STATISTICAL ANALYSIS PLAN	
Study Code	D134BC00001
Edition Number	4
Date	30-Aug-2024

A Phase III, Multicentre, International Study with a Parallel, Randomised, Double-blind, Placebo-controlled, 2 Arm Design to Assess the Efficacy and Safety of Selumetinib in Adult Patients with NF1 who have Symptomatic, Inoperable Plexiform Neurofibromas (KOMET).

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation or Specialised Term	Definition
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
bid	Twice daily
BOR	Best objective response
BSA	Body surface area
cCR	Confirmed complete response
CI	Confidence interval
COA	Clinical outcome assessment
cPR	Confirmed partial response
CR	Complete response
CRF	Case report form
CRO	Contract research organisation
CSP	Clinical study protocol
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DCO	Data cut-off
DOR	Duration of Response
eCDF	Empirical cumulative distribution function
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EOT	End of treatment
EQ-5D-5L	EuroQoL five dimensions, five level health state utility index
EQ-VAS	EuroQol-visual analogue scale
FAS	Full analysis set
gSD	Geometric standard deviation
HRQoL	Health related quality of life

Abbreviation or Specialised Term	Definition
ICF	Informed consent form
ICR	Independent central review
IF	Information fraction
IPD	Important protocol deviation
IRC	Independent review charter
ITT	Intention-To-Treat
KM	Kaplan Meier
LLOQ	Lower limit of quantification
LPD	Last patient dosed
LS	Least-squares
LVEF	Left ventricular ejection fraction
MAR	Missing at random
Max	Maximum
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
Min	Minimum
MMRM	Mixed Model Repeated Measures
MNAR	Missing not at random
MRI	Magnetic resonance imaging
n	Number of observations
NE	Not evaluable
NF1	Neurofibromatosis type 1
OAE	Other significant adverse event
ORR	Objective response rate
PAINS-pNF	PAin INtensity Scale for Plexiform Neurofibroma
PD	Progressive disease
PedsQL	Paediatric Quality of Life Inventory
PFS	Progression-free survival
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	Pharmacokinetic(s)
PlexiQoL	Plexiform Neurofibroma Quality of Life scale
PN	Plexiform neurofibroma

Abbreviation or Specialised Term	Definition
PR	Partial response
PRO	Patient-reported outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
PT	Preferred term
Q1	1st quartile
Q3	3rd quartile
QoL	Quality of life
RDI	Relative dose intensity
REiNS	Response Evaluation in Neurofibromatosis and Schwannomatosis
REML	Restricted maximum likelihood
SAE	Serious adverse event
SAF	Safety analysis set
SAS	Statistical Analysis Software
SD	Stable disease
SoA	Schedule of assessments
SOC	System organ class
TEAE	Treatment-emergent adverse event
TELC	Treatment-emergent laboratory change
TFL	Tables, figures, and listings
TTP	Time to progression
TTPP	Time to chronic target PN pain palliation
TTR	Time to response
ULN	Upper limit of normal
URC	Unblinded review committee
WHO	World Health Organisation

AMENDMENT HISTORY

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
N/A	16Feb2022	Initial approved SAP	N/A	N/A
Wording around the interim analysis not being performed if 2 DCOs are approximately 4 months apart	21Dec2023	Clarification provided in Sections 1.1 and 5	Yes	To align with CSP v4
Change in estimand of 1st key secondary endpoint	21Dec2023	Estimand updated, primary analysis changed from ANCOVA to MMRM, and supplementary analyses updated	Yes	To align with CSP v4
Upgraded PlexiQoL to 2nd key secondary endpoint	21Dec2023	- PlexiQoL elevated from secondary to 2nd key secondary - Change in MTP (Section 3.3.7) - Change in sample size determination (Section 3.4) - Change in estimand	Yes	To align with CSP v4
Added new exploratory objective on the fed state	21Dec2023	New analysis set in Section 3.2.11 (Fed Measurable FAS) and new endpoint and analyses in Section 4.2.23 (related new objective in Section 4.2)	Yes	To align with CSP v4
Analysis set changed for best percentage change from baseline in target PN	21Dec2023	Now on Measurable PN FAS versus before on FAS	N/A	To align with primary endpoint
Definition of analysis periods	21Dec2023	Better definition of analysis periods and treatment-emergent events in Section 3.3.2	N/A	To align with study design
Visit windowing	21Dec2023	Clarification on EOT visit being included in windows	N/A	To align with study design
PN pain medication scoring and change in analysis of PN pain medication	21Dec2023	Clarification on scoring algorithm and re assessment of PN analyses	Yes	To align with CSP v4

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Attributability of the reason for discontinuation during the randomised period	21Dec2023	Added attributability of the reason for discontinuation during the randomised period in Section 3.3.9	N/A	For sensitivity analyses on key secondary endpoints
PRO completion	21Dec2023	Simplification of the analyses	N/A	To simplify interpretability of the results
PRO analyses	21Dec2023	All PROs are now analysed with MMRM	Yes	To align with key secondary Estimands
Analysis of AEs	21Dec2023	Added exposure-adjusted incidence rates for summaries of on-selumetinib TEAEs based on Section 4.6.2.2	Yes	To align with study design
Primary endpoint	21Dec2023	Reassessment of intercurrent events for the ORR	Yes	To align with CSP v4
		Section 4.2.1.1: Sentence corrected for the interpretation of 20% (e.g., increase from nadir not applicable) and swimmer plot details updated.	N/A	Clarification on derivation
		Reduction of number of supplementary analyses	N/A	To simplify interpretability of the results
Duration of response	21Dec2023	Added a few items to derive	N/A	To align with the US label
Time-to-event endpoints (DOR, PFS, TTP, TTR)	21Dec2023	Changed unit from cycles to months	N/A	To align with the US label
Chronic pain palliation and time to chronic pain palliation	21Dec2023	Reduction of number of supplementary analyses	N/A	To simplify interpretability of the results
Spike PN pain intensity	21Dec2023	Added analyses	N/A	To align with study objectives
IP compliance	21Dec2023	Added IP compliance and added more details on how to derive actual exposure and RDI quantities	N/A	To align with company standards

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Multiple imputation technique	21Dec2023	Added method details	N/A	For sensitivity analyses on key secondary endpoints
Added clarification on how to derive ECG value and overall evaluation in case of a triplicate test	21Dec2023	The average of the 3 values and the worse evaluation (ie, abnormal, clinically significant) will be taken	N/A	To align with company standards
Added 3 new sensitivity analyses of the 1st key secondary endpoint	21Dec2023	Summary statistics and primary analysis (MMRM) will be rerun on different subsets of patients who have different numbers of missing scores during the cycle.	N/A	Requested by the FDA for the validation of the PAINS-pNF questionnaire
Added eCDF plot for the 1st key secondary endpoint as a supplementary analysis	21Dec2023	The empirical cumulative distribution function of change from baseline by study intervention arm will be plotted at each cycle during the randomised period for the Pain FAS (Section 4.2.2.1.7)	N/A	Requested by the FDA for the validation of the PAINS-pNF questionnaire
PFS, TTR, and TTP derived from randomisation date	21Dec2023	To align with the start of the randomised period and with company standards, efficacy time-to-events are derived from the randomisation date and not from the date of first dose.	N/A	To align with company standards
Clarification that box plots of interocular pressure should be done by eye	21Dec2023	2 sets of box plots in the same output, one for the left eye and one for the right eye	N/A	Clarification on derivation
Change in the derivation of partial volumes for following DCOs	21Dec2023	In order to keep responses reported at DCO1 unchanged for patients with partial volumes after DCO1 (similar for DCO2)	N/A	To keep consistency with interim and primary analyses
Added PROMIS total score	21Dec2023	To have a total score to helps the interpretation of physical function results	N/A	To have a physical function summary

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Updated PK Analysis Set definition	21Dec2023	To match the protocol	Yes	To match the protocol
Removed EQ5D plots and clog log plot for TTPP	21Dec2023	Not necessary for interpretability of results	N/A	Unnecessary for interpretability of results
Added summaries of PGIS and PGIC	12 Jan 2024	Added shift table and summary statistics of the distribution of score by treatment over time	No	PGIS and PGIC are collected in the study to support the psychometric validation of the PAINS-pNF questionnaire. However, following a request by the clinical and HrQoL teams, these analyses have also been included as additional exploratory clinical outcome assessments for the CSR.
Writing and formatting QC	18 Jan 2024	Writing and formatting QC Checks by Synchrogenix	N/A	As per company standards
Stratification factors	25 Apr 2024	Average baseline chronic target PN pain group missing as stratification factor from the primary analysis for the following endpoints: PlexiQoL (Section 4.2.2.2.5); PII-pNF pain interference total score (Section 4.2.12.5); PROMIS (Section 4.2.13.5); PedsQL (Section 4.2.14.5); EQ-5D-5L (Section 4.2.15.5).	No	Omitted from model in error
Long-term period	25 Apr 2024	Section 3.3.2: Removal of definition of long-term period, as this is not required for analysis purposes	N/A	No long-term period-specific outputs will be produced

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Definition of treatment emergent	25 Apr 2024	Section 3.3.2.4: Updated end of treatment emergent period from 'one day prior to the date of Cycle 13 Day 1' to 'last dose of Cycle 12'	N/A	Clarification on derivation
Definition of TEAEs in the on-selumetinib safety period	25 Apr 2024	Section 3.3.2.4: Clarify definition of identifying TEAEs in the on-selumetinib period as a result of a worsening of AE	Yes	Add clarity for programming purposes
General considerations – imputation of missing data	25 Apr 2024	Section 3.3.5.1: Added more details on rules for imputation of missing AE/concomitant medication start dates when either day is missing or day and month is missing.	N/A	To align with standards and clarification for programming purposes
Multiple testing procedure	25 Apr 2024	Section 3.3.7: Updated to reflect that the information fractions for the key secondary endpoints at DCO1 will be 100%	Yes	Clarification of significance levels to be used
Protocol deviations	25 Apr 2024	Section 4.1.3.1: Extra summary table of IPDs for randomised period only added	N/A	To aid interpretability of results
Distribution of stratification factors	25 Apr 2024	Section 4.1.4.2: Add distribution of baseline chronic target PN pain score stratification factor and the combined distribution of both stratification factors to demography table	N/A	To aid interpretability of results
Disease characteristics	25 Apr 2024	Section 4.1.6.1: Additional variables for reasons for the PN not being operable; presence of non-target PNs (yes/no); and PN symptoms added to summary table	N/A	To aid interpretability of results
Prior/concomitant medications and procedures	25 Apr 2024	Section 4.1.8.1: Removal of outputs related to post-study intervention discontinuation medications	Yes	Data are not collected

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Prior/concomitant medications and procedures	25 Apr 2024	Section 4.1.8.2: Analysis set for table summarising medications during the on-selumetinib period updated to use the extended selumetinib FAS	N/A	To ensure correct denominators are used for percentage calculations
Table of endpoint analysis	02 May 2024	Section 4.2: Updated to reflect that the exploratory analysis for 'Over Duration of Study' was removed for PAINS-pNF chronic Target PN pain intensity score; PAINS-pNF spike Target PN pain intensity score; chronic target PN pain palliation; PII-pNF pain interference total score; pain medication.	N/A	Not required for interpretability of results.
Partial PN volumes	25 Apr 2024	Section 4.2.1.3: Sentence added to reflect that there may also be partial PN volumes for the non-target lesion	N/A	Add clarity for programming purposes
MRI assessments during prolonged study intervention interruption	25 Apr 2024	Section 4.2.1.6: Added clarification that the rules around using MRI volumetric assessments performed during a prolonged interruption of study intervention apply only when a while-on-treatment approach is taken for this intercurrent event	N/A	Add clarity for programming purposes
Primary endpoint	02 May 2024	Section 4.2.1.9: Clarification of data to use for primary analysis at DCO2 if interim analysis (DCO1) takes place.	N/A	Add clarity for programming purposes
Supplementary analysis of primary endpoint	25 Apr 2024	Section 4.2.1.11: Added supplementary analysis to calculate the BOR by end of Cycle 16 by randomised treatment group for the measurable PN FAS	N/A	In support of the primary endpoint
Supplementary analysis of primary endpoint	25 Apr 2024	Section 4.2.1.11: Removed supplementary analysis of ORR by end of Cycle 16 on the FAS.	N/A	Analysis not required

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Sensitivity analyses for 1st and 2nd key secondary endpoints	25 Apr 2024	Sections 4.2.2.1.6.1, 4.2.2.2.6: Added details on how to identify the first cycle eligible for imputation for each intercurrent event Added further details on implementation reversion-to-baseline approach to align with paper by Ratitch et al. (2013)	N/A	Add clarity for programming purposes
Best objective response definition	25 Apr 2024	Section 4.2.3.2 (Table 11): Table updated to add in more scenarios for achieving BOR of cCR CR, cPR, PR and table title updated for clarity.	N/A	Updated to cover more scenarios
Time-to-event endpoints: DOR, PFS, TTP, TTR	25 Apr 2024	Sections 4.2.4.5, 4.2.5.5, 4.2.6.5, 4.2.7.5: Individual subject-level data to be summarised in an overall time-to-event listing for DCO2 and DCO3	N/A	To aid interpretability of results
Intercurrent events for REiNS-related secondary endpoints	01 May 2024	Sections 4.2.3.3, 4.2.4.3, 4.2.5.3, 4.2.6.3, 4.2.7.3: Clarification of intercurrent events for single-arm ORR, DOR, PFS, TTP and TTR endpoints.	N/A	Updated for clarity.
Time to pain palliation	1 May 2024	Section 4.2.10.5: Specification of variables to adjust for in the Cox model for time to pain palliation.	N/A	Missing in error from previous versions of SAP.
Analyses of pain medication	25 Apr 2024	Section 4.2.11.6: Tables of chronic PN pain and spike PN pain medication by drug name will be dropped and the drug names will be listed instead	N/A	Drug names from eDiary cannot be cleaned to provide meaningful values for table
Analyses of pain medication	25 Apr 2024	Section 4.2.11.6: Column charts of strongest analgesic ladder category per cycle and shift table of strongest analgesic ladder category as compared to baseline at each cycle over the randomised period and long-term period removed	N/A	Not required for interpretability of results

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Analyses of pain medication	25 Apr 2024	Section 4.2.11.6: Listing of chronic and spike PN pain medications added.	N/A	To aid interpretability of results
Analyses of PROMIS	25 Apr 2024	Section 4.2.13.6: Stacked column charts of distribution of response and distribution of change in response by treatment group at each assessment removed. Replaced with plots of mean scores and change from baseline with 95% CI by cycle for randomised period added	N/A	To align figures with data presented in table of descriptives.
Additional analyses of EQ-5D-5L	01 May 2024	Section 14.2.15.6: Removed descriptive analysis for number of completed questionnaires at each visit and number and proportion responding to each dimension of the EQ-5DL-5L.	N/A	Not required for interpretability of results – instrument completion rate already covered in Section 4.1.1
PAINS-pNF spike target PN pain intensity	25 Apr 2024	Section 4.2.17.4: Added figure of LS means and 95% CIs Section 4.2.17.5: Added plots of means scores and change from baseline for randomised period	N/A	To aid interpretability of results
Exposure	25 Apr 2024	Section 4.6.1.1: DCO date added as potential cut-off when deriving exposure time.	N/A	Add clarity for programming purposes
Adverse events	25 Apr 2024	Section 4.6.2.2: Remove the following tables for all treatment periods: • SAEs leading to discontinuation of study intervention • SAEs leading to discontinuation of study intervention, possibly related to study intervention	N/A	Not required for interpretability of results – repetition of data presented in other tables
Adverse events	1125 Apr 2024	Section 4.6.2.2: Threshold for summary of most common AEs by preferred term changed from >=5% to >=10%.	N/A	To align with other Koselugo studies

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Vital signs – height	1125 Apr 2024	Section 4.6.6.2: Clarification of how height should be summarised in tables and listings and the baseline value to use	Yes	To align with CSP v4.
China and Japan subpopulation analyses	25 Apr 2024	Appendices detailing analyses to be performed for the subpopulations of Japan (Section 7) and China (Section 8) added.	N/A	Provide details on specific subpopulation analyses.
Primary analysis of primary endpoint and supplementary endpoints on Measurable PN FAS	30 Aug 2024	The primary analysis of the primary endpoint and related supplementary endpoints (Tumour volumes and BOR) will be performed on the FAS instead of the Measurable PN FAS, in order to analyse all randomised participants (i.e. ITT population). Measurable PN FAS has been removed as an analysis population.	No	To align with FDA feedback on the SAP v3
Update of the definition of the analysis sets of selumetinib FAS, extended selumetinib FAS, fed measurable FAS.	30 Aug 2024	The requirement of having a measurable target PN at baseline per ICR has been removed from the definition of the analysis set to align with the definition of the FAS. The 'Fed measurable PN FAS' was renamed to 'Fed FAS'.	No	To align with FAS definition
PFS During the Randomised Period	30 Aug 2024	Correction of an inconsistency from previous version: analysis aligned with all other time-to-event endpoints using the time unit of Months rather than Cycles.	N/A	Correction of an inconsistency
Confirmation of target lesion response	30 Aug 2024	Correction of an inconsistency from previous version: confirmed CR/PR will be achieved on a consecutive scan within 3 to 6 months and not within 3 to 6 (± 2weeks) as it was reported in Section 4.2.1.2. in line with CSP.	Yes	Correction of an inconsistency

AE, adverse event; ANCOVA, analysis of covariance; BOR, best objective response; eCDF, Empirical cumulative distribution function; CI, confidence interval; cPR, confirmed partial response; CR, complete response; CSP, clinical study protocol; CSR, clinical study report; DCO, data cut-off; DOR, duration of response; ECG, electrocardiogram; EDC, electronic data capture; EOT, end of treatment; EQ-5D, EuroQoL five dimensions; EQ-5D-5L, EuroQoL five dimensions, five level health state utility index; FAS, full analysis set; FDA, Food and Drug Administration; HrQoL, health related quality of life; IP, investigational product;

MMRM, Mixed Model Repeated Measures; MRI, magnetic resonance imaging; MTP, Multiple Testing Procedure; N/A, not applicable; ORR, objective response rate; PAINS-pNF, Pain Intensity Scale for Plexiform Neurofibroma; PedsQL, Paedriatric Quality of Life Inventory; PFS, progression free survival; PGIC, Patient Global Impression of Change; PGIS, Patient Global Impression of Severity; PII-pNF, Pain Interference Index for Plexiform Neurofibroma; PK, pharmacokinetic; PlexiQoL; Plexiform Neurofibroma Quality of Life scale; PN, plexiform neurofibroma; PR, partial response; PRO, patient-reported outcome; PROMIS, Patient-Reported Outcomes Measurement Information System; QC, quality control; RDI, relative dose intensity; SAE, serious adverse event; SAP, statistical analysis plan; TEAE, treatment-emergent adverse event; TTP, time to progression; TTPP, time to chronic target PN pain palliation; TTR, time to response; US, United States.

1 INTRODUCTION

The purpose of this document is to give details for the statistical analysis of study D134BC00001 supporting the clinical study report (CSR). The reader is referred to the clinical study protocol (CSP) and the case report form (CRF) for details of study conduct and data collection. This statistical analysis plan is based on Amendment 3 of the CSP dated 14 December 2023. Details of the statistical analysis to support submissions in Japan and China are also provided in Sections 7 and 8, respectively. Details of interim analyses to be performed for this study can be found in the interim analysis SAP.

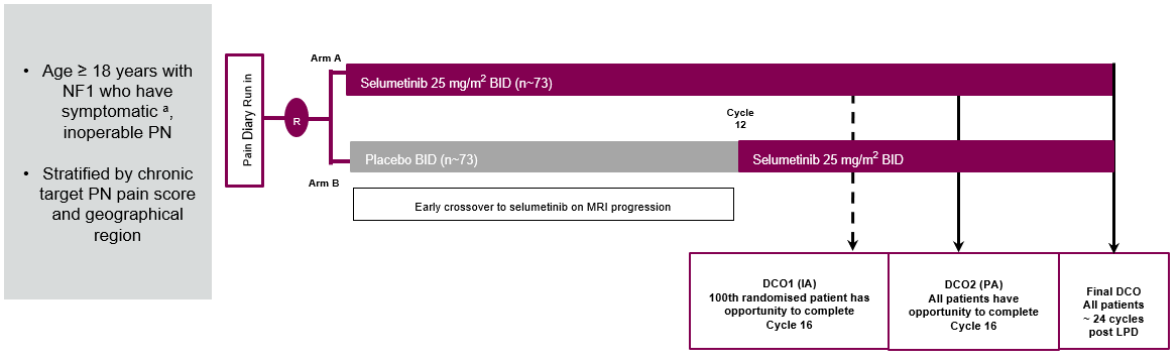
1.1 Study Design

This is a Phase III, multicentre, international study with a parallel, randomised, double-blind, placebo-controlled, 2-arm design to assess the efficacy and safety of selumetinib in adult patients with neurofibromatosis type 1 (NF1) who have symptomatic, inoperable plexiform neurofibroma (PN).

The primary objective is to compare the effect of selumetinib relative to placebo by assessment of confirmed partial (PR) and complete response (CR) rate (objective response rate [ORR]) by the end of Cycle 16 using volumetric magnetic resonance imaging (MRI) analysis as determined by independent central review (ICR) (per Response Evaluation in Neurofibromatosis and Schwannomatosis [REiNS] criteria). The study population is patients with NF1 who have symptomatic, inoperable PN.

An outline of the study design is shown in Figure 1.

Figure 1: Study Design



[a] Symptoms may include (but are not limited to) pain, motor morbidity, and disfigurement.

bid, twice daily; DCO, data cut-off; IA, interim analysis; LPD, last patient dosed; n, number of observations; NF1, neurofibromatosis type 1, PA, primary analysis; PN, plexiform neurofibroma.

It is estimated that approximately 212 adult patients will be enrolled at approximately 46 sites across 13 countries to achieve approximately 146 patients randomised to study intervention. Patients will be randomised 1:1 to receive selumetinib or placebo (i.e. 73 patients per arm). Randomisation will be stratified at the central level by average baseline PAin INtensity Scale for Plexiform Neurofibroma (PAINS-pNF) chronic target PN pain score (< 3 and ≥ 3) and geographical region (Europe, China, Japan, and Rest of World). A blocked randomisation will be generated, and randomisation will be balanced within the IRT at the central level. The number of patients randomised will be capped at approximately 106 patients with an average baseline PAINS-pNF chronic target PN pain score ≥ 3 and approximately 40 patients with an average baseline PAINS-pNF chronic target PN pain score < 3. Patients randomised to selumetinib will receive selumetinib 25 mg/m^2 orally twice daily (bid) (based on body surface area [BSA], capped at 50 mg bid when BSA is ≥ 1.9 m^2) in 28-day cycles until a selumetinib discontinuation criterion is met. Patients randomised to placebo will receive placebo orally bid and will receive selumetinib treatment after the end of Cycle 12. Selumetinib capsules will be administered using BSA-based dosing. If it is anticipated that 20% or more of patients (across blinded study intervention groups) withdraw prior to the end of Cycle 16 for reasons other than progression, then recruitment rates withstanding, enrolment will continue to randomise at least 50% of the total number of dropout patients to ensure that the primary endpoint is adequately powered.

The interim analysis (data cut-off [DCO]1) will occur after the 100th randomised patient has had the opportunity to complete their end of Cycle 16 assessment, and the primary analysis (DCO2) will occur after the last patient dosed (LPD; the last patient receiving first dose) has had the opportunity to complete their end of Cycle 16 assessment. If the interim analysis DCO (DCO1) is due to take place within approximately 4 months of the primary analysis DCO (DCO2), then the interim analysis may not be performed and only the primary analysis will be performed. The primary analysis of the ORR will take place at the end of the Cycle 16 landmark and will be assessed in all randomised patients with measurable target PN at baseline per ICR. The 1st key secondary endpoint of change from baseline in PAINS-pNF chronic target PN pain intensity at Cycle 12 will be assessed in patients with a PAINS-pNF chronic target PN pain intensity score of ≥ 3 at baseline and at least one post baseline average cycle PAINS-pNF chronic target PN pain intensity score. The 2nd key secondary endpoint of change from baseline in Plexiform Neurofibroma Quality of Life scale (PlexiQoL) total score at Cycle 12 will be assessed in full analysis set (FAS) patients with a baseline PlexiQoL total score and at least one post-baseline PlexiQoL

total score. The analysis of the primary endpoint and of the two key secondary endpoints will be comparative between selumetinib and placebo by the end of the randomised period.

The final analysis (final DCO) will occur approximately 24 cycles post LPD. The proposed duration for the study is approximately 24 months from the LPD to provide efficacy data as well as further safety and tolerability data. Following the end of the study, a mechanism will be in place to ensure that patients will be able to continue taking selumetinib as long as they derive clinical benefit, as judged by the investigator and in the absence of discontinuation criteria. After study intervention discontinuation, all patients will undergo an end-of-treatment visit and will be followed up for safety assessments 30 (+ 7) days after their last dose of study intervention (i.e., the safety follow-up visit).

2 CHANGES TO PROTOCOL PLANNED ANALYSES

- The split of the safety analysis set (SAF) into the Randomised Period SAF and the On-Selumetinib SAF is not listed in protocol.

- Patient Global Impression of Severity (PGIS) and Patient Global Impression of Change (PGIC) are collected in the study to support the psychometric validation of the PAINS-pNF questionnaire. However, following a request by the clinical and health related quality of life (HrQoL) teams, summary statistics of PGIS and PGIC by treatment over time have been included as additional exploratory clinical outcome assessments (COAs) for the CSR.

- The stratification factor ‘average baseline chronic target PN pain group’ has been omitted from primary analysis model in error in the CSP for the following endpoints: PlexiQoL (Section 4.2.2.2.5); PII-pNF pain interference total score (Section 4.2.12.5); PROMIS (Section 4.2.13.5); PedsQL (Section 4.2.14.5); EQ-5D-5L (Section 4.2.15.5).

- Subgroup analyses by baseline PAINS-pNF chronic target PN score group for PII-pNF, PROMIS, PedsQL and EQ-5D-5L will not be performed but these analyses will be supplementarily run on the set of patients with a baseline PAINS-pNF chronic target PN pain intensity score ≥ 3 (pain FAS).

- The primary endpoint (ORR by end of Cycle 16) and related supplementary endpoints (Target PN volumes and BOR) will be analyzed in the FAS (i.e. ITT population) and not on the Measurable PN FAS defined in CSP, which requires participants to have a measurable target PN at baseline per ICR, following FDA feedback on SAP version 3. Subsequently, the definition of the following analysis sets have changed to remove the requirement of having a measurable target PN at baseline per ICR: selumetinib FAS, extended selumetinib FAS, fed measurable FAS.

3 DATA ANALYSIS CONSIDERATIONS

3.1 Timing of Analyses

There are three planned DCOs for this study. A CSR will be written for the interim analysis (DCO1) and two addendum CSRs for the primary analysis (DCO2) and final analysis (final DCO), provided that the AstraZeneca unblinded review committee (URC) makes the decision to unblind the study team at the interim analysis. If, at the interim analysis, the primary endpoint is not statistically significant or the AstraZeneca URC makes the decision not to unblind the study team at the interim analysis, then only two CSRs (primary analysis and final DCO) will be written. In the latter scenario, the study team will remain blinded until the primary analysis.

Interim Analysis - DCO1

The DCO1 is planned after the 100th randomised patient has had the opportunity to complete their end of Cycle 16 assessment.

Primary Analysis - DCO2

The DCO2 is planned after the patient LPD has had the opportunity to complete their end of Cycle 16 assessment (ie, following the end of recruitment).

Final DCO

The final DCO is planned approximately when all patients have had the opportunity to achieve Cycle 24.

3.2 Analysis Populations

3.2.1 Enrolled

All patients who signed the informed consent form (ICF).

3.2.2 Full Analysis Set (FAS)

All patients who are randomised to study intervention in the study. Treatment groups will be compared on the basis of randomised study intervention, regardless of the study intervention actually received. Patients who were randomised but did not subsequently receive study intervention are included in the analysis in the treatment group to which they were randomised.

3.2.3 Pain FAS

All patients with a baseline PAINS-pNF chronic target PN pain intensity score ≥ 3. Treatment groups will be compared on the basis of randomised study intervention, regardless of the intervention actually received. Patients who were randomised but did not subsequently receive study intervention are included in the analysis in the treatment group to which they were randomised

3.2.4 Selumetinib FAS

All patients who are randomised to selumetinib.

3.2.5 Extended Selumetinib FAS

All patients who are randomised to study intervention in the study who have received at least one dose of selumetinib, that is, including patients randomised to placebo who crossover to selumetinib treatment.

3.2.6 Safety Analysis Set (SAF)

All enrolled patients who received any amount of study intervention. Safety data will be summarised according to the treatment received. Patients will be assigned to either the selumetinib group or the placebo/selumetinib group.

3.2.7 Randomised Period SAF

The randomised safety analysis set will consist of all enrolled patients who received any amount of study intervention in the randomised period as defined in Section 3.3.2.2. Safety data will be summarised according to the treatment received. If a patient receives any amount of selumetinib in the randomised period, they will be summarised in the selumetinib group. If a patient only receives placebo in the randomised period, they will be summarised in the placebo group.

3.2.8 On-Selumetinib SAF

The on-selumetinib safety analysis set will consist of all enrolled patients who received any amount of selumetinib in the on-selumetinib period as defined in Section 3.3.2.3. Safety data will be summarised according to the treatment received.

3.2.9 Pharmacokinetics Analysis Set

All patients assigned to study intervention who take at least one dose of study intervention for whom any post dose reportable pharmacokinetic (PK) concentration is available and who do not violate or deviate from the protocol in ways that would significantly affect the PK analyses.

3.2.10 Fed FAS

All patients who are randomised to study intervention in the study who have received at least one dose of selumetinib in the fed state, and who have at least one evaluable scan per ICR assessment in the fed state at the end of Cycle 30 or later.

"Fed state" is defined from the CSP v4.0 dated 14 December 2023 that from the end of Cycle 24 (Cycle 25 Day 1) patients will not be required to continue to observe the fasting restriction. Patients who reach the end of Cycle 24 prior to the approval of CSP v4.0 dated 14 December 2023 will be considered in the fed state from the timepoint when they will reconsent to participate in the study without the fasting restriction, and the date of their ICF signature for the CSP v4.0 after Cycle 25 Day 1 will be the start date of the fed state.

3.2.11 Summary of Outcome Variables and Analysis Sets

The analysis sets for each outcome variable are provided in Table 1. For some outcome variables, the analysis set will use a further subset (e.g., patients who are evaluable for the analysis of duration of response (DOR) are those who responded in the ORR analysis) which are described in detail in the endpoint analysis sections in Section 4.2.

Table 1 Summary of Outcome Variables and Analysis Sets

Outcome variable	Analysis set
Efficacy data	
ORR by the end of Cycle 16 (primary endpoint)	FAS
Change from baseline in PAINS-pNF chronic target PN pain intensity score at Cycle 12 (1st key secondary endpoint)	Pain FAS
Change from baseline in PlexiQoL total score at Cycle 12 (2nd key secondary endpoint)	FAS and Pain FAS (supplementary)
Chronic target PN pain palliation, time to chronic target PN pain palliation	Pain FAS
ORR (single arm), DOR (single arm), PFS (single arm), time to progression (single arm), time to response (single arm), and BOR (single arm)	Selumetinib FAS, Extended Selumetinib FAS, and FAS (selumetinib arm only)
PFS during the randomised period (exploratory endpoint)	FAS
Best percentage change from baseline in target PN volume during the randomised period	FAS
Pain medication, PII-pNF pain interference total score, PROMIS Physical Function items, Peds QL Inventory (NF1 Module Acute Version 3.0 – Adult Report), Skin Sensations Domain, EQ-5D-5L, EQ-VAS, PAINS-pNF spike target PN pain intensity score	FAS

Table 1 Summary of Outcome Variables and Analysis Sets

Outcome variable	Analysis set
Percentage change from baseline in target PN volume at Cycle 30 or later.	Fed FAS
Study Population/Demography Data	
Demography characteristics (e.g., age, sex, etc.)	FAS
Baseline and disease characteristics	FAS
Important deviations	FAS
Medical/surgical history	FAS
Previous anti-PN therapy	FAS
Concomitant medications/procedures	FAS
Subsequent anti-PN therapy	FAS
PK Data	
Selumetinib and N-desmethyl selumetinib plasma concentration data and PK parameters	PK
Safety data	
Exposure	Randomised period SAF and On-selumetinib SAF
Adverse events	Randomised period SAF and On-selumetinib SAF
Laboratory measurements	Randomised period SAF and On-selumetinib SAF
Vital signs	Randomised period SAF and On-selumetinib SAF
ECGs	Randomised period SAF and On-selumetinib SAF
Ophthalmologic assessment	Randomised period SAF and On-selumetinib SAF
Physical Examinations	Randomised period SAF and On-selumetinib SAF
ECHO	Randomised period SAF and On-selumetinib SAF

BOR, best objective response; DOR, duration of response; ECG, electrocardiogram; ; ECHO, echocardiogram; EQ-5D-5L, EuroQoL five dimensions, five level health state utility index; EQ-VAS, EuroQol-visual analogue scale; FAS, full analysis set; NF1, neurofibromatosis type 1; ORR, objective response rate; PAINS-pNF, PAin INtensity Scale for Plexiform Neurofibroma; PFS, progression free survival; PII-pNF, Pain interference index – plexiform neurofibroma; PK, pharmacokinetic(s); PlexiQoL, Plexiform Neurofibroma Quality of Life scale; PN, plexiform neurofibroma; PROMIS, Patient-Reported Outcomes Measurement Information System; SAF, Safety Analysis Set.

3.3 General Considerations

The general principles mentioned below will be followed throughout the study:

- Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarised by the number of observations (n), mean, standard deviation, minimum (min), 1st quartile (Q1), median, 3rd quartile (Q3), and maximum (max). Categorical variables will be summarised by frequency counts and percentage for each category.
- For continuous data, the mean and median will be rounded to one additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.
- For categorical data, percentages will be rounded to one decimal place with the exception of 100%, which is presented as a whole number. For 0 results, the percentages will not be included and only 0 will be presented as result of the categorical variable.
- Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment group. Overall totals will be calculated for baseline summaries only.
- Descriptive summaries for all endpoints collected by visit will be based on the planned visit schedule unless otherwise stated.
- Statistical Analysis Software (SAS®) version 9.3 or above will be used for all analyses.

3.3.1 Baseline Definition and Change From Baseline

There will be two baseline definitions for the study.

In general, for efficacy (excluding the Extended Selumetinib FAS) and PRO) endpoints (see also Section 3.3.8 for further information on derivation of baseline for certain PRO endpoints), the last observed measurement prior to randomisation will be considered the baseline measurement. However, if an evaluable assessment is only available after randomisation but before the first dose of randomised treatment, then this assessment will be used as the baseline. For the Extended Selumetinib FAS (defined in Section 3.2.5), baseline will be defined as the last non-missing value obtained prior to the first dose of selumetinib.

For safety endpoints (excluding the on-selumetinib safety analysis set), the last observation before the first dose of study intervention will be considered the baseline measurement unless otherwise specified.

If two or more measurements are recorded on the same day with no time (or on the same day with the same time/timepoint) then the average will be used as the baseline. If these measurements occur on the day where a patient crosses over from placebo to selumetinib, the pre-crossover value will be used as the baseline. If there are multiple pre-crossover measurements, then the average of the pre-crossover measurements will be used as the baseline. For non-numeric laboratory tests (i.e., some of the urinalysis parameters) where taking an average is not possible, the best value would be taken as the baseline as this is the most conservative.

For the on-selumetinib safety analysis set (defined in Section 3.2.8), baseline will be defined as the last non-missing value obtained prior to the first dose of selumetinib.

In all summaries, change from baseline variables will be calculated as the post-treatment value minus the value at baseline. The percentage change from baseline is calculated as (post-baseline value – baseline value) / (baseline value) * 100.

3.3.2 Study Periods

3.3.2.1 On-Treatment MRI Volumetric Assessments

On-treatment MRI volumetric assessments data will be defined as data after the date of first dose of study intervention until the study intervention discontinuation date or DCO, whichever occurs first. For analyses using the Extended Selumetinib FAS, volumetric assessments prior to the first dose of selumetinib will not be considered on-treatment MRI volumetric assessments.

Unscheduled scans will be considered for the on-treatment MRI volumetric assessments if those scans are not excluded due to prolonged study intervention interruption as described in Section 4.2.1.6.

3.3.2.2 Randomised Period

The randomised period for patients randomised to selumetinib will be defined as the date of first dose of selumetinib until the first date when the selumetinib kit(s) are dispensed as recorded in the drug accountability form at or after the Cycle 12 Day 28 visit date or the date of study intervention discontinuation (whichever occurs first).

The randomised period for patients randomised to placebo will be defined as the date of first dose of placebo until the date when the first selumetinib kit(s) are dispensed as recorded in the drug accountability form (crossover date) or date of study intervention discontinuation (whichever occurs first).

Efficacy assessments taken on the date of the first open-label selumetinib kit dispensed as recorded in the drug accountability form mentioned above will be included in the randomised period.

3.3.2.3 On-Treatment Safety Period

For the purposes of summarising safety data assessed at visits, in addition to baseline data, only on-treatment data are included in the summary tables. On-treatment data will be defined as data from the first dose of study intervention until the last dose of study intervention + 30 days.

The On-Treatment Safety Period will be further subset into the On-treatment Safety Randomised Period and On-selumetinib Safety Period.

If the On-Treatment Safety Period of a patient is still ongoing at the time of the analysis, their on-treatment safety period will be truncated at the DCO date, and any measurement/event with a start date after the DCO date will be presented at a later DCO.

On-treatment Safety Randomised Period.

The on-treatment safety randomised period for patients receiving selumetinib in the randomised period will be defined as the date from the first dose of study intervention until the first date when the selumetinib kit(s) are dispensed as recorded in the drug accountability form at or after the Cycle 12 Day 28 visit date or 30 days after study intervention discontinuation or up to one day prior to start of subsequent therapy, whichever occurs first.

The on-treatment safety randomised period for patients receiving only placebo in the randomised period will be defined as the date from the first dose of study intervention until the date when the first selumetinib kit(s) are dispensed as recorded in the drug accountability form (crossover date) or 30 days after placebo discontinuation or up to one day prior to the start of subsequent therapy, whichever occurs first.

If the on-treatment randomised period of a patient is still ongoing at the time of the analysis, their on-treatment randomised period will be truncated at the DCO date, and any measurement/event with a start date after the DCO date will be presented at a later DCO.

On-Selumetinib Safety Period

The on-selumetinib safety period will be defined as the date from the first dose of selumetinib until 30 days after the last dose of selumetinib or up to one day prior to the start of subsequent therapy, whichever occurs first.

The baseline definition for patients receiving placebo in the randomised period and then crossing over to selumetinib either on Cycle 13 Day 1 or earlier in the case of progression will differ for the on-treatment safety randomised period and the on-selumetinib period as described in Section 3.3.1.

If the on-selumetinib period of a patient is still ongoing at the time of the analysis, their on-selumetinib period will be truncated at DCO date, and any measurement/event with a start date after the DCO date will be presented at a later DCO.

Table Presentation

Tables for the on-treatment during the randomised period will be presented by patients who received only placebo and patients who received any selumetinib.

Tables for the On-selumetinib Safety Period will be presented by patients who received selumetinib from the start of the randomised period and patients who received selumetinib after crossing over from placebo.

The baseline definition for patients receiving placebo and then starting on selumetinib treatment (post crossover) will differ for the on-treatment during the randomised period and the on-selumetinib period as described in Section 3.3.1.

3.3.2.4 Treatment-emergent Periods

The analysis of treatment-emergent adverse events (TEAEs) and concomitant medication will be performed separately for TEAEs occurring/concomitant medication taken during the safety randomised period and TEAEs occurring/concomitant medication taken during the on-selumetinib period. While the analysis of TEAEs in the safety randomised period will be comparative of the two treatment groups, TEAEs presented for the on-selumetinib period will be not comparative due to the different exposure periods. To note, the two safety analysis periods are not mutually exclusive.

TEAEs and Concomitant Medication during the Safety Randomised Period

For patients receiving any selumetinib in the safety randomised period, an adverse event (AE) or a concomitant medication is considered treatment emergent during the safety randomised period if the start date/time of the event or worsening date/time (only for AEs) occurs between the date of first dose of study intervention (Cycle 1 Day 1) and last dose of Cycle 12 or date of study intervention discontinuation (end of treatment [EOT]) + 30 days or DCO date, whichever comes first.

For patients receiving placebo in the safety randomised period, an AE or a concomitant medication is considered treatment emergent during the safety randomised period if the start date/time of the event, or worsening date/time of an AE occurs between the date of first dose of placebo (Cycle 1 Day 1) and the crossover date or date of early placebo discontinuation + 30 days or DCO, whichever comes first.

A worsening is defined as having any on-treatment severity during the safety randomised period worse than the baseline severity. The reported severity of an AE that is treatment emergent during the safety randomised period is given by the maximum Common

Terminology Criteria for Adverse Events (CTCAE) grade recorded any time on-treatment during the safety randomised period. The reported outcome of an AE will be the last outcome corresponding to the maximum CTCAE grade recorded on-treatment in the safety randomised period. Seriousness and causality during the randomised period will also be reported as the worst value recorded in the safety randomised period.

TEAEs and Concomitant Medication during the On-Selumetinib Period

An AE or a concomitant medication is considered treatment-emergent during the on-selumetinib period if the start date/time of the event, or worsening date/time of a AE occurs between the first dose of selumetinib (which will be Cycle 1 Day 1 for patients receiving selumetinib in the safety randomised period and the date of the crossover for patients receiving placebo in the safety randomised period) until 30 days after the last dose of selumetinib or up to DCO, whichever comes first.

A worsening is defined as having any on-treatment severity during the on-selumetinib period worse than the severity of the safety randomised period. Worsening in the on-selumetinib period will be assessed by comparing to the last-recorded value in the safety randomised period.

For AEs of placebo patients starting during the safety randomised period and worsening during the On-Selumetinib period, the stop date of the event during the safety randomised period will be one day prior to the date of the first worsening after crossing over, while the start date of the event during the on-selumetinib period will be equal to the date of first worsening after crossing over.

3.3.3 Visit Window

Time windows will be defined for any presentations that summarise values by visit. The following conventions will apply:

- Study day references the date of first dose of study intervention as Day 1.
- The time windows will be exhaustive so that data recorded at any timepoint have the potential to be summarised. Inclusion within the time window will be based on the actual date and not the intended date of the visit.
- The window for visits following baseline is constructed in such a way that the upper limit of the interval falls halfway between the two visits (the lower limit of the first post-baseline visit is Day 2). If an even number of days exists between two consecutive visits, then the upper limit is taken as the midpoint value minus 1 day.
- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment during the randomised period

and the on-selumetinib period (as defined in Section 3.3.2.3) will be used (regardless of where it falls in an interval).

- Listings display all values contributing to a timepoint for a patient.
- For visit-based summaries/analyses, if there is more than one value per patient within a time window, then the closest value to the scheduled visit date will be summarised or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the patient that contributed to the summary table, wherever feasible.

Note: In summaries of extreme values, all post-baseline values collected are used, including those collected at unscheduled visits regardless of whether the value is closest to the scheduled visit date.

Note: For the purpose of the visit-based summaries of efficacy data (except tumour volumes and response), the EOT visit will be mapped into one of the planned visit windows. For visit-based safety summaries and for tumour volume and response summaries, the EOT and/or Safety Follow-up Visit will be summarised apart.

3.3.4 Handling of Unscheduled Visits

Unscheduled visits are included in the method of assigning data to scheduled visits described in Section 3.3.3. Unscheduled visits are not included as separate visits in the summary tables.

For summaries at the patient level, such as those of extreme values, all post-baseline values collected are used to derive a patient level statistic including those collected at unscheduled visits and regardless of whether they appear in the corresponding visit-based summary.

3.3.5 Missing Data

In general, other than for partial dates, missing data will not be imputed and will be treated as missing unless specifically described in an analysis section. However, safety assessment values of the form of "< x" (i.e., below the lower limit of quantification) or > x (i.e., above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings.

3.3.5.1 Missing AE or Concomitant Medication Dates

For missing AE/concomitant medication dates, the imputation of dates for AEs and concomitant medications is performed to determine if an AE is treatment emergent and whether a medication is concomitant. Flags are retained in the database indicating where any programmatic imputation has been applied, and in such cases, any durations are not calculated.

For missing AE/concomitant medication start dates, the following will be applied:

- Missing day: Impute the 1st of the month unless the month is the same as month of the first dose of study drug and the end date is on or after the first dose of study drug or ongoing, then impute first dose date.
- Missing day and month: Impute 1st January unless year is the same as first dose date and the end date is on or after the first dose of study drug or ongoing, then impute first dose date.
- Completely missing: Impute first dose date unless the end date suggests it started prior to this, in which case impute 01 January of the same year as the end date.

An imputed start date of an AE/concomitant medication must be prior to the end date of the AE.

For missing AE/concomitant medication end dates, the following will be applied:

- Missing day: Impute the last day of the month unless the month is same as month of last dose of study drug, then impute last dose date.
- Missing day and month: Impute 31 December unless the year is the same as last dose date, then impute last dose date.
- Completely Missing: Assume that AE/concomitant medication is still ongoing (ie, do not impute a date).

3.3.5.2 Missing Dates of Diagnosis

For missing dates of diagnosis, if the day and/or month are missing use 01 and/or January. If the year is missing, put the complete date to missing.

3.3.5.3 Missing Death Dates

If a patient is known to have died where only a partial death date is available, then the date of death will be imputed as the latest of the last date known to be alive + 1 from the database and the death date using the available information provided:

- For missing day only: Use the 1st of the month.
- For missing day and month: Use 01 January.

Note: The last date known to be alive for each individual patient is defined as the latest date among all start/end/assessments/collections dates unrelated to death.

3.3.5.4 AEs With Missing Causality

AEs that have missing causality (after data querying) will be assumed to be related to study drug.

3.3.6 COVID-19 Impact

Summaries of data relating to the impact of the global/country COVID-19 situation will be generated, including the following:

- Discontinuations of study intervention and withdrawal from study due to global/country situation
- Important protocol deviations (IPDs) related to the global/country situation
- Global/country situation study disruptions
- Listing for patients affected by the global/country situation
- Listing for patients with reported issues in the Clinical Trial Management System due to the global/country situation

3.3.7 Multiplicity/Multiple Comparisons

The MTP will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoint (ORR by the end of Cycle 16), 1st key secondary endpoint (change from baseline in PAINS-pNF chronic target PN pain intensity score at Cycle 12), and 2nd key secondary endpoint (change from baseline in PlexiQoL total score at Cycle 12) intended for label claims and is described in Figure 2.

Significance levels will be adjusted for each confirmatory endpoint in order to preserve the overall type 1 error (familywise error rate) at 5% (two-sided) in the strong sense by defining a gatekeeping strategy where the families of hypotheses are tested in a sequential manner. Each family will be defined by one endpoint and the two associated tests (DCO1 and DCO2): the first family based on the primary endpoint, the second family based on the 1st key secondary endpoint, and the third family based on the 2nd key secondary endpoint.

Full alpha of 0.05 (two-sided) will be initially assigned to the primary endpoint. If statistical significance for the primary endpoint is reached (either at DCO1 or at DCO2), the overall alpha of 0.05 (two-sided) will be sequentially re-assigned to test the 1st key secondary endpoint. If statistical significance is reached for the 1st key secondary endpoint, then the overall alpha of 0.05 (two-sided) will be sequentially re-assigned to test the 2nd key secondary endpoint. If statistically significance is not reached for the primary endpoint, the p-value of the key secondary endpoints will only be nominal. If statistical significance is not reached for the 1st key secondary endpoint either at DCO1 or at DCO2, the p-value of the 2nd key secondary endpoint will only be nominal.

The interim analysis (i.e., when the 100th randomised patient has had the opportunity to have their end of Cycle 16 assessment) will take place with approximately 68.5% of the primary endpoint information expected at the primary analysis. For the first family, a split-alpha strategy will be conducted where an alpha of 0.003 (two-sided) will be initially allocated to test at DCO1. If statistical significance is not reached at DCO1, the primary endpoint will be tested again at DCO2 with the remaining alpha of 0.047 (two-sided). If statistical significance of the primary endpoint is reached at DCO1 or DCO2, the overall

alpha of 0.05 (two-sided) will be allocated to the 1st key secondary endpoint. Similarly, if statistical significance of the 1st key secondary endpoint is reached at DCO1 or DCO2, the overall alpha of 0.05 (two-sided) will be allocated to the 2nd key secondary endpoint. For the second and third families, the Haybittle-Peto approach (Haybittle 1971, Peto et al 1976) will be applied within the family.

The expected IF for each key secondary endpoint is 100%. In this case, the overall alpha of 0.05 (two-sided) will be available for DCO1 for the 2nd and 3rd families respectively. Therefore, the alpha split shown in Figure 2 should be considered as a reference example when IF < 100%. The actual IF for each key secondary endpoint will be calculated at the time of the analysis, and adjustments to the alpha will be made accordingly by applying the Haybittle-Peto approach (Haybittle 1971, Peto et al 1976) within the family.

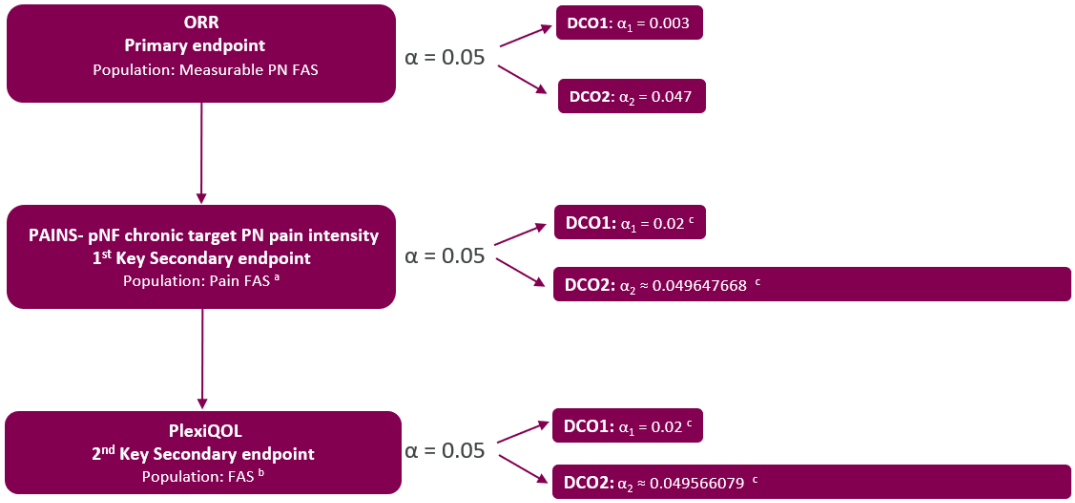
The testing of the 1st key secondary endpoint will proceed as follows according to the below two scenarios regarding significance on the primary endpoint:

1. If the primary endpoint is not statistically significant at DCO1, then the key secondary variables are not tested at DCO1. If the primary endpoint is statistically significant at DCO2, then the 1st key secondary endpoint will be tested at 0.05 (two-sided), given an expect IF of 100% at DCO1

2. If the primary endpoint is statistically significant at DCO1, then the 1st key secondary endpoint will be tested at 0.05, given an expected IF of 100% at DCO1.

The same approach will be applied to the testing of the 2nd key secondary endpoint based on the significance of the 1st key secondary endpoint, using their corresponding IF and alpha levels.

Figure 2: Multiple Testing Procedure



[a] Includes Pain FAS patients who have at least one post-baseline average cycle PAIN-pNF chronic target PN pain intensity score.

[b] Includes FAS patients who have a baseline PlexiQOL total score and at least one post-baseline PlexiQOL total score.

[c] Significance levels assuming 95.2% and 94.8% information fraction at DCO1 for 1st and 2nd key secondary endpoints, respectively. Final α_2 will be determined at DCO1 based on the actual information fraction using the Haybittle-Peto procedure in EAST. With 100% information fraction at DCO1, $\alpha_1 = 0.05$.

DCO, data cut-off; FAS, full analysis set; ORR, objective response rate; PAINS-pNF, PAin INtensity Scale for Plexiform Neurofibroma; PlexiQoL, Plexiform Neurofibroma Quality of Life scale; PN, plexiform neurofibroma.

3.3.8 PN Pain Medication Modified WHO Analgesic Ladder and PN Pain Medication Score Changes

The eDiary PN pain medications reported daily by the patient will be assigned by medical review to a drug class and analgesic ladder score identified in Table 2. In case a combination of one or more drug is reported in a daily diary entry, the strongest drug class and its related daily dose will be selected during the medical adjudication process. If all the medications of the combination belong to the same class, precedence will be given to a drug already taken by the patient. Otherwise, the first drug of the combination will be selected.

The strongest drug class and the highest analgesic ladder score at each cycle will be programmatically identified separately for both chronic PN medications and spike PN medications.

Table 2: Modified WHO Analgesic Ladder

Score	Drug class	Examples
0	No Analgesia	
1	Non-opioids	Acetaminophen, NSAIDs
2	Neuropathic/ other non-typical pain medications	Pregabalin, Gabapentin, Amitriptyline, Duloxetine Muscle relaxants
3	Weak opioids	Codeine
4	Strong opioids	Morphine

Additional class of pain medications added to include typical pain medications used by NF1-PN patients

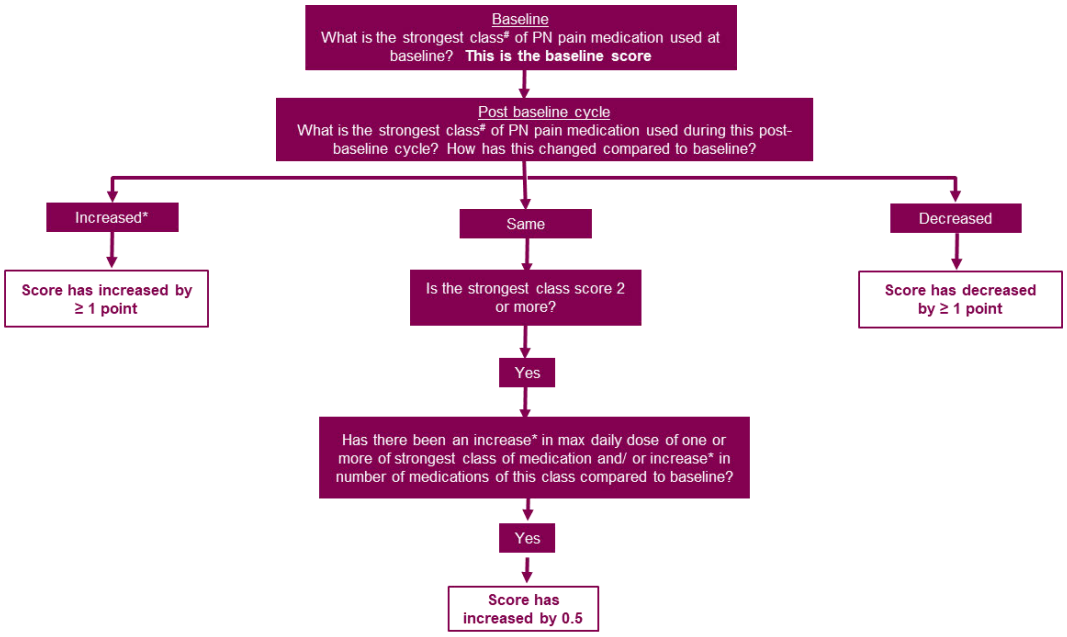
NF1, neurofibromatosis type 1; NSAID, nonsteroidal anti-inflammatory drug; PN, plexiform neurofibroma; WHO, World Health Organisation.

For chronic PN pain medication, changes from baseline will be derived. This definition is used for the derivation of PRO endpoints such as decrease in PN pain medication and chronic target PN pain palliation, as well as the intercurrent event "increase in chronic PN pain medication" of the key secondary PAINS-pNF chronic target PN pain intensity estimand as described in Section 4.2.2.1. No changes from baseline are derived for the spike PN chronic medication.

Following the identification of the strongest drug class in chronic PN pain medication at each cycle, the change in score at each cycle compared to baseline will be programmatically identified following the algorithm outlined in Figure 3. If changes in the strongest drug class from baseline do not follow under the categories of Figure 3. of "Increased" or "Decreased," the category "No Change" will be assigned. For the cycles where no PN chronic medication has been recorded in the eDiary, the change will be set to "Missing".

A secondary definition for change from baseline in PN pain medication (Increased/Decreased/No Change) is derived using the investigator analgesic requirement assessment recorded on the e-CRF, indicating whether the reason is PN pain or other pain.

Figure 3: PN Pain Medication Score Changes Algorithm



Based on modified WHO analgesic ladder
* Minimum requirement of increase is for specific drug to be taken at least 3 days in 28-day cycle.

PN, plexiform neurofibroma; WHO, World Health Organisation.

3.3.9 Derivation of the Attributability of Reason for Discontinued Randomised Study Intervention

In order to explore the robustness of the inference of the Mixed Model Repeated Measures (MMRM) models of the key secondary endpoints to deviate from the missing at random (MAR) assumption, the number and the reasons for early treatment discontinuations prior to Cycle 12 will be assessed.

Reasons for discontinuation of the randomised study intervention prior to Cycle 12 will be classified as attributable or non-attributable, depending on the potential relationship with study treatment (ie, lack of efficacy or tolerability issues), as per Table 3.

Table 3: Attributable Assessment for Reason of Study Intervention Discontinuation

Non-attributable	Evaluated on case-by-case	Attributable
Technical problems	Non-treatment related adverse event	Treatment -related adverse event
Lost to follow-up	Withdrawal by subject	Progressive disease

Pregnancy	Development of withdrawal criteria	Death
	Important protocol deviation	
	Physician decision	
	Study terminated by sponsor	
	Other	

A complete review of the individual reasons for discontinuation prior to Cycle 12 will be undertaken during a data review meeting of the clinical study team prior to study unblinding to determine the final categorisation.

3.4 Sample Size Determination

Approximately 212 patients will be enrolled to achieve approximately 146 patients randomly assigned to study intervention (either selumetinib or placebo) with a 1:1 ratio (73 patients in each arm).

Randomisation will be stratified by average baseline PAINS-pNF chronic target PN pain score (< 3 and ≥ 3) and geographical region (Europe, China, Japan, and Rest of World). The number of patients randomised will be capped at approximately 106 patients with an average baseline PAINS-pNF chronic target PN pain score ≥ 3 and approximately 40 patients with an average baseline PAINS-pNF chronic target PN pain score < 3.

Note: “Enrolled” means a patient’s agreement to participate in a clinical study following completion of the informed consent process. Potential patients who are screened for the purpose of determining eligibility for the study but are not randomly assigned/assigned in the study are considered “screen failures,” unless otherwise specified by the protocol.

With a sample size of 73 patients per arm, a Fisher’s exact test with a two-sided alpha of 5% will have >99% power to detect the difference between the selumetinib ORR of 20% and the placebo ORR of 0%. The ORR of 20% in the selumetinib arm by the end of Cycle 16 is assumed from ad hoc modelling performed using the SPRINT NCI and ICR data and the adult NF1 NCI study response rates. The ORR of 0% in the placebo arm is assumed because at the end of Cycle 16, patients randomised to placebo will only have received 4 cycles of selumetinib post-crossover, therefore any response detected post-crossover at the

end of Cycle 12 will be unconfirmed. Spontaneous shrinkage ≥20% during the first 12 cycles of placebo treatment is also not expected (Akshintala et al 2020).

Forty-two patients per arm are required for the study to have 90% power to detect a treatment difference at Cycle 12 of ≥ -2 in the 1st key secondary endpoint change from baseline of PAINS-pNF chronic target PN pain score (assuming a standard deviation of 2.8) in favour of selumetinib compared with placebo at a two-sided alpha level of 5%. To allow for approximately 20% drop out (i.e., patients without at least one post-baseline average cycle PAINS-pNF chronic target PN pain score), 106 patients with a baseline PAINS-pNF chronic target PN pain score ≥ 3 will be randomised in a 1:1 selumetinib:placebo allocation.

There is no formal sample size calculation regarding the randomisation of 40 patients with baseline PAINS-pNF chronic target PN pain score of < 3; however, this is deemed sufficient to ensure that the target population of adults with NF1 and symptomatic, inoperable PN (including those with little or no baseline PN pain) are represented.

By assuming 20% drop out (i.e., patients without at least one post-baseline PlexiQoL total score), 58 patients per arm will provide at least 80% power to detect a treatment difference at Cycle 12 in the 2nd key secondary endpoint change from baseline of PlexiQoL total score (assuming a standard deviation of 2.3) of at least -1.2 in favour of selumetinib compared with placebo at a two-sided alpha level of 5%. Refer to Table 4 for some scenarios for the statistical power for the 2nd key secondary endpoint with different mean differences in the change from baseline of the PlexiQoL total score at Cycle 12 between the two treatment groups. This range of possible treatment effects for selumetinib that could be observed in a multinational population was extrapolated based on the differences observed in other HRQoL scores in other rare disease studies as compared to an adult selumetinib cohort (D1346C00011).

Table 4: Scenarios for the Power of the Second Key Secondary Endpoint

Mean treatment difference	Statistical power
-0.7	37%
-0.8	47%
-0.9	56%
-1	65%
-1.1	73%
-1.2	80%
-1.3	86%
-1.4	91%
-1.5	94%
-1.6	96%
-1.7	98%

If it is anticipated that 20% or more of patients (across blinded study intervention groups) withdraw prior to the end of Cycle 16 for reasons other than progression, then recruitment rates withstanding, enrolment will continue to randomise at least 50% of the total number of dropout patients to ensure the primary endpoint is adequately powered.

4 STATISTICAL ANALYSIS

This section provides information on definitions, derivation, and analysis/data presentation per domain.

4.1 Study Population

The domain study population covers subject disposition, analysis sets, protocol deviations, demographics, baseline characteristics, medical history, prior and concomitant medication, and study drug compliance.

4.1.1 Subject Disposition and PRO Completion Status

4.1.1.1 Definitions and Derivations

Geographical region will be separated into China, Japan, Europe (including France, Germany, Italy, Poland, Russia, Spain, and United Kingdom), and Rest of World (including Australia, Brazil, Canada, and United States).

4.1.1.2 Presentation

Subject disposition will be summarised, presenting the number of patients enrolled; the number and percentage of patients randomised; and the number and percentage of patients who did and did not receive study intervention, who were randomised to placebo and crossed over to selumetinib – with the breakdown of whether the crossover occurred after the end of the Cycle 12 visit or earlier, who discontinued study intervention, who discontinued study intervention prior to the end of the Cycle 12 visit, and who discontinued the study. A breakdown by reason for patients discontinuing study intervention, discontinuing study intervention prior to the end of the Cycle 12 visit, and discontinuing from the study will be included in this summary. The number of patients ongoing on treatment and in the study at the time of the DCO will also be presented.

Disposition summaries will be presented for all enrolled patients.

A summary of recruitment by geographical region, country, and site will be produced for all patients in the FAS. Listings presenting details of individual patient disposition events will be produced for those patients in the FAS discontinuing study intervention and/or discontinuing from the study.

Distribution of the eDiary entries will be presented for the FAS and the Pain FAS. For the PAINS-pNF (chronic target PN pain and spike PN pain) and the PN pain medication eDiaries, summary tables of eDiary completion rates per cycle (including baseline) and study intervention group will be presented at each cycle (including baseline). The number and percentage (ie, eDiary compliance) of expected daily diary entries is defined as the sum of the number of daily entries expected prior to study treatment discontinuation for all patients on treatment. A daily entry will be considered as completed for PAINS-pNF if a daily PN pain score is recorded. A diary entry will be considered as completed for PN pain medication if the question "Did you take any medication for your tumour pain from the time you went to bed last night until now (including overnight)?" was answered and if the items after it are answered. Additionally, for PAINS-pNF chronic target PN pain, the number of expected daily diary entries and eDiary compliance will be presented by 0, 1, 2, 3, or 4 non-overlapping 7-day periods with at least 4 daily pain scores at baseline and during each 28-day cycle, which is the minimum compliance requirement to calculate a cycle average chronic PN pain intensity score.

For the rest of the PRO instruments, the number of expected forms and the following instrument completion rates at each cycle will be reported by treatment group:

- The number and percentage of patients with all questions completed.
- The number and percentage of patients meeting at least the minimum requirements for scoring of the instrument.
- The number and percentage of patients with at least 1 question completed.

All completion rates will be presented for the FAS, and the denominator will be the number of patients who are expected to have PRO assessments.

4.1.2 Analysis Sets

4.1.2.1 Definitions and Derivations

The definitions and derivations of the analysis set are described in Section 3.2.

4.1.2.2 Presentation

The number of patients in each analysis set and the reason for exclusion from each analysis set will be summarised by treatment group overall for all enrolled patients.

A listing of individual patients not included in each analysis set will be provided.

4.1.3 Protocol Deviations

4.1.3.1 Definitions and Derivations

Protocol deviations will be defined as any change, divergence, or departure from the study design in the CSP. Prospective approval of protocol deviations to recruitment and enrolment criteria, also known as protocol waivers or exemptions, is not permitted.

IPDs are defined in the protocol deviation plan and will be finalised at the data review meeting before database lock. IPDs are identified from the complete set of protocol deviations. IPDs are those that may significantly impact the reliability of the study data or that may significantly affect a patient's rights, safety, or well-being.

4.1.3.2 Presentation

Summary tables will be produced to show the number and percentage of patients with any IPD by category of IPD for each study intervention arm. Two separate tables will be produced; one to display IPDs in the randomised period only and another for IPDs at any time while on-treatment

The individual patient data for IPDs will also be listed.

IPD data will be presented for all patients in the FAS.

4.1.4 Demographics

4.1.4.1 Definitions and Derivations

Age is recorded in the CRF, and this is used in the analysis rather than deriving age from the date of birth.

4.1.4.2 Presentation

A summary table of demographic data of age, sex, race, ethnicity and the stratification factors, geographical region and baseline chronic target PN pain score, as well as the combined distribution of the stratification factors, will be produced, and demographic data will be listed.

Demographic data will be presented for all patients in the FAS.

4.1.5 Baseline Characteristics

Characteristics of height, weight, and BSA at baseline will be summarised and listed.

Baseline characteristic data will be presented for all patients in the FAS.

4.1.6 Disease Characteristics

4.1.6.1 Definitions and Derivations

Missing diagnostic dates will be imputed for the summary table as described in Section 3.3.5.2.

4.1.6.2 Presentation

A summary table of characteristics of the time from diagnosis of NF1 to the start of study intervention, time from diagnosis of inoperable PN to the start of study intervention, reasons for the PN not being operable and NF1 diagnostic criteria (eg, any café-au-lait spots, freckles in axilla or groin, etc.) will be produced by study intervention arm. Presence of non-target PNs (yes/no), PN overall location, PN laterality, PN measurability, incomplete coverage, and the reasons for incomplete coverage, together with PN symptoms for both target and non-target PN will be included.

A listing of disease characteristics will be produced for individual patients.

Disease characteristic data will be presented for all patients in the FAS.

4.1.7 Medical History and Concomitant Disease

4.1.7.1 Definitions and Derivations

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA; using the latest MedDRA version).

Any medical history that is ongoing at the time of informed consent is considered a current medical history; otherwise, it will be considered past medical history.

4.1.7.2 Presentation

The number of patients with prior PN therapy, prior radiotherapy, and surgical history will be summarised. A summary table of prior PN therapy will be produced by Anatomical Therapeutic Chemical (ATC) codes, and a summary table of past and current medical history will be produced by system organ class (SOC) and preferred term (PT).

The individual patient data for medical history (past and current), prior PN therapy, prior radiotherapy, and surgical history will also be listed.

Data will be presented for all patients in the FAS.

4.1.8 Prior/Concomitant Medications and Procedures

4.1.8.1 Definitions and Derivations

Medications received prior to, concomitantly, or post study intervention discontinuation are coded using the latest World Health Organisation (WHO) Drug Dictionary.

For the purpose of inclusion in prior and/or concomitant medication or therapy summaries, incomplete medication start and stop dates are imputed as detailed in Section 3.3.5.1.

Prior, concomitant medications are defined based on imputed start and stop dates as follows:

- Prior medications are those taken prior to study intervention with a stop date prior to the first dose of study intervention.
- Concomitant medications are those with a stop date on or after the first dose date of study intervention (and could have started prior to or during treatment).

Missing coding terms are listed and summarised as "Not coded".

Procedures will be coded using MedDRA (using the latest MedDRA version).

4.1.8.2 Presentation

The following summaries will be produced by treatment group:

- Summary of prior medications
- Summary of allowed concomitant medications
- Summary of disallowed concomitant medications

The summary tables of allowed and disallowed concomitant medications will be presented for medications during the on-treatment safety randomised period using the FAS and the on-selumetinib safety period using the Extended Selumetinib FAS. Definitions of the safety randomised period and on-selumetinib safety period are given in Section 3.3.2.3.

All concomitant medications will be listed.

Procedures will not be summarised but only listed.

Medication and procedure data will be presented for all patients in the FAS.

4.1.9 Study Drug Compliance

Study drug compliance is covered in Section 4.6.1.

4.2 Endpoint Analyses

This section covers details related to the endpoint analyses such as primary, secondary, and exploratory.

Statistical category	Endpoint	Analysis Set	Intercurrent event strategy	Population level summary (analysis)	Details in section
To compare the effect of selumetinib relative to placebo by assessment of confirmed partial and complete response rate (ORR) by the end of Cycle 16 using volumetric MRI analysis as determined by ICR (per REiNS criteria) in patients with NF1 who have symptomatic, inoperable PN.					
Primary	Objective response rate (ORR) is defined as the proportion of patients who have confirmed CR or confirmed PR by the end of Cycle 16 as determined by ICR per REiNS criteria	FAS	Post randomised study intervention discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 continuous days of no study intervention): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	ORR compared using a Fisher's exact test. Estimate and two-sided 95% CI based on the Clopper-Pearson method	4.2.1
To compare the effect of selumetinib relative to placebo by assessment of change in chronic target PN pain intensity from baseline in patients with a PAINS-pNF chronic target PN pain score of ≥ 3 at baseline.					
First key secondary	Change from baseline in PAINS-pNF chronic target PN pain intensity score at Cycle 12	Subset of the Pain FAS with at least one evaluable post-baseline assessment	Post-randomised study intervention discontinuation: hypothetical strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; change in chronic PN pain medication: treatment policy strategy; early crossover due to documented progression on imaging: while on-treatment strategy; target PN resection: treatment policy strategy	Difference of the means in the change from baseline in PAINS-pNF chronic target PN pain intensity score at Cycle 12 between selumetinib and placebo using MMRM with geographical region, treatment, cycle as a factor, treatment-by-cycle, and baseline score-by-cycle as interactions and covariates including baseline chronic PAINS-pNF intensity score	4.2.2.1

To compare the effect of selumetinib relative to placebo by assessment of change in HRQoL from baseline in patients with NF1 who have symptomatic, inoperable PN.					
Second key secondary	Change from baseline in PlexiQoLtotal score at Cycle 12	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	Post randomised study intervention discontinuation: hypothetical strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; early crossover due to documented progression on imaging: while on-treatment strategy; target PN resection: treatment policy strategy	Difference of the means in the change from baseline in PlexiQoL total score at Cycle 12 between selumetinib and placebo using MMRM with geographical region, average baseline chronic target PN pain group, treatment, cycle as a factor, treatment-by-cycle, and baseline score-by-cycle as interactions and covariates including baseline PlexiQoL total score	4.2.2.2
To demonstrate the effectiveness of selumetinib by assessment of confirmed partial and complete response rate (ORR) using volumetric MRI analysis as determined by ICR (per REiNS criteria) in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	ORR (single arm]	Selumetinib FAS, Extended Selumetinib FAS	Post randomised study intervention discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	ORR. Estimate and two-sided 95% CI based on the Clopper-Pearson method.	4.2.3
To demonstrate the effectiveness of selumetinib by assessment of DOR in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	Duration of response (DOR) [single arm]	Patients who have a confirmed response in Selumetinib FAS,	Treatment discontinuation: while on-treatment strategy; prolonged study intervention	Median DOR, 25th and 75th percentiles, and 95% CI, and n (%) of patients remaining in	4.2.4

		Extended Selumetinib FAS	interruption (greater than 28 days): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	response and estimated percentage remaining in response + 95% CI using the KM method	
To demonstrate the effectiveness of selumetinib by assessment of PFS in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	Progression-free survival (PFS) [single arm]	Selumetinib FAS, Extended Selumetinib FAS, and FAS (selumetinib arm only)	Treatment discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment; strategy; target PN resection: treatment policy strategy; sbsequent NF1-PN treatment: while on-treatment strategy	Percentage PFS at specified timepoints	4.2.5
To demonstrate the effectiveness of selumetinib by assessment of TTP in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	Time to progression (TTP) [single arm]	Selumetinib FAS, Extended Selumetinib FAS, and FAS (selumetinib arm only)	Treatment discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	Percentage TTP at specified timepoints	4.2.6
To demonstrate the effectiveness of selumetinib by assessment of time to response (TTR) in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	Time to response	Patients who have a confirmed response in	Not applicable for descriptive analysis	Percentage TTR at specified timepoints	4.2.7

	(TTR) [single arm]	Selumetinib FAS, Extended Selumetinib FAS, and FAS (selumetinib arm only)			
To demonstrate the effect of selumetinib relative to placebo by assessment of percentage change from baseline in target PN volume in patients with NF1 who have symptomatic, inoperable PN.					
Secondary	Best percentage change from baseline in target PN volume during the randomised period	Patients in the FAS with measurable target PN at baseline per ICR and at least one post-baseline target PN volume during the randomised period	Treatment discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	Mean difference between selumetinib and placebo in best percentage change from baseline in target PN volume using ANCOVA with covariates including baseline target PN volume, geographical region and average baseline chronic target PN pain group.	4.2.8
To compare the effect of selumetinib relative to placebo by assessment of chronic target PN pain palliation and time to chronic target PN pain palliation in patients with a PAINS-pNF chronic target PN pain score of ≥ 3 at baseline.					
Secondary	Proportion of patients with chronic target PN pain palliation during the randomised period	Subset of the Pain FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	Post randomised study intervention discontinuation: while on-treatment strategy; study intervention interruption greater than 28 days: while on-treatment strategy; change in (chronic PN) pain medication: composite strategy; early crossover due to documented progression on imaging: while on-treatment strategy; target PN resection: treatment policy strategy	Odds ratio between selumetinib and placebo in proportion of patients with chronic target PN pain palliation using random-effect logistic regression with geographical region, treatment and cycle as a factors and covariates including baseline PAINS-pNF pain intensity score and baseline pain medication score	4.2.9

Secondary	Time to chronic target PN pain palliation	Subset of the Pain FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	Post randomised study intervention discontinuation: while on-treatment strategy; study intervention interruption greater than 28 days: while on-treatment strategy; change in pain medication: composite strategy; early crossover due to documented progression on imaging: while on-treatment strategy; target PN resection: treatment policy strategy	KM method, stratified log-rank test, Cox proportional hazards model adjusted for treatment and geographical region as fixed effects and baseline chronic target PN pain intensity score and baseline chronic PN pain medication modified WHO analgesic ladder score as covariates	4.2.10
To compare the effect of selumetinib relative to placebo by assessment of pain medication compared with baseline.					
Secondary	Change from baseline in pain medication use (as reported using the eDiary and as assessed by the investigator) at post-baseline cycles and overall, over the randomised period	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	Post randomised study intervention discontinuation: hypothetical strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; early crossover due to documented progression on imaging: while on-treatment strategy; target PN resection: treatment policy strategy	Odds ratio between selumetinib and placebo in decrease in pain medication using random-effect logistic regression model	4.2.11
To compare the effect of selumetinib relative to placebo by assessment of pain interference compared with baseline.					
Secondary	Change from baseline in PII-pNF pain interference total score at	Subset of the FAS with an evaluable baseline assessment and at least one evaluable	As for the 2nd key secondary objective	Mean difference between selumetinib and placebo in change from baseline in PII-pNF pain	4.2.12

	post-baseline cycles and overall, over the randomised period	post-baseline assessment		interference total score using MMRM with covariates including baseline pain interference total score	
To compare the effect of selumetinib relative to placebo by assessment of physical functioning compared with baseline					
Secondary	Change from baseline in PROMIS Physical Function items at post-baseline cycles and overall, over the randomised period	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	As for the 2nd key secondary objective	Mean difference between selumetinib and placebo in change from baseline in PROMIS Physical Function items and total score using MMRM with covariates including baseline PROMIS item/total score	4.2.13
To compare the effect of selumetinib relative to placebo by further assessment of HRQoL compared with baseline					
Secondary	Change from baseline in the Skin Sensations domain from the PedsQL (NF1 module acute version 3.0 – adult report) at post-baseline cycles and overall, over the randomised period	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	As for the 2nd key secondary objective	Mean difference between selumetinib and placebo in change from baseline in PedsQL Skin Sensations domain score using MMRM with covariates including baseline PedsQL score	4.2.14
To compare the effect of selumetinib relative to placebo by assessment of health status compared with baseline.					
Secondary	Change from baseline in EQ-5D-5L at post-baseline cycles and overall, over the	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	As for the 2nd key secondary objective	Mean difference between selumetinib and placebo in change from baseline in EQ-5D-5L score using MMRM with covariates including	4.2.15

	randomised period			baseline EQ-5D-5L score	
	Change from baseline in EQ-VAS at post-baseline cycles and overall, over the randomised period	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	As for the key secondary objective	Mean difference between selumetinib and placebo in change from baseline in EQ-VAS score using MMRM with covariates including baseline EQ-VAS score	4.2.15
To evaluate PFS during the randomised period on selumetinib and placebo in patients with NF1 who have symptomatic, inoperable PN					
Exploratory	PFS during the randomised period in selumetinib and placebo	FAS	Treatment discontinuation: while on-treatment strategy; prolonged study intervention interruption (greater than 28 days): while on-treatment strategy; target PN resection: treatment policy strategy; subsequent NF1-PN treatment: while on-treatment strategy	Percentage PFS at specified timepoints	4.2.16
To compare the effect of selumetinib relative to placebo by assessment of spike target PN pain intensity.					
Exploratory	PAINS-pNF spike target PN pain intensity in all post-baseline cycles during the randomised period	Subset of the FAS with an evaluable baseline assessment and at least one evaluable post-baseline assessment	As for the 1st key secondary objective	Mean difference between selumetinib and placebo in maximum scores during the randomised period using MMRM with geographical region as a factor and covariates including baseline spike pain intensity score	4.2.17
To evaluate the efficacy under fed dosing of selumetinib.					
Exploratory	Percentage change from baseline in target PN at	Fed FAS	Not applicable for descriptive analysis	Descriptive analysis	4.2.18

	Cycle 30 and Cycle 36				
To evaluate the Patient Global Impression of Severity (PGIS) of symptoms and the Patient Global Impression of Change (PGIC) in symptoms					
Exploratory	PGIS and PGIC	FAS	Not applicable for descriptive analysis	Descriptive analysis	4.2.19
To characterise the PK of selumetinib and N-desmethyl selumetinib.					
Secondary	Plasma concentration data and PK parameters	PK analysis set	All data from the PK analysis set are included.	Descriptive analysis	4.4

ANCOVA, analysis of covariance; CI, confidence interval; CR, complete response; DOR, duration of response; EQ-5D-5L, EuroQoL five dimensions, five level health state utility index; EQ-VAS, EuroQol-visual analogue scale; FAS, full analysis set; HRQoL, health related quality of life; ICR, independent central review; KM, Kaplan-Meier; MMRM, Mixed Model Repeated Measures; MRI, magnetic resonance imaging; NF1, neurofibromatosis type 1; ORR, objective response rate; PAINS-pNF, PAin INtensity Scale for Plexiform Neurofibroma; PedsQL, Paediatric Quality of Life Inventory; PFS, progression free survival; PGIC, Patient Global Impression of Change; PGIS, Patient Global Impression of Severity; PII-pNF, Pain interference index – plexiform neurofibroma; PK, pharmacokinetic; PlexiQoL, Plexiform Neurofibroma Quality of Life scale; PN, plexiform neurofibroma; PR, partial response; PROMIS, patient-reported outcomes measurement information system; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis; TTP, time to progression; TTR, time to response.

4.2.1 Primary Endpoint

4.2.1.1 PN Volumes and Derivation of REiNS Tumour Response

For all patients, the REiNS tumour data will be used to determine each patient's visit response according to the REiNS criteria. Tumour assessments will be measured by volumetric MRI.

Baseline radiological tumour assessments will be performed during the 28-day screening period. Unless clinically indicated otherwise, tumour assessments of the target and non-target PN (if relevant) will be obtained at screening and every 4 cycles (16 ± 1 weeks) relative to the date of the first dose for the first 24 cycles. From the end of Cycle 24, tumour assessments will be performed every 6 cycles (24 ± 1 week) as long as the patient remains on study intervention.

Volumetric MRI assessment will be performed by AstraZeneca's appointed imaging contract research organisation (CRO) according to the REiNS criteria. A double read of all MRI scans will be performed. The ICR reviewers will be blinded to the study intervention group. Further details of the ICR will be documented in the independent review charter (IRC).

All volumetric MRI scans for all patients (including those at unscheduled visits or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca-appointed imaging CRO for central analysis. The investigator will select the target PN and non-target PN (if relevant) based on the results of the screening PN assessments. Details about target PN location and non-target PN location (if relevant) will also be collected and sent to the imaging CRO to ensure that the independent reviewer also follows the most clinically relevant PN. The imaging CRO will determine the tumour measurements of the target PN and non-target PN.

If an unscheduled assessment is performed (e.g., to investigate clinical signs/symptoms of progression) and the patient has not progressed, the subsequent assessments at the next scheduled visit must be performed. This schedule is to be followed in order to minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

To determine the level of response, the TL and NTL volume will be assessed as the average PN volumes from the double read, and follow-up scans will be compared to the baseline scan or the scan at the time of best response after documenting a CR/PR for the same PN.

If the interim analysis takes place, then:

- At DCO1, if only part of a PN was imaged on post-baseline scans, the ICR will outline that same part of the PN (partial PN volume) at all timepoints. The detailed handling of partial volumes in the ICR data can be found in the IRC.

- At the primary analysis (DCO2), two sets of parameters will be produced from the ICR due to different handling of the partial volumes. The first set of parameters will contain results that are not adjusted for partial volumes occurring after DCO1, that is, any partial volume for timepoints after DCO1 which captures a different PN part than the outlined PN part used at DCO1 will be assessed as NE. Any partial volume for timepoints after DCO1 which captures the same PN part as the outlined PN part used for the partial volume at DCO1 will be assessed. The second set of parameters will contain results that are adjusted for partial volumes after DCO1, that is, if there is a partial volume after DCO1 which captures a different PN part than the outlined PN part used at DCO1, then all timepoints will be adjusted to outline the same part of the PN at all timepoints.

In order to keep responses of the patients included in both the interim and the primary analyses consistent across review periods (i.e., responses reported at DCO1 will remain unchanged at DCO2 for patients having partial volumes after DCO1), a third set of parameters will be programmatically derived by proportionally scaling up partial volumes after DCO1. At each post-baseline visit after DCO1, partial volumes will be re-derived as follows: (post-baseline partial volume * baseline full volume) / baseline partial volume. Partial volumes can be selected from the second set of parameters, while baseline full volume will be provided in the first set of parameters.

All the analyses using PN volumes performed at DCO2 will use the just mentioned scaled partial volumes after DCO1 data (third set of parameters) and the non-adjusted partial volume after DCO1 (first set of parameters) will be supportive.

Otherwise, if the interim analysis does not take place, then at the primary analysis (DCO2), all the analyses using PN volumes performed at DCO2 will use the first set of parameters.

At the final DCO, two sets of parameters will be produced from the ICR due to different handling of the partial volumes. The first set of parameters will contain results that are not adjusted for partial volumes occurring after DCO2, that is, any partial volume for timepoints after DCO2 which capture a different PN part than the outlined PN part used at DCO2 will be assessed as NE. The second set of parameters will contain results that are adjusted for partial volumes after DCO2, that is, if there is a partial volume after DCO2 which captures a different PN part than the outlined PN part used at DCO2 then all timepoints will be adjusted to outline the same part of the PN at all timepoints.

At the final DCO, similar to the approach at DCO2, post-DCO2 partial volumes will be scaled up to make sure that responses previously reported will remain unchanged and post-DCO2 partial volumes are proportionally adjusted. Each post-DCO2 partial volume will be multiplied by baseline partial volume and then divided by baseline full volume.

The primary endpoint analysis performed at the final DCO will use the scaled partial volumes after DCO2 data (third set of parameters), and a supplementary analysis will be performed using the unadjusted partial volume after DCO2 (first set of parameters).

All the analyses using PN volumes performed at final DCO will use the (third set of parameters with scaled partial volumes after DCO2 data, while the non-adjusted partial volume after DCO2 (first set of parameters) will be supportive.

Using the REiNS criteria, a change in volume from baseline of 20% is used to indicate a clinically significant increase or decrease of PN volume from baseline. For all patients, the target PN (see Section 4.2.1.2) and non-target PN (see Section 4.2.1.3), if applicable, will be assessed, both will be given a response (see Table 5 and Table 6), and any new PNs (see Section 4.2.1.4) will be recorded.

If the two independent readers disagree on the presence of a new lesion, an adjudication will be performed, and the adjudicator decision regarding the presence or absence of the new lesion will be used for the assessment. At each visit, patients will be programmatically assigned an overall REiNS visit response of CR, PR, stable disease (SD), or progressive disease (PD) using the information from target PNs, non-target PNs, and new PNs and depending on the status of their disease compared with baseline and previous assessments (see Table 7). If a patient has a PN assessment that cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE).

4.2.1.2 Target PN Response

The target PN is selected at screening. The target PN will be defined as the most clinically relevant PN that is also measurable by volumetric MRI analysis. If there is a second PN that is also considered clinically relevant and is measurable, this may be identified as a non-target PN at baseline; only one non-target PN can be selected (see Section 4.2.1.3).

The target PN visit response will be derived as per Table 5 using the average over both baseline PN volume assessments and the average over both visit PN volume assessments from ICR. Partial target PN volumes will be handled following the strategy in Section 4.2.1.1. Table 5 is based on the recommendations from Dombi et al 2013.

Table 5: Target PN Visit Responses (REiNS)

Visit responses	Description
Complete response (CR)	Disappearance of the target PN.
Partial response (PR)	Decrease in the volume of the target PN by 20% or more compared to baseline.
Progressive disease (PD)	Increase in the volume of the target PN by 20% or more compared to baseline or the time of best response after documenting a PR. The appearance of new PN (with the exception of new discrete subcutaneous neurofibromas), which is unequivocally and completely distinct and separate from the target PN and the non-target PN, or unequivocal progression of an existing non-target PN is also considered PD. See Table 6 for definition of unequivocal progression of non-target PN.
Stable disease (SD)	Insufficient volume change to qualify for either PR or PD (a < 20% increase or < 20% decrease in the volume of the target PN).
Not evaluable (NE)	Data unavailable for target PN assessment.

PD, progressive disease; PN, plexiform neurofibroma; PR, partial response; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

Confirmed CR (cCR) and confirmed PR (cPR) will be derived as CR/PR achieved on a consecutive scan within 3 to 6 months with no missed or non-evaluable visits as per assessment schedule.

4.2.1.3 Non-Target PN Response

Only one non-target PN may be selected. Non-target PN must also be clinically relevant and measurable by volumetric MRI. Partial non-target PN volumes will be handled following the strategy in Section 4.2.1.1. This section provides the definitions of the criteria used to determine and record a response for the non-target PN at each MRI assessment.

Non-target response will be derived using the PN volume over both ICR assessments as follows:

Table 6: Non-Target PN Visit Responses (REiNS)

Visit responses	Description
Progressive disease (PD)	Unequivocal progression of an existing non-target PN. In this study, unequivocal progression is defined as an increase in the volume of the non-target PN by 20% or more compared to baseline.
Non-progressive disease (non-PD)	Insufficient volume change to qualify for PD.
Not applicable (NA)	No non-target PN recorded at baseline.
Not evaluable (NE)	Data unavailable for non-target PN assessment.

PD, progressive disease; PN, plexiform neurofibroma; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

Table 6 is based on the recommendations from Dombi et al 2013.

4.2.1.4 New PN Response

Details of any new PNs will also be recorded in the CRF with the date of the first scan that revealed the new PN(s). The appearance of a new PN (with the exception of new discrete subcutaneous neurofibromas), which is unequivocally and completely distinct and separate from the target PN and the non-target PN, if applicable, or unequivocal progression of an existing non-target PN also qualifies for PD.

The new PN must be confirmed by scans with PN details recorded in the CRF. Once a new PN has been identified, at subsequent timepoints, it will be classified as either absent, present, or NE (in the case where image quality issues prevent a full assessment of the previously identified new PN). For new PNs assessed by ICR, if during the double read both independent reviewers disagree on the verdict, a third radiologist will perform adjudication. The assessment of the adjudicator will be used to define the overall visit response as described in the next section. Further details on the adjudication process can be found in the IRC.

4.2.1.5 Overall Visit Response

Table 7 defines how the previously defined target PN and non-target PN visit responses will be combined with new PN information to give an overall visit response.

Table 7: Overall Visit Responses

Target PN	Non-target PN	New PN	Overall visit response
CR	Non-PD or NE or NA	No or NE	CR
PR	Non-PD or NE or NA	No or NE	PR
SD	Non-PD or NE or NA	No or NE	SD
PD	Any	Any	PD
NE	Non-PD or NE or NA	No or NE	NE
Any	Any	Yes	PD
Any	PD	Any	PD

CR, complete response; NA, not applicable; NE, not evaluable; PD, progressive disease; PN, plexiform neurofibroma; PR, partial response; SD, stable disease.

The overall visit response will be derived programmatically based on the derived target PN and non-target PN responses as well as the adjudicated new PN response as described in the previous sections. In addition, the ICR will provide two overall visit responses, one response for each reviewer of the double read. The programmatically derived overall visit

response will be used in the summary tables and figures. The ICR overall visit response will only be listed.

4.2.1.6 MRI Assessments During Prolonged Study Intervention Interruption

Patients may require prolonged interruption of treatment to undergo a surgery or manage treatment toxicity. A prolonged study intervention interruption is defined as greater than 28 continuous days of no study intervention.

If a while-on-treatment strategy to prolonged study intervention interruption is chosen, the following rules will be applied for MRI volumetric assessments performed during a prolonged interruption of the study intervention prior to the derivation of the target PN response:

- Any MRI assessments performed within the first 28 continuous days of prolonged study intervention interruption will be included in the analysis.
- Any MRI assessments performed after the first 28 continuous days of study intervention interruption and during ongoing prolonged study intervention interruption will be excluded from the analysis.
- Any MRI assessments performed at a timepoint when the patient has received less than 28 days of study intervention since study intervention recommencing following the prolonged study intervention interruption will be excluded from the analysis.
- Any MRI assessments performed at a timepoint when the patient has received greater than or equal to 28 days of study intervention since study intervention recommencing will be included in the analysis.

4.2.1.7 Intercurrent Events

The intercurrent event strategy for the primary estimand is described in Table 8.

Table 8: Intercurrent Event Strategy for Primary Estimand

Intercurrent event	Strategy	Details
Post randomised study intervention discontinuation (including early crossover from placebo[1] and progression) prior to the	While on-treatment strategy	Scans after the randomised study intervention discontinuation will not be included in the analysis.

end of cycle 16 due to any reason		
Prolonged study intervention interruption (greater than 28 days[2])	While on-treatment strategy	Scans after the first 28 days of a prolonged study intervention interruption until study intervention has recommenced for at least 28 days will not be included in the analysis. See Section 4.2.1.6 for further details.
Target PN resection	Treatment policy strategy	Scans after a target PN resection will be included in the analysis.
Subsequent NF1-PN treatment	While on-treatment strategy	Scans after subsequent NF1-PN treatment will not be included in the analysis.

[1] Crossover from placebo to selumetinib prior to Cycle 12 in case of progression on imaging as determined by ICR per REiNS criteria.
[2] 28 Continuous days of no study intervention
NF1, neurofibromatosis type 1; PN, plexiform neurofibroma.

4.2.1.8 ORR by the End of Cycle 16 Definition

The ORR by the end of Cycle 16 will be defined as the proportion of patients who have a cCR (defined as disappearance of the target PN, confirmed by a consecutive scan within 3 to 6 months after the first response; see Section 4.2.1.2) or cPR (defined as a target PN volume decrease ≥ 20% compared to baseline, confirmed by a consecutive scan within 3 to 6 months after the first response; see Section 4.2.1.2) by the end of Cycle 16 as determined by ICR per REiNS criteria.

The ORR by the end of Cycle 16 will be derived using while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

Any cCR or cPR at or prior to the end of Cycle 16 which occurred after a subsequent NF1-PN treatment (following study intervention discontinuation) will not be included in the numerator for the ORR calculation (where the FAS will be the denominator). Patients with no post-baseline MRI assessments will be considered as non-responders (i.e., not having a cCR or cPR) if they are included in the FAS.

The confirmation by a consecutive scan within 3 to 6 months will be derived as CR/PR achieved on consecutive visit with no missed visits as per assessment schedule. If MRI assessments are excluded due to prolonged study intervention interruption (as described in Section 4.2.1.6), then the visit will be considered as missed visit.

4.2.1.9 Primary Analysis of the Primary Endpoint

The primary analysis will include all randomised patients (FAS). Data obtained using on-treatment MRI volumetric assessments (see Section 3.3.2.1), from first dose up until progression (if progression occurs prior to the end of Cycle 16), or the last evaluable assessment up to and including the end of Cycle 16 assessment in the absence of progression will be included in the assessment of the ORR and by excluding MRI during prolonged study intervention interruption (while-on-treatment strategy). For the complete intercurrent event strategy, refer to Section 4.2.1.7.

If the interim analysis takes place, the ORR by the end of Cycle 16 at DCO2 will be analysed using the ICR data with the scaled partial volume after DCO1 as described in Section 4.2.1.1 (third set of parameters).

The formal statistical analysis will be performed to test the following hypotheses for the primary analysis:

- H_0: ORR selumetinib = ORR placebo.
- H_1: ORR selumetinib ≠ ORR placebo.

The ORR will be compared at the landmark end of Cycle 16 between selumetinib versus placebo using a Fisher's exact test with the significance level stated in Section 3.3.7. The ORR by the end of Cycle 16 in each treatment group will be presented with corresponding 2-sided exact 95% confidence interval (CI) based on the Clopper-Pearson method (Clopper and Pearson 1934). The risk difference and 95% CI based on the Miettinen-Nurminen (score) method (Miettinen and Nurminen 1985) will also be presented.

4.2.1.10 Sensitivity Analyses of the Primary Endpoint

Sensitivity analyses of the primary endpoint will be performed:

The ORR by the end of Cycle 16 will be calculated based on all MRI volumetric assessments without excluding scans during study intervention interruption. MRI volumetric assessments data will be defined as data after the date of first dose of study intervention until the study intervention discontinuation date or DCO, whichever occurs first. The analysis will follow the same methods as for the primary analysis.

4.2.1.11 Supplementary Analyses of the Primary Endpoint

Supportive evidence of the primary endpoint of ORR will include the following:

- If the interim analysis takes place, a supplementary analysis performed at DCO2 using the ICR data with the non-adjusted partial volume after DCO1 (if applicable) as described in Section 4.2.1.1. The analysis will follow the same methods as for the primary analysis.

- Target PN volumes (absolute value, changes and percentage changes from baseline) will be summarised descriptively over time (even beyond Cycle 16) in the FAS. A line plot will display individual target PN volumes over time (even beyond Cycle 16), and a line plot will show individual percentage target PN volume changes over time (even beyond Cycle 16). Percentage change from baseline in target PN volume over time (even beyond Cycle 16) will be also presented using box plots. For patients randomised to placebo, this will include data prior and post selumetinib intervention initiation. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo patients who have already crossed-over to selumetinib.
- The best objective response (BOR) by Cycle 16 by randomised treatment group will be calculated based on the overall visit responses from each on-treatment MRI assessment, as described in Section 0 for the FAS. It is the best response a patient has had by the end of Cycle 16 following the start of intervention, but prior to starting any subsequent NF1-PN therapy and up to and including progression/death or the last evaluable MRI assessment in the absence of progression/death. Patients with no post-baseline MRI assessments will be considered as non-responders. The categorisation of BOR will be based on REiNS using the following response categories: cCR, CR, cPR, PR, SD, PD, and NE. The different ways of achieving a BOR of cCR, CR, cPR, or PR are shown in Table 11. The intercurrent event strategy described in Table 8 will be followed.

4.2.2 Key Secondary Endpoints

4.2.2.1 PAINS-pNF chronic target PN pain intensity score

4.2.2.1.1 Definition

The 1st key secondary estimand is the difference of the means in the change from baseline in PAINS-pNF chronic target PN pain intensity score at Cycle 12 between selumetinib and placebo amongst patients with a PAINS-pNF chronic target PN pain intensity score ≥ 3 at baseline (Pain FAS) and at least one post-baseline average cycle PAINS-pNF chronic target PN pain intensity score.

The randomised period is defined in Section 3.3.2.2.

4.2.2.1.2 Derivations

The response variable for a cycle will be the average cycle PAINS-pNF chronic target PN pain score change from baseline, defined as the average of the available daily PAINS-pNF chronic target PN pain scores for the 28 days up to and including the last day of the cycle minus the baseline PAINS-pNF chronic target PN pain score.

The average cycle PAINS-pNF chronic target PN pain score will only be derived if the patient meets the criteria of having at least 4 daily pain scores in at least 3 non-overlapping 7-day periods in the 28-day cycle. Baseline PAINS-pNF chronic target PN pain is defined as the average of the available daily PAINS-pNF chronic target PN pain scores in the screening period. During this time, patients must complete their pain diary for at least 4 days in at least 2 non-overlapping 7-day periods in order to determine the patient's average baseline chronic target PN pain intensity score.

4.2.2.1.3 Handling of Dropouts and Missing Data

Missing PAINS-pNF chronic target PN pain scores after treatment discontinuation are handled by the "while on-treatment" approach, using the change from baseline in PAINS-pNF of cycles where a patient was on treatment, as described in the previous section.

4.2.2.1.4 Intercurrent Events

The intercurrent event strategy for the 1st key secondary estimand is described in Table 9.

Table 9: Intercurrent Event Strategy for 1st Key Secondary Estimand

Intercurrent event	Strategy	Details
Changes to patients' chronic PN pain medication	Treatment policy strategy	PAINS-pNF chronic target PN pain intensity scores collected following changes to patients' chronic PN pain medication will be included in the analysis.
Randomised study intervention discontinuation	Hypothetical strategy	PAINS-pNF chronic target PN pain intensity scores following randomised study intervention discontinuation will not be collected and will be modelled through direct likelihood techniques.
Early crossover from placebo to selumetinib in patients with documented progression on imaging (as determined by ICR per REiNS criteria)	While on-treatment strategy	PAINS-pNF chronic target PN pain intensity scores after early crossover from placebo to selumetinib will be set to missing.
Prolonged study intervention interruption (defined as a study intervention interruption greater than or equal to 28	While on-treatment strategy	The following rules will be followed to align with the while on-treatment strategy: - PAINS-pNF chronic target PN pain intensity scores during the first 28 days

continuous days of no study intervention)		of prolonged treatment interruption will be included. - PAINS-pNF chronic target PN pain intensity scores after the first 28 days of prolonged treatment interruption will be excluded. - PAINS-pNF chronic target PN pain intensity scores from the day of study intervention recommencement will be included.
Target PN resection	Treatment policy strategy	PAINS-pNF scores after a target PN surgical resection will be included in the analysis.

ICR, independent central review; PAINS-pNF, PAin INtensity Scale for Plexiform Neurofibroma; PN, plexiform neurofibroma; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

4.2.2.1.5 Primary Analysis of the 1st Key Secondary Endpoint

The difference in the mean change from baseline in PAINS-pNF chronic target PN pain score at Cycle 12 between treatment groups (selumetinib minus placebo), standard error, 95% CI and p-value will be assessed by least-squares (LS) means of a MMRM with randomised study intervention, cycle number, and geographical region as categorical fixed effects; the baseline PAINS-pNF chronic target PN pain intensity score as a continuous covariate; and treatment-by-cycle number and baseline PAINS-pNF score-by-cycle number interactions. Parameters will be estimated with the restricted maximum likelihood (REML) approach (Brown and Prescott 2006), and the Kenward-Roger approximation is used to estimate the degrees of freedom.

The analysis will include PAIN-FAS patients with at least one post-baseline average cycle PAINS-pNF chronic target PN pain intensity score. Data obtained during the randomised period (see Section 3.3.2.2) will be included in the analysis, and the intercurrent event strategy is described in, Section 4.2.2.1.4.

The analysis will be conducted using PROC MIXED in SAS.

Missing PAINS-pNF chronic target PN pain intensity scores will be modelled through direct likelihood techniques, which use the information from the observed pain intensity scores via the within-patient correlation structure (covariance matrix) to provide information about the unobserved pain intensity scores. Random (subject-specific) effects will be not explicitly modelled, and an unstructured covariance matrix will be used to

model the within-patient correlations between the repeated measurements and to allow for unequal treatment variance.

An unstructured modelling of within-patient correlations removes one layer of assumptions on the random effects and often provides the best fit to the data. In case of missing data, the use of an unstructured covariance matrix in the MMRM can guarantee a more robust control of type I error by better detecting even small treatment differences compared with the single imputation methods such as last observation carried forward, even in case of unequally distributed dropout rates across treatment arms and regardless of the form of the true covariance matrix (Mallinckrodt et al 2008). If the fit of the unstructured covariance structure fails to converge, every attempt should be made to ensure convergence is obtained from the unstructured correlation structure. Parameters will be tried to be estimated by means of the Fisher's scoring algorithm rather than the default Newton-Raphson algorithm.

If convergence is still not reached after the change in the interaction algorithm, a more parsimonious model will be implemented using one of the following structured covariance structures in order until convergence is reached: toeplitz with heterogeneity, autoregressive with heterogeneity, toeplitz, autoregressive, and compound symmetry. All main effects and the interaction terms will remain in the model, regardless of significance.

In addition to the Cycle 12 estimate, the model will present LS mean estimates for each treatment group and their differences, standard errors, 95% CIs, and p-values (where applicable and with statistical significance effects being interpreted as nominally significant) for mean changes from baseline to each cycle during the randomised period (based on the treatment-by-cycle interaction coefficient) and the average over the randomised period (based on the treatment coefficient). A plot of the LS means accompanied by the 95% CI will be produced. Variables listed as categorical in the list above will be included in the CLASS statement of the procedure. The unique patient identifier will also be included as a class variable. A REPEATED statement over the visits will be included with the unique patient identifier as the SUBJECT variable in the REPEATED statement.

4.2.2.1.6 Sensitivity Analyses of the 1st Key Secondary Endpoint

Analyses to assess the robustness to assumptions of the 1st key secondary estimand.

The analyses use MI and apply the planned primary analysis to the imputed outcomes, combining the statistics using Rubin's rules. These three supplementary analyses only vary as to the scenarios implemented for the intercurrent events of treatment discontinuation and change in pain medication, as follows:

- Multiply impute reversion to baseline after discontinuation of treatment up to Cycle 12: this analysis provides an estimate for the estimand that targets the hypothetical approach

using the assumption that subjects who discontinue treatment gain no improvement in changes from baseline.

- MI using attributable and non-attributable dropout reasons: hypothetical strategy with missing not at random (MNAR) with reversion to baseline for scores after study intervention discontinuation due to attributable dropout reasons and MAR after study intervention discontinuation due to non-attributable dropout reasons. The reason for study intervention discontinuation will be categorised as attributable or non-attributable as outlined in Section 3.3.9.
- Multiply impute reversion to baseline for a cycle in which there is an increase in chronic PN pain medication (see Section 3.3.8 for the definition of this: this analysis provides an estimate for the estimand that targets the treatment policy approach using the assumption that, for the cycle in which an increase in chronic PN pain medication occurs, a subject's health experience is such that they lose any improvement over baseline that they may have gained in that cycle.

The planned primary analysis will be repeated for the following 3 supplementary analyses to assess the robustness to the missing cycle score rule of having at least 4 daily pain scores in at least 3 non-overlapping 7-day periods in the 28-day cycle:

- Post-baseline scores will only be derived if the patient meets the criteria of having at least 4 non-missing daily pain scores per week during each of the 4 weeks of the 28-day cycle, and baseline scores will only be derived if the patient meets the criteria of having at least 4 non-missing daily pain scores in at least 2 non-overlapping 7-day periods (first secondary definition).
- Both baseline and post-baseline scores will only be derived if the patient meets the criteria of having at least 4 non-missing daily pain scores in at least 3 non-overlapping 7-day periods in the 28-day cycle (second secondary definition).
- Both baseline and post-baseline scores will only be derived if the patient meets the criteria of having at least 4 non-missing daily pain scores in at least 2 non-overlapping 7-day periods in the 28-day cycle (third secondary definition).

4.2.2.1.6.1 Multiple Imputation

Before implementing multiple imputation methods, the first step will be to identify the first cycle(s) eligible for reversion-to-baseline. For the intercurrent event of treatment discontinuation, if a subject has a missing PAINS-pNF chronic target pain cycle score due to not having at least 4 daily scores in at least 3 non-overlapping 7-day periods in the 28-day cycle (per Section 4.2.2.1.2) in the same cycle as treatment discontinuation occurs, then the missing cycle score for that cycle and all further cycles up to and including cycle 12 will be imputed. Otherwise, the observed data for that cycle will be retained and the cycle

score will be imputed from the next cycle after treatment discontinuation occurs onwards up to and including cycle 12.

For the intercurrent event of increase in pain medication, the first date of increase in chronic pain medication meeting the minimum requirement for increase per section 3.3.8 will be identified. All daily PAINS-pNF scores on and after the first date of increase in pain medication will be set to missing for the cycle in which the increase in pain medication occurs. After setting these daily values to missing, if there is sufficient (i.e., meets the criteria of needing at least 4 daily scores in at least 3 non-overlapping 7-day periods in the 28-day cycle [per Section 4.2.2.1.2]) observed daily data prior to increase in medication to derive the cycle PAINS-pNF score, then these data will be used to derive the cycle score. This derived score will be used to overwrite the actual observed cycle score. Otherwise, the cycle score will be set to missing to be imputed under MNAR using reversion-to-baseline. This method will be repeated for each cycle in which an increase in pain medication occurs up to and including cycle 12. The resulting dataset from repeating this process can contain non-monotone missing data for 2 reasons: (i) missing score in cycle x due to an increase in pain medication (and a cycle score cannot be derived using daily data prior to the intercurrent event), but cycle score is not missing for cycle x+1 and; (ii) missing cycle score due to reasons not related to an increase in pain medication (e.g. subject's eDiary broke during the cycle).

MI methods can be summarised in three steps: imputation phase, analysis phase, and pooling phase.

- Imputation Phase

By means of the SAS procedure PROC MI, 25 copies (assuming 10% to 35% of missing data) of the original dataset will be simulated, where missing continuous endpoint values will be replaced with a 2-step imputation process.

By means of the MI SAS procedure PROC MI, 25 copies of the original dataset will be simulated with random seed = 54321, where missing endpoint values will be replaced with a 2-step imputation process.

The first step will be to impute non-monotone missing data ("holes" while on-treatment), which will be imputed under the MAR assumption by means of a multivariate joint Gaussian imputation model using the stochastic Markov Chain Monte Carlo (MCMC) method.

For the intercurrent event of increase in pain medication, as described above, there can be non-montone missing data due to the intercurrent event. Therefore, after implementation of the MCMC method to impute non-monotone missing data, these values (where intercurrent

event has occurred) will be set back to missing to be imputed using the reversion-to-baseline approach described below.

The second step is imputing the remaining missing values using a reversion-to-baseline approach. The reversion-to-baseline approach is based on the MNAR assumption that, after the occurrence of the intercurrent event, patients in the experimental arm will no longer benefit from treatment and will be assumed to revert their efficacy to baseline.

Reversion-to-baseline will be implemented following the approach in Ratitch et al. (2013). For the template code to implement the reversion-to-baseline, see the supplementary appendix to Ratitch et al 2013), under the code heading "Pattern imputation based on reasons for discontinuation".

After the cycle(s) eligible for imputation using reversion-to-baseline have been identified, per the approach described above, for each of the 25 copies of the original dataset following imputation of non-monotone missing data with MCMC (and setting back to missing in case of increase in pain medication), the next step will be to create the input datasets containing all subjects and available baseline data to proceed with reversion-to-baseline. For subjects experiencing the intercurrent event, there will be records corresponding to non-missing baseline data and one or more records with missing values pertaining to each of the cycles to be imputed after the intercurrent event occurs, as described above. These missing values will be imputed using PROC MI by values sampled from the estimated model of the baseline distribution. The model will contain the baseline variables geographical region and baseline PAINS-pNF chronic target PN pain intensity score. For sensitivity analysis 1 (treatment discontinuation) and 3 (increase in pain medication), once missing data have been imputed, proceed to the analysis phase, described below.

For sensitivity analysis 2 (dropout due to attributable and non-attributable reasons), further steps are required to impute missing data for subjects who discontinued due to non-attributable dropout reasons. These missing data will be imputed under MAR using a sequential (visit-by-visit) regression model. The input dataset will be the 25 copies of the original dataset following imputation of non-monotone missing data containing all subjects. PROC MI will be used to impute missing data following discontinuation due to non-attributable dropout reasons up to and including Cycle 12. Once these missing data have been imputed, data for subjects who dropped out due to attributable dropout reasons will be blanked out (ie, set to missing). Imputed data for these subjects, obtained via the reversion-to-baseline approach under MNAR, will be merged in by subject ID and imputation number to overwrite the blanked-out data. This will create 25 fully imputed datasets to be used in the analysis phase described below.

For non-monotone imputation and imputation under MAR using sequential regression, all imputation models should be at least as complex as the analysis model. All fixed effects used

to adjust the analysis model will therefore be included in the MCMC model and the sequential regression models. Since the MCMC method models all variables as continuous, it will be necessary to code all factors as numerical.

- Analysis Phase

After the imputation phase, the same MMRM of the main analysis will be applied to the 25 fully imputed datasets to estimate the LS means and LS mean differences, with one set of estimates for each imputed dataset.

- Pooling Phase

The treatment effects and treatment differences for each visit from the results from 25 datasets will be combined to produce the study results, using the SAS procedure PROC MIANALYZE.

4.2.2.1.7 Supplementary Analyses of the 1st Key Secondary Endpoint

Additionally, PAINS-pNF chronic target PN pain intensity actual scores as well as the change from baseline at each cycle will be summarised descriptively by visit and study intervention arm for FAS and Pain FAS. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib.

Summary statistics will include n, mean, standard deviation, min, Q1, median, Q3, and max, where n is the number of subjects contributing to the summary statistics for the visit.

This descriptive analysis in the Pain FAS will be repeated only for the randomised period by considering the three missing cycle score rules mentioned in Section 4.2.2.1.6.

Summary statistics will also be presented by attributable versus non-attributable reasons (see Section 3.3.9) only for the randomised period.

Plots of the mean scores, as well as the changes from baseline and the associated 95% CIs by cycle will be produced for the randomised period.

The empirical cumulative distribution function (eCDF) of change from baseline by study intervention arm will be plotted at Cycle 12 for the Pain FAS. The eCDF will be the cumulative percent frequency (Cum_Pct) of the change from baseline in the PAINS-pNF chronic target PN pain intensity actual score by study intervention arm saved by default in the output dataset of the SAS procedure proc freq.

4.2.2.2 Plexiform Neurofibroma Quality of Life Scale

4.2.2.2.1 Definition

The PlexiQoL is a patient-derived quality of life (QoL) measure specific to adults with NF1-associated PNs. It assesses the impact of PNs on patients' ability to fulfil their human needs. The scale adopts the needs-based model of QoL and joins a large portfolio of high-quality outcome measures that are widely used in international clinical studies (see for example Marzo-Ortega et al 2005; McKenna et al 2006; Tay et al 2011). The scientifically rigorous methodology employed in the development of the PlexiQoL ensures accurate and valid measurement of the impact of the condition and the value of potential treatments to the lives of patients.

The PlexiQoL questionnaire is planned to be collected at baseline and Cycles 2, 4, 8, and 12 during the randomised period as well as at Cycles 16, 20, 24, 30, and every 6 weeks onwards as per the Schedule of Activities (see Section 1.3 of the CSP).

The randomised period is defined in Section 3.3.2.2.

4.2.2.2.2 Derivations

The measure consists of 18 dichotomous items with 0 ="Not True" and 1 = "True". Scores are summed to a maximum of 18, with lower scores indicating better quality of life. According to the PlexiQoL user manual, it is recommended that for respondents with between one and three missing responses (that is, cases with no more than 20% missing data), the total score is calculated as follows: T=18*x / (18-m), where T is the final total score, x is the item summation score, and m is the number of missing items. Cases with more than three missing responses cannot be allocated a total score.

The change from baseline to the end of each cycle in PlexiQoL total score will be derived as the PlexiQoL total score at the specific cycle minus baseline PlexiQoL total score.

4.2.2.2.3 Intercurrent Events

The intercurrent event strategy for the 2nd key secondary estimand is described in Table 10.

Table 10: Intercurrent Event Strategy for the 2nd Key Secondary Estimand

Intercurrent event	Strategy	Details
Randomised study intervention discontinuation	Hypothetical strategy	PlexiQoL total scores following randomised study intervention discontinuation will not be collected and will be modelled through direct likelihood techniques.

Early crossover from placebo to selumetinib in patients with documented progression on imaging (as determined by ICR per REiNS criteria)	While on-treatment strategy	PlexiQoL total scores after early crossover from placebo to selumetinib will be set to missing.
Prolonged study intervention interruption (defined as a study intervention interruption greater than or equal to 28 continuous days of no study intervention)	While on-treatment strategy	The following rules will be followed to align with the while on-treatment strategy: - PlexiQoL total scores during the first 28 days of prolonged treatment interruption will be included. - PlexiQoL total scores after the first 28 days of prolonged treatment interruption will be excluded. - PlexiQoL total scores from the day of study intervention recommencement will be included.
Target PN resection	Treatment policy strategy	PlexiQoL total scores after a target PN surgical resection will be included in the analysis.

ICR, independent central review; PlexiQoL, Plexiform Neurofibroma Quality of Life scale; PN, plexiform neurofibroma; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

4.2.2.2.4 Handling of Dropouts and Missing Data

Missing items within an assessment will be handled as in the previous section.

4.2.2.2.5 Primary Analysis of the 2nd Key Secondary Endpoint

The difference in the mean change from baseline in PlexiQoL total score at Cycle 12 between treatment groups (selumetinib minus placebo), standard error, 95% CI, and p-value will be assessed by LS means of an MMRM with randomised study intervention, cycle number, geographical region and average baseline chronic target PN pain intensity group as categorical fixed effects; the baseline PlexiQoL total score as a continuous covariate; and treatment-by-cycle number and baseline PlexiQoL total score-by-cycle number interactions. Parameters will be estimated with the REML approach, and the Kenward-Roger approximation is used to estimate the degrees of freedom.

The analysis will be performed on a subset of the FAS population with an evaluable baseline assessment and at least one evaluable post-baseline assessment. Evaluable assessment is defined as a PlexiQOL total score. Data collected during the randomised period (see Section 3.3.2.2) will be included, adhering to the intercurrent event strategy (refer to Section 4.2.2.2.3).

The analysis will be conducted using PROC MIXED in SAS using the same model specifications as specified for the 1st key secondary endpoint (see Section 4.2.2.1.5). All main effects and the interaction terms will remain in the model, regardless of significance.

In addition to the Cycle 12 estimate, the model will present LS mean estimates for each treatment group and their differences, standard errors, 95% CIs, and p-values (where applicable and with statistical significance effects being interpreted as nominally significant) for mean changes from baseline to each cycle during the randomised period (based on the treatment-by-cycle interaction coefficient) and the average over the randomised period (based on the treatment coefficient). A plot of the LS means accompanied by the 95% CI will be produced.

4.2.2.2.6 Sensitivity Analyses of the 2nd Key Secondary Endpoint

Analyses to assess the robustness to assumptions of missing data.

The analyses use MI and apply the planned primary analysis to the imputed outcomes, combining the statistics using Rubin's rules. These two supplementary analyses only vary as to the scenarios implemented for the intercurrent event of treatment discontinuation, as follows:

- Multiply impute reversion to baseline after discontinuation of treatment (Ratitch et al 2013) up to Cycle 12: this analysis provides an estimate for the estimand that targets the hypothetical approach, using the assumption that subjects who discontinue treatment gain no improvement in change of baseline. For the template code to implement the reversion to baseline, see the supplementary appendix to Ratitch et al 2013, under the code heading "Pattern imputation based on reasons for discontinuation".
- MI using attributable and non-attributable dropout reasons: hypothetical strategy with MNAR with reversion to baseline for scores after study intervention discontinuation due to attributable dropout reasons and MAR after study intervention discontinuation due to non-attributable dropout reasons. The reason for study intervention discontinuation will be categorised as attributable or non-attributable as outlined in Section 3.3.9.

Refer to Section 4.2.2.1.6.1 for details on identifying cycles eligible for imputation using reversion-to-baseline and for a summary of the approach.

4.2.2.2.7 Supplementary Analyses of the 2nd Key Secondary Endpoint

PlexiQoL total score, as well as the change from baseline at each cycle will be summarised descriptively by study intervention arm. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib. Summary statistics will include n, mean, standard deviation, min, Q1, median, Q3 and max.

Graphical plots of the mean scores, as well as the change from baseline and associated 95% CIs by cycle will be produced for the randomised period.

Summary statistics will also be presented by attributable versus non-attributable reasons (see Section 3.3.9) only for the randomised period.

The analysis will be performed on the FAS and the Pain FAS.

4.2.3 ORR (Single-Arm)

4.2.3.1 ORR (Single-Arm) Definition

ORR will be defined as the proportion of patients who have a cCR (defined as the disappearance of the target PN, confirmed by a consecutive scan within 3 to 6 months after the first response; see Section 4.2.1.2) or cPR (defined as a target PN volume decrease ≥ 20% compared to baseline, confirmed by a consecutive scan within 3 to 6 months after the first response; see Section 4.2.1.2) as determined by ICR per REiNS criteria.

The ORR will be derived using while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

Any cCR or cPR which occurred after a subsequent NF1-PN treatment (following study intervention discontinuation) will not be included in the numerator for the ORR calculation (where the selumetinib FAS will be the denominator). Patients with no post-baseline MRI assessments will be considered as non-responders (not having a cCR or cPR).

The confirmation by a consecutive scan within 3 to 6 months will be derived as CR/PR achieved on consecutive visits with no missed visits as per assessment schedule. If MRI assessments are excluded due to prolonged study intervention interruption (as described in Section 4.2.1.6), then the visit will be considered a missed visit.

If the interim analysis takes place, at DCO2 the same partial volume strategy defined in 4.2.1.1 will be used for this endpoint and the ORR will be analysed using the scaled partial

volumes after DCO1 (third set of parameters), and a supplementary analysis will be performed using the non-adjusted partial volume after DCO1 (first set of parameters).

At the final DCO, the same partial volume strategy defined in 4.2.1.1 will be used for this endpoint. At the final DCO, the ORR will be analysed using the scaled partial volumes after DCO2 (third set of parameters), and a supplementary analysis will be performed using the unadjusted partial volume after DCO2 (first set of parameters).

4.2.3.2 Best Objective Response Definition

The best objective response (BOR) will be calculated based on the overall visit responses from each on-treatment MRI assessment, as described in Section 0. It is the best response a patient has had following the start of intervention, but prior to starting any subsequent NF1-PN therapy and up to and including progression or the last evaluable MRI assessment in the absence of progression. The categorisation of BOR will be based on REiNS using the following response categories: cCR, CR, cPR, PR, SD, PD, and NE.

The different ways of achieving a BOR of cCR, CR, cPR, or PR are shown in Table 11.

Table 11: Scenarios for Best Objective Response

Post-baseline Visit N	Post-baseline Visit N+1	BOR
PR	PR/CR	Confirmed PR (cPR)
PR	SD/PD/NE	Unconfirmed PR
CR	CR	Confirmed CR (cCR)
CR	PR/SD/PD/NE	Unconfirmed CR
SD/NE	CR	Unconfirmed CR
SD/NE	PR	Unconfirmed PR
NE	NE	NE

BOR, best objective response; CR, complete response; NE, not evaluable; PD, progressive disease; PR, partial response; SD, stable disease.

BOR will be determined programmatically based on REiNS from the overall visit response using all data up until the first progression event or the last evaluable MRI assessment in the absence of progression. BOR will be derived using the on-treatment MRI volumetric assessments (see Section 3.3.2.1). The denominators will be consistent with those used in the single-arm ORR analysis.

Patients with no post-baseline MRI assessments will be considered as non-responders. For patients whose progression event is death, the BOR will be calculated based on all evaluable MRI assessments prior to death using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

At DCO2, the BOR will be analysed using the ICR data with the scaled partial volume after DCO1 (third set of parameters) as described in Section 4.2.1.1 .

At the final DCO, the BOR will be analysed using the ICR data with the scaled partial volume after DCO2 (third set of parameters) as described in Section 4.2.1.1.

4.2.3.3 Intercurrent Events

The approach for intercurrent events mostly follows the same as outlined for the primary endpoint in Section 4.2.1.7 However, the first intercurrent event for the primary endpoint, as detailed in Table 8, will change to be 'Post randomised study intervention discontinuation due to any reason' since early crossover from placebo is not relevant to the analysis of single-arm ORR, as analysis is performed on patients randomised to selumetinib (see Section 4.2.3.4). And any post randomised study intervention discontinuation should not be restricted to prior to Cycle 16, as the single-arm ORR is evaluated over the whole while on-treatment MRI volumetric assessments period (Section 3.3.2.1).

4.2.3.4 Primary Analysis of Single-Arm ORR

The primary analysis will include all patients randomised to selumetinib (selumetinib FAS), that is, a single arm assessment of ORR. Data obtained from the first selumetinib dose up until progression or the last evaluable assessment in the absence of progression, and while on-treatment MRI volumetric assessments (see Section 3.3.2.1) will be included in the assessment of single-arm ORR. Data are included as outlined in Section 4.2.3.1, adhering to the intercurrent event strategy (refer to Section 4.2.3.3).

The ORR will be presented with a corresponding 2-sided exact 95% CI based on the Clopper-Pearson method (Clopper and Pearson 1934).

Analyses will occur at DCO1, DCO2, and the final DCO. At DCO2 and the final DCO the partial volume strategy defined in Section 4.2.1.1 will be used for this endpoint. At DCO2, the ORR will be analysed using the scaled partial volumes after DCO1 (third set of parameters). At the final DCO, the ORR will be analysed using the scaled partial volumes after DCO2 (third set of parameters).

4.2.3.5 Supplementary Analyses of ORR

Supportive evidence of the primary endpoint of ORR will include the following:

- The ORR will be calculated based on all MRI volumetric assessments without excluding scans during study intervention interruption. MRI volumetric assessments data will be defined as data after the date of first dose of study intervention until the study intervention discontinuation date or DCO, whichever occurs first. The analysis will follow the same methods as for the primary analysis.

- A supplementary analysis will be performed including all randomised patients who received at least one dose of selumetinib (Extended Selumetinib FAS). In this analysis, the baseline MRI scan for patients randomised to placebo will be defined as the last MRI scan performed on placebo intervention prior to starting on selumetinib intervention. The ORR will be derived using while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- The BOR, defined as the BOR (see Section 4.2.3.2) recorded from the start of study intervention until prior to starting any subsequent NF1-PN therapy or up to and including progression or the last evaluable MRI assessment in the absence of progression, will also be summarised by category (cCR, CR, cPR, PR, SD, PD, or NE) in the selumetinib FAS and Extended Selumetinib FAS. BOR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

4.2.4 Duration of Response

4.2.4.1 Definition

For the subset of patients who have a cCR or cPR, DOR will be defined as the time from the date of the first documented response (which is subsequently confirmed) until the date of documented progression as assessed by ICR per REiNS criteria or death due to any cause. DOR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1). The end of response should coincide with the MRI assessment of progression or death from any cause used for the progression-free survival (PFS) endpoint (see Section 4.2.5). The date of the first documented response will be defined as the latest of the dates contributing towards the first response of PR or CR which is subsequently confirmed.

4.2.4.2 Derivations and Censoring Rules

If a patient does not progress following a response, then their DOR will use the PFS censoring (Section 4.2.5.2). The derivation formula for the DOR is date of PFS event or censoring – date of first response + 1. The DOR will be derived based on the actual MRI assessment dates and not visit dates.

4.2.4.3 Intercurrent Events

The approach for intercurrent events mostly follows the same as outlined for the primary endpoint in Section 4.2.1.7. However, the first intercurrent event for the primary endpoint, as detailed in Table 8, will change to be ‘Post randomised study intervention discontinuation due to any reason’ since early crossover from placebo is not relevant to the analysis of duration of response, as analysis is performed on patients randomised to selumetinib (see Section 4.2.4.4). And any post randomised study intervention discontinuation should not be restricted to prior to Cycle 16, as duration of response is evaluated over the whole while on-treatment MRI volumetric assessments period (Section 3.3.2.1).

4.2.4.4 Primary Analysis of DOR

Kaplan-Meier (KM) plots of DOR will be presented including a patients-at-risk table. Median DOR, 25th and 75th percentiles, and 95% CI will also be calculated together with the estimated percentage remaining in response and 95% CI using the KM method and presented in a summary table. The number and percentage of responding patients remaining in response at 8, 12, 16, 20, and 24 months post-response and every 6 months thereafter will also be summarised.

Only patients randomised to selumetinib (Selumetinib FAS) with a cCR or cPR prior to selumetinib discontinuation will be included in this analysis. Data are included as outlined in Sections 4.2.4.1 and 4.2.4.2, adhering to the intercurrent event strategy (refer to Section 4.2.4.3).

Swimmer plots that clearly show the profile of each patient who responds will also be produced. Additional graphical symbols to depict the start of the objective response, PD or death, and subsequent therapy will be added. “On treatment” will be designated with different colour from “post-treatment.”

4.2.4.5 Supplementary Analyses of DOR

- DOR will be calculated based on all MRI volumetric assessments without excluding scans during study intervention interruption. MRI volumetric assessments data will be defined as data after the date of first dose of study intervention until the study intervention discontinuation date or DCO, whichever occurs first. The analysis will follow the same methods as for the primary analysis of DOR.
- If there is at least one patient who progresses or dies after two or more consecutive missed MRI assessments, then the DOR analysis will be repeated without the specific PFS censoring rule (i.e., also including progression or death events after two or more consecutive missed MRI assessments). DOR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- A supplementary analysis will be performed including all randomised patients who received at least one dose of selumetinib (Extended Selumetinib FAS) with a cCR or cPR prior to selumetinib discontinuation (e.g., the date of the first documented response is the date of the first documented response to selumetinib). In this analysis, the baseline MRI scan for patients randomised to placebo will be defined as the last MRI scan performed on placebo intervention prior to starting on selumetinib intervention. The DOR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- Individual subject-level DOR (months) for each patient who responds will be summarised in a time-to-event listing for DCO2 and DCO3 for the Extended Selumetinib FAS.

4.2.5 Progression-Free Survival

4.2.5.1 Definition

PFS will be defined as the time from the date of randomisation until the date of progression by ICR per REiNS criteria or death due to any cause. PD is defined in Section 4.2.1.2. The appearance of new PN (with the exception of new discrete subcutaneous neurofibromas) or unequivocal progression of existing clinically relevant non-target PN is also considered PD as described in Section 4.2.1.3. PFS will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

4.2.5.2 Derivations and Censoring Rules

Patients who have not progressed or died at the time of analysis will be censored at their last evaluable MRI assessment. However, if the patient progresses or dies after two or more consecutive missed MRI assessments, the patient will be censored at the time of the latest evaluable MRI assessment prior to the missed visits. The details of censoring rules are listed in Table 12.

The derivation formula for PFS is date of PFS event or censoring – date of randomisation + 1.

If the patient has no evaluable MRI assessments post-baseline, they will be censored at Day 1 unless they die within two cycles from baseline. The PFS time will be derived based on actual MRI assessment dates and not visit dates.

Table 12: Summary of Censoring Guidelines for PFS

Assessment	Date of event, death or censoring	PFS outcome
Death, PD	MRI assessment of earliest sign of PD or death date if the event is death	Event
No PD or death at time of analysis or lost to follow-up	Latest evaluable MRI assessment	Censored
Death or PD after ≥ 2 missed assessments	Latest evaluable MRI assessment prior to the two missed assessments	Censored
No evaluable MRI assessment or no baseline data and no death within two cycles from baseline	Day 1	Censored

MRI, magnetic resonance imaging; PD, progressive disease; PFS, progression-free survival.

Based on the scheduled visit assessment scheme (i.e., every 4 cycles until Cycle 24 and every 6 cycles from then onwards), the definition of 2 missed visits will change over time and is calculated as the protocolled time between 2 subsequent scans + the protocol-allowed visit window for an early visit at the previous assessment + the protocol-allowed visit window for a late visit at the expected assessment:

Table 13: Two Missed Visits Window Rule

Scheduled assessments	Previous REiNS assessment	Two missed visits window
Q16w ± 1 week up to Week 80	Up to Week 15 (<Day 105)	33 weeks (231 days)
	Week 15 – Week 79 ≥105 - <553	34 weeks (238 days)
	Week 79 – Week 95 ≥ 553 - <665	42 weeks (294 days)
Q24w ± 1 week	Week 95 ≥665	50 weeks (350 days)

REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

4.2.5.3 Intercurrent Events

The approach for intercurrent events mostly follows the same as outlined for the primary endpoint in Section 4.2.1.7. However, the first intercurrent event for the primary endpoint, as detailed in Table 8, will change to be 'Post randomised study intervention discontinuation due to any reason' since early crossover from placebo is not relevant to the analysis of PFS, as analysis is performed on patients randomised to selumetinib (see Section 4.2.5.4). And any post randomised study intervention discontinuation should not be restricted to prior to Cycle 16, as PFS is evaluated over the whole while on-treatment MRI volumetric assessments period (Section 3.3.2.1).

4.2.5.4 Primary Analysis of PFS

A KM plot of PFS will be presented including a patients-at-risk table. The percentage of PFS at Months 4, 8,12, 16, 20, and 24 and every 6 months thereafter will be summarised based on the KM method. If more than 50% of patients progress in either arm, the median PFS and 95% CIs will also be calculated using the KM method and presented in the summary table.

Data are summarised and analysed using the Selumetinib FAS. Data are included as outlined in Sections 4.2.5.1 and 4.2.5.2, adhering to the intercurrent event strategy (refer to Section 4.2.5.3).

4.2.5.5 Supplementary Analyses of PFS

The following supplementary analyses will be performed:

- PFS will be calculated based on all on-treatment MRI volumetric assessments without excluding scans during study intervention interruption. The analysis will follow the same methods as the primary analysis of PFS.
- If there is at least one patient who progresses or dies after two or more consecutive missed MRI assessments, then the same PFS analysis will be done without this specific censoring rule (i.e., also including progression or death results after two or more consecutive missed MRI assessments). PFS will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- A supplementary analysis will be performed including all randomised patients who received at least one dose of selumetinib (Extended Selumetinib FAS). In this analysis, the baseline MRI scan for patients randomised to placebo will be defined as the last MRI scan performed on placebo intervention prior to starting on selumetinib intervention. PFS will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- Individual subject-level PFS (months) will be summarised in a time-to-event listing for DCO2 and DCO3 for the Extended Selumetinib FAS.

4.2.6 Time to Progression

4.2.6.1 Definition

Time to progression (TTP) will be defined as the time from the date of randomisation until the date of the first documented objective disease progression by ICR per REiNS criteria. Objective disease progression is defined as PD as per overall visit response. TTP will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

4.2.6.2 Derivations and Censoring Rules

Patients who have not progressed at the time of analysis will be censored at their last evaluable MRI assessment. However, if the patient progresses after two or more missed assessments, the patient will be censored at the latest evaluable MRI assessment prior to the two missed assessments.

The derivation formula for TTP is date of progression or censoring – date of randomisation + 1.

If the patient has no evaluable MRI assessments post-baseline, they will be censored at Day 1. The TTP will always be derived based on actual MRI assessment dates and not visit dates.

Refer to Section 4.2.5.2 for further details on the derivation of two or more missed cycles.

4.2.6.3 Intercurrent Events

The approach for intercurrent events mostly follows the same as outlined for the primary endpoint in Section 4.2.1.7. However, the first intercurrent event for the primary endpoint, as detailed in Table 8, will change to be 'Post randomised study intervention discontinuation due to any reason' since early crossover from placebo is not relevant to the analysis of TTP, as analysis is performed on patients randomised to selumetinib (see Section 4.2.6.4). And any post randomised study intervention discontinuation should not be restricted to prior to Cycle 16, as TTP is evaluated over the whole while on-treatment MRI volumetric assessments period (Section 3.3.2.1).

4.2.6.4 Primary Analysis of TTP

A KM plot of TTP will be presented including a patients-at-risk table. The percentage TTP at Months 4, 8, 12, 16, 20, and 24 and every 6 months thereafter will be summarised based on the KM method. If more than 50% of patients progress in either arm, the median TTP and 95% CIs will also be calculated using the KM method and presented in the summary table.

Data are summarised and analysed using the Selumetinib FAS. Data are included as outlined in Sections 4.2.6.1 and 4.2.6.2, adhering to the intercurrent event strategy (refer to Section 4.2.6.3).

4.2.6.5 Supplementary Analyses of TTP

- TTP will be calculated based on all on-treatment MRI volumetric assessments without excluding scans during study intervention interruption. The analysis will follow the same methods as the primary analysis of TTP.
- If there is at least one patient who progresses or dies after two or more consecutive missed MRI assessments, then the same TTP analysis will be done without this specific censoring rule (i.e., also including progression results after two or more consecutive missed MRI assessments). TTP will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- A supplementary analysis will be performed including all randomised patients who received at least one dose of selumetinib (Extended Selumetinib FAS). In this analysis, the baseline MRI scan for patients randomised to placebo will be defined as the last MRI scan performed on placebo intervention prior to starting on selumetinib intervention. TTP will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- Individual subject-level TTP (months) will be summarised in a time-to-event listing for DCO2 and DCO3 for the Extended Selumetinib FAS.

4.2.7 Time to Response

4.2.7.1 Definition

Time to response (TTR) will be defined as the time from the date of randomisation until the date of the first documented objective response (which is subsequently confirmed; cCR or cPR) as determined by ICR per REiNS criteria (i.e., date of first response – date of randomisation+ 1). TTR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1). The date of first documented response should coincide with that used for the DOR endpoint (see Section 4.2.4).

4.2.7.2 Derivations and Censoring Rules

The TTR will always be derived based on the actual MRI assessment dates and not visit dates.

Only patients in the Selumetinib FAS who have achieved a cCR or a cPR prior to selumetinib discontinuation will be evaluated for TTR.

The derivation formula for TTR is date of first documentation of cCR or cPR – date of randomisation + 1.

4.2.7.3 Intercurrent events

The approach for intercurrent events mostly follows the same as outlined for the primary endpoint in Section 4.2.1.7. However, the first intercurrent event for the primary endpoint, as detailed in Table 8, will change to be 'Post randomised study intervention discontinuation due to any reason' since early crossover from placebo is not relevant to the analysis of TTR, as analysis is performed on patients randomised to selumetinib (see Section 4.2.7.4). And any post randomised study intervention discontinuation should not be restricted to prior to Cycle 16, as TTR is evaluated over the whole while on-treatment MRI volumetric assessments period (Section 3.3.2.1).

4.2.7.4 Primary Analysis of Time to Response

A KM plot of TTR will be presented including a Patients-at-Risk table. The percentage TTR at months 4, 8, 12, 16, 20 and 24, and every 6 months thereafter will be summarised based on the KM method. Median TTR and 95% CIs will also be calculated using the KM method and presented in the summary table.

Data is summarised and analysed using the Selumetinib FAS. Data is included as outlined in Section 4.2.7.1 adhering to the intercurrent event strategy (refer to Section 4.2.7.3).

4.2.7.5 Supplementary Analyses of TTR

- TTR will be calculated based on all on-treatment MRI volumetric assessments without excluding scans during study intervention interruption. The analysis will follow the same methods as for the primary analysis of TTR.

- A supplementary analysis will be performed including all patients in the Extended Selumetinib FAS with a cCR or cPR prior to selumetinib discontinuation. In this analysis, the baseline MRI scan for patients randomised to placebo will be defined as the last MRI scan performed on placebo intervention prior to starting on selumetinib intervention. TTR will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).
- Individual subject-level TTR (months) will be summarised in a time-to-event listing for DCO2 and DCO3 for the Extended Selumetinib FAS.

4.2.8 Best Percentage Change From Baseline in Target PN Volume During the Randomised Period

4.2.8.1 Definition

The randomised period is defined in Section 3.3.2.2 and target PN in Section 4.2.1.2.

4.2.8.2 Derivations

The absolute change and percentage change from baseline in target PN volume at each assessment will be calculated. For each assessment, two target PN volumes will be available from two independent radiologists. The average of the two target baseline PN volumes and the average of each post-baseline target PN volume will be used to calculate the absolute change and percentage change from baseline. The best change in volume is the largest decrease from baseline or the smallest increase from baseline in the absence of a reduction and will include all on-treatment MRI volumetric assessments (see Section 3.3.2) during the randomised period prior to the earliest of death in the absence of progression, any evidence of progression, or the last evaluable REiNS assessment if the patient has not died or progressed.

The best percentage change from baseline in target PN volume will be derived using the best evaluable while on-treatment MRI volumetric assessment (see Section 3.3.2.1) during the randomised period.

4.2.8.3 Intercurrent Events

Follow the same approach as outlined for the primary endpoint in Section 4.2.1.7

4.2.8.4 Primary Analysis of Best Percentage Change From Baseline in Target PN Volume During the Randomised Period

The effect of selumetinib on best percentage change from baseline in target PN volume during the randomised period will be estimated from an analysis of covariance (ANCOVA) model including covariates for baseline target PN volume, baseline PAINS-pNF chronic target PN pain intensity group (<3, ≥ 3), and geographical region. The number of patients,

unadjusted mean, and LS means for each treatment group should be presented together with the difference in LS means, 95% CI, and corresponding p-value.

Data will be summarised and analysed using FAS. Data collected during the randomised period (see Section 3.3.2.1) adhering to the intercurrent event strategy (refer to Section 4.2.8.3) are included.

4.2.8.5 Supplementary Analyses of Best Percentage Change From Baseline in Target PN Volume During the Randomised Period

Percentage change from baseline will be calculated based on all on-treatment MRI volumetric assessments without excluding scans during study intervention interruption. The analysis will follow the same methods as the primary analysis of percentage change from baseline during the randomised period. Data will be summarised and analysed using a subset of the FAS including patients with measurable target PN at baseline per ICR who received at least one dose of study intervention and had a least one post-baseline target PN volume during the randomised period.

The absolute values, change in target PN volume from baseline, and percentage change in target PN volume from baseline will be summarised using descriptive statistics and presented at each timepoint during the randomised period by treatment group.

Waterfall plots and histograms (Mercier et al 2019) will summarise the best percentage change from baseline in target PN volume.

For the waterfall plot, each patient's best percentage change in PN volume will be represented as a separate bar, with the bars ordered from the largest increase to the largest decrease.

For the histograms, each histogram bar will represent the number of patients within specific ranges of best percentage change in target PN size from baseline.

The waterfall plot and histograms will be summarised and analysed using a subset of the FAS including patients with measurable target PN at baseline per ICR who received at least one dose of study intervention and had at least one post-baseline target PN volume during the randomised period.

The analysis will be repeated using the same method but will not be restricted to the randomised period using the Selumetinib FAS.

4.2.9 Chronic Target PN Pain Palliation During the Randomised Period

4.2.9.1 Definition

Chronic target PN pain palliation at a cycle is defined as the occurrence of both of the following events:

1) Improvement of ≥ 2 in average PAINS-pNF chronic target PN pain intensity score

2) No increase in chronic PN pain medication compared to baseline for that cycle

   i. Assuming that the patient had non-missing chronic PN pain medication data (see Section 4.2.9.4 for definition of "missing") and did not meet the criteria of eDiary chronic PN pain medication score increase of ≥1 compared to baseline (primary definition; see Section 4.2.9.2).
   ii. Assuming that the patient had non-missing chronic PN pain medication data (see Section 4.2.9.4 for definition of "missing") and did not meet the criteria of eDiary chronic PN pain medication score increase of ≥0.5 compared to baseline (first sensitivity definition; see Section 4.2.9.2).
   iii. Assuming that the patient had a non-missing investigator electronic case report form (eCRF) assessment (see Section 4.2.9.4 for definition of "missing") and did not meet the criteria of "increased" due to "PN pain" compared to baseline (second sensitivity definition; see Section 4.2.9.2).

Table 14 summarises the chronic target PN pain definitions that will be examined.

Table 14: Definitions of Chronic Target PN Pain Palliation

Definition	Chronic target PN pain intensity threshold for palliation	Non-missing pain medication data and pain medication increase based on
Primary	2	eDiary chronic PN pain medication score increase of ≥ 1
Sensitivity 1	2	eDiary chronic PN pain medication score increase of ≥ 0.5
Sensitivity 2	2	Investigator's assessment of "increased" "due to increased PN pain"
Sensitivity 3	From psychometric work*, if different than 2	eDiary chronic PN pain medication increase of ≥ 1
Sensitivity 4	From psychometric work*, if different than 2	eDiary on chronic PN pain medication score of ≥ 0.5
Sensitivity 5	From psychometric work*, if different than 2	Investigator's assessment of "increased" "due to increased PN pain"

PN, plexiform neurofibroma.
*Psychometric validation of the PAINS-pNF instrument in this study will be performed based on blinded data before DCO1, and a clinically meaningful within-patient change will

be derived based on interim data. In case this value is different than 2, an additional definition of chronic target PN pain palliation will be an improvement of greater or equal than the clinically meaningful within-patient change derived from the psychometric analysis.

Pain palliation will be assessed in all post-baseline cycles during the randomised period in the Pain FAS.

The randomised period is defined in Section 3.3.2.2.

4.2.9.2 Derivations

Detailed derivations for PAINS-pNF chronic target PN pain intensity average cycle score is defined in Section 4.2.2.1.2.

An increase in eDiary-reported chronic PN pain medication per cycle compared to baseline is defined as one of the following:

Primary definition: Increase from baseline ≥ 1 point (increase to a stronger class of analgesic medication on the WHO modified analgesic ladder [see Section 3.3.8, Table 2]) with the same drug of the strongest score taken for at least 3 days in a 28-day cycle (refer to Section 3.3.8,

- Figure 3).
- Sensitivity definition 1: Increase from baseline ≥ 0.5 point (increase to a stronger class of medication on the WHO modified analgesic ladder [see Section 3.3.8 Table 2] or an increase in the daily dose or number of pain medications within the same analgesic class) with the same drug taken for at least 3 days in a 28-day cycle (refer to Section 3.3.8,

- Figure 3)
- Sensitivity definition 2: "No increase in PN pain medication at a specific cycle compared to baseline" will be identified as those patients who had available eCRF data on investigator assessment of analgesic requirement but did not have an investigator's assessment of "increased" analgesic requirement "due to PN pain".

Following the definitions in Section 4.2.9.1, chronic target PN pain palliation values will be 1 for patients experiencing pain palliation and 0 for patients that do not.

4.2.9.3 Intercurrent Events

The intercurrent event strategy for the chronic target PN pain palliation is described in Table 15.

Table 15: Intercurrent Event Strategy for Chronic Target PN Pain Palliation

Intercurrent event	Strategy	Details
Changes to patients' chronic PN pain medication	N/A	Included in endpoint definition.
Randomised study intervention discontinuation	Hypothetical strategy	PAINS-pNF chronic target PN pain intensity scores following randomised study intervention discontinuation will not be collected and will be modelled through direct likelihood techniques.
Early crossover from placebo to selumetinib in patients with documented progression on imaging (as determined by ICR per REiNS criteria)	While on-treatment strategy	PAINS-pNF chronic target PN pain intensity scores and pain medication data after early crossover from placebo to selumetinib will be set to missing.
Prolonged study intervention interruption (defined as a study intervention interruption greater than or equal to 28 days)	While on-treatment strategy	The following rules will be followed to align with the while on-treatment strategy: - PAINS-pNF chronic target PN pain intensity scores and pain medication data during the first 28 days of prolonged treatment interruption will be included. - PAINS-pNF chronic target PN pain intensity scores and pain medication

		after the first 28 days of prolonged treatment interruption will be excluded. - PAINS-pNF chronic target PN pain intensity scores and pain medication data from the day of study intervention recommencement will be included.
Target PN resection	Treatment policy strategy	PAINS-pNF chronic target PN pain intensity scores and pain medication data after a target PN surgical resection will be included in the analysis.

ICR, independent central review; N/A, not applicable; PAINS-pNF, PAin INtensity Scale for Plexiform Neurofibroma; PN, Plexiform Neurofibroma; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

4.2.9.4 Handling of Dropouts and Missing Data

Missing components required for the determination of chronic target PN pain palliation may result if there is:

- Missing average cycle chronic PN pain intensity score for a specific cycle, that is, patients does not have at least 4 out of 7 PAINS-pNF chronic target PN pain intensity scores in at least 3 non-overlapping 7-day periods (same definition as for 1st key secondary endpoint; see Section 4.2.2.1.2);

and/or

- Missing pain medication:
  1) For eDiary, missing chronic PN pain medication for all days* (i.e., 28-day cycle) in that specific cycle.
  2) For investigator assessment, missing eCRF investigator assessment of analgesic use at that specific cycle.

*For the purposes of this analysis, pain medication data are considered missing *for a day* if the gatekeeper item "Did you take any medication for your usual (chronic) tumour pain from the time you went to bed last night until now (including overnight)?" is "yes" but no pain medications have been entered or the gatekeeper item is missing.

4.2.9.5 Primary Analysis of Chronic Target PN Pain Palliation

The pain palliation response during the randomised period is analysed using a random-effect logistic regression model including treatment, geographical region, and cycle number as fixed effects; baseline chronic target PN pain intensity score and baseline

chronic PN pain medication modified WHO analgesic ladder score as covariates; and the following interaction terms: treatment-by-cycle number, baseline chronic target PN pain intensity score-by-cycle number, and baseline chronic PN pain medication score-by-cycle number.

The analysis will be conducted using PROC GLIMMIX in SAS with a random intercept also included to account for variability between patients. Random (subject-specific) effects will be not explicitly modelled, and an unstructured covariance matrix will be used to model the within-patient correlations between the repeated measurements and to allow for unequal treatment variance. An unstructured modelling of within-patient correlations removes one layer of assumptions on the random effects and often provides the best fit to the data.

If the fit of the unstructured covariance structure fails to converge, every attempt should be made to ensure that convergence is obtained from the unstructured correlation structure. Parameters will be tried to be estimated by means of the Fisher's scoring algorithm rather than the default Newton-Raphson algorithm. If convergence is still not reached after the change in the interaction algorithm, a more parsimonious model will be implemented using one of the following covariance structures in order until convergence is reached: toeplitz with heterogeneity, autoregressive with heterogeneity, toeplitz, autoregressive. and compound symmetry. All main effects and the interaction terms will remain in the model, regardless of significance.

The results of the analysis are presented in terms of an odds ratio (selumetinib versus placebo) together with its 95% CI and p-value at each cycle of the randomised period (estimated using the treatment-by-cycle interaction coefficient) and overall, over the randomised period (based on the treatment coefficient). For each treatment arm, the odds will be presented along with the number of non-missing observations used in the estimation (n), the observed number and percentage of responder, and the corresponding 95% CI.

The analysis will be performed on a subset of the Pain FAS population with evaluable baseline assessments and at least one evaluable post-baseline assessment. An evaluable baseline assessment is defined as having both baseline PAINS-pNF chronic target PN pain intensity score (same definition as the 1st key secondary endpoint; refer to Section 4.2.2.1.2) and non-missing pain medication (refer to Section 4.2.9.4 for missing pain medication derivation). An evaluable assessment at baseline or a specific cycle is defined as having both average chronic target PN pain intensity score (same definition as the 1st key secondary endpoint; refer to Section 4.2.2.1.2) and non-missing pain medication (refer to Section 4.2.9.4 for missing pain medication derivation).

This analysis will be repeated using both sensitivity definitions outlined in Section 4.2.9.2. For the sensitivity using investigator assessment of analgesic use, the random-effect logistic

regression model will include treatment, geographical region, and cycle number as fixed effects; baseline chronic target PN pain intensity score and baseline chronic PN pain medication score as covariates; and the following interaction terms: treatment-by-cycle number, baseline chronic target PN pain intensity score-by-cycle number, and baseline chronic target PN pain medication score-by-cycle number.

4.2.10 Time to Chronic Target PN Pain Palliation During the Randomised Period

4.2.10.1 Definition

Time to chronic target PN pain palliation (TTPP) is defined as time to first chronic target PN pain palliation.

The date of the last day of the cycle will be used for calculating the time to event; therefore, time to event will be measured in months. TTPP will be defined as the duration of time in months from randomisation to the date of the last day of the first cycle where chronic target PN pain palliation is achieved.

Chronic target PN pain palliation is defined in Section 4.2.9.1.

The main definitions for chronic target PN pain palliation are proposed to be the primary, sensitivity 1, and sensitivity 2 definitions, resulting in 3 time-to-event endpoints. However, other definitions in Table 14 might be considered, especially should the psychometric analysis demonstrate a clinically meaningful within-patient change different than the pre-specified value of 2.

The randomised period is defined in Section 3.3.2.2.

4.2.10.2 Derivations

Patients who did not experience chronic target PN pain palliation will be censored at the end date of the last non-missing cycle as defined in Section 4.2.9.1, that is, at the last cycle with non-missing average chronic target PN pain intensity score and “eligible” pain medication data. Patients with no post-baseline average cycle chronic target PN pain palliation scores will be excluded from the analysis.

Only PAINS-pNF chronic target PN pain scores and pain medication data during the randomised period (i.e., up to and including Cycle 12) will be considered.

4.2.10.3 Intercurrent Events

Follow the same approach as outlined for chronic target PN pain palliation in Section 4.2.9.3.

4.2.10.4 Handling of Dropouts and Missing Data

Not applicable.

4.2.10.5 Primary Analysis of TTPP

KM curves will be used to estimate the distribution of TTPP. The 25th, 50th, and 75th percentiles of KM estimates will be used to estimate TTPP. A two-sided 95% CI will be provided for these estimates. TTPP distributions between the two study intervention arms will be compared using a log-rank test stratified by geographical region.

A KM plot will be produced.

If there is an incident of interruption of randomised intervention that results in cycle outcomes omitted from the analysis for any patient, the KM analysis will be performed with interval censoring using PROC ICLIFETEST (SAS procedure) and a Cox proportional hazard regression model will be performed using PROC ICPHREG (SAS procedure). If there is no such incident, the SAS procedures PROC LIFETEST and PROC PHREG will be used for the analyses. The Cox proportional hazards model will include treatment and geographical region as fixed effects and baseline chronic target PN pain intensity score and baseline chronic PN pain medication modified WHO analgesic ladder score as covariates.

Separate analyses will be performed for each of the time-to-event definition (defined in Section 4.2.10.1).

The analysis will be performed on a subset of the Pain FAS population with an evaluable baseline assessment and at least one evaluable post-baseline cycle, that is, both average cycle chronic target PN pain intensity score (same definition as 1st key secondary endpoint; refer to Section 4.2.2.1.2) and chronic PN pain medication score (see Section 4.2.9.4 for missing pain medication derivation).

4.2.11 PN Pain Medication, Change in PN Pain Medication and PN Pain Medication Decrease During the Randomised Period

4.2.11.1 Definition

See Section 3.3.8 for definitions.

4.2.11.2 Derivations

For both spike and chronic PN pain medication, the analgesic ladder class and the strongest analgesic ladder score for each cycle will be derived following the scoring algorithm presented in Section 3.3.8. For chronic PN pain, changes from baseline (Increased/Decreased/No Change) will also be derived as outlined in Section 3.3.8.

At each cycle, a binary variable will be derived, set to 1 if a decrease in chronic PN pain medication is observed as per definition in Section 3.3.8 and 0 otherwise. A sensitivity definition of this binary endpoint is derived using the secondary definition of decrease from baseline as assessed by the investigator when the reason for the decrease in analgesic requirement is PN pain.

4.2.11.3 Intercurrent Events

The intercurrent event strategy for the chronic target PN pain palliation is described in Table 16

Table 16: Intercurrent Event Strategy for Decrease in Pain Medication

Intercurrent event	Strategy	Details
Changes to patients' chronic PN pain medication	N/A	Included in endpoint definition.
Randomised study intervention discontinuation and early crossover from placebo to selumetinib in patients with documented progression on imaging (as determined by ICR per REiNS criteria)	While on-treatment strategy	The strongest analgesic ladder scores following randomised study intervention discontinuation will not be included.
Target PN resection	Treatment policy strategy	The strongest analgesic ladder scores after a target PN surgical resection will be included in the analysis.

ICR, independent central review; N/A, not applicable; PN, Plexiform Neurofibroma; REiNS, Response Evaluation in Neurofibromatosis and Schwannomatosis.

4.2.11.4 Handling of Dropouts and Missing Data

For the chronic PN pain medication based on the eDiary, the strongest class of pain medications reported over the 28 days of each cycle will be based on observed scores, that is, days with missing information will be ignored in the derivation.

For purposes of the analysis of decrease in pain medication, 1) if the pain medication data provide sufficient evidence to rule out a decrease for a cycle (see Section 4.2.9.2), that cycle is considered not to have missing pain medication data; 2) if, due to missing *daily*

pain medication data as defined in Section 4.2.9.4, the cycle does not provide sufficient evidence to rule out either an increase or a decrease in pain medication for a cycle, that cycle will be considered to have missing pain medication data.

For the investigator's assessment of change in pain medication, analyses will be based on observed data.

4.2.11.5 Primary Analysis of Pain Medication

The pain medication decrease during the randomised period will be analysed using a random-effect logistic regression model including treatment, geographical region, and cycle number as fixed effects; baseline chronic PN pain medication score as covariates; and the following interaction terms: treatment-by-cycle number and baseline chronic PN pain medication score-by-cycle number.

A random intercept will also be included to account for variability between patients. An unstructured covariance matrix will be used to allow for unequal treatment variance and to model the correlation between different treatment measurements within the same subject. If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be used in order until convergence is reached: toeplitz with heterogeneity, autoregressive with heterogeneity, toeplitz, autoregressive, and compound symmetry. All main effects and the interaction terms will remain in the model regardless of significance.

The results of the analysis are presented in terms of an odds ratio (selumetinib versus placebo) together with its 95% CI and p-value at each cycle (estimated using the treatment-by-cycle interaction coefficient) and overall, over the randomised period (based on the treatment coefficient). For each treatment arm, the odds will be presented along with the number of non-missing observations used in the estimation (n), the observed number and percentage of responders, and the corresponding 95% CIs. The analysis will be performed on a subset of the FAS population with an evaluable baseline and at least one post-baseline chronic PN pain medication score available. Data collected during the randomised period (see Section 3.3.2.2) adhering to the intercurrent event strategy (refer to Section 4.2.11.3) are included.

This analysis will be repeated using the sensitivity definition outlined in Section 4.2.11.2. For the sensitivity using investigator assessment of analgesic use, the random-effect logistic regression model will include treatment, geographical region, and cycle number as fixed effects; baseline chronic PN pain medication score as covariate; and the following interaction terms: treatment-by-cycle number and baseline chronic target PN pain medication score-by-cycle number.

4.2.11.6 Additional Analyses of Pain Medication

Descriptive analyses of chronic PN pain medication will be produced for the FAS as follows:

- Strongest analgesic ladder category and score at baseline and at each post-baseline cycle,
- A shift table with medication that "increased," "decreased," or "No Change" as compared to baseline at each cycle,

Descriptive analyses of spike PN pain medication will be produced for the FAS as follows:

- Strongest analgesic ladder category at baseline and each post-baseline cycle.

Chronic and spike PN pain medications will be listed.

For the investigator assessment of pain medication during the randomised period, descriptive analyses of the number of patients whose medication "increased," "decreased," or "No Change" as compared to baseline at each cycle will be shown. The number of patients whose pain medication increased or decreased will also be shown separately by reason.

The analyses on PN pain medication, change in PN pain medication and PN pain medication decrease during the randomised period (except for spike medication) will be repeated for patients in Pain FAS.

4.2.11.7 Subgroup Analyses

The above analyses on the FAS and Pain FAS will be repeated per geographical region (Europe, China, Japan, Rest of World). For spike medication, the FAS analyses are repeated per geographical region and per baseline PAINS-pNF chronic target PN score categories.

4.2.12 PII-pNF Pain Interference Total Score

4.2.12.1 Definition

The PII-pNF assesses the extent to which PN pain interferes with daily functioning. All items ask the patient to consider the pain related to their PN in the past 7 days. Patients are asked how much their PN pain has made it hard for them to engage in physical activities (e.g., challenging physical activities and self-care), affected their physiological processes (e.g., sleep and energy) or impacted their social-emotional functioning (e.g., enjoyment of activities and mood). Items are rated on a 7-point Likert scale (from 0 = not at all to 6 = completely), and the total score is the mean of the completed items.

The PII-pNF is planned to be collected at the end of Cycles 1, 2, 4, 6, 8, 10, and 12 during the randomised period and at Cycles 16, 20, 24, 30, and every 6 cycles onwards.

The randomised period is defined in Section 3.3.2.2.

4.2.12.2 Derivations

The mean change from baseline in PII-pNF total score during the randomised period will be calculated using the patient's cycle PII-pNF score minus the baseline score.

4.2.12.3 Intercurrent Events

Follow the same approach as outlined for the 2nd key secondary endpoint in Section 4.2.2.2.3.

4.2.12.4 Handling of Dropouts and Missing Data

Similar to the 2nd key secondary endpoint, missing PII-pNF total scores after treatment discontinuation are handled by the "while on-treatment" approach.

4.2.12.5 Primary Analysis of PII-pNF Pain Interference Total Score

The primary analysis for PII-pNF pain interference total score will be similar to the primary analysis of the 2nd key secondary endpoint as described in Section 4.2.2.2.5, that is, the patient PII-pNF pain interference total score change from baseline at each cycle during the randomised period will be compared between study intervention arms using an MMRM with treatment, geographical region, average baseline chronic target PN pain intensity group and cycle number as factors as well as baseline PII-pNF pain interference total score as a covariate.

Adjusted mean change from baseline estimates per treatment group over the randomised period of study treatment and at each cycle up to Cycle 12 are presented along with 95% CIs. The corresponding treatment differences, 95% CIs, and p-values are also presented.

The analysis will be performed on a subset of the FAS population with an evaluable baseline assessment and at least one evaluable post-baseline assessment. An evaluable assessment is defined as a PII-pNF pain interference total score. Data collected during the randomised period (see Section 3.3.2.2) adhering to the intercurrent event strategy (refer to Section 4.2.12.3) are included.

4.2.12.6 Additional Analyses of PII-pNF Pain Interference Total Score

PII-pNF pain interference total actual scores as well as the changes from baseline at each cycle will be summarised descriptively by study intervention arm. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib. Summary statistics will include n, mean,

standard deviation, min, Q1, median, Q3, and max. Graphical plots of the mean scores as well as change from baseline and associated 95% CIs by cycle will be produced for the randomised period.

Analysis will be performed on the FAS and the Pain FAS.

4.2.13 Patient-Reported Outcomes Measurement Information System Physical Function Items

4.2.13.1 Definition

The Adult Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function items measure self-reported capability rather than the actual performance of physical activities. Items assess upper extremity function (dexterity), lower extremity function (walking or mobility), and central regions (neck and back) as well as activities of daily living such as running errands. Patients will be assessed using the following items from the PROMIS item bank that map to physical function concepts relevant to patients with symptomatic PN:

- Does your health now limit you in bending, kneeling, or stooping?
- How much difficulty do you have doing your daily physical activities, because of your health?
- To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?

Items are rated on a 5-point Likert rating scale. The 7-day recall Patient-Reported Outcomes Measurement Information System Physical Function Short Form v2.0 – Physical Function 8c (PROMIS PF8c) instrument is used.

The adult PROMIS Physical Function items are planned to be collected at Baseline and at Cycles 2, 4, 8, and 12 during the randomised period as well as at Cycles 16, 20, 24, 30, and every 6 cycles onwards as per the Schedule of Activities (see Section 1.3 of the CSP).

The randomised period is defined in Section 3.3.2.2.

4.2.13.2 Derivations

Three items selected from the PROMIS physical function item bank will be used in this study. The items capture physical function impacts separate from those related to pain interference.

A total score will be derived by summing up the responses of the three items. Change from baseline to the end of each cycle in PROMIS item and total scores will be derived.

4.2.13.3 Intercurrent Events

Follow the same approach as outlined for the 2nd key secondary endpoint in Section 4.2.2.2.3.

4.2.13.4 Handling of Dropouts and Missing Data

Missing items within an assessment will be handled as described in Section 4.2.13.2.

4.2.13.5 Primary Analysis of PROMIS Physical Function Items

The primary analysis for PROMIS Physical Function items (including total score) will be similar to the primary analysis of the 2nd key secondary endpoint as described in Section 4.2.2.2.5, that is, the PROMIS item/total score change from baseline at each cycle during the randomised period will be compared between study intervention arms using an MMRM with treatment, geographical region, average baseline chronic target PN pain intensity group and cycle number as factors as well as baseline PROMIS item/total score as a covariate.

Adjusted mean changes from baseline estimates per treatment group over the randomised period of study treatment and at each cycle up to Cycle 12 are presented along with 95% CIs. The corresponding treatment differences, 95% CIs, and p-values are also presented.

The analysis will be performed on a subset of the FAS population with an evaluable baseline assessment and at least one evaluable post-baseline assessment. An evaluable assessment is defined as a PROMIS Physical Function item/total score. Data collected during the randomised period (see Section 3.3.2.2) adhering to the intercurrent event strategy (refer to Section 4.2.13.3) are included.

4.2.13.6 Additional Analyses of PROMIS Physical Function Items

PROMIS Physical Function actual item/total scores as well as the change from baseline at each cycle will be summarised descriptively by study intervention arm. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib. Plots of the mean scores, as well as change from baseline, and associated 95% CI by cycle will be produced for the randomised period.

The analysis will be performed on the FAS and the Pain FAS.

4.2.14 Paediatric Quality of Life Inventory NF1 Module – Adult Form

4.2.14.1 Definition

The adult version of the Paediatric Quality of Life Inventory (PedsQL) NF1 Module can be used to understand the multidimensional nature of NF1 on the HRQoL patients with this disorder and may assist in medical decision making. The instrument demonstrates initial feasibility, reliability, and discriminant validity. The PedsQL NF1 Module 3.0: Adult self-report instrument comprises 18 domains/subscales and patients will be assessed using the Skin Sensations domain (3 items).

The PedsQL questionnaire is planned to be collected at Baseline, Cycles 2, 4, 8, and 12 during the randomised period as well as at Cycles 16, 20, 24, 30, and every 6 cycles onwards as per the Schedule of Activities (see Section 1.3 of the CSP).

The randomised period is defined in Section 3.3.2.2.

4.2.14.2 Derivations

Items are rated on a 5-point Likert scale from 0 (“Never”) to 4 (“Almost always”). Items are reverse-scored and linearly transformed to a 0 to 100 scale as follows: 0 = 100, 1 = 75, 2 = 50, 3 = 25, 4 = 0. Domain scores are calculated by the sum of all the items over the number of items answered. If more than 50% of the items in the scale are missing, the scale score should not be computed. If 50% or more items are completed, missing scores will then be imputed with the mean of the completed items in a scale.

The change from baseline to the end of each cycle in the PedsQL Skin Sensation domain score will be derived.

4.2.14.3 Intercurrent Events

Follow the same approach as outlined for the 2nd key secondary endpoint in Section 4.2.2.2.3.

4.2.14.4 Handling of Dropouts and Missing Data

Missing items within an assessment will be handled as described in Section 4.2.14.2.

4.2.14.5 Primary Analysis of PedsQL (NF1 Module – Adult Form)

The primary analysis for the PedsQL Skin Sensations domain score will be similar to the primary analysis of the 2nd key secondary endpoint as described in Section 4.2.2.2.5, that is, the PedsQL Skin Sensations domain score change from baseline at each cycle during the randomised period will be compared between study intervention arms using an MMRM with treatment, geographical region, average baseline chronic target PN pain intensity group and cycle number as factors as well as baseline PedsQL Skin Sensations domain score as a covariate.

Adjusted mean change from baseline estimates per treatment group over the randomised period of study treatment and at each cycle up to Cycle 12 are presented, along with 95% CIs. The corresponding treatment differences, 95% CIs, and p-values are also presented.

The analysis will be performed on a subset of the FAS population with an evaluable baseline assessment and at least one evaluable post-baseline assessment. Evaluable assessment is defined as a PedsQL Skin Sensations domain score. Data collected during the randomised period (see Section 3.3.2.2) adhering to the intercurrent event strategy (refer to Section 4.2.15.3) are included.

4.2.14.6 Additional Analyses of PedsQL (NF1 Module – Adult Form)

The PedsQL Skin Sensations domain actual scores as well as the change from baseline at each cycle will be summarised descriptively by study intervention arm. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib. Summary statistics will include n, mean, standard deviation, min, Q1, median, Q3, and max. Graphical plots of the mean scores as well as change from baseline and associated 95% CIs by cycle will be produced for the randomised period.

The analysis will be performed on the FAS and the Pain FAS

4.2.15 EuroQoL Five Dimensions, Five Level Health State Utility Index

4.2.15.1 Definition

The EuroQoL five dimensions, five level health state utility index (EQ-5D-5L) will be used to explore the impact of treatment and disease state on health state utility.

The EQ-5D-5L, developed by the EuroQoL Group, is a generic questionnaire that provides a simple descriptive profile of health and a single index value for health status for economic appraisal (EuroQoL 2019). The EQ-5D-5L questionnaire comprises 6 questions that cover 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) over 5 levels of increasing severity (“no problems,” “slight problems,” “moderate problems,” “severe problems,” and “extreme problems”). Patients indicate their current health state by selecting the most appropriate level in each of the 5 dimensions.

Respondents also assess their health today using the EQ-VAS, which ranges from 0 (worst imaginable health) to 100 (best imaginable health).

The EQ-5D-5L questionnaire is planned to be collected at baseline and at Cycles 2, 4, 8, and 12 during the randomised period as well as at Cycles 16, 20, 24, 30, and every 6 cycles onwards, as per the Schedule of Activities (see Section 1.3 of the CSP).

The randomised period is defined in Section 3.3.2.2.

4.2.15.2 Derivations

A unique EQ-5D-5L health state, termed the EQ-5D-5L profile, is reported as a five-digit code with a possible 3,125 health states. For example, state 11111 indicates no problems on any of the five dimensions. Respondents also assess their health today using the EQ-VAS, which ranges from 0 (worst imaginable health) to 100 (best imaginable health).

The EQ-5D profile will be converted into a weighted health state utility value, termed the EQ-5D index, by applying a country-specific equation to the EQ-5D-5L profile that represents the comparative value of health states. This equation is based on national valuation sets elicited from the general population, and the base case will be the United Kingdom perspective. Where a valuation set has not been published, the EQ-5D-5L profile will be converted to the EQ-5D index using a crosswalk algorithm (

Van Hout et al).

The EQ-VAS is reported separately. No derivation is needed for the EQ-VAS.

4.2.15.3 Intercurrent Events

Follow the same approach as outlined for the 2nd key secondary endpoint in Section 4.2.2.2.3.

4.2.15.4 Handling of Dropouts and Missing Data

Missing cycle assessments will not be imputed, that is, only observed data will be used for this analysis.

4.2.15.5 Primary Analysis of EQ-5D-5L

The primary analysis for EQ-5D-5L total score will be similar to the primary analysis of the 2nd key secondary endpoint as described in Section 4.2.2.2.5, that is, the EQ-5D-5L total score change from baseline at each cycle during the randomised period will be compared between study intervention arms using an MMRM with treatment, geographical region, average baseline chronic target PN pain intensity group and cycle number as factors as well as baseline EQ-5D-5L total score as a covariate and baseline EQ-5D-5L total score-by-cycle and treatment-by-cycle as interactions.

The adjusted mean changes from baseline estimates per treatment group over the randomised period of study treatment and at each cycle up to Cycle 12 are presented, along with 95% CIs. The corresponding treatment differences, 95% CIs, and p-values are also presented. A plot of the LS means accompanied by the 95% CI will be produced.

The analysis will be performed on a subset of the FAS population with an evaluable baseline assessment and at least one evaluable post-baseline assessment. An evaluable assessment is defined as an EQ-5D-5L total score. Data collected during the randomised period (see Section 3.3.2.2) adhering to the intercurrent event strategy (refer to Section 4.2.15.3) are included.

The above analysis will be repeated for EQ-VAS.

4.2.15.6 Additional Analyses of EQ-5D-5L

Summary statistics (eg, n, mean, standard deviation, min, Q1, median, Q3, and max) will be reported for the EQ-5D index score and the EQ-VAS score, and the change from baseline for the EQ-5D index score and the EQ-VAS score. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib.

The analysis will be performed on the FAS.

To support submissions to payers, additional analyses may be undertaken, and these will be outlined in a separate Payer Analysis Plan.

4.2.16 PFS During the Randomised Period

4.2.16.1 Definition

If the data show evidence of early progression, then the PFS during the Randomised Period will be analysed. Evidence of early progression will be defined as at least 5 events of progression or death due to any cause during the Randomised Period over both treatment groups.

The randomised period is defined in Section 3.3.2.2, and PFS is defined in Section 4.2.5.

4.2.16.2 Derivations

PFS will be derived using the while on-treatment MRI volumetric assessments (see Section 3.3.2.1).

The 2-missed-visit rule as described in the PFS section is not applicable for this analysis.

4.2.16.3 Primary Analysis of PFS During the Randomised Period

A KM plot of PFS will be presented. The percentage of PFS at Months 4, 6, 8, 10, and 11 will be summarised based on the KM method. If more than 50% of patients progress in either arm, the median PFS and 95% CIs will also be calculated using the KM method and presented in the summary table.

Data will be summarised and analysed using the FAS.

4.2.16.4 Additional Analyses of PFS During the Randomised Period

PFS will be calculated based on all on-treatment MRI volumetric assessments without excluding scans during study intervention interruption. The analysis will follow the same methods as the primary analysis of PFS during the randomised period.

4.2.17 PAINS-pNF Spike Target PN Pain Intensity

4.2.17.1 Definition

The primary definition will be the patient's mean maximum cycle score in PAINS-pNF spike target PN pain intensity at each cycle and will be assessed amongst all randomised patients. This definition will exclude cycles for which the patient's maximum score was 0 or missing and will not interpolate for these.

4.2.17.2 Derivations

The baseline PAINS-pNF spike target PN pain score is defined as the maximum PAINS-pNF spike target PN pain score over the screening period, regardless of the number of missing scores in the screening period.

The maximum cycle PAINS-pNF spike target PN pain score is defined as the maximum PAINS-pNF spike target PN pain score over the 28 days up to and including the last day of the cycle, regardless of the number of missing scores in the 28 days.

For the spike pain definition, cycles where the patients' maximum score is 0 for spike pain will be excluded.

At each cycle, the PAINS-pNF spike target PN pain score change from baseline is defined as the cycle PAINS-pNF spike target PN pain score at the specific cycle minus the baseline PAINS-pNF spike target PN pain score. The randomised period is defined in Section 3.3.2.2.

4.2.17.3 Handling of Dropouts and Missing Data

Missing PAINS-pNF data after treatment discontinuation is handled by the "while on-treatment" approach, using the PAINS-pNF of cycles where a patient was on treatment.

4.2.17.4 Primary Analysis of PAINS-pNF Spike Target PN Pain Intensity

The primary analysis for PAINS-pNF spike target PN pain intensity score will mirror the primary analysis of the 2nd key secondary endpoint, as described in Section 4.2.2.2.5, that is, the maximum score during the randomised period will be compared between arms using an MMRM model with treatment and geographical region as factors as well as baseline spike pain intensity score as a covariate and baseline spike pain intensity score-by-cycle and treatment-by-cycle as interactions. The LS mean (95% CI) in each study intervention arm as well as the treatment difference and p-value will be presented at each cycle and

overall, in the randomised period. A plot of the LS means accompanied by the 95% CI will be produced.

The analysis will be performed on the FAS population with a baseline score and at least one post-baseline assessment for spike target PN pain intensity.

4.2.17.5 Additional Analyses of PAINS-pNF Spike Target PN Pain Intensity

PAINS-pNF spike target PN pain intensity actual scores and change from baseline will be summarised descriptively by study intervention arm. If there are instances of early crossover, then some cycles prior to Cycle 12 will summarise data of placebo subjects who have already crossed over to selumetinib. Summary statistics will include n, mean, standard deviation, min, Q1, median, Q3, and max. Plots of the mean scores, as well as the changes from baseline and the associated 95% CIs by cycle will be produced for the randomised period.

The analysis will be performed on the FAS.

4.2.18 Percentage Change From Baseline in Target PN Volume in the Fed State

4.2.18.1 Definition

See Section 4.2.1.2.

4.2.18.2 Derivations

Since the date of switching to the fed state from Cycle 25 Day 1 is not at a fixed timepoint, the time from the fed state will be derived cycles which will be re-numbered in relation to the switch point. Any cycle after the switch will be re-labelled as “Fed Cycle 1” (i.e, the first cycle after the switch), “Fed Cycle 2” (i.e., the second cycle after the switch), etc. Any cycle prior to the switch will be re-labelled as “ Fed Cycle -1” (i.e., the last cycle before the switch), “Fed Cycle 2” (i.e., the second-to-last cycle before switch), etc.

The date of switching to the fed state (see Section 3.2.10 for the definition of the fed state) will be the maximum date between Cycle 25 Day 1 (based on the drug accountability form) and the day after the ICF signature for the CSP v4.0.

In an ideal situation when the patient switch date to the fed state corresponds to Cycle 25 Day1, the following map yields:

	Time from randomisation (timepoint)	Time from the fed state (timepoint)
Fasting state	Baseline	Baseline
	Cycle 1 Day 28	Fed Cycle -24
	Cycle 8 Day 28	Fed Cycle -16

	Cycle 12 Day 28	Fed Cycle -12
	Cycle 16 Day 28	Fed Cycle -8
	Cycle 20 Day 28	Fed Cycle -4
	Cycle 24 Day 28	Fed Cycle -1
Fed state	Cycle 30 Day 28	Fed Cycle 6
	Cycle 36 Day 28	Fed Cycle 12

4.2.18.3 Handling of Dropouts and Missing Data

Analyses will be based on available data.

4.2.18.4 Primary Analysis

The percentage change from baseline in target PN volume over time from the fed state switch will be presented using box plots and summary statistics in the Fed FAS. The individual percentage change from baseline over time from the fed state switch will also be shown in a line plot. For patients randomised to placebo, these will include data prior and post selumetinib intervention initiation.

The time from the fed state will be listed together with the time from randomisation in the listing of target PN visit response.

4.2.19 PGIS and PGIC

4.2.19.1 Definition

The PGIS is a self-reported one-item instrument that evaluates target PN symptom severity (with separate single items for chronic pain and spike pain) over the past 7 days on a 4-point scale (1 = No pain, 2= Mild, 3 = Moderate, and 4 = Severe).

The PGIC is a self-reported two-item instrument that evaluates the patient's perspective of changes in their PN-related pain (with separate two items for chronic pain and spike pain) on a 7-point scale (from 1 = Much better to 7 = Much worse) since their last visit to the study site and since starting the study medication.

The PGIS and the PGIC are planned to be collected at the end of Cycles 1, 2, 4, 6, 8, 10, and 12. The PGIS is also collected at baseline.

4.2.19.2 Analysis of PGIS and PGIC

Descriptive statistics (counts and percentages) of the item score for each question (chronic pain and spike pain) will be produced by treatment and visit for PGIS and PGIC (one table for each PRO).

A shift table will also be provided for the PGIS, showing the distribution of the responses by treatment at the end of Cycle 12 by baseline PGIS response.

The analyses will be performed on the FAS population.

4.3 Pharmacodynamic Endpoint(s)

See Section 4.4.

4.4 Pharmacokinetics

This section covers the non-compartmental PK endpoints and analyses.

Details of population PK exposure response/safety analyses will be described in the modelling analysis plan that will be finalised before database lock. The population PK exposure response/safety analyses will be presented separately from the main CSR.

4.4.1 Calculation or Derivation of PK Parameters

The PK parameters for selumetinib and N-desmethyl selumetinib will be derived using non-compartmental methods in Phoenix® WinNonlin® Version 8.1 or higher (Certara).

The PK parameters will be derived according to AstraZeneca PK standards.

PK analysis will, where data allow, be carried out using actual elapsed times determined from the PK sampling and dosing times recorded in the database. If actual elapsed times are missing, nominal times may be used.

Where data allow, the following PK parameters for selumetinib and N-desmethyl selumetinib will be derived from Cycle 1 Day 8 plasma concentrations.

PK Parameters

Cmax	Maximum observed concentration
tmax	Time to reach maximum observed concentration following drug administration
AUC(0-6)	Partial area under the concentration-time curve from time 0 to 6 hours
AUC(0-8)	Partial area under the concentration-time curve from time 0 to 8 hours
AUClast	Area under the concentration-curve from time 0 to the last quantifiable concentration
CL/F	Apparent total body clearance of drug from plasma after extravascular administration (selumetinib only)
Vss/F	Volume of distribution (apparent) at steady state following extravascular administration (selumetinib only)
MPAUC	Metabolite: parent ratio based on AUC[a]
MPCmax	Metabolite: parent ratio based on Cmax
tlast	Time of last observed concentration

[a] Will be based on either AUC(0-6) or AUC(0-8) depending on which has the most reportable values.

Additional PK parameters may be determined where appropriate.

If an entire concentration-time profile is not quantifiable, the profile will be excluded from the PK analysis.

The minimum requirement for the calculation of area under the concentration-time curve values will be the inclusion of at least 3 consecutive quantifiable concentrations. Where there are only 3 quantifiable concentrations at least one of these should follow the peak concentration.

4.4.2 Presentation of PK Data

The plasma selumetinib and N-desmethyl selumetinib concentrations and the PK parameters will be listed and presented in tabular and graphical form as appropriate according to the version of the AstraZeneca Corporate CSRHLD TFL templates and reporting standards as documented in the TFL shells, that includes applicable descriptive statistics, defined handling of individual concentrations below the Lower Limit of Quantification (LLOQ), and precision/rounding rules for concentrations and PK parameter data.

Exclusion of concentration and/or parameter data from PK summaries may apply, this will be flagged in the listings with the reason(s) for exclusion.

PK sampling is targeted for Day 8, however for operational reasons the PK sampling can take place on any day between Day 4 and Day 14 as long as the patient has received 3 consecutive days of dosing immediately prior to the PK Day. Every effort should be made to conduct PK sampling within the Day 4 - Day 14 window. However, if for unforeseen reasons this is not possible, PK sampling can be performed at any time within the first cycle provided the patient has received 3 consecutive days of dosing immediately prior to the PK day and all data will be reported and summarised as multiple dose PK data.

4.4.2.1 Plasma Concentrations

Selumetinib and N-desmethyl selumetinib plasma concentrations for each scheduled timepoint will be listed and summarised by analyte based on the PK Analysis Set. A listing of all concentration-time data, i.e., PK scheduled times, actual sample collection times, sample actual relative times, as well as derived sampling time deviations will be presented based on the Selumetinib FAS.

4.4.2.2 PK Parameters

Individual PK parameters for selumetinib and N-desmethyl selumetinib PK parameters will be listed and summarised by analyte, based on the PK Analysis Set.

4.4.2.3 Graphical Presentation

Individual concentration-time data will be graphically presented on linear and semi-logarithmic scales, for the PK Analysis Set. Each individual’s plot will include the

concentration versus actual time profiles for both analytes overlaid on the same plot. Combined individual plasma concentration versus actual times grouped by treatment will be plotted separately by analyte, on both the linear and semi-logarithmic scale in the PK Analysis Set.

Figures of geometric mean concentrations versus schedule time will be presented separately by analyte, on both a linear scale with geometric Standard Deviation (gSD) error bars and semi-logarithmic scale (no error bars), based on the PK Analysis Set.

4.5 Immunogenicity

Not applicable.

4.6 Safety Analyses

The domain safety covers exposure, AEs, clinical laboratory, vital signs, electrocardiogram (ECG), ophthalmological assessments, physical examination, and Eastern Cooperative Oncology Group (ECOG) performance score.

Tables are provided for both the randomised period safety analysis set and the on-selumetinib safety analysis set. Listings are provided for the safety analysis set only.

4.6.1 Exposure

4.6.1.1 Definitions and Derivations

Exposure will be defined as follows:

- Total (or intended) exposure = min (last dose date of selumetinib/placebo where dose > 0 mg, date of death, DCO Date) – first dose date of selumetinib/placebo + 1

- Actual exposure = Total (or intended) exposure – total duration of dose interruptions, where total (or intended) exposure will be calculated as above. A dose interruption is defined as either any day with dose = 0 mg, or any half day where planned frequency is bid but actual frequency is daily, during the treatment period.

This calculation will use the individual start and stop dates on the EX form of the CRF, which will account for any dose interruptions. The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

Exposure will also be measured by the number of cycles received. A cycle corresponds to a period of 28 days. A cycle will be counted if study intervention is started even if the full dose is not delivered.

Exposure will be derived for selumetinib and placebo. Patients who only received selumetinib will only have exposure records for selumetinib, while patients who received placebo and selumetinib will have exposure records for selumetinib and placebo.

Patients Who Permanently Discontinue During a Dose Interruption

If a patient permanently discontinues study intervention during a dose interruption, then the date of last administration of study medication will be used in the calculation of exposure. In this case, the interruption would not be included in the summary tables, but it will be included in the listings.

Relative Dose Intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to study intervention discontinuation. RDI will be defined as follows:

RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the day of treatment discontinuation or the actual last day of dosing. D is the total dose that would be delivered if the patient received the planned dose as per the CRF. Intended cumulative dose will be calculated by summing the individual doses that should have been received up to and including the last day of treatment according to the protocol planned dose and schedule. This is given by:

Integer part (rounded up) of ((min(date of last dose date where dose > 0, date of death, date of DCO) – first dose date +1) × intended dose according to the protocol)/0.5.

The intended dose for Selumetinib is 25 mg/m^2 bid.

The actual cumulative dose will be calculated by summing ((End date of study drug administration - Start date of study drug administration + 1) × Dose)/0.5 for each period of study drug administration recorded on the study drug exposure form up to min(date of last dose where dose > 0, date of death, date of DCO).

Compliance

Percentage compliance is the percentage of the actual cumulative dose delivered relative to the cumulative dose the patient should have received according to the protocol including protocol allowable dose reductions/dose interruptions up to and including the last day of day of treatment.

4.6.1.2 Presentation

Total (or intended) exposure, actual exposure, number of cycles received, RDI, and number of dose interruptions, dose reductions, and dose modifications (either a dose interruption or

a dose reduction) will be summarised for each treatment group and treatment. Exposure will be summarised for the randomised period and on-selumetinib safety analysis set. For the randomised period safety analysis set, the exposure will be separated by placebo and selumetinib. For the on-selumetinib safety analysis set, the exposure will be separated by treatment arm and will include an overall summary.

Exposure data will be listed. Date of first exposure to treatment, total exposure (days), actual exposure (days), number of cycles received, RDI and compliance to IP will be summarised at period-level for the randomised period and on-selumetinib period. Additionally, details on all dose interruptions and reductions will be summarised including, reason for action taken, start and end date of interruption/reduction and dose per administration.

Percentage treatment compliance summarised as a continuous variable and counts using categories: eg, 100%, ≥95% - <100%, ≥90% - <95%, and ≥80% - <90%, etc.

4.6.2 Adverse Events

4.6.2.1 Definitions and Derivations

The MedDRA (using the latest MedDRA version) is used to code AEs. AEs are graded according to the CTCAE version 5.0.

Treatment-emergent AEs will be assigned to treatment-emergent periods, as described in Section 3.3.2.4. The same treatment-emergent AE can be assigned to both treatment periods (e.g., AE starts during the randomised period but worsens during the on-selumetinib period).

AEs of Special Interest

Some clinical concepts (including some selected individual PTs and higher-level terms) have been considered “AEs of special interest” (AESIs) to the selumetinib program.

The AESIs in CSP Table 9 have been identified as a list of categories by the patient safety team.

Other Significant Adverse Events (OAEs)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as serious AEs (SAEs) and AEs leading to discontinuation. Based on the expert’s judgement, significant AEs of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered OAEs and reported as such in the CSR.

A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs. Examples of these could be marked haematological and other laboratory

abnormalities, and certain events that lead to intervention (other than those already classified as serious), dose reduction, or significant additional treatment.

4.6.2.2 Presentation

Summary tables of the number and percentage of patients with AEs by SOC and PT will be produced for below categories. An overall summary table will include the number and percentage of patients in each category.

- All AEs
- AEs possibly related to study intervention
- AEs of CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, possibly related to study intervention
- AEs with an outcome of death
- AEs with an outcome of death possibly related to study intervention
- All SAEs
- SAEs possibly related to study intervention
- AEs leading to discontinuation of study intervention
- AEs leading to discontinuation of study intervention, possibly related to study intervention
- AEs leading to dose interruption of study intervention
- AEs leading to dose reduction of study intervention
- AESIs
- OAEs

Patients with more than one event in any category are only counted once in each category (i.e., multiple events are not accounted for).

Additionally, the most common AEs, which are those AEs that occur in at least 10% of patients overall, are summarised by PT and by decreasing frequency. This cut-off may be modified after review of the data. When applying a cut-off (eg, 10%), the raw percentage should be compared to the cut-off, and no rounding should be applied first (ie, an AE with frequency 9.9% does not appear if the cut-off is 10%).

A separate table will present non-serious AEs occurring in more than 5% of patients in any treatment group.

CTCAE grades and summaries of the number and percentage of patients will be provided by SOC, PT, and maximum reported CTCAE grade. A summary table of AESIs by maximum reported CTCAE grade will be produced.

Each summary table above will be presented for TEAEs during the randomised period and for TEAEs during the on-selumetinib period as defined in Section 3.3.2.4.

For the summaries of on-selumetinib TEAEs, in addition to number and number and percentage of patients [n(%)], exposure-adjusted incidence rates (per 100 person-years) will be reported. These are derived as:

$$\frac{Number\ of\ participants\ with\ AE}{Total\ on-selumetinib\ person-years}\ x\ 100$$

where the total on-selumetinib person-years is given by the sum of all the individual durations of exposure until experiencing the first occurrence of the AE across all patients in the on-selumetinib period. For those patients experiencing the AE, duration of exposure will be calculated as (first occurrence AE start date – first selumetinib dose date +1)/365.25. For those patients not experiencing the AE, it will be equal to (min(last selumetinib dose date +30 days, date of study discontinuation, DCO) – first selumetinib dose date +1)/365.25.

The number of patients in the placebo treatment group might differ for the randomised period and on-selumetinib period. Patients who were randomised and treated with placebo but discontinued prior to them being able to start on selumetinib will not be counted in the on-selumetinib period but will be counted in the randomised period.

All reported AEs will be listed for the safety analysis set. The key listings which will be produced are the following:

- AEs with an outcome of death – key subject information
- SAEs – key subject information
- AEs leading to discontinuation of study intervention – key subject information
- AEs leading to dose modification (dose interruption and/or dose reduction of study intervention) – key subject information
- AESIs – key subject information
- All AEs – listing of key information

Key listings will be split by treatment group using subheadings.

4.6.3 Clinical Laboratory, Blood Sample

4.6.3.1 Definitions and Derivations

Blood samples for the determination of clinical chemistry and haematology will be taken at the visits indicated in the CSP.

The rules described in Sections 3.3.1, 3.3.3, 3.3.4, and 3.3.5 of this document considering definition of baseline, visit windows, and handling multiple records and records above or below the limit of quantification are followed.

Absolute values are compared to the reference range and classified as low (below range), normal (within range or limits of range), and high (above range). Local reference ranges are used for the primary interpretation of laboratory data.

Changes from baseline in clinical chemistry and haematology will be calculated for each post-dose visit on treatment, where on treatment includes all data collected between the first dose of selumetinib until the earlier of last dose of study intervention + 30 days, or DCO.

All laboratory results will be converted to standard units, and CTCAE grades (version 5.0) will be derived at each visit according to the CTCAE grade criteria. Parameters for which CTCAE grades are defined may have both high and low range values, CTCAE grades will be calculated for each set of high and low values in those cases. Laboratory parameters that will be summarised, along with details of whether maximum and/or minimum are of interest and availability of CTCAE grading, are given in Table 17.

Table 17: Haematology and Clinical Chemistry Parameters – Minimum/Maximum of Interest

Laboratory assessment	Maximum of interest	Minimum of interest	CTCAE grading available
ALT	Yes		Yes
AST	Yes		Yes
Alkaline phosphatase	Yes		Yes
Total bilirubin	Yes		Yes
Magnesium	Yes	Yes	Yes
Sodium	Yes	Yes	Yes
Potassium	Yes	Yes	Yes
Creatinine	Yes		Yes
Total calcium	Yes	Yes	No
Blood urea nitrogen	Yes		No
Creatine kinase	Yes		Yes
Haemoglobin	Yes	Yes	Yes
Lymphocytes (absolute)	Yes	Yes	Yes

Table 17: Haematology and Clinical Chemistry Parameters – Minimum/Maximum of Interest

Laboratory assessment	Maximum of interest	Minimum of interest	CTCAE grading available
Neutrophils (absolute)		Yes	Yes
Leukocytes (absolute)		Yes	Yes
Total red blood cell count		Yes	No
Platelets		Yes	Yes
Albumin		Yes	Yes
Total Protein		Yes	No
Amylase	Yes		Yes
Gamma-glutamyl transferase	Yes		Yes
Phosphate	Yes	Yes	No

ALT, alanine aminotransferase; AST, aspartate aminotransferase; CTCAE, Common Terminology Criteria for Adverse Events.

For parameters with no CTCAE grading that are listed in Table 17, any increase/decrease and treatment-emergent laboratory change (TELC) is derived. An on-treatment increase is defined as an increase to a value above the upper local laboratory reference limit at any time on treatment for patients with a value below the upper local laboratory reference limit at baseline. An on-treatment decrease is defined as a decrease to any value below the local laboratory reference range limit at any time on treatment for patients with a value above the lower local laboratory reference limit at baseline. A TELC is defined as any on treatment increase or decrease from baseline.

4.6.3.2 Presentations

Only on treatment laboratory data as defined in Section 3.3.2.3 are included in the summary tables. The same section describes the treatment period definitions and provides further information on the presentation of the tables. Haematology and clinical chemistry data will be summarised by treatment group and scheduled visit using descriptive statistics over time in terms of absolute values and changes from baseline for each treatment period.

Shift tables showing CTCAE grades from baseline to worst CTCAE grade on treatment will be produced for each treatment period for parameters with available CTCAE grades (see Table 17). Hyper- and/or hypo-directionality will be indicated in the parameter's labels. Percentages are based on the number of patients with a baseline value and an on-treatment value.

For parameters in Table 17 with no CTCAE grading ("No" for CTCAE grading available), the number and percentage of patients with any on treatment increase from baseline, any on treatment decrease from baseline, and a TELC is summarised. Percentages are based on the

number of patients with a baseline value below/above the local laboratory upper/lower reference limit and an on-treatment value for the any increase/decrease summaries, respectively. Percentages for a TELC are based on the number of patients with a baseline value and an on-treatment value.

Haematology and clinical chemistry data in Table 17 will also be presented graphically for each treatment period using boxplots for absolute values and changes from baseline by scheduled visit.

A table will be provided for patients with potential Hy's law (alanine aminotransferase [ALT] or aspartate aminotransferase [AST] ≥3× upper limit of normal [ULN], and total bilirubin ≥2× ULN), where the onset date of ALT or AST elevation is prior to or on the date of total bilirubin elevation.

Laboratory data for the variables listed in the CSP Table 8 will be listed. Site reference ranges will also be listed. Flags (H or L) will be applied to values falling outside the reference ranges (which will be explicitly noted on these listings where applicable), and for the CTCAE grade for parameters for which CTCAE grading applies.

4.6.4 Clinical Laboratory, Urinalysis

4.6.4.1 Definitions and Derivations

Urine samples for determination of urinalysis are collected as described in the CSP.

The rules described in Sections 3.3.1, 3.3.3, 3.3.4, and 3.3.5 of this document considering definition of baseline, visit windows, and handling multiple records and records above or below the limit of quantification are followed.

4.6.4.2 Presentations

Urinalysis data will only be listed, and no summary tables will be produced. Site reference ranges will also be listed. Flags (H or L) will be applied to values falling outside the reference ranges (which will be explicitly noted on these listings where applicable), and to values for which CTCAE grading applies.

4.6.5 Other Laboratory Evaluations

4.6.5.1 Definitions and Derivations

Pregnancy tests (serum at screening, and serum or urine at other timepoints) will be performed for women of childbearing potential.

4.6.5.2 Presentations

Pregnancy test data will only be listed, and no summary tables will be produced.

4.6.6 Vital Signs

4.6.6.1 Definitions and Derivations

Vital signs will be assessed at timelines specified in the CSP.

The rules described in Sections 3.3.1, 3.3.3, and 3.3.4 of this document considering definition of baseline, visit windows, and how to handle multiple records are followed.

4.6.6.2 Presentations

Only on-treatment vital signs data as defined in Section 3.3.2.3 are included in the summary tables. The same section describes the treatment period definitions and further information on the presentation of the tables.

All vital signs (temperature, systolic and diastolic blood pressure, pulse rate, respiratory rate, and oxygen saturation) and growth parameters (weight and BSA) will be summarised using descriptive statistics over time in terms of absolute values and change from baseline by treatment group and scheduled visit for each treatment period. As per the schedule of assessments in the CSP, height will only be summarised at baseline in the table of descriptives; however, all available data will be listed. Baseline for this parameter will be calculated using screening assessment. If this value is unavailable, then the last value before first dose on Cycle 1 Day 1 will be used.

Vital signs data will also be presented graphically for each treatment period using boxplots for absolute values and changes from baseline by treatment group and scheduled visit.

All vital signs data will be listed.

4.6.7 Electrocardiogram

4.6.7.1 Definitions and Derivations

12-lead ECGs will be recorded at visits specified in the CSP.

The rules described in Sections 3.3.1, 3.3.3, and 3.3.4 of this document considering definition of baseline, visit windows, and handling multiple records are followed.

For ECGs taken on the same day in triplicate, the average of the three readings is calculated.

The overall evaluation of an ECG is either “normal” or “abnormal”, with abnormalities categorised as either “clinically significant” or “not clinically significant”. For ECGs taken on the same day in triplicate, the worse evaluation will be selected, i.e. “abnormal” versus “normal” and “clinically significant” versus “not clinically significant”.

4.6.7.2 Presentations

A summary table of absolute values will be presented for baseline.

A frequency table showing the number of patients with their interpretation of the ECG reading (normal; abnormal, clinically not significant; and abnormal, clinically significant) at baseline by treatment group will be presented.

All ECG data will be listed.

4.6.8 Echocardiogram/Cardiac MRI

4.6.8.1 Definitions and Derivations

An echocardiogram (ECHO) or cardiac MRI scan to assess left ventricular ejection fraction (LVEF), end diastolic, and end systolic left ventricular volumes will be performed at visits specified in the CSP.

The rules described in Sections 3.3.1, 3.3.3, and 3.3.4 of this document considering definition of baseline, visit windows, and handling multiple records are followed.

4.6.8.2 Presentations

Only on-treatment laboratory data as defined in Section 3.3.2.3 are included in the summary tables. The same section describes the treatment period definitions and further information on the presentation of the tables.

A summary table of absolute values and change from baseline of LVEF, end diastolic, and end systolic left ventricular volumes will be presented by treatment group and scheduled visit for each treatment period.

For ejection fraction decrease, a shift table showing CTCAE grades from baseline to maximum on-treatment will be produced for each treatment period.

All ECHO and cardiac MRI data will be listed.

4.6.9 Ophthalmologic Assessments

4.6.9.1 Definitions and Derivations

An ophthalmologic examination (best corrected visual acuity, intraocular pressure, and slit-lamp fundoscopy) will be evaluated at the times outlined in the CSP and as clinically indicated whilst the patient is on study intervention.

The rules described in Sections 3.3.1, 3.3.3, and 3.3.4 f this document considering definition of baseline, visit windows, and handling multiple records are followed.

4.6.9.2 Presentations

Only intraocular pressure will be presented graphically using box plots for absolute values and changes from baseline by treatment group and scheduled visits for each treatment period, split by right eye and left eye within the same plot.

Only on-treatment ophthalmologic examination data as defined in Section 3.3.2.3 are included in the figure. The same section describes the treatment period definitions and further information on the presentation of the tables.

Ophthalmology examination data will be listed for patients, including a description of the abnormalities and relevant baseline result.

4.6.10 Physical Examinations

4.6.10.1 Definitions and Derivations

A full physical examination will be performed at screening and baseline, including assessments of the following: general appearance, respiratory, cardiovascular, abdomen, skin, head and neck (including ears, eyes, nose, and throat), lymph nodes, thyroid, musculoskeletal (including spine and extremities), urogenital, dermatological, gastrointestinal, endocrine, haematologic/lymphatic, and neurological systems. Targeted physical examinations will be performed at subsequent visits and will include, at a minimum, assessments of the general appearance, respiratory and cardiovascular systems, skin, and abdomen (liver and spleen).

4.6.10.2 Presentations

Physical examination data will be listed, including a description of the abnormalities and relevant baseline findings.

4.6.11 ECOG Performance Status

4.6.11.1 Definitions and Derivations

An assessment of ECOG performance status score will be performed at the visits as shown in the CSP.

The ECOG performance status scores range from 0 to 5, with lower scores indicating greater patient activity:

0 Fully active; able to carry out all usual activities without restrictions.

1 Restricted in strenuous activity, but ambulatory and able to carry out light work or work of a sedentary nature (e.g., light housework or office work).

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

3 Capable of only limited self-care; confined to bed or chair more than 50% of waking hours.

4 Completely disabled; unable to carry out any self-care and totally confined to bed or chair.

5 Dead

The rules described in Sections 3.3.1, 3.3.3, and 3.3.4 of this document considering definition of baseline, visit windows, and handling multiple records are followed.

4.6.11.2 Presentations

A shift table showing each category from baseline to maximum on-treatment record will be produced for each treatment period.

Only on-treatment ECOG data as defined in Section 3.3.2.3 are included in the shift table. The same section describes the treatment period definitions and further information on the presentation of the tables.

The ECOG performance status at each visit indicated in the SoA will be listed.

5 INTERIM ANALYSIS

The interim analysis (DCO1) is expected to occur after the 100th randomised patient has had the opportunity to complete their end of Cycle 16 assessment. This corresponds to approximately 68.5% of the primary endpoint information expected at the primary analysis (DCO2). As it is expected that all 145 patients will have completed Cycle 12 by the time the100th randomised patient has completed Cycle 16, the information fractions for the key secondary endpoints are estimated to be 100% at the interim analysis. See Section 3.3.7 for details on the multiplicity testing approach.

The interim analyses will be performed by a team of unblinded statisticians and programmers independent from AstraZeneca and reviewed by an Unblinded Review Committee. If the URC recommends the study team to be unblinded, the results of these analyses may form the basis of submissions for regulatory approval. If the interim analysis DCO (DCO1) is due to take place within approximately 4 months of the primary analysis DCO (DCO2), then the interim analysis may not be performed and only the primary analysis will be performed.

For further details of the interim analysis including which outputs will be provided to the URC, please refer to the URC Charter and the Interim Analysis SAP.

6 REFERENCES

Akshintala et al 2020
Akshintala S, Baldwin A, Liewehr DJ, Goodwin A, Blakeley JO, Gross AM, et al. Longitudinal evaluation of peripheral nerve sheath tumors in neurofibromatosis type 1: Growth analysis of plexiform neurofibromas and distinct nodular lesions. Neuro Oncol. 2020;22(9):1368-78.

Brown and Prescott 2006
Brown H, Prescott R (2006). Applied mixed models in medicine. 2nd ed. <place of publication>: John Wiley and Sons Ltd; 2006.

Clopper and Pearson 1934
Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934;26(4):404-13.

Dombi et al 2013
Dombi E, Ardern-Holmes SL, Babovic-Vuksanovic D, Barker FG, Connor S, Evans DG, et al. Recommendations for imaging tumor response in neurofibromatosis clinical trials. REiNS International Collaboration Neurology. 2013 Nov;81 (21 supplement 1):33-40.

Haybittle 1971
Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol. 1971 Oct;44(526):793-7.

Mallinckrodt et al 2008
Mallinckrod CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. Drug Information Journal. 2008;42(4):303-319.

Marzo-Ortega et al 2005
Marzo-Ortega H, McGonagle D, Jarrett S, Haugeberg G, Hensor E, O'Connor P,et al. Infliximab in combination with methotrexate in active ankylosing spondylitis: a clinical and imaging study. Ann Rheum Dis. 2005;64(11):1568-1575.

McKenna et al 2006
McKenna SP, Whalley D, Prost Y, Staab D, Huels J, Paul CF, et al. Treatment of paediatric atopic dermatitis with pimecrolimus (Elidel®, SDZ ASM 981): impact on quality of life and health-related quality of life. J Eur Acad Dermatol Venereol. 2006;20:248-254.

Mercier et al 2019
Mercier F, Consalvo N, Frey N, Phipps A, Ribba B. From waterfall plots to spaghetti plots in early oncology clinical development. Pharm Stat. 2019;18(5):526-32.

Miettinen and Nurminen 1985
Miettinen O, Nurminen M. Comparative analysis of two rates. Stat Med. 1985;4(2):213-26.

Peto et al 1976
Peto R., Pike M. C., Armitage P., Breslow N. E., Cox D. R., Howard S. et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer. 1976 Dec;34(6):585–612.

Ratitch et al 2013
Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat. 2013;12(6):337-47, and Appendix: SAS® implementation for analyses discussed in the paper.

Tay et al 2011
Tay EL, Peset A, Papaphylactou M, Inuzuka R, Alonso-Gonzalez R, Giannakoulas G, et al. Replacement therapy for iron deficiency improves exercise capacity and quality of life in patients with cyanotic congenital heart disease and/or the Eisenmenger syndrome. Int J Cardiol. 2011;151:307-312.

Van Hout et al 2012
Van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health. 2012 Jul 1;15(5):708-15.

7 APPENDIX 1 – ANALYSIS OF DATA FROM JAPAN

7.1 Introduction

To support registration in Japan, in addition to the evaluation of the global study data for primary, secondary and safety objectives, evaluation of consistency in efficacy and safety in the Japanese population is required to facilitate the benefit-risk assessment for Japanese patients. The Japan subpopulation analyses will not be part of the CSR.

This appendix outlines the pre-specified analyses of study D134BC00001 to be conducted for the Japan subpopulation.

7.1.1 Definition of Japan Subpopulation

The Japan subpopulation will include all patientsenrolled at sites in Japan.

7.2 Analysis Sets

The definition of the analysis sets for the Japan subpopulation will be the same as the global study but including only patients enrolled in Japan. Definitions of analysis sets for DCO1 are provided in Section 2.2 of the interim SAP. Definitions of analysis sets for DCO2 and final DCO are provided in Section 3.2 of the main study SAP.

Analysis sets for Japan subpopulation analysis include the following:

- Japan-only enrolled subjects
- Japan-only FAS
- Japan-only pain FAS
- PK analysis set (Japanese subpopulation vs non-Japanese subpopulation)
- Japan-only randomised period SAF
- Japan-only on-selumetinib SAF

The analysis sets used for each outcome will also be the same as defined in Section 3.2.11 of the main study SAP.

7.3 Analysis Variables

Endpoints for the Japan subpopulation include the following:

- Study population:
    - Subject disposition
    - Demographic characteristics, subject characteristics, and disease characteristics at baseline

- REiNS-related efficacy:
    - Objective response rate (ORR) by end of Cycle 16, on-treatment MRI volumetric assessments period for comparison of selumetinib vs placebo

- Actual value and change from baseline in target PN volume over time

- Other efficacy:
  - Actual value and change from baseline in PAINS-pNF chronic target PN pain intensity scores over time
  - Mean change from baseline in PAINS-pNF chronic target PN pain intensity scores over randomised period
  - Mean change from baseline in PlexiQoL total score over the randomised period

- Pharmacokinetic:
  - Summary of plasma concentrations of selumetinib and N-desmethyl selumetinib for multiple dose Cycle 1, Day 8 (Japanese subpopulation vs non-Japanese subpopulation)
  - Combined individual plasma concentrations of selumetinib versus time and N-desmethyl selumetinib versus time for multiple dose Cycle 1, Day 8 (Japanese subpopulation vs non-Japanese subpopulation) (linear and semi-logarithmic scales)
  - Summary of pharmacokinetic parameters of selumetinib and N-desmethyl selumetinib for multiple dose Cycle 1, Day 8 (Japanese subpopulation vs non-Japanese subpopulation)

- Safety:
  - Duration of exposure
  - Treatment-emergent adverse events (AE) as follows: (i) in any category; (ii) by system organ class (SOC) and preferred term (PT); (iii) assessed by investigator as possibly related to study intervention by SOC and PT; (iv) CTCAE grade 3 or higher by SOC and PT; (v) with outcome of death by SOC and PT
  - Adverse events of special interest
  - Treatment-emergent serious adverse events (TEAE) as follows: (i) by SOC and PT; (ii) leading to discontinuation of study intervention by SOC and PT

7.4 Analysis Methods and Considerations

The timing of analyses for the Japan subpopulation will follow that of the global population. This includes three planned data cut-offs (DCOs). These correspond to an interim analysis (DCO1), a primary analysis (DCO2) and a final analysis (final DCO). Full details of each DCO are provided in Section 3.1 of the main study SAP and Section 2.1 of the interim SAP. The list of TFLs to be provided for the Japan subpopulation is provided in

Section 7.5. If any TFLs are omitted for a particular DCO for the global study, then the corresponding TFL for the Japan subpopulation will also be omitted if included in the subset of TFLs for Japan. Lists of TFLs specific to each DCO are presented in Sections 6.2 and 6.3 of the interim SAP.

The same analysis methods described in Section 3 of the interim SAP for DCO1 and Section 4 of the main study SAP for DCO2 and final DCO will be applied to the Japan subpopulation accordingly. However, the analyses for the efficacy endpoints for the Japan subpopulation include the following differences compared to those described in the interim and main study SAPs:

- All statistical analyses will be considered exploratory in the Japan subpopulation.
- Results of all statistical analyses will be presented using a 95% CI. P-values will not be displayed. No adjustment for multiplicity will be made and the multiple testing procedure will not be applied.
- Region or ethnicity-related stratification factors, if any, will be omitted from the model adjustments in the Japan subpopulation.
- No subgroup analyses are planned.
- No sensitivity or supplementary analyses are planned.
- Numbering of TFLs for the Japan subpopulation will be the same as the global study but adding the suffix ‘J’ to the end with ‘(Japanese)’ added to the end of the title. For example, ‘Table 14.1.1.1.5 Demographics and Baseline Characteristics’ for global study will be appear as ‘Table 14.1.1.1.5.J Demographics and Baseline Characteristics (Japanese)’ for the Japan subpopulation.

No separate listings of individual patient in the Japan subpopulation will be produced.

7.5 TFLs for Japan Subpopulation

Details of the specific analyses to be performed for the Japan subpopulation are presented in Table 18.

Table 18 TFLs for Japan subpopulation

Type	Title	Additional Notes
Table	Subject disposition (Japanese)	• Japan-only enrolled patients
Table	Demographic characteristics (Japanese)	• Japan-only FAS
Table	Subject characteristics (Japanese)[a]	• Japan-only FAS
Table	Disease characteristics at baseline (Japanese)	• Japan-only FAS
Table	Objective response rate by end of Cycle 16, primary analysis, on-treatment MRI volumetric assessments period (Japanese)	• Japan-only FAS • p-value not to be displayed
Table	Target PN volume, actual value and change from baseline (Japanese)	• Japan-only FAS
Table	PAINS-pNF chronic target PN pain intensity scores: actual value and change from Baseline (Japanese)	• Japan-only pain FAS
Table	PlexiQoL total scores over time: actual value and change from Baseline (Japanese)	• FAS
Table	Summary of plasma concentrations (ng/mL) of selumetinib - Multiple dose Cycle 1, Day 8 (Japanese)	• PK analysis set – Japan subpopulation
Table	Summary of plasma concentrations (ng/mL) of selumetinib - Multiple dose Cycle 1, Day 8 (non-Japanese)	• PK analysis set - non-Japan subpopulation
Table	Summary of plasma concentrations (ng/mL) of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (Japanese)	• PK analysis set – Japan subpopulation
Table	Summary of plasma concentrations (ng/mL) of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (non-Japanese)	• PK analysis set - non-Japan subpopulation
Figure	Combined individual plasma concentrations (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Japanese vs non-Japanese)	• PK analysis set - Japan and non-Japan subjects to be displayed on the same plot using different colours for each subpopulation
Figure	Combined individual plasma concentrations (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8	• PK analysis set - Japan and non-Japan patients to be displayed on the same

Table 18 TFLs for Japan subpopulation

	(semi-logarithmic scale) (Japanese vs non-Japanese)	plot using different colours for each subpopulation
Figure	Combined individual plasma concentrations (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Japanese vs non-Japanese)	• PK analysis set - Japan and non-Japan subjects to be displayed on the same plot using different colours for each subpopulation
Figure	Combined individual plasma concentrations (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (semi-logarithmic scale) (Japanese vs non-Japanese)	• PK analysis set - Japan and non-Japan patients to be displayed on the same plot using different colours for each subpopulation
Table	Summary of pharmacokinetic parameters of selumetinib - Multiple dose Cycle 1, Day 8 (Japanese)	• PK analysis set – Japan subpopulation
Table	Summary of pharmacokinetic parameters of selumetinib - Multiple dose Cycle 1, Day 8 (non-Japanese)	• PK analysis set - non-Japan subpopulation
Table	Summary of pharmacokinetic parameters of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (Japanese)	• PK analysis set – Japan subpopulation
Table	Summary of pharmacokinetic parameters of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (non-Japanese)	• PK analysis set - non-Japan subpopulation
Table	Duration of exposure, randomised period (Japanese)	• Japan-only randomised period SAF
Table	Duration of exposure, on-selumetinib period (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with treatment emergent adverse events in any category (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with treatment emergent adverse events in any category (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with adverse events by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF

Table 18 TFLs for Japan subpopulation

Table	Number of subjects with adverse events, assessed by investigator as possibly related to study intervention by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events, assessed by investigator as possibly related to study intervention by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with adverse events of CTCAE grade 3 or higher by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events of CTCAE grade 3 or higher by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with adverse events of special interest, by grouped term and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events of special interest, by grouped term and preferred term (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with adverse events with outcome of death, by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events with outcome of death, by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with serious adverse events, by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with serious adverse events, by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF
Table	Number of subjects with adverse events leading to discontinuation of study intervention, by system organ class and preferred term (Japanese)	• Japan-only randomised period SAF
Table	Number of subjects with adverse events leading to discontinuation of study intervention, by system organ class and preferred term (Japanese)	• Japan-only on-selumetinib SAF

CTCAE, Common Terminology Criteria for Adverse Events; FAS, full analysis set; MRI, magnetic resonance imaging; PAINS-PNF , PAin INtensity Scale for Plexiform Neurofibroma; PK ,

Table 18 TFLs for Japan subpopulation

pharmacokinetic; PN plexiform neurofibroma; PLEXIQOL Plexiform Neurofibroma Quality of Life scale; SAF, safety analysis set.

8 APPENDIX 2 – ANALYSIS OF DATA FROM CHINA

8.1 Introduction

To support registration in China, in addition to the evaluation of the global study data for primary, secondary and safety objectives, evaluation of consistency in efficacy and safety in Chinese population is required to facilitate the benefit-risk assessment for Chinese patients.

This appendix outlines the pre-specified analysis of study D134BC00001 to be conducted for the China subpopulation of study to support submission in China. Interpretation of results from analyses of the China subpopulation will be included in a supplementary CSR.

8.1.1 Definition of China Subpopulation

The China subpopulation will include all patients enrolled at the sites in China.

8.2 Analysis Sets

Unless otherwise specified, the definition of the analysis sets will be the same as the global study, as defined in the study main SAP Section 3.2 with only patients from China subpopulation.

Analysis sets for China subpopulation analysis include the following:

- China-only enrolled subjects
- China-only full analysis set (FAS
- China-only pain FAS
- China-only selumetinib FAS
- China-only extended selumetinib FAS
- China-only randomised period safety analysis set (SAF)
- China-only on-selumetinib SAF
- Pharmacokinetic (PK) analysis set (Chinese sub-population vs non-Chinese subpopulation)

The analysis sets used for each outcome will also be the same as defined in the study main SAP Section 3.2.11.

8.3 Analysis Variables

Endpoints for the China subpopulation include the following:

- Study population:
    - Subject disposition
    - Analysis sets
    - Protocol deviations

  - Demographic characteristics at baseline
  - Subject characteristics and disease characteristics at baseline
  - Medical history and previous PN-directed treatment modalities and medical treatments
  - Prior medications, allowed and disallowed concomitant medications during the randomised period, allowed and disallowed concomitant medications during the on-selumetinib period
  - Distribution of eDiary entries for the following: PAINS-pNF chronic target PN pain; PN pain medication
  - Instrument completion rate for PlexiQoL

- REiNS-related efficacy:
  - Objective response rate (ORR) by end of Cycle 16, on-treatment MRI volumetric assessments period for comparison of selumetinib vs placebo
  - Actual value, change from baseline and percentage change from baseline in target PN volume over time
  - Single-arm ORR, best objective response (BOR), duration of response (DOR), progression-free survival (PFS), time to progression (TTP) and time to response (TTR)
  - PFS during randomised period
  - Best percentage change from baseline in target PN volume during randomised period and on-treatment MRI volumetric assessments period

- Other efficacy:
  - Actual value, change from baseline and mean scores for PAINS-pNF chronic target PN pain intensity scores over time
  - Mean change from baseline in PAINS-pNF chronic target PN pain intensity scores over randomised period
  - Actual value and change from baseline in PlexiQoL total scores over time
  - Mean change from baseline for PlexiQoL total score over randomised period
  - Chronic target PN pain palliation and time to chromic target PN pain palliation during randomised period
  - Chronic PN pain medication:; strongest analgesic ladder class and score by cycle; change from baseline
  - Decrease in pain medication over the randomised period, as reported in the e-Diary and as reported by investigator
  - Distribution of Patient Global Impression of Severity (PGIS) and Patient Global Impression of Change (PGIC) over time and shift in PGIS responses at cycle 12

- Pharmacokinetic:
    - Summary of plasma concentrations of selumetinib and N-desmethyl selumetinib for multiple dose Cycle 1, Day 8 (Chinese subpopulation vs non-Chinese subpopulation)
    - Combined individual plasma concentrations of selumetinib versus time and N-desmethyl selumetinib versus time for multiple dose Cycle 1, Day 8 (linear and semi-logarithmic scales)
    - Geometric mean plasma concentration of selumetinib vs time and N-desmethyl selumetinib versus time for multiple dose Cycle 1, Day 8 (Chinese subpopulation vs non-Chinese subpopulation) (linear and semi-logarithmic scales)
    - Summary of pharmacokinetic parameters of selumetinib and N-desmethyl selumetinib for multiple dose Cycle 1, Day 8 (Chinese subpopulation vs non-Chinese subpopulation)

- Safety:
    - Duration of exposure
    - Drug interruptions and dose reductions
    - Treatment-emergent adverse events (AE) as follows: (i) in any category; (ii) by system organ class (SOC) and preferred term (PT); (iii) assessed by investigator as possibly related to study intervention by SOC and PT; (iv) CTCAE grade 3 or higher by SOC and PT; (v) by SOC, PT and maximum reported CTCAE grade; (vi) leading to study intervention dose interruption by SOC and PT; (vii) leading to study intervention dose reduction by SOC and PT; (viii) most common (frequency of >=10%) by PT; (ix) with outcome of death by SOC and PT; (x) leading to discontinuation of study intervention by SOC and PT
    - Adverse events of special interest by SOC and PT
    - Treatment-emergent serious adverse events (AE) by SOC and PT
    - Haematology and clinical chemistry over time
    - Haematology and clinical chemistry CTCAE grade change from baseline to worst CTCAE grade on-treatment
    - Treatment-emergent abnormalities by pre-defined criteria on-treatment for haemagology and clinical chemistry
    - Left ventricular ejection fraction decrease, CTCAE grade change from baeline to worse CTCAE grade on-treatment
    - Shift from baseline in ECOG performance status to maximum record on-treatment

The genomics initiative sample data are not collected for China subpopulation and thus the corresponding endpoints will not be derived or analysed.

8.4 Analysis Methods and Considerations

China subpopulation analysis will be performed at the same time as the first global study DCO as described in the study main SAP. Additional DCO for China subpopulation may be performed if deemed necessary, eg, to fulfil local health authority's request.

The same analysis methods described in Section 4 of the study main SAP will be applied to the China subpopulation accordingly. However, the analyses for the efficacy endpoints for the China subpopulation include the following differences compared to those described in the study main SAP:

- All statistical analyses will be considered non-inferential in the China subpopulation. Only descriptive statistics will be presented.
- Results of all statistical analysis will be presented using a 95% CI. The calculated p-values are considered as nominal only. No adjustment for multiplicity will be made and the multiple testing procedure will not be applied.
- Unstratified/unadjusted analysis will be used as the primary analysis method for China subpopulation due to the risk of low number of patients in a stratum. Stratified/adjusted analysis may also be performed as supporting analysis where data allow, using the same adjustment strategy described in the study main SAP. Region or ethnicity-related stratification factors, if any, will be omitted from the model adjustments in the China subpopulation.
- No subgroup analyses are planned.
- No sensitivity analyses are planned.
- Numbering of TFLs for the China subpopulation will be the same as the global study but adding the suffix 'Chn' to the end with '(Chinese)' added to the end of the title. For example, 'Table 14.1.1.1.5 Demographics and Baseline Characteristics' for global study will be appear as 'Table 14.1.1.1.5 Chn Demographics and Baseline Characteristics (Chinese)' for the Chinese subpopulation.

No separate listings of individual patients in the China subpopulation will be produced.

8.5 TFLs for China Subpopulation

Details of the specific analyses to be performed for the China subpopulation are presented in Table 19.

Table 19 TFLs for China subpopulation

Type	Title	Additional Notes
Table	Subject disposition (Chinese)	China-only enrolled patients
Table	Analysis sets (Chinese)	China-only enrolled patients
Table	Important protocol deviations (Chinese)	China-only FAS
Table	Demographic characteristics (Chinese)	China-only FAS
Table	Subject characteristics (Chinese)	China-only FAS
Table	Disease characteristics at baseline (Chinese)	China-only FAS
Table	Previous PN-directed treatment modalities (Chinese)	China-only FAS
Table	Previous PN-directed medical treatments (Chinese)	China-only FAS
Table	Medical history (past and current) (Chinese)	China-only FAS
Table	Prior medications (Chinese)	China-only FAS
Table	Allowed concomitant medications during the randomised period (Chinese)	China-only FAS
Table	Allowed concomitant medications during the on-selumetinib period (Chinese)	China-only extended selumetinib FAS
Table	Disallowed concomitant medications during the randomised period (Chinese)	China-only FAS
Table	Disallowed concomitant medications during the on-selumetinib period (Chinese)	China-only extended selumetinib FAS
Table	Distribution of eDiary entries for PAINS-pNF chronic target PN pain (Chinese)	China-only pain FAS
Table	Distribution of eDiary entries for PN pain medication (Chinese)	China-only FAS
Table	Distribution of eDiary entries for PN pain medication (Chinese)	China-only pain FAS
Table	Instrument completion rate for PlexiQoL (Chinese)	China-only FAS
Table	Objective response rate by end of Cycle 16, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only FAS
Table	Target PN volume, actual value and change from baseline (Chinese)	China-only FAS

Table 19 TFLs for China subpopulation

Figure	Target PN volume, percentage change from baseline, individual subject data over time (Chinese)	China-only FAS
Table	PAINS-pNF chronic target PN pain intensity scores: actual value and change from Baseline (Chinese)	China-only pain FAS
Figure	Mean scores for PAINS-pNF chronic target PN pain intensity scores (Chinese)	China-only pain FAS
Table	Mean change from baseline for PAINS-pNF chronic target PN pain intensity score over the randomised period: Primary analysis (Chinese)	China-only pain FAS
Figure	LS mean change from baseline in PAINS-pNF chronic target PN pain intensity score over the randomised period (Chinese)	China-only pain FAS
Table	PlexiQoL total scores over time: actual value and change from Baseline (Chinese)	China-only FAS
Figure	Mean change from baseline for PlexiQoL total scores (Chinese)	China-only FAS
Table	Mean change from baseline for PlexiQoL total score over the randomised period: Primary analysis (Chinese)	China-only pain FAS
Figure	LS mean change from baseline in PlexiQoL total scores over the randomised period (Chinese)	China-only FAS
Table	Objective response rate - single arm, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Objective response rate - single arm, primary analysis, using non-scaled partial volumes after the second data cut-off, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Best objective response - single arm, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Figure	Duration of response, Kaplan-Meier plot, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Duration of response, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS

Table 19 TFLs for China subpopulation

Figure	Duration of response, swimmer plot, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Figure	Progression free survival, Kaplan-Meier plot, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Progression free survival, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Figure	Time to progression, Kaplan-Meier plot, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Time to progression, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Time to response, descriptive statistics, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only selumetinib FAS
Table	Best percentage change from baseline in target PN volume, primary analysis, on-treatment MRI volumetric assessments period, randomised period (Chinese)	China-only FAS
Figure	Target PN volume, best percentage change during the randomised period, waterfall plot (Chinese)	China-only FAS
Figure	Target PN volume, best percentage change during the randomised period, histogram (Chinese)	China-only FAS
Figure	Target PN volume, best percentage change during the on-treatment MRI volumetric assessments period, waterfall plot (Chinese)	China-only FAS
Figure	Target PN volume, best percentage change during the on-treatment MRI volumetric assessments period, histogram (Chinese)	China-only FAS
Table	Proportion of cycles during the randomised period with chronic target pain palliation - Primary definition (Chinese)	China-only pain FAS
Table	Time to first chronic target PN pain palliation during the randomised period (Chinese)	China-only pain FAS
Figure	Time to first chronic target PN pain palliation during the randomised period, Kaplan-Meier plot (Chinese)	China-only pain FAS

Table 19 TFLs for China subpopulation

Table	Chronic PN pain medication strongest analgesic ladder class and score by cycle (Chinese)	China-only FAS
Table	Change from baseline in chronic PN pain medication (Chinese)	China-only FAS
Table	Change from baseline in chronic PN pain medication (Chinese)	China-only pain FAS
Table	Proportion of cycles during the randomised period with pain medication decrease as reported in the e-Diary (Chinese)	China-only FAS
Table	Proportion of cycles during the randomised period with pain medication decrease as reported by the investigator (Chinese)	China-only FAS
Figure	Progression free survival during the randomised period, Kaplan-Meier plot, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only FAS
Table	Progression free survival during the randomised period, primary analysis, on-treatment MRI volumetric assessments period (Chinese)	China-only FAS
Table	Distribution of Patient Global Impression of Severity (PGIS) responses over time (Chinese)	China-only FAS
Table	Distribution of Patient Global Impression of Change (PGIC) responses over time	China-only FAS
Table	Shift table of Patient Global Impression of Severity (PGIS) responses at Cycle 12 (Chinese)	China-only FAS
Table	Summary of plasma concentrations (ng/mL) of selumetinib - Multiple dose Cycle 1, Day 8 (Chinese)	PK analysis set – Chinese subpopulation
Table	Summary of plasma concentrations (ng/mL) of selumetinib - Multiple dose Cycle 1, Day 8 (non-Chinese)	PK analysis set - non-Chinese subpopulation
Table	Summary of plasma concentrations (ng/mL) of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (Chinese)	PK analysis set – Chinese subpopulation
Table	Summary of plasma concentrations (ng/mL) of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (non-Chinese)	PK analysis set - non-Chinese subpopulation

Table 19 TFLs for China subpopulation

Figure	Combined individual plasma concentrations (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Chinese)	PK analysis set (China only)
Figure	Combined individual plasma concentrations (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8 (semi-logarithmic scale) (Chinese)	PK analysis set (China only)
Figure	Combined individual plasma concentrations (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Chinese)	PK analysis set (China only)
Figure	Combined individual plasma concentrations (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (semi-logarithmic scale) (Chinese)	PK analysis set (China only)
Figure	Geometric mean (gSD) plasma concentration (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Chinese vs non-Chinese)	• PK analysis set - Chinese and non-Chinese patients to be displayed on the same plot using different colours for each subpopulation
Figure	Geometric mean plasma concentration (ng/mL) of selumetinib versus time - Multiple dose Cycle 1, Day 8 (semi-logarithmic scale) (Chinese vs non-Chinese)	• PK analysis set - Chinese and non-Chinese patients to be displayed on the same plot using different colours for each subpopulation
Figure	Geometric mean (gSD) plasma concentration (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (linear scale) (Chinese vs non-Chinese)	• PK analysis set - Chinese and non-Chinese patients to be displayed on the same plot using different colours for each subpopulation
Figure	Geometric mean plasma concentration (ng/mL) of N-desmethyl selumetinib versus time - Multiple dose Cycle 1, Day 8 (semi-logarithmic scale) (Chinese vs non-Chinese)	• PK analysis set - Chinese and non-Chinese patients to be displayed on the same plot using different colours for each subpopulation
Table	Summary of pharmacokinetic parameters of selumetinib - Multiple dose Cycle 1, Day 8 (Chinese)	PK analysis set – Chinese subpopulation
Table	Summary of pharmacokinetic parameters of selumetinib - Multiple dose Cycle 1, Day 8 (non-Chinese)	PK analysis set - non-Chinese subpopulation

Table 19 TFLs for China subpopulation

Table	Summary of pharmacokinetic parameters of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (Chinese)	PK analysis set – Chinese subpopulation
Table	Summary of pharmacokinetic parameters of N-desmethyl selumetinib - Multiple dose Cycle 1, Day 8 (Chinese)	PK analysis set - non-Chinese subpopulation
Table	Duration of exposure, randomised period (Chinese)	China-only randomised period SAF
Table	Duration of exposure, on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Drug interruptions and dose reductions for study intervention, randomised period (Chinese)	China-only randomised period SAF
Table	Drug interruptions and dose reductions for study intervention, on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with treatment emergent adverse events in any category (Chinese)	China-only randomised period SAF
Table	Number of subjects with treatment emergent adverse events in any category (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events, assessed by investigator as possibly related to study intervention by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events, assessed by investigator as possibly related to study intervention by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events of CTCAE grade 3 or higher by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events of CTCAE grade 3 or higher by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF

Table 19 TFLs for China subpopulation

Table	Number of subjects with adverse events by system organ class, preferred term and maximum reported CTCAE grade (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events leading to study intervention dose interruption, by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events leading to study intervention dose interruption, by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events leading to study intervention dose reduction, by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events leading to study intervention dose reduction, by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events of special interest, by grouped term and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events of special interest, by grouped term and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events, most common (frequency of >=10%), by preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events, most common (frequency of >=10%), by preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with adverse events with outcome of death, by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events with outcome of death, by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Number of subjects with serious adverse events, by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with serious adverse events, by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF

Table 19 TFLs for China subpopulation

Table	Number of subjects with adverse events leading to discontinuation of study intervention, by system organ class and preferred term (Chinese)	China-only randomised period SAF
Table	Number of subjects with adverse events leading to discontinuation of study intervention, by system organ class and preferred term (Chinese)	China-only on-selumetinib SAF
Table	Haematology laboratory variables over time, on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	Haematology laboratory variables over time, on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Clinical chemistry laboratory variables over time, on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	Clinical chemistry laboratory variables over time, on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Haematology CTCAE grade change from baseline to worst CTCAE grade on treatment, randomised period (Chinese)	China-only randomised period SAF
Table	Haematology CTCAE grade change from baseline to worst CTCAE grade on treatment, on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Clinical chemistry CTCAE grade change from baseline to worst CTCAE grade on treatment, randomised period (Chinese)	China-only randomised period SAF
Table	Clinical chemistry CTCAE grade change from baseline to worst CTCAE grade on treatment, on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Treatment emergent haematology abnormalities by pre-defined criteria, on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	Treatment emergent haematology abnormalities by pre-defined criteria, on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	Treatment emergent clinical chemistry abnormalities by pre-defined criteria, on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	Treatment emergent clinical chemistry abnormalities by pre-defined criteria, on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF

Table 19 TFLs for China subpopulation

Table	Left Ventricular Ejection Fraction decrease, CTCAE grade change from baseline to worst CTCAE grade on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	Left Ventricular Ejection Fraction decrease, CTCAE grade change from baseline to worst CTCAE grade on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF
Table	ECOG performance status, shift table from baseline to maximum record on-treatment randomised period (Chinese)	China-only randomised period SAF
Table	ECOG performance status, shift table from baseline to maximum record on-treatment on-selumetinib period (Chinese)	China-only on-selumetinib SAF

CTCAE, Common Terminology Criteria for Adverse Events; ECOG, Eastern Cooperative Oncology Group; FAS, full analysis set; LS, least squares; MRI, magnetic resonance imaging; PAINS-PNF , PAin INtensity Scale for Plexiform Neurofibroma; PK , pharmacokinetic; PN plexiform neurofibroma; PLEXIQOL Plexiform Neurofibroma Quality of Life scale; SAF, safety analysis set.

Signature Page for VV-RIM-05949212 System v5.0
D134BC00001 Statistical Analysis Plan Edition 4

Approve: Document Level Task Verdict: Approved	PPD oval 04-Sep-2024 11:59:04 GMT+0000

Approve: Document Level Task Verdict: Approved	PPD al 04-Sep-2024 12:56:02 GMT+0000

Signature Page for VV-RIM-05949212 System v5.0
D134BC00001 Statistical Analysis Plan Edition 4