Concomitant Therapy 16

6. EFFICACY ANALYSES 17

6.1 Primary Efficacy Analysis 17
6.1.1 IRLS total score change from baseline to Day 42 17

6.1.2 Proportion of patients rated as much or very much improved with the Clinical Global Impression (CGI) performed by Investigator (CGI-I) on Day 42. 17
6.1.3 Sensitivity Analyses 17
6.1.4 Examination of Subgroups 19
6.2 SECONDARY EFFICACY ANALYSES 19
6.2.1 CGI-S on Day 42 19
6.2.2 RLS-QLI, MOS Sleep Scale, and Fatigue Linear Analog Scale Change From Baseline to Day 42 19
6.2.3 Time off pre-enrollment prescribed RLS medications 19
6.3 OTHER EFFICACY ENDPOINT(S) 20
6.3.1 Proportion of responders based on CGI-I / CGI-S at each time point 20
6.3.2 IRLS, RLS-QLI, MOS Sleep Scale, and Fatigue Linear Analog Scale Change from baseline at each time point 20
6.3.3 CGI-I and CGI-S scores at each time point 20
6.3.4 Augmentation assessment by Investigator change between baseline, Day 42, Day 168 and end of study (Day 365) 20
6.3.5 Proportion of subjects required intervention for RLS 21
6.3.6 Time from Day 5 to the next dose of study drug 21
7. SAFETY ANALYSIS 21
7.1 ADVERSE EVENTS 21
7.1.1 Deaths, Serious and Other Significant Adverse Events 22
7.2 CLINICAL LABORATORY PARAMETERS 23
7.3 VITAL SIGNS, PHYSICAL EXAMINATION FINDINGS, AND C-SSRS 23
7.3.1 Vital Signs 23
7.3.2 Physical Examination 24
7.3.3 Columbia-Suicide Severity Rating Scale (C-SSRS) 24
7.4 OTHER ANALYSES 24
8. INTERIM ANALYSES AND DATA AND SAFETY MONITORING BOARD (DSMB) 25
8.1 INTERIM ANALYSES 25
8.2 DATA AND SAFETY MONITORING BOARD (DSMB) 25
9. SUMMARY OF MAJOR CHANGES IN THE PLANNED ANALYSES 26
REFERENCES 27

ABBREVIATIONS

Abbreviation	Term
ALT	Alanine Aminotransferase
ANCOVA	Analysis Of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BUN	Blood Urea Nitrogen
CGI-I	Clinical Global Impression-Global Improvement
CGI-S	Clinical Global Impression- Subject
CI	Confidence Intervals
C-SSRS	Columbia-Suicide Severity Rating Scale
DSMB	Data And Safety Monitoring Board
eCRF	Electronic Case Report Form
FAS	Full Analysis Set
GGT	Gamma-Glutamyl Transferase
Hct	Hematocrit
Hgb	Hemoglobin
IRLS	International Restless Legs Syndrome
ITT	Intent-To-Treat
IV	Intravenous
LDH	Lactate Dehydrogenase
LOCF	Last Observation Carried Forward
MAR	Missing At Random
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measures
MOS	Medical Outcomes Study
NSS	Normal Sterile Saline
OC	Observed Cases
PCS	Potentially Clinically Significant
PT	Preferred Term
RBC	Red Blood Cells
RDW	Red Cell Distribution Width
RLS	Restless Leg Syndrome
RLS-QLI	Restless Legs Syndrome – Quality Of Life

Abbreviation	Term
ROS	**Rest of World**
SAEs	Serious AEs
SAP	Statistical Analysis Plan
SD	Standard Deviation
SI	Système International
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TIBC	Total Iron Binding Capacity
TSAT	Percentage Serum Transferrin Saturation
US	United States
WBC	White Blood Cells
WHO	World Health Organization

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for Protocol 1VIT14037 (Amendment IV, Date: 02 June 2016).

1.1 STUDY OBJECTIVES

1.1.1 Primary Objective(s)

The primary objective of the study was to evaluate the efficacy and safety of an IV Injectafer® in subjects with Restless Leg Syndrome (RLS).

1.2 STUDY ENDPOINTS

1.2.1 Primary Efficacy Endpoint(s)

The co-primary efficacy variables include:

- IRLS total score change from baseline to Day 42.
- Proportion of subjects rated as much or very much improved with the Clinical Global Impression (CGI) performed by Investigator (CGI-I) on Day 42.

1.2.2 Secondary Efficacy Endpoint(s)

The major secondary endpoints will be tested in a hierarchical order, that is, statistical significance for major secondary efficacy endpoint will be declared only if the primary efficacy test (IRLS and CGI-I on Day 42) is statistically significant. Because a hierarchical testing scheme is used, no adjustments of the alpha level will be needed at each stage of testing. All tests will be done at an alpha ≤ 0.05.

The major secondary efficacy endpoints in ranked order of testing include:

1. Clinical Global Impression (CGI) performed by Subject (CGI-S) on Day 42.
2. Restless Legs Syndrome Quality of Life (RLS-QLI) change from baseline to Day 42.
3. Medical Outcome Study (MOS) Sleep Scale change from baseline to Day 42.
4. Fatigue Linear Analog Scale change from baseline to Day 42.
5. Time off pre-enrollment prescribed RLS medications.

1.2.3 Other Efficacy Endpoint(s)

The Other efficacy endpoints include:

1. Proportion of responders (any improvement) based on CGI-I at each time point.
2. IRLS total score change from baseline at each time point.
3. Proportion of responders based on CGI-S at each time point.
4. CGI-I and CGI-S scores at each time point and by baseline RLS treatment group.

5. RLS-QLI scores change from baseline at each time point.
6. MOS Sleep scale total score change from baseline at each time point.
7. Fatigue Linear Analog Scale total score change from baseline at each time point.
8. Augmentation assessment by Investigator change between baseline, Day 42, Day 168 and end of study (Day 365).
9. Proportion of subjects required intervention for RLS.
10. Time from Day 5 to the next dose of study drug.

1.2.4 Safety Endpoints

Safety endpoints include:

1. Incidence of treatment emergent adverse events and incidence of serious adverse events
2. Change in clinical laboratory tests
3. Change in vital signs
4. Change in C-SSRS

1.3 SUMMARY OF THE STUDY DESIGN

1.3.1 General Study Design and Plan

This is a Phase III, double blinded, multi-center, randomized, placebo-controlled study evaluating the efficacy and safety of an IV Injectafer® in subjects with Restless Leg Syndrome (RLS).

Injectafer® is a stable type I polynuclear iron (III)-hydroxide carbohydrate complex developed as an IV iron replacement therapy for the treatment of iron deficiency anemia. RLS has been reported to be associated with the reduction of iron in central nervous system [1]. In a previous phase II trial Injectafer® was found to be safe and well tolerated in the treatment of RLS [2]. In this study, 200 eligible subjects will be randomized in a 1:1 ratio to receive either a 750 mg undiluted blinded dose of IV Injectafer® at 100mg/minute or a blinded IV Placebo (15 ml of Normal Sterile Saline [NSS]) at 2ml/minute during clinic site visits. The randomization will be stratified by Augmentation at baseline (no augmentation, uncertain augmentation, definitive augmentation). The study duration is 12 months, including a 5 day treatment phase (Dosing on Day 0 and Day 5), and a 12 months efficacy and safety follow-up phase. Subjects will visit the clinic on Days 14, 42, 168, and 365. In between the clinic visits subjects will be contacted remotely by phone on Days 28, 84, 126, 210, 252, 294, and 336. Subjects will be considered to have reached the end of the study for efficacy when intervention is required for their RLS symptoms; those subjects will continue the follow-up phase for safety. Subjects who have not had an intervention may receive additional blinded study drug treatment after Day 42 at the discretion of the Investigator. The additional dosing will be the same blinded study medication as previously randomized to receive during the treatment phase, and the dosing days should mirror the original treatment/follow-up period (dosing 5 days apart with follow-up visits, to include laboratory assessments only, on 14 and 42 Days post the first dose). No additional iron maybe administered between Day 320 and the Day 365 visit.

The Schedule of Events is presented by Table 1.

Table 1 Schedule of Event

	Scr.	Treatment					Efficacy and Safety Follow-up ** *Contact completed remotely via phone*							
Study Day	-7	**0**	**5**	**14**	*28***	**42**	*84***	*126***	**168**	*210***	*252***	*294***	*336***	**365**
Informed Consent	x													
Eligibility	x	x												
Medical History[9]	x													
D/C oral iron products	x													
Site contact IRT	x[1]	x[2]												x
Physical exam	x					x			x					x
Vital Signs	x[3]	x[4]	x[4]			x[3]			x[3]					x[3]
Height and Weight without shoes	x													
Augmentation Subject Questionnaire	x[13]													
Augmentation Investigator Assessment	x[14]					x			x					x
RLS Diary[16]	x	x	x	x	x	x								
Phone contact treatment phase					x									
IRLS Scale	x	x[5]	x		x	x	*x*	*x*	x	*x*	*x*	*x*	*x*	x
CGI-I and CGI-S			x[LS]		*x*[5]	x[LS]	*x*[5]	*x*[5]	x[LS]	*x*[5]	*x*[5]	*x*[5]	*x*[5]	x[LS]
RLS QLI	x	x[5]				x			x					x
MOS Sleep Scale	x	x[5]	x			x			x					x
Fatigue Linear Analog Scale	x	x[5]	x			x			x					x
C-SSRS[15]	x					x			x					x
Hematology	x		x	x		x			x					x
Iron Indices	x		x	x		x			x					x
Chemistry	x		x	x		x			x					x[8]
Pregnancy test[10]	x													
Transferrin receptor	x[6]					x[6]			x[6]					
Study Drug Dosing		x	x				**See section 6.3.5 for additional study drug dosing requirements**							
RLS Treatment Stability / Tapering review / assessment	x	x	x[11]	x	x	x	x	x	x	x	x	x	x	x
Concomitant Meds	x	x	x		x	x	*x*	*x*	x	*x*	*x*	*x*	*x*	x
Adverse Events[7]		x	x	x	x	x	*x*	*x*	x	*x*	*x*	*x*	*x*	x

1 Contact IRT for screening number assignment during screening, and on Day 365 to complete subject from the study.
2 Contact IRT for randomization assignment, unblinded site personnel
3 Vital Signs includes sitting BP and heart rate.
4 On study drug dosing days sitting vital signs including blood pressure and heart rate should be collected immediately pre dosing, immediately and 30 minutes post dosing. Body temperature will also be collected pre dosing on Days 0 and 5.
5 IRLS, RLS QLI, MOS Sleep Scale, and Fatigue Linear Analog Scale may be completed on Day 0 prior to randomization.
6 Fasting, if indicated.
7 Adverse event assessments will start at the time of the first dose of study drug. All events noted prior to the 1st dose of study drug should be considered history and captured on the medical history page of the eCRF.
8 If the phosphorous is below the LLN at the time of early termination or end of the study (Day 365) the subject should return (as directed by the Investigator) for repeat phosphorous until the value is back WNL's or the subject's baseline.

[9] To include past and present treatment for RLS.
[10] Female subjects only.
[11] After dosing on Day 5 the site will review the tapering requires with the subject.
[12] Assess compliance of tapering. Once the taper period is complete assess stability (off meds, decrease dose, baseline dose).
[13] Augmentation questionnaire should be completed at screening prior to the Investigator's Augmentation Assessment.
[14] Augmentation assessment can be done either at screening or prior to dosing on Day 0 but after the subject has completed the augmentation questionnaire.
[15] The Columbia-Suicide Severity Rating Scale will be used throughout the study to assess the subjects suicide ideation or behavior. This scale is an assessment tool, it is ultimately the Investigators responsibility to evaluate each subject's risk and treat as appropriate.
[16] Subject Diary reviewed for compliance at Day 0,5,14 and 28. The completed diary will be returned by the subject to the site on Day 42.
[I,S] Investigator and Subject Assessment.
[S] Subject only Assessment.

1.3.2 Randomization and Blinding

Eligible subjects will be randomly assigned in a 1:1 ratio to IV Injectafer® or IV Placebo via an IRT system. Randomization will be stratified by baseline RLS medication-related augmentation (no augmentation, uncertain augmentation, definitive augmentation). The randomization schedule was generated by FMD K&L Inc. The schedule was approved by Sponsor before being loaded into the IRT system. Other than designated randomization personnel, all subjects, investigators, and study conduct personnel are blinded to study drug assignment. The blinded investigators are also blinded to the IRLS score and subject CGI score for each subject after randomization.

1.3.3 Sample Size and Statistical Power Considerations

Sample size estimates were based on the mITT population for Study 1VIT05009. The last observation carried forward (LOCF) and observed cases (OC) results on Day 28 were used for the co-primary efficacy endpoints (CGI) and first ranked secondary endpoint (IRLS). It is unclear how treatment differences on Day 42 results might differ from Day 28, apart from a higher dropout rate, but it is reasonable to assume that treatment differences will decrease. Therefore, power calculations were based on smaller treatment differences and larger standard deviations.

Co-primary endpoint (IRLS): The smallest treatment difference for change from baseline to Day 28 was 4.4 with a standard deviation of 8.5. A sample size of 100/group provides 95% power under this scenario.

Co-primary efficacy endpoint (CGI-I): The smallest treatment difference on Day 28 was 45% versus 15% of subjects in the FCM and placebo groups, respectively, who reported much or very much improvement on the patient CGI-S. A sample size of 100/group provides >99% power. A sample size of 100/group has at least 90% power when the treatment difference is as small as 45% versus 22%.

2. STATISTICAL METHODS

2.1 GENERAL CONSIDERATIONS

In general, continuous variables will be summarized by number of subjects, mean, standard deviation (SD), median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of subjects in each category. If applicable, hypothesis testing will be carried out at the two-sided α=0.05 level unless otherwise specified; 2-sided 95% confidence intervals (CIs) will be presented, where specified.

In general, summary tables will present data by treatment group. Source data for the summary tables and statistical analyses will be presented as subject data listings, which include data collected on the electronic case report forms (eCRFs as well as any derived variables for all enrolled subjects.

Baseline will be defined as the last non-missing value obtained prior to the first dose of study drug.

2.1.1 Reporting Precision

Summary statistics will be presented to the following degree of precision, unless otherwise specified:

Statistics	**Degree of Precision**
Mean, Geometric mean, Median, Quartiles, Confidence limit boundaries	One decimal place more than the raw data.
Standard deviation, Standard error	Two decimal places more than the raw data.
Minimum, Maximum	The same as the raw data.
p-value	Rounded to 4 decimal places and therefore presented as 0.xxxx; p-values smaller than 0.0001 as '<0.0001'; p-values greater than 0.9999 as '>0.9999'.
Percentage	One decimal place. A percentage of 100% will be reported as 100%. Percentages of zero will be reported as 0.

For weight and height, one decimal place will be used for summary statistics. Fractional numeric values will be presented with a zero to the left of the decimal point (for example, 0.12 – 0.30).

2.2 DEFINITIONS OF ANALYSIS POPULATIONS (ANALYSIS SETS)

2.2.1 Safety Population

All subjects who received at least one dose of randomized treatment.

2.2.1 Full Analysis Set Population

All subjects who received at least one dose of randomized treatment and had at least one post-randomization measurement of the IRLS and Clinical Global Impression performed by Investigator (CGI-I).

2.3 TIME WINDOWS FOR ANALYSIS

For the following analyses no windows will be applied, including safety shift, efficacy shift, and safety potential clinical significant (PCS) summaries.

For safety and efficacy by visit descriptive statistical summaries (mean, standard deviation, median, minimum and maximum), all CRF collected observations (including scheduled and unscheduled visits) will be summarized with adjusted analysis-defined visit windows based on the scheduled days of protocol. Variables will be windowed to the closest scheduled visit for that variable. Visit windows have been constructed to ensure every post-dose observation collected by CRF can be allocated to a particular visit. No windows will be applied for the screening visit. The adjusted analysis-defined windows are summarized in Table 2.

Table 2 Visit windows for assessments

Assessment	Defined Windows Visit	Scheduled Study Day	Maximum Windows
Vital Signs*			
	Day 0	1	Study Day=1
	Day 5	6	2 ≤Study Days≤ 24
	Day 42	43	25 ≤Study Days≤ 105
	Day 168	169	106 ≤Study Days≤ 267
	Day 365	366	268 ≤Study Days

Assessment	Defined Windows Visit	Scheduled Study Day	Maximum Windows
C-SSRS, Physical Exam, Augmentation Investigator Assessment, RLS QLI			
	Day 42	43	1 ≤Study Days≤ 105
	Day 168	169	106 ≤Study Days≤ 267
	Day 365	366	268 ≤Study Days
CGI-I, MOS Sleep Scale, Fatigue Linear Analog Scale,			
	Day 5	6	1 ≤Study Days≤ 24
	Day 42	43	25 ≤Study Days≤ 105
	Day 168	169	106 ≤Study Days≤ 267
	Day 365	366	268 ≤Study Days
IRLS Scale, CGI-S			
	Day 5	6	1 ≤Study Days≤ 17
	Day 28	29	18 ≤Study Days≤ 35
	Day 42	43	36 ≤Study Days≤ 63
	Day 84	85	64 ≤Study Days≤ 105
	Day 126	127	106 ≤Study Days≤ 147
	Day 168	169	148 ≤Study Days≤ 189
	Day 210	211	190 ≤Study Days≤ 231
	Day 252	253	232 ≤Study Days≤ 273
	Day 294	295	274 ≤Study Days≤ 315
	Day 336	337	316 ≤Study Days≤ 351
	Day 365	366	352 ≤Study Days

* Vital signs will be summarized by time-point of each visit, i.e. immediately post dose and 30 minutes post dose.

For assignment of post-dose data to adjusted analysis-defined visit windows, study day will be defined as follows:

Date of assessment – date of first treatment + 1

By this definition, the first dose day will be study day 1 for example.

If multiple readings are recorded within a single adjusted visit window, please refer to the rules below.

- If there are 2 or more valid, non-missing observations within the same visit window, then the non-missing one closest to the scheduled visit will be used in the analysis.

- If 2 valid observations are equidistant from the scheduled visit, then the scheduled visit (if one of the two observations is) will be used in the analysis, or else the non-missing observation with the earlier collection date will be used.

- If 2 valid observations are collected on the same day, then the scheduled visit (if one of the two observations is) will be used in the analysis, or else the non-missing observation with the earlier collection time will be used.

If a visit window does not contain any observations, then the data will remain missing, excepting for the efficacy inferential analyses where LOCF will be used for missing data imputation.

Any repeated or additional assessments performed will be included in the individual patient data listings.

2.4 POOLING OF CENTERS

This study was a multi-center study. For statistical inferences (models), sites will be pooled into 2 geographic regions (United States [US] and Europe). For summaries using descriptive statistics, data of all sites will be combined. Subgroup analyses by pooled regions (US and Europe) will be performed for the co-primary efficacy endpoints using descriptive statistics.

2.5 HANDLING OF MISSING DATA

Missing data will not be imputed for descriptive statistical summaries in safety or efficacy analyses.

For statistical inferences (models) analyses, missing data will be imputed using last observation carried forward (LOCF) method for CGI and other efficacy endpoints. e.g. for primary endpoint, IRLS total score, when Day 42 data is missing (subject has discontinued before Day 42, or measurement not taken at Day 42 though subject was not discontinued), Day 28 result will be carried forward. If the post-baseline value of the first scheduled visit is missing, the worst value obtained from the same time point of all subjects in FAS population will be used for both treatment groups. The worst value is defined as “Non-responder” for CGI responder rate analysis, the highest score for IRLS, RLS-QLI, MOS sleeping scale, Fatigue Linear Analog Scale and CGI, and “Definitive Augmentation” for augmentation scale.

Sensitivity analyses will assess the impact of missing values on inferences based on primary efficacy endpoints (Section 6.1.1).

2.6 ANALYSIS SOFTWARE

All summaries and statistical analyses will be generated using SAS® version 9.2 or later.

3. STUDY SUBJECTS

3.1 DISPOSITION OF SUBJECTS

Disposition will be summarized by treatment group and overall for all enrolled subjects.

The disposition will include the following:

- Subjects who are randomized
- Subjects who are in the Safety Population
- Subjects who are in the FAS
- Subjects who complete the study
- Subjects who complete the Day 42 visit
- Subjects who discontinue the study

The number and percent of subjects will be summarized for each reason for premature discontinuation. For subjects who are prematurely discontinued from the study, the reasons for discontinuation as recorded on the disposition page in the CRF will also be presented.

A listing of dispositions will be provided for all subjects.

3.2 PROTOCOL DEVIATIONS

The clinical team will identify deviations and the deviations will be identified and classified into types in the database. The number and percent of subjects will be summarized for each type of deviation. A subject data listing of clinically important protocol deviations will be listed.

4. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

4.1 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographics and baseline characteristics (e.g., gender, age, race, ethnicity, body height, weight, and the status of iron intolerance) will be summarized with descriptive statistics by treatment group and overall for the Safety Population and FAS Population.

A subject data listing of demographics and baseline characteristics will be provided.

4.2 MEDICAL HISTORY

Medical history will be coded with the Medical Dictionary for Regulatory Activities (MedDRA) Terminology. The number and percent of subjects with clinically significant medical history at screening will be summarized by system organ class (SOC) and preferred term (PT) by treatment group and overall for Safety Population.

5. STUDY DRUG AND EXPOSURE

5.1 TREATMENT COMPLIANCE AND EXTENT OF EXPOSURE

Treatment compliance will not be summarized, as each subject will receive a single dose of study drug on Day 0 and Day 5 in an inpatient setting.

The total number of study drug injections received will be summarized via descriptive statistics by treatment group. The number (percentage) of subjects who received IV injections of study drug will be summarized by the number of doses received by treatment group in safety population.

A subject data listing of treatment exposure will be provided by treatment.

5.2 PRIOR AND CONCOMITANT THERAPY

Prior medications are defined as medications that started prior to the first dose of study drug. Concomitant medications are defined as medications (other than the study drug) taken on or after the first does of the study drug during the study. Medications started before the first dose of study drug and continuing at the time of the first dose of study drug are considered both prior medication and concomitant medication.

All investigator terms for medications recorded on the CRF will be coded using the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) drug dictionary (Version WHO Drug Dictionary Enhanced, WHO DDE March 2014). The number (percentage) of subjects who took prior and concomitant medications will be summarized on the Safety Population by treatment group, ATC Classification, and WHO Drug preferred term.

A subject data listing of priori and concomitant therapy will be provided by treatment.

6. EFFICACY ANALYSES

All efficacy analyses will be based on the FAS Population. Missing data will be imputed per Section 2.5 for statistical inference analyses (no imputation for descriptive statistical summaries). Subject who has an intervention will no longer be evaluated for efficacy starting at the time of the intervention (these subjects will remain eligible for the safety evaluation).

An intervention is defined as either

(1) An increase in dosage from the RLS medication at study entry
(2) The initiation of a new RLS medication
(3) Resumption of the previous medication prescribed for RLS (the previous prescribed RLS medication should remain unchanged until the tapering process begins post Day 5 treatment with study drug).

6.1 PRIMARY EFFICACY ANALYSIS

6.1.1 IRLS total score change from baseline to Day 42

The actual values of IRLS total score on baseline and Day 42, and the change from baseline to Day 42 will be summarized via descriptive statistics by treatment group. Treatment group difference for change in IRLS total score will be assessed with the analysis of covariance (ANCOVA), with region (United States [US], Europe), and baseline RLS medication-related augmentation and treatment as fixed factors and baseline IRLS total score as a covariate. Missing value will be imputed using LOCF as specified in Section 2.5. P-values to test treatment effect will be determined. Point estimates, associated 95% confidence interval and p-values will be reported.

6.1.2 Proportion of patients rated as much or very much improved with the Clinical Global Impression (CGI) performed by Investigator (CGI-I) on Day 42.

Responder is defined as subjects rated as much or very much improved with the CGI-I on Day 42. The number (percentage) of CGI-I responder will be summarized by treatment group. Treatment differences for proportions will be assessed with the Logistic regression with treatment, region (US, Europe), and baseline RLS medication related augmentation as fixed factors, and baseline IRLS total score as a covariate. Missing value will be imputed using LOCF as specified in Section 2.5 . Odds ratio for treatment comparison and associated 95% confidence interval and p-values will be presented.

A subject data listing will be provided by treatment for each endpoint.

6.1.3 Sensitivity Analyses

To assess the robustness of the primary efficacy analysis, sensitivity analyses will be performed using the following approaches:

MMRM Analysis of Change in IRLS Total Score:

A mixed model repeated measures (MMRM) model will be conducted to provide an analysis under the assumption that data are missing at random (MAR). The MMRM model will include changes in IRLS total score from baseline to Day 5, 28 and 42 as the dependent variables; and baseline RLS medication-related augmentation, treatment group, visit, treatment group by visit interaction, and regions (US, Europe) as fixed effects; and subjects within treatment group as random effects; and baseline IRLS total score as covariate. An unstructured covariance matrix will be used to allow for unequal variances between visits. If the model does not converge with unstructured variance – covariance matrix, then the Toeplitz, first-order autoregressive, compound symmetric, variance components structures will be tried and the covariance structure will be decided based on model convergence status and the AKaike information criterion. The Kenward-Roger approximation will be used to calculate the denominator of degrees of freedom for the fixed effect. If the interaction term is not significant, it will be dropped from the model. No imputation of missing data other than that inherent in the MMRM model will be performed. P-values to test treatment effect on Day 42 will be determined. Point estimates, associated 95% confidence interval and p-values will be reported.

Logistic Regression Based on Observed Cases for Proportion of Responders based on CGI-I:

A Logistic regression model will be used to assess the treatment group difference in proportion of subjects rated as much or very much improved with the CGI-I on Day 42. The model will have treatment, region (US, Europe), and baseline RLS medication related augmentation as fixed factors, and baseline IRLS total score as a covariate. If the Day 42 data are missing, no imputation will be performed. Subjects with missing data will be excluded from the statistical testing. Odds ratio for treatment comparison and associated 95% confidence interval and p-values will be presented.

Logistic Regression with Missing Data Imputed as Non-Responder for Proportion of Responders based on CGI-I:

The same Logistic regression model as the primary analysis will be used to assess the treatment group difference in proportion of subjects rated as much or very much improved with the CGI-I on Day 42. Missing data of Day 42 will be imputed as "Non-Responder". Odds ratio for treatment comparison and associated 95% confidence interval and p-values will be presented.

CMH with Missing Data Imputed as Non-Responder for Proportion of Responders based on CGI-I:

Cochran-Mantel-Haenszel test will be used to assess the treatment group difference in proportion of responder (i.e. subjects rated as much or very much improved) with the CGI-I on Day 42, using baseline RLS medication-related augmentation and region (US, Europe) as the stratification factors. Subjects with missing data will be imputed as non-responder. Point

estimate of the treatment difference and associated 95% confidence interval and p-values will be presented.

6.1.4 Examination of Subgroups

The co-primary endpoints will be summarized descriptively by region, baseline augmentation, sex, race, and age (<65 years, ≥65 years).

6.2 SECONDARY EFFICACY ANALYSES

6.2.1 CGI-S on Day 42

The number (percentage) of subjects of each CGI-S category on Day 42 will be summarized by treatment group. Treatment group differences for CGI-S scores will be assessed with the Cochran-Mantel-Haenszel test, with LOCF (Section 2.5), using baseline RLS medication-related augmentation and region (US, Europe) as the stratification factors at each time point. P-values to test treatment effect will be determined.

6.2.2 RLS-QLI, MOS Sleep Scale, and Fatigue Linear Analog Scale Change From Baseline to Day 42

The actual values on baseline and Day 42, and the change from baseline to Day 42 will be summarized via descriptive statistics by treatment group for RLS-QLI, MOS sleep scale and fatigue linear analog scale, separately. Treatment group difference for change from baseline will be assessed with the analysis of covariance (ANCOVA) using LOCF (Section 2.5), with treatment, and region (US, Europe) and baseline RLS medication-related augmentation as fixed factors, and baseline score as a covariate, for each endpoint. P-values to test treatment effect will be determined. Point estimates, associated 95% confidence interval and p-values will be reported.

A subject data listing will be provided by treatment for each endpoint.

6.2.3 Time off pre-enrollment prescribed RLS medications

Subjects will start tapering of pre-enrollment prescribed RLS medications after Day 5 study drug administration until an intervention will be given. An intervention is defined as Section 6. Therefore, time off pre-enrollment prescribed RLS medications will be estimated with time from Day 5 to the next RLS intervention. The endpoint will be analyzed with log-rank test stratified by baseline RLS medication augmentation and region (US, Europe). P-value of the treatment effect will be reported. Time from Day 5 to the next RLS intervention will be calculated as the date of the next RLS intervention minus date of Day 5 plus 1. Subjects who discontinue or complete the study before an intervention will be censored at last study visit. If the subject does not return to the clinic after Day 5, the subject will be censored at Day 5. Kaplan-Meier curves will be presented for each of the treatment groups.

6.3 OTHER EFFICACY ENDPOINT(S)

6.3.1 Proportion of responders based on CGI-I / CGI-S at each time point

A responder for CGI-I / CGI-S is defined as any improvement seen at each time point. The number (percentage) of responders will be summarized for each time point (clinical visits and telephone follow-ups) by treatment group, for CGI-I and CGI-S, separately. Treatment differences for proportions will be assessed using the Cochran-Mantel-Haenszel (CMH) test, with baseline RLS medication-related augmentation and region (US, Europe) as the stratification factors. Missing data at each time point will be imputed using LOCF as specified in Section 2.5. Point estimate of the treatment difference and associated 95% confidence interval and p-values will be presented.

6.3.2 IRLS, RLS-QLI, MOS Sleep Scale, and Fatigue Linear Analog Scale Change from baseline at each time point

The actual values and the changes from baseline will be summarized for each time point (clinical visits and telephone follow-ups) via descriptive statistics by treatment group, for IRLS, RLS-QLI, MOS sleep scale, and Fatigue Linear Analog Scale, separately. Treatment group difference for change from baseline will be assessed with the analysis of covariance (ANCOVA) using LOCF (Section 2.5), with treatment, and region (US, Europe) and baseline RLS medication-related augmentation as fixed factors, and baseline score as a covariate, for each endpoint. P-values to test treatment effect will be determined. Point estimates, associated 95% confidence interval and p-values will be reported.

6.3.3 CGI-I and CGI-S scores at each time point

The number (percentage) of subjects of each CGI category will be summarized for each time point (clinical visits and telephone follow-ups) by treatment group, for CGI-S and CGI-I, separately. Treatment group differences for CGI scores will be assessed at each time point via Cochran-Mantel-Haenszel with LOCF (Section 2.5), using baseline RLS medication-related augmentation and region (US, Europe) as the stratification factors. P-value of treatment group difference will be determined and reported.

6.3.4 Augmentation assessment by Investigator change between baseline, Day 42, Day 168 and end of study (Day 365)

The number (percentage) of subjects of each category (no augmentation, uncertain augmentation, and definitive augmentation) will be summarized for baseline, Days 42, 168 and 365 by treatment group.

The change from baseline will be summarized by a shift table for each post baseline visit above. Treatment group differences will be assessed at each post baseline time point with the Cochran-Mantel-Haenszel test (row mean score), with LOCF (Section 2.5), using baseline RLS

medication-related augmentation and region (US, Europe) as the stratification factors. P-value of treatment group difference will be determined and reported.

6.3.5 Proportion of subjects required intervention for RLS

The number (percentage) of subjects required intervention after the first dose of study drug will be summarized by treatment group. Treatment differences for proportions will be assessed using continuity-corrected chi-square test. p-values of treatment effect will be determined.

6.3.6 Time from Day 5 to the next dose of study drug

Time from Day 5 to the next dose of study drug will be analyzed with log-rank test stratified by region (US, Europe) and baseline RLS medication augmentation. P-value of the treatment effect will be reported. Time from Day 5 to the next dose of study drug will be calculated as the date of the next dose of study drug minus Day 5 date plus 1. Subjects who discontinue or complete the study before receiving a second dose of study drug will be censored at last study visit. If the subject does not return to the clinic after Day 5, the subject should be censored at Day 5. The Kaplan-Meier curve will be presented for each of the treatment groups.

A subject data listing will be provided by treatment for each efficacy endpoint.

7. SAFETY ANALYSIS

All safety analyses will be performed on the Safety Population. No formal statistical testing will be performed for safety. Safety assessments include:

- Incidence of treatment emergent adverse events and incidence of serious adverse events
- Clinical laboratory tests
- Vital signs
- Columbia-Suicide Severity Rating Scale (C-SSRS)

7.1 ADVERSE EVENTS

Adverse events (AEs) will be coded by System Organ Class (SOC) and Preferred Term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA®, version 18.0/AECODE). The verbatim term will be included in the AE listings.

A treatment-emergent adverse event (TEAE) is defined as an AE that that occur or worsen on or after the first dose of study drug. Only TEAEs will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

The incidence of TEAEs will be summarized as the number (percentage) of subjects with TEAEs within SOC and PT by treatment group. Subjects who report the same PT on multiple occasions will be counted once for the PT: under the highest severity (severe > moderate > mild) when summarized by severity and under the closest relationship (probably related > possibly related > unlikely related > none) to study drug/RLS treatment tapering when summarized by relationship. If a subject reports multiple PT for a SOC, the subject will be counted only once for that SOC. Treatment related AEs are defined as those events recorded on the CRF as 'Probably Related' or 'Possibly Related', others will not be related AEs.

TEAEs will be summarized as below.

- An overview table, including number of subjects with
    - TEAEs
    - serious AEs (SAEs)
    - study drug related TEAEs
    - RLS treatment tapering related TEAEs
    - TEAEs by severity
    - TEAEs leading to study discontinuation
    - TEAEs leading to death
- TEAE by SOC and PT
- TEAE by SOC, PT, and Severity
- Study drug related TEAEs by SOC, PT
- RLS treatment tapering related TEAEs by SOC and PT
- SAEs by SOC and PT
- TEAEs leading to study discontinuation by SOC and PT

All AE tables will be sorted by SOC and PT in decreasing frequency of the number and percentage of subjects in the treatment group.

7.1.1 Deaths, Serious and Other Significant Adverse Events

The listings of serious AEs, AE leading to study discontinuation, and subjects who died during the study will be listed.

7.2 CLINICAL LABORATORY PARAMETERS

Laboratory assessments include hematology, clinical chemistry, and iron indices/phosphorus:

- Hematology: Hgb, Hct, RBC, WBC, MCV, MCH, MCHC, RDW, platelets, differential count.
- Clinical chemistry: sodium, potassium, chloride, BUN, creatinine, albumin, alkaline phosphatase, total bilirubin, gamma-GGT, AST, ALT, LDH, calcium, glucose, bicarbonate.
- Iron indices and Phosphorus: serum iron, serum ferritin, and total iron binding capacity (TIBC), percentage serum transferrin saturation (TSAT), transferrin receptor, Phosphorus.

All laboratory parameters will be presented in conventional units. Quantitative results (including actual value, and change from baseline to each smallest value after baseline, largest value after baseline, and end of study) will be summarized using descriptive statistics by treatment, for each laboratory test group above. All laboratory data will be included in the listings.

Laboratory test results will be assigned an LNH classification according to whether the value was below (L), within (N), or above (H) the laboratory parameter's reference range. Shifts from baseline to the smallest value after baseline, largest value after baseline, and end of study in low/normal/high classification for each parameter in hematology and clinical chemistry will be summarized by treatment group.

The number and percent of subjects with treatment-emergent potentially clinically significant (PCS) laboratory values at any time after baseline will be summarized by treatment. The denominator is all subjects with normal baseline and at least one post baseline assessment in the safety population and the numerator is the number of subjects with PCS (i.e., meets Grade III or Grade IV toxicity criteria from the National Cancer Institute Common Terminology Criteria) at post-baseline.

7.3 VITAL SIGNS, PHYSICAL EXAMINATION FINDINGS, AND C-SSRS

7.3.1 Vital Signs

Vital signs include sitting body temperature, blood pressure (BP) and heart rate. On study drug dosing days BP and heart rate will be collected immediately pre dosing, immediately and 30 minutes post dosing, and body temperature will only be collected immediately pre dosing.

All vital signs will be presented in standard units. Frequency and percentage of subjects with values considered PCS occurring at any time post-baseline in dosing days will be summarized by visit (and time point) and by treatment. The denominator is all subjects with a baseline assessment in the safety population and the numerator is the number of subjects with PCS at post-baseline. Criteria for PCS are presented below Table 3.

Table 3 Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion	Definition of PCS at Any Post-baseline Time Point
Systolic blood pressure	Low	Value ≤90 mmHg and decreased ≥20 mmHg [1]
	High	Value ≥180 mmHg and increased ≥20 mmHg [1]
Diastolic blood pressure	Low	Value ≤50 mmHg and decreased ≥15 mmHg [1]
	High	Value ≥105 mmHg and increased ≥15 mmHg [1]
Pulse	Low	Value ≤50 bpm and decreased ≥15 bpm [1]
	High	Value ≥ 120 bpm and increased ≥15 bpm [1]

[1] For immediately and 30 minutes post-dose assessments on Dosing Days, the change (decrease or increase) is based on the pre-dose value obtained on each corresponding dosing day.

All vital sign data will be included in the listings

7.3.2 Physical Examination

Physical examination (PE) results were collected at screening, Days 42, 168 and 365, including five body systems, i.e. skin, cardiovascular, pulmonary, abdominal, and central nervous system. Each component of the baseline physical examination will be recorded as normal or abnormal. Each component of the post baseline physical examinations will be recorded as No Significant Change from Previous PE or Significant Change from Previous PE.

Physical examination results will be listed.

7.3.3 Columbia-Suicide Severity Rating Scale (C-SSRS)

The Columbia-Suicide Severity Rating Scale (C-SSRS) questionnaire will be completed at screening (baseline), Days 42, 168 and 365. The number (percentage) of subjects with a “yes” response will be summarized and listed for each question each visit by treatment group.

7.4 OTHER ANALYSES

Pregnancy test result (assessed at screening) will be listed.

8. INTERIM ANALYSES AND DATA AND SAFETY MONITORING BOARD (DSMB)

8.1 INTERIM ANALYSES

An interim analysis (IA) of efficacy will take place after all 200 subjects have completed Day 84. The efficacy endpoints analyzed will include:

- The treatment group difference in the IRLS total score change from baseline to Day 28, Day 42 and Day 84 estimated using the same analysis method as in Section 6.1.1, i.e. ANCOVA.
- The proportion of patients rated as much or very much improved with CGI-I on Day 42, and the treatment differences for proportions will be assessed with Logistic regression as in Section 6.1.2.
- The number (percentage) of subjects who receive a second dose of study product between Day 42 and Day 84, and the number (percentage) of subjects who discontinue the study between Day 42 and Day 84. Treatment difference for proportion will be assessed using Logistic regression with treatment, region (US, Europe), and baseline RLS medication related augmentation as fixed factors, and baseline IRLS total score as a covariate.

All endpoints will be analyzed by treatment group in overall FAS population. For IRLS total score and CGI-I, additional subgroup analyses will be conducted:

- by region (US vs. Rest of World)
- by baseline RLS medication related augmentation (no augmentation, uncertain augmentation, definitive augmentation)

Data from Day 0 to Day 84 will be frozen to generate IA results. Therefore, no adjustment to Type I error will be made.

8.2 DATA AND SAFETY MONITORING BOARD (DSMB)

The DSMB will be established and composed of approximately 3-5 senior academic individuals, including the DSMB Chair. They will have high-level expertise in neurology, hematology, cardiology and/or statistics. A senior statistician assigned to the trial from the group performing data management services for this trial will oversee the provision of interim data reports for use by the DSMB. The data management group for this trial will transfer pre-agreed datasets to the DSMB. During the Open Session of the DSMB meetings, the Study Chair or Luitpold representatives may present updates on the trial status or the safety profile of Ferric Carboxymaltose, but will not be privy to discussions of the data conducted during the Closed Sessions and will not vote. Proceedings and minutes of the Closed Session will be held in strict confidence and will not be shared outside the DSMB while the trial is ongoing. The DSMB will be responsible for the interests of the subjects and, to this end, will undertake reviews of the

safety data. The DSMB will have access to an agreed subset of the study data as listed in the DSMB charter (updated as necessary during the trial) throughout the study duration. In addition, the DSMB will evaluate the data approximately (as outlined in the Charter) either by face to face meeting or teleconference. The DSMB will determine if it believes the trial should be terminated early because clear evidence of a significant safety concern exists.

If the DSMB finds it necessary to recommend actions regarding interruption of the study or changes to the protocol based on medical rationale that would make it unethical to continue the study in its present form, those recommendations will be forwarded to the Study Chair and Sponsor. The details of the DSMB's functions and the early stopping rules will be delineated in a separate DSMB charter.

9. SUMMARY OF MAJOR CHANGES IN THE PLANNED ANALYSES

Compared with the study protocol (June 02, 2016 Amendment IV), below items are the major changes made in this SAP:

1. Efficacy analyses:
    - IRLS domain scores will not be analyzed, because only IRLS total score will be collected in EDC.
    - Baseline RLS medication-related augmentation will be included in statistical models where appropriate. In the protocol, it is not included.
    - CMH is the primary analysis method for CGI-I responder rate on Day 42 in study protocol. To get baseline covariate adjustment, Logistic regression model is proposed in this SAP. CMH will be performed as sensitivity analysis.

2. Interim analysis: The IA timelines have been updated from what is currently specified in the protocol. Originally, the IA was planned to take place once 100 subjects had completed Day 42, but due to a greater than anticipated rate of enrollment the IA will be performed after all 200 subjects have completed Day 84 to allow for a more accurate interpretation of the data. The change will be documented and submitted to FDA but not in a form of protocol amendment.

3. Disposition will be summarized for all enrolled subjects vs. for safety and FAS populations proposed in protocol. Number (percentage) of subjects who complete the Day 42 visit will be included in disposition summary that is not specified in protocol.

4. Medical history will be summarized for safety population vs. for all enrolled subjects proposed in protocol.

REFERENCES

1. Earley CJ. Horska A, Mohamed MA, Barker PB, Beard JL, Allen RP. A randomized, double-blind, placebo-controlled trial of intravenous iron sucrose in restless legs syndrome. Sleep Medicine 2009;10:206-11.

2. Allen RP, Adler CH, Du W, Butcher A, Bregman DB, Earley CJ. Clinical efficacy and safety of IF ferric carboxymaltose (FCM) treatment of RLS: A multi-centered, placebo-controlled preliminary clinical trial. Sleep Medicine 2011Oct;12(9):906-13.