Official Title: A Multicenter, Multinational, Extension Study to Evaluate the Long-Term Efficacy and Safety of BMN 190 in Patients with CLN2 Disease

NCT Number: NCT02485899

Applicant/MAH: BioMarin Pharmaceutical Inc.

Version Date: 10 Mar 2021

105 Digital Drive
Novato, CA 94949

Final Statistical Analysis Plan

Study BMN 190-202

A Multicenter, Multinational, Extension Study to Evaluate the Long-Term Efficacy and Safety of BMN 190 in Patients with CLN2 Disease

10 March 2021

1 APPROVALS

Title: A Multicenter, Multinational, Extension Study to Evaluate the Long-Term Efficacy and Safety of BMN 190 in Patients with CLN2 Disease

Protocol: 190-202

Date: 10 March 2021

DocuSigned by:
PI

PI
Director, Biometrics
Signing Reason: I approve this document
Signing Time: 12-Apr-2021 | 10:56 AM PDT
730FAE52504F4671AEDCC5E64085D1D2

Date

DocuSigned by:
PI

PI, ScD
Executive Director, Biostatistics, Biometrics
Signing Reason: I approve this document
Signing Time: 12-Apr-2021 | 4:30 PM PDT
0407169F552349068C56CF4CD8FD9ED6

Date

DocuSigned by:
PI

PI, MD
Executive Director, Global Medical Affairs
Signing Reason: I approve this document
Signing Time: 13-Apr-2021 | 6:19 AM PDT
F661E71EDB2C4B4CA7E014E624D27C9D

Date

DocuSigned by:
PI

PI
Director, Regulatory Affairs
Signing Reason: I approve this document
Signing Time: 12-Apr-2021 | 2:10 PM PDT
A49E99923BCC4E6C8EFFA0A95B660857

Date

DocuSigned by:
PI

PI, MD
Medical Director, Drug Safety Surveillance
Signing Reason: I approve this document
Signing Time: 12-Apr-2021 | 8:05 AM PDT
0113E866001145BCA1EDED4F75F14ADA

Date

DocuSigned by:
PI

PI, PhD
Senior Scientist 2, Immunogenicity Assessment

Signer Name: PI
Signing Reason: I have reviewed this document
Signing Time: 11-Apr-2021 | 2:47 PM PDT
BB47E0174A194E089B500C5E2AFD77AD

Date

DocuSigned by:
PI

PI, MD
VP, Medical Director, Clinical Science

Signer Name: PI
Signing Reason: I approve this document
Signing Time: 12-Apr-2021 | 12:15 AM PDT
C730A310B17242FF8F029CC5EFE93CB8

Date

DocuSigned by:
PI

PI, PhD
VP, Head of Biostatistics

Signer Name: PI
Signing Reason: I approve this document
Signing Time: 11-Apr-2021 | 6:29 PM PDT
027DA7AC705E447B9682C8910CF2CB1D

Date

TABLE OF CONTENTS

LIST OF APPENDICES

2 SUMMARY OF CHANGES

Study 202 is a study of extended treatment in patients who completed Study 201. Analyses and reporting have combined the data of Study 201 and Study 202 and have followed the SAP of Study 201. The final CSR for Study 202 will be based on combined data of Study 201 and Study 202 and will follow a new SAP with rationale given below.

Study 201 Version 2: The statistical analysis plan (SAP) for Study 201 was amended so that the population for the primary analysis of the primary endpoint was closer to ITT and consistent with the population for the primary analysis of the primary endpoint outlined in the integrated summary of efficacy (ISE). The original SAP required that the primary analysis include only subjects who have a CLN2 score between 3 and 6 inclusive just prior to initiation of 300 mg dosing, with a score of at least 1 in each of the two domains.
This requirement, an inclusion criteria for the study, was no longer a requirement for the primary analysis but was retained as a sensitivity analysis of the primary endpoint.

Study 201 Version 3: After discussions with regulatory agencies, the SAP was amended to place the responder analysis as the primary endpoint. The analysis of slopes remains an important endpoint and was performed as well. The primary analysis for each endpoint was based on the ITT population. Analyses were also performed after removing patients who entered the study with a CLN2 score of 6 that did not drop below 6 during follow-up.

The primary endpoint was defined as the absence of an unreversed 2-point decline or score of zero in CLN2 score by Week 48 (Study Day 340 relative to first 300 mg infusion).
An unreversed 2-point decline was clarified as “any decline of 2-points or more that had not reverted to a 1-point decline (or better) as of the last recorded observation”.

Study 202 Version 1: The Study 202 SAP is initiated as a revision of the Study 201 Version 3 SAP. The key change is that the primary endpoint of unreversed ML 2-point decline or score of zero has been revised (1) to evaluate over the full duration of follow-up instead of the initial 48 weeks of Study 201 (2) to compare against patient data from a natural history database (Study 901; two sample analysis) instead of against a fixed response rate of 50% (one sample analysis). The method of analysis for the primary endpoint is the Kaplan-Meier method and Cox proportional hazards model with covariate adjustment.
The rationale for this change is that with extended follow-up the focus of evaluation is long term and not the first 48 weeks on treatment. With longer-term and variable follow-up, patients will have variable follow-up (censoring) and the Cox model allows for evaluation of response rates with censoring. The Cox model also allows comparison to a natural history

cohort (Study 901) since it is unclear what response rate is an appropriate fixed-point comparison for a one sample analysis.

Study 202 Version 2: The Study 202 SAP is amended to include a survival analysis as an exploratory endpoint.

3 LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
AESI	adverse event of special interest
ANCOVA	analysis of covariance
ATC	Anatomical Therapeutic Chemical (classification system)
CLN2	late-infantile neuronal ceroid lipofuscinosis disease, also known as classical late-infantile CLN2, cLINCL, or Jansky-Bielschowsky disease, a form of Batten Disease
CSF	cerebrospinal fluid
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
ECG	electrocardiogram
EEG	electroencephalogram
HAE	hypersensitivity adverse event
HLGT	high-level group term
HRQL	health-related quality of life
ICV	intracerebroventricular,intracranial volume
IgE	immunoglobulin E
ITT	intent-to-treat
L	language subscale of the CLN2 disease rating scale
M	motor subscale of the CLN2 disease rating scale
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed effects model for repeated measures
mg	milligram
ML	Combined score of motor and language subscales on the CLN2 disease rating scale
MLV	Combined score of motor, language, vision subscales on the CLN2 disease rating scale
MLVS	Combined score of motor, language, vision, seizure subscales on the CLN2 disease rating scale
MRI	magnetic resonance imaging
NAb	neutralizing anti-TPP1 antibody
OCT	optical coherence tomography
PedsQL	measurement model for Pediatric Quality of Life Inventory
PLT	preferential looking test
PT	preferred term
QOL	quality-of-life

Abbreviation	Definition
qow	every other week
S	seizure subscale of the CLN2 disease rating scale
PT	preferred term (MedDRA or WHO drug)
SAE	serious adverse event
SAP	statistical analysis plan
SMQ	Standardized MedDRA Query
SOC	system organ class (MedDRA)
TAb	total anti-TPP1 antibody
TEAE	treatment-emergent adverse event
TRE	temporally-related event
WBV	whole brain volume
WHO	World Health Organization

4 INTRODUCTION

Study 190-202 is an extension of parent Study 190-201. It is a multi-center, multinational, extension study to evaluate BMN 190 treatment in patients with CLN2 who completed 190-201 (Study 201). All patients who have completed 48 weeks in Study 201 were eligible to enroll in Study 901-202 (Study 202).

The purpose of this Statistical Analysis Plan (SAP) is to present a description of the planned analyses for safety and efficacy data in Study 202. However, in order to characterize the safety and efficacy of longer-term treatment with BMN 190, Study 201 will be combined with the Study 202 to support analyses, i.e. data from Study 201 through study completion/ database lock will be pooled into Study 202. Furthermore, efficacy endpoint of CLN2 scores from the combined study (Study 201/202, treated subjects) will be compared with the Study 901 natural history controls (untreated subjects, herein Study 901).

For a detailed analysis plan for Study 201, please refer to Study 201 SAP version 3 dated on 25 March 2016.

Protocol history for Study 202 is as follows:

- Original protocol (3 October 2014)
- Amendment 1 (12 August 2015)
- Amendment 2 (16 November 2015)
- Amendment 3 (26 February 2016)
- Amendment 4 (17 March 2017)
- Amendment 5 (05 May 2017)
- Amendment 6 (17 December 2018)

The first subject was consented on 13 September 2013 for Study 201. The first subject enrolled in Study 202 on 02 February 2015.

4.1 Objectives of Study

The primary objectives of this study include the following:

- To evaluate the long-term safety of BMN 190 administration at 300 mg qow in patients with CLN2.
- To assess change in motor and language (ML) subscales of the CLN2 disease rating scale in patients with CLN2 receiving BMN 190 at 300 mg qow.

Secondary objectives of this study include the following:

- To assess changes in quantitative assessment of magnetic resonance imaging (MRI).
- To assess change in CLN2 disease scale total score.
- To evaluate quality of life (QoL) with long-term BMN 190 administration.

Exploratory objectives of this study include the following:

- To evaluate age-appropriate developmental milestones with long-term BMN 190 administration.
- To evaluate the impact of treatment on disease-related biomarkers from cerebrospinal fluid (CSF) and blood.

4.2 Study Design

This is a Phase 1/2, open-label, multi-center, multinational, extension study to evaluate the safety and efficacy of BMN 190 treatment in subjects with CLN2 who completed Study 201. All subjects who completed 48 weeks of BMN 190 treatment in Study 201 were eligible to enroll in Study 202.

The Screening period for Study 202 started simultaneously with the Week 47 visit in Study 201. Baseline values for Study 202 were recorded on the day of the first infusion (Week 1, Day 1) of Study 202, for all subjects on active treatment. The first dose of BMN 190 in Study 202 (Week 1/Study Day 1) was given following the Week 49 study assessments in Study 201. This study was open label, with all subjects continuing on treatment with BMN 190 300 mg qow.

4.3 Study Population

All patients who complete 48 weeks in the 201 study may be eligible to enroll in Study 202.

4.3.1 Key Inclusion Criteria

The key inclusion criterion is: Must have completed 48 weeks in Study 201.

4.3.2 Key Exclusion Criteria

If any of the following exclusion criteria apply, a patient will not be eligible to participate in the study:

- Has had a loss of 3 or more points in the combined motor and language components of the Hamburg CLN2 rating scale between Baseline of Study 201 and the Study Completion visit in Study 201 and would not benefit from enrolling in the study in the Investigator's discretion.

- Has a score of 0 points on the combined motor and language components of the Hamburg CLN2 rating scale.

4.4 Study Dosage and Administration

All study subjects will be administered BMN 190 300 mg by intracerebroventricular (ICV) infusion every other week.

4.5 Sample Size Determination

The sample size will be determined by the number of subjects who complete Study 201 and decide to roll into Study 202. Twenty-three subjects enrolled into Study 202.

4.6 Blinding and Randomization Methods

This is an open-label, non-randomized study.

4.7 Interim Analysis

Accruing study data will be summarized periodically for joint review by BioMarin and the Data Monitoring Committee (DMC). All summaries will be descriptive. Early stopping will be considered on the basis of safety only and will be based on clinical judgment; no inferential stopping rules will be employed.

Interim analyses of safety and efficacy were performed in September 2014 and January 2015 with plans to file data early to regulatory authorities if compelling efficacy were observed.

Interim analyses of safety and efficacy were performed based on data cut-off dates of:

15 October 2015: interim CSR for Biologics License Application

3 June 2016: 120 Safety Update

01 November 2016: interim CSR update for drug application

5 GENERAL ANALYSIS CONSIDERATIONS

Analyses will be conducted based on the combined data of Study 201 and Study 202 (Study 201/202). Efficacy analyses of CLN2 rating scales will also include comparison with a natural history database (Study 901), based on matching.

5.1 Analysis Populations

5.1.1 Efficacy

5.1.1.1 Intent-to-treat Analysis Population

The intent-to-treat (ITT) analysis population comprises all subjects from Study 201/202 who received any amount of study drug and report any efficacy results but excluding one subject (1287-1007) who terminated from the Study 201 after a single infusion of study drug due to unwillingness to continue with study procedures.

5.1.1.2 Study 901 Intent-to-treat Analysis Population

The Study 901 ITT population comprises the Study 901 subjects who satisfy the Study 201 inclusion criterion: Age ≥ 3, and at least one ML score ≥3 at age ≥3 years.

5.1.1.3 Study 901 Evaluable Population

The evaluable population will include Study 901 subjects in ITT population who have at least two CLN2 assessments with values within the range 1 - 5 and at least 6 months apart.

5.1.2 Safety Analysis Population

The Safety analysis population will comprise all subjects in Study 201/202 who had an ICV reservoir implanted in Study 201.

5.2 Treatment Group Presentation

Study 201/202 subjects will be pooled into a single group for the efficacy and safety analyses. Analyses that compare CLN2 scores for Study 201/202 subjects to Study 901 natural history controls will be presented as two groups and include the Study 201/202 ITT patients and Study 901 evaluable patients.

5.3 Pooling of Data from Sites with Small Enrollment

Given the small sample size of this study, subjects will be pooled across all sites for the primary efficacy and safety analyses.

Results may be examined within site on an exploratory basis.

5.4 Study Day Derivation

For dates on or after initiation of study drug, "Study Day" is defined as follows:

- Study Day = [Date – (date of first study drug)] + 1

For dates preceding initiation of study drug, "Study Day" is defined as follows:

- Study Day = [Date – (date of first study drug)]

Thus the day of first study drug is Study Day 1, and the day preceding first study drug is Study Day -1. There is no Study Day "0".

Date of first study drug is the first dosing date in Study 201.

For Study 901, the time of events is presented in terms of the subjects' age, in units of months. The Study 901 assessment that is used as a match to a Study 202/202 subject will be assigned as Study Day 1. Study Day is assigned to subsequent records assuming 30-day months, e.g., Study Day 31, 61, 91, etc.

5.5 Definition of Baseline

The definition of baseline will depend on context. The following general principles will apply:

- For evaluation of safety that includes the safety of the ICV reservoir, baseline will be the last observation preceding ICV implantation, if available. If no such pre-implantation observation exists, baseline will be the first value following ICV implantation, assuming that this is clinically reasonable.
- For evaluation of safety that is focused on the safety of BMN 190, baseline will be the last observation preceding the first infusion of BMN 190.
- For evaluation of efficacy of the 300 mg dose of BMN 190 with respect to CLN2 ratings and MRI measurements, baseline will be the last observation preceding the first 300 mg infusion.
- For evaluation of efficacy of the 300 mg dose of BMN 190 with respect to quality-of-life assessments (the disease-based QOL and the two PedsQL assessments) and developmental milestones (Denver II), baseline will, in general be the last observation preceding the first infusion of BMN 190.
- For the assessments conducted in Study 202 only (e.g. visual acuity testing, optical coherence tomography, electrocardiogram (ECG; 3- or 5-Lead.), neurological examination, hypersensitivity labs, ED-5D-5L QoL instrument etc), baseline will be the first available assessment in the 202 study.

For analyses of CLN2 scales based on matching, the baseline for subjects in Study 901 is taken as the assessment at the age (in months) for the match.

5.6 Visit Windows for Analysis

Patient visits for assessment of CLN2 scales and MRI is synchronized to a schedule relative to the first dose of 300 mg. CLN2 scales and MRI data will be windowed according to the schedule (see Appendix 1). For CLN2 scales in Study 201, the visit window defined for the 300 mg dose will be utilized (see Study 201 SAP dated on 25 March 2016).

Other efficacy and safety assessments are not synchronized and windowing will not be used.

5.7 Handling of Dropouts and Missing Data

PI . PI For example, AE dates will be imputed so as to ensure that an AE is considered treatment emergent and has the longest possible duration, assuming these are ambiguous with the missing/partial date.

For the primary efficacy analysis (responder) of CLN2 ratings, missing data will be handled by the Kaplan-Meier and Cox proportional hazards model methods. For the analysis of slopes, missing data is not expected to be an issue since the analysis is based on estimated slopes using baseline and last assessment.

A linear imputation algorithm is used for some of the analyses that compare Study 201/202 with Study 901. The CLN2 assessment schedule for Study 901 is not regular and does not generally correspond to CLN2 assessment frequency in Study 201/202. Summaries of CLN2 score using a window grid would have highly variable N which is difficult to interpret. For this reason, interpolation will be used. Summaries of Study 201/202 and Study 901 CLN2 score data will be based on data that has been imputed to the Study 201/202 visit target, using linear interpolation between Study 901 subject visits. Details are provided in Appendix 2.

Additional analyses will be conducted as appropriate to evaluate the impact of the COVID-19 pandemic on the study conduct and results, especially for the treatment effect as estimated in the trial.

6 SUBJECT DISPOSITION

The following will be summarized:

- the number of subjects enrolled
- the number of subjects who did not fulfill all inclusion/exclusion criteria
- the number of subjects who received study drug
- reason for termination from study
- length of time on Study 201/202

7 DISCONTINUATION AND COMPLETION

This is addressed in Section 6, Subject Disposition.

8 PROTOCOL DEVIATIONS

The number of subjects with protocol deviations in Study 201/202 will be summarized with respect to deviation class (major or minor) and deviation category (e.g., out of window, procedure not done etc.)

9 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Demographic variables (age, sex, ethnicity, and race), height, and weight will be descriptively summarized.

Baseline disease characteristics (e.g., age at diagnosis, genotype, and CLN2 scores at screening and prior to initiation of study drug) will be descriptively summarized.

10 MEDICAL HISTORY

Medical history will be coded in accordance with the most current version of MedDRA at the time of coding.

Medical history will be summarized by system organ class (SOC) and preferred term (PT).

11 PRIOR AND CONCOMITANT MEDICATIONS/PROCEDURES

Prior and concomitant medications in Study 201/202 will be defined as follows:

- A "prior" medication is any medication that was taken prior to the initiation of study drug in Study 201, regardless of dose level of study drug.
- A "concomitant" medication is any medication that was taken on or after the initiation of study drug in Study 201, regardless of the dose level of study drug. This can occur in two ways:
    - The medication was initiated on or after the initiation of study drug.
    - The dosing of a prior medication extended into the study drug dosing period. If the stop date of a prior medication is reported as "continuing", the prior medication will be assumed to be a concomitant medication as well.

Prior and concomitant medications will be mapped in accordance with the most current version of World Health Organization (WHO) Drug at the time of coding.

Prior and concomitant medications will be summarized separately, in terms of the number of subjects taking medications within each Anatomical-Therapeutic-Chemical class (ATC) and preferred term medication.

12 COMPLIANCE

The total amount of study drug (in mg) will be calculated as defined in the next section (Section 13, Extent of Exposure to Study Drug).

Measures of compliance will include:

- Actual study drug intake relative to planned study drug intake, expressed as a proportion:

$$100\% \times \frac{\text{Actual intake (mg)}}{\text{Planned intake (mg)}}$$

- Number (%) of missed infusions

These measures will be described descriptively over (1) the entire study period in Study 201/202 and (2) the 300 mg dosing period in Study 201/202.

13 EXTENT OF EXPOSURE TO STUDY DRUG

Exposure to study drug in Study 201/202 will be descriptively summarized in terms of the following:

- Total number of infusions, regardless of dose level
- Number of infusions at each dose level
- Total duration of dosing, regardless of dose level (time from first infusion to last infusion, in weeks)
- Duration of dosing at 300 mg (time from first 300 mg infusion to last infusion, in weeks)
- Total exposure over all infusions, regardless of dose level (in mg)
- Total exposure to 300 mg (in mg)

Exposure (mg) should be calculated as follows:

Exposure (mg) = [Volume (mL)] × [Study drug concentration (mg/mL)]

- If "volume" is explicitly recorded on the CRF, then use the recorded volume.
  If "volume" is NOT recorded on the CRF, and if "entire volume infused" is indicated, then assume the volume is 10 mL, as specified by protocol.
- "Study drug concentration" is determined from the "Dose" that is recorded on CRF page 18. Dose should be 30, 100, or 300 mg. (If otherwise, then flag.) We will assume that the concentration of prepared study drug for the three dose levels is as follows:

Dose	Concentration
30 mg	3 mg/mL
100 mg	10 mg/mL
300 mg	30 mg/mL

(Reference: Pharmacy Manual v2 10-Oct-2013)

If there are multiple records for a given infusion, e.g., then "exposure" is calculated separately for each record, and exposures are then summed across records.

14 EFFICACY EVALUATIONS

Study 201/202 efficacy will be evaluated from the first dose of 300 mg. Efficacy variables will be summarized by visit unless otherwise specified. For the CLN2 disease rating scales, comparison of Study 201/202 will be made with natural history data (Study 901).

For CLN2 scales, two sets of analyses comparing the Study 201/202 patients to Study 901 will be considered. The first set will compare the Study 201/202 ITT population and the Study 901 evaluable population. The second set will be based on 1-1 matching between the Study 201/202 ITT population and Study 901 evaluable cohort. The matched analyses will be based on a single matched set and the algorithm for 1-1 matching is described in Appendix 3. The matching criteria are:

- Equal baseline ML score
- Baseline age within 3 months
- Genome: equal number of common alleles (c.622C→T, c.509.1G→C)

We refer to the two sets of analyses as the full and matched sets. All analyses of CLN2 scales will use the full set and key analyses of CLN2 scales will also use the matched set as specified below.

14.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the time to unreversed ML 2-point decline or score of zero.

14.1.1 Primary Analysis Method

Time to unreversed ML two-point decline or score of zero will be analyzed using the Kaplan-Meier method and the Cox proportional hazards model. The Cox model will include baseline ML score and baseline age as continuous covariates and genome (common alleles) and sex as categorical covariates. The analyses will be performed on the full and matched sets.

14.1.2 Supportive Analysis

14.1.2.1 Time to Unreversed ML Score of Zero

Time to unreversed ML score of zero will be analyzed similarly as the time to unreversed ML 2-point decline or score of zero (Section 14.1.1).

14.1.2.2 Rate of decline

An important measure of treatment effect is the rate of decline in the 0-6 point ML score, scaled to a 48-week time period. Since the endpoint is a rate of decline, as opposed to a rate

of change, it is generally expected to be a positive number, with larger values representing a steeper deterioration of clinical status over time. The rate of decline will be estimated for each subject. Estimation of rate of decline is described in Appendix 4.

The mean rate of decline in Study 201/202 subjects will be compared to the mean rate of decline in Study 901 subjects using a 2-sample t-test. The null and alternative hypotheses to be tested are:

H_0: $\mu_{201/202} = \mu_{901}$

H_1: $\mu_{201/202} \neq \mu_{901}$

where $\mu_{201/202}$ represents the treated population mean rate of decline from the Study 201/202 and μ_{901} represents the untreated population mean rate of decline from Study 901. Treatment effect will be estimated as the difference in the mean of the slopes: Study 901 versus Study 201/202 subjects. Testing will use the two-sample t-test at significance level of α=0.05. The analyses will be produced for the full and matched sets.

An additional analysis of covariance (ANCOVA) model will be run for the full set. The model will include baseline ML score and baseline age as continuous covariates, genome (number of common alleles) and sex as a categorical covariate. Least square means and 95% CI for treatment versus natural history will be provided.

An additional mixed effects model for repeated measures (MMRM) will be run for the full set. Slope on treatment versus untreated will be estimated and tested using a random slope and intercept model. The model will include baseline ML score and baseline age as continuous covariates, genome (number of common alleles) and sex as a categorical covariate.

14.1.2.3 Motor (M) Scale

Time to unreversed motor 2-point decline or score of zero in Motor scale will be analyzed using the Kaplan-Meier method and the Cox proportional hazards model similarly as the primary endpoint (see Section 14.1.1).

The mean rate of decline of M scale in Study 201/202 subjects will be compared to the mean rate of decline in Study 901 subjects using a 2-sample t-test as described in Section 14.1.2.2.

The analyses will be performed on the full set only.

14.1.2.4 Language (L) Scale

Language (L) scale will be analyzed similarly to Motor (M) scale (Section 14.1.2.3).

14.1.2.5 Descriptive Summaries

For comparative summary of CLN2 scores, additional many-to-one matching between the Study 901 evaluable population and Study 201/202 ITT population will be conducted using criteria described in Appendix 3a.

Summary by Visit

To enable a descriptive assessment of CLN2 scores at common time points in which all subjects contribute to the descriptive statistics, each subject of Study 201/202 and Study 901 will have their CLN2 scores imputed to those time points using linear interpolation (see Section 5.7). The time points for interpolation will be in accord with the Study 201/202 schedule of CLN2 assessments.

Summary statistics of interpolated values and change from baseline by visit will be produced. A graph of mean change from baseline and standard error by visit will be produced. These analyses will be produced for the individual CLN2 scale domains motor (M), language (L), vision (V), seizure (S), and the composites ML, MLV, MLVS. The analyses will be produced for the full and matched sets (1-1 and many-to-one).

Summary statistics will also be produced for non-interpolated 201/202 data using the full set.

Individual Subject Plots

Data of each of the Study 201/202 subjects will be plotted with the matched Study 901 subject's data based on 1-1 and many-to-one matching.

14.2 Secondary Efficacy Endpoint(s)

The secondary efficacy endpoints will be analyzed for Study 201/202 using the ITT population.

14.2.1 MRI parameters

The key secondary endpoint, in concept, is brain atrophy. This will be evaluated with one or more of the following MRI measurements:

- Whole brain volume (WBV) (in mm^3)
- Volume of cerebrospinal fluid (in mm^3 and as a percentage of WBV)
- Volume of total cortical gray matter (in mm^3 and as a percentage of WBV)
- Total white matter volume (in mm^3 and as a percentage of WBV)
- Whole brain apparent diffusion coefficient (mm^2/s)

Values and change from baseline will be described descriptively over time.

14.3 Exploratory Endpoints

Exploratory endpoints will be analyzed using the ITT population.

14.3.1 Survival

Time to death will be analyzed using the Kaplan-Meier method and the Cox proportional hazards model. The analyses will be performed on the full and matched sets. For the analysis based on the matched set, survival will be measured from time of baseline to time of death (event) or time of last CLN2 assessment (censored). The Cox model will include baseline ML score and baseline age as continuous covariates and genome (common alleles) and sex as categorical covariates. For the analysis based on the full set, survival will be measured from birth to time of death (event) or time of last CLN2 assessment (censored). The Cox model will include genome (common alleles) and sex as categorical covariates.

14.3.2 CSF/Plasma Biomarkers

Analysis of disease-related biomarkers from CSF and blood will be conducted if usable biomarkers have been identified.

14.3.3 Developmental Milestones

Achievement of developmental milestones is assessed by the Denver II Development Scale for Study 201 and the Denver II Development Scale Updated for Study 202.

The Denver II Development Scale includes domain of Language and Gross Motor. Overall test interpretation is provided. Each domain includes the number of cautions and number of delays. The updated version in Study 202 includes domain of Personal Social, Fine Motor Adaptive, Language and Gross Motor. Each domain includes the number of passes, number of fails, number of cautions, number of delays and developmental age of most advanced pass. Overall test interpretation is provided as well.

The overall interpretation of the test – normal, suspect, or untestable – will be summarized at scheduled assessment times will be summarized for Study 201 and 202 separately.
For assessments done by Denver II Development Scale Updated, age equivalent performance (month) on personal social, fine motor adaptive, language and gross motor will be plotted over time using the age at assessment (month) for each subject. Age equivalence and change from baseline will be summarized. The ratio of change in age equivalence and change in age will be summarized at each visit as well as the last visit.

14.3.4 Anti-epileptic treatment

Graphical displays for each patient will include start and end dates for anti-epileptic treatment, with indication whether prophylactic or acute use. Convulsion adverse events (AEs) per SMQ and the 0-3 point S score will also be plotted on the patient chart.

14.3.5 Quality of Life

Quality of life is assessed with the following survey instruments:

14.3.5.1 PedsQL - Parent Report for Toddlers

This will be summarized as described in the PedsQL documentation, "Scaling and scoring of the Pediatric Quality of Life Inventory PedsQL". The following text briefly summarizes the scoring of the Parent Report for Toddlers:

The Parent Report for Toddlers is composed of 21 items comprising 4 dimensions:

- Physical Functioning (8 items)
- Emotional Functioning (5 items)
- Social Functioning (5 items)
- School Functioning (3 items)

Each item is scored on a 5-point Likert scale from 0 (Never) to 4 (Almost always).

The scoring procedure is a follows: Items are reverse-scored and transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0.

Score by dimension: If more than 50% of the items in the scale are missing, the scale scores should not be computed. The mean score = Sum of the items over the numbers of items answered.

Psychosocial Health Summary Score: This is the sum of the items over the number of items answered in the Emotional, Social, and School Functioning scales.

Total Score: This is the sum of all the items over the number of items answered on all the scales.

The various scores (by dimension, psychosocial health summary, and total scores) and change from baseline will be descriptively summarized.

14.3.5.2 PedsQL - Parent Family Impact

This will be summarized as described in the PedsQL documentation, "Scaling and scoring of the Pediatric Quality of Life Inventory PedsQL." The following text briefly summarizes the scoring of the Parent Report for Toddlers:

The full PedsQL Family Impact Module is composed of 36 items comprising 8 dimensions. In Study 201, BioMarin is collecting 5 of the 8 dimensions:

- Physical Functioning (6 items)
- Emotional Functioning (5 items)
- Social Functioning (4 items)
- Cognitive Functioning (5 items)
- Communication (3 items)

Each item is scored on a 5-point Likert scale from 0 (Never) to 4 (Almost always).

The scoring procedure is as follows: Items are reverse-scored and transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0.

Score by dimension: If more than 50% of the items in the scale are missing, the scale scores should not be computed. The mean score = Sum of the items over the numbers of items answered.

Parent HRQL Summary score (20 items): This is the sum of the items over the number of items answered in the Physical, Emotional, Social, and Cognitive scales.

(Modified) Total Score: This is the sum of all the items over the number of items answered on all the scales.

The various scores (by dimension, Parent HRQL, and the (modified) total scores) and change from baseline will be descriptively summarized.

14.3.5.3 CLN2 Disease-based Quality of Life

This instrument was developed by BioMarin, modeled upon the PedsQL instruments, and will be summarized similarly to the PedsQL instruments, as described below.

The CLN2 Disease-based QOL instrument is composed of 28 items comprising 6 dimensions:

- Seizures (6 items)
- Feeding / No G-tube (4 items)
- Feeding / with G-tube (3 items)
- Sleep (5 items)
- Behavior (6 items)
- Daily activities (4 items)

Each item is scored on a 5-point Likert scale from 0 (Never) to 4 (Almost always).

The scoring procedure is a follows: Items are reverse-scored and transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0.

Score by dimension: If more than 50% of the items in the scale are missing, the scale scores should not be computed. The mean score = Sum of the items over the numbers of items answered.

Total Score: This is the sum of all the items over the number of items answered on all the scales.

The scores (by dimension, and total) and change from baseline will be descriptively summarized.

14.3.6 Retinal Anatomy using Optical Coherence Tomography (OCT)

The retinal thickness parameters were assessed by Optical coherence tomography (OCT) for each eye in Study 202:

- Central Foveal
- Central Retinal
- Retinal – 3mm nasal to the fovea
- Retinal – 3mm temporal to the fovea

PI

14.3.7 Visual Acuity

All subjects undergo Preferential Looking Testing (PLT). In addition, for those children who retain the cognitive function to adhere to testing, the Lea Vision Test or E Hook (or Tumbling E) Vision Test are also performed during the same assessment.

The assessment of visual acuity was only conducted in Study 202. PI

14.3.8 EQ-5D-5L

The EQ-5D-5L instrument is a self-reported questionnaire designed to measure general health status (The EuroQol Group, 1990, Health Policy), (Brooks, 1996, Health Policy). The EQ-5D-5L is composed of 2 parts: a descriptive system that assesses 5 levels of perceived problems (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) in 5 dimensions and the EQ visual analogue scale (EQ VAS) assessment for overall health.

The assessment of EQ-5D-5L was conducted in Study 202 but was discontinued with the implementation of protocol amendment 6 (dated on 17 December 2018). PI

14.3.9 Electroencephalograms (EEG)

Electroencephalograms (EEG) will be summarized in terms of the proportion of subjects with epileptiform activity and/or frequency slowing, in combination with the activity's location (focal vs. generalized), at baseline and at any time after initiation of study drug.
The proportion of subjects showing new such activity (defined by the combination of activity and location) relative to baseline will be summarized. EEG evaluations from local and central vendor will be analyzed separately.

14.4 Examination of Efficacy by Subgroup

The impact of genotype, with a focus on the c.622C>T and c.509-1G>C mutations, on outcome may be examined.

Demographic characteristics, investigative site, or geographical region may be examined.

Various sub-intervals of the treatment period may be examined, e.g., the entire dosing experience, experience after initiation of 300 mg dosing, experience after entering Study 202.

All subgroup analyses are exploratory.

15 SAFETY EVALUATIONS

Safety will be assessed by examination of adverse events (including serious adverse events [SAEs], hypersensitivity adverse events (HAEs), infusion-associated reactions, and adverse events of special interest), clinical laboratory results (including chemistry, hematology, urinalysis, and CSF), vital signs, ECGs, and immunogenicity.

The Safety analysis population is defined in Section 5.1.2.

Summarization of safety data will be descriptive. No formal inference will be conducted.

Unless stated otherwise, baseline is generally defined as the last measurement prior to initiation of study drug, regardless of dose level in Study 201. However, as warranted by context, some summaries will use a value recorded prior to implantation of the ICV reservoir as baseline.

15.1 Adverse Events

Adverse events will be coded in accordance with the most current version of MedDRA at the time of coding.

Only treatment-emergent adverse events (TEAEs) will be summarized. A TEAE is defined as any AE that newly appeared, increased in frequency, or worsened in severity following ICV surgery. If the onset date of an AE is missing or indeterminate, the AE will be considered treatment-emergent.

15.1.1 All Adverse Events

Adverse events will be summarized in terms of incidence during the study period, by system organ class (SOC) and preferred term (PT).

Adverse events will be further summarized by severity (CTCAE grade). If a subject reports the same event more than once, with different severities, the greatest severity will be summarized.

The following subsets of adverse events will also be summarized:

- AEs scored by the investigator as “related” to study drug and/or study device
- AEs with outcome of death
- AEs that did not resolve, or that resolved with sequelae
- AEs resulting in termination from study and/or permanent discontinuation of study drug
- AEs resulting in dose reduction

15.1.2 Deaths and Serious Adverse Events

Serious adverse events (SAEs) will be summarized in a similar manner to general adverse events. PI .

15.1.3 Adverse Events Causing Premature Discontinuation

Adverse events that caused premature discontinuation of study or study drug will be listed.

15.1.4 Convulsion Events

Convulsion events are adverse events that map to the broad Convulsions Standardized MedDRA Query (SMQ). Convulsion events will be summarized by SOC, PT and severity. The convulsion events will also be summarized by 24 week study intervals: 0-24 weeks, 24-48 weeks, etc.

15.1.5 Hypersensitivity Adverse Events

A hypersensitivity adverse event (HAE) is defined as any adverse event that maps into either:

- the broad "hypersensitivity" Standardized MedDRA query (SMQ), or
- the broad algorithmic "anaphylactic reaction" SMQ

  Note: The algorithm requires either one or more of the following: an "A" event, or "B & C" events, or "B & D" events, or "C & D" events, with onset within 24 hours of start of a study drug infusion.

HAEs will be summarized in a similar manner to general adverse events.

15.1.6 Temporally-related events (TRE)

A temporally-related events (TRE) is defined as any adverse event with onset after initiation of a study drug infusion and within 24 hours of the start or restart of the infusion.

TREs will be summarized in a similar manner to general adverse events.

15.1.7 Adverse Events of Special Interest

"Adverse events of special interest" (AESIs) are defined in the study protocol amendment 6 (Section 10.3) to include:

- hypersensitivity adverse events (defined above)
- temporally-related events (defined above)
- status epilepticus
    - Status epilepticus is defined to comprise the following MedDRA preferred terms: petit mal epilepsy, status epilepticus.

- hydrocephalus (communicating and noncommunicating)
    - Hydrocephalus is defined to comprise the following MedDRA preferred terms: hydrocephalus, congenital hydrocephalus.
- meningitis
    - Meningitis is defined to comprise the MedDRA preferred terms presented in Appendix 5.
- unexpected rapid decline on CLN2 disease scale not attributable to other causes.
    - This will be identified by clinical review.
- Device-related events (e.g., infection, prophylactic ICV replacement, malfunction with an associated AE such as leaking reservoir or a problem that ends the administration of study drug for that visit, etc.)
- Cardiac and ECG events
    - It is defined to comprise the preferred terms within the SOCs of vascular disorders and cardiac disorders, the High Level Term of ECG investigations, and the HLGT of cardiac and vascular investigations.

Adverse events of special interest will be summarized in a similar manner to general adverse events.

15.2 Clinical Laboratory Tests

Chemistry, hematology, and urinalysis will be summarized by presentation of summary statistics (e.g., means, medians, standard deviations, etc.) for baseline values, for values observed after initiation of study drug in Study 201 (e.g., within-subject medians, minima and maxima over the follow-up period), and changes from baseline in these values.

These labs will also be summarized in terms of results relative to lab reference ranges (within, below, or above reference range). The baseline distribution will be presented. The within-subject percentage of follow-up lab values that are above reference range will be summarized; one summary will include all study subjects regardless of baseline values; a second summary will exclude subjects whose baseline values are above reference range. Similarly, the within-subject percentage of follow-up labs that are below reference range will be summarized in a manner analogous to that described above.

PI .

15.3 Vital Signs

Vital signs (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, and temperature) are recorded approximately every 2 weeks, prior to, during, and after each infusion of study drug. See the protocol Schedule of Events for the exact schedule.

Each vital sign will be summarized by presentation of summary statistics (e.g., means, medians, standard deviations, etc.) for baseline values, for values observed after initiation of study drug in Study 201 (e.g., within-subject medians, minima and maxima over the follow-up period), and changes from baseline in these values.

15.4 Height and Weight

Height and weight will be summarized in a similar manner as vital signs.

15.5 Physical Examinations

Clinically significant abnormalities observed during physical examinations are to be recorded under Medical History at Screening in Study 201 or as adverse events thereafter. Therefore, physical examinations will not be summarized but will be presented in listings only.

15.6 Electrocardiograms

Quantitative ECG (12-lead) parameters include heart rate, RR, PR, QRS and QT. ECG will be summarized by presentation of quantitative summary statistics (e.g., means, medians, etc.) for baseline values and values observed after initiation of study drug in Study 201 (e.g., within-subject medians, minima and maxima over the follow-up period) .
Changes from baseline at these timepoints will be similarly summarized.

QTcF will be further summarized in terms of the number of subjects with observed values >450, >480, or >500 msec, or with changes from baseline >30 or >60 msec.

The number of subjects with one or more ECGs that are judged abnormal, and clinically significantly abnormal, will be summarized.

The continuous ECG monitoring (3- or 5-lead) were performed for all subjects for at least one infusion of BMN 190 (and preferably the next infusion) in Study 202, The ECG should begin 15 (± 5 minutes) prior to infusion start, continue through infusion of BMN 190, and end after infusion of flushing solution. If a 12-lead ECG is required during this time, continuous monitoring should be interrupted in order to obtain the 12-lead ECG. The 3- or 5-lead ECG data in Study 202 will be provided in a PI The overall interpretation from investigator ECGs will be cross tabulated using a shift table at baseline vs. the worst post-baseline results and the last visit.

15.7 Neurological Examination

The neurological examination was conducted every 12 weeks in Study 202. It includes mental status, speech language, cranial nerve examination, motor strength, motor tone, abnormal movements, reflexes, sensory, gait, romberg`s test, nystagmus and coordination. The neurological examination results (normal vs abnormal) will be listed only.

15.8 Immunogenicity

Immunogenicity tests will be performed using validated immunogenicity assays. Routine immunogenicity tests will include total antibody (TAb) and neutralizing antibody (NAb) in the CSF and serum. NAb testing will not be performed if the TAb is negative. A subset of serum and CSF NAb samples were tested to detect inhibition of cerliponase alfa enzyme activity. For all assays conducted, incidence and titer summary statistics will be provided for serum TAb, CSF TAb, serum NAb and CSF NAb in table format and will include mean, median, standard deviation, and minimum/maximum titer values at each study visit. Titer results will be presented in a cumulative manner to include 190-201 data and summarized by visit as well as by change at study timepoint from study baseline in 190-201. Potential impact of anti-drug antibodies (ADA) on efficacy and safety will be explored.

In the event of serious or severe (≥Grade 3) hypersensitivity AE, a blood sample will be collected no sooner than 8 hours after the event (or before the next infusion) for drug-specific IgE testing. Potential associations between immunogenicity results and hypersensitivity adverse events might be analyzed.

16 REFERENCES

Brooks, R. EuroQol: the current state of play. Health Policy 37[1], 53-72. 1996.

Denver II Training Manual, Denver Developmental materials, Inc., P.O. Box 371075, Denver, CO 80237-5075 (last copyright: 2009, Wilhelmine R. Frankenburg)

Outbound Data Delivery Specification; Document version: Draft_v0.1_02Apr2014; Protocol Identifier: 190-201. ICON Medical Imaging, 2800 Kelly Road, Suite 200, Warrington, PA.

Pharmacy Manual v2 10-Oct-2013, Clinical protocol number 190-201. BioMarin Pharmaceutical Inc.

The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 16[3], 199-208. 1990.

17 APPENDICES

Appendix 1: Windows

Assessment	Derived Visit	Scheduled Visit Day [b]	Window (day)
CLN2 disease rating scales i=1,2…	Week 49 [a]	337	[-, 365]
	Week 49 + i*8	337 + i*56	[310+i*56, 365+i*56]
MRI i=1,2…	Week 9	57	[-, 113]
	Week 25	169	[114, 253]
	Week 49	337	[254, 421]
	Week 49 + i*24	337+i*168	[254+i*168, 421+i*168]

[a]: Week 49 starts at the first day in Study 202

[b]: It is the study day relative the first dose in Study 201

Appendix 2: Imputation of CLN2 Scores at Nominal Time points

To enable a descriptive assessment of CLN2 scores at common time points in which all subjects contribute to the descriptive statistics, each subject of the 201/202 and 901 studies will have their CLN2 scores imputed to those time points.

Nominal time points are Weeks 4, 8, 12, … 144, 168 with corresponding target Analysis Days 29, 57, 85, ... 1009, 1177.

The imputed CLN2 value at each target Analysis Day is calculated by linear interpolation. If a target Analysis Day is not bracketed on both sides by CLN2 scores, but there is a single CLN2 assessment within 28 days, inclusive, of the target Analysis Day, then that single assessment is used for imputation (observation carried forward). The imputed data of the 901 patients is truncated to have no longer follow-up than the corresponding imputed data for the Study 201/202 patient matched.

Appendix 3: Matching of Study 901 and Study 201/202 patients (1-1)

This matching algorithm is based on maximizing the number of Study 201/202 subjects matched to Study 901 subjects and satisfying several criteria (baseline ML score equal, genome: equal number of common alleles, baseline age close and no more than 3 months apart). The data of Study 901 subjects will be restricted and includes the assessment at the age of the match as the baseline assessment. Duration of follow-up is measured with respect to this baseline. Follow-up assessments up to the largest duration that is less than or equal to

the full follow-up duration of the matched 201/202 subject are included for the matched analysis. If this derived duration of follow-up for the Study 901 subject is not of duration 6 months or greater then matching at this age of assessment will not be considered.
For Study 901, where the first assessment of ML has the value 6, backwards imputation of the value 6 to earlier ages is allowed.

Distance is defined as the absolute value of the difference in baseline age for the potential match. Study 201/202 subjects are paired off from first through last. No Study 901 subject is matched more than once. To maximize the number of matched pairs, and overall low mean squared distance, at each pairing:

- Identify the Study 201/202 subject who has not yet been matched and has the least potential Study 901 candidates for pairing based on the requirement for equal baseline ML score, equal number of common alleles and distance ≤ 3. If greater than one Study 201/202 subject is identified, break the tie by considering the number of potential Study 901 candidates for matching based on the requirement of distance ≤ 2. Potentially there are still ties and repeat as needed using distance criteria based on thresholds ≤ 1 and ≤ 0.
- For the selected Study 201/202 subject, match with the Study 901 subject who has not yet been matched and has a (ML, age of assessment) combination that minimizes the distance measure. There may be greater than one Study 901 subject that satisfies minimal distance in which case choose the 901 Subject who has fewest potential Study 201/202 matches based on the distance ≤ 3 (and ≤ 2, ≤ 1, ≤ 0 as needed).
If there are no Study 901 subjects who satisfy distance ≤ 3 then there is no match for the Study 201/202 subject.
- It is possible that there remains greater than one 901 patients selected. In this situation, ties are broken down based on the following criteria ordered: exact genome, sex, seizure age, onset age, country.
- Repeat till all Study 201/202 subjects have been attempted for match. This completes the 1-1 matching.

Appendix 3a: Matching of Study 901 and Study 201/202 patients (many-one)

To facilitate comparison of Study subjects to Natural History subjects with respect to the clinical course over time, as represented by the 0- to 6-point CLN2 scores vs. time, Study 201/202 subjects will be matched to Study 901 Natural History subjects on the basis of study subjects' baseline CLN2 scores. Once the matches based on subject baseline CLN2 scores have been identified, more refined matching will be done by further matching on the basis of

(1) subject age at baseline and (2) genotype. That is, three sets of matching will be done, based on these matching criteria:

- Study subject's baseline CLN2 score
- Study subject's baseline CLN2 score and age at baseline
- Study subject's baseline CLN2 score and genotypic classification

The matching algorithms can result in zero, one, or multiple Study 901 Natural History matches for any given Study 201/202 subject.

Algorithm for matching on Study subject's baseline CLN2 score

For a given Study 201 subject, identify the baseline CLN2 score and analysis day.

Then identify each Study 901 subject who has one or more records with a CLN2 score equal to the given Study 201 subject's baseline CLN2 score. For each such Study 901 subject, determine the mid-point of these records that match Study 201 subject baseline score, and consider this to be the "baseline" record for this Study 901 subject when paired with this Study 201 subject. Assign an "analysis day" time value to this Study 901 baseline record that equals the Study 201 subject's baseline analysis day. Assign "analysis days" to subsequent Study 901 records relative to this pseudo-baseline record; use the convention that each "month" has 30 days.

Algorithm for further matching on Study subject's age at baseline

The algorithm is the same as matching on baseline CLN2 score alone , with the additional requirement that the age (in years) of the Study 901 subject's matched baseline is equal to the given Study 201 subject's age at enrollment.

Algorithm for further matching on Study subject's genotype

For purpose of matching on genotype, "genotype" will be defined as a trichotomous variable, based on the presence or absence of the two predominant mutations, c.622>T and c.509-1G>C, as follows:

- Both: Both alleles are predominant mutations – that is, two of one mutation, or one of each mutation
- One: Exactly one allele is a predominant mutation. The other allele is not a predominant mutation.
- None: Neither allele is a predominant mutation.

Note: We assume understanding is that the IVS5-1G>C mutation identified in the Study 901 dataset is the same as the c.509-1G>C mutation in the Study 201 database.

Similarly to matching on (baseline CLN2 score and) baseline age, to further match on Study subject's genotype, the algorithm starts with the matching obtained by matching on subjects' baseline CLN2 scores. For a given Study 201 subject, the Study 901 matches will be culled to retain only those Study 901 subjects whose genotypic classification is equal to the Study subject's genotypic classification.

Appendix 4: Estimation of Rate of decline

The purpose of this appendix is to describe the algorithm for estimation of rate of decline in study subjects.

General Algorithm

The rate of decline is calculated as follows:

1. Identify a starting point and an ending point, where a "point" is a bivariate observation comprised of (1) a CLN2 score and (2) a timepoint.
2. Determine the slope of the line connecting the two points:

$$\text{Slope} = \frac{(\text{Ending CLN2 score}) - (\text{Starting CLN2 score})}{(\text{Ending date}) - (\text{Starting date})}$$

3. Calculate the rate of decline as the negative of the line's slope, scaled to a 48-week time period:

$$\text{Rate of decline} = (-1) \times (48 \times 7) \times \text{Slope}$$

A General Consideration in Selecting the Starting Point

When a subject with a baseline CLN2 score <6 enrolls in the study, we assume that the subject enrolls at a random time point within the time period at which the subject "resides" at that value. This assumption has implications for the determination of the starting point for subjects with an initial CLN2 score of 6.

If the subject enrolling with CLN2 = 6 does not subsequently have a drop in CLN2 score, then the starting point is the baseline assessment. Otherwise we must wait until the subject's score drops to <6 before using the subject in analysis. If the subject has multiple records at that lower score, it would be inappropriate to select the earliest such observation, given the assumption in the previous paragraph. To consistently select the earliest (or last) observation could introduce bias, by systematically inflating (or decreasing) the time until a subsequent drop. To mitigate the introduction of bias in this situation, we therefore select the midpoint between the first and last of any multiple observations.

Estimation of Rate of Decline in Study Subjects

The starting assessment will be defined by the following algorithm:

- Identify the baseline CLN2 assessment appropriate for the given analysis.
- If the baseline CLN2 score is less than 6, then use this assessment; skip the next bullet.
- If the baseline CLN2 score is a 6:
    - If the subject did not have a decline in CLN2 score during follow-up then use this assessment.
    - If the subject drops below a 6, move forward in time to the earliest CLN2 rating that is less than or equal to 5. Determine the set of contiguous visits with this CLN2 score; this may be a single observation or multiple observations. Define the starting time as the mid-point between the first and last of these visits.

The ending assessment is the last observation with a CLN2 score greater than 0.

Once starting and ending points have been identified, calculate the rate of decline and scale it to 48 weeks as described above.

Appendix 5: MedDRA Preferred Terms Corresponding to Meningitis

The study protocol identifies "meningitis" as an adverse event of special interest (AESI). Meningitis is defined to comprise the MedDRA preferred terms presented below:

Central nervous system infection	Meningitis histoplasma
Meningitis	Meningitis leptospiral
Meningitis aseptic	Meningitis listeria
Meningitis aspergillus	Meningitis meningococcal
Meningitis bacterial	Meningitis mumps
Meningitis borrelia	Meningitis neonatal
Meningitis candida	Meningitis noninfective
Meningitis chemical	Meningitis pneumococcal
Meningitis coccidioides	Meningitis salmonella
Meningitis coxsackie viral	Meningitis staphylococcal

Meningitis cronobacter
Meningitis cryptococcal
Meningitis echo viral
Meningitis enterococcal
Meningitis enteroviral
Meningitis eosinophilic
Meningitis exserohilum
Meningitis fungal
Meningitis gonococcal
Meningitis haemophilus
Meningitis herpes
Meningitis streptococcal
Meningitis toxoplasmal
Meningitis trypanosomal
Meningitis tuberculous
Meningitis viral
Herpes zoster meningitis
Herpes simplex meningitis
Pachymeningitis
Propionibacterium acnes
Pseudomonas aeruginosa meningitis