Statistical Analysis Plan	
Study Code	D1346C00011
Edition Number	3.0
Date	8th September 2022

A Phase 1 Open Label Study to Assess the Safety, Tolerability, Pharmacokinetics and Clinical Efficacy of Selumetinib, a Selective Mitogen Activated Protein Kinase Kinase (MEK) 1 Inhibitor, in Chinese Paediatric and Adult Subjects with Neurofibromatosis Type 1 (NF1) and Inoperable Plexiform Neurofibromas (PN)

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUC	Area under the curve
BID	*Bis in die* (Twice daily)
BLQ	Below the limit of quantification
BOR	Best overall response
BP	Blood pressure
BSA	Body surface area
C_{max}	Maximal plasma concentrations
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
CV	Coefficient of variation
DBL	Database lock
DCO	Data cut-off
DoR	Duration of response
ECG	Electrocardiogram
ECHO	Echocardiogram
eCRF	Electronic case report form
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
FAS	Full Analysis Set
FLACC	Face, Legs, Activity, Cry, Consolability

Abbreviation or special term	Explanation
HRQoL	Health related quality of life
ICF	Informed consent form
ICR	Independent central review
IOP	Intraocular pressure
IP	Investigational product
IRC	Independent Review Charter
IPD	Important protocol deviation
KM	Kaplan-Meier
LLoQ	Lower limit of quantification
Ln	Natural logarithm or logarithm to the base e
LVEF	Left ventriculation ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
MEK	Mitogen Activated Protein Kinase Kinase
MMRM	Mixed model repeated measures
MRI	Magnetic resonance imaging
NA	Not applicable
NC	Not calculated
NE	Not evaluable
NF1	Neurofibromatosis Type 1
NIH	National Institute of Health
NQ	Not quantifiable
NR	Not reportable
NRS-11	Numeric Rating Scale
NS	No sample
NTL	Non-target lesion
ORR	Objective response rate
PD	Progressive disease
PedsQL	Paediatric Quality of Life Inventory
PFS	Progression free survival
PGIC	Patient's global impression of change
PGIS	Patient's global impression of severity
PII	Pain Interference Index
PK	Pharmacokinetics

Abbreviation or special term	Explanation
PlexiQoL	PN quality of life scale
PN	Plexiform Neurofibromas
PR	Partial response
PRO	Patient-reported outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
QTcF	QT interval correct for heart rate using Fridericia's formula
REiNS	Response Evaluation in Neurofibromatosis and Schwannomatosis
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SAS®	A commercially available integrated system of software products, commonly used for reporting and analysis of Clinical Studies
SD	Stable disease
SRC	Safety Review Committee
$t_{1/2}$	Half life
TCP	Temporal change parameter
t_{max}	Time to peak plasma concentration
TFL	Tables, Figures and Listings
TL	Target lesion
TTP	Time to progression
TTR	Time to response
ULN	Upper limit of normal
ULoQ	Upper limit of quantification

AMENDMENT HISTORY

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Primary or secondary endpoints	03 Mar 2022	Revise DCO timepoints. Applies to figure 1, sections 4.2, 4.2.4.2, 4.2.4.3, 4.2.4.4, 4.2.4.5, and 4.2.7.	Y (3.0)	To align with CSP
	03 Mar 2022	Adding complete list of PK parameters. Applies to table 9.	Y (3.0)	To align with CSP
	03 Mar 2022	Inclusion of interim analysis. Applies to section 5.	Y (3.0)	To align with CSP
	03 Mar 2022	Remove reference to other significant adverse events. Applies to sections 3.11 and 4.2.2.2.	Y (3.0)	To align with CSP
Derivation of primary or secondary endpoints	08 Sep 2022	Update to derivation of CTCAE grades for liver parameters to use project reference ranges. Applies to section 3.1.2.	NA	Correction to derivation
	08 Sep 2022	Inclusion of total observation duration. Applies to sections 3.1.12 and 4.2.2.10.	NA	Additional summary of interest
	03 Mar 2022	Update AESI categories. Applies to table 5.	Y (3.0)	To align with CSP
	03 Mar 2022	Modified liver summaries text to mention alkaline phosphatase. Applies to section 3.1.2.	Y (3.0)	To improve previous intent
	03 Mar 2022	Update duration of exposure derivation. Removing date of death. Adding derivation for days Applies to section 3.1.12.	NA	To improve previous intent
	03 Mar 2022	Introduce new derivation of cycles for efficacy outputs. Applies to sections 3.3.3, 3.3.4, 3.3.5, 3.3.6, 4.2.4.2, 4.2.4.3, 4.2.4.4, and 4.2.4.5.	NA	To improve previous intent
	03 Mar 2022	Update definition of two missed visits in progression free survival. Applies to section 3.3.4.	NA	To improve previous intent
	03 Mar 2022	Text updated to add financial difficulties item into EORTC QLQ-C30 summary. Applies to sections 3.4.2.2, and 4.2.6.2.2.	NA	To improve previous intent
	03 Mar 2022	Remove reference to the use of publicly available data to inform death after DCO.	Y (3.0)	To align with CSP
Data presentation	08 Sep 2022	Additional tables and figures to illustrate ECG and echocardiogram data. Applies to section 4.2.2.5.	NA	Additional summary of interest
	08 Sep 2022	Addition of tumour size reference lines to the average tumour volume plot. Applies to section 4.2.4.6.	NA	Additional summary of interest

	08 Sep 2022	Modified tumour size outputs to remove reference of type of PN being marked. Applies to section 4.2.4.6.	NA	Correction
	08 Sep 2022	Addition of outputs relating to prior medications. Applies to section 4.2.8.	NA	Additional summary of interest
	03 Mar 2022	Modified laboratory summary to mention “change from baseline” rather than change over time. Applies to section 3.1.2.	NA	To improve previous intent
	03 Mar 2022	Update text mentioning bi-directional laboratory parameters. Applies to section 3.1.2, 4.2.2.3.	NA	To improve previous intent
	03 Mar 2022	Remove reference to minimum CTCAE grade in relation to laboratory parameter summaries, remove table 17. Applies to sections 3.1.2, 4.2.2.3.	NA	To improve previous intent
	03 Mar 2022	Remove “reason for screen failure” from summary as data noted collected. Applies to section 4.2.1.	NA	To improve previous intent
	03 Mar 2022	Update text in relation to table of most common AEs, to suggest splitting by cohort. Applies to section 4.2.2.2.	Y (3.0)	To align with CSP
	03 Mar 2022	Update text of laboratory summaries indicating plots will be produced for change from baseline rather than maximum/minimum	NA	To improve previous intent
	03 Mar 2022	Updated outputs to be presented for the clinical outcome assessment. Applies to section 4.6.	NA	To improve previous intent
	03 Mar 2022	Edit text to mention summary of both self-report and parent report scores. Applies to sections 4.2.6.2.1, and 4.2.6.3.	NA	To improve clarity
	03 Mar 2022	Remove reference to important protocol deviation from definition of Pharmacokinetic analysis set. Applies to section 2.1.2.	Y (3.0)	To align with CSP
	03 Mar 2022	Remove text explicitly mentioning the inclusion of T-score outputs for physical functioning assessment. Applies to sections 3.4.3 and 4.2.6.3.	NA	To improve previous intent
Editorial	08 Sep 2022	Remove albumin from urinalysis. Applies to table 6.	NA	To align with data collected

08 Sep 2022	Additional reference to 95% CI only being included in DoR, PFS, and TTP outputs if evaluable. Applies to section 3.3.3, 3.3.4, and 3.3.5.	NA	To improve clarity
08 Sep 2022	Update to pain medication survey. Removing reference to number and dose on analgesics. Applies to section 4.2.6.1.5.	NA	To improve previous intent
08 Sep 2022	Added section 3.5.2 to describe the data and derivations involved amongst the prior and concomitant treatments. Applies to sections 3.5.2 and 4.2.8.	Y (3.0)	To improve clarity
08 Sep 2022	Revise statement relating to pain medication survey. Pain medication after start of treatment was collected as part of the eCRF. Applies to sections 3.5.2 and 4.2.8.	N	To improve clarity
03 Mar 2022	List of abbreviations: Added definition of TCP	NA	To improve clarity
03 Mar 2022	Updated version number of the CSP from 2 to 3. Applies to section 1.	Y (3.0)	To align with CSP
03 Mar 2022	Revise wordings for the study details for paediatric and adult cohorts. Applies to section 1.	Y (3.0)	To align with CSP
03 Mar 2022	Replace “causally” with “possibly” in reference to sentence: AEs possibly related to study treatment. Applies to section 4.2.2.2	NA	To improve previous intent
03 Mar 2022	Added AUC(0-6) to list of PK parameters for completeness. Applies to section 4.2.3	NA	To improve clarity
03 Mar 2022	Updated derivation of Gmean ± gSD. Applies to section 4.2.3.	NA	To improve clarity
03 Mar 2022	Edit text to specify whether rules for precision and rounding apply to listings or tables. Applies to section 4.2.3	NA	To improve clarity
03 Mar 2022	Add detail of the modelling method used in clinical outcome assessment. Applies to section 4.2.6	NA	To improve clarity
03 Mar 2022	Update text related to profile of swimmer plots to include all subjects (not just responders)	NA	To improve previous intent

1. STUDY DETAILS

This statistical analysis plan (SAP) contains a more detailed description of the analyses in the clinical study protocol (CSP). This SAP is based on version 3.0, 30th September 2021, of the CSP.

The target population is Chinese paediatric and adult subjects with Neurofibromatosis Type 1 (NF1) and inoperable Plexiform Neurofibromas (PN) that require treatment due to symptoms or with the potential to develop significant clinical complications. The evaluations of pediatric cohort and adult cohort will be conducted separately.

1.1 Study Objectives

1.1.1 Primary Objectives

The primary objectives for this study, as outlined in the CSP, and the corresponding endpoints/variables are presented in Table 1.

Table 1: Primary Objectives

Primary objectives:	**Endpoints/variables:**
To assess the safety and tolerability of selumetinib in Chinese paediatric and adult subjects with NF1 and inoperable PN	**Paediatric and adult cohorts:** Safety and tolerability will be evaluated in terms of adverse events (AEs), clinical safety laboratory assessments, physical examination, vital signs, height/weight, electrocardiogram (ECG), echocardiogram (ECHO), ophthalmologic assessment and performance status. **Paediatric only:** Safety and tolerability will also be evaluated in terms of bone growth and Tanner stages. **Assessments related to AEs will include:** • Occurrence/frequency. • Relationship to investigational product (IP) as assessed by the investigator. • Common Terminology Criteria for Adverse Events (CTCAE) grade. • Seriousness. • Death. • AEs leading to discontinuation of IP. • AEs of special interest.
To characterise the pharmacokinetics (PK) of selumetinib and its metabolite (N-desmethyl selumetinib) in Chinese paediatric and adult subjects with NF1 and inoperable PN	PK parameters for selumetinib and N-desmethyl selumetinib will be derived following single dose and multiple doses. These may include, but are not limited to: **After a single dose:** • Area under the concentration-time curve from zero to infinity (AUC).

Primary objectives:	Endpoints/variables:
	• Area under the concentration-time curve from zero to 12 hours (AUC_{0-12}). • Area under the concentration-time curve from zero to the last measurable concentration (AUC_{0-t}). • Maximum plasma concentration (C_{max}). • Time to maximum plasma concentration (t_{max}). • Terminal half-life ($t_{1/2}$). **After multiple doses:** • Area under the concentration-time curve from zero to 12 hours at steady-state ($AUC_{0-12,ss}$). • Maximum steady-state plasma concentration ($C_{max,ss}$). • Accumulation ratio (Rac).

1.1.2 Secondary Objectives

The secondary objectives for this study and the corresponding endpoints/variables are shown in Table 2.

Table 2: Secondary Objectives

Secondary objectives:	Endpoints/variables:
To evaluate the clinical efficacy of selumetinib in Chinese paediatric and adult subjects with NF1 and inoperable PN on objective response rate (ORR), duration of response (DoR), progression-free survival (PFS), time to progression (TTP), and time to response (TTR)	ORR is defined as the proportion of subjects who have a complete response or confirmed partial response (defined as a target PN volume decrease ≥20% compared to baseline, confirmed by a consecutive scan within 3 to 6 months after first response), as determined by the investigator and independent central review per Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) criteria. DoR is defined as the time from the date of first documented response (which is subsequently confirmed) until the date of documented progression or death in the absence of disease progression, as determined by the investigator and independent central review per REiNS criteria. PFS is defined as the time from the date of first dose until progression per REiNS criteria, as assessed by the investigator and independent central review, or death due to any cause. TTP is defined as the time from the date of first dose until progression per REiNS criteria, as assessed by the investigator and independent central review.

Secondary objectives:	Endpoints/variables:
	TTR is defined as the time from the date of first dose until the date of first documented response (which is subsequently confirmed), as determined by the investigator and independent central review per REiNS criteria.
To evaluate the effect of selumetinib on pain in Chinese paediatric and adult subjects with NF1 and inoperable PN	• Face, Legs, Activity, Cry, Consolability (FLACC) scale (3 years of age). • Faces pain scale - revised (4 to 17 years of age). • Numeric Rating Scale (NRS-11; adult cohort). • Pain Interference Index (PII; adult cohort; self- and parent-reported in the paediatric cohort). • Pain Medication Survey (self-reported in the adult cohort; parent-reported in the paediatric cohort).
To determine the effect of selumetinib on health-related quality of life	• Paediatric Quality of Life Inventory (PedsQL; paediatric cohort – self- and parent-reported). • European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30; adult cohort) and PN quality of life scale (PlexiQoL; adult cohort).
To determine the effect of selumetinib on physical functioning	• Patient-Reported Outcomes Measurement Information System (PROMIS; upper extremity; self- and parent-reported in the paediatric cohort). • PROMIS (mobility; self- and parent-reported in the paediatric cohort). • PROMIS Physical Function - Short Form 8c 7-day (adult cohort).

1.1.3 Exploratory Objectives

The exploratory objectives for this study and corresponding endpoints/variables are shown in Table 3.

Table 3: Exploratory Objectives

Exploratory objectives:	Endpoints/variables
To evaluate patient's global impression of severity (PGIS) of symptoms	PGIS (adult cohort; self- and parent-reported in the paediatric cohort).
To evaluate patient's global impression of change (PGIC) in symptoms	PGIC (adult cohort; self- and parent-reported in the paediatric cohort).

To determine the effect of selumetinib on disfigurement	Photographic evaluation.

1.1.4 Safety Objectives

The safety objectives are contained within the primary objectives.

1.2 Study Design

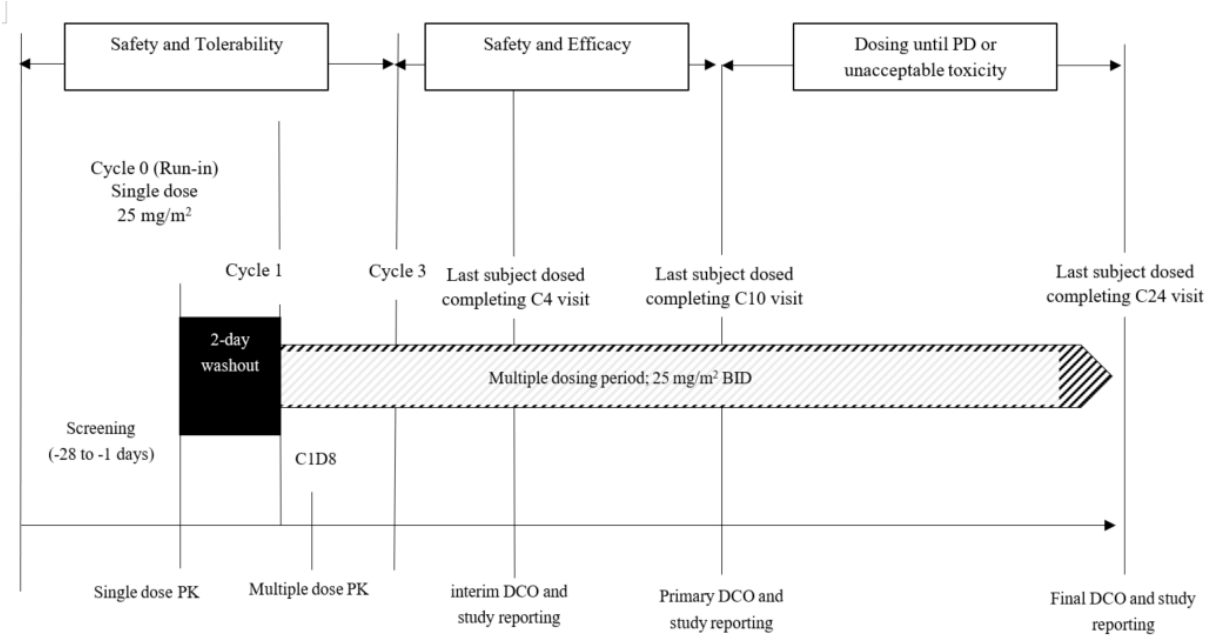
This is an open label, single arm Phase 1 study with 2 independent cohorts to assess the safety, tolerability, PK and clinical efficacy of selumetinib in Chinese paediatric and adult subjects with NF1 and inoperable PN.

The primary objectives are to assess the safety, tolerability and PK of selumetinib in Chinese subjects with NF1 and inoperable PN and the secondary objectives are to assess the efficacy of selumetinib in Chinese subjects with NF1 and inoperable PN. The study population is subjects with a clinical diagnosis of NF1 per National Institutes of Health (NIH) guidelines (NIH Consensus Development Conference Statement) and inoperable measurable PNs that require treatment due to symptoms or with the potential to develop significant clinical complications. Approximately 16 subjects will be enrolled into each of the two cohorts, paediatric and adult subjects, and the cohorts will be analysed independently.

The Safety Review Committee (SRC) will evaluate preliminary tolerability, safety and available PK data after the first 6 subjects in both cohorts have been treated for approximately 3 cycles. Additional enrolment will be initiated per SRC recommendation. Detailed information will be provided in the SRC charter.

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

Figure 1: Study design



BID=twice daily; CnDn=Cycle n Day n; DCO=data cut-off; PD=progressive disease; PK=pharmacokinetics; SRC=Safety Review Committee.
Cohort 1: Paediatric cohort = 16 subjects.
Cohort 2: Adult cohort = 16 subjects.
The SRC will evaluate preliminary tolerability and safety data, as well as PK data (if available) after the first 6 subjects in both cohorts have been treated for approximately 3 cycles. Additional enrolment will be initiated per SRC recommendation.
Long-term post-treatment safety follow-up assessments will be conducted for 1 year (at 6 and 12 months post-treatment) for paediatric subjects only.

1.2.1 Drug Administration

Subjects will initially receive a single oral dose of selumetinib 25 mg/m^2 on Cycle 0 Day 1, followed by a 2-day washout period. From Cycle 1 Day 1 subjects will receive selumetinib 25 mg/m^2 twice daily (approximately every 12 hours) on a continuous schedule for 28-day cycles, with no rest periods between cycles. Subjects should be instructed to take the dose of selumetinib on an empty stomach (no food or drink other than water for 2 hours before and 1 hour after dosing) with water only.

The dose will be based on body surface area (BSA) for both the paediatric and adult cohorts, and rounded to the nearest 5 or 10 mg using a dose nomogram. The dose of selumetinib will be capped at 50 mg when BSA is ≥ 1.9 m^2. The dose will be adjusted for changes in BSA at every cycle and at follow-up evaluations.

Subjects must keep a drug diary recording the specific time each dose was taken, and to record reasons for any missed doses. If a dose of selumetinib is missed, it should only be taken if it is more than 6 hours until the next scheduled dose.

1.2.2 Duration of Treatment

Dosing of selumetinib will continue until either disease progression or unacceptable drug-related toxicity, whichever occurs first. Subjects in both cohorts will have a safety follow-up visit 30 days after their last dose/end of treatment, and subjects in the paediatric cohort will be followed up for safety for an additional year, with assessments at 6 and 12 months post-treatment. For subjects who discontinue study treatment due to reasons other than disease progression, tumour assessments will continue for approximately 1 year, if clinically applicable, and the assessment schedule should continue per the regular tumour assessment interval.

1.3 Number of Subjects

Approximately 16 paediatric and 16 adult subjects with NF1 and inoperable PN will be enrolled in the study. Assignment to the treatment cohort and evaluation of efficacy and safety will be performed independently for each cohort. Per the Chinese Phase 1 PK guideline, at least 8 to 12 subjects per cohort are required. To allow for dropouts and nonevaluable cases, 16 subjects in each cohort will be enrolled. There must be a minimum of 6 subjects each in the 3 to 11 and 12 to 17 year age groups at the time of enrolment.

2. ANALYSIS SETS

2.1 Definition of Analysis Sets

2.1.1 Safety Analysis Set (SAF)

The Safety analysis set will consist of all subjects who have received at least one dose of selumetinib.

2.1.2 Pharmacokinetic Analysis Set

The Pharmacokinetic analysis set (PK) will include all subjects who receive at least one dose of selumetinib per protocol, for whom there are at least one evaluable post-dose PK concentration.

The analysis sets for each outcome variable are provided in Table 4.

Table 4: Summary of outcome variables and analysis sets

Outcome variable	*Analysis set*
Safety data	
Adverse events	Safety
Laboratory measurements	Safety
Vital signs	Safety

Outcome variable	*Analysis set*
Height/weight	Safety
ECGs/ECHOs	Safety
Ophthalmologic assessment	Safety
Performance status	Safety
Bone growth (paediatric cohort only)	Safety
Tanner stages (paediatric cohort only)	Safety
PK data	
PK data	PK
Efficacy data	
ORR, DoR[1], PFS, TTP, TTR	Safety
Disfigurement	Safety[2]
Clinical outcome assessment data	
Pain scales	Safety
HRQoL	Safety
Physical functioning	Safety
Study Population/Demography data	
Demography characteristics (e.g. age, sex etc.)	Safety
Baseline and disease characteristics	Safety
Important deviations	Safety
Medical/surgical history	Safety
Previous PN therapy	Safety
Concomitant medications/procedures	Safety

[1]*Subjects who are evaluable for the analysis of DoR are those who responded in the ORR analysis.*
[2]*This will be a subset of the Safety analysis set, including those subjects with disfigurement/visible PN at baseline.*

2.2 Protocol Deviations

Protocol deviations will be defined as any change, divergence or departure from the study design in the protocol. No protocol deviations to recruitment or enrolment criteria will be permitted in this study.

Protocol deviations that impact on PK may result in data being excluded from the PK Analysis Set (see Section 2.1.2) but will still be reported in the study result listings.

During the study, a list of important protocol deviations (IPDs) will be developed, which will include those deviations which are likely to have an impact on the safety and PK analyses.

These are detailed below. Note that the contents of these tables are not an exhaustive list. A complete list of anticipated protocol deviations (including important protocol deviations) will be compiled separately and finalised prior to database lock.

Inclusion criteria deviations:

- Capable of giving signed informed consent which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in the CSP.
  Paediatric cohort: parent/legal guardian consent is required, where possible involving the paediatric subject in the assent process with appropriate documentation.
- Provision of signed and dated written ICF prior to any mandatory study-specific procedures, sampling, and analyses.
  Paediatric cohort: mandatory provision of signed and dated parent/legal guardian consent for the study along with the paediatric assent form, when applicable. For subjects who reach the age of legal consent during the clinical study, notification may be required and a new consent form may need to be signed by the subject.
- Paediatric cohort: Chinese subjects ≥ 3 years and < 18 years of age with a BSA ≥ 0.55 m^2 at the time of study enrolment who are able to swallow whole capsules.
- Adult cohort: Chinese subjects ≥ 18 years of age at the time of study enrolment who are able to swallow whole capsules.
- Subjects' NF1 and inoperable PN require treatment due to actual symptoms or has the potential to develop significant clinical complications, as judged by the investigator.
- Subjects must have at least one measurable typical or nodular PN, defined as a lesion of at least 3 cm measured in one dimension, which can be seen on at least 3 imaging slices and have a reasonably well-defined contour.
- The target PN will be defined as the clinically most relevant PN, which has to be amenable to volumetric MRI analysis and classified as either typical or nodular (i.e. must not be solitary nodular).

Exclusion criteria deviations:

- Evidence of malignant peripheral nerve sheath tumour.
- Current or past history of retinal pigment epithelial detachment/central serous retinopathy or retinal vein occlusion.
- Have had prior treatment with a MEK, Ras or Raf inhibitor (including, but not limited to, vemurafenib).
- Supplementation with vitamin E greater than 100% of the daily recommended dose. Any multivitamin containing vitamin E must be stopped prior to initiation of selumetinib.
- Have had recent major surgery within a minimum of 4 weeks prior to starting study treatment, with the exception of surgical placement for vascular access. Have planned major surgery during the treatment period.

Key study assessment/procedure-related deviations:

- Baseline imaging (MRI) is missing as specified in CSP.

- Non-target lesion is also considered clinically relevant PN by the investigator; only one non-target lesion can be selected. Non-target lesions must also be measurable and either typical or nodular.
- MRI imaging is not done or of poor quality for ICR tumour assessment for cycle 4, cycle 8 and/or cycle 12 for enrolled subjects.

Received the wrong treatment or incorrect dose:

- Subjects are assigned to treatment but did not receive study treatment.
- Incorrect or incomplete dose during PK period (e.g. the study product is not administered under fasting condition during intensive PK period (Cycle 0 Day 1 and Cycle 1 Day 8, morning dose); a dose is missed, it is not taken within 3 hours after the scheduled time during PK period; the missed dose is subsequently taken > 3 hours after the scheduled time during PK period; times and dates of the IP doses on the day of blood sampling and/or the times and dates of the doses on the day before blood sampling are not recorded during PK period; subject vomits within 2 hours after dosing during intensive PK period.

This will be used as a guiding principle for the PK scientist review of all incorrect dosing prior to database lock to identify those likely to have an impact on the PK profile.

Excluded medications taken:

- Prior and concomitant medication taken which affects PK evaluation (to be reviewed by the PK scientist and study physician).

PK sample deviations:

- Incomplete PK profile collected with which accurate PK parameters cannot be evaluated. PK sample and/or data is missing during Cycle 0 and on Cycle 1 Day 8 (e.g. PK sample collection times for dates or time for each sample were not recorded; any PK sample is not collected per CSP requirement, during treatment period; any PK sample which goes missing after being obtained and before being transferred to central lab).
- Sample collection/processing errors that lead to inaccurate bioanalytical results.

This will be used as a guiding principle for the PK scientist review of all PK sample deviations prior to database lock to identify those likely to have an impact on the PK profile.

All IPDs will be summarised. A by-subject listing of all protocol deviations will be provided.

3. PRIMARY AND SECONDARY VARIABLES

The primary endpoints include the safety, tolerability and PK of selumetinib. The secondary endpoints assess the clinical efficacy. The endpoints relating to tumour response, such as duration of response and progression free survival, will be derived per the REiNS criteria

using both the Investigator's assessment and independent centrally read 3D MRI volumetric analysis.

3.1 Safety Variables

Safety and tolerability will be assessed by adverse events, clinical safety laboratory assessments, physical examination, vital signs data, height and weight, ECG, echocardiogram, ophthalmologic assessment and performance status as recorded in the CRF. For the paediatric cohort, bone growth and Tanner stages will also be used for assessment.

Safety assessments will be performed at screening and at the end of each cycle up to the end of Cycle 4, at the end of every 2 cycles up to the end of Cycle 12, at the end of every 4 cycles up to the end of Cycle 24, and every 6 cycles thereafter, unless otherwise stated, as long as the subject remains on study treatment. They will also be performed at other visits during Cycles 0 and 1. The screening visit should be performed as close as possible to Cycle 0 Day 1 and no more than 28 days prior to starting treatment. A 30-day follow-up visit will be conducted at 30 days post-last dose/end of treatment for both paediatric and adult cohorts. Long-term post-treatment safety follow-up assessments will be conducted for 1 year (at 6 and 12 months post-treatment) for paediatric subjects only to collect AE data, height and weight, and Tanner stages.

3.1.1 Adverse Events

AEs and serious adverse events (SAEs) will be collected throughout the study, from date of informed consent until 30 days after the last dose of study treatment. Events will be defined as treatment emergent if they onset or worsen (by investigator report of a change in intensity) during the treatment period as defined in the protocol, which includes during the 30-day follow-up period. For the paediatric cohort, after the 30-day follow-up visit until the end of the long-term post-treatment safety follow-up visit, all SAEs regardless of causality, and all AEs causally related to selumetinib will be collected. Any AE occurring before the start of study treatment, which does not worsen during treatment, will be summarised separately to treatment emergent AEs.

The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (using the CTCAE version referenced in the Clinical Study Protocol).

AEs of special interest

Some clinical concepts (including some selected individual preferred terms and higher-level terms) have been considered "AEs of special interest" (AESI) to the selumetinib program.

These AESIs have been identified as a list of categories provided by the subject safety team:

Table 5: Adverse events of special interest (AESIs)

AESI	MedDRA Preferred Terms (PT) Defining the AESIs
Ocular toxicity	Chorioretinopathy (central serous retinopathy [CSR]); Retinal detachment; Retinal tear; Vision blurred; Visual impairment; Vitreous floaters; Photopsia; Eye disorder; Photophobia; Retinal vein occlusion (RVO); Detachment of retinal pigment epithelium (Retinal pigment epithelial detachment [RPED]).
Hepatotoxicity	Drug-induced liver injury; ALT increased; AST increased.
Muscular toxicity	Blood creatine phosphokinase increased; Musculoskeletal pain; Muscular weakness; Myalgia; Rhabdomyolysis; Myoglobin blood increased; Myoglobin urine present; Acute kidney injury; Myopathy.
Cardiac toxicity	Ejection fraction decreased; Oedema peripheral; Peripheral swelling; Oedema; Left ventricular dysfunction; Ventricular dysfunction.
For Paediatric population only	
Physeal dysplasia	Metaphyseal dysplasia, Multiple epiphyseal dysplasia, Arthralgia, Joint stiffness, Joint hyperextension, Gait disturbance, Short stature.
Choking on the capsule	Choking, Retching.

AE=adverse event; MedDRA=Medical Dictionary for Regulatory Activities.

Other categories may be added as necessary or existing terms may be merged. An AstraZeneca medically qualified expert after consultation with the Global Safety Physician has reviewed the AEs of interest and identified which higher-level terms and which preferred terms contribute to each AESI. Further reviews may take place prior to database lock (DBL) to ensure any further terms not already included are captured within the categories. Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI grouping. Groupings will be based on preferred terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

3.1.2 Laboratory Evaluations

Blood and urine samples for determination of haematology, clinical chemistry and urinalysis will be collected throughout the study, from screening through to 30-day follow-up. Additional samples may be collected as clinically indicated. Table 6 below presents the laboratory variables to be summarised and listed. Qualitative urinalysis (Glucose, Haemoglobin, Qualitative protein) will be also listed.

Change from baseline in haematology and clinical chemistry variables will be calculated for each post-dose visit on-treatment.

All laboratory results will be converted to standard units, and CTCAE grades (version 5.0) will be defined at each visit according to the CTCAE grade criteria. Some parameters may have both high and low ranges, CTCAE grades will be calculated for each set of high and low

values in those cases. Site reference ranges will be used to derive CTCAE grades in all laboratory summaries except liver chemistries.

For summaries of liver chemistries, alanine aminotransferase (ALT), aspartate aminotransferase (AST), Alkaline phosphatase (ALP), and total bilirubin, lower limit of normal (LLN) will be defined at the project level. This was due to the site level lower limit being greater than 0. A response value less than this was not considered to be clinically meaningful. Therefore, the project level lower limit of 0 was applied.

The upper limit of normal (ULN) will be defined at the site level for each cohort, as each cohort has one site and the two cohorts will be analysed separately. For paediatric cohort, ALP has additional separate site level ULN based on gender.

Absolute values will be compared to the reference range and classified as low (below range), normal (within range or on limits of range) and high (above range). Summaries of laboratory results will include CTCAE grade changes from baseline, and listings will include reference range indicators to highlight results considered as high or low. The maximum on-treatment post-dose value will be defined for each laboratory parameter. The following parameters have CTCAE grades defined for both high and low values: Haemoglobin, lymphocytes, sodium, potassium, and magnesium. For these parameters, high and low CTCAE grades will be calculated.

The denominator used in laboratory summaries of CTCAE grades will only include evaluable subjects, in other words those who had sufficient data to have the possibility of an abnormality, as defined below.

For example:

- If a CTCAE criterion involves a change from baseline, evaluable subjects would have both a pre-dose and at least 1 post-dose value recorded.
- If a CTCAE criterion does not consider changes from baseline, to be evaluable the subject need only have 1 post-dose value recorded.

Table 6: Laboratory safety variables

Clinical chemistry* (Serum/Plasma)	Haematology (Blood)	Urinalysis
Albumin	**Complete Blood Count (CBC):**	**Quantitative:**
Alanine aminotransferase (ALT)	Total white blood cells (Leukocytes)	Protein
Amylase	Total red blood cells (Erythrocytes)	Glucose
Alkaline phosphatase	Haemoglobin	
Aspartate aminotransferase (AST)	Neutrophils	Haemoglobin/ erythrocytes/blood
Urea nitrogen [a]	Eosinophils	
Calcium (total)	Basophils	
Creatine kinase (CK) [b]	Lymphocytes	
Creatinine	Monocytes	
Gamma-glutamyl transferase		
Magnesium	**Differential:**	
Phosphate	Leukocytes (absolute)	
Potassium		
Total bilirubin	**Platelet:**	
Total protein	Platelets	
Sodium		
Uric acid		
Troponin [c]		

[a] Urea or blood urea nitrogen based on local site practice.
[b] CK to be assessed at screening and if clinically indicated (e.g. muscle symptoms)
[c] Troponin (isoform per site norm), should be assessed at screening, and performed when there is a significant drop in LVEF (of ≥10 percentage points relative to baseline and to an absolute LVEF below the institution's LLN on study treatment) or for any cardiorespiratory events with no obvious diagnosis. If troponin assessments are not available, per local practice, CK-MB isoform should be assessed. Subjects should be managed according to the CSP.
* Fasting.

3.1.3 Physical Examinations

A complete physical examination will be conducted prior to the start of selumetinib dosing (i.e. baseline) and at the planned time points outlined in Section 3.1, including an assessment 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) and neurological systems. Abnormal clinically significant findings will be recorded as medical history if found during screening, and as an adverse event if it occurs or worsens after screening.

3.1.4 Vital Signs

The following vital signs will be collected prior to blood collection: axillary temperature (ºC), pulse rate (beats/min), oxygen saturation (%) by pulse oximetry, respiratory rate

(breaths/min), and blood pressure (BP) (mmHg). Change from baseline will be calculated for each scheduled assessment on treatment.

3.1.5 Height and Weight

Height will be recorded in centimetres and weight will be recorded in kilograms. For paediatric subjects all available documented previous height and weight measurements should be collected at screening.

3.1.6 Electrocardiograms

Single 12-lead electrocardiograms (ECGs) will be obtained locally at the site. ECGs will be evaluated at the planned time points outlined in Section 3.1 using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT and corrected QT intervals. Subjects should be supine and at rest 5 minutes prior to recording the ECG. RR interval and QTcF will be calculated by AstraZeneca from the data provided.

If an abnormal ECG finding at the screening assessment is considered to be clinically significant by the investigator, it should be reported as a concurrent condition. During the study, clinically significant abnormal ECG findings not present at screening should be reported as an AE. If present, the clinical signs and symptoms associated with the abnormal finding should be reported as an AE, with the ECG abnormality given as explanatory information.

3.1.7 Echocardiograms

Echocardiograms will be performed and analysed locally at the site. Echocardiograms will be evaluated at baseline and at the planned time points outlined in Section 3.1. Left ventriculation ejection fraction (LVEF), end diastolic and end systolic left ventricular volumes should be recorded at each echocardiogram assessment. The subject should be examined using the same machine and operator throughout the study wherever possible.

3.1.8 Ophthalmologic Examination

All subjects will have a detailed ophthalmologic examination (best corrected visual acuity, intraocular pressure [IOP] and slit-lamp fundoscopy) at baseline and at the planned time points outlined in Section 3.1, and as clinically indicated whilst on treatment.

If a subject experiences symptoms of visual disturbance (including blurring of vision), a complete ophthalmological examination, including a slit-lamp examination, must be performed. If an abnormality is detected, fundus photography and an optical coherence tomography scan can also be performed where required.

If a retinal abnormality prior to or at time of selumetinib discontinuation is observed, a repeat ophthalmological examination is to be performed 30 days after discontinuation of selumetinib in order to document reversibility.

3.1.9 Performance Status

Performance status will be measured using the Karnofsky scale for subjects aged over 16 years at the time of assessment, and Lansky scoring will be used for subjects 16 years old or younger at the time of assessment. Both scales have a score range of 10 to 100, with 10 being the worst score and 100 being the best.

3.1.10 Bone Growth in Paediatric Subjects

Yearly growth assessments require integration of information including height measurements, puberty staging (Tanner staging, see Section 3.1.11), and bone age if clinically indicated. Children with impaired growth velocity should be referred to a paediatric endocrinologist as clinically indicated for evaluation, and should be followed until they reach final adult height.

An X-ray of the hand and wrist will be performed at screening and at other visits/cycles if clinically indicated to monitor growth plates and the potential risk of physeal dysplasia.

The investigator should regularly look for signs of physeal dysplasia (e.g. joint pain and fatigue after exercising, abnormal gait, and spine irregularities) in the paediatric population. Children who show signs of physeal dysplasia at the time of selumetinib discontinuation should have a follow-up assessment after 30 days to evaluate the potential for reversibility.

3.1.11 Tanner Stages in Paediatric Subjects

For paediatric subjects, Tanner stages will be assessed. This looks at sexual maturity rating and is assessed according to the local site.

3.1.12 Duration of Exposure

The total duration of exposure (months) to selumetinib will be calculated as:

Total (or intended) exposure of selumetinib:

- Total (or intended) exposure (days) = min(last dose date where dose > 0 mg, date of DCO) – first dose date + 1.
- Total (or intended) exposure (months) = Total exposure (days) / (365.25/12).

The actual exposure of selumetinib will be the total duration of exposure excluding dose interruptions (as recorded on the exposure page of the CRF):

- Actual exposure = total exposure – total duration of dose interruptions, where total exposure will be calculated as above and a dose interruption is defined as any day with dose = 0 mg.

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. For subjects still continuing treatment at DCO, the date of DCO will be used as the date of last dose and the total treatment duration and actual treatment duration calculated in the same way.

Total observation duration (months) will be calculated as:

- Total observation duration = (min(study discontinuation date, date of DCO) – first dose date + 1) / (365.25/12).

Subjects who permanently discontinue during a dose interruption

If a subject permanently discontinues study treatment during a dose interruption, then the date of last administration of study medication will be used in the calculation of exposure.

3.2 Pharmacokinetic Variables

A pharmacokinetic evaluation of selumetinib and its metabolite N-desmethyl selumetinib will be performed after the first dose and at steady state of selumetinib administration for all consenting adult and paediatric subjects. This analysis will be performed by Covance (Contract Research Organization), on behalf of AstraZeneca R&D.

The PK blood sampling time points are shown in Table 7 (single dose) and Table 8 (multiple doses).

Table 7: Pharmacokinetic sampling schedule (single dose)

Cycle 0	Day	Pre-dose [a]	30 min [b]	1 hr [c]	1.5 hr [c]	3 hr [c]	6 hr [c]	8 hr [c]	12 hr [c]	24 hr [c]	30 hr [c]
	1	x	x	x	x	x	x	x	x	x	x

[a] Within 10 minutes prior to dosing.
[b] Time allowance: ± 5 minutes.
[c] Time allowance: ± 10 minutes.
hr = hour; min = minute.

Table 8: Pharmacokinetic sampling schedule (multiple doses)

Cycle 1	Day	Pre-dose [a]	30 min [b]	1.5 hr [c]	3 hr [c]	6 hr [c]	12 hr [c]
	8	x	x	x	x	x	x

[a] Within 10 minutes prior to dosing.
[b] Time allowance: ± 5 minutes.
[c] Time allowance: ± 10 minutes.
hr = hour; min = minute.

The following PK parameters (see Table 9) will be determined where possible from the plasma concentration of selumetinib and its metabolite N-desmethyl selumetinib using non compartmental methods with Phoenix® WinNonlin® (Certara, L.P., Version 8.1 or higher).

Table 9: PK parameters definitions

PK Parameter	Definition
Single dose	
AUCinf	Area under plasma concentration-time curve from time zero to infinity
AUClast	Area under the plasma concentration-time curve from time zero to time of last quantifiable concentration, calculated using the linear/log trapezoidal rule (linear-up-log-down)
AUC(0-6)	Area under the plasma concentration-time curve from zero to 6 hours post-dose
AUC(0-12)	Area under the plasma concentration-time curve from zero to 12 hours post-dose
Cmax	Maximum observed plasma concentration
tmax	Time to reach maximum observed plasma concentration
t½λz	Half-life associated with terminal slope (λz) of a semi-logarithmic concentration-time curve, calculated as ln2/λz
λz	Terminal elimination rate constant
CL/F	Apparent total body clearance of drug from plasma after extravascular administration, calculated as Dose/AUCinf (selumetinib only)
Vz/F	Apparent volume of distribution during the terminal phase after extravascular administration, calculated as Dose/(λz *AUCinf) (selumetinib only)
Dose normalised Cmax	Dose normalised maximum observed plasma concentration, calculated by Cmax divided by selumetinib dose
Dose normalised AUClast	Dose normalised area under the plasma concentration-time curve from time zero to time of last quantifiable concentration, calculated by AUClast divided by selumetinib dose
Dose normalised AUCinf	Dose normalised area under plasma concentration-time curve from time zero to infinity, calculated by AUCinf divided by selumetinib dose
Dose normalized AUC(0-6)	Dose normalized area under plasma concentration-time curve from zero to 6 hours post- dose, calculated by AUC(0-6) divided by selumetinib dose
Dose normalised AUC(0-12)	Dose normalised area under plasma concentration-time from zero to 12 hours post-dose, calculated by AUC(0-12) divided by selumetinib dose
Dose and BSA normalised Cmax	Dose and body surface area (BSA) normalized maximum observed plasma concentration, calculated by Cmax divided by selumetinib dose and then multiplied by subject BSA. The unit is (ng/mL/mg*m2)

PK Parameter	Definition
Dose and BSA normalised AUClast	Dose and BSA normalized area under the plasma concentration-time curve from time zero to time of last quantifiable concentration, calculated by AUClast divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2)
Dose and BSA normalised AUCinf	Dose and BSA normalized area under plasma concentration-time curve from time zero to infinity, calculated by AUCinf divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2)
Dose and BSA normalised AUC(0-6)	Dose and BSA normalized area under plasma concentration-time curve from zero to 6 hours post-dose, calculated by AUC(0-6) divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2).
Dose and BSA normalised AUC(0-12)	Dose and BSA normalized area under plasma concentration-time curve from zero to 12 hours post-dose, calculated by AUC(0-12) divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2).
Metabolite:Parent ratio	Metabolite:parent ratio, calculated by AUClast(metabolite)/AUClast(parent), Cmax(metabolite)/Cmax(parent)
Multiple dose	
Cmax,ss	Maximum observed plasma concentration at steady state
tmax,ss	Time to reach maximum observed plasma concentration at steady state
AUC(0-6),ss	Area under the plasma concentration-time curve from zero to 6 hours post-dose at steady state
AUC(0-12),ss	Area under the plasma concentration-time curve from zero to 12 hours post-dose at steady state
Rac AUC	Accumulation ratio based on AUC(0-12), calculated as the ratio of the AUC(0-12),ss following multiple dosing by the AUC(0-12) following the first dose
Rac Cmax	Accumulation ratio based on Cmax, calculated as the ratio of the Cmax,ss following multiple dosing by the Cmax following the first dose
Dose normalised Cmax,ss	Dose normalised maximum observed plasma concentration at steady state, calculated by Cmax,ss divided by selumetinib dose
Dose normalised AUC(0-6),ss	Dose normalized area under the plasma concentration-time curve from zero to 6 hours post-dose at steady state, calculated by AUC(0-6),ss divided by selumetinib dose
Dose normalised AUC(0-12),ss	Dose normalised area under the plasma concentration-time curve from zero to 12 hours post-dose at steady state, calculated by AUC(0-12),ss divided by selumetinib dose
Dose and BSA normalized Cmax,ss	Dose and BSA normalized maximum observed plasma concentration at steady state, calculated by Cmax,ss divided by selumetinib dose and then multiplied by subject BSA. The unit is (ng/mL/mg*m2)
Dose and BSA normalized AUC(0-6),ss	Dose and BSA normalized area under plasma concentration-time curve from zero to 6 hours post-dose at steady state, calculated by AUC(0-6),ss divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2)

PK Parameter	Definition
Dose and BSA normalized AUC(0-12),ss	Dose and BSA normalized area under plasma concentration-time curve from zero to 12 hours post-dose at steady state, calculated by AUC(0-12),ss divided by selumetinib dose and then multiplied by subject BSA. The unit is (h*ng/mL/mg*m2)
Metabolite:Parent ratio	Metabolite:parent ratio, calculated by AUC(0-12),ss(metabolite)/ AUC(0-12),ss (parent), Cmax,ss(metabolite)/Cmax,ss(parent)
TCP	Temporal change parameter in systemic exposure, calculated by AUC(0-12),ss/first dose AUCinf, Cmax,ss/first dose Cmax

The following diagnostic parameters (see Table 10) for plasma PK analysis will be listed, but not summarised:

Table 10: Diagnostic parameters for plasma PK analysis definitions

Diagnostic parameter(s)	Definition
λz lower	Lower (earlier) t used for λz determination
λz upper	Upper (later) t used for λz determination
λzN	Number of data points used for λz determination
λz span ratio	Time period over which λz was determined as ratio of t½λz
Rsq adj	Statistical measure of fit for the regression used for λz determination adjusted for the number of used data points (n obs)
AUCextr	Extrapolated area under the curve from tlast to infinity, expressed as percentage of AUCinf
tlast	Time to last quantifiable concentration

Additional PK parameters may be determined where appropriate.

3.3 Efficacy Variables

3.3.1 Derivation of REiNS Tumour Response

For all subjects, the REiNS tumour response data will be used to determine each subject's visit response according to REiNS criteria. Tumour assessments will be measured by 3D volumetric magnetic resonance imaging (MRI). Baseline radiological tumour assessments are to be performed in the 28-day screening period. Unless clinically indicated otherwise, tumour assessments of the target and non-target PN will be obtained at screening and every 4 cycles (16±1 weeks) relative to the date of first dose for the first 2 years (24 cycles). From the end of Cycle 24, tumour assessments will be performed every 6 cycles (24±1 weeks) as long as the subject remains on study treatment or until disease progression. For subjects who discontinue study treatment due to reasons other than disease progression, tumour assessments will

continue for approximately 1 year, if clinically applicable, and the assessment schedule should continue per the regular tumour assessment interval.

To minimise bias, tumour assessments will be done by both the investigator and by independent central review (ICR) according to REiNS criteria. All radiological scans for all subjects (including those at unscheduled visits, or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca appointed imaging Contract Research Organisation (CRO) for central analysis. The details about target lesion location and non-target lesion location (if relevant) will also be collected and sent to an AZ appointed imaging CRO to ensure the independent reviewer also follows the most clinically relevant PN. Further details of the ICR will be documented in the Independent Review Charter (IRC).

The results of the assessments by the investigator and ICR will be summarised separately, and all analyses detailed below will be repeated for both the investigator data and the ICR data, unless otherwise stated.

If an unscheduled assessment is performed, and the subject has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits. This schedule is to be followed in order to minimise any unintentional bias caused by some subjects being assessed at a different frequency than other subjects.

To determine the level of response, follow-up scans will be compared to the pre-treatment scan, or the scan at the time of best response after documenting a partial response, for the same tumour. Using the REiNS criteria, a change in volume of 20% is used to indicate an increase or decrease of tumour size. For all subjects, the target lesion (TL – see Section 3.3.1.1) and non-target lesion (NTL – see Section 3.3.1.2), if applicable, will be assessed and both given a response (see Table 11 and Table 12), and any new PNs will be recorded. At each visit, subjects will be programmatically assigned an overall REiNS visit response of complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD), using the information from TLs, NTLs and new PNs, and depending on the status of their disease compared with baseline and previous assessments (see Table 13). If a subject has a tumour assessment that cannot be evaluated, then the subject will be assigned a visit response of not evaluable (NE).

3.3.1.1 Target Lesions (TLs)

The target lesion is selected at screening. The target PN is defined as the most clinically relevant PN that is also measurable and either typical or nodular and must not be solitary nodular. If there is a second PN that is also considered clinically relevant, this may be identified as a non-target lesion at baseline.

Table 11: Target lesion visit responses (REiNS)

Visit Responses	Description
Complete response (CR)	Disappearance of the TL.
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.
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 TL assessment.

3.3.1.2 Non-target Lesions (NTLs) and New PNs

Only one NTL may be selected. NTLs must also be clinically relevant, measurable and either typical or nodular and must not be solitary nodular. This section provides the definitions of the criteria used to determine and record a response for the NTL at each MRI assessment.

NTL response will be defined as follows:

Table 12: Non-target lesion 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 lesion recorded at baseline.
Not evaluable (NE)	Data unavailable for NTL assessment.

Details of any new PNs will also be recorded in the CRF with the date of the first scan that revealed the new lesion(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 TL and the NTL, if applicable, also qualifies for disease progression, irrespective of

the TL and NTL response. The new PN must be confirmed by scans with lesion details recorded in the CRF. Once a new PN has been identified, at subsequent time points 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, further details can be found in the IRC.

Worsening of existing symptoms or the appearance of new symptoms that persist for more than 7 days and that are felt to be definitely related to the PN should be evaluated by repeating the MRI per the protocol. Subjects should not be classified as having PD solely on the basis of new or increased symptoms.

3.3.1.3 Overall Visit Response

Table 13 defines how the previously defined TL and NTL visit responses will be combined with new PN information to give an overall visit response.

Table 13: Overall visit responses

TARGET	NON-TARGET	NEW PN	OVERALL VISIT RESPONSE
CR	Non-PD or NE or NA	No or NE	**CR**
CR	Non-PD or NE or NA	Yes	**PD**
CR	PD	Any	**PD**
PR	Non-PD or NE or NA	No or NE	**PR**
PR	Non-PD or NE or NA	Yes	**PD**
PR	PD	Any	**PD**
SD	Non-PD or NE or NA	No or NE	**SD**
SD	Non-PD or NE or NA	Yes	**PD**
SD	PD	Any	**PD**
PD	Any	Any	**PD**
NE	Non-PD or NE or NA	No or NE	**NE**
NE	Non-PD or NE or NA	Yes	**PD**
NE	PD	Any	**PD**

NA = Not applicable, for subjects where no non-target lesion was recorded at baseline.

NE = Not evaluable, for lesions which do not have available data at the particular visit.

The overall visit response will be derived programmatically based on the site investigator data for TL, NTL and new PNs as described in the previous sections. ICR will define the overall visit response (i.e. the response obtained overall at each visit by assessing TLs, NTLs and new PNs) and no programmatic derivation of visit response will be necessary.

In the absence of full volume at a post-baseline investigator assessment, partial volumes of the target or non-target lesion at that post-baseline assessment will be measured. The following steps must be taken in this situation:

1. If a full volume measurement cannot be collected at any post-baseline assessment (Visit x), the investigator must identify which section of the lesion is measurable, and take a partial measurement.
2. The investigator must then go back to the baseline scan, identify the section of the lesion which remains visible at Visit x and recalculate the corresponding partial volume at baseline. This partial volume at baseline can be used to verify if the Visit x partial volume meets the criteria for CR/PR/SD/PD.
3. If there is a post-baseline visit prior to Visit x, when the subject had a response of PR and the volume is nadir, the investigator will need to go back to this visit and again identify the corresponding section of the lesion which remains visible at Visit x, take the corresponding partial volume at nadir, and use this to verify if the measurable volume at Visit X qualifies for PD.

Further details of the handling of partial volumes in the ICR data can be found in the IRC.

3.3.2 Objective Response Rate (ORR)

The ORR is defined as the proportion of subjects in the SAF who have a CR or confirmed partial response (PR, defined as a target PN volume decrease ≥ 20% compared to baseline) per REiNS criteria. A confirmed response is defined as a response that is recorded at one visit and confirmed by a consecutive scan within 3 to 6 months after the initial first response.

Data obtained up until progression or last evaluable MRI assessment in the absence of progression will be included in the assessment of ORR.

However, any CR or a confirmed partial response (cPR) which occurred after a subsequent NF1 PN treatment (following study treatment discontinuation) will not be included in the numerator for the ORR calculation (where the SAF will be the denominator).

3.3.2.1 Best overall response (BOR)

Best overall response (BOR) is calculated based on the overall visit responses from each MRI assessment, described in Section 3.3.1.3. It is the best response a subject has had following the start of treatment, but prior to progression or the last evaluable MRI assessment in the absence of progression, or the start of subsequent NF1 PN treatment following study treatment discontinuation. Categorisation of BOR will be based on REiNS using the following response categories: CR, cPR, PR, SD, PD and NE.

The different ways of achieving a BOR of CR, cPR, or PR are shown in Table 14.

Table 14: BOR Categories

Visit N	Visit N+1	BOR
PR	PR	Confirmed PR
PR	SD/PD/NE	Unconfirmed PR
CR	CR/PD/NE	CR
PR/SD/NE	CR	CR
NE	NE	NE

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, or the start of subsequent NF1 PN treatment following study treatment discontinuation. The denominators will be consistent with those used in the ORR analysis.

For subjects whose progression event is death, BOR will be calculated based upon all evaluable MRI assessments prior to death.

3.3.3 Duration of Response (DoR)

For the subset of subjects who have a CR or cPR, duration of response is defined as the time from the date of the MRI assessment of the first documented response (which is subsequently confirmed in the case of PR) until the date of the MRI assessment of documented progression or death (by any cause in the absence of progression) regardless of whether the subject withdrew from study treatment or received another NF1 PN treatment (following study treatment discontinuation) prior to progression (i.e. date of PFS event or censoring – date of first response + 1), per REiNS criteria. The end of response should coincide with the MRI assessment of progression or death from any cause used for the PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first response of PR or CR. Duration of response is also defined in terms of cycles (i.e. date of PFS event or censoring – date of first response + 1)/28.

If a subject does not progress following a response, then their DoR will use the PFS censoring. The DoR will be derived based on the actual MRI assessment dates and not visit dates.

3.3.4 Progression Free Survival (PFS)

Progression free survival is defined as the time from the date of first dose of selumetinib until the MRI assessment of objective disease progression per REiNS criteria, as assessed by the investigator and ICR, or death (by any cause in the absence of progression) regardless of whether the subject withdrew from study treatment or received another NF1 PN treatment (following study treatment discontinuation) prior to progression (i.e. date of PFS event or censoring – date of first dose + 1). Progression free survival is also defined in terms of cycles (i.e. date of PFS event or censoring – date of first dose + 1)/28.

Subjects who have not progressed or died at the time of analysis will be censored at their last evaluable MRI assessment. However, if the subject progresses or dies after two or more consecutive missed MRI assessments, the subject 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 15.

If the subject has no evaluable MRI assessments or does not have baseline data, 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 15: 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

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:

- If the previous assessment is prior to study day 107 (i.e. the two missed visits are those immediately after the baseline assessment), two missing visits will equate to 233 days since the last assessment (2 x 16 weeks + 1 week for a late assessment + 2 days allowing for Cycle 0).
- If the previous assessment is on or after study day 107 and prior to study day 555, two missing visits will equate to 34 weeks since the last assessment (2 x 16 weeks + 1 for early missing visit + 1 for a late assessment).
- If the two missed visits occur over the period when the scheduled frequency of assessments changes from every 4 cycles to every 6 cycles this will equate to 42 weeks (i.e. take the average of 16 and 24 weeks which gives 20 weeks and then apply same

rationale, hence 2 x 20 weeks + 1 week for an early assessment + 1 week for a late assessment = 42 weeks). The time period for the previous assessment will be from study days 555 to 666.
- From Cycle 24 (day 667) onwards (when the scheduling changes to every 6 cycles), two missing visits will equate to 50 weeks (i.e. 2 x 24 weeks + 1 week for an early assessment + 1 week for a late assessment = 50 weeks.

MRI assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- For both ICR and investigational assessments, the date of progression will be determined based on the earliest of the dates of the component that triggered the progression.
- For both ICR and investigational assessments, when censoring a subject for PFS the subject will be censored at the latest of the dates contributing to a particular overall visit assessment.

3.3.5 Time to Progression (TTP)

Time to progression is defined as the time from the date of first dose of selumetinib until the MRI assessment of objective disease progression regardless of whether the subject withdrew from study treatment or received another NF1 PN treatment (following study treatment discontinuation) prior to progression (i.e. date of PD – date of first dose + 1). Time to progression is also defined in terms of cycles (i.e. date of PD – date of first dose + 1)/28.

Subjects who have not progressed at the time of analysis will be censored at their last evaluable MRI assessment. However, if the subject progresses after two or more missed cycles, the subject will be censored at the latest evaluable MRI assessment.

If the subject has no evaluable MRI assessments or does not have baseline data, they will be censored at day 1. The TTP time will always be derived based on actual MRI assessment dates and not visit dates.

3.3.6 Time to Response (TTR)

Time to response is defined as the time from the date of first dose until the date of the MRI assessment of the first documentation of CR or a subsequently confirmed PR as determined per REiNS criteria (i.e. date of first response – date of first dose + 1). The date of first documented response should coincide with that used for the DoR endpoint. Time to response is also defined in terms of cycles (i.e. date of first response – date of first dose + 1)/28.

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

Only subjects who have achieved a complete response or a confirmed partial response will be evaluated for TTR.

3.3.7 Change in TL and NTL Tumour Size

Change in PN growth will also be evaluated using the baseline MRI scan and the scans through treatment, for both the TL and NTL. Both the absolute and percentage change in PN volume from baseline will be calculated for each assessment. See Section 4.1.1 for baseline definition.

A windowing rule will be applied and will follow the protocol allowed visit window; therefore, any REiNS scan performed within ± 1 week of the protocol scheduled visit will be used for that visit.

The best percentage change from baseline will also be summarised, where the best change is the largest decrease in volume from baseline, or the smallest increase in the absence of a decrease and will include all assessments prior to the earliest of progression, death in the absence of progression, the start of subsequent NF1 PN treatment, or the last evaluable MRI assessment if the subject has not died or progressed. All assessments should be included in the derivation of the best percentage change from baseline, including unscheduled assessments.

3.4 Clinical Outcome Assessment (COA) Variables

In this study clinical outcome assessments (COAs) will include both patient-reported outcomes and observer-reported outcomes. Health-related quality of life, pain evaluations, functional patient reported outcome (PRO) assessments and functional evaluations relevant to the target PN will be performed in all subjects at baseline (except PGIC) and then at the end of every cycle until Cycle 4, at the end of every second cycle up to the end of Cycle 12, at the end of every fourth cycle up to Cycle 24, and every six cycles thereafter, as long as the subject remains on study treatment. COAs will be performed electronically using on-site tablets, except the pain medication survey which will be done on paper.

Table 16 provides an overview of expected subjects COA questionnaires completion by age of subject at enrolment. A parent (or the legal guardian) of subjects in the paediatric cohort will complete the parent proxy measures.

Table 16: Overview of eligible subjects completing COA questionnaires, by age of subject and respondent

Clinical outcome assessment questionnaire	Respondent	Age
FLACC	Parent/guardian-reported	3 years
Faces pain scale - revised	Self-reported	4 to 17 years

Clinical outcome assessment questionnaire	**Respondent**	**Age**
NRS-11	Self-reported	Adults
Pain Interference Index	Self-reported	8 to 17 years Adults
	Parent/guardian-reported	5 to 17 years
Pain Medication Survey	Self-reported	Adults
	Parent/guardian-reported	Paediatric cohort
PedsQL	Self-reported	5 to 7 years
		8 to 12 years
		13 to 17 years
	Parent/guardian-reported	3 to 4 years
		5 to 7 years
		8 to 12 years
		13 to 17 years
EORTC QLQ-C30	Self-reported	Adults
PlexiQoL	Self-reported	Adults
PROMIS (mobility)	Self-reported	8 to 17 years
	Parent/guardian-reported	5 to 17 years
PROMIS (upper extremity function)	Self-reported	8 to 17 years
	Parent/guardian-reported	5 to 17 years
PROMIS (Physical Function – Short Form 8c 7-day)	Self-reported	Adults
PGIS	Self-reported	8 to 17 years

Clinical outcome assessment questionnaire	Respondent	Age
		Adults
	Parent/guardian-reported	Up to 17 years
PGIC	Self-reported	8 to 17 years
		Adults
	Parent/guardian-reported	Up to 17 years

3.4.1 Pain

The effect of selumetinib on pain will be evaluated using the FLACC scale (children aged 3 years), Faces pain scale - revised (children aged 4 to 17 years), NRS-11 (adult cohort only), PII (adult cohort; self- and parent-reported in the paediatric cohort) and pain medication survey (all subjects).

3.4.1.1 FLACC pain scale (3 years)

The FLACC scale (Merkel et al, 1997) is an observer-reported measurement used to assess pain in children aged from 2 months to 7 years, or for individuals that are unable to communicate their pain. In this study, parents/guardians will receive a short training on how to do the assessments of pain for their child. The scale is scored from 0 to 10, with 0 representing no pain. The scale has 5 criteria (face, legs, activity, cry, and consolability), which each are assigned a score of 0 to 2. In this study, the FLACC scale will be used to assess pain in children 3 years of age. Changes in score compared with baseline will be evaluated.

3.4.1.2 Faces pain scale (4 to 17 years)

The Faces pain scale - revised is a self-reported measure of pain intensity developed for children (Hicks et al, 2001). Facial images depicting no pain (score of 0) to very much pain (score of 10) are presented to the child, who will then pick the face that matches their own discomfort over the past week, giving a score. In this study, the Faces pain scale will be used to assess pain in children aged from 4 to 17 years of age. Changes in score compared with baseline will be evaluated.

3.4.1.3 NRS-11 (adult cohort only)

The NRS-11 consists of 3 questions scored on an 11-point scale, where subjects are asked to select the one number from 0 to 10 that best describes their worst pain over the past 7 days,

with 0 representing “no pain” and 10 representing “worst pain you can imagine” (Hawker et al, 2011). The 3 questions to describe pain on the NRS-11 are in relation to:

1) Physician selected target-tumour pain.
2) Overall tumour pain.
3) Overall pain.

The primary outcome for the self-reported NRS-11 will be the rating of pain in the target tumour selected by the physician at baseline. Only adult subjects will complete the NRS-11 for this study. Changes in score compared with baseline will be evaluated.

3.4.1.4 Pain Interference Index (PII)

The PII is a 6-item measure that assesses the extent to which pain has interfered with daily activities in the past week (Martin et al, 2015). A parallel parent version of the PII was also created (Kemani et al, 2016). For this study, adults and children from 8 to 17 years of age will complete self-reported PII, and parents or the legal guardian of children from 5 to 17 years of age will complete the parent proxy PII.

Items are rated on a 7-point Likert scale (0=not at all to 6=completely), and the total score is the mean of the completed items. The total score will be computed if more than 50% of items are answered (e.g. 4 out of 6). Changes in score compared with baseline will be evaluated.

3.4.1.5 Pain Medication Survey

A pain medication survey will be completed. Information on daily and as-needed pain medications will be collected, including dates, pain medication name, dose and frequency.

3.4.2 Health-related Quality of Life (HRQoL)

The effect of selumetinib on HRQoL will be evaluated using the PedsQL (paediatric cohort) and EORTC QLQ-C30 and PlexiQoL (adult cohort).

3.4.2.1 PedsQL (paediatric cohort)

The PedsQL Core version 4 will be used to assess HRQoL in the paediatric cohort (Varni et al, 2001). The PedsQL 4.0 Generic Core Scales are multidimensional child self-reported and parent proxy reported scales. PedsQL is a brief standardised paediatric HRQoL scale with good reliability and validity, which includes both generic and disease-specific modules. It consists of 23 questions across 4 subscales and yields a total score to give a measure of global HRQoL. The 4 subscales are:

- physical functioning (8 items)
- emotional functioning (5 items)
- social functioning (5 items)
- school functioning (5 items)

Each of the 23 items is scored on a 5-point Likert scale (0=never a problem; 1=almost never a problem; 2=sometimes a problem; 3=often a problem; 4=almost always a problem). For self-reported and parent/guardian reported separately, items are reverse-scored and linearly transformed to a 0-100 scale (0=100, 1=75, 2=50, 3=25, 4=0), where higher scores indicate better HRQoL. The scale score is calculated as the sum of the items divided by the number of items answered (to account for missing data), and a Total Scale Score is the average across all items answered across all the 4 scales. If more than 50% of the items in the scale are missing, the scale score is not computed (Nutakki et al, 2018). In addition, scale scores will be calculated using the raw item scores (without the linear transformation, but still reversed). Change in scores from baseline will also be calculated.

The PedsQL Core version 4 has been developed in age-specific versions. In this study, the self-reported versions are for subjects aged 5 to 7 years, 8 to 12 years, and 13 to 17 years. In addition, the corresponding parent-reported versions will also be included. For subjects 3 to 4 years of age, only the parent-reported version will be included. Acute versions of the PedsQL questionnaires (7-day recall period) will be used in this study.

The primary outcomes for HRQoL will be the Total Scale Score of the self-reported PedsQL for children ≥5 years of age and the Total Scale Score of the parent PedsQL administered to parents of children 3 to 4 years of age.

Secondary outcomes for HRQoL will be the mean scores of the 4 domains (physical, emotional, social, and school) of the self-reported scale completed by children ≥5 years of age and the 4 domain mean scores from the parent-reported scale given to parents of children 3 to 4 years of age.

3.4.2.2 EORTC QLQ-C30 and PlexiQoL (adult cohort)

For the adult cohort, the validated measures EORTC QLQ-C30 (Aaronson et al, 1993) and PlexiQoL (Heaney et al, 2019) will be used for assessments of HRQoL.

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce the following scales:

- 5 multi-item functional scales (physical, role, emotional, cognitive, and social)
- 3 multi-item symptom scales (fatigue, pain, and nausea/vomiting)
- a 2-item global measure of health status
- 6 single items assessing additional common symptoms (dyspnoea, loss of appetite, insomnia, constipation, diarrhoea, and financial difficulties)

An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales, functional scales, 6 single items (dyspnoea, loss of appetite, insomnia,

constipation, diarrhoea, and financial difficulties) and the global health status scale according to the EORTC QLQ-C30 Scoring Manual (Fayers et al, 2001).

Higher scores on the global health status/QoL and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity.

Changes in score compared with baseline will be evaluated. See Section 4.1.1 for the definition of baseline. For each subscale, if < 50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al, 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing.

The PlexiQoL may be considered ancillary to the EORTC QLQ-C30. It contains 18 items measuring quality of life. The impacts captured by this instrument, including shame/discomfort/embarrassment, negative body image, and impact on role, are not captured by the EORTC QLQ-C30.

The PlexiQoL consists of several questions that are each answered as “true” if the statement applies or “false” if the statement does not apply. The time frame for the assessment is “at the moment.” These statements assess whether the subject avoids crowds, are unable to join activities with family and friends, etc. Each statement is given a score of 1 = True or 0 = Not true. A response of ‘true’ indicates adverse quality of life. All item scores are then summed to give a total score from 0 (good QoL) to 18 (poor QoL). If less than 4 of the responses are missing, the total score is divided by the number of non-missing items and multiplied by 18. If more than 3 of the items in the scale are missing, the total score is not computed. Changes in score from baseline will also be evaluated.

3.4.3 Physical Functioning

PROMIS is a set of person-centred measures that evaluates and monitors physical, mental, and social health in adults and children. The PROMIS Physical Functioning Scales (PROMIS website, 2016) assess the domains of Mobility and Upper Extremity Function and include mobility items such as “I can walk upstairs without holding on to anything” and upper extremity items such as “I can button my shirt or pants.” Parent proxy items are parallel to child items.

Self-reported short forms will be administered to children from 8 to 17 years of age. Parallel parent proxy forms will be administered for children aged from 5 to 17 years. The PROMIS paediatric short forms consists of 8 items using a 5-point Likert scale format (i.e. 1 = unable to do, 5 = can do without any difficulty). Data permitting, raw scores are converted to a T-score, which is based on reference data from the US general population, where mean = 50, standard deviation (SD) = 10. Raw scores will be analysed. Higher scores indicate better physical

functioning. Change from baseline (items scores and raw domain scores) will also be calculated at each assessment. In addition, T-scores will be calculated as above, and analysis for raw scores will be repeated for T-scores when appropriate.

In the adult cohort, a separate newly developed Physical Function Short Form 8c 7-day scale will be used, also with 8 items measured on a 5-point scale as above. Raw scores will be analysed. Changes in score compared with baseline (item scores and raw domain scores) will be evaluated. In addition, T-scores will be calculated as above, and analysis for raw scores will be repeated for T-scores when appropriate.

3.4.4 Patient's Global Impression of Symptom Severity

The PGIS instrument is a single-item scale that evaluates symptom severity (separate single items for tumour pain, overall pain and tumour-related problems) at different points in time. Global impression of symptom severity will be measured by asking the subject or parent to best describe how his/her symptoms are on a 6 point scale (0 = no symptoms, 1 = very mild, 2 = mild, 3 = moderate, 4 = severe, 5 = very severe) where a lower score demonstrates a better result.

The PGIS instrument will be parent-reported for children aged up to 17 years and self-reported for children aged 8 to 17 years and adults.

3.4.5 Patient's Global Impression of Change

The PGIC instrument is a single-item scale that evaluates the subject's perspective of changes in their tumour pain, overall pain and tumour-related problems. An adapted version of the PGIC instrument will be used in this study at all visits post baseline. A parallel parent proxy form will also be used. On the adapted PGIC, subjects (and their parents separately) will give the overall impression of change of the subject's: 1) tumour pain, 2) overall pain, 3) tumour related problems from before initiation of selumetinib to the current evaluation point.

PGIC will be parent-reported for children aged up to 17 years and self-reported for children aged 8 to 17 years and adults.

3.4.6 Compliance

Summary measures of compliance over time will be derived for each COA, respectively. These will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least one individual item completed.
- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time e.g. a questionnaire from a subject who has not withdrawn from the study at the scheduled assessment time. For subjects that have progressed, the latest of progression and safety follow-up will be used to assess whether the subject is

still under COA follow-up at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.

- Evaluable questionnaire = a questionnaire with a completion date and at least one domain, subscale or total score that is non-missing.
- Overall COA compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.

Compliance over time will be calculated separately for each visit, including baseline, as the number of subjects with an evaluable questionnaire at the time point (as defined above), divided by number of subjects still expected to complete questionnaires. Summaries will also be provided for the percentage of received questionnaires out of those expected, and the percentage of fully completed questionnaires out of those expected.

3.5 Other Outcomes

3.5.1 Disfigurement

Consenting subjects will have photographic documentation of visible PNs. Disfigurement assessments will be performed at screening and then at the end of Cycles 8, 16 and 24 as long as the subject remains on study treatment.

3.5.2 Prior and concomitant treatments

Any medication or vaccine including over-the-counter or prescription medicines, vitamins, and/or herbal supplements that the subject is receiving within the 28 days prior to initiation of study treatment or receives during the study will be recorded on the eCRF. Details include reason for use, dates of administration including start and end dates, dosage information including dose and frequency. Pain medications are recorded at enrolment in the pain medication survey, and during the study in the concomitant medications CRF.

Prior medications are those which started and stopped prior to study treatment initiation.

Concomitant medications are those which started post study treatment initiation or ongoing medications which started prior to study treatment initiation.

Incomplete medication start and stop dates will be imputed as detailed in Section 4.2.2.1. Missing coding terms should be listed and summarised as "Not coded".

4. ANALYSIS METHODS

4.1 General Principles

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

- For all summaries, data will be summarised by cohort unless otherwise specified.
- Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarised by the number of non-missing observations, mean, standard deviation, median, minimum, and maximum. For log-transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarised by frequency counts and percentages for each category.
- Unless otherwise stated, percentages will be calculated out of the population total.
- For continuous data, the mean and median will be rounded to 1 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 1 decimal place.
- Descriptive summaries for all endpoints will be based on the planned visit schedule unless otherwise stated (e.g. baseline, week 1, week 4, week 8 etc.).
- Where summaries are over time, study day will be calculated in relation to the date of first treatment.
- For all tumour response analyses, investigator assessments and ICR assessments will be analysed separately.
- SAS® version 9.4 or above will be used for all analyses.

There is no formal hypothesis testing on this study.

Demographic data and baseline subject characteristics will be analysed for the safety analysis set. PK will be analysed based on the PK analysis set. Efficacy, clinical outcome assessment, exposure and safety data will be analysed based on the safety analysis set.

4.1.1 Baseline Measurements and Change from Baseline Variables

Baseline is defined as the last non missing value obtained prior to the first dose of selumetinib.

Assessments on the day of the first dose where time is not captured will be considered prior to the first dose if such procedures are required by the protocol to be conducted before the first dose.

If two measures are recorded on the same day with no time (or on the same day with same time/timepoint) then the average 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 then the best value would be taken as baseline as this is the most conservative.

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 will be calculated as (post-baseline value - baseline value) / baseline value x 100.

4.2 Analysis Methods

Unless specifically mentioned, the following methods will be applied for each cohort separately. There will be a primary DCO when the last subject dosed has had the chance to complete their visit at the end of cycle 10. A final DCO will occur when the last subject dosed has had the chance to complete their visit at the end of cycle 24. In addition, an interim analysis will be conducted when last subject dosed has had the chance to complete their visit at the end of cycle 4 to assess PK, safety, tolerability and efficacy. Data analyses will be performed for each DCO and results will be reported.

4.2.1 Demographics and Baseline Characteristics

The following will be summarised by cohort for all subjects in the SAF:

- Subject disposition (including screening failure)
- Important protocol deviations
- Inclusion in analysis sets
- Demographics (age, sex, race and ethnicity)
- Subject characteristics at baseline (height, weight, BSA)
- Previous chemotherapy prior to this study
- Disease characteristics at baseline (including disease diagnosis of NF1 and PN)
- Medical history (past and current)
- Relevant surgical history
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation NF1 PN treatment

The medications will be coded following AZ standard drug dictionary / WHO Drug dictionary as applicable.

4.2.2 Safety

The safety analysis set will be used for all tables, figures and listings except where expressly noted.

4.2.2.1 General Considerations for Safety Assessments

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

- The time windows will be exhaustive so that data recorded at any time point has 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.
- All unscheduled visit data have the potential to be included in the summaries.
- The window for the visits following Cycle 1 Day 1 will be 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 Cycle 1 Day 1 visit will be Day 4). If an even number of days exists between two consecutive visits, then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for vital signs data (with 4 weeks between scheduled assessments) are:

- Cycle 1 Day 1: Study Day 3
- Cycle 1 Day 8: Study Day 10, visit window 4 – 20
- Cycle 1 Day 28: Study Day 30, visit window 21 – 44
- Cycle 2 Day 28: Study Day 58, visit window 45 – 72
- Cycle 3 Day 28: Study Day 86, visit window 73 – 100
- Cycle 4 Day 28: Study Day 114, visit window 101 – 128
- etc.

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- Listings should display all values contributing to a time point for a subject.
- For visit based summaries
    - If there is more than one value per subject 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 subject 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 or not the value is closest to the scheduled visit date.
- For summaries at a subject level, all values will be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a subject level statistic such as a maximum.

The following considerations are made for missing safety data, diagnostic dates, AE dates and concomitant medication dates:

- 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.
- Adverse events that have missing causality (after data querying) will be assumed to be related to study drug.
- For missing diagnostic dates, if day and/or month are missing use 01 and/or Jan. If year is missing, put the complete date to missing.
- For missing AE/concomitant medication start dates, the following will be applied:

    a. Missing day - impute the 1st of the month unless month is the same as month of the first dose of study drug then impute first dose date.

    b. Missing day and month - impute 1st January unless year is the same as first dose date then impute first dose date.

c. Completely missing - impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st of January of the same year as the end date.

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

  a. Missing day - Impute the last day of the month unless month is same as month of last dose of study drug then impute last dose date.

  b. Missing day and month – impute 31st December unless year is the same as last dose date then impute last dose date.

  c. Completely Missing – assume that AE is still present (i.e. do not impute a date).

- For all missing start/end dates, flags will be retained in the analysis datasets indicating where any programmatic imputation has been applied, and in such cases, any durations would not be calculated.

If a subject 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 – using the 1st of the month.
- For Missing day and month – using the 1st of January.

4.2.2.2 Adverse Events

Adverse events will be coded using the latest of current version of the MedDRA dictionary and graded according to CTCAE version 5.0. Data from all cycles of treatment will be combined in the presentation of adverse events, unless stated otherwise. All AEs, both in terms of MedDRA system organ class, preferred term and CTCAE grade, will be listed and summarised descriptively by count (n) and percentage (%). An event that occurred one or more times during the relevant period for the table will contribute one observation to the numerator of the proportion (unless otherwise stated). The denominator of the proportion will comprise all subjects in the safety set.

The summary tables, unless otherwise stated, will include any treatment-emergent AEs, which includes those that occurred on or after the date of the first dose of study treatment and within 30 days after the last dose of study treatment, and those which started before the first dose of study treatment but increased in severity on or after the date of the first dose of study treatment and within 30 days after the last dose of study treatment.

Summary information (the number and percent of subjects and the number of events) will be tabulated (in a single table) for:

- All AEs
- AEs possibly related to study treatment
- AEs of CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to study treatment
- AEs with an outcome of death*
- All serious adverse events
- SAEs causally related to study treatment
- AEs leading to discontinuation of study treatment
- AEs leading to dose interruption and reduction (separately)
- AEs of special interest
- AEs of special interest with CTCAE grade 3 or higher

Separate AE summaries of the number and percentage of subjects with AEs in each of the categories above, excluding those highlighted with an asterisk (*), will be produced by system organ class and preferred term. There will be an additional summary table for all AEs by system organ class, preferred term and maximum CTCAE grade.

A separate summary table will present all prior AEs, any AE occurring before the start of study treatment, which did not worsen during the study, by system organ class and preferred term. These will be included and identified as 'pre-treatment' in the AE listings.

For the paediatric cohort only, there will be an additional listing and summary tables by system organ class and preferred term for all serious adverse events (SAEs) irrespective of causality and for all AEs causally related to selumetinib occurring during the long-term safety follow-up period. This includes those which begin more than 30 days after the end of study treatment, until 1 year following the end of study treatment.

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

A truncated AE table of most common AEs, showing all events that occur in at least 10% of any cohort, will be summarised by preferred term, by decreasing frequency.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved), seriousness, investigator's assessment of CTCAE grade, action taken (interruption/reduction/discontinuation) and relationship to study drug. The key listings which will be produced are listed below:

- Adverse events with an outcome of death – key subject information
- Serious adverse events - key subject information
- Adverse events leading to discontinuation of investigational product – key subject information
- All adverse events - Listing of key information

- All AEs – long term safety follow up – listing of key information (Paediatric cohort only)

4.2.2.3 Laboratory Assessments

Laboratory data (haematology, clinical chemistry and quantitative urinalysis) will be summarised by scheduled visit using descriptive statistics over time in terms of absolute values and changes from baseline.

Shift tables showing CTCAE grade changes from baseline to maximum on treatment will be produced by laboratory parameters for haematology and clinical chemistry, where on treatment includes all data collected between the first dose of selumetinib until the earlier of last dose of study treatment + 30 days (+ 7-day window), or DCO. Laboratory parameters with bi-directional (low and high) CTCAE categorisations will be displayed separately from parameters with non bi-directionality.

Continuous data (excluding urinalysis) will also be presented graphically using boxplots for absolute values and change from baseline over each scheduled visit.

For liver chemistries, a listing will be produced for subjects with potential Hy's Law, where subjects have combined elevations (ALT or AST ≥ 3xULN and total bilirubin ≥ 2xULN) on-treatment, where the elevation of ALT or AST precedes or is at the same time as the elevation in total bilirubin.

Laboratory data for the variables listed in Table 6 and qualitative urinalysis (Glucose, Erythrocytes, Haemoglobin, Blood, Qualitative protein) will be listed. Site reference ranges will also be listed. Flags (H or L) will be applied to values falling outside reference ranges (which will be explicitly noted on these listings where applicable), and to values for which CTCAE grading applies.

4.2.2.4 Vital Signs and Growth

All vital signs and growth parameters (height, weight and BSA) will be summarised using descriptive statistics over time in terms of absolute values and change (absolute and percentage) from baseline by scheduled visit. All vital signs data will be listed.

A summary of baseline and maximum on-treatment blood pressure measurements will also be produced, where on treatment includes all data collected between the first dose of selumetinib until the earlier of last dose of study treatment + 30 days (+ 7-day window), or DCO.

Vital signs data will also be presented graphically using boxplots for absolute values and changes over each scheduled visit.

4.2.2.5 ECGs and Echocardiograms

ECG mean heart rate and echocardiogram LVEF will be summarised using descriptive statistics over time in terms of absolute values and change (absolute and percentage) from baseline by scheduled visit.

Shift tables showing overall assessment changes from baseline to worst on treatment value will be produced for ECG(normal/abnormal-not clinically significant/abnormal-clinically significant) and echocardiogram (normal/abnormal), where on treatment includes all data collected between the first dose of selumetinib until the earlier of last dose of study treatment + 30 days (+ 7-day window), or DCO.

All individual ECG/ECHO data will be listed, including a description of the abnormalities and relevant baseline result.

Echocardiogram LVEF will also be presented graphically using boxplots for absolute values and changes over each scheduled visit.

4.2.2.6 Performance Status

Performance status scores (Lansky and Karnofsky) will be summarised using descriptive statistics and listed, along with a text description of the score as described in Appendix G of the CSP.

4.2.2.7 Ophthalmologic Assessments

Ophthalmology examination data will be listed for subjects with abnormal results, including a description of the abnormalities and relevant baseline result.

4.2.2.8 Bone Growth Assessments

Data from monitoring bone growth will be listed for subjects in the paediatric cohort with abnormal results.

4.2.2.9 Tanner Stages

Tanner stages will be summarised using descriptive statistics and listed for subjects in the paediatric cohort.

4.2.2.10 Exposure

Total exposure (defined in 3.1.12) to selumetinib will be summarised by mean, standard deviation, median, minimum and maximum. The following summaries related to study treatment will be produced:

- Total exposure to selumetinib
- Actual exposure to selumetinib

- Number of subjects with at least
    - one dose interruption
    - one dose reduction

Dose interruptions and reductions will be summarised based on the information recorded in the exposure page of the CRF.

Total observation duration (defined in 3.1.12 will also be summarised using descriptive statistics.

4.2.3 Pharmacokinetic Data

All PK data and analyses will be summarised based on the PK analysis set. The data excluded from the PK analysis set will be flagged in subject listing only.

Calculation or Derivation of Pharmacokinetic Parameters

Pharmacokinetic analysis will, where possible, be carried out using actual times recorded in the raw data. If actual times are missing, nominal times may be used at the discretion of the PK scientist with approval from AstraZeneca.

In the PK analysis, plasma concentrations that are below the limit of quantification (BLQ) prior to the first quantifiable concentration will be set to a value of zero. After the first quantifiable concentration, any BLQ plasma concentrations will be set to missing for all concentration profiles. Where 2 or more consecutive concentrations are BLQ at the end of a profile, the profile will be deemed to have terminated and therefore any further quantifiable concentrations will be set to missing for the calculation of the PK parameters unless it is considered to be a true characteristic of the profile of the drug.

If an entire concentration-time profile is BLQ, the profile will be excluded from the PK analysis.

Cmax, tmax and tlast will be determined by visual inspection of the concentration-time profiles.

Where there are sufficient data, λz will be calculated by log-linear regression of the terminal portion of the concentration-time profiles and t½λz will be calculated as ln2/λz. For the determination of λz, the start of the terminal elimination phase for each subject will be defined by visual inspection and will be the first time point at which there is no systematic deviation from the log-linear decline in plasma concentrations (λz lower). The last point (λz upper) will be the time of the last quantifiable plasma concentration.

The choice of data points used to estimate λz will follow the general guidelines:

- If there is more than 1 phase, use only observations from the terminal phase.
- In general, a minimum of 3 quantifiable concentrations are required (n obs ≥ 3) and the recommended duration of time over which λz is evaluated should be at least three times the subsequently estimated terminal half-life (t½λz). Where t½λz is estimated over less than three half-lives, the value will be flagged in the data listings.
- Include the last quantifiable concentration.
- Include only observations after peak concentration.

The Rsq_adj value will also be calculated to show the goodness-of-fit of the log-linear regression taking into consideration the number of points used in the estimation (λzN). λz and associated PK parameters (e.g., t½λz, AUCinf, CL/F, and Vz/F) will only be reported when the Rsq_adj value is ≥ 0.8; λz and associated PK parameters will be flagged in the data listings when Rsq_adj value is < 0.8.

AUCs [AUC(0-6), AUC(0-6),ss, AUC(0-12), AUC(0-12),ss, AUClast, and AUCinf] will be calculated using the linear trapezoidal method when concentrations are increasing and the logarithmic trapezoidal method when concentrations are decreasing (linear up, log down). AUCinf is estimated by AUClast + Clast/λz where Clast is the observed last quantifiable drug concentration. The AUCinf value where the percentage extrapolation (AUCextr) is greater than 20% will be flagged in the data listings and excluded from descriptive statistics and inferential statistical analyses (if applicable) by agreement with AstraZeneca.

The minimum requirement for the calculation of AUCs will be the inclusion of at least 3 consecutive plasma concentrations above the lower limit of quantification (LLOQ), with at least one of these concentrations following Cmax.

Presentation of Pharmacokinetic Data

Pharmacokinetic concentration and parameter summaries will be presented based on PK analysis set, unless otherwise specified. Individual PK concentration and parameter listings will be presented based on safety analysis set. Data from any subjects not included in the PK analysis set or excluded from the descriptive summaries and inferential statistical analyses (if applicable) will be included in the individual data listings and flagged with an appropriate footnote.

A listing of PK blood sample collection times, as well as derived sampling time deviations and concentrations at each protocol scheduled time point will be provided for selumetinib and N-desmethyl selumetinib.

Plasma concentrations will be listed and summarised by cohort, analyte, dose level (if applicable), and nominal sample time using the following summary statistics:

- The geometric mean (Gmean, calculated as exp [μ], where μ is the mean of the data on a logarithmic scale)
- Geometric coefficient of variation (GCV; GCV% calculated as 100 √ [exp(s2) - 1], where s is the standard deviation of the data on a log scale)
- Gmean ± geometric standard deviation (gSD) (calculated as exp[arithmetic mean in log scale ± standard deviation in log scale])
- Arithmetic mean calculated using untransformed data (mean)
- Standard Deviation (SD) calculated using untransformed data
- Coefficient of variation (CV; CV% calculated using untransformed data)
- Median
- Minimum
- Maximum
- Number of observations (n)
- Number of observations below the lower limit of quantification (LLOQ; n below LLOQ)

The derived PK parameters except tmax will be summarised by cohort, analyte, and dose level (if applicable) using the following summary statistics:

- Gmean (calculated as exp [μ], where μ is the mean of the data on a logarithmic scale)
- GCV (GCV%, calculated as 100√(exp(s2) - 1), where s is the SD of the data on a log scale)
- Gmean ± gSD
- Arithmetic mean calculated using untransformed data (mean)
- SD calculated using untransformed data
- CV (CV%, calculated using untransformed data)
- Median
- Minimum
- Maximum
- Number of observations (n)

Geometric statistics (Gmean, Gmean ± gSD, and GCV%) will not be presented for t½λz, λz, CL/F, and Vz/F.

The following summary statistics will be presented for single dose and/or steady state tmax:

- Median
- Minimum
- Maximum
- Number of observations (n)

Diagnostic parameters (e.g. tlast, λz lower, λz upper, λzN, λz span ratio, Rsq_adj and AUCextr) will be listed only and not summarised.

For the calculation of summary statistics of PK parameters, all NR and NC values will be set to missing. Three values are required as a minimum for a PK parameter to be summarised.

Two values are presented as a minimum and maximum with the other summary statistics as NC. If one or more values for a given parameter is zero, then no geometric statistics will be calculated for that parameter and the results for geometric statistics will be set to NA (Not Applicable).

Handling of Non-Quantifiable Concentrations

Individual concentrations below the LLOQ of the bioanalytical assay will be presented as NQ (Not Quantifiable) in the listings with the LLOQ defined in the footnotes of the relevant tables, figures and listings (TFLs). Individual plasma concentrations that are Not Reportable will be reported as NR and those that are missing will be reported as NS (No Sample) in the listings. Plasma concentrations that are NQ, NR or NS will be handled as follows for the provision of descriptive statistics:

- Any values reported as NR or NS will be excluded from the summary tables and corresponding figures.
- At a time point where less than or equal to 50% of the concentration values are NQ, all NQ values will be set to the LLOQ, and all descriptive statistics will be calculated accordingly.
- At a time point where more than 50% (but not all) of the values are NQ, the mean, SD, Gmean, Gmean ± gSD, and GCV% will be set to NC (Not Calculated). The maximum value will be reported from the individual data, and the minimum and median will be set to NQ.
- If all values are NQ at a time point, no descriptive statistics will be calculated for that time point. The Gmean, mean, minimum, median and maximum will be reported as NQ, and GCV% and Gmean ± gSD will be reported as NC.
- The number of values below LLOQ (n < LLOQ) will be reported for each time point together with the total number of collected values (n).
- Three observations > LLOQ are required as a minimum for a plasma concentration to be summarised. Two observations > LLOQ are presented as minimum and maximum with the other summary statistics as NC.

Precision and Rounding Rules for Pharmacokinetic Data

For concentration data, the listings will be presented to the same number of significant figures as the data received from the bioanalytical laboratory; for PK parameters, the listings will be presented according to the following rules:

- Cmax and Cmax,ss – will be presented to the same number of significant figures as received from the bioanalytical laboratory
- tmax, tlast, λz lower and λz upper – will be presented as received in the data, usually to 2 decimal places
- AUCinf, AUClast, AUC(0-6), AUC(0-6),ss, AUC(0-12), AUC(0-12),ss, t½λz, CL/F, Vz/F, Rac AUC, Rac Cmax, dose normalised and BSA normalised Cmax and Cmax,ss,

dose normalised and BSA normalised AUCs [AUClast, AUCinf, AUC(0-6), AUC(0-6),ss, AUC(0-12), AUC(0-12),ss], AUCextr, Rsq_adj, and λz span ratio will be presented to 3 significant figures
- λz – will be presented to 4 significant figures
- λzN – will be presented as an integer (no decimals)

For PK concentration data, all descriptive statistics will be presented to 4 significant figures with the exception of the minimum and maximum which will be presented to 3 significant figures, and n and n < LLOQ which will be presented as integers.

For PK parameters data, the descriptive statistics will be presented according to the following rules and presented by cohort, analyte, and dose level (if applicable):

- Cmax, Cmax,ss, AUCinf, AUClast, AUC(0-6), AUC(0-6),ss, AUC(0-12), AUC(0-12),ss, t½λz, CL/F, Vz/F, Rac AUC, Rac Cmax, dose normalised and BSA normalised Cmax and Cmax,ss, dose normalised and BSA normalised AUCs [AUClast, AUCinf, AUC(0-6), AUC(0-6),ss, AUC(0-12), AUC(0-12),ss] – all descriptive statistics will be presented to 4 significant figures with the exception of the minimum and maximum which will be presented to 3 significant figures
- λz – all descriptive statistics will be presented to 5 significant figures with the exception of the minimum and maximum which will be presented to 3 significant figures
- tmax – all descriptive statistics will be presented as received in the data, usually to 2 decimal places
- Number of values (n) - present as an integer

Graphical presentation of PK Data

The plasma concentrations of selumetinib and N-desmethyl selumetinib from the first dose in Cycle 0 Day 1 and at steady state in Cycle 1 Day 8 in Chinese paediatric and adult subjects will be displayed graphically as appropriate.

Individual plasma concentrations versus actual time will be plotted in linear and semi-logarithmic scale with all PK days overlaid on the same plot and separate plots for each subject and analyte, based on safety analysis set.

Combined individual plasma concentration versus actual times will be plotted in linear and semi-logarithmic scale with separate plots for cohort, PK day and analyte, based on the PK analysis set will be presented.

Geometric mean plasma concentration (±gSD) versus nominal sampling time will be plotted in linear and semi-logarithmic scale with all PK days overlaid on the same figure with separate plots for each analyte, based on the PK analysis set will be presented.

For consistency, the plasma concentration values used in the mean plots will be those given in the summary table for each time point.

Plasma concentrations that are NQ will be handled as follows for display in figures:

- For Gmean concentration-time plots: NQ concentrations will be handled as described for the descriptive statistics. If this handling results in a Gmean of "NQ", then the value plotted at that time point will be zero for linear plots and set to missing for semi logarithmic plots.
- For individual plots and combined individual plots: NQ values prior to the first quantifiable concentration in that profile will be set to zero (linear plots only); after the first quantifiable concentration of the profile any NQ values will be set to missing.

Additional graphical displays may be generated as appropriate.

4.2.4 Efficacy

Efficacy of selumetinib based on tumour response is a secondary endpoint of the study. Efficacy endpoints (ORR, DoR, PFS, TTR and TTP) based on independent central review and investigator assessments will be analysed separately.

4.2.4.1 Objective Response Rate (ORR)

ORR is defined as the percentage of subjects with complete response (CR) or confirmed partial response (PR, defined as PN decrease ≥ 20% compared to baseline) and will be based on subjects who received at least one dose of selumetinib. SAF subjects with no imaging assessments will be counted as non-responders. The ORR will use all scans regardless of whether they were scheduled or not.

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

The overall response (at DCO) will also be summarised by category (CR, Unconfirmed PR, Confirmed PR, SD, PD or NE).

The best objective response (BOR), defined as the best response recorded from the start of the treatment until progression or the last evaluable MRI assessment in the absence of progression, will also be summarised by category (CR, cPR, PR, SD, PD or NE) and descriptive statistics provided.

4.2.4.2 Duration of Response (DoR)

Kaplan Meier (KM) plots of DoR will be presented. Median DoR and 95% CI (if evaluable) will also be calculated using the KM method and presented in a summary table. The number and percentage of responding subjects remaining in response at 4 months, 8 months, 12

months, 16 months, 20 months, 24 months will also be summarised based on the KM method. The same will be summarised based on every 4 cycles to 24 cycles.

Only subjects who have a complete response or confirmed partial response will be included in this analysis.

Swimmer plots that show the profile of each subject will also be produced. Additional graphical symbols to depict the start of complete response, confirmed partial response, disease progression, ongoing confirmed partial response or stable disease at DCO and treatment discontinuation will be added.

4.2.4.3 Progression Free Survival (PFS)

A KM plot of PFS will be presented. Median PFS and 95% CIs (if evaluable) will also be calculated using the KM method and presented in a summary table. The percentage PFS at 4 months, 8 months, 12 months, 16 months, 20 months, 24 months will be summarised. The same will be summarised based on every 4 cycles to 24 cycles.

4.2.4.4 Time to Progression (TTP)

A KM plot of TTP will be presented. Median TTP and 95% CIs (if evaluable) will also be calculated using the KM method and presented in a summary table. The percentage TTP at 4 months, 8 months, 12 months, 16 months, 20 months, and 24 months will be summarised. The same will be summarised based on every 4 cycles to 24 cycles.

4.2.4.5 Time to Response (TTR)

The TTR will be summarised (i.e. number of subjects (%) based upon the number of responders for each cohort) by the MRI assessment timepoint at which the response was first observed. The timepoints will include 4 months, 8 months, 12 months, 16 months, 20 months, 24 months. The same will be summarised based on every 4 cycles to 24 cycles. Additionally, descriptive summary statistics (i.e. minimum, maximum, median, Q1 and Q3) will also be presented.

4.2.4.6 Change in TL and NTL Tumour Size

Changes in PN growth will be evaluated descriptively by summaries of percentage and absolute change in PN volume from baseline and presented for each time point for both the TL and the NTL.

The average tumour volume with its associated 95% CI will also be displayed graphically across time (post-cycle assessments). The data displays presented by scheduled MRI assessment will include summaries of all scheduled MRI assessments where data were collected until DCO.

The best percentage change from baseline will be summarised descriptively and presented graphically using waterfall plots. The best percentage change will be derived as the maximum reduction from baseline or (in the absence of reduction) the minimum increase from baseline to DCO. Each subject's best percentage change in PN volume will be represented as a separate bar and the bars ordered from the largest increase to the largest decrease.

A spaghetti (also called spider) plot showing individual subject tumour volumes will also be produced to show the volume of PN over time (post-cycle assessments).

Additionally, spaghetti plots will be produced, showing each subject's percentage change in tumour volume as a line over time and progression will be indicated.

Reference lines at the +20% and –20% change in tumour size levels which correspond with the definitions of progression and partial response respectively, will be added to the average tumour volume, waterfall and spider plots presenting percentage change in tumour volume.

Tumour volume (observed and change from baseline) will also be listed including the PN primary lesion site and response assessment. The best response will be flagged on the listing.

4.2.4.7 Sensitivity Analyses

4.2.4.7.1 Discordance Between Investigator and ICR

The discordance between the investigator and the independent central review will be provided for ORR to assess the robustness of the single reviewer. The reviews will be derived using the same method. Cross-tabulation summaries of the tumour response (CR, PR, SD, PD) by the investigator versus the tumour response (CR, PR, SD, PD) from the ICR will be presented.

4.2.5 Disfigurement

For the analysis of disfigurement, anonymised photographs of PN-related morbidity will be provided in the individual subject reviews, where appropriate. This is beyond the scope of this SAP.

4.2.6 Clinical Outcome Assessment Analyses

All functional and clinical outcome assessment analyses will be performed on the safety analysis set unless otherwise noted. Summaries and analyses will be done for the domain/total scores only, and not the individual item scores, unless stated otherwise. The primary analysis of the clinical outcome assessments and functional outcomes will be based on descriptive statistics, including change from baseline, calculated for each scheduled visit in the study. In addition, change from baseline will be analysed using a mixed model repeated measures (MMRM) approach, with baseline score and scheduled visit included in the model as fixed explanatory variables, where there are at least ten patients at a visit. An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger

approximation will be used to estimate the degrees of freedom. 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.

The analyses described below will be provided for each COA endpoint, unless otherwise specified.

A table with descriptive statistics (N, mean, SD, median, minimum, maximum) for the score and change from baseline will be provided.

For each of the COA raw items, a table of responses over time will be presented by scheduled visit. In addition, stacked column charts will be produced to display the percentage in shift from baseline to change in post-baseline scores, displaying summaries categorised by worsened/stable/improved..

Line graphs of mean values over time with 95% confidence intervals and mean actual change from baseline over time with 95% confidence intervals will be presented separately. In addition, stacked column charts of the responses over time will be presented by scheduled visit.

Descriptive statistics for compliance for each questionnaire will include the number of subjects, the number of questionnaires received and completed.

All COA data will be fully listed, including scores and changes from baseline, in terms of domain, and total (item-total) scores.

For analysis of the clinical outcome assessments, the self-reported and parent/guardian reported questionnaires will be analysed and presented separately. All age groups within each domain will also be analysed separately.

The general considerations and handling of missing data outlined in Section 4.2.2.1 should be applied.

4.2.6.1 Pain

4.2.6.1.1 FLACC Pain Scale (3 years)

Pain intensity in 3 year olds will be assessed using the Face, Legs, Activity, Cry, Consolability (FLACC) scale.

At each post-baseline assessment, the absolute change in pain intensity total score from baseline will be calculated as the post-baseline value minus baseline value. Individual FLACC scale scores and changes from baseline will be listed, and a line graph will be plotted with the

mean values and mean values of change from baseline and corresponding 95% CI for each scheduled visit.

4.2.6.1.2 Faces Pain Scale – Revised (4 to 17 years)

For children aged 4 to 17 years, pain intensity will be measured using the Faces pain scale – revised, which uses images to give a score on a scale of 0 (no pain) to 10 (very much pain).

All analyses described in Section 4.2.6 will be performed for the Faces pain scale score.

Change from baseline in the primary outcome score will be further analysed using a MMRM analysis as described in Section 4.2.6 if there are enough data.

4.2.6.1.3 NRS-11 (adult cohort only)

Pain intensity in the adult cohort will be measured by the Numerical Rating Scale-11 (NRS-11), consisting of 3 questions scored on a scale 0=no pain to 10=worst pain you can imagine.

The analyses described in Section 4.2.6 will be performed for each of the three questions of the NRS-11 scale.

For each of the 3 questions, change from baseline in the primary outcome scores will be further analysed using a MMRM in a similar way as described in Section 4.2.6 if there are enough data.

4.2.6.1.4 Pain Interference Index

Pain interference will be measured by the Pain Interference Index (PII). Higher scores indicate more interference with daily activities. The following analyses will be performed for self-reported scores and parent-reported scores separately.

The analyses described in Section 4.2.6 will be performed for the total PII score (calculated as the mean of the completed items).

Change from baseline in the total PII score will be further analysed using a MMRM in a similar way as described in Section 4.2.6 if there are enough data.

4.2.6.1.5 Pain Medication Survey

Information on all analgesics used by subjects for pain control will be collected on the pain medication survey. Pain medication usage will be listed.

4.2.6.2 Health-related Quality of Life

General health-related quality of life (HRQoL) will be measured using the generic Paediatric Quality of Life Inventory (PedsQL™ 4.0 Generic Core Scales) for the paediatric cohort, and EORTC QLQ-C30 and PlexiQoL for the adult cohort.

4.2.6.2.1 PedsQL

At each post-baseline assessment, the absolute change in PedsQL scores from baseline will be calculated as the post-baseline value minus baseline value for each domain and the total score. The following analyses will be performed for self-reported scores and parent-reported scores separately.

For both the primary and secondary outcomes (total scores and domain scores respectively), the analyses described in Section 4.2.6 will be performed. Outcomes will be produced describing both reversed and transformed scores.

Change from baseline in the primary outcome (transformed) scores will be further analysed using a MMRM in a similar way as described in Section 4.2.6, if there are enough data.

4.2.6.2.2 EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 questions which are combined into various scales and an overall global health measure, as described in Section 3.4.2.2.

At each post-baseline assessment, the absolute change in EORTC QLQ-C30 scores from baseline will be calculated as the post-baseline value minus baseline value for each symptom scale, functional domain, 6 single items (dyspnoea, loss of appetite, insomnia, constipation, diarrhoea, and financial difficulties), and the global HRQoL score.

The analyses described in Section 4.2.6 will be performed for each of these scores.

Change from baseline in the primary outcome scores will be further analysed using a MMRM in a similar way as described in Section 4.2.6 if there are enough data.

4.2.6.2.3 PlexiQoL

PlexiQoL assesses quality of life using a range of questions with “true” or “false” as the answer. At each post-baseline assessment, the absolute change in total score (derived from the 18 questions) from baseline will be calculated as the post-baseline value minus baseline value.

The analyses described in Section 4.2.6 will be performed for the total PlexiQoL score.

4.2.6.3 Physical Functioning

At each post-baseline assessment, the absolute change from baseline will be calculated as the post-baseline value minus the baseline value for PROMIS mobility and upper extremity items

and raw domain scores for the paediatric cohort, and for the Physical Function Short Form items and raw domain scores for the adult cohort. In relation to the mobility, upper extremity scores, the following analyses will be performed for self-reported scores and parent-reported scores separately.

For the mobility, upper extremity and Physical Function Short Form raw domain scores , the analyses described in Section 4.2.6 will be performed. For the item scores descriptive statistics will be provided.

Change from baseline in the mobility, upper extremity and Physical Function Short Form raw domain scores will be further analysed using a MMRM in a similar way as described in Section 4.2.6, where there are enough data.

Analyses based on raw scores will be repeated for T-scores when appropriate.

4.2.6.4 Patient's Global Impression of Symptom Severity

To evaluate the severity of a subject's symptoms, the PGIS instrument will be used to give a score on a 6-point scale (0 = no symptoms, 1 = very mild, 2 = mild, 3 = moderate, 4 = severe, 5 = very severe) for each of tumour pain, overall pain, and tumour-related problems.

The subject self-report and parent report ratings for each of the three items will be reported separately.

The following will be provided for each of the 3 questions:

- Descriptive statistics (counts and percentages) for each response option at each scheduled visit; a stacked column chart showing the distribution of shift from baseline to change in post-baseline responses at each scheduled visit will also be provided.

4.2.6.5 Patient's Global Impression of Change

To evaluate the subject's perspective of change, an adapted version of the PGIC instrument will be used to assess tumour pain, overall pain, and tumour-related problems, compared to baseline. It is scored on a 7-point scale (1 = Very much improved to 7 = Very much worse) and is both self-assessed by subjects and also assessed by parents.

The subject self-report and parent report ratings for each of the three items will be reported separately.

The following will be provided for each of the 3 questions:

- Descriptive statistics (counts and percentages) for each response option at each scheduled visit.

4.2.7 Data Cut-offs

Primary DCO will take place when the last subject dosed has had the chance to complete their visit at the end of cycle 10. Final DCO will occur when the last subject dosed has had the chance to complete their visit at the end of cycle 24. In addition, an interim analysis will be conducted when last subject dosed has had the chance to complete their visit at the end of cycle 4.

4.2.8 Prior, Concomitant and Other Treatments

All prior and concomitant medications, as defined in section 3.5.2, will be listed along with the reason for use, route of administration, dates and dosage. Pain medications recorded at enrolment and during the study will be summarised by the coded terms.

5. INTERIM ANALYSES

An interim analysis will be conducted when the last subject dosed has had the chance to complete their visit at the end of cycle 4 to assess PK, safety, tolerability and efficacy of selumetinib.

Additionally, a Safety Review Committee (SRC) will review the safety, tolerability and PK of selumetinib. They will meet initially after the first 6 subjects of each cohort have been treated for approximately 3 cycles. There will be further SRC meetings if necessary during the study.

6. CHANGES OF ANALYSIS FROM PROTOCOL

There are currently no changes of analysis from the protocol.

7. REFERENCES

Aaronson et al. (1993). Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*, 85(5):365-76.

Basch et al. (2014). Basch E, Trentacosti AM, Burke LB, Kwitkowski V, Kane RC, Autio KA, et al. Pain palliation measurement in cancer clinical trials: the US Food and Drug Administration perspective. *Cancer*, 120(5):761-7.

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

Farrar et al. (2000). Farrar, J.T., et al. Defining the clinically important difference in pain outcome measures. *Pain 88(3)*, 287-94.

Fayers et al. (2001). Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organization for Research and Treatment of Cancer.

Hawker et al. (2011). Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale. *Arthritis Care Res (Hoboken)*, 63(Suppl 11):S240-52.

Heaney et al. (2019). Heaney A, Wilburn J, Langmead S, Blakeley J, Huson S, Jim C, et al. A qualitative study of the impact of plexiform neurofibromas on need fulfilment in adults with neurofibromatosis type 1. *SAGE Open Med*, 7:2050312119829680.

Hicks et al. (2001). Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale - Revised: toward a common metric in pediatric pain measurement. *Pain*, 93(2):173-83.

Kemani et al. (2016). Kemani MK, Zetterqvist V, Kanstrup M, Holmström L, Wicksell RK. A validation of the pain interference index in adults with long-standing pain. *Acta Anaesthesiol Scand*, 60(2):250-8.

Kendrick and Strout. (2005). Kendrick, D.B. and T.D. Strout. The minimum clinically significant difference in patient-assigned numeric scores for pain. *Am J Emerg Med 23(7)*, 828-32.

Martin et al. (2015). Martin S, Nelson Schmitt S, Wolters PL, Abel B, Toledo-Tamula MA, Baldwin A, et al. Development and validation of the English Pain Interference Index and Pain Interference Index - Parent Report. *Pain Med*, 16(2):367-73.

Merkel et al. (1997). Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. *Pediatr Nurs*, 23(3):293-7.

Nutakki et al. (2018). Nutakki K, Varni JW, Swigonski NL. PedsQL neurofibromatosis type 1 module for children, adolescents and young adults: feasibility, reliability, and validity. *J Neurooncol*, 137(2):337–47.

PROMIS website. (2016). *Parent Proxy Mobility – Short Form 8a.* Retrieved from PROMIS website: http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20Parent%20Proxy%20SF%20v2.0%20-%20Mobility%208a%207-29-2016.pdf

PROMIS website. (2016). *Parent Proxy Upper Extremity – Short Form 8a.* Retrieved from PROMIS website: http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20Parent%20Proxy%20SF%20v2.0%20-%20Upper%20Extremity%208a%207-29-2016.pdf

PROMIS website. (2016). *Pediatric Mobility – Short Form 8a.* Retrieved from PROMIS website: http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20Pediatric%20SF%20v2.0%20-%20Mobility%208a%207-28-2016.pdf.

PROMIS website. (2016). *Pediatric Upper Extremity – Short Form 8a.* Retrieved from PROMIS website: http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20Pediatric%20SF%20v2.0%20-%20Upper%20Extremity%208a%207-28-2016.pdf

Salaffi et al. (2004). Salaffi, F., et al. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain 8(4)*, 283-91.

Varni et al. (2001). Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care*, 39(8):800-12.

Voepel-Lewis et al. (2011). Do 0–10 Numeric Rating Scores Translate into Clinically Meaningful Pain Measures for Children? *Anesth Analg. 112(2)*, 415-21.

Wicksell et al. (2009). Wicksell RK, Melin L, Lekander M, Olsson GL. Evaluating the effectiveness of exposure and acceptance strategies to improve functioning and quality of life in longstanding pediatric pain - a randomized controlled trial. *Pain*, 141(3):248-257.