octapharma

Statistical Analysis Plan

Study ID	NGAM-08
Study title	Prospective, Double-Blind, Randomized, Multicenter Phase III Study Evaluating Efficacy and Safety of Three Different Dosages of Newgam in Patients With Chronic Inflammatory Demyelinating Poly(Radiculo)Neuropathy (ProCID study)
Study phase	III

Document author	[redacted], Manager Biometrics, Octapharma
Document version	Final 03
Supersedes	2.0
Effective date	06-Jun-2019
Doc ID	2969

Protocol version this SAP is based upon	Version 04 (dated 08-JUN-2018)

Approved by

Name *Function, affiliation*	*Date and signature*
[redacted] Manager Biometrics, Octapharma	[redacted]
[redacted] International Medical Director, Octapharm	[redacted]
[redacted] Principal Statistician, [redacted]	

Document History

Version	Date	Author	Description
Final v1	17-Dec-2015	[redacted]	New document
Final v2	16-Dec-2016	[redacted]	Incorporation of protocol changes based on FDA feedback to protocol version 1. Specifications with respect to the randomization strata.
Final v3	06-Jun-2019	[redacted]	Incorporation of protocol changes (v4) Details added on planned statistical analyses: Primary endpoint: Logistic regression Secondary endpoints: ANCOVA for changes from baseline in I-RODS total score; grip strength; MRC sums score; FSS total score; NCS measurements; QoL scores; ANCOVA models will be used for LSMEANS calculation (+95% CI) in various subgroups Logistic regression for responses based on I-RODS; based on grip strength; based on MRC sum score For INCAT disability total score and PI-NRS 95% confidence intervals for the median are added. Kaplan-Meier analysis for time-to response variables Deterioration after Wash-out/during dose evaluation added

Abbreviations

AE	Adverse Event
BMI	Body Mass Index
CI	Confidence Interval
CIDP	Chronic Inflammatory Demyelinating Poly(Radiculo)Neuropathy
CSR	Clinical Study Report
DB	Database
ECG	Electrocardiogram
(e)CRF	(Electronic) Case Report Form
EFNS	European Federation of Neurological Societies
FAS	Full Analysis Set
FDA	Food and Drug Administration

ITT	Intention-To-Treat
I-RODS	Inflammatory Rasch-built Overall Disability Scale
IWRS	Interactive Web Response System
MADSAM	Acquired Demyelinating Sensory And Motor Neuropathy
MCS	Mental Composite Score
MedDRA	Medical Dictionary for Regulatory Activities
MRC	Medical Research Council
NAT	Nucleic Acid Test
NCS	Nerve Conduction Studies
PCS	Physical Composite Score
PNS	Peripheral Nerve Society

FSS	Fatigue Severity Scale
GCP	Good Clinical Practice
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICH	International Conference on Harmonisation
IgG	Immunoglobulin G
IGIV	Intravenous Immunoglobulin
IMP	Investigational Medicinal Product
INCAT	Inflammatory Neuropathy Cause and Treatment

PP	Per Protocol
PT	Preferred Term
QoL	Quality of Life
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event

Table of Contents

1. Preface

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for Octapharma protocol NGAM-08 (the *ProCID* study):

Prospective, Double-Blind, Randomized, Multicenter Phase III Study Evaluating Efficacy and Safety of Three Different Dosages of Newgam in Patients With Chronic Inflammatory Demyelinating Poly(Radiculo)Neuropathy

This phase III study is conducted to assess efficacy and safety of three different dosages of Newgam in adult patients with a documented diagnosis of CIDP, who fulfill the definite or probable EFNS/PNS criteria for CIDP, and who have an adjusted INCAT disability score between 2 and 9.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the FDA and International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials[1].

The following documents were reviewed in preparation of this SAP:

- Clinical Study Protocol NGAM-08, Version 04

The reader of this SAP is encouraged to also read the clinical protocol for details on the planned conduct of this study. Operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analyses.

[1] International Conference on Harmonization. (1998). Guidance on Statistical Principles. ICH Topic E9 (Statistical Principles for Clinical Trials) (p. 37). London: International Conference on Harmonization.

2. Purpose

This SAP outlines all statistical analyses to be performed on data collected in study NGAM-08, and the resulting output that will be compiled to support the completion of the Clinical Study Report (CSR).

The planned analyses identified in this SAP will be included in regulatory submissions and/or future manuscripts. Exploratory analyses not necessarily identified in this SAP may be performed to support the clinical development program. Any post-hoc or unplanned analyses performed that are not identified in this SAP will be clearly identified in the respective CSR.

The statistical output provided to the medical writer of the CSR will closely follow the ICH guideline for industry on topic E3 (Structure and Content of Clinical Study Reports[2]) to facilitate the subsequent compilation of the CSR.

This statistical output will consist of tables, figures and listings, including

- Tables, figures and listings used or referenced in, or appended to the CSR as detailed in the remainder of this SAP (section 14 of the CSR)
    - Demographic data summary figures and tables
    - Efficacy data summary figures and tables
    - Safety data summary figures and tables
- Listings provided as appendices to the CSR
    - Patient data listings (section 16.2 of the CSR)
    - Individual patient data listings (section 16.4 of the CSR) will be covered by inclusion of SAS datasets into the electronic submission to the authorities

A detailed list of all tables, figures and listings will be supplied in a separate document later when all feedback from authorities will be available.

[2] International Conference on Harmonization. (1996). Structure and Content of Clinical Study Reports. Structure and Content of Clinical Study Reports (Guideline for Industry) (S. 37). London: International Conference on Harmonization.

3. Study Objectives and Endpoints

3.1. Study Objectives

3.1.1. Primary Objective

- To provide confirmatory data on the effect of 1.0 g/kg NewGam every three weeks in patients with active CIDP based on the percentage of responders at Week 24, which should corroborate the existing evidence on efficacy of IGIV in CIDP as known from published literature.

3.1.2. Secondary Objectives

Secondary objectives are

- To assess the effect of 0.5 g/kg and 2.0 g/kg NewGam every three weeks in patients with active CIDP based on the percentage of responders at Week 24 compared to patients on 1.0 g/kg NewGam every three weeks
- To evaluate the safety of NewGam administration using various dosages in patients with CIDP
- To further evaluate the beneficial effect of three NewGam dosages in patients with CIDP by assessing different parameters/scores/scales

3.1.3. Exploratory Objectives

- To assess the primary and secondary objectives at 3 weeks after having provided rescue medication (if applicable)
- To further evaluate the beneficial effect of NewGam administration in patients with CIDP by additional assessments/scores including quality of life (QoL) measures

3.2. Study Endpoints (Target Variables)

3.2.1. Primary Target Variables

The primary (efficacy) endpoint is the proportion of responders in the 1.0 g/kg NewGam arm at Week 24.

A responder is defined as a patient with a decrease of at least 1 point on the adjusted INCAT disability score relative to baseline (week 0).

The adjusted INCAT (Inflammatory Neuropathy Cause and Treatment) disability score is a measure of activity limitation ranging from 0 (healthy) to 10 (unable to make any purposeful movements with arms and/or legs), and is frequently used as a primary endpoint in CIDP trials. Participants in study NGAM-08 are required to have an adjusted INCAT disability score between 2 and 9 at study entry.

The proportion of responders will be calculated as (# of responders) / (# of patients in analysis set) and presented together with the two-sided 95% Wilson-score confidence interval (CI).

In case the primary endpoint can't be evaluated for a patient due to missing data, this patient will be considered as non-responder in the intent-to-treat (ITT) analysis based on the full analysis set (FAS), whereas the same patient would be excluded from the per-protocol (PP) analysis.

3.2.2. Secondary and Exploratory Target Variables

- Efficacy
    - Proportion of responders in the 0.5 g/kg and 2.0 g/kg NewGam arms at Week 24 relative to baseline at Week 0 compared to the 1.0 g/kg arm, based on the adjusted INCAT disability score
    - Proportion of responders in the 0.5 g/kg and 2.0 g/kg NewGam arms at Week 24 relative to baseline at Week 0 compared to the 1.0 g/kg arm, based on the grip strength (Martin Vigorimeter) using the previously published minimum clinically important difference (MCID) cut-off of 8 kPa
    - Proportion of responders in the 0.5 g/kg and 2.0 g/kg NewGam arms at Week 24 relative to baseline at Week 0 compared to the 1.0 g/kg arm, based on the I-RODS scores using the MCID concept related to the varying standard errors (MCID-SE) as recently demonstrated (please refer to protocol)
    - Time to first confirmed worsening on the adjusted INCAT disability scale by at least 1 point or more from the value at Baseline
    - Mean change from Baseline (Week 0) to Termination Visit in
        - grip strength of both hands (assessed by Martin Vigorimeter)
        - Inflammatory Rasch-built overall disability sum score (I-RODS using the concept of MCID-SE as recently reported; Appendix 1) and number of improvers
        - sum of the distal evoked amplitude of 4 right sided and 4 left sided motor nerves (peroneal, tibial, ulnar and median)
        - Pain Intensity Numeric Rating Scale (PI-NRS)
    - Time to first confirmed worsening on the I-RODS scale
    - Time to 1 point decrease (improvement of disability) in adjusted INCAT disability score
    - Time to decrease in I-RODS scores
    - Mean change from baseline (Week 0) to Termination Visit in:
        - modified Fatigue Severity Scale (FSS)
        - SF-36 Health Survey physical composite score (PCS), mental composite score (MCS) and their 8 health domains
        - Additional nerve conduction studies (NCS) analyses (e.g. individual nerve analysis)
    - For MRC the following will be done:
        - Time to decrease in MRC sum score to or below baseline value after temporary improvement (increase)
        - mean change from baseline (Week 0) to Week 12 and to Termination Visit
        - number of improvers by at least 4 points from baseline (Week 0) to Week 12 and to Week 24

- Safety:
    - Throughout the entire 24-week Dose-evaluation Phase:
    - Occurrence of all adverse events (AEs)
    - Short term tolerance parameters including vital signs
    - Physical/neurological examination

- Laboratory parameters: Serum IgG, hematology, clinical chemistry, urinalysis, pregnancy test, and tests for viral safety

4. Study Methods

4.1. Overall Study Design and Plan

Study NGAM-08 is designed as a prospective, double-blind, randomized, multicenter phase III study. The rationale for conducting this study is to confirm the published clinical results obtained with the 1.0 g/kg dosing in order to apply for market authorization of NewGam, and to evaluate different dosing options on both sides of the standard dose with respect to efficacy and safety.

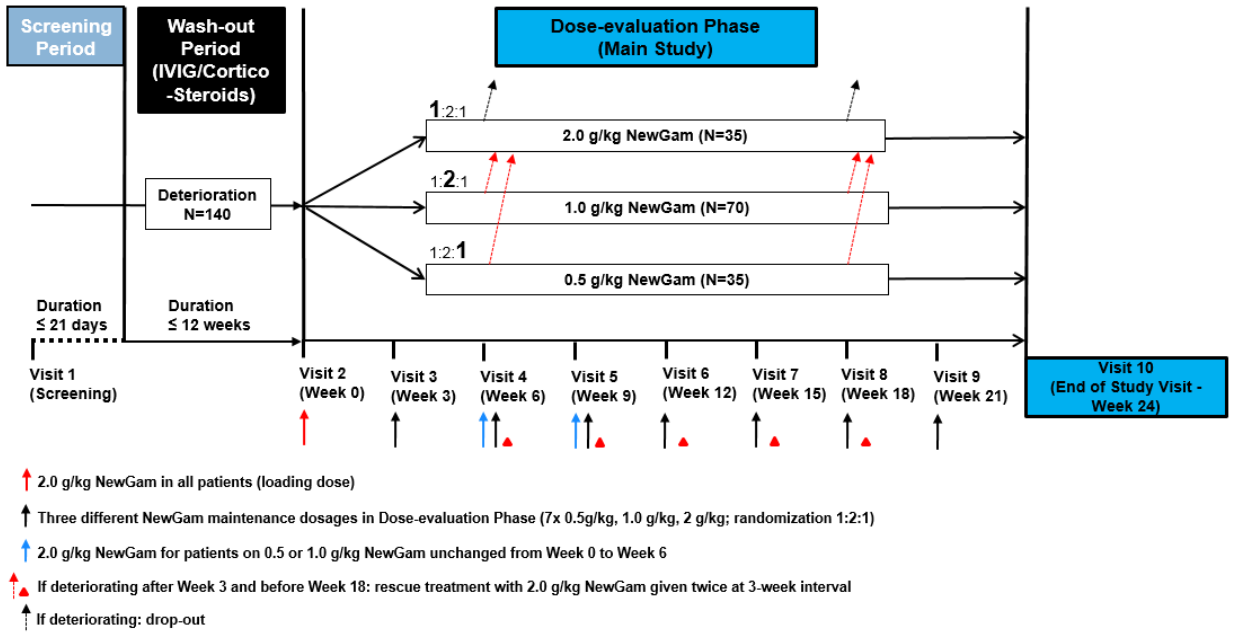
Patients will be screened for eligibility and, if eligible, thereafter have to decrease their current medication (Immunoglobulins by 25% from infusion to infusion or corticosteroids as per the discretion of the Investigator at a rate to expect study entry within 6-12 weeks) for a maximum of 12 weeks (Wash-out Phase). Only patients who deteriorate in this phase, i.e. active patients who need treatment, will be randomized 1:2:1 between three different dosages of NewGam (0.5 g/kg, 1.0 g/kg or 2.0 g/kg NewGam), and enter the Dose-Evaluation Phase.

The two allowed types of treatment prior to enrollment – Immunoglobulins or Corticosteroids – will be considered as strata in a stratified block design randomization with fixed block size. An electronic IWRS tool will be used to realize the randomization ratio of 1:2:1 with respect to the dosage groups, in each of the strata independently.

During the 24-week Dose-evaluation Phase all patients will first receive an initial loading dose of 2.0 g/kg NewGam, followed by 7 maintenance dosages of 0.5 g/kg, 1.0 g/kg or 2.0 g/kg NewGam every 3 weeks as they were randomized to.

Patients not deteriorating in the 12-week Wash-out Phase will not be enrolled into the ProCID study as they are not having active CIDP.

In addition the following rescue treatment option will be available: two consecutive infusions of 2.0 g/kg NewGam at 3-week intervals for all patients in the 0.5 and 1.0 g/kg NewGam arms who are either stable at Week 6 or deteriorate after Week 3 and before Week 18. Patients in the 2.0 g/kg NewGam arm will drop-out if they are stable at Week 6 or deteriorating after Week 3 and before Week 21.



4.2. Selection of Study Population

The study population consists of patients of both sexes, ≥ 18 years of age, with documented diagnosis of CIDP by a neurologist specialized/experienced in neuromuscular diseases. Eligible patients need to fulfill the definite or probable EFNS/PNS criteria for CIDP, and have to have an adjusted INCAT disability score between 2 and 9 at screening (a score of 2 has to be exclusively from leg disability).

Please refer to the protocol for the complete list of all eligibility criteria.

5. Sequence of Planned Analyses

5.1. Interim Analyses

No interim analysis is planned in this study.

5.2. Final Analyses and Reporting

Once the last patient has completed the study, data validation will be completed and the database will be locked according to the applicable standard operating procedures. This process includes (blinded) data review, medical coding and review, the identification and classification of any protocol deviations as detailed in section 7, and thus the patient disposition with respect to the analysis populations. After completion of all these procedures and the finalization of the statistical analysis plan, the formal database lock can be conducted. Only after that the blind will be broken, and the individual treatment assignment will be added to the clinical database for analysis.

All final, planned, analyses identified in the protocol and in the SAP will then be performed.

Key statistics and study results will be made available to the study team following database lock and prior to completion of the final CSR by means of tables, figures and listings.

Any, post-hoc, exploratory analyses completed to support planned study analyses, which were not identified in the final SAP, will be documented and reported in appendices to the CSR. Any results from these unplanned analyses will also be clearly identified in the text of the CSR.

6. Sample Size Determination

The sample size calculation is based on the proportion of responders in the 1.0 g/kg dose group. This dose regimen is the currently most commonly used for maintenance IGIV doses in CIDP. Recent experiences with this dosing regimen are available from the ICE study and the PRIMA trial (please refer to the protocol for more details). In both of these studies, the lower confidence limit of the 95% Wilson-Score confidence interval for the proportion of responders (R) was 42% if rounded to the nearest integer, and thus we have also chosen 42% as the threshold for the following pair of hypotheses for evaluation of the primary endpoint:

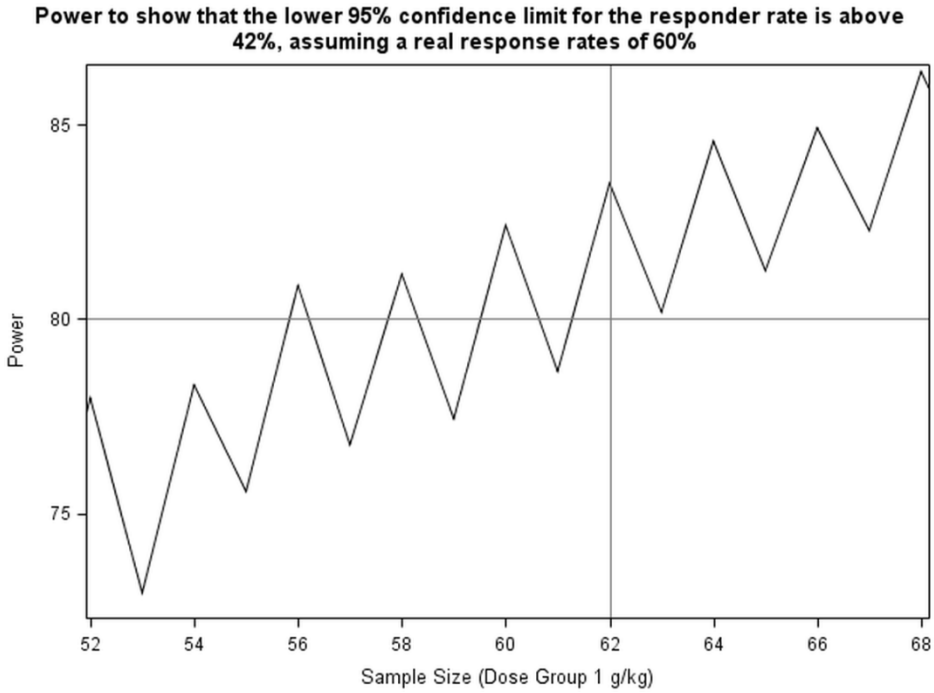
$$H_0: R < 0.42$$

$$H_1: R \geq 0.42$$

This will be tested by comparing the lower limit of the 95% Wilson-Score confidence interval for the observed proportion of responders with the pre-defined threshold of 0.42.

We estimate the true percentage of responders in CIDP patients treated with NewGam as 60%. This corresponds to the 60.7% of responders observed in the PRIMA study, but is higher than the outcome in the ICE (54.2%), that in the opinion of many experts underestimates the true percentage of responders.

We applied these parameters to a computer simulation realized in SAS to determine the statistical power associated with various group sizes:



To achieve a power of at least 80%, a minimum of 62 evaluable patients in the 1 g/kg dose group will be needed; to account for possible dropouts we plan to enroll 70 patients into this group.

In order to allow for the comparison between dose groups, we plan to enroll half the same number of patients also into the other groups (0.5 g/kg and 2.0 g/kg), resulting in a total of 124 evaluable patients, and an enrollment target of 140 patients overall.

6.1. Patient Replacement Policy

Only in case of a withdrawal portion in excess of 10%, patients will be replaced, unless the planned number of patients in the respective dose group has been achieved already. A replacement patient will be assigned to the same treatment group as the withdrawn patient.

Patients withdrawn from the study because of safety reasons will not be replaced.

6.2. Premature Termination of the Study

If early termination of the study becomes necessary, the sponsor and the investigator will ensure that adequate consideration is given to the protection of the patients' interests. Premature termination will be notified in accordance with applicable regulatory requirements. Early termination might concern the study as a whole or individual centers. In any case the blind will be maintained until all database locking procedures are completed and the database has been locked; this includes finalization of the statistical analysis plan, which will be adapted to cover any issues arising from the premature termination.

7. Analysis Populations

The following populations will be considered for the statistical analysis:

The **safety set (SAF)** will include all randomized patients who received at least part of one infusion of IMP.

The **full analysis set (FAS)** is defined according to the intention-to-treat principle, and will include all patients of the SAF for whom any data was collected post infusion of IMP. Every treated subject will be considered in the analysis according to his randomized treatment/dose assignment.

The **per-protocol set (PPS)** – being a subset of the FAS — will exclude patients with significant protocol deviations, which may have an impact on the evaluation of the primary study outcome parameter(s). This is the set of patients who participated in the study as intended and for whom the primary efficacy endpoint can be evaluated as planned.

All protocol deviations documented during the conduct of the study or identified at the data review process prior to DB lock will be reviewed and classified as minor or major and with respect to its effect on the planned analysis. Only major protocol deviations with the potential to significantly affect the study results or to invalidate the interpretation of the data obtained will lead to exclusion of patients from the PP set. This classification of protocol deviations is the joint responsibility of the clinical study manager, the study statistician, and Octapharma's responsible medical expert, and will be performed and documented before the database is locked and the statistical analyses are performed.

The analysis of safety will be based on the safety set.

The evaluation of primary objective(s) will be performed for the FAS (intention-to-treat analysis) and for the PPS (per-protocol analysis). The ITT analysis is considered to be the most relevant, and will be presented first in the statistical report.

The evaluation of secondary objectives will be performed for the FAS, and in addition for the PPS in case the two populations differ by more than 5%.

7.1. Subgroups

In addition to combined presentations over all strata, evaluation of the primary endpoint will be presented for each randomization stratum (prior treatment with Immunoglobulins or Corticosteroids) separately to facilitate the identification of noteworthy differences in the clinical responses between these two groups.

The different CIDP variants eligible for enrollment, such as Multifocal Acquired Demyelinating Sensory And Motor Neuropathy (MADSAM) or pure motor CIDP, will be compared with respect to dose group and proportion of response as well.

Additional analyses, such as logistic regression or ANCOVA, will be deployed for primary and secondary endpoints to further examine differences between subgroups in an exploratory manner. The ANCOVA models will also be used for least square means calculation including 95% confidence intervals in various subgroups (see section 10.2).

8. General Issues for Statistical Analysis

Descriptive summaries will be presented for each of the primary, secondary and exploratory variables. In general, summaries will be completed for all patients overall and by dose group, and the data will be presented by means of standard summary statistics according to the different types of data:

- Binary data (whether or not an event has occurred): counts and proportions
- Count data (the frequency of an event in a set time period): rate (count per unit time)
- Continuous data (measurements on a continuous scale, including quasi-continuous variables): arithmetic mean, standard deviation, median, minimum, maximum
- Scales data (Ordinal and Non-ordinal): absolute and relative frequencies
- Time-to-event data (how long it takes to observe the outcome of interest: time to event or last evaluation (censored data in case patients are lost to follow-up) and event rate). These parameters might not be tabulated separately, but can be included as an inset into a Kaplan-Meier plot of the product-limit survival function estimates.

Additional descriptive and exploratory statistics, such as geometric means, quartiles or confidence intervals, are included as appropriate. If not mentioned otherwise, confidence intervals are to be understood as two-sided, 95% confidence intervals.

In general the denominator for calculating percentages will be the total number of patients in the analysis population unless otherwise specified.

8.1. Analysis Software

Statistical analyses will be performed using SAS Software version 9.1 or later.

8.2. Withdrawals

Patients who withdraw from the study prematurely will be considered in all data presentations for which they contribute data; in particular patients who fail to perform an INCAT evaluation at week 24 will be considered to be non-responders in the analysis of the primary endpoint in the FAS.

8.3. Handling of Missing Data

In general, missing data will not be imputed, and all data derivations will be based on observed values only; exceptions are:

In case of missing body weight, the last available weight measurement will be used for calculating the dose per kg bodyweight (last observation carried forward (LOCF)).

Because the overall disability sum score (I-RODS) cannot be calculated if not all activities are evaluated, missing items will be replaced by the latest available score of the same activity (LOCF).

If missing values occur in the confirmatory analysis of the primary endpoint in the FAS, they will be imputed by worst observed values, i.e. the patient concerned will be analyzed as non-responder.
No analyses of the patterns of missing data will be done.

For adverse events the following will be applied:

An Adverse Event (AE) is defined as treatment-emergent, if first onset or worsening is after start of the first infusion of *NewGam*.

If the start date and time of an AE are partially or completely missing, the AE will be assumed to be treatment-emergent if it cannot be definitely shown that the AE did not occur or worsen after start of the first infusion of *NewGam*. Missing dates and times will not be replaced.

For medications the following will be applied: A medication will be assumed to be concomitant if it cannot be definitely shown that the medication was not administered after start of the first infusion of *NewGam*. Missing dates will not be replaced.

8.4. Derived and Computed Variables

The following derived and computed variables have been initially identified as important for the analysis of the primary and secondary target variables. It is expected that additional variables will be required. The SAP will not be amended for additional variables that are not related to the primary target or key secondary target variables. Any additional derived or computed variables will be identified and documented in the SAS programs that create the analysis files. If the SAP is not amended, further derivations related to primary and secondary target variables will be described in the CSR.

- **Body Mass Index**: BMI = (Body weight) / Height^2 [Unit: kg/m^2]
- Response based on the adjusted INCAT disability score:

  Yes, if [Score at week 24] ≤ [Score at Week 0] - 1 and patient did not receive rescue treatment; No otherwise (including patients discontinued before week 24 and patients receiving rescue treatment)

  The same definition will be applied to other visits than week 24 to allow for an analysis of the time to first response. Time to first response will be censored at start date of rescue medication if a patient received rescue medication and did not show a response before. In case a patient did not show any response until termination visit and did not receive rescue medication, time to first response will be censored at termination visit.

- Worsening with respect to INCAT disability score:

  Worsening occurs if a patient score increases to at least 1 point above the Week 0 value.

- Response based on grip strength (Martin Vigorimeter):

  Evaluate both hands. Yes, if [Strength at Week 24] ≥ [Strength at Week 0] + 8kPa in the dominant or the non-dominant hand and patient did not receive rescue treatment; No otherwise (including patients discontinued before week 24 and patients receiving rescue treatment).

- Response/Worsening based on the I-RODS score:

  The individual Minimally Clinically Important Difference (MCID)-related SE (MCID-SE) score is derived for each visit and patient individually, and categorized as follows:

MCID-SE ≥ 1.96	clinically important improvement
0 < MCID-SE < 1.96	clinically unimportant improvement
MCID-SE = 0	no change
-1.96 < MCID-SE < 0	clinically unimportant deterioration
MCID-SE ≤ -1.96	clinically important deterioration

  The individual MCID-SE values are computed through the Rasch person location values (logits) **R** associated with the I-RODS Summed Raw Scores as follows:

[redacted]

Overall patient response is defined as follows:

Yes, if clinically important improvement (i.e. MCID-SE ≥ 1.96) at week 24 and patient did not receive rescue treatment; No otherwise (including patients discontinued before week 24 and patients receiving rescue treatment).

- Response based on the original MRC sum score [at Week 24] :

  Yes, if increase from Week 0 by at least 4 points and patient did not receive rescue treatment; No otherwise (including patients discontinued before week 24 and patients receiving rescue treatment)

  An analogue definition will be applied for response at Week 12.

  Worsening based on the original MRC sum score:

  Decrease to or below baseline value after temporary improvement by at least 4 points. Patients who did not temporarily improve by at least 4 points will be excluded from the analysis of time to first worsening.

- Changes from baseline to termination visit in scales/scores:

  When calculating those changes, values under rescue treatment will not be taken into account. Changes will be calculated to last value before start of rescue treatment.

- MRC sum score:

  Though mentioned in the protocol, a modified version of the MRC sum score will not be calculated, as this modified version could not be established in other studies.

- NCS measurements:

  The following variables will be calculated from NCS measurements:

  - Sum of all distal amplitudes: the sum will be calculated from 4 nerves, both right and left sides, for a total of 8 amplitudes. The distal sites are the ankle (lower limbs, peroneal and tibial nerves) and wrist (upper limbs, ulnar and median nerves).
  - Sum of the proximal amplitudes: this will be a total sum of 6 parameters, as no proximal amplitude was measured for the tibial nerve. Proximal sites are at the fibular head (lower limb peroneal nerve) and elbow (upper limbs).
  - Sum of conduction velocities (6 in total, as no conduction velocity was measured for tibial nerve)
  - Sum of durations – distal locations (8 in total)
  - Sum of areas – distal locations (8 in total)

  For the last three variables, in case the result is NR (No response) in the data set, the worst value in the data set for the corresponding parameter and for the respective nerve will be used. Worst value is the highest value for durations, lowest value for area and lowest value for conduction velocity.

9. Study Subjects and Demographics

9.1. Disposition of Subjects and Withdrawals

All patients enrolled in the study will be accounted for. Descriptive summaries of population data will be provided overall, by randomization stratum, and by dose group; these will include

- The frequency and percent of patients in each analysis population, sex and age distribution
- The disposition of patients (including number of patients enrolled, number of patients randomized, number of patients treated, number of completers)
- study withdrawals by reason of withdrawal

9.2. Protocol Deviations

Protocol deviations will be documented throughout the study, and checked for on complete data for each patient by manual review as well as by electronic validation checks.

The final decision regarding inclusion/exclusion of patients from the analysis sets will be taken based on protocol adherence reports during data review meetings before database lock, data release and analysis, applying the definitions in section 7.

Major protocol deviations will be summarized by type of deviation and by dose group. Individual patients with these protocol deviations will be listed.

9.3. Demographics and Other Baseline Characteristics

Descriptive summaries of the demographic and other baseline characteristics will be completed for the populations specified below, overall and by dose group. Some particular tables will be repeated also by randomization stratum and/or by CIDP variant. Baseline tables include:

- Demographics (SAF, FAS, PP)
  Age, Gender, Race/Ethnicity, Height, Weight, BMI
  These parameters will also be summarized by investigational site
- Medical History (SAF)
  Medical history will be coded with the Medical Dictionary for Regulatory Activities (MedDRA, according to the version specified in the Data Management Plan). Incidences of findings in medical history will be summarized by MedDRA system organ class (SOC) and preferred term (PT)
- Prior and Concomitant Medications (SAF)
  Medications will be coded using the WHO Drug Dictionary (according to the version specified in the Data Management Plan). Incidences of prior and concomitant medications will be summarized by ATC level 2 and ATC level 4
- Screening Physical Examination
- ECG at screening
- Disease status (FAS) at screening:
  Adjusted INCAT disability score, I-RODS disability score, grip strength, MRC sum score, modified FSS (fatigue scale), NCS (electrophysiological recordings)
- Deterioration after the wash-out phase:

Any deterioration after the wash-out phase will be listed. The frequencies of deteriorations after the wash-out phase will be presented by dose group and category (Patients´ Global Impression of Change, adjusted INCAT disability score, grip strength, I-RODS) as well as in total.

9.4. Measurement of Treatment Compliance

The following parameters will be listed and summarized per patient and/or per infusion:

- Body weight
- Actual dose (total and per kg body weight, based on the latest available weight measurement)
- Total dose of *NewGam* administered
- Total number of infusions administered
- Total volume of solution administered
- Infusion times
- Overall amount of product administered (only included in data listings)
- Infusion rates

Deviations from the planned treatment schedule will be summarized by counting the number of infusions that deviate from the scheduled intervals by more than 4 days, and by listing all cases with more than 4 days deviation individually.

10. Efficacy Analysis

The primary objective of this study is to confirm the efficacy of NewGam in patients with CIDP that were observed in published clinical studies by evaluating the proportion of responders in the 1.0 g/kg dose group.

This study will furthermore compare all efficacy endpoints between the different dosing groups.

Formal planned analyses are described below. It may be necessary for additional exploratory analyses to be performed after the results from the planned analysis are completed; in this case full details of additional analyses will be given in the CSR.

The evaluation of the primary endpoint will be performed for the FAS and for the PPS. The evaluation of secondary endpoints will be performed for the FAS, and in addition for the PPS in case the two populations differ by more than 5%. The FAS analysis is considered to be the primary assessment of efficacy, and will be presented first in the report.

10.1. Primary Efficacy Variable Analysis

The primary efficacy variable is the proportion of responders in the 1.0 g/kg NewGam arm at Week 24 (Termination Visit) relative to baseline (Week 0). A responder being defined as a patient with a decrease of at least 1 point on the adjusted INCAT disability score.

Confirmatory statistical testing of the primary endpoint will be done by constructing the 95% Wilson-Score confidence interval for the proportion of responders, and comparing the lower confidence limit with the predefined threshold of 0.42. This corresponds to a one-sided test of the following pair of hypotheses at the α=0.025 confidence level.

$$H_0: R < 0.42$$

$$H_1: R \geq 0.42$$

Where *R* is the proportion of responders i.e. the number of patients in the FAS who are responders according to the abovementioned definition, divided by the total number of patients in the FAS.

This evaluation will also be presented for the PPS to assess the robustness of the primary study endpoint with respect to protocol deviations.

To assess and compare the primary efficacy endpoint also between the different dose groups, and also between the randomization strata (by dose group and over all dose groups) the proportion of responders will be presented side-by-side, including the 95% confidence intervals.

The time to first response and time to first worsening will be examined and presented by means of a Kaplan-Meier analysis.

Additional analyses, such as logistic regression- will be deployed to further examine any noteworthy differencesin an exploratory manner. The logistic regression model for analysis of response will include dose group, randomization stratum and CIDP variant (typical vs. atypical) as factors and baseline score as covariate. Odds ratios will be calculated to compare response rates between the factor levels, these will be presented together with 95% confidence intervals.

Moreover the proportion of responders after having provided rescue medication will be calculated with respect to all patients who received rescue medication and will be presented by dose group and overall, together with 95% confidence intervals.

10.2. Secondary and Exploratory Efficacy Variable Analysis

All efficacy endpoints detailed in section 3.2 will be summarized by means of descriptive statistics, presented graphically as appropriate, and listed in full detail. Tables and graphs will allow for direct comparison of the different dose groups. If data are tabulated by dose group, the presentations will also include a total category, summarizing the data of all subjects in any of the dose groups. Tables for changes from baseline will include 95% confidence intervals of the mean for continuous secondary endpoints (I-RODS total score; grip strength; MRC sum score; FSS total score; NCS measurements; QoL scores) and 95% confidence intervals for the median for ordinal scales (INCAT disability sore, PI-NRS), calculated with PROC univariate option CIQUANTDF.

Similar analyses as for the primary endpoint will be performed for response variables defined based on secondary endpoints (I-RODS, grip strength, MRC sum score). These will include analysis of response rates and 95% confidence intervals, logistic regression models using the same factors as for the primary endpoint, Kaplan-Meier analyses for time to response and time to first worsening and response rates after having provided rescue medication.

Moreover, ANCOVA models will be applied for continuous secondary endpoints (I-RODS total score; grip strength; MRC sums score; FSS total score; NCS measurements; QoL scores). These ANCOVA models for changes from baseline to end of study will include dose group, randomization stratum and CIDP variant as factors and baseline value of the score/measurement as covariate. ANCOVA models will be used for calculation of least square means in various subgroups as well as for calculation of difference in least square means between the factor levels. All least square means will be presented together with 95% confidence intervals.

Any deterioration during the dose evaluation phase will be listed. Moreover, deteriorations during dose evaluation will be presented by dose group and by their effect on subject disposition (discontinuation, rescue medication).

11. Safety and Tolerability Analyses

The analysis of safety assessments in this study will be presented for the SAF.

11.1. Adverse Events

All reported AEs will be coded according to MedDRA.

An Adverse Event (AE) is defined as treatment-emergent, if first onset or worsening is after start of the first infusion of *NewGam*. Only treatment-emergent AEs (TEAE) are accounted for in the analysis.

A TEAE will be considered to be temporally associated with the infusion (or 'infusional') if it starts during or within 72 hours of the end of the infusion.

All treatment-emergent AEs for each patient, including multiple occurrences of the same event, will be listed in full detail, including reported term, MedDRA preferred term and system organ class, onset, duration, time to the AE occurrence from last dose, causality, dosage, severity, seriousness and actions taken.

For each dose group and for the study as a whole the following will be given:

- Total number of TEAEs reported
- Number of temporally associated TEAEs
- Number and percentage of infusions temporally associated with one or more TEAE
- Number of temporally associated TEAEs divided by the total number of infusions

On a per patient basis the incidence of TEAEs will be summarized by system organ class, preferred term and maximum severity. If a patient experiences an adverse event more than once the event at the most severe occurrence will be considered. Patients will be included only once under each system organ class and only once in the overall totals under the most severe event occurrence.

TEAEs will be summarized by strongest relationship to study medication by system organ class and preferred term. If a patient experiences an adverse event more than once the event most related to study medication will be considered. Patients will be included only once under each system organ class and only once in the overall totals under the most related to study medication occurrence.

Incidences of TEAEs (given as the number and percentage of patients) will be summarized overall and for each dose group as follows:

- Any TEAE
- Any serious TEAE
- Any TEAE probably or possibly related to the trial drug
- Any severe TEAE
- Any withdrawal due to TEAE
- Any significant TEAE
- Any TEAE by MedDRA System Organ Class (SOC) and MedDRA Preferred Term (PT)

Significant AEs fulfill at least one of the following criteria:

- any concomitant medication or therapy started due to the AE
- Study treatment was stopped/withdrawn or dose changed

Narratives will be prepared describing each death, each other serious AE and those of the other significant AEs that are judged to be of special interest because of clinical importance .

11.2. Laboratory Safety Tests

The safety laboratory data (Clinical chemistry, hematology, and urinalysis) will be converted to standard units during the data management process. The laboratory data will be listed with suitable flags indicating abnormal values (L=Lower than reference range, H=Higher than reference range) and clinical significance as judged by the investigator whenever available.

Summary statistics for the laboratory values as well as their changes from baseline at each time point will be tabulated

Summary statistics for the laboratory values as well as their changes from baseline at each time will be tabulated for all laboratory parameters.

11.3. Viral Markers

Viral marker samples will be taken before the first NewGam infusion at screening and at the (early) termination visit. All viral marker results will be listed with suitable flags indicating positive results, and shift tables will be used to identify any possible seroconversion during the course of the study.

11.4. Vital Signs

To evaluate short-term tolerance, vital signs will be monitored at all visits, including measurements before, at least once during and after infusion. Vital signs parameters will be summarized by visit and measurement time, using the standard set of summary statistics for both absolute values and changes from baseline. For pulse, respiratory rate, blood pressure and body temperature the baseline value is the pre-infusion measurement, whereas the weight measurements will be compared to the weight at screening.

Fever is defined as a body temperature >37.8 °C orally, >38.2 °C rectally, >38 °C axillary, >38.2 °C tympanic, and has to be documented as an AE.

Height is measured at baseline only, as is the standard ECG.

11.5. Physical Examination

A general physical examination will be performed at the screening visit, and relevant findings will be documented; the physical examination will be repeated at visits 2, 6, and at the end of study.

Changes in results of physical examination will be summarized by body system and visit. Any clinically relevant worsening from the baseline status will be documented as an AE.

12. Reporting Conventions

The following reporting conventions will be adopted for the presentation of study data. These conventions will enhance the review process and help to standardize presentation with common notations.

12.1. General Reporting Conventions

- All tables and data listings will be developed in landscape orientation, unless presented as part of the text in a CSR.
- Figures will in general also be presented in landscape orientation, unless presented as part of the text in a CSR.
- Legends will be used for all figures with more than one variable or item displayed.
- Figures will be in black and white, unless color figures have been identified as useful for discriminating presentation in the figure. Lines in figures should be wide enough to view the line after being photocopied.
- Specialized text styles, such as bolding, italics, borders, shading, superscripted and subscripted text will not be used in tables, figures and data listings unless they add significant value to the table, figure, or data listing.
- Only standard keyboard characters should be used in tables and data listings. Special characters, such as non-printable control characters, printer specific, or font specific characters, will not be used on a table, figure, or data listing. Hexadecimal character representations are allowed (e.g., μ, α, β).
- The ICH numbering convention is to be used for all tables, figures and data listings.
- All footnotes will be left justified and placed at the bottom of a page. Footnotes must be present on the page where they are first referenced. Footnotes should be used sparingly and must add value to the table, figure, or data listing. If more than four footnote lines are planned then a cover page may be used to display footnotes.
- Missing values for both numeric and character variables will be presented as blanks in a table or data listing. A zero (0) may be used if appropriate to identify when the frequency of a variable is not observed.
- All date values will be presented as DDMMMYYYY (e.g., 29AUG2001) format.
- All observed time values will be presented using a 24-hour clock HH:MM format (e.g. 15:26).
- Time durations will be reported in HH:MM notation. The use of decimal notation to present (display) time durations should be avoided (e.g. 0.083h = 5min) unless it is necessary to show the computation of time differences in a table, figure, or data listing, in which case both notations may be used to display the time duration.
- All tables, figures and data listings will have the name of the program, and a date stamp on the bottom of each output.

12.2. Population Summary Conventions

- Population(s) represented on the tables or data listings will be clearly identified in the title as “Population: <name of population>” where <name of population> is any of the analysis population names or abbreviations defined in section 7 (safety analysis set (SAF), full analysis set (FAS or ITT), per-protocol set (PPS).

- Sub-population(s) or special population(s) descriptions will provide sufficient detail to ensure comprehension of the population (e.g., ITT Females, Per-Protocol Males >60 years of age) used for analysis in a table or figure.
- Population sizes may be presented for each treatment or dosing category as totals in the column header as (N=xxx), where appropriate.
- Population sizes shown with summary statistics are the samples sizes (n) of patients with non-missing values.
- All population summaries for continuous variables will include: N, mean, SD, median, minimum and maximum.
- All percentages are rounded and reported to a single decimal point (xx.x%).

13. Tables, Listings and Figures

To be supplied in a separate document later when all feedback from authorities will be available.