Official Title: A Phase 2, Open-Label, Multicenter Study to Evaluate Safety, Tolerability, and Efficacy of Intracerebroventricular BMN 190 in Pediatric Patients < 18 years of age with CLN2 Disease

NCT Number: NCT02678689

Applicant/MAH: BioMarin Pharmaceutical Inc.

Version Date: 18 April 2022

16.1.9 Documentation of Statistical Methods

The following materials are provided in this appendix:

- 190-203 Interim SAP (25 June 2020; Version 1.0)
- 190-203 Final SAP (18 April 2022; Version 2.0)
- 190-203 Final Database Errata

BMN 190-203 Final Database Errata

The following discrepancies were noted, and per BioMarin, these data points will not be updated/changed in the database.

Errata:

Subject	Source	Visit	Form	Field	Discrepancy Description	Actual Database Value	Confirmed Value	Justification
PI	EDC	Week 83	Major protocol deviations	Deviation category	*Field was completed incorrectly.*	Procedure not done	Out of window	Participant discharge at Week 83, without overnight, due to COVID-19 precautions should have been categorized as Out of Window. Procedures for this visit were completed.
PI	Eurofins	Week 97			CSF cell uptake NAb sample not tested			CSF TAb samples were tested per protocol, however, CSF NAb samples from TAb positive participants were collected and not tested because additional NAb testing did not inform decision making related to continuing therapy of clinical trial participants with Brineura or alter the interpretation of immunogenicity risk assessment of Brineura. TAb testing is sufficient to inform the benefit to risk assessment of Brineura in this patient population.
PI	Eurofins	Week 109			CSF cell uptake NAb sample not tested			Same as above
PI	Eurofins	Week 121			CSF cell uptake NAb			Same as above

Subject	Source	Visit	Form	Field	Discrepancy Description	Actual Database Value	Confirmed Value	Justification
					sample not tested			
PI	Eurofins	Week 133			CSF cell uptake NAb sample not tested			Same as above
PI	Eurofins	Study Completion/ ETV			CSF cell uptake NAb sample not tested			Same as above
PI	Eurofins	Week 49			CSF cell uptake NAb sample not tested			Sample Quantity Not Sufficient (QNS)
PI	Eurofins	Week 97			CSF cell uptake NAb sample not tested			CSF TAb samples were tested per protocol, however, CSF NAb samples from TAb positive subjects were collected and not tested because additional NAb testing did not inform decision making related to continuing therapy of clinical trial participants with Brineura or alter the interpretation of immunogenicity risk assessment of Brineura. TAb testing is sufficient to support the benefit - risk assessment of Brineura in this patient population.
PI	BAS	Week 97, 4 hour EOI		Plasma PK		236000 ng/mL		Incorrect dilution factor. Datapoint not included in analysis.

Subject	Source	Visit	Form	Field	Discrepancy Description	Actual Database Value	Confirmed Value	Justification
PI	BAS	Week 97, 72 hour EOI		Plasma PK		39.5 ng/mL		%CV > 20. Inadvertently not retested. Datapoint not included in analysis
PI	EDC		N/A	Adverse Event	Duplicate events	N/A	N/A	Duplicate events were captured in EDC for the following three AEs: Generalised tonic-clonic seizure (AE # 20 & 21), Partial seizures (AE # 25 & 26), Partial seizures (AE # 27 & 28) Note: These duplicate events led to overcounting of 3 more events of seizure, but did not have any impact on incidence numbers. Seizure is a known risk for study drug. There was no impact on safety conclusions due to this error.
PI	EDC		N/A	Adverse Event	Duplicate event	N/A	N/A	Duplicate events were captured in EDC for following AE: Hypersensitivity (AE # 4 & 3) Note: This duplicate event led to overcounting of 1 event of hypersensitivity, but did not have any impact on incidence numbers. Hypersensitivity is a known risk for study drug. There was no impact on safety conclusions due to this error.

Subject	Source	Visit	Form	Field	Discrepancy Description	Actual Database Value	Confirmed Value	Justification
Pl	EDC	Weeks 93, 133	Study Drug Administration	Entire Volume Infused	Field value not confirmed	No	UNK	Reason for drug interruption was listed as "Brineura instead of study drug was administered at Rotterdam site." This does not provide a reason for incomplete or interrupted infusion, and it is unclear if the full dose was administered or if the dose was interrupted.
Pl	EDC	Weeks 117, 119, 121, 123, 125, 127, 129, 131, 135, 137, 139, 141, 143	Study Drug Administration	Entire Volume Infused	Field value not confirmed	No	UNK	Reason for drug interruption was listed as "COVID-19 related travel restrictions; therefore, Brineura treatment in home country" This does not provide a reason for incomplete or interrupted infusion, and it is unclear if the full dose was administered or if the dose was interrupted.

AE, adverse event; BAS, Bioanalytical Sciences; CSF, cerebrospinal fluid; CSR, clinical study report; EDC, electronic data capture; EOI, end of infusion; N/A, not applicable; NAb, neutralizing anti-TPP1 antibody; PK, pharmacokinetics; QNS, quantity not sufficient; UNK, unknown

Final Statistical Analysis Plan

Study BMN 190-203

A Phase 2, Open-Label, Multicenter Study to Evaluate Safety, Tolerability, and Efficacy of Intracerebroventricular BMN 190 in Pediatric Patients < 18 years of age with CLN2 Disease

Version 2

18 April 2022

SIGNATURE AND APPROVAL PAGE

Please see attached signature page.

TABLE OF CONTENTS

LIST OF APPENDICES

1 LIST OF ABBREVIATIONS

Abbreviation	Definition
ADA	Anti-drug antibodies
AE	adverse event
AESI	adverse events of special interest
ATC	anatomical therapeutic chemical
BMI	body mass index
BP	blood pressure
CI	confidence interval
CRF	case report form
CSF	cerebrospinal fluid
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events (v4.0)
ECG	electrocardiogram
EEG	electroencephalograms
eCRF	electronic case report form
HAE	hypersensitivity adverse event
ICV	intracerebroventricular
ITQOL	infant toddler quality of life questionnaire
ITT	intent-to-treat
L	language
M	motor
MedDRA	Medical Dictionary for Regulatory Activities
ML	motor-language
MLV	motor-language-vision
MLVS	motor-language-vison-seizure
MRI	magnetic resonance imaging
NAb	neutralizing anti-TPP1 antibody
OCT	optical coherence tomography
PD	pharmacodynamic
PK	pharmacokinetic
PP	per protocol
PR	pulse rate
PT	preferred term
S	seizure
SAE	serious adverse event

SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SMQ	standardized MedDRA query
SOC	system organ class
TAb	total anti-TPP1 antibody
TEAE	treatment-emergent adverse event
TLGs	tables, listings, and graphs
TRE	temporally-related events
V	vision
VAS	visual analogue scale
WBV	whole brain volume
WHO	World Health Organization

2 SUMMARY OF CHANGES

The initial statistical analysis plan (SAP) was signed on 25 June 2020 for an interim clinical study report requested by the European Medicines Agency Pediatric Committee. The changes made to the SAP for the final 190-203 clinical study report (CSR) are described below.

Appendix 3: Matching of Study 190-901 and Study 190-203 subjects

- Matching criteria were updated (1) to add the tie-breaking criterion of most recent birthdate; and (2) to add onset and first seizure age as the first two tie-breaking criteria, with the further requirement that subjects with onset and first seizure ages greater than the age of their match are set to missing.

Section of 13.1.2 Sensitivity Analysis Methods(s)

- An additional sensitivity analysis was added to repeat the primary analysis, but without the use of weights, for the rate of ML decline.
- Time to unreversed ML score of zero was added as an analysis.

Section 13.3.3 Denver II Developmental scale

- Description for Denver Development Scale was updated. Additional summary for overall interpretation was added.

Section 13.3 Exploratory Endpoints

- All patient reported outcome endpoints, except for Denver II, will be listed only.

Section 13.4 Examination of Efficacy by Subgroup.

- The summary of ML scores and MRI parameters over study time will be repeated for the subgroups of baseline age <2, <3, and ≥3 years.

Section 14.4 Electrocardiogram (ECG)

- Additional analyses have been added, including summary of the subjects with one or more abnormal ECGs, and a data listing for subjects with clinically significantly abnormal 3-/5-lead ECG data.

3 INTRODUCTION

This document describes the statistical methods to be implemented in the analysis of data collected under clinical study protocol 190-203, "A Phase 2, Open-Label, Multicenter Study to Evaluate Safety, Tolerability, and Efficacy of Intracerebroventricular BMN 190 in Pediatric Patients < 18 years of age with CLN2 Disease" (Amendment 6, 05 February 2019). Analyses of pharmacokinetics data are outlined in a separate document. This statistical analysis plan (SAP) contains definitions of analysis populations, derived variables, and statistical methods for the analyses of efficacy and safety. There will be two sets of efficacy analyses. The first set are descriptive summaries of the Study 190-203 subjects at baseline and follow-up visits. The second set are based on matching Study 190-203 subjects with historical controls and includes the primary efficacy analysis. The historical controls are from Study 190-901: DEM-CHILD Multi-Center Clinical NCL Database.

3.1 Objectives of Study

The primary objectives of this study include the following:

- evaluate safety and tolerability of BMN 190 administered via intracerebroventricular (ICV) device
- to evaluate treatment effectiveness as a delay in progression of motor-language (ML) score on the Hamburg CLN2 clinical rating scale

Secondary objectives of this study include the following:

- assess immunogenicity of BMN 190 in CSF and serum
- to characterize the pharmacokinetics of BMN 190 in CSF and plasma
- measure MRI parameters of disease progression
- assess impact of treatment on the total Hamburg clinical rating scale
- assess the time to disease manifestation for asymptomatic patients

Exploratory objectives of this study include the following:

- assess development achievement
- assess abnormal involuntary movements
- evaluate retinal anatomy using optical coherence tomography (OCT)
- determine seizure onset, type and frequency

- determine change in seizure activity
- anti-epileptic treatment
- assess quality of life metrics
- assess changes in EEG
- assess changes in visual acuity
- analysis of disease-related biomarkers from CSF and blood

3.2 Study Design

This is a Phase 2 open-label, multicenter study in pediatric patients < 18 years of age with CLN2 disease. The study is designed to assess disease progression in CLN2 patients treated with BMN 190 compared to natural history data from untreated historical controls (data from a CLN2 disease patient registry and data from siblings with CLN2 disease, when available). The comparison of treated subjects with untreated controls is from the time of the initiation of study drug.

The planned enrollment for this study is at least 10 subjects: at least 5 subjects with a Hamburg ML score ≥ 5 points, at least 5 subjects with a ML score < 5 points, and at least 5 subjects < 2 years of age. Study eligibility and enrollment occur before (≤ 21 days) subjects are admitted to the hospital for surgical implantation of an ICV access device. The first infusion is administered at least 14 days from surgery to implant the ICV reservoir and no more than 28 days after surgery. The Baseline visit is to collect clinical scores and clinical laboratory parameters and is to be completed no more than 2 days before the first infusion.

BMN 190 are administered every 14 days from the date of the first infusion (±3 days) for at least 144 weeks according to the subject's age. At study completion or early termination visit, subjects who have ICV device removal are required to have a device safety follow-up visit at 4 weeks (±3 days) after ICV device removal. All subjects require a safety follow-up visit at 6 months (±1 week) after last dose.

CLN2 assessments are scheduled every 4 weeks, study completion/early termination, and safety follow-up visits. Cranial MRI is assessed every 24 weeks and at study completion/early termination visit. Immunogenicity assessment is based on CSF and serum lab draws collected at baseline, every 12 weeks, study completion/early termination, and safety follow-up visits (serum only). PK assessment is based on CSF and serum lab draws every 24 weeks and at study completion/early termination visit.

The safety and tolerability of treatment will be assessed by collection of adverse events, physical findings, vital signs, ECG, EEG, and clinical laboratory tests.

3.3 Study Population

This is a study in pediatric patients < 18 years of age with CLN2 disease. Key inclusion and exclusion criteria are as follows:

Inclusion Criteria

- Clinical assessment of the Hamburg motor-language aggregate score 3-6 at Screening

Exclusion Criteria

Another inherited neurologic disease, e.g., other forms of CLN or seizures unrelated to CLN2 disease (patients with febrile seizures may be eligible)

Another neurological illness that may have caused cognitive decline (e.g., trauma, meningitis, hemorrhage) or interfere with disease rating (autism) before Screening

Has received stem cell, gene therapy, or ERT

Episode of generalized motor status epilepticus within 4 weeks before the First Dose visit

Severe infection (e.g., pneumonia, pyelonephritis, or meningitis) within 4 weeks before the First Dose visit (enrollment may be postponed)

Has received any investigational medication within 30 days before the first infusion of study drug or is scheduled to receive any investigational drug other than BMN 190 during the course of the study

3.4 Study Dosage and Administration

BMN 190 is administered every 14 days from the date of the first infusion ($\pm$3 days) for at least 144 weeks according to the subject's age:

- birth to < 6 months: 100 mg
- 6 months to < 1 year: 150 mg
- 1 year to < 2 years: 200 mg (first 4 doses), 300 mg (subsequent doses)
- ≥2 years: 300 mg

3.5 Sample Size Determination

Sample size was determined on the basis of clinical judgment. Statistical power was not a consideration.

3.6 Blinding and Randomization Methods

This is a single-arm, open-label study; there is no blinding to treatment or randomization.

3.6.1 Interim Analysis

No formal interim analyses were planned. Analyses were produced for a biologics licensing application to FDA based on visits through November 01, 2016. Analyses were produced for purpose of regulatory reporting and congress presentations based on data through April, years 2018, 2019, 2020, and 2021. An interim clinical study report (CSR) based on visits up to 26 April 2020 was produced at the request of the European Medicines Agency Paediatric Committee. This interim report included safety and the following efficacy analyses:

- Summary statistics ML rate of decline
- Summary statistics ML rate of decline subset to baseline ML<6
- Summary statistics ML rate of decline subset to baseline ML=6
- Kaplan-Meier time to unreversed ML 2-point decline or score of zero
- Kaplan-Meier time to MLVS<12 for patients with baseline MLVS=12
- ML summary statistics by visit
- M, L, V (Vision), and S (Seizure) score summary statistics by visit
- MLV and MLVS (Motor-Language-Vision-Seizure) summary statistics by visit
- MRI parameter summary statistics by visit
- Immunogenicity
- Adverse events
- Clinical Laboratory Tests
- Vitals signs
- Electrocardiogram (ECG)
- Electroencephalograms (EEG)
- Neurological Examination

- Cerebrospinal Fluid Surveillance (CSF)

The analyses of ML, MLVS, S also included matched analyses with natural history data.

There is no Data Monitoring Committee. However, BioMarin has convened a Data Review Board of internal and external experts to assist the Medical Monitor in assessing data.

Early stopping is to be considered based on clinical judgment; no inferential stopping rules are being employed.

4 GENERAL ANALYSIS CONSIDERATION

Safety and efficacy variables will be summarized descriptively. Descriptive statistics include subject count, mean, median, standard deviation, minimum, and maximum for continuous variables and count and percentage for categorical variables. The 95% confidence interval for mean and percentiles may be included, if appropriate.

4.1 Analysis Populations

4.1.1 Efficacy

The Intent-to-Treat Analysis Population (ITT) is defined as all enrolled subjects who received study drug. Efficacy analyses will be based on the ITT population.

The Per Protocol Population (PP) is defined as all enrolled subjects who receive study drug and satisfy major inclusion and exclusion criteria. All enrolled subjects received study drug, and 1 subject (PI) was screened with an ML score of 1 in violation of inclusion/exclusion criteria. This subject is not included in the PP population. The PP population will be used for sensitivity analyses of the primary efficacy endpoint.

4.1.2 Safety

The Safety Analysis Population is defined as all subjects who had an ICV reservoir implanted.

4.1.3 Study 190-901 Evaluable Population

Study 190-901 is a natural history study of patients diagnosed with CLN2 disease. The Study 190-901 evaluable population excludes Study 190-901 subjects who enrolled to Study 190-203. Observations at age 0 are omitted. Further requirements correspond to the inclusion/exclusion criteria of Study 190-203 and a follow-up duration requirement to allow reliable estimation of ML rate of decline.

- At least one assessment ML ≥ 3
- At least two ML assessments in the range 1 – 6 and at least 6 months apart

4.2 Treatment Group Presentation

This is a single-arm study. Subjects < 2 years of age received lower than 300 mg dosing according to protocol and then received dosing of 300 mg after their 2nd birthday. Efficacy analyses will be for the full group of subjects, regardless of dose level received. Descriptive summaries of safety data will be provided by age category (separate columns) as well as full group (total column):

- Baseline age < 2 years
- Baseline age < 3 years
- Baseline age ≥ 3 years

4.3 Study Day Derivation

A study day will be obtained by subtracting the initial study drug start date from a visit date plus 1 if the visit date occurs after the initial study drug start date. Otherwise, the study day will be the visit date minus the initial study drug start date. Therefore, Study Day 1 will be the same date as the initial study drug start date.

For Study 190-901, the time of events is presented in terms of the subjects' age, in units of months. The Study 190-901 assessment that is used as a match to a Study 190-203 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.

4.4 Baseline Value

The baseline value of an assessment is defined as the last available measurement prior to the administration of study drug unless otherwise specified. For the matched analyses using Study 190-901, the baseline is the value at the matching assessment.

4.5 Visit Windows for Analysis

Clinical variables assessed at scheduled visits will be summarized by visit. Assessments will be classified to visit windows centered on the target scheduled visit study day. If more than one assessment falls within a window, the closest assessment to the target is selected for analysis. If two assessments are equally closest to the target, then the later assessment is selected. Details for the visit windows definition are provided in Appendix 1.

4.6 Handling of Dropouts and Missing Data

Missing ML assessments for symptomatic patients (baseline ML< 6 points) will not have much impact on the assessment of the primary endpoint since the disease progression on ML score is close to linear, estimation of rate of decline in ML is reliable for follow-up periods of greater than 6 months and all subjects had been followed > 6 months. Impact of missing ML assessments for asymptomatic subjects (baseline ML = 6 points) is also expected to be low since the follow-up of matched Study 190-901 subjects is limited to the length of follow-up of the Study 190-203 subject matched. 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.

Statistical models for time-to-event analyses accommodate variable follow-up arising from staggered accrual or early study discontinuation, using an assumption of missing at random.

The CLN2 assessment schedule for Study 190-901 is not regular and does not generally match CLN2 assessment frequency in Study 190-203. Summaries of CLN2 score over a window grid would have highly variable N which is difficult to interpret. For this reason, interpolation will be used. Summaries of matched Study 190-203 and Study 190-901 CLN2 data will be based on data that has been imputed to the Study 203 visit target, using linear interpolation between Study 190-901 subject visits. Details are provided in Appendix 4.

Partial or missing dates for adverse events and medications will be imputed conservatively by assignment to the earliest and latest possible dates for start and stop dates, respectively.

Some patients only have year of birth recorded as permitted by the site Institutional Review Board. Age in integer months is provided at each visit for Denver assessment. For subjects with incomplete date of birth provided, the date of birth will be imputed by back-calculation of birth date from the date of Denver assessment: (date of Denver) – round ((age in months) x (365.25/12)). For subjects with multiple Denver assessments, the average of the imputed date of births is used.

5 SUBJECT DISPOSITION

The following will be summarized:

- Number of subjects enrolled
- Number of subjects who did not fulfill all inclusion/exclusion criteria
- Number of subjects who received study drug
- Length of time on study

6 DISCONTINUATION AND COMPLETION

For subjects who prematurely discontinued study, the primary reason for discontinuation will be summarized. A similar summary will be provided for subjects who discontinue study drug.

- Number of subjects who completed study
- Number of subjects who completed treatment
- Reason for study discontinuation
- Reason for treatment discontinuation

7 PROTOCOL EXEMPTIONS AND DEVIATIONS

The Study 190-203 trial's guideline for managing protocol deviations defines protocol deviations, including whether they are minor or major. Major protocol deviations will be summarized for the ITT population. A data listing of protocol deviations will be provided as well.

8 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Summaries of the following variables will be performed for the ITT population

- Age (years)
- Age group (<2, 2 - <3, ≥3 years)
- Sex
- Ethnicity
- Race
- Height (cm)
- Weight (kg)
- CLN2 Motor scale score
- CLN2 Language scale score
- CLN2 Vision scale score
- CLN2 Seizure scale score
- CLN2 Motor-Language (ML) score
- CLN2 Motor-Language-Vison-Seizure (MLVS) score
- Age of diagnosis (years)
- Age of first symptoms (years)

9 MEDICAL HISTORY

Medical history will be coded in accordance with the most current version of Medical Dictionary for Regulatory Activities (MedDRA) at the time of coding. Summaries will be by system organ class (SOC) and preferred term (PT).

10 PRIOR AND CONCOMITANT MEDICATIONS/PROCEDURES

Prior and concomitant medications are defined as follows:

- Prior medication – any medication with a stop date prior to the date of the first infusion
- Concomitant medication – any medication delivered on or after the date of the first infusion.
    - Medication has start date on or after date of first infusion
    - Medication is listed as 'continuing'
    - Medication has start date prior to the date of the first infusion and stop date on or after date of first infusion

All medications will be coded using the current version of the World Health Organization Drug (WHO Drug) Dictionary. Prior and concomitant medication use will be separately summarized by Anatomical Therapeutic Chemical (ATC) medication class (Level 4) and preferred name (i.e., generic medication name). A subject reporting the same medication more than once will be counted once when calculating the number and percentage of subject who received the medication.

11 COMPLIANCE

Measures of compliance will be summarized:

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

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

Number (%) of missed infusions

Number (%) of infusions administered more than 3 days outside the target dosing day.

Number (%) of interrupted infusions

12 EXTENT OF EXPOSURE TO STUDY DRUG

Exposure to study drug will be summarized:

- Number of infusions
- Duration of dosing
- Exposure (mg)
- Total exposure over all infusions

Exposure (mg) will be calculated as follows:

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

13 EFFICACY EVALUATIONS

The efficacy analyses will be based on the ITT population. The PP population will be used for sensitivity analyses of the primary endpoint.

There will be two sets of efficacy analyses. The first set will be descriptive summaries of the Study 190-203 subjects at baseline and follow-up visits for all efficacy endpoints. The second set of analyses are for the CLN2 scales M, L, V, S, ML, MLV, and MLVS and based on matching Study 190-203 subjects to Study 190-901 historical controls. These will include descriptive summaries and statistical model results for treatment comparison to natural history. If natural history data are available on siblings with disease, data of sibling pairs will be presented in listings and graphical displays.

The primary efficacy endpoint is the rate of decline in the 0 to 6-point ML score and the primary analysis is based on matching with Study 190-901 subjects. All matched analyses will be based on a single matched set and the algorithm for matching has been prespecified prior to any matched analyses and is described in Appendix 3. The matching will be a 3-1 matching of the Study 190-901 Evaluable population and the Study 190-203 ITT population. The matching criteria are:

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

For all matched analyses, Study 190-901 subjects will be weighted inversely to the number of Study 190-901 subjects matched to a given Study 190-203 subjects. The weights for 3, 2, and 1 Study 190-901 subject(s) matched to a given Study 190-203 subject will be 1/3 ,1/2, and 1 respectively and the weights are normalized to the number of Study 190-901 subjects matched to Study 190-203 subjects.

To demonstrate comparability of the matched groups, summaries of the following variables will be produced for each group, with Study 190-901 subjects weighted according to number of matches.

- Age (years, continuous)
- Sex (Female/Male)
- Genome
- Site / Region
- M, L, ML, MLVS scale scores

- Age of diagnosis (years)
- Age of first symptoms (years)

13.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint is the rate of decline in the 0 to 6-point ML score, measured in units of 48 weeks. Since the endpoint is a rate of decline, as opposed to a rate of change, it is 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.

13.1.1 Primary Analysis Method

The mean rate of decline in treated subjects will be compared to the mean rate of decline in untreated subjects using a 2-sample t-test. The null and alternative hypotheses to be tested are:

H_0: $\mu_{190\text{-}203} = \mu_{190\text{-}901}$

H_1: $\mu_{190\text{-}203} \neq \mu_{190\text{-}901}$

where $\mu_{190\text{-}203}$ represents the population mean rate of decline in the BMN 190-treated population and $\mu_{190\text{-}901}$ represents the population mean rate of decline in Study 901.

Testing will use the two-sample T-test with unequal variance and will be conducted at the two-sided α=0.05 significance level. The T-test will weight Study 190-901 subjects according to the number of matches: weights 1/3, 1/2, 1 for 3, 2, 1 matches respectively. The weights will be normalized to sum to the total number of Study 190-901 subjects matching to Study 190-203 subjects. Treatment effect will be estimated as the difference in the mean rate of decline: Study 190-901 versus Study 190-203 subjects. The primary analysis will be based on the matched ITT population. Details of the slope estimation is given in Appendix 2.

13.1.2 Sensitivity Analysis Method(s)

The rate of ML decline endpoint will also be analyzed as follows:

- Repeat primary analysis (T-test) on the PP population by removing any matched set where the Study 190-203 member is not within the PP population.
- Test difference in rate of ML decline using the Sign test (Binomial) for the ITT and PP populations. The slope for each matched Study 190-203 patient is compared to the average slope of the Study 190-901 subject (s) of the 3-1 matched set.
- Repeat primary analysis on the baseline ML<6 population and baseline ML=6 population

- If not all Study 190-203 subjects have at least one Study 190-901 match, produce summary statistics for the rate of ML decline for the ITT and matched Study 190-203 populations.
- Repeat primary analysis but without the use of weights (equal weighting)

Analyses of other measures of ML change are outlined in the next subsections and are based on the ITT population with matching.

13.1.3 Time to Unreversed 2-point Decline or Score of Zero

Response defined as the absence of an unreversed two-point decline or score of zero in ML score by last assessment relative to baseline, will be analyzed using Kaplan-Meier methods and the Cox proportional hazards model. An unreversed two-point decline is defined 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. An unreversed score of zero is a score of zero that did not subsequently increase to a score greater than zero as of the last recorded observation.

The Cox model will include an indicator variable for study group and no additional covariates since the study groups are matched on the important variables. Kaplan-Meier curves will be produced for the two matched cohorts. Time from baseline to the first unreversed 1-point decline in ML score will also be displayed by Kaplan-Meier method. The Cox model and Kaplan-Meier curves will weight Study 190-901 subjects according to the number of matches: weights 1/3, 1/2, 1 normalized to the number of Study 190-901 subjects matched to Study 190-203 subjects, for 3, 2, 1 matches, respectively.

13.1.4 Time to Unreversed Score of Zero

Time to an unreversed ML score of zero will be analyzed similarly as time to unreversed ML 2-point decline or score of zero.

13.1.5 Summary by Visit

Summary statistics of ML score (continuous variable) and change from baseline at each windowed visit will be produced. A tabulation of ML score (categorical variable) and change from baseline at Weeks 48, 96, 144, and last assessment will also be provided.

For comparison with Study 190-901, to enable a descriptive assessment of ML scores at common time points where all subjects contribute to the descriptive statistics, each subject of the matching cohorts will have their ML scores imputed to those time points using linear interpolation. The time points for interpolation will be in accord with the Study 190-203 schedule of CLN2 assessments. Appendix 4 describes the imputation procedure in greater detail. Summary statistics

of interpolated values and change from baseline by visit will be produced for the matched groups. A graph of mean change from baseline and standard error by visit will be produced.

13.1.6 Subscales M and L

The rate of decline, time-to-event (unreversed 2-point decline or score of zero, unreversed 1-point decline), summary by time point, and categorical change from baseline analyses will be performed for the separate M and L scales. Patients with baseline M=0 (or L=0) will be excluded from the M (or L) rate of decline and time-to-event analyses.

13.1.7 Comparisons with Siblings

If natural history data are available on siblings with disease, ML, M, and L scales of sibling pairs may be graphed.

13.2 Secondary Efficacy Endpoint(s)

13.2.1 Time to Disease Manifestation

For analysis purposes, asymptomatic patients are defined as the patients with baseline MLVS = 12. Subsequent disease manifestation is defined as post-baseline consecutive measurements of M, L, V, or S scores that are less than 3 and at least 22 days apart. The time of disease manifestation is defined as the time of the first of the two measurements demonstrating the deficit. Kaplan-Meir analysis will be performed.

For patients who have not presented with CLN2 disease symptoms prior to initiation of treatment (baseline), onset of disease symptoms captured in AE reporting will be described.

13.2.2 MLV and MLVS

The MLV and MLVS scores will be summarized and interpolated similarly to the ML endpoint (Section 13.1.5). The categorical variable summaries will not be produced.

13.2.3 MRI parameters

Brain atrophy 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 cortical gray matter (in mm^3 and as a percentage of WBV)
- White matter volume (in mm^3 and as a percentage of WBV)

- Whole brain apparent diffusion coefficient (mm^2/s)
- Results will be summarized descriptively over time.

Comparisons with natural history may be produced if MRI data are available. If natural history data are available on siblings with disease, MRI parameters of sibling pairs may be graphed.

13.3 Exploratory Endpoints

13.3.1 PedsQL

The PedsQL™ Generic Core Scales are designed to measure Quality of Life in children and adolescents. The assessments are brief, practical and developmentally appropriate. It will be performed at Baseline visit prior to first infusion, every 12 weeks till study completion / early termination.

13.3.1.1 PedsQL - Parent Report

The four parent reports cover the ages from 1-12 months, 13-24 months, 2-4 years (toddlers) and 5-7 years (young child). It includes questions regarding physical, emotional, and social functioning, with school functioning where applicable.

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:

The Parent Report for various ages is composed of 21 (or 23 items) comprising 4 dimensions:

- Physical Functioning (8 items)
- Emotional Functioning (5 items)
- Social Functioning (5 items)
- School Functioning (5 items for young child or 3 items for toddlers)
- 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) will be provided in a listing.

13.3.1.2 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 Family Impact:

The full PedsQL Family Impact Module is composed of 36 items comprising 8 dimensions:

- Physical Functioning (6 items)
- Emotional Functioning (5 items)
- Social Functioning (4 items)
- Cognitive Functioning (5 items)
- Communication (3 items)
- Daily Activities (3 items)
- Family Relationships (5 items)
- Worry (5 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.

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.

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 total scores) will be provided in a listing.

13.3.1.3 PedsQL – Young Child (5-7 years) Self-reported

The young child ages (5-7 years) self-reported is composed of 23 items comprising 4 dimensions:

- Physical Functioning (8 items)
- Emotional Functioning (5 items)
- Social Functioning (5 items)
- School Functioning (5 items)
- Each item is scored on a 3-point scale: 0 (Not at all), 2 (Sometimes) and 4 (A lot) for the ages 5-7 years child report.

Young child self-report has been removed from the eCRF from 22 October 2018. Data collected in database will be provided in a listing.

13.3.2 CLN2-specific Quality of Life Questionnaire

The CLN2-specific QoL Questionnaire is a disease specific supplement to the PedsQL using the same format and quantitation. It will be performed at Screening, at the Baseline visit prior to first infusion, and every 12 weeks until study completion / early termination.

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 -specific 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 provided in a listing.

13.3.3 Denver II Developmental scale

The Denver II is a revision and update of the Denver Developmental Screening Test. Both tests were designed to monitor the development of infants and preschool-aged children. The tests cover four general functions: personal social (such as smiling), fine motor adaptive (such as grasping and drawing), language (such as combining words), and gross motor (such as walking). It 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 by visit.

Age equivalent performance, number of passes, and number of fails on personal social, fine motor adaptive, language and gross motor will be plotted over time using the age at assessment 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.

13.3.4 Modified Unified Batten Disease Rating Scale Involuntary Movement Scale (mUBDRS-Movement)

The modified Unified Batten Disease Rating Scale Involuntary Movement (mUBDRS-Movement) inventory is a rating scale that measures the type, frequency and severity of 5 common involuntary movements (7 items) associated with CLN2 disease:

- Myoclonus (maximal) (2 items: frequency and severity)
- Dystonia (maximal) (2 items: frequency and severity)
- Dysmetria
- Chorea
- Tics/Stereotypy

Each item of the frequencies of Myoclonus and Dystonia is scored on a 4-point Likert scale: 0=“Not present during exam” or “Absent during exam”, 1=“Present, but rarely occurs”, 2=“Periods of activity with pauses” and 3=“Occurs regularly”.

Each item of the severities of Myoclonus and Dystonia is scored on a 4-point Likert scale: 0="Not present during exam" or "Absent during exam", 1="No impairment of activity", 2="Some impairment of activity" and 3="Often impairment of activity".

The item of Dysmetria is scored on a 4-point Likert scale: 0="Normal finger-to-nose or reach to object", 1="Accurate, mild loss direction/velocity", 2="Occasional miss of target" and 3="Unable to touch/reach target".

Each item of Chorea and Tics/stereotypy is scored on a 4-point Likert scale: 0="Absent during exam", 1="Slight and intermittent", 2="Mild/moderate and intermittent" and 3="Pronounced and common".

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

If more than 50% of the items in the items are missing, the total scores should not be computed.

Total score and change from baseline will be provided in a listing.

13.3.5 Modified Unified Batten Disease Rating Scale Seizure Inventory (mUBDRS-Seizure)

The mUBDRS–Seizure inventory measures the type and frequency of seizures and the presence of seizure complications in CLN2 patients in the prior 3-month interval. The inventory is completed with the aid of caregiver/family member recall in the period between study visits.

This measurement is composed of 10 items as follows:

- Frequencies of primary generalized / atonic / myoclonic / complex partial / simple partial seizures (5 items)
- Presences of seizure complications (5 items): secondary generalization, absence, drops, Prolonged Post-ICTAL Period, and medication change/initiation

Each item of the frequencies is scored on a 4-point Likert scale: 0="None in prior 3 months", 1="1-2 seizures in 3 months", 2="3-5 seizures in 3 months" and 3="> 5 seizures in 3 months".

Each item of the presences is scored on a 2-point Likert scale: 0="None in prior 3 months" and 1="Present in prior 3 months at least once".

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

If more than 50% of the items in the items are missing, the total scores should not be computed.

Total score and change from baseline will be provided in a listing.

13.3.6 Retinal Anatomy using Optical Coherence Tomography (OCT)

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

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

Rate of decline in thickness will be computed for each eye and each parameter based on the baseline and last assessments, scaled in units of 24 weeks. Summaries will be for each of the eyes and the average of the eyes. Thickness parameters and change from baseline will be summarized by visit. Individual patient data on thickness parameters will be plotted. The overall OCT assessment (normal, abnormal) will be summarized by visit.

The Weill Cornell Ophthalmic Scale will be performed on the optic nerve, vasculature, macula, and periphery for each eye. The overall severity scale is based on both eyes and has values ranging from 1 to 5. Overall severity and change from baseline will be summarized by visit. Individual patient severity data will be plotted.

A listing which combines the thickness and overall assessment by OCT, and overall severity scale from Weil Cornell Ophthalmic Scale will be provided.

13.3.7 Ophthalmologic Assessment

The measurements and the overall comparison of ophthalmoscopy to most recent assessment will be provided in a listing.

13.3.8 Visual Acuity

All subjects undergo Preferential Looking Testing (PLT) every 12 weeks. 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 will be provided in a listing.

13.3.9 CLN2 Scale Vision Score

The 0-3 point Vision score (V) will be summarized and interpolated similarly to the ML endpoint (Section 13.1.5).

13.3.10 CLN2 Scale Seizure Score

The 0-3 point Seizure score (S) will be summarized and interpolated similarly to the ML endpoint (Section 13.1.5).

13.3.11 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 per SMQ and the 0-3 point S score will also be plotted on the patient chart.

13.3.12 Infant Toddler Quality of Life Questionnaire (ITQOL)

The Infant Toddler Quality of Life Questionnaire (ITQOL) is to assess levels of health and wellbeing in children aged between 2 months to 5 years. This tool asks parents of pre-school children to reflect on physical and psychosocial domains for development, pain, moods, and impact of child health on parents, including overall health, physical abilities, overall growth and development, discomfort/pain, temperament and moods, general behavior, global behavior, getting along with others, general health perceptions, change in health, parent impact - emotional, parental impact - time, family cohesion,

The assessment of ITQOL has been terminated from the study (refer the protocol amendment dated on 17 December 2018). The transformed raw score for each domain, as well as change from baseline, will be provided in a listing.

13.3.13 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 (1-100).

The assessment of EQ-5D-5L has been terminated from the study (refer the protocol amendment dated on 17 December 2018). The 5 levels of perceived problems will be provided in a listing. The EQ VAS and change from baseline will be provided in a listing as well.

13.3.14 CSF/ Blood Biomarkers

Samples of blood and CSF will be collected prior to infusion and may be used to evaluate biochemical, molecular, cellular, and genetic aspects relevant to CLN2 disease, and to develop the assays used for these evaluations.

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

13.3.15 Electroencephalograms (EEG)

A standard electroencephalogram (EEG) will be recorded with 48 hours before infusion. The EEG data will be listed.

13.4 Examination of Efficacy by Subgroups

Analyses for asymptomatic patients (ML < 6) is described in Section 13.2.1. The primary endpoint analyses will be performed on the baseline ML=6 and baseline ML < 6 subgroups (Section 13.1.2).

ML score and MRI parameters will be summarized by visit for the age groups <2, <3, and ≥3 years, similarly to the overall population analysis (Section 13.1.5 and Section 13.2.3).

14 SAFETY EVALUATIONS

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

Safety variables will be summarized descriptively. No formal inference will be conducted. The summarization will be based on the safety analysis population.

14.1 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.

14.1.1 All Adverse Events

The incidence and number of events for all TEAEs will be summarized by system organ class (SOC), preferred term (PT) and severity (CTCAE grade). For AEs that occurred more than once during the study, the maximum severity will be used.

Adverse events that did not resolve or resolved with sequelae will be listed.

14.1.2 Drug-Related Adverse Events

All TEAEs assessed by investigator as study drug related will be summarized by SOC, PT and severity.

14.1.3 Deaths and Serious Adverse Events

Serious adverse events (SAEs) will be summarized in a similar manner to general adverse events. A listing of AEs for subjects who died and a listing of all SAEs will be provided.

14.1.4 Adverse Events Causing Premature Discontinuation

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

14.1.5 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. The analyses will be repeated for subgroups ML=6 and ML<6 at baseline (pre-symptomatic and symptomatic) separately.

14.1.6 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.

14.1.7 Adverse Events of Special Interest (AESI)

"Adverse events of special interest" (AESIs) are defined in the study protocol (Section 10.1.3) to include:

- Status epilepticus
- Status epilepticus is defined to comprise the following MedDRA preferred term: status epilepticus.
- Hydrocephalus (communicating and non-communicating)
- 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 scale not attributable to other causes
- This will be identified by clinical review.
- Hypersensitivity (see Section 14.1.6)
- Temporally-related events (TRE) (e.g. AEs with onset after initiation of a study drug infusion and within 24 hours of the start or restart of the infusion)

- 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.

14.2 Clinical Laboratory Tests

Clinical laboratory tests include chemistry, hematology, and urinalysis. Clinical laboratory test values and change from Baseline will be summarized descriptively by visit. Laboratory test values will be presented over time using boxplots.

Laboratory tests will also be summarized in terms of results relative to the lab reference ranges (low, high). Shift tables cross-tabulating the normal/abnormal results at Baseline vs. post-baseline visits will be provided as well. A supportive listing of abnormal test values will be produced.

14.3 Vital Signs

Vital signs include systolic blood pressure (BP), diastolic BP, heart rate, respiratory rate, and temperature. Vital signs will be summarized descriptively by visit using the maximum/minimum value for multiple assessments (e.g. during and pre- and post-infusion) at each visit. Vital signs will also be presented every 8 weeks using boxplots.

14.4 Electrocardiogram (ECG)

Quantitative ECG (12-lead) parameters (heart rate, RR, PR, QRS, QTcB, and QTcF) and change from baseline will be summarized by visit. QTcF and QTcB 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 overall interpretation from investigator ECGs will be cross tabulated using a shift table at baseline vs. post-baseline results.

For the first infusion of BMN 190, continuous ECG monitoring (3- or 5-lead) is performed for all subjects. In the event that the subject has already received the first infusion of BMN 190, the next infusion is monitored. If a 12-lead ECG is required during this time, continuous monitoring will be interrupted in order to obtain the 12-lead ECG. The 3- or 5-lead ECG data in Study will be provided in a listing.

14.5 Physical Examinations

Clinically significant abnormalities observed during physical examinations will be summarized by visit and presented in listings.

Body weight, height and body mass index (BMI) will be summarized by visit.

14.6 Pregnancy Test

At Screening and at any time during the study, a female subject judged by the investigator to be of childbearing potential (as defined by onset of menses) will be tested for pregnancy with a urine pregnancy test; additional urine tests will be performed whenever pregnancy is in question. A serum pregnancy test will be performed if a urine test result is positive or equivocal.

The test results will be provided in a listing.

14.7 Neurological Examination

A complete neurological examination includes level of consciousness, speech, language, cranial nerves, motor strength, motor tone, abnormal movements, reflexes, upper extremity sensation, lower extremity sensation, gait, Romberg, nystagmus, and coordination.

The neurological examination results (normal vs abnormal) will be listed only.

14.8 Cerebrospinal Fluid Surveillance (CSF)

Samples of standard clinical laboratory CSF for routine surveillance (cell count with differential, protein, glucose, and culture) will be summarized in a similar manner to clinical lab tests.

15 IMMUNOGENICITY ASSESSMENT

Immunogenicity tests will be performed using validated immunogenicity assays. Routine immunogenicity tests will include total antibody (TAb) and neutralizing antibody (NAb) in the serum and CSF. NAb testing will not be performed if the TAb is negative. 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. Immunogenicity will be summarized by visit as well as by change at study timepoint from baseline. 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 PHARMACOKINETICS AND PHARMACODYNAMICS

The analysis of pharmacokinetics (PK) and pharmacodynamics (PD) will be finalized prior to final database lock as a separate document.

17 APPENDICES

Appendix 1: Windows

Assessment	Derived Visit	Scheduled Visit Day	Window (day)
Device patency/infection, Brief physical examination	Baseline	1	<=1
	Day 6	6	[2, 10]
	Week 3	15	[11, 22]
i=1,2….71	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
CSF (cell count, protein, glucose, and culture),	Baseline	1	<=1
	Day 6	6	[2, 10]
i=1,2….70	Week 3	15	[11, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 145	1016	Day 1003 and later
Vital Signs	Baseline	1	Pre-dose
	Day 1	1	Post-dose
i=1,2….70	Week 3	15	[2, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 145	1016	[1003, 1092]
	Week 169		Day 1093 and later
CLN2 disease rating scale*	Screening		Screening visit
	Baseline	<1	<=1
i=1,2….34	Week 5	29	[2, 43]
	Week 5 + i*4	29 + i*28	[16+i*28, 43+i*28]
	Week 145	1016	[996, 1092]
	Week 169		Day 1093 and later
Modified Unified Batten Disease Rating Scale Involuntary Movement Scale, Modified Unified Batten Disease Rating Scale Seizure Inventory	Baseline	<1	<=1
	Week 13	85	[2, 127]
i=1,2….11	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
PedsQL, Denver II Developmental Scale,	Baseline	<1	<=1
	Week 13	85	[2, 127]

CLN2-specific QoL questionnaire, Visual acuity testing, Neurological examination CSF for immunogenicity ITQOL 4Q-5D-5L i=1,2….10	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
	Week 145	1016	Day 968 and later
Serum for immunogenicity i=1,2….10	Baseline	<1	<=1
	Week 13	85	[2, 127]
	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
	Week 145	1016	[968, 1092]
	Week 169		Day 1093 and later
ECG, 3- or 5-lead	Baseline	1	Pre-dose
	Day 1	1	Post-dose
ECG, 12-lead** i=1,2,…4	Baseline	<1	<1
	Day 1 - 30 minutes	1	1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	[926, 1092]
	Week 169		Day 1093 and later
Cranial MRI i=1,2,3,4	Screening		Screening visit
	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	Day 926 and later
CSF / plasma for PK	Baseline	1	Pre-dose
	Day 1	1	Post-dose
	Week 25	169	[2, 253]
	Week 49	337	[254, 505]
	Week 97	673	Day 506 and later
CSF for disease-related biomarkers	Baseline	1	<=1

i=1,2..33 j=1,2,..8	Week 3	15	[2, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 71	491	[485, 498]
	Week 79	547	[499, 575]
	Week 79 + j*8	15 + j*56	[520+j*56, 575+j*56]
	Week 151	1058	Day 1024 and later
Blood for disease-related biomarkers i=1,2,3,4 j=1,2,3,4,5,6	Baseline	1	Pre-dose
	Week 9	57	[1, 85]
	Week 9 + i*8	57 + i*56	[30+i*56, 85+i*56]
	Week 49	337	[310, 379]
	Week 61	421	[380, 463]
	Week 61 + j*12	421 + j*84	[380+j*84, 463+j*84]
	Week 145	1016	Day 968 and later
EEG, standard awake, Optical coherence tomography[r], i=1,2,3,4	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	Day 926 and later
Height and weight assessments i=1,2,..4	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	[926, 1092]
	Week 169		Day 1093 and later
Ophthalmologic assessments	Baseline	1	<=1
	Week 49	337	[2, 505]
	Week 97	673	[506, 841]
	Week 145	1016	Day 842 and later
Complete physical examination	Baseline	1	<=1
	Week 49	337	[2, 505]
	Week 97	673	[506, 841]
	Week 145	1016	[842, 1092]

	Week 169		Day 1093 and later
Blood/urine for clinical lab tests i=1,2,3 j=1,2,3,…6	Baseline	<=1	<=1
	Day 2	2	2
	Week 5	29	[3, 43]
	Week 9	57	[44, 71]
	Week 17	113	[72, 141]
	Week 17 + i*8	113 + i*56	[86+i*56, 141+i*56]
	Week 49	337	[310, 379]
	Week 61	421	[380, 463]
	Week 61 + j*12	421 + j*84	[380+j*84, 463+j*84]
	Week 145	1016	[968, 1092]
	Week 169		Day 1093 and later

*: only motor, language and motor-language scores will be summarized for screening visit.

**:In patients with present or past bradycardia, conduction disorders, or with structural heart disease, ECG will be performed within 30 minutes before the start of infusion (±5 minutes), at 2 hours (±15 minutes) during infusion, 30 (±5) minutes after infusion end, and 12 hours (±3 hours) after infusion end. In this case, please use the prior infusion assessment on Day 1 as baseline.

6-month Safety follow-up (Week 169) is defined using the window week of 156 to infinite.

Appendix 2: Estimation of Rate of decline for a Subject

The rate of decline on ML is based on selecting a starting and ending ML assessment. For Study 190-203 and Study 190-901, the starting assessment is the baseline ML assessment and the ending assessment is the last ML score > 0. Note that for Study 190-901, the baseline ML assessment is defined as the assessment of matching to the Study 190-203 subject.

The rate of decline is calculated as follows:

1. Determine the slope of the line connecting the two points:

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

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

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

Appendix 3: Matching of Study 190-901 and Study 190-203 Subjects

The population for the primary analysis is based on the ITT population and a 3-1 matching algorithm. This matching algorithm is based on maximizing the number of Study 190-203 subjects matched to Study 190-901 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 190-901 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 190-203 subject are included for the matched analysis. If this derived duration of follow-up for the Study 190-901 subject is not of duration 6 months or greater, then matching at this age of assessment will not be considered. For Study 190-901, 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 190-203 subjects are paired off from first through last and there are potentially 14 matched pairs. No Study 190-901 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 190-203 subject who has not yet been matched and has the least potential Study 190-901 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 190-203 subject is identified, break the tie by considering the number of potential Study 190-901 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 190-203 subject, match with the Study 190-901 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 190-901 subject that satisfies minimal distance, in which case choose the 190-901 subject who has fewest potential Study 190-203 matches based on the distance ≤ 3 points (and ≤ 2, ≤ 1, ≤ 0 as needed). If there are no Study 190-901 subjects who satisfy distance ≤ 3, then there is no match for the Study 190-203 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: onset age, seizure age, exact genome, sex, country, most recent date of birth. For Study 203 patients missing onset (seizure) age, matching 901 patient(s) must have missing onset (seizure) age or onset (seizure) age after age of match.

Repeat till all Study 190-203 subjects have been attempted for match. This completes a 1-1 matching. To attain 3-1 matching, cycle through the Study 190-203 subjects twice more. Study 190-203 subjects will have 3, 2, 1, or 0 Study 190-901 subjects matched. We refer to the group as a matched set. For analysis purposes, Study 190-203 subjects will have a weight of one and Study 190-901 subjects will be weighted inversely to the number of matches (1/3, 1/2, 1) and the weights will be normalized to the number of Study 190-901 subjects matched to Study 190-203 subjects.

Appendix 4: 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 190-203 and 190-901 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 190-901 subjects is truncated to have no longer follow-up than the corresponding imputed data for the Study 190-203 subject matched.

Appendix 5: Medra 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
- Meningitis aseptic
- Meningitis aspergillus
- Meningitis bacterial
- Meningitis borrelia
- Meningitis candida
- Meningitis chemical
- Meningitis coccidioides
- Meningitis coxsackie viral
- Meningitis cronobacter
- Meningitis cryptococcal
- Meningitis echo viral
- Meningitis enterococcal
- Meningitis enteroviral
- Meningitis eosinophilic
- Meningitis exserohilum
- Meningitis fungal
- Meningitis gonococcal
- Meningitis haemophilus
- Meningitis herpes
- Meningitis histoplasma
- Meningitis leptospiral
- Meningitis listeria
- Meningitis meningococcal
- Meningitis mumps
- Meningitis neonatal
- Meningitis noninfective
- Meningitis pneumococcal
- Meningitis salmonella
- Meningitis staphylococcal
- Meningitis streptococcal
- Meningitis toxoplasmal
- Meningitis trypanosomal
- Meningitis tuberculous
- Meningitis viral
- Herpes zoster meningitis
- Herpes simplex meningitis
- Pachymeningitis
- Propionibacterium acnes
- Pseudomonas aeruginosa meningitis

Signature Page for VV-CLIN-003401 v3.0

Approval	PI PI
Approval	PI veillance PI
Approval	PI nt PI PI
Approval	PI PI Clinical Science PI
Approval	PI Representative PI PI
Approval	PI i sentative PI

Signature Page for VV-CLIN-003401 v3.0

Signature Page for VV-CLIN-003401 v3.0

Approval	PI [redacted]

Signature Page for VV-CLIN-003401 v3.0

Statistical Analysis Plan

Study BMN190-203

A Phase 2, Open-Label, Multicenter Study to Evaluate Safety, Tolerability, and Efficacy of Intracerebroventricular BMN 190 in Pediatric Patients < 18 years of age with CLN2 Disease

Version Final Draft

25 June 2020

SIGNATURE AND APPROVAL PAGE

DocuSigned by:
PI
0407169F552349068C56CF4CD8FD9ED6

PI
PI , Biostatistics
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
Signer Name
Signing Reas
Signing Time
PI
CF12A6516C7144E5A6ABD272E4D686EB

PI , MD, PhD
Clinical Science Fellow
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
730FAE52504F4671AEDCC5E64085D1D2

PI
PI , Statistical Programming
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
0113E866001145BCA1EDED4F75F14ADA

PI , MD
PI Drug Safety Surveilla ment
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
BB47E0174A194E089B500C5E2AFD77AD

PI PhD
PI Translational Scien ment
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
C730A310B17242FF8F029CC5EFE93CB8

PI MBBS, PhD
PI Clinical Scien
BioMarin Pharmaceutical Inc.

Date

DocuSigned by:
PI
Signer Name
Signing Reas
Signing Time
PI
027DA7AC705E447B9682C8910CF2CB1D

PI PhD
PI
BioMarin Pharmaceutical Inc.

Date

Note: additional e-signature approvals are on file.

TABLE OF CONTENTS

1 LIST OF ABBREVIATIONS

Abbreviations Definitions

AE	Adverse event
AEoSI	Adverse Events of Special Interest
ATC	Anatomical therapeutic c
BMI	Body mass index
BP	Blood pressure
CI	Confidence interval
CRF	Case report form
CSF	Cerebrospinal fluid surveillance
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events (v4.0)
ECG	Electrocardiogram
EEG	Electroencephalograms
eCRF	Electronic case report form
HAE	Hypersensitivity adverse event
ICV	Intracerebroventricular
ITQOL	Infant toddler quality of life questonnaire
ITT	Intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
ML	Motor-language
MLVS	Motor-language-vison-seizure
MRI	Magnetic resonance imaging
OCT	Optical coherence tomography
PD	Pharmacodynamic
PK	Pharmacokinetic
PP	Per Protocol
PR	Pulse rate
PT	Preferred term
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SE	Standard error
SMQ	Standardized MedDRA query
SOC	System organ class
TEAE	Treatment-emergent adverse event
TLGs	Tables, listings, and graphs
TRE	Temporally-related events
WHO	World Health Organization

2 INTRODUCTION

This document describes the statistical methods to be implemented in the analysis of data collected under clinical study protocol 190-203, "A Phase 2, Open-Label, Multicenter Study to Evaluate Safety, Tolerability, and Efficacy of Intracerebroventricular BMN 190 in Pediatric Patients < 18 years of age with CLN2 Disease" (Amendment 6, 05 February 2019). Analyses of pharmacokinetics data is outlined in a separate document. This statistical analysis plan (SAP) contains definitions of analysis populations, derived variables, and statistical methods for the analyses of efficacy and safety. There will be two sets of efficacy analyses. The first set will be descriptive summaries of the Study 203 patients at baseline and follow-up visits. The second set of analyses are based on matching Study 203 patients to historical controls and includes the primary efficacy analysis. The historical controls are from Study 190-901: DEM-CHILD Multi-Center Clinical NCL Database.

2.2 Objectives of Study

The primary objectives of this study include the following:

- evaluate safety and tolerability of BMN 190 administered via intracerebroventricular (ICV) device
- to evaluate treatment effectiveness as a delay in progression of motor-language (ML) score on the Hamburg CLN2 clinical rating scale
- assess immunogenicity of BMN 190 in CSF and serum
- Secondary objectives of this study include the following:
- to characterize the pharmacokinetics of BMN 190 in CSF and plasma
- measure MRI parameters of disease progression
- assess impact of treatment on the total Hamburg clinical rating scale
- assess the time to disease manifestation for asymptomatic patients
- Exploratory objectives of this study include the following:
- assess development achievement
- assess abnormal involuntary movements
- evaluate retinal anatomy using optical coherence tomography (OCT)
- determine seizure onset, type and frequency
- determine change in seizure activity
- anti-epileptic treatment
- assess quality of life metrics

- assess changes in EEG
- assess changes in visual acuity
- analysis of disease-related biomarkers from CSF and blood

2.3 Study Design

This is a Phase 2 open-label, multicenter study in pediatric patients < 18 years of age with CLN2 disease. The study is designed to assess disease progression in CLN2 patients treated with BMN 190 compared to natural history data from untreated historical controls (data from a CLN2 disease patient registry and data from siblings with CLN2 disease, when available). The comparison of treatment with untreated controls will be compared from the time of the initiation of study drug.

The planned enrollment for this study is at least 10 patients: at least 5 patients with a Hamburg ML score ≥ 5 points, at least 5 patients with a ML score < 5 points, and at least 5 patients < 2 years of age. Study eligibility and enrollment will occur before (≤ 21 days) subjects are admitted to the hospital for surgical implantation of an ICV access device. The first infusion will occur at least 14 days from surgery to implant the ICV reservoir and no more than 28 days after surgery. The Baseline visit to collect clinical scores and clinical laboratory parameters will be completed no more than 2 days before the first infusion.

BMN 190 will be administered every 14 days from the date of the first infusion (±3 days) for at least 144 weeks according to the patient's age (see Section 2.5). At study completion or early termination visit, patients who have ICV device removal are required to have a device safety follow-up visit at 4 weeks (±3 days) after ICV device removal. All patients require a safety follow-up visit at 6 months (±1 week) after last dose.

CLN2 assessments are scheduled every 4 weeks, study completion/early termination, and safety follow-up visits. Cranial MRI is assessed every 24 weeks and at study completion/early termination visit. Immunogenicity assessment is based on CSF and serum lab draws collected at baseline, every 12 weeks, study completion/early termination, and safety follow-up visits (serum only). PK assessment is based on CSF and serum lab draws every 24 weeks and at study completion/early termination visit.

2.4 Study Population

This is a study in pediatric patients < 18 years of age with CLN2 disease, confirmed by deficiency of TPP1 enzyme activity and mutation of the CLN2 gene. Key inclusion and exclusion criteria:

Inclusion Criteria

Clinical assessment of the Hamburg motor-language aggregate score 3-6 at Screening

Exclusion Criteria

Another inherited neurologic disease, e.g., other forms of CLN or seizures unrelated to CLN2 disease (patients with febrile seizures may be eligible)

Another neurological illness that may have caused cognitive decline (e.g., trauma, meningitis, hemorrhage) or interfere with disease rating (autism) before Screening

Has received stem cell, gene therapy, or ERT

Episode of generalized motor status epilepticus within 4 weeks before the First Dose visit

Severe infection (e.g., pneumonia, pyelonephritis, or meningitis) within 4 weeks before the First Dose visit (enrollment may be postponed)

Has received any investigational medication within 30 days before the first infusion of study drug or is scheduled to receive any investigational drug other than BMN 190 during the course of the study

2.5 Study Dosage and Administration

BMN 190 will be administered every 14 days from the date of the first infusion (±3 days) for at least 144 weeks according to the patient's age:

- birth to < 6 months: 100 mg
- 6 months to < 1 year: 150 mg
- 1 year to < 2 years: 200 mg (first 4 doses), 300 mg (subsequent doses)
- ≥2 years: 300 mg

2.6 Sample Size Determination

Sample size was determined on the basis of clinical judgment. Statistical power was not a consideration.

2.7 Blinding and Randomization Methods

This is a single arm open label study and there is no blinding or randomization.

2.7.1 Interim Analysis

No formal interim analyses are planned. An interim clinical study report based on visits up to April 26, 2020 will be produced at the request of the European Medicines Agency Paediatric Committee. This interim report will include safety and the following efficacy analyses:

- Summary statistics ML rate of decline
- Summary statistics ML rate of decline subset to baseline ML<6
- Summary statistics ML rate of decline subset to baseline ML=6
- Kaplan-Meier time to unreversed ML 2-point decline or score of zero
- Kaplan-Meier time to MLVS<12 for patients with baseline MLVS=12
- ML summary statistics by visit
- M, L, V (Vision), and S (Seizure) score summary statistics by visit
- MLV and MLVS (Motor-Language-Vision-Seizure) summary statistics by visit
- MRI parameter summary statistics by visit
- Immunogenicity
- Adverse events
- Clinical Laboratory Tests
- Vitals signs
- Electrocardiogram (ECG)
- Electroencephalograms (EEG)
- Neurological Examination
- Cerebrospinal Fluid Surveillance (CSF)

The analyses of ML, MLVS, S will also include matched analyses with natural history data.

As a single-arm, open-label study, study data will be available for review on an ongoing basis.

There will be no Data Monitoring Committee. However, BioMarin will convene a Data Review Board of internal and external experts to assist the Medical Monitor in assessing data.

Early stopping will be considered based on clinical judgment; no inferential stopping rules will be employed.

3 GENERAL ANALYSIS CONSIDERATION

Safety and efficacy variables will be summarized descriptively. Descriptive statistics include subject count, mean, median, standard deviation, minimum, and maximum for continuous variables and count and percentage for categorical variables. The 95% confidence interval for mean and percentiles may be included, if appropriate.

3.1 Analysis Populations

3.1.1 Efficacy

The Intent-to-Treat Analysis Population (ITT) is defined as all enrolled subjects who received study drug. Efficacy analyses will be based on the ITT population.

The Per Protocol Population (PP) is defined as all enrolled subjects who receive study drug and satisfy major inclusion and exclusion criteria. All enrolled patients received study drug and one patient (PI) screened with an ML score of 1 in violation of inclusion/exclusion criteria. This patient is not included in the PP population. The PP population will be used for sensitivity analyses of the primary efficacy endpoint.

3.1.2 Safety

The Safety Analysis Population is defined as all subjects who had an ICV reservoir implanted.

3.1.3 Study 901 Evaluable Population

The Study 901 evaluable population excludes Study 901 patients who enrolled to Study 203. Observations at age 0 are omitted. Further requirements correspond to the inclusion/exclusion criteria of Study 203 and a follow-up duration requirement to allow reliable estimation of ML rate of decline.

- At least one assessment ML ≥ 3
- At least two ML assessments in the range 1 – 6 and at least 6 months apart

3.2 Treatment Group Presentation

This is a single arm study. Patients < 2 years of age will receive lower than 300 mg dosing according to protocol and then receive dosing of 300 mg after their 2nd birthday. Efficacy analyses will be for the full group of patients, regardless of dose level received. Descriptive summaries of safety data will be provided by age category (separate columns) as well as full group (total column):

- Baseline age < 2 years
- Baseline age < 3 years

- Baseline age ≥ 3 years

3.3 Study Day Derivation

A study day will be obtained by subtracting the initial study drug start date from a visit date plus 1 if the visit date occurs after the initial study drug start date. Otherwise, the study day will be the visit date minus the initial study drug start date. Therefore, Study Day 1 will be the same date as the initial study drug start date.

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 203 patient 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.

3.4 Baseline Value

The baseline value of an assessment is defined as the last available measurement prior to the administration of study drug unless otherwise specified. For the matched analyses using Study 901, the baseline is the value at the matching assessment.

3.5 Visit Windows for Analysis

Clinical variables assessed at scheduled visits will be summarized by visit. Assessments will be classified to visit windows centered on the target scheduled visit study day. If more than one assessment falls within a window, the closest assessment to the target is selected for analysis. If two assessments are equally closest to the target, then the later assessment is selected. Details for the visit windows definition are provided in Appendix 1.

3.6 Handling of Dropouts and Missing Data

Missing ML assessments for symptomatic patients (baseline ML<6) will not have much impact on the assessment of the primary endpoint since the disease progression on ML score is close to linear, estimation of rate of decline in ML is reliable for follow-up periods of greater than 6 months and all patients had been followed >6 months. Impact of missing ML assessments for asymptomatic patients (baseline ML=6) is also expected to be low since the follow-up of matched Study 901 patients is limited to the length of follow-up of the Study 203 patient matched. 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.

Statistical models for time-to-event analyses accommodate variable follow-up arising from staggered accrual or early study discontinuation, using an assumption of missing at random.

The CLN2 assessment schedule for Study 901 is not regular and does not generally match CLN2 assessment frequency in Study 203. Summaries of CLN2 score over a window grid

would have highly variable N which is difficult to interpret. For this reason, interpolation will be used. Summaries of matched Study 203 and Study 901 CLN2 data will be based on data that has been imputed to the Study 203 visit target, using linear interpolation between Study 901 subject visits. Details are provided in Appendix 4.

Partial or missing dates for adverse events and medications will be imputed conservatively by assignment to the earliest and latest possible dates for start and stop dates respectively.

Some patients only have year of birth recorded as permitted by the site Institutional Review Board. Age in integer months is provided at each visit for Denver assessment. For patients with incomplete date of birth provided, the date of birth will be imputed by back-calculation of birth date from the date of Denver assessment: (date of Denver) – round ((age in months) x (365.25/12)). For patients with multiple Denver assessments, the average of the imputed date of births is used.

4 SUBJECT DISPOSITON

The following will be summarized:

- Number of subjects enrolled
- Number of subjects who did not fulfill all inclusion/exclusion criteria
- Number of subjects who received study drug
- Length of time on study

5 DISCONTINUATION AND COMPLETION

For subjects who prematurely discontinued study, the primary reason for discontinuation will be summarized. A similar summary will be provided for subjects who discontinue study drug.

- Number of subjects who completed study
- Number of subjects who completed treatment
- Reason for study discontinuation
- Reason for treatment discontinuation

6 PROTOCOL EXEMPTIONS AND DEVIATIONS

The Study 203 trial's guideline for managing protocol deviations defines protocol deviations, including whether they are minor or major. Major protocol deviations will be summarized for the ITT population. A data listing of protocol deviations will be provided as well.

7 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Summaries of the following variables will be performed for the ITT population

- Age (year)
- Age group (<2, 2 - <3, ≥3)
- Sex
- Ethnicity
- Race
- Height (cm)
- Weight (kg)
- CLN2 Motor scale score
- CLN2 Language scale score
- CLN2 Vision scale score
- CLN2 Seizure scale score
- CLN2 Motor-Language (ML) score
- CLN2 Motor-Language-Vison-Seizure (MLVS) score
- Age of diagnosis (year)
- Age of first symptoms (year)

8 MEDICAL HISTORY

Medical history will be coded in accordance with the most current version of Medical Dictionary for Regulatory Activities (MedDRA) at the time of coding. Summaries will be by system organ class (SOC) and preferred term (PT).

9 PRIOR AND CONCOMITANT MEDICATION/PROCEDURES

Prior and concomitant medications are defined as follows:

- Prior medication – any medication with a stop date prior to the date of the first infusion
- Concomitant medication – any medication delivered on or after the date of the first infusion.
    - Medication has start date on or after date of first infusion
    - Medication is listed as 'continuing'
    - Medication has start date prior to the date of the first infusion and stop date on or after date of first infusion

All medications will be coded using the current version of the World Health Organization Drug (WHO Drug) Dictionary. Prior and concomitant medication use will be separately summarized by Anatomical Therapeutic Chemical (ATC) medication class (Level 4) and preferred name (i.e., generic medication name). A subject reporting the same medication more than once will be counted once when calculating the number and percentage of subject who received the medication.

10 COMPLIANCE

Measures of compliance will be summarized:

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

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

Number (%) of missed infusions

Number (%) of infusions administered more than 3 days outside the target dosing day.

11 EXTENT OF EXPOSURE TO STUDY DRUG

Exposure to study drug will be summarized:

- Number of infusions
- Duration of dosing (time from first infusion to last infusion, in weeks)
- Exposure (mg)

12 EFFICACY EVALUATIONS

The efficacy analyses will be based on the ITT population. The PP population will be used for sensitivity analyses of the primary endpoint.

There will be two sets of efficacy analyses. The first set will be descriptive summaries of the Study 203 patients at baseline and follow-up visits for all efficacy endpoints. The second set of analyses are for the CLN2 scales M, L, V, S, ML, MLV, and MLVS and based on matching Study 203 patients to Study 901 historical controls. These will include descriptive summaries and statistical model results for treatment comparison to natural history. If natural history data are available on siblings with disease, data of sibling pairs will be presented in listings and graphical displays.

The primary efficacy endpoint is the rate of decline in the 0 to 6-point ML score and the primary analysis is based on matching with Study 901 patients. All matched analyses will be based on a single matched set and the algorithm for matching has been prespecified prior to any matched analyses and is described in Appendix 3. The matching will be a 3-1 matching of the Study 901 Evaluable population and the Study 203 ITT population. The matching criteria are:

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

For all matched analyses, Study 901 patients will be weighted inversely to the number of Study 901 subjects matched to a given Study 203 patient. The weights for 3, 2, and 1 Study 901 patient(s) matched to a given Study 203 patient will be 1/3 ,1/2, and 1 respectively and the weights are normalized to the number of Study 901 patients matched to Study 203 patients.

To demonstrate comparability of the matched groups, summaries of the following variables will be produced for each group, with Study 901 patients weighted according to number of matches.

- Age (years, continuous)
- Sex (Female/Male)
- Genome
- Site / Region
- M, L, ML, MLVS scale scores
- Age of diagnosis (year)
- Age of first symptoms (year)

12.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint is the rate of decline in the 0 to 6-point ML score, measured in units of 48 weeks. Since the endpoint is a rate of decline, as opposed to a rate of change, it is 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.

12.1.1 Primary Analysis Method

The mean rate of decline in treated subjects will be compared to the mean rate of decline in untreated subjects using a 2-sample t-test. The null and alternative hypotheses to be tested are:

$$H_0: \quad \mu_{190\text{-}203} = \mu_{190\text{-}901}$$

$$H_1: \quad \mu_{190\text{-}203} \neq \mu_{190\text{-}901}$$

where $\mu_{190\text{-}203}$ represents the population mean rate of decline in the BMN 190-treated population and $\mu_{190\text{-}901}$ represents the population mean rate of decline in Study 901.

Testing will use the two-sample T-test with unequal variance and will be conducted at the two-sided α=0.05 significance level. The T-test will weight Study 901 patients according to the number of matches: weights 1/3, 1/2, 1 for 3, 2, 1 matches respectively. The weights will be normalized to sum to the total number of Study 901 patients matching to Study 203 patients. Treatment effect will be estimated as the difference in the mean rate of decline: Study 901 versus Study 203 subjects. The primary analysis will be based on the matched ITT population. Details of the slope estimation is given in Appendix 2.

12.1.2 Sensitivity Analysis Method(s)

The rate of ML decline endpoint will also be analyzed as follows:

- Repeat primary analysis (T-test) on the PP population by removing any matched set where the Study 203 member is not within the PP population.
- Test difference in rate of ML decline using the Sign test (Binomial) for the ITT and PP populations. The slope for each matched Study 203 patient is compared to the average slope of the Study 901 patient(s) of the 3-1 matched set.
- Repeat primary analysis on the baseline ML<6 population and baseline ML=6 population.
- If not all Study 203 patients have at least one Study 901 match, produce summary statistics for the rate of ML decline for the ITT and matched Study 203 populations.

Analyses of other measures of ML change are outlined in the next subsections and are based on the ITT population with matching.

12.1.2.2 Time to Unreversed 2-point Decline or Score of Zero

Response defined as the absence of an unreversed two-point decline or score of zero in ML score by last assessment relative to baseline, will be analyzed using Kaplan-Meier methods and the Cox proportional hazards model. An unreversed two-point decline is defined 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. An unreversed score of zero is a score of zero that did not subsequently increase to a score greater than zero as of the last recorded observation.

The Cox model will include an indicator variable for study group and no additional covariates since the study groups are matched on the important variables. Kaplan-Meier curves will be produced for the two matched cohorts. Time from baseline to the first unreversed 1-point decline in ML score will also be displayed by Kaplan-Meier method. The Cox model and Kaplan-Meier curves will weight Study 901 patients according to the number of matches: weights 1/3, 1/2, 1 normalized to the number of Study 901 patients matched to Study 203 patients, for 3, 2, 1 matches respectively.

12.1.2.3 Summary by Visit

Summary statistics of ML score (continuous variable) and change from baseline at each windowed visit will be produced. A tabulation of ML score (categorical variable) and change from baseline at Weeks 48, 96, 144, and last assessment will also be provided.

For comparison with Study 901, to enable a descriptive assessment of ML scores at common time points where all subjects contribute to the descriptive statistics, each subject of the matching cohorts will have their ML scores imputed to those time points using linear interpolation. The time points for interpolation will be in accord with the Study 203 schedule of CLN2 assessments. Appendix 4 describes the imputation procedure in greater detail. Summary statistics of interpolated values and change from baseline by visit will be produced for the matched groups. A graph of mean change from baseline and standard error by visit will be produced.

12.1.2.4 Subscales M and L

The rate of decline, time-to-event (unreversed 2-point decline or score of zero, unreversed 1-point decline), summary by time point, and categorical change from baseline analyses will be performed for the separate M and L scales. Patients with baseline M=0 (or L=0) will be excluded from the M (or L) rate of decline and time-to-event analyses.

12.1.2.5 Comparisons with Siblings

If natural history data are available on siblings with disease, ML, M, and L scales of sibling pairs may be graphed.

12.2 Secondary Efficacy Endpoint(s)

12.2.1 Time to Disease Manifestation

Asymptomatic patients are defined as the patients with baseline MLVS = 12. Subsequent disease manifestation is defined as post-baseline consecutive measurements of M, L, V, or S scores that are less than 3 and at least 22 days apart. The time of disease manifestation is defined as the time of the first of the two measurements demonstrating the deficit. Kaplan-Meir analysis will be performed.

12.2.2 MLV and MLVS

The MLV and MLVS scores will be summarized and interpolated similarly to the ML endpoint (Section 12.1.2.3). The categorical variable summaries will not be produced.

12.2.3 MRI parameters

Brain atrophy 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)

Results will be summarized descriptively over time.

Comparisons with natural history may be produced if MRI data are available. If natural history data are available on siblings with disease, MRI parameters of sibling pairs may be graphed.

12.3 Exploratory Endpoints

12.3.1 PedsQL

The PedsQL™ Generic Core Scales are designed to measure Quality of Life in children and adolescents. The assessments are brief, practical and developmentally appropriate. It will be performed at Baseline visit prior to first infusion, every 12 weeks till study completion / early termination.

12.3.1.1 PedsQL - Parent Report

The four parent reports cover the ages from 1-12 months, 13-24 months, 2-4 years and 5-7 years. It includes questions regarding physical, emotional, and social functioning, with school functioning where applicable.

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:

The Parent Report for various ages is composed of 21 (or 23 items) comprising 4 dimensions:

- Physical Functioning (8 items)
- Emotional Functioning (5 items)
- Social Functioning (5 items)
- School Functioning (3 items or 5 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) will be descriptively summarized.

12.3.1.2 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 Family Impact:

The full PedsQL Family Impact Module is composed of 36 items comprising 8 dimensions:

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

- Daily Activities (3 items)
- Family Relationships (5 items)
- Worry (5 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.

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.

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 total scores) will be descriptively summarized.

12.3.1.3 PedsQL – Young Child (5-7 years) Self-reported

The young child ages (5-7 years) self-reported is composed of 23 items comprising 4 dimensions:

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

Each item is scored on a 3-point scale: 0 (Not at all), 2 (Sometimes) and 4 (A lot) for the ages 5-7 years child report.

Young child self-report has been removed from the eCRF from 22 October 2018. Data collected in database will be provided in listing.

12.3.2 CLN2-specific Quality of Life Questionnaire

The CLN2-specific QoL Questionnaire is a disease specific supplement to the PedsQL using the same format and quantitation. It will be performed at Screening, Baseline visit prior to first infusion, every 12 weeks till study completion / early termination.

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 -specific 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) will be descriptively summarized.

12.3.3 Denver II Developmental scale

The Denver II is a revision and update of the Denver Developmental Screening Test. Both tests were designed to monitor the development of infants and preschool-aged children. The tests cover four general functions: personal social (such as smiling), fine motor adaptive (such as grasping and drawing), language (such as combining words), and gross motor (such as walking). Ages covered by the tests range from birth to 6 years. The test will be performed at Baseline visit prior to first infusion, every 12 weeks till study completion / early termination.

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.

12.3.4 Modified Unified Batten Disease Rating Scale Involuntary Movement Scale (mUBDRS-Movement)

The modified Unified Batten Disease Rating Scale Involuntary Movement (mUBDRS-Movement) inventory is a rating scale that measures the type, frequency and severity of 5 common involuntary movements (7 items) associated with CLN2 disease:

- Myoclonus (maximal) (2 items: frequency and severity)
- Dystonia (maximal) (2 items: frequency and severity)
- Dysmetria
- Chorea
- Tics/Stereotypy

Each item of the frequencies of Myoclonus and Dystonia is scored on a 4-point Likert scale: 0="Not present during exam" or "Absent during exam", 1="Present, but rarely occurs", 2="Periods of activity with pauses" and 3="Occurs regularly".

Each item of the severities of Myoclonus and Dystonia is scored on a 4-point Likert scale: 0="Not present during exam" or "Absent during exam", 1="No impairment of activity", 2="Some impairment of activity" and 3="Often impairment of activity".

The item of Dysmetria is scored on a 4-point Likert scale: 0="Normal finger-to-nose or reach to object", 1="Accurate, mild loss direction/velocity", 2="Occasional miss of target" and 3="Unable to touch/reach target".

Each item of Chorea and Tics/stereotypy is scored on a 4-point Likert scale: 0="Absent during exam", 1="Slight and intermittent", 2="Mild/moderate and intermittent" and 3="Pronounced and common".

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

If more than 50% of the items in the items are missing, the total scores should not be computed.

Total score and change from baseline will be summarized by visit.

12.3.5 Modified Unified Batten Disease Rating Scale Seizure Inventory (mUBDRS-Seizure)

The mUBDRS–Seizure inventory measures the type and frequency of seizures and the presence of seizure complications in CLN2 patients in the prior 3 month interval.
The inventory is completed with the aid of caregiver/family member recall in the period between study visits.

This measurement is composed of 10 items as follows:

- Frequencies of primary generalized / atonic / myoclonic / complex partial / simple partial seizures (5 items)
- Presences of seizure complications (5 items): secondary generalization, absence, drops, Prolonged Post-ICTAL Period, and medication change/initiation

Each item of the frequencies is scored on a 4-point Likert scale: 0="None in prior 3 months", 1="1-2 seizures in 3 months", 2="3-5 seizures in 3 months" and 3="> 5 seizures in 3 months".

Each item of the presences is scored on a 2-point Likert scale: 0="None in prior 3 months" and 1="Present in prior 3 months at least once".

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

If more than 50% of the items in the items are missing, the total scores should not be computed.

Total score and change from baseline will be summarized by visit.

12.3.6 Retinal Anatomy using Optical Coherence Tomography (OCT)

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

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

Rate of decline in thickness will be computed for each eye and each parameter based on the baseline and last assessments, scaled in units of 24 weeks. Summaries will be for each of the eyes and the average of the eyes. Rate of decline will be summarized. Thickness parameters and change from baseline will be summarized by visit. Individual patient data on thickness parameters will be plotted. The overall OCT assessment (normal, abnormal) will be summarized by visit.

The Weill Cornell Ophthalmic Scale will be performed on the optic nerve, vasculature, macula, and periphery for each eye. The overall severity scale is based on both eyes and has values ranging from 1 to 5. Overall severity and change from baseline will be summarized by visit. Individual patient severity data will be plotted.

A listing which combines the thickness and overall assessment by OCT, and overall severity scale from Weil Cornell Ophthalmic Scale will be provided.

12.3.7 Ophthalmologic Assessment

The measurements and the overall comparison of ophthalmoscopy to most recent assessment will be provided in a listing.

12.3.8 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 will be provided in a listing. CLN2 Scale Vision Score

The 0-3 point Vision score (V) will be summarized and interpolated similarly to the ML endpoint (Section 12.3.1.3).

12.3.9 CLN2 Scale Seizure Score

The 0-3 point Seizure score (S) will be summarized and interpolated similarly to the ML endpoint (Section 12.1.2.3).

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 per SMQ and the 0-3 point S score will also be plotted on the patient chart.

12.3.10 Infant Toddler Quality of Life Questionnaire (ITQOL)

The Infant Toddler Quality of Life Questionnaire (ITQOL) is to assess levels of health and wellbeing in children aged between 2 months to 5 years. This tool asks parents of pre-school children to reflect on physical and psychosocial domains such as development, pain, moods, and impact of child health on parents.

The assessment of ITQOL has been terminated from the study (refer to the protocol amendment dated 17 December 2018). The data collected in database will be provided in the listing.

12.3.11 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 has been terminated from the study (refer the protocol amendment dated on 17 December 2018). The data collected in database will be provided in the listing.

12.3.12 CSF/ Blood Biomarkers

Samples of blood and CSF will be collected prior to infusion and may be used to evaluate biochemical, molecular, cellular, and genetic aspects relevant to CLN2 disease, and to develop the assays used for these evaluations.

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

12.4 Examination of Efficacy by Subgroups

Analyses for asymptomatic patients is described in Section 12.2.1. The primary endpoint analyses will be performed on the baseline ML=6 and baseline ML<6 subgroups (Section 12.1.2).

13 SAFETY EVALUATIONS

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

Safety variables will be summarized descriptively. No formal inference will be conducted. The summarization will be based on the safety analysis population which is defined in Section 3.1.

13.1 Adverse Events

Adverse event 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.

13.1.1 All Adverse Events

The incidence and number of events for all TEAEs will be summarized by system organ class (SOC), preferred term (PT) and severity (CTCAE grade). For AEs that occurred more than once during the study, the maximum severity will be used.

Adverse events that did not resolve or resolved with sequelae will be listed.

13.1.2 Drug-Related Adverse Events

All TEAEs assessed by investigator as study drug related will be summarized by SOC, PT and severity.

13.1.3 Deaths and Serious Adverse Events

Serious adverse events (SAEs) will be summarized in a similar manner to general adverse events. A listing of AEs for subjects who died and a listing of all SAEs will be provided.

13.1.4 Adverse Events Causing Premature Discontinuation

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

13.1.5 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.

13.1.6 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.

13.1.7 Adverse Events of Special Interest (AEoSI)

"Adverse events of special interest" (AESIs) are defined in the study protocol (Section 10.1.3) to include:

- Status epilepticus

Status epilepticus is defined to comprise the following MedDRA preferred terms: petit mal epilepsy, status epilepticus.

- Hydrocephalus (communicating and non-communicating)

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 scale not attributable to other causes

This will be identified by clinical review.

- Hypersensitivity (see Section 13.1.6
- Temporally-related events (TRE) (e.g. AEs with onset after initiation of a study drug infusion and within 24 hours of the start or restart of the infusion)
- 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.

13.2 Clinical Laboratory Tests

Clinical laboratory tests include chemistry, hematology, and urinalysis. Clinical laboratory test values and change from Baseline will be summarized descriptively by visit. Lab test values will be presented over time using boxplots.

Lab tests will also be summarized in terms of results relative to the lab reference ranges (low, high). Shift tables cross-tabulating the normal/abnormal results at Baseline vs. post-baseline visits will be provided as well. A supportive listing of abnormal test values will be produced.

13.3 Vital Signs

Vital signs include systolic BP, diastolic BP, heart rate, respiratory rate, and temperature. Vital signs will be summarized descriptively by visit using the maximum/minimum value for multiple assessments (e.g. during and pre- and post-infusion) at each visit. Vital signs will also be presented every 8 weeks using boxplots.

13.4 Electrocardiogram (ECG)

Quantitative ECG (12-lead) parameters (heart rate, RR, PR, QRS, QTcB, and QTcF) and change from baseline will be summarized by visit. QTcF and QTcB 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 overall interpretation from investigator ECGs will be cross tabulated using a shift table at baseline vs. post-baseline results.

13.5 Electroencephalograms (EEG)

Electroencephalograms will be summarized in terms of the number of subjects with epileptiform activity and/or frequency slowing, in combination with the activity's location (focal vs. generalized), at baseline and post-baseline. The number of subjects showing new such activity (defined by the combination of activity and location) relative to baseline will be summarized.

13.6 Physical Examinations

Clinically significant abnormalities observed during physical examinations will be presented in listings. Body weight, height and body mass index (BMI) will be summarized by visit.

13.7 Neurological Examination

The neurological examination will be listed only.

13.8 Cerebrospinal Fluid Surveillance (CSF)

Samples of standard clinical laboratory CSF for routine surveillance (cell count with differential, protein, glucose, and culture) will be summarized in a similar manner to clinical lab tests.

14 IMMUNOGENICITY ASSESSMENT

Immunogenicity tests will be performed using validated immunogenicity assays. Routine immunogenicity tests will include total antibody (TAb) and neutralizing antibody (NAb) in the serum and CSF. NAb testing will not be performed if the TAb is negative. 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. Immunogenicity will be summarized by visit as well as by change at study timepoint from baseline. 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.

15 PHARMACOKINETICS AND PHARMACODYNAMICS

The analysis of pharmacokinetics (PK) and pharmacodynamics (PD) will be finalized prior to final database lock as a separate document.

16 APPENDICES

Appendix 1: Windows

Assessment	Derived Visit	Scheduled Visit Day	Window (day)
Device patency/infection, Brief physical examination i=1,2….71	Baseline	1	<=1
	Day 6	6	[2, 10]
	Week 3	15	[11, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
CSF (cell count, protein, glucose, and culture), i=1,2….70	Baseline	1	<=1
	Day 6	6	[2, 10]
	Week 3	15	[11, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 145	1016	Day 1003 and later
Vital Signs i=1,2….70	Baseline	1	Pre-dose
	Day 1	1	Post-dose
	Week 3	15	[2, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 145	1016	[1003, 1092]
	Week 169		Day 1093 and later
CLN2 disease rating scale*	Screening		Screening visit
	Baseline	<1	<=1
	Week 5	29	[2, 43]
i=1,2….34	Week 5 + i*4	29 + i*28	[16+i*28, 43+i*28]
	Week 145	1016	[996, 1092]
	Week 169		Day 1093 and later
Modified Unified Batten Disease Rating Scale Involuntary Movement Scale, Modified Unified Batten Disease Rating Scale Seizure Inventory i=1,2….11	Baseline	<1	<=1
	Week 13	85	[2, 127]
	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
PedsQL, Denver II Developmental Scale, CLN2-specific QoL questionnaire, Visual acuity testing, Neurological examination	Baseline	<1	<=1
	Week 13	85	[2, 127]
	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
	Week 145	1016	Day 968 and later

Assessment	Derived Visit	Scheduled Visit Day	Window (day)
CSF for immunogenicity i=1,2….10			
Serum for immunogenicity i=1,2….10	Baseline	<1	<=1
	Week 13	85	[2, 127]
	Week 13 + i*12	85 + i*84	[44+i*84, 127+i*84]
	Week 145	1016	[968, 1092]
	Week 169		Day 1093 and later
ECG, 3- or 5-lead	Baseline	1	Pre-dose
	Day 1	1	Post-dose
ECG, 12-lead** i=1,2,…4	Baseline	<1	<1
	Day 1 - 30 minutes	1	1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	[926, 1092]
	Week 169		Day 1093 and later
Cranial MRI i=1,2,3,4	Screening		Screening visit
	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	Day 926 and later
CSF / plasma for PK	Baseline	1	Pre-dose
	Day 1	1	Post-dose
	Week 25	169	[2, 253]
	Week 49	337	[254, 505]
	Week 97	673	Day 506 and later
CSF for disease-related biomarkers i=1,2..33 j=1,2,..8	Baseline	1	<=1
	Week 3	15	[2, 22]
	Week 3 + i*2	15 + i*14	[9+i*14, 22+i*14]
	Week 71	491	[485, 498]
	Week 79	547	[499, 575]

Assessment	Derived Visit	Scheduled Visit Day	Window (day)
	Week 79 + j*8	15 + j*56	[520+j*56, 575+j*56]
	Week 151	1058	Day 1024 and later
Blood for disease-related biomarkers i=1,2,3,4 j=1,2,3,4,5,6	Baseline	1	Pre-dose
	Week 9	57	[1, 85]
	Week 9 + i*8	57 + i*56	[30+i*56, 85+i*56]
	Week 49	337	[310, 379]
	Week 61	421	[380, 463]
	Week 61 + j*12	421 + j*84	[380+j*84, 463+j*84]
	Week 145	1016	Day 968 and later
EEG, standard awake, Optical coherence tomographyr, i=1,2,3,4	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	Day 926 and later
Height and weight assessments i=1,2,..4	Baseline	<1	<=1
	Week 25	169	[2, 253]
	Week 25 + i*24	169 + i*168	[86+i*168, 253+i*168]
	Week 145	1016	[926, 1092]
	Week 169		Day 1093 and later
Ophthalmologic assessments	Baseline	1	<=1
	Week 49	337	[2, 505]
	Week 97	673	[506, 841]
	Week 145	1016	Day 842 and later
Complete physical examination	Baseline	1	<=1
	Week 49	337	[2, 505]
	Week 97	673	[506, 841]
	Week 145	1016	[842, 1092]
	Week 169		Day 1093 and later
Blood/urine for clinical lab tests i=1,2,3 j=1,2,3,…6	Baseline	<=1	<=1
	Day 2	2	2
	Week 5	29	[3, 43]
	Week 9	57	[44, 71]

Assessment	Derived Visit	Scheduled Visit Day	Window (day)
	Week 17	113	[72, 141]
	Week 17 + i*8	113 + i*56	[86+i*56, 141+i*56]
	Week 49	337	[310, 379]
	Week 61	421	[380, 463]
	Week 61 + j*12	421 + j*84	[380+j*84, 463+j*84]
	Week 145	1016	[968, 1092]
	Week 169		Day 1093 and later

*: only motor, language and motor-language scores will be summarized for screening visit.

**:In patients with present or past bradycardia, conduction disorders, or with structural heart disease, ECG will be performed within 30 minutes before the start of infusion (±5 minutes), at 2 hours (±15 minutes) during infusion, 30 (±5) minutes after infusion end, and 12 hours (±3 hours) after infusion end. In this case, please use the prior infusion assessment on Day 1 as baseline.

6-month Safety follow-up (Week 169) is defined using the window week of 156 to infinite.

Appendix 2: Estimation of Rate of decline for a Patient

The rate of decline on ML is based on selecting a starting and ending ML assessment. For Study 203 and Study 901 the starting assessment is the baseline ML assessment and the ending assessment is the last ML score > 0. Note that for Study 901, the baseline ML assessment is defined as the assessment of matching to the Study 203 patient.

The rate of decline is calculated as follows:

1. Determine the slope of the line connecting the two points:

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

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

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

Appendix 3: Matching of Study 901 and Study 203 patients

The population for the primary analysis is based on the ITT population and a 3-1 matching algorithm. This matching algorithm is based on maximizing the number of Study 203 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 203 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 203 subjects are paired off from first through last and there are potentially 14 matched pairs. 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 203 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 203 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. If still ties, select to maximize mean distance of Study 901 patients with distance ≤ 3.
- For the selected Study 203 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 203 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 203 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, birthdate.

Repeat till all Study 203 subjects have been attempted for match. This completes a 1-1 matching. To attain 3-1 matching, cycle through the Study 203 subjects twice more. Study 203 patients will have 3, 2, 1, or 0 Study 901 patients matched. We refer to the group as a matched set. For analysis purposes Study 203 patients will have a weight of one and Study 901 patients will be weighted inversely to the number of matches (1/3, 1/2, 1) and the weights will be normalized to the number of Study 901 patients matched to Study 203 patients.

Appendix 4: 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 203 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 203 patient matched.

Appendix 5: Medra 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:

- Meningitis
- Meningitis aseptic
- Meningitis aspergillus
- Meningitis bacterial
- Meningitis borrelia
- Meningitis candida
- Meningitis chemical
- Meningitis coccidioides
- Meningitis coxsackie viral
- Meningitis cronobacter
- Meningitis cryptococcal
- Meningitis echo viral
- Meningitis enterococcal
- Meningitis enteroviral
- Meningitis eosinophilic
- Meningitis exserohilum
- Meningitis fungal
- Meningitis gonococcal
- Meningitis haemophilus
- Meningitis herpes
- Meningitis histoplasma
- Meningitis leptospiral
- Meningitis listeria
- Meningitis meningococcal
- Meningitis mumps
- Meningitis neonatal
- Meningitis noninfective
- Meningitis pneumococcal
- Meningitis salmonella
- Meningitis staphylococcal
- Meningitis streptococcal
- Meningitis toxoplasmal
- Meningitis trypanosomal
- Meningitis tuberculous
- Meningitis viral
- Herpes zoster meningitis
- Herpes simplex meningitis
- Pachymeningitis
- Pseudomonas aeruginosa meningitis