Protocol B7461024

A PHASE 2, MULTI-CENTER, OPEN-LABEL, DUAL-COHORT STUDY TO EVALUATE THE EFFICACY AND SAFETY OF LORLATINIB(PF-06463922) MONOTHERAPY IN ALK INHIBITOR-TREATED LOCALLY ADVANCED OR METASTATIC ALK-POSITIVE NON-SMALL CELL LUNG CANCER PATIENTS IN CHINA

Statistical Analysis Plan
(SAP)

Version: 2.0

Date: 11-Sep-2020

TABLE OF CONTENTS

LIST OF TABLES 4
LIST OF FIGURES 5
1. VERSION HISTORY 6
2. INTRODUCTION 7
2.1. Study Objectives 8
2.2. Study Design 8
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 9
3.1. Primary Endpoint(s) 9
3.2. Secondary Endpoint(s) 9
CCI [redacted]
3.4. Baseline Variables 10
3.5. Safety Endpoints 11
3.5.1. Adverse Events 11
3.5.2. Adverse Events of Special Interest (AESI) 11
3.5.3. Laboratory Data 11
4. ANALYSIS SETS 12
4.1. Full Analysis Set 12
4.2. Per Protocol (PP) Analysis Set 12
4.3. Safety Analysis(SA) Set 12
4.4. Other Analysis Sets 12
4.4.1. Pharmacokinetics Analysis Set 12
CCI [redacted]
5. GENERAL METHODOLOGY AND CONVENTIONS 12
5.1. Hypotheses and Decision Rules 12
5.1.1. Hypotheses and Sample Size Section 12
5.1.2. Decision Rules 13
5.2. General Methods 13
5.2.1. Data Handling after the Cut-off Date 13
5.2.2. Pooling of Data by Center 13
5.2.3. Definition of Study Day 14
5.2.4. Definition of Cycle and Cycle Day 14

5.2.5. Definition of Start of New Anti-cancer Drug Therapy 14
5.2.6. Definition of Start of New Anti-cancer Therapy 15
5.2.7. Date of Last Contact 15
5.2.8. Measurable Disease 16
5.2.9. Tumor Assessment Date 16
5.2.10. Sum of Lesion Diameters 17
5.2.11. Adequate Baseline 17
5.2.12. Adequate Post Baseline Tumor Assessment 17
5.2.13. Unscheduled Assessments 17
5.2.14. Standard Derivations and Reporting Conventions 17
5.2.15. Analyses for Continuous Data 18
5.2.16. Analyses for Categorical Data 18
5.2.17. Analyses for Time to Event Data 18
5.3. Methods to Manage Missing Data 18
5.3.1. Missing Dates 18
5.3.2. Missing Toxicity Grade of Adverse Events 21
5.3.3. Missing Pharmacokinetic (PK) data 22
5.3.4. Missing ECG Data 23
CCI
6. ANALYSES AND SUMMARIES 23
6.1. Primary Endpoint(s): Objective Response in Cohort 1 23
6.1.1. Sensitivity/Robustness Analyses (if needed) 25
6.2. Secondary Endpoint(s) 25
6.2.1. OR in Cohort 2 25
6.2.2. Intracranial Objective Response (IC-OR) in Both Cohorts 25
6.2.3. Duration of Response (DoR) in Both Cohorts 26
6.2.4. Duration of Intracranial Response (IC-DoR) in Both Cohorts 26
6.2.5. Time to Tumor Response (TTR) in Both Cohorts 26
6.2.6. Progression-free Survival (PFS) in Both Cohorts 26
CCI
6.2.7. Overall Survival (OS) in Both Cohorts 30
6.2.8. Pharmacokinetics 31
6.3. Other Endpoint(s) 32

CCI
6.4. Baseline and Other Summaries 33
6.4.1. Baseline Summaries 33
6.4.2. Study Conduct and Subject Disposition 34
6.4.2.1. Patient Disposition 34
6.4.2.2. Protocol Deviations 35
6.4.3. Study Treatment Exposure 35
6.4.3.1. Exposure to lorlatinib 35
6.4.3.2. Dose Reductions, Interruptions, and Delays 36
6.4.4. Concomitant Medications and Non-Drug Treatments 36
6.4.5. Subsequent Anti-Cancer Therapies/Procedures 37
6.5. Safety Summaries and Analyses 37
6.5.1. Adverse Events 37
6.5.1.1. Adverse Events of Special Interest 39
6.5.2. Deaths 39
6.5.3. Laboratory Data 39
6.5.4. Vital Signs 42
6.5.5. Electrocardiogram 43
6.5.6. Physical Examination 45
6.5.7. Left Ventricular Ejection Fraction (LVEF) 45
6.5.8. Performance Status 45
6.5.9. Mood and Suicidal Ideation and Behavior Analyses 45
7. INTERIM ANALYSES 46
8. CONSIDERATION ON COVID-19 IMPACT 46
9. REFERENCES 47

LIST OF TABLES

Table 1. Summary of Major Changes in SAP Amendments 6
Table 2. Outcome and event dates for PFS analyses 28
Table 3. Censoring Reasons and Hierarchy 29
CCI

LIST OF FIGURES

Figure 1. Study B7461024 Design 9

APPENDICES

Appendix 1. Derived Objective Response per mRECIST v1.1 based on ICR Data 48

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study B7461024 is based on the protocol amendment dated 11 October 2019.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Date	Change	Rationale
1.0	6-Jun-2018	Original SAP	Not Applicable
2.0	11-Sep-2020	The PCD time has been modified from 12 months after last subject enrolled in Cohort1 to 8 months after last subject first dose in study B7461024. Updated the definition of on-treatment and adding new-anti cancer therapy in Section 3.5. Updated the definition of treatment emergent adverse events(TEAE) in Section 3.5.1. Updated Per Protocol Analysis Set in Section 4.2. Updated Cohort 2 in Section 5.1.1. Added definition of start of new anti-cancer drug therapy and definition of start of new anti-cancer therapy in Section 5.2.5 and Section 5.2.6. Updated adequate baseline in Section 5.2.11 (corresponding to Section 5.2.9 in the 1st version SAP). Updated Sensitivity/Robustness analysis in Section 6.1.1. Removed the unnecessary analysis in Section 6.2.2. and Section 6.2.4.	The rationale for this change is per protocol amendment and study requirement

		Updated censoring method for anti-cancer therapies in Section 6.2.6. Updated the Pharmacokinetics analysis in Section 6.2.8. Removed the summary of relative dose(%) in Section 6.4.3.1 and adding descriptions of dose interruption in the case of multiple dose interruptions Section 6.4.3.2. Updated Adverse Events summary in Section 6.5.1. Removed unnecessary wordings and analysis in Section 6.5.3. Removed the unnecessary analysis for Vital Signs in Section 6.5.4. Corrected Fridericia's correction (QTcF) formula and removed unnecessary analysis in Section 6.5.5. Modified the summary of LVEF% in Section 6.5.7. Remove the unnecessary analysis in Section 6.5.9. Added consideration on covid-19 impact in Section 8. Added sensitivity analysis for primary endpoint ORR by independent central radiology(ICR) review in Appendix 1.	

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B7461024. This analysis plan is meant to supplement the study protocol. Any deviations from this analysis plan will be described in the Clinical Study Report.

The primary analysis will include all data up to a data cutoff date which will be determined by the data maturity and is currently estimated to be around 8 months after last subject first dose in study B7461024. All summaries and analyses will include all data pertaining to visits/assessments performed up to and including the data cutoff date.

2.1. Study Objectives

Primary Objective(s):

- To evaluate the anti-tumor effect of lorlatinib as a single agent as measured by objective response rate (ORR) per RECIST v1.1 in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after crizotinib treatment.

Secondary Objective(s):

- To evaluate the anti-tumor effect of lorlatinib as a single agent as measured by ORR per RECIST v1.1 in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after ALK inhibitor treatment other than crizotinib.
- To evaluate the progression-free survival (PFS) and overall survival (OS) in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after crizotinib and other ALK inhibitor treatment, respectively.
- To evaluate other antitumor activities in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after crizotinib and other ALK inhibitor treatment, respectively.
- To evaluate the safety and tolerability of lorlatinib treatment in locally advanced or metastatic ALK+ NSCLC patients whose disease has progressed after crizotinib and other ALK inhibitor treatment.
- To evaluate the pharmacokinetics of lorlatinib and potential pharmacokinetic/pharmacodynamics relationship for lorlatinib if appropriate.

CCI

[redacted]

[redacted]

2.2. Study Design

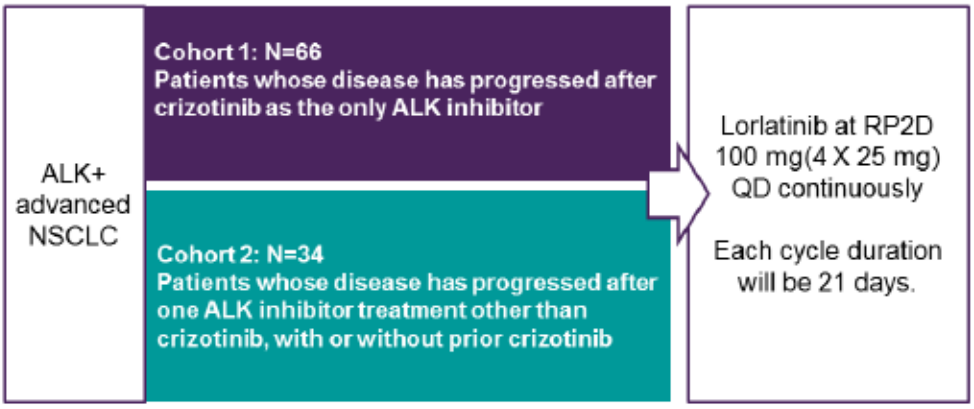
This is an open-label, multi-center, Phase 2, dual-cohort study of evaluating the efficacy and safety of PF-06463922 as a monotherapy in patients with ALK-positive locally advanced or metastatic NSCLC.

Approximately 100 patients with locally advanced or metastasis ALK-positive NSCLC will be enrolled to receive lorlatinib monotherapy. In Cohort 1 approximately 66 patients whose disease has progressed after crizotinib will be enrolled, and in Cohort 2, approximately 34 patients whose disease has progressed after ALK inhibitor other than crizotinib.

A cycle duration will be 3 weeks (**21 days**) and will always be considered 3 weeks irrespective of any dose delays/dosing interruptions or missed doses which may affect nominal days of each cycle.

The study design is illustrated in Figure 1 below.

Figure 1. Study B7461024 Design



Study treatment may continue until confirmed disease progression assessed by ICR, subject refusal, subject lost to follow-up, unacceptable toxicity, or the study is terminated by the sponsor, whichever comes first. Subjects who develop radiological disease progression confirmed by ICR assessment but are otherwise continuing to derive clinical benefit from study treatment will be eligible to continue with the treatment they have been assigned to, provided that the treating physician has determined that the benefit/risk for doing so is favorable.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

- Objective Response (OR) in patients whose disease have progressed after crizotinib (Cohort 1) by independent central radiology (ICR) assessment per RECIST v1.1. OR is defined as a complete response (CR) or partial response (PR) recorded from enrollment until disease progression or start of new anti-cancer therapy. Both CR and PR must be confirmed by repeat assessments performed no less than 4 weeks after the criteria for response are first met.

3.2. Secondary Endpoint(s)

- OR in patients whose disease have progress after ALK inhibitor treatment other than crizotinib (Cohort 2) as assessed by RECIST v1.1 per ICR assessment.

- Progression-free survival (PFS) based on ICR assessment and investigator assessment by RECIST v.1.1 is defined as the time from first dose to first documentation of objective disease progression or to death due to any cause, whichever comes first in both Cohort 1 and Cohort 2.

- Overall Survival (OS) is defined as the time from first dose to the date of death due to any cause.

- Intracranial objective response(IC-OR)

  IC-OR is defined the same as OR, but limited to Intra Cranial lesions only on patients with CNS metastases (ie, Best Overall Intracranial Response as confirmed CR or confirmed PR considering only the lesions having disease site as Brain).

- Duration of response (DoR)

  DoR is defined as the time from the first documentation of CR or PR to the first documentation of disease progression or death due to any cause, whichever occurs first.

- Duration of intracranial response(IC-DoR)

  IC-DoR is defined the same as DoR, but limited to Intra Cranial lesions only on patients with CNS metastases (ie, Best Overall Intracranial Response as confirmed CR or confirmed PR considering only the lesions having disease site as Brain).

- Time to tumor response (TTR)

  TTR is defined as the time from first dose to first documentation of objective tumor response (CR or PR).

- PK: parameters on day 1 of cycle 1 and at steady state: C_{max}, T_{max}, AUC_t, AUC_{tau} at steady state, AUC_{inf}, CL/F, Vz/F, $t_{1/2}$ and R_{ac} as data permit.

CCI

[redacted]

3.4. Baseline Variables

Any deviations from the protocol specified window will be documented as protocol deviations. A separate definition of adequate baseline will be provided for tumor assessment related efficacy endpoints.

For efficacy analyses, safety (including Eastern Cooperative Oncology Group (ECOG) performance status) and baseline characteristics associated with tumor assessments (eg, number of sites of disease at baseline), the last assessment prior to first dose will serve as the baseline assessment.

Triplicate ECGs are collected; therefore the baseline for each ECG measurement is the average of the pre-dose measurements on the baseline day. Unscheduled assessments will not be included in the calculation of the average. Most of the ECG parameters will not be derived as they are provided by sites on the CRF. QTcF (Fridericia's correction) and QTcB (Bazett's correction) will be derived based on RR and QT. The average of the replicate measurements will be determined after the derivation of the individual parameters at each timepoint.

3.5. Safety Endpoints

Safety endpoints will be summarized based on the on-treatment period unless otherwise specified.

On-treatment is defined as the time from the first dose of study treatment through the end of study follow-up (infinite lag) or start of new anti-cancer drug therapy, whichever occurs first. Adverse events occurring on the same day as the first dose of study treatment will be considered to have occurred during the on-treatment period. All other assessments which occur on the same day as the first dose of study treatment will be considered baseline assessments (see Section 3.4 for the definition of baseline).

Safety data collected outside the on-treatment period as described above will be listed but not summarized.

3.5.1. Adverse Events

An adverse event is considered treatment emergent (TEAE) if the event occurs during the on-treatment period.

3.5.2. Adverse Events of Special Interest (AESI)

AESIs include events such as CNS (mood, speech, cognition) events. These events will be defined based on a list of MedDRA Preferred Terms specified in the Safety Review Plan for lorlatinib. A final list will be provided to programming prior to database release.

3.5.3. Laboratory Data

Hematology, chemistry, lipids , coagulation and urinalysis result will be programmatically graded according to the NCI CTCAE version 4.03 for relevant parameters. A shift summary of baseline grade by maximum post-baseline grade will be presented. Parameters which cannot be graded will be summarized relative to the normal range (ie, normal range high or normal range low). Additional details are provided in Section 6.5.3.

4. ANALYSIS SETS

Data for all subjects will be assessed to determine if subjects meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per standard operating procedures.

4.1. Full Analysis Set

The full analysis (FA) set will include all patients who are enrolled, regardless of whether or not treatment was received.

4.2. Per Protocol (PP) Analysis Set

The per protocol analysis set will include all enrolled subjects with ALK positive NSCLC, received a prior ALK inhibitor as protocol required, and who received at least 1 dose of lorlatinib:

- In Cohort 1, received crizotinib as the only ALK inhibitor.
- In Cohort 2, received one ALK inhibitor other than crizotinib, with or without prior crizotinib treatment.

4.3. Safety Analysis(SA) Set

The safety analysis set includes all enrolled patients who receive at least one dose of lorlatinib.

4.4. Other Analysis Sets

4.4.1. Pharmacokinetics Analysis Set

The PK parameter analysis population is defined as all enrolled patients who receive at least one dose of study medication and have sufficient information to estimate at least 1 of the PK parameters of interest.

CCI

[redacted]

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

5.1.1. Hypotheses and Sample Size Section

Cohort 1:

The primary objective of the study is to evaluate the anti-tumor effect of lorlatinib as a single agent as measured by ORR per RECIST v1.1 in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after crizotinib treatment, ie, Cohort 1.

The study is designed to test the null hypothesis H0: ORR<=30% vs. HA: ORR>50%. Assuming at least a 20% increase of ORR with lorlatinib treatment, 66 patients will be required to provide 90% power to reject the null hypothesis at a significance level of 0.025 (one-sided).

Cohort 2:

No formal statistical hypothesis testing is planned for Cohort 2. This cohort will explore the safety and efficacy of lorlatinib after one ALK inhitibor treatment other than crizotinib with 34 patients planned to be enrolled.

The sample size of 34 patients will provide the estimated ORR with a maximum width of the 95% CI of 35%, observed with 17 responses out of 34 patients. In order to collect more information of anti-tumor activity, if the patients received study treatment but without adequate baseline tumor assessment or without post baseline tumor assessments, additional patients may be allowed to be enrolled.

With approximately 66 and 34 patients in Cohort 1 and 2 respectively, this study will provide adequate sample size for safety evaluation with a minimal number of 100.

5.1.2. Decision Rules

Cohort 1 will test the null hypothesis H0: ORR<=30%.

If at least 28 responses are observed among the 66 patients enrolled, then the null hypothesis will be rejected, and it will be concluded that the study has demonstrated that the true ORR exceeds 30%.

However, if the actual number of patients is not equal to 66 at the time of final analysis of ORR in Cohort 1, the testing will depend on the exact number of patients enrolled. If the p-value of the test <0.025, then the null hypothesis will be rejected.

5.2. General Methods

Unless otherwise specified, baseline data will be summarized by Cohort. Disposition, efficacy, exposure (including concomitant therapies), and safety data will also be summarized by Cohort and Total.

5.2.1. Data Handling after the Cut-off Date

Data after the cut-off date may not undergo the cleaning process and will not be displayed in any listings or used for summary statistics, statistical analyses or imputations.

5.2.2. Pooling of Data by Center

In order to provide overall estimates of treatment effects, data will be pooled across centers. The 'center' factor will not be considered in statistical models or for subset analyses due to the high number of participating centers in contrast to the anticipated small number of patients treated at each center.

5.2.3. Definition of Study Day

Start day of study treatment is the day of the first dose of study treatment.

The study day for assessments occurring on or after the first dose of study treatment (eg, adverse event onset, laboratory date, tumor measurement) will be calculated as:

Study day = Date of the assessment/event - start date of study treatment + 1.

The study day for assessments occurring prior to the first dose of study treatment (eg, baseline characteristics, medical history) will be negative and calculated as:

Study day = Date of the assessment/event –start date of study treatment.

The study day will be displayed in all relevant data listings.

5.2.4. Definition of Cycle and Cycle Day

Cycle start and end dates are derived per patient.

Lorlatinib is administered every 3 weeks in a cycle; therefore the nominal cycle length is 21 days.

For Cycle X, the actual cycle start date for each subject is the earliest start date of dosing (dose>0 at that visit) in Cycle X visit CRF exposure page.

- For all but the last cycle;
- Actual cycle stop date is calculated as the start date of the next cycle minus one day.
- Actual cycle duration is calculated from Day 1 of a cycle to the day prior to Day 1 of the next cycle, as follows:

  Actual Cycle Duration (weeks) = (cycle stop date – cycle start date + 1) / 7

- For the last cycle, actual cycle duration is calculated as follows:

  Actual Cycle Duration (weeks) = (last date of study treatment– cycle start date + 1) / 7

The cycle day will be calculated as:

Cycle day = Date of the assessment/event – cycle start date + 1.

5.2.5. Definition of Start of New Anti-cancer Drug Therapy

Start date of new anti-cancer drug therapy is used to determine the end of the on-treatment period (see Section 3.5).

The start date of new anti-cancer drug therapy is the earliest start date of anti-cancer drug therapy recorded in the 'Follow-up Cancer Therapy' eCRF pages that is after the first dose of study treatment. When start date of anti-cancer drug therapy is missing or partially missing, the imputation rules described in Section 5.3.4 should be applied using only data from the 'Follow-up Cancer Therapy' eCRF pages.

5.2.6. Definition of Start of New Anti-cancer Therapy

Start date of new anti-cancer therapy (drug, radiation, surgery) is used for censoring in efficacy analyses (see Section 6.2.6).

The start date of new anti-cancer therapy is the earliest date after first dose amongst the following:

- Start date of anti-cancer drug therapy recorded in the 'Follow-up Cancer Therapy' eCRF pages.
- Start date of radiation therapy recorded in 'Concomitant Radiation Therapy', and 'Follow-up Radiation Therapy' eCRF pages with 'Treatment Intent' = 'Curative in intent'.
- Surgery date recorded in 'Concomitant Surgery', and 'Follow-up Surgery' eCRF pages when 'Surgery Outcome' = 'Resected' or 'Partially Resected'.

When start date of anti-cancer therapy is missing or partially missing, the imputation rules described in Section 5.3.1 should be applied using 'Follow-up Cancer Therapy', 'Concomitant Radiation Therapy', 'Follow-up Radiation Therapy', 'Concomitant Surgery', and 'Follow-up Surgery' eCRF pages.

5.2.7. Date of Last Contact

The date of last contact will be derived for patients not known to have died at the data cutoff date using the latest complete date (ie, imputed dates will not be used in the derivation) among the following:

- All patient assessment dates (eg, blood draws (laboratory, Pharmacokinetics (PK)), vital signs, physical exam, performance status, ECG, Echocardiograms (ECHO)/multigated acquisition (MUGA) scans, tumor assessments, Mood and Suicidal Ideation assessment);
- Start and stop dates of concomitant therapies including non-drug treatments or procedures;
- Start and end dates of anti-cancer therapies administered after study treatment discontinuation including systemic therapy, radiation, and surgeries;
- AE start and end dates;

- Last date of contact collected on the 'Survival Follow-up' CRF (do not use date of survival follow-up assessment unless status is 'alive');
- Study treatment start and end dates;
- Date of discontinuation on disposition CRF pages (do not use if reason for discontinuation is lost to follow-up or death).

Only dates associated with actual examinations of the patient will be used in the derivation. Dates associated with a technical operation unrelated to patient status such as the date a blood sample was processed or dates when data were entered into the CRF will not be used. Assessment dates after the data cutoff date will not be applied to derive the last contact date.

5.2.8. Measurable Disease

A patient will be considered to have measureable disease if there is at least one target lesion identified at baseline meeting the following criteria:

- Non lymph node lesions with longest diameter ≥10 mm by CT scan ;
- Non lymph node lesions with longest diameter ≥10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as non-measurable);
- CNS lesions with longest diameter ≥5 mm provided by gadolinium contrast enhanced MRI;
- Lymph nodes with short axis ≥15 mm when assessed by CT.

5.2.9. Tumor Assessment Date

Tumor assessment dates will be assigned differently for ICR assessment and for investigator's assessment of tumor data.

For analyses of ICR assessments, the date of tumor assessment will be determine based on the radiologist assessments at each nominal timepoint and will be provided on listings as the associated date of response/progression at nominal timepoints. These dates will also be used for censoring in time to event analyses; however, the date of progression and response as provided by the independent radiologist will be used in time to event endpoint analyses as the independent radiologist provided an overall assessment of best response and date of progression based on a review of each nominal timepoint and additional clinical information available for the determination of the date of response/progression.

For analyses based on investigator's assessment when response/progression are derived programmatically from the target lesions measurements, non-target lesions status, and new lesions recorded on the CRF, the date of tumor assessment should be derived as the earliest scan/assessment date. Since tumor assessments are captured on a log CRF page, a clustering algorithm for grouping scans will be used:

Each cluster represents an actual visit. For each patient, the number of clusters is equal to the maximum number of assessments available among all target and non-target lesions. SAS procedure, Proc Fastclus, is applied to a variable that represents the days from the date of first dose to the date of the scan for each target and non-target lesion (date of scan – date of first dose +1). Then the assessments of target and non-target lesions that occurred close to each other in time will be assigned to the same cluster.

5.2.10. Sum of Lesion Diameters

For lesions that are assessed as 'too small to measure', 5 mm will be imputed and used in the calculation of the sum of the lesion diameters.

5.2.11. Adequate Baseline

Adequate baseline for tumor data is defined using the following criteria:

- All baseline assessments must be within 28 days prior to date of first dose.
- All documented lesions must have non-missing assessments (ie, non missing measurements for target lesions and non missing lesions status at baseline for non-target lesions).

5.2.12. Adequate Post Baseline Tumor Assessment

An adequate assessment is defined as an assessment where a response of CR, PR, Stable Disease (SD), non-CR/non-PD, or PD has been provided by the investigator or ICR for the analyses by investigator or ICR respectively. Timepoints where the response is not evaluable or no assessment was performed will not be used for determining the censoring date.

5.2.13. Unscheduled Assessments

Unless otherwise specified, unscheduled assessments will not be displayed in summary tables by nominal visit/timepoint. Unscheduled assessments will be used when deriving baseline and worst case on-treatment for safety analyses. Additionally, unscheduled tumor assessments will be used for efficacy analyses (eg, defining date of progression/censoring, best overall response, date of last contact).

5.2.14. Standard Derivations and Reporting Conventions

The following conversion factors will be used to convert days into weeks, months or years: 1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

Percentages will be reported to one decimal place. The rounding will be performed to closest integer/first decimal using the common mid-point between the two consecutive values. Eg, 5.1 to 5.4 will be rounded to an integer of 5, and 5.5 to 5.9 will be rounded to an integer of 6.

5.2.15. Analyses for Continuous Data

Continuous variables will be summarized using descriptive statistics ie, number of non-missing values and number of missing values [ie, n (missing)], mean, median, standard deviation (SD), minimum, maximum and first and third quartile (Q1 and Q3). CCI [redacted] PK summaries will also include coefficient of variation percent (%CV).

In case the analysis refers only to certain visits, percentages will be based on the number of patients with an assessment at that visit, unless otherwise specified.

5.2.16. Analyses for Categorical Data

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise specified, the calculation of proportions will include the missing category. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

In case the analysis refers only to certain visits, percentages will be based on the number of patients with an assessment at that visit, unless otherwise specified.

5.2.17. Analyses for Time to Event Data

Time to event endpoints will be summarized using the Kaplan-Meier method and estimated survival curves will be displayed graphically when appropriate. Graphs will describe the number of patients at risk over time. The median, first and third quartiles, and probabilities of an event at particular points in time will be estimated by the Kaplan-Meier method. Confidence intervals for medians and quartiles are based on the Brookmeyer-Crowley[1] method. Confidence intervals for the estimated probability of event at a particular timepoint will be generated using the log(-log) method with back transformation to a confidence interval on the untransformed scale. Summaries of the number and percentages of patients with an event will also be provided on summary tables and/or figures.

5.3. Methods to Manage Missing Data

Unless otherwise specified, all data will be evaluated as observed, and no imputation method for missing values will be used.

Any imputations will occur at the analysis dataset level. Additionally, in all patient data listings imputed values will be presented and flagged as imputed.

Missing statistics, eg, when they cannot be calculated, should be presented as 'ND' for not done, 'NR' for not reached or 'NA' for not applicable. For example, if N=1, the measure of variability cannot be computed and should be presented as 'ND' or 'NA'.

5.3.1. Missing Dates

For purposes of data listings, dates will reflect only the information provided by the investigator on the CRF.

If start dates for adverse events or concomitant medications are completely missing a worst case approach will be taken whereby the events will be considered treatment emergent and the medications will be considered concomitant. If only partial information are available (eg, only a month and year or only a year) and the partial information provide sufficient information to indicate the dates are prior to the start of study treatment (eg, month/year less than month/year of first dose) then these will be considered to have started prior to treatment; otherwise a similar worst case approach will apply and these will be considered to have started after treatment.

Date of Last Dose of Study Treatment

No imputation will be done for first dose date. Date of last dose of study treatment, if unknown or partially unknown, will be imputed as follows:

- If the last date of study treatment is completely missing and there is no End of Treatment (EOT) CRF page and no death date, the patient should be considered to be ongoing and use the data cutoff date for the analysis as the last dosing date; or

- If the last date of study treatment is completely or partially missing and there is EITHER an EOT CRF page OR a death date available (on or prior to the data cutoff date), then impute this date as the last dose date:

  = 31DECYYYY, if only Year is available and Year < Year of min (EOT date, death date),

  = Last day of the month, if both Year and Month are available and Year = Year of min (EOT date, death date) and Month < the month of min (EOT date, death date), or

  = min (EOT date, death date), for all other cases.

Missing or Partial Death Dates

Missing or partial death dates will be imputed based on the last contact date:

- If the entire date is missing it will be imputed as the day after the date of last contact (see derivation of date of last contact in Section 5.2.7); or

If the day or month is missing, death will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the following:

Missing day: 1st day of the month and year of death, or

Missing day and month: January 1st of the year of death.

Date of Start of New Anti-cancer Therapy

Incomplete dates for new anti-cancer therapy will be imputed as follows and will be used to determine censoring dates for efficacy analyses:

- The end date of new anti-cancer therapy will be included in the imputation for start date of new anti-cancer therapy. If the end data of new anti-cancer therapy is:
  - Completely missing then it will be ignored in the imputations below,
  - Partially missing with only year available then the imputations below will consider 31DECYYYY as the end date of the new anti-cancer therapy, or
  - Partially missing with only month and year available then the imputations below will consider the last day of the month for MMMYYYY as the end date of the new anti-cancer therapy.
- For patients who have not discontinued study treatment at the time of the data cutoff date, last dose of study treatment is set to the data cutoff date in the imputations below.
- If the start date of new anti-cancer therapy is completely or partially missing then the imputed start date of new anti-cancer therapy is:

  = 31DECYYYY, if only Year is available and Year < Year of min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

  = Last day of the month, if both Year and Month are available and

  Year = Year of min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

  Month < Month of min [max (PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]

  = min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy], for all other cases.

AE Onset Date:

The following imputation rules apply if the event is unique for a patient or it is the first of a series of similar events; otherwise, the AE Onset Date will not be imputed:

- If the AE Collection Date is not missing, is less than the Date of First Exposure to Treatment, and is less than the AE Stop Date, then AE Onset Date is set to the Date of AE Collection.
- If the Previous Visit Date is greater than the Date of First Exposure to Treatment and less than the AE Stop Date, the AE Start Date is set to the previous visit date.

- If the Date of First Exposure to Treatment is greater than the previous visit date and less than the AE Stop Date, the AE Onset Date is set to the Date of First Exposure to Treatment.

- Otherwise AE Onset date is set to the AE Stop date.

AE Stop Date:

Ongoing events will have the AE Stop Date set to one of the following values:

- Date of Death, if the patient died and a date of death exists.

- Maximum of (Patient Withdraw date, AE Onset Date, AE Collection Date) if the patient withdrew from the study and a date of withdraw exists.

- Maximum of (AE Onset Date, Subject Summary Collection Date, AE Collection Date) if the Disposition CRF page exists but a date of withdraw does not exists.

- Maximum of (Last Treatment Date, AE Onset Date) if no Disposition page exits.

Imputation will only occur if event is unique for the patient, or it is the last of a series of similar events; otherwise the Stop Date will not be imputed. Adverse Events are deemed similar if they have the same verbatim term.

Resolved events will have the AE Stop Date set to the maximum of the AE collection date and the AE Onset date.

Other Missing or Partial Dates

Imputation methods for other partial dates as follows:

- If the day of the month is missing for a start date used in a calculation, the first day of the month will be used to replace the missing date.

- If both the day and month are missing, the first day of the year is used.

- For stop dates, the last day of the month, or last day of the year is used if the day or day and month are missing, respectively.

- If the date is completely missing, no imputation will be performed.

5.3.2. Missing Toxicity Grade of Adverse Events

Prior to Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event prior to the first study treatment, then Grade 1 will be assumed for purposes of defining a baseline grade for assessing if further occurrences are treatment emergent. However, if the patient experiences multiple episodes of the same AE prior to

study treatment, then the maximum toxicity grade observed prior to study treatment will be utilized in assessing if further occurrences are treatment emergent.

During Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event during the study treatment, then the event will be considered treatment emergent unless a baseline event was reported as Grade 4.

In summaries which present maximum toxicity grade, the maximum of non-missing grades will be displayed. Missing grade will only be displayed for cases where a patient reported only one event and the grade is missing.

5.3.3. Missing Pharmacokinetic (PK) data

Concentrations below the limit of quantification

For all calculations, figures, and estimation of individual pharmacokinetic parameters, all concentrations assayed as below the level of quantification (BLQ) will be set to zero. In log-linear plots these values will not be represented. The BLQ values will be excluded from calculations of geometric means and their confidence intervals. A statement similar to 'All values reported as BLQ have been replaced with zero' should be included as a footnote to the appropriate tables and figures. In listings BLQ values will be reported as below limit of quantification ("<LLOQ"), where LLOQ will be replaced with the corresponding value from the analytical assay used.

Deviations, Missing Concentrations and Anomalous Values

In summary tables and plots of median profiles, concentrations will be set to missing if one of the following cases is true:

- A concentration has been reported as ND (ie, not done) or NS (ie, no sample);
- A deviation in sampling time is of sufficient concern or a concentration has been flagged as anomalous by the clinical pharmacologist.

Summary statistics will not be presented at a particular timepoint if more than 50% of the data are missing. For analysis of pharmacokinetic concentrations, no values will be imputed for missing data.

Actual PK sampling times will be used in the derivation of PK parameters. If a PK parameter cannot be derived from the concentration data, the parameter will be coded as NC (ie, not calculated). NC values will not be generated beyond the day that a patient discontinues.

In summary tables of concentration-time profiles or PK parameters, statistics will be calculated by setting NC values to missing; and statistics will not be presented for the treatment if more than 50% of the data are not collected, not calculated, or below LLOQ. For statistical analyses (ie, analysis of variance), PK parameters coded as NC will also be set to missing.

If an individual patient has a known biased estimate of a PK parameter (due for example to a deviation for the assigned dose level), this will be footnoted in summary tables and will not be included in the calculation of summary statistics or statistical analyses.

5.3.4. Missing ECG Data

For ECG analyses, no values will be imputed for missing data (except RR values that can be derived from HR values, if present. If both RR and HR values are missing QTcB will not be determined). In case of missing RR value, RR (msec) will be derived as (60/HR (bpm))* 1000, if HR is collected.

If one or two of the triplicate measurements for an ECG parameter are missed, the average of the remaining two measurements or the single measurement can be used in the analyses. If all triplicate measurements are missing at a timepoint for an ECG parameter, no values will be imputed for this timepoint. If the triplicate needs to be repeated because of an artifact, then the repeated triplicate will be reported on an unscheduled CRF page. Based on a review of the data these unscheduled assessments may be used in place of the assessments at the nominal time. Data review and consultation with the study team is required to flag these cases.

CCI

[redacted]

6. ANALYSES AND SUMMARIES

All efficacy analyses will be performed by Cohort on the SA set.

Radiographic images and clinical information collected on study will be reviewed by an ICR. ICR assessment will be used for the primary analysis of OR and for all secondary endpoints based on radiological assessments of tumor burden (ie, IC-OR, DoR, IC-DoR, TTR and PFS). These endpoints will also be derived using the local radiologist's/investigator's assessment.

6.1. Primary Endpoint(s): Objective Response in Cohort 1

The primary efficacy analysis will evaluate the anti-tumor effect of lorlatinib as a single agent as measured by OR per RECIST v1.1 in locally advanced or metastatic ALK-positive NSCLC patients whose disease has progressed after crizotinib treatment.

OR is defined as a complete response (CR) or partial response (PR) recorded from enrollment until disease progression or start of new anti-cancer therapy. Both CR and PR must be confirmed by repeat assessments performed no less than 4 weeks after the criteria for response are first met.

Objective Response Rate (ORR) is defined as the percentage of patients with a best overall confirmed response of CR or PR according to RECIST v1.1 relative to the total patients in the analysis population. Patients without documented CR or PR will be considered as non-responders. Additionally, patients with inadequate data for tumor assessment (eg, no baseline assessment or no follow-up assessments) will be considered as non-responders. The primary evaluation of ORR will be based on ICR's review relative to the SA set.

The exact test will be performed to test the null hypothesis H0: ORR<=30%. One-sided p-value will be provided together with the estimate of ORR and the corresponding exact 95% CI.

Best overall response (BOR) will be assessed based on reported overall responses at different evaluation timepoints by the independent radiologist and by the investigator from the date of first dose until documented disease progression or start of new anti-cancer therapy, according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart and documented before progression and start of new anti-cancer therapy.
- PR = at least two determinations of PR or better (and not qualifying for a CR) at least 4 weeks apart and before progression and start of new anti-cancer.
- SD (for patients with at least one measurable lesion at baseline)= at least one SD assessment (or better and not qualifying for CR or PR) ≥6 weeks after date of first dose and before progression and the start of new anti-cancer therapy.
- Non-CR/Non-PD (for patients with only non-target disease at baseline) = at least one Non-CR/Non-PD assessment (or better and not qualifying for CR or PR) ≥6 weeks after date of first dose and before progression and the start of new anti-cancer therapy.
- PD = progression ≤14 weeks after date of first dose and not qualifying for CR, PR or SD.
- Not Evaluable (NE) = all other cases.

Clinical deterioration will not be considered as documented disease progression.

Patients with BOR of NE will be summarized by reason for having NE status. The following reasons will be used:

- Early death defined as death prior to 6 weeks after date of first dose;
- No post-baseline assessments, for reasons other than early death;
- All post-baseline assessments have overall response NE;
- New anti-cancer therapy started before first post-baseline assessment;

- SD of insufficient duration (<6 weeks after first dose);
- PD too late (>14 weeks after first dose).

Special and rare cases where BOR is NE due to both early SD and late PD will be classified as 'SD of insufficient duration'.

The frequency (number and percentage) of patients with BOR of CR, PR, SD, PD, non-CR/non-PD (applicable only to patients with non-measurable disease at baseline), and NE (not-evaluable) will be tabulated.

Waterfall plots displaying the best percentage change from baseline in tumor size will be presented. The best percent change from baseline will be calculated from start of study treatment up to first visit with disease progression or to the last visit available prior to the start of new anti-tumor therapy (for patients who start new anti-tumor therapy prior to progression or who are progression-free patients at the time of analysis).

A spider plot displaying the % change from baseline across visits in tumor size will be presented. The % change from baseline will be calculated from start of study treatment to each visit up to first visit with disease progression or to the last visit available prior to the start of new anti-tumor therapy (for patients who start new anti-tumor therapy prior to progression or who are progression free patients at the time of analysis).

6.1.1. Sensitivity/Robustness Analyses (if needed)

The Disagreement Rate between investigator's assessment and ICR's assessment for OR will be calculated in Cohort 1 in the safety analysis population.

Derived ORR per mRECIST v1.1 based on tumor assessment per ICR will be calculated in Cohort 1 in the safety analysis population(details can be found in Appendix 1).

Sensitivity analyses may be conducted calculating the same ORRs relative to the PP set, if applicable.

6.2. Secondary Endpoint(s)

6.2.1. OR in Cohort 2

The OR in Cohort 2, will be summarized in a similar way as OR in Cohort 1. The estimate of ORR and the corresponding 95% CI will be provided. The frequency (number and percentage) of patients with BOR of CR, PR, SD, PD, non-CR/non-PD (applicable only to patients with non-measurable disease at baseline), and NE (not-evaluable) will be tabulated. The waterfall plot and spider plot will be presented.

6.2.2. Intracranial Objective Response (IC-OR) in Both Cohorts

The IC-OR, will be summarized similar to OR as described above in the subset of the SA population with at least 1 intracranial lesion.

Intracranial ORR is defined as the percent of patients with intracranial response (ie, Best Overall Intracranial Response as confirmed CR or confirmed PR considering only the Lesions having Disease Site=Brain) relative to patients with Brain lesions at study entry, and will be provided along with the corresponding 95% confidence interval. This includes patients for whom the brain lesions have been choosen as RECIST target lesions or not. Surgery or radiotherapy of extracranial lesions will not affect the determination of IC-OR.

6.2.3. Duration of Response (DoR) in Both Cohorts

DoR is defined as the time from the first documentation of CR or PR to the first documentation of disease progression or death due to any cause, whichever occurs first.

DoR will only be calculated for the subgroup of patients with a confirmed objective tumor response. DoR will be summarized in this subgroup using Kaplan-Meier methods and will be displayed graphically where appropriate. The median event time (if applicable) and 2-sided 95% CI for the median will be calculated according to Brookmeyer and Crowley.[2] If the number of patients with a confirmed CR or PR is small and the use of Kaplan-Meier methods may be limited, the DoR will be listed.

In case the number of patients with Progressive Disease after a confirmed CR or PR is small, the use of Kaplan-Meier method may be limited so the DR and IC-DR will be summarized using Number (%) of Subjects Censored with DR less than 6 months, ≥6 - <9 , ≥9-<12, ≥12-<15, ≥15-<18, ≥18-<21, ≥21-<24, >=24 months.

6.2.4. Duration of Intracranial Response (IC-DoR) in Both Cohorts

The IC-DoR, will be summarized similar to DoR as described above in the subset of the SA population with at least 1 intracranial lesion. Surgery or radiotherapy of extracranial lesions will not affect the determination of IC-DoR.

6.2.5. Time to Tumor Response (TTR) in Both Cohorts

TTR is defined as the time from first dose to first documentation of objective tumor response (CR or PR). TTR will only be summarized for the subgroup of patients with an objective response. Descriptive statistics mean, median and range will be provided.

6.2.6. Progression-free Survival (PFS) in Both Cohorts

Progression-Free Survival (PFS) is defined as the time from date of first dose to the date of the first documentation PD per RECIST v1.1 as assessed by the independent oncologist and investigator or death due to any cause, whichever occurs first and will be summarized in months:

PFS (months) = [date of event or censoring–date of first dose+1]/30.4375

PFS data will be censored as follows:

- For patients who start a new anti-cancer therapy (as defined in Section 5.2.6) prior to an event, censoring will be at the last adequate tumor assessment (see Section 5.2) prior to the start of new anti-cancer therapy. Note: if date of progression occurs on the same date as the start of new anti-cancer therapy, the progression will be counted as an event.

- For patients with documented progression or death after two or more missing tumor assessments, censoring will occur at the last adequate tumor assessment prior to the missing assessments. In this study antitumor activity will be assessed through radiological tumor assessments conducted at screening and every 6 weeks until disease progression regardless of initiation of subsequent anti-cancer therapy. The allowable time window for disease assessments is ±7 days while on treatment and whenever disease progression is suspected (eg, symptomatic deterioration). Therefore time without adequate assessment is defined as 98 days (12 weeks plus 2 weeks).

- For patients who do not have an adequate baseline tumor assessment or who do not have any post-baseline tumor assessments, censoring will occur on the date of first dose unless death occurred on or before the time of the second planned tumor assessment (ie, on or before Day 98) in which case the death will be considered an event. Note for patients who died without any post baseline assessments and meet the definition of two or more missing assessment the reason for censoring will be documented as two or more missed assessments and censoring will occur on the date of first dose.

- All other patients alive without objective progression will be censored on the date of the last adequate tumor assessment.

The date of tumor response at each nominal timepoint based on the independent radiologist assessments will be used for determining the dates of last adequate for censoring.

The date of tumor response at each nominal timepoint based on the independent radiologist assessments will be used for determining the dates of last adequate assessment for censoring purposes.

The censoring and event date options to be considered for the PFS analysis are presented in Table 2.

Table 2. Outcome and event dates for PFS analyses

Scenario	Date of event/censoring	Outcome
No adequate baseline assessment	Date of first dose	Censored[a]
Progression or death ≤98 days after last adequate tumor assessment or 98 days after date of first dose	Date of progression or death	Event
Progression or death >98 days after the last adequate tumor assessment[b]	Date of last adequate assessment b documenting no PD prior to anti-cancer therapy or missed assessments	Censored
No progression		
New anti-cancer therapy given prior to PD		

a. If the patient dies ≤98 days after date of first dose the death is an event on the death date.

b. If there are no adequate post-baseline assessments prior to PD or death, then the time without adequate assessment should be measured from the date of first dose; if the criteria were met the censoring will be on the start date.

Kaplan-Meier estimates (product-limit estimates) will be presented by Cohort together with a summary of associated statistics including the median PFS time with two-sided 95% CIs. The PFS rate at clinical meaningful timepoints will be estimated with corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley[2] and the CIs for the survival function estimates at the timepoints defined above will be derived using the log(-log) method according to Kalbfleisch and Prentice[3] (conftype=loglog default option in SAS Proc LIFETEST). The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of patients with each event type (PD or death) and censoring reasons will be presented by Cohort along with the overall event and censor rates.

Reasons for censoring should be summarized according to the categories in Table 3. If a patient meets multiple definitions for censoring the list will be used to define the hierarchy.

Table 3. Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No adequate baseline assessment	No adequate baseline assessment
2	Start of new anti-cancer therapy before event.	Start of new anti-cancer therapy
3	Event more than 98 days from last adequate post-baseline tumor assessment/start date	Event after missing assessments[a]
4	No event and [withdrawal of consent date ≥ start date OR End of study (EOS) = Subject refused further follow-up]	Withdrawal of consent
5	No event and lost to follow-up in any disposition page	Lost to follow-up
6	No event and [EOS present OR disposition page for any EPOCH after screening says patient will not continue into any subsequent phase of the study] and no adequate post-baseline tumor assessment	No adequate post-baseline tumor assessment
7	No event and none of the conditions in the prior hierarchy are met	Ongoing without an event

[a] more than 98 days after last adequate tumor assessment.

The PFS time or censoring time and the reasons for censoring will also be presented in a patient listing.

Time of Follow-Up for PFS

A Kaplan-Meier summary table for PFS follow-up duration will also be generated to assess the follow-up time in the treatment arms reversing the PFS censoring and event indicators.

CCI [redacted]

CCI [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

CCI [redacted]

CCI [redacted]

6.2.7. Overall Survival (OS) in Both Cohorts

Overall survival (OS) is defined as the time from date of first dose to the date of death due to any cause. Patients last known to be alive will be censored at date of last contact (see Section 5.2.7). OS will be summarized in months:

OS (months) = [date of death or censoring–start date +1]/30.4375

OS time associated with Cohort will be summarized using the Kaplan- Meier method (product-limit estimates) and displayed graphically where appropriate. CIs for the 25th, 50th, and 75th percentiles will be reported.

The OS rate at clinical meaningful timepoints will be estimated with corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley[2] and the CIs for the survival function estimates at the timepoints defined above will be derived using the log(-log) method according to Kalbfleisch and Prentice[3] (conftype=loglog default option in SAS Proc LIFETEST). The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of patients with an event (death) and censoring reasons will be presented by treatment arm. Censoring reasons are as follows:

- Alive;
- Withdrawal of consent;
- Lost to follow-up.

The OS time or censoring time and the reasons for censoring will also be presented in a patient listing.

Time of Follow-Up for OS

A Kaplan-Meier summary table for OS follow-up duration will also be generated to assess the follow-up time in Cohort reversing the OS censoring and event indicators.

6.2.8. Pharmacokinetics

For lorlatinib, concentrations will be summarized descriptively (n, mean, SD, CV, median, minimum, maximum, geometric mean and its associated CV) by cohort, cycle, day and nominal time. Individual patient and median profiles of the concentration-time data will be plotted by cohortcohortcycle and day (single dose and steady-state). For individual patient plots by time, the actual PK sampling time will be used. For summary statistics and mean/median plots by sampling time, the nominal PK sampling time will be used. MedianMedian profiles will be presented on both linear-linear and log-linear scales.

Presentations for concentrations will include but not be limited to:

- Listing of all concentrations of lorlatinib sorted by cohort, day of assessment, patient ID and nominal time post dose. The listing of concentrations will include the actual times. Deviations from the nominal time will be given in a separate listing.
- Summary of concentrations by cohort, day of assessment and nominal time post dose, where the set of statistics will include n, mean, standard deviation, coefficient of variation (CV), median, minimum, maximum, geometric mean and its associated CV and the number of concentrations above the lower limit of quantification.

- Linear and semi-log plots of median concentrations against nominal time post dose by cohort, day of assessment (based on the summary of concentrations by day of assessment and time post dose).

- Linear plots of individual concentrations against actual time post dose by cohort and day of assessment.

- Trough concentrations will be plotted using a box-whisker plot by cohort, cycle and day in order to assess the attainment and maintaining of steady-state.

Plasma pharmacokinetic parameters of lorlatinib on Day 1 of Cycle 1 (C1D1) and at steady state including the maximum plasma concentration (C_{max}), time to maximum plasma concentration (T_{max}), area under the plasma concentration versus time profile from time 0 to time t (AUC_t) and area under the plasma concentration versus time profile within a dose interval (AUC_{tau}) for lorlatinib will be estimated using non-compartmental analysis. If data permit or if considered appropriate, area under the plasma concentration versus time curve to infinity (AUC_{inf}), terminal elimination half-life ($t_{1/2}$), oral plasma clearance (CL/F), apparent volume of distribution (V_z/F) and accumulation ratio (R_{ac}) will be also estimated. The C1D1 and steady-state PK parameters will be summarized descriptively by cycle and day and will include the set of summary statistics as specified in the table below:

Parameter	Summary Statistics
AUC_{last}, AUC_{inf}*, AUC_{tau}, C_{max}, CL/F*, V_z/F*,	N, arithmetic mean, median, cv%, standard deviation, minimum, maximum, geometric mean, geometric cv%.
T_{max}	N, median, minimum, maximum.
$t_{1/2}$, Rac*	N, arithmetic mean, median, cv%, standard deviation, minimum, maximum.

* if data permits.

Pharmacokinetic and Exposure-Response Analysis

Pharmacokinetic and pharmacodynamics data from this study may be analyzed using compartmental or mixed-effect modeling approaches and may also be pooled with other study results. PK/PD modeling may be attempted to investigate any causal relationshipbetween lorlatinib exposure and efficacy, CCI [redacted] or significant safety endpoints. The results of these analyses, if performed, will be reported separately.

CCI [redacted]

CCI

[redacted]

6.4. Baseline and Other Summaries

6.4.1. Baseline Summaries

- **Demographic and Physical Characteristics**

  The following demographic and baseline characteristics will be summarized by number and percentage:

  - Gender(male, female);
  - Age <(18-<45; 45- <65; ≥65) >;
  - Eastern Cooperative Oncology Group (ECOG) Performance status;
  - Smoking classification (never smoker, former smoker, smoker).

Age (continuous), height (cm), weight (kg) will be summarized with descriptive statistics (mean, median, standard deviation, minimum, and maximum).

- **Disease Characteristics**

  The following baseline disease characteristics will be summarized by number and percentage:

  - Histopathological Classification TMN stage at initial diagnosis and current stage;
  - Extent of disease (locally/advanced, metastatic);
  - Measurable disease at baseline (yes/no) (see Section 5.2.8);
  - Adequate baseline assessment (yes/no) (see Section 5.2.11);
  - Involved tumor sites at baseline.

Time since diagnosis (months), defined as (date of first dose– date of diagnosis)/30.4375, will be summarized by descriptive statistics (mean, median, standard deviation, minimum, and maximum).

Involved tumor sites at baseline will be derived from target and non target lesions at baseline. Each patient will be counted once per organ. Similarly,number of sites of disease at baseline will be derived by counting the number of unique organ sites from target and non target lesions at baseline. "Other" will be counted as one organ site.

- **Medical History**

  - Medical history will be coded using the most current version of MedDRA and summarized by MedDRA's System Organ Class (SOC) and PT. Each patient will be counted only once within each PT or SOC. Summaries will be ordered by primary SOC and PT in descending order of frequency by Cohort. Separate summaries will be provided for past and present conditions.

- **Prior Anti-Cancer Treatments**

  Prior anti-cancer treatments include systemic therapy, radiation, and surgery.

  The number and percentage of patients in each of the following anti-cancer therapy categories will be tabulated:

  - Patients with at least one type of prior anti-cancer treatment;Patients with different type of previous ALK inhibitor;
  - Patients with at least one prior anti-cancer drug therapy;
  - Patients with at least one prior anti-cancer radiotherapy;
  - Patients with at least one prior anti-cancer surgery.

  Prior anti-cancer drug therapy will be summarized as follows based on the number and percentage of patients:

  - Number of prior anti-cancer therapy regimens: 0/ 1 / >1;
  - Intent of Therapy: Neo-Adjuvant/Adjuvant/Advanced – Metastatic.

6.4.2. Study Conduct and Subject Disposition

The following analyses will be based on the FA population overall and separately by Cohort.

6.4.2.1. Patient Disposition

A summary of the number of patients enrolled by site will be provided by Cohort as well as overall for the FA population.

Discontinuations from study treatment will be summarized using the SA set.

Discontinuations from study treatment due to adverse events will be identified as either related or not related to study treatment. If causality is missing the event will be considered related to treatment. If multiple events lead to study treatment discontinuation and at least one was considered related, discontinuation will be reported as related to study treatment.

6.4.2.2. Protocol Deviations

Protocol deviations will be compiled prior to database closure and will be listed and summarized by category (n(%)) for the FA set by Cohort. Categories will be assigned by the study Clinician.

6.4.3. Study Treatment Exposure

The following analyses will be based on the SA population.

6.4.3.1. Exposure to lorlatinib

The summary of treatment exposure for lorlatinib will include the following information:

- Treatment duration (weeks);
- Cumulative dose (mg);
- Dose intensity (mg/week);
- Relative dose intensity (%);

The duration of lorlatinib (in weeks) is defined as:

$$\text{Treatment duration (weeks)} = (\text{last dose date} - \text{first dose date} + 1)/7$$

The cumulative dose (mg) of lorlatinib is the sum of the actual dose levels that the patient received (ie, total dose administered (mg)). Lorlatinib is given once daily, administered as 4 x 25 mg oral tablets, continuously; therefore if all doses (eg, d mg/day) were received in a week the cumulative dose would be 7× 100 mg.

The dose intensity (DI) and the relative dose intensity (RDI) of lorlatinib will be calculated for each patient during the study. The DI (mg/week) of lorlatinib during the study is defined as

$$\text{DI (mg/week)} = [\text{cumulative dose (mg)}]/[\text{treatment duration (weeks)}]$$

The RDI of lorlatinib is defined as the ratio of the DI and planned dose intensity and expressed in %

$$\text{RDI (\%)} = 100\times [\text{DI (mg/week]}/[7\times 100 \text{ (mg/week)}].$$

6.4.3.2. Dose Reductions, Interruptions, and Delays

A dose reduction is defined as a non-zero dose that is less than the prior dose.

The number and percentage of patients with at least one dose reduction as well as a breakdown of dose reductions (1/2/3) will be summarized by Cohort.

Reasons for dose reductions will also be summarized. Patients can contribute to more than one reason if multiple dose reductions occurred for different reasons, but will only be counted once per reason. Percentages will be calculated based on the total number of patients in SA set.

An interruption is defined a 0 mg dose administered for more than 1 day. The number and percentage of patients with dose interruptions and the corresponding reasons will be summarized by Cohort. Patients can contribute to more than one reason if multiple dose interruptions occurred for different reasons, but will only be counted once per reason. Percentages will be calculated based on the total number of patients in the SA set.

What follows defines how dose interruptions will be counted in the case of consideratios for multiple dose interruptions:

- If an interruption occurs consecutively for at least two days due to the same reason, then it will be counted only once.
- If an interruption occurs consecutively for at least two days due to different reasons, then it will be counted for each reason.
- If an interruption occurs for more than one day due to the same reason, but the days are not consecutive, ie, there is at least one dosing day in between, then each dose interruption will be counted as a different occurrence.

6.4.4. Concomitant Medications and Non-Drug Treatments

Concomitant medications and non-drug treatments received by patients during the study will be summarized for the SA set by Cohort.

Concomitant medications refer to all medications which started prior to first dose of study treatment and continued during the on-treatment period (see Section 5.2) as well as those started during the on-treatment period. Concomitant medications will be coded in the WHO Drug coding dictionary and will be tabulated by Anatomical Therapeutic Chemical (ATC) Classification level 2 and preferred term in descending order of frequency in the lorlatinib arm. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. A patient will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. Preferred Terms will be reported under each ATC class that it is included under within WHO Drug (no primary path is available in WHO Drug).

Concomitant non-drug treatments refer to all non-drug treatments administered during the on-treatment period. Non-drug treatments will be coded with the most current version of MedDRA and will be summarized by MedDRA's SOC and PT in descending order of frequency in the lorlatinib arm. Patients will be counted only once per PT even if he/she received the same treatment multiple times.

Any medications or non-drug treatments, aside from anti-cancer treatments described in Section 6.4.1 which were only administered prior to treatment start will be listed but not summarized.

6.4.5. Subsequent Anti-Cancer Therapies/Procedures

Subsequent Anti-Cancer Therapies and Procedures are defined as therapies entered on the 'Follow-up Cancer Therapy', 'Follow-up Radiation Therapy', and 'Follow-up Surgery' CRF pages. The number and percentage of patients within each category (medication therapy, radiation therapy, and surgeries) will be provided by Cohort.

Medications will be coded using the WHO Drug coding dictionary and will be tabulated by preferred term in descending order of frequency on the study treatment.

Analyses will be based on the FA population by Cohort.

6.5. Safety Summaries and Analyses

Unless otherwise specified, summaries of AEs and other safety parameters will be based on the safety analysis set by Cohort and Total.

6.5.1. Adverse Events

All analyses will be based on treatment emergent events unless otherwise specified. Treatment emergent is defined in Section 3.5.1. AEs not considered treatment emergent will be flagged in data listings.

A high level summary of adverse events will include the number and percent of patients with:

- Any Adverse Event;
- Serious AE;
- Adverse Events with CTCAE Grade 3-4;
- Grade 5 events;
- AEs leading to dose interruptions;
- AEs leading to dose reductions;
- AEs leading to withdraw.

Additionally, the number of any AE will be provided. Each unique adverse event at the PT level in each Cohort of the study for a patient is included in the count.

Seriousness, toxicity grade, action taken (interruption, reduction, and withdraw) are as reported by the investigator on the adverse event CRF.

Summaries by Cohort, SOC and PT in decreasing frequency based on the frequencies observed in each Cohort will be provided for:

- Treatment Emergent Events by Maximum Toxicity Grade (All Causality);
- Treatment Emergent Events by Maximum Toxicity Grade (Treatment Related);
- Serious Treatment Emergent Events (All Causality);
- Serious Treatment Emergent Events (Treatment Related).

An event will be considered treatment related if the investigator considered the event related to the study drug or this information is unknown.

The following summaries will be provided by Cohort and PT/clusters of AEs (ie, summaries will not include SOC) in decreasing frequency based on the frequencies observed in the study treatment for:

- Treatment Emergent Events (All Causality) by Preferred Term (including Clusters of Preferred Term)and Maximum Toxicity Grade;
- Treatment Emergent Events (Treatment Related) by Preferred Term(including Clusters of Preferred Term) and Maximum Toxicity Grade;
- Treatment Emergent Grade 3-5 Events (All Causality) by Preferred Term (including Clusters of Preferred Term) and Maximum Toxicity Grade;
- Treatment Emergent Adverse Events Leading to Dose Interruptions by Maximum Toxicity Grade (All Causality);
- Treatment Emergent Adverse Events Leading to Dose Reductions by Maximum Toxicity Grade (All Causality);
- Treatment Emergent Adverse Events Leading to Permanent Withdraw by Maximum Toxicity Grade (All Causality);
- Serious Treatment Emergent Events (All Causality).
- Serious Treatment Emergent Events (Treatment Related).

Each patient will be counted only once within each PT.

As described in Section 5.3 in case a patient has events with missing and non missing toxicity grades, the maximum of the non-missing grade will be displayed. Missing grade will only be displayed in the event that only one event has been reported for a patient and the grade is missing.

6.5.1.1. Adverse Events of Special Interest

These analyses will be performed for treatment emergent AEs of special interest as specified in Section 3.5.2.

- **Time to AE Onset** (in days) is defined as the time from the date of the first dose to the onset date of the AE, regardless of grade. If a patient has multiple episodes of an AE, the date of the first occurrence is used. Time to AE onset (in days) will be calculated as (AE start date – first dose date +1). Time to onset is calculated for the subgroup of patients who had the specific AE

- **Time to Grade 3 or 4 AE Onset** (in days) is defined similarly as time to AE onset for Grade 3 or 4 AEs.

- **Duration of AE** (in days) is defined as the cumulative duration across episodes of the AE, regardless of grade, where duration for each episode is the time from the AE start date to the AE end date. For one episode, duration (in days) = AE end date – AE start date + 1. If a patient has multiple episodes of an AE, cumulative duration across all episodes will be used adjusting for any overlap. If a patient has an AE that was ongoing at the time of analysis, the time is censored at the last available on treatment visit date. Duration is calculated for the subgroup of patients who had the specified AE.

Descriptive statistics will be presented for time to AE onset (days), time to Grade 3 or 4 AE onset, and duration of AEs for the subgroup of patients with the AE.

6.5.2. Deaths

The frequency (number and percentage) of patients in the safety analysis set who died and who died within 28 days after last dose of study treatment as well as the primary reason for death, will be tabulated based on information from the 'Notice of Death' and 'Survival Follow-Up' CRFs, by Cohort.

The frequency (number and percentage) of patients in the safety analysis set who died during follow-up period after 28 days after the last dose of study treatment will also be provided.

Date and cause of death will be provided in individual patient data listing together with selected dosing information (study treatment received, date of first/last administration, dose).

6.5.3. Laboratory Data

Laboratory results will be converted to International System of Units (Système International d'unités, SI) units which will be used for applying toxicity grades and for all summaries.

As described in Section 3.4, baseline will defined as the last assessment performed on or prior to date of the first dose of study treatment. If there are multiple assessments that meet the baseline definition on the same day without the ability to determine which was truly last, then the worst grade will be assigned as the baseline grade. Since a few CTCAE terms (including Hypo/Hypercalcemia and Activated Partial Thromboplastin) can be derived using several laboratory tests (analytes) refer to the section 2.3.7 of the "Pfizer Oncology CTCAE Grading Implementation Guidance for Laboratory Data" for determination of baseline CTCAE grade in this situation.

Additionally, laboratory results will be programmatically classified according to NCI-CTCAE version 4.03 grade. Non-numerical qualifiers will not be taken into consideration in the derivation of grade (eg, hypokalemia Grade 1 and Grade 2 are only distinguished by a non-numerical qualifier and therefore Grade 2 will not be derived). In summary statistics the number and percentage of patients corresponding to grades that only include non-quantitative criteria will be displayed as a blank or NA (not assessed) rather than 0. If there is any overlap between grade criteria (eg, CTCAE grading criteria for Creatinine Increased - a value can fall into one range based on comparison to ULN and another range based on comparison to baseline), the highest (worst) grade would be assigned to that record. Grade 5 is defined in the CTCAE criteria guidance as an event with an outcome of death. Since laboratory data does not collect an outcome, Grade 5 is not used when programmatically grading laboratory data.

Grade 0 or Outside Toxicity Reference (OTR) is not defined specifically by in the CTCAE guidance. However, programmatically this is used as a category to represent those patients who did not meet any of the Grades 1 to 4 criteria. If the laboratory value is evaluable for CTCAE criteria grading (numeric value is present, valid units and ranges are present as required to allow conversion to standard units and grading), and does not qualify for any of the Grade 1-4 criteria for a given lab test, then the value is assigned as Grade 0 or OTR.

Abnormalities will be described using the worst grade by scheduled timepoint and overall. Worst grade by scheduled timepoint will be determined using only central laboratory results. Worst case overall will be determined using both central and local laboratory results from scheduled and unscheduled visits. Several laboratory tests have bi-directional grading criteria defined so that both low (hypo) and high (hyper) values can be graded separately. Each criterion will be summarized separately. In the cases where a value is graded as a Grade 1, 2, 3, or 4 for one of the directions, that value will also be assigned as a Grade 0 for the opposite direction for that test. For example, a value meeting the criteria for Grade 3 Hypercalcemia will be classified as a Grade 0 Hypocalcemia. For CTCAE terms that can be derived using one of several laboratory tests, the maximum post-baseline grade for a given patient and CTCAE term will be the maximum across all possible laboratory tests.

For **WBC differential counts** (total neutrophil [including bands], lymphocyte, monocyte, eosinophil, and basophil counts), the absolute value will be used when reported by the lab. When only percentages are available (this is mainly applicable for neutrophils and lymphocytes, because the CTCAE grading is based on the absolute counts), the absolute value is derived as follows:

Derived differential absolute count = (WBC count) × (Differential %value/100)

If the investigator reports both the absolute and % value for Neutrophils or Lymphocytes from the same laboratory sample date and patient, ONLY the absolute value will be graded. The % value will not be graded in this scenario.

If the % value is converted to the differential absolute count for grading and the LLN for the differential absolute count is not available (only LLN for % is available) then Grade 1 will be assigned if the following conditions are met:

- Lymphocyte count decreased:
  - derived absolute count does not meet Grade 2-4 criteria, and
  - % value < % LLN value, and
  - derived absolute count ≥800/mm^3.
- Neutrophil count decreased:
  - derived absolute count does not meet Grade 2-4 criteria, and
  - % value < % LLN value, and
  - derived absolute count ≥1500/mm^3.

For **calcium**, CTCAE grading is based on Corrected Calcium and Ionized Calcium. Corrected Calcium is calculated from Albumin and Calcium as follows:

- Corrected Calcium (mg/dL) = Calcium (mg/dL) – 0.8 [Albumin (g/dL)-4].

Laboratory toxicities will be tabulated using descriptive statistics (number of patients and percentages):

- Shift table will summarize baseline CTCAE grade versus the worst post-baseline CTCAE grade.

Liver function tests: Alanine aminotransferase (ALT), aspartate aminotransferase (AST), Alkaline Phosphatase (ALP), and total bilirubin (TBILI) are used to assess possible drug induced liver toxicity. The ratios of test result over the ULN will be calculated and classified for these three parameters during the on-treatment period.

Summary of liver function tests will include the following categories. The number and percentage of patients with each of the following during the on-treatment period will be summarized by treatment arm:

- ALT ≥3×ULN, ALT ≥5xULN, ALT ≥10×ULN, ALT ≥20×ULN;

- AST ≥3×ULN, AST ≥5xULN, AST ≥10×ULN, AST ≥20×ULN;
- (ALT or AST) ≥3×ULN, (ALT or AST) ≥5×ULN, (ALT or AST) ≥10×ULN, (ALT or AST) ≥20×ULN;
- TBILI ≥2×ULN;
- Concurrent ALT ≥3×ULN and TBILI ≥2×ULN;
- Concurrent AST ≥3×ULN and TBILI ≥2×ULN;
- Concurrent (ALT or AST) ≥3×ULN and TBILI ≥2×ULN;
- Concurrent (ALT or AST) ≥3×ULN and TBILI ≥2×ULN and ALP >2×ULN;
- Concurrent (ALT or AST) ≥3×ULN and TBILI ≥2×ULN and ALP ≤2×ULN or missing.

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, ie, a patient with an elevation of AST ≥10×ULN will also appear in the categories ≥5×ULN and ≥3×ULN. Liver function elevation and possible Hy's Law cases will be summarized using frequency counts and percentages.

An evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot will also be created, with different symbols for different treatment arms, by graphically displaying:

- Peak serum ALT(/ULN) vs peak total bilirubin (/ULN) including reference lines at ALT=3×ULN and total bilirubin=2×ULN.
- Peak serum AST(/ULN) vs peak total bilirubin (/ULN) including reference lines at AST=3×ULN and total bilirubin=2×ULN.

In addition, a listing of all TBILI, ALT, AST and ALP values for patients with a post-baseline TBILI ≥2×ULN, ALT ≥3×ULN or AST ≥3×ULN will be provided.

6.5.4. Vital Signs

Vital signs data includes weight, pulse, systolic blood pressure, and diastolic blood pressure. Measurements were only to be provided once per timepoint. If multiple assessments are provided per timepoint, the maximum value will be used for reporting.

The number and percent of patients in each of the following minimum and maximum blood pressure, body weight, and pulse categories will be presented:

- Increase in Systolic Blood Pressure ≥40 mmHg;
- Decrease in Systolic Blood Pressure ≥40 mmHg;

- Increase in Diastolic Blood Pressure ≥20 mmHg;
- Decrease in Diastolic Blood Pressure ≥20 mmHg;
- Increase in Diastolic Blood Pressure ≥60 mmHg;
- Decrease in Body Weight ≥10%;
- Increase in Body Weight between 10% and 20%;
- Increase in Body Weight (increase) ≥20%;
- Maximum Pulse Rate >120 bpm;
- Minimum Pulse Rate <50 bpm;
- Maximum increase in pulse rate ≥30 bpm;
- Maximum decrease in pulse rate ≥30 bpm.

All assessments, including unscheduled assessments will be considered. A patient can be included in multiple categories if different criteria are met at different timepoints.

6.5.5. Electrocardiogram

Triplicate ECGs were required at each assessment. A mean score is calculated for any replicate measurements having the same nominal visit. The mean measurement is reported.

ECG summaries will include all ECG assessments from the on-treatment period. All ECG assessments will be listed, and those collected outside the on-treatment period will be flagged in the listing.

The analysis of QT data is complicated by the fact that the QT interval is highly correlated with heart rate. Because of this correlation, formulas are routinely used to obtain a corrected value, denoted QTc, which is independent of heart rate. This QTc interval is intended to represent the QT interval at a standardized heart rate.

Fridericia's correction (QTcF) will be programmatically derived using the following formula:

$$\text{QTcF(m sec)} = \text{QT(m sec)} / \sqrt[3]{RR(sec)}$$

and Bazett's correction (QTcB) will be programmatically derived using the following formula:

$$QTcB(m\,\text{sec}) = \frac{QT(m\,\text{sec})}{\sqrt{RR(\text{sec})}},$$

where RR represents the RR interval of the ECG, in seconds.

RR used in the above formula will be derived by (60/HR (bpm)) if HR is collected. Otherwise, RR will be directly used for calculation.

Additionally QTcB and QTcF (and QTcS if applicable) will be summarized by maximum on-treatment values using the following categories:

- ≤450 msec;
- >450 msec but ≤480 msec;
- >480 msec but ≤500 msec;
- >500 msec.

Unscheduled assessments will be utilized in addition to planned assessments.

Shift tables will be provided for baseline value versus worst on-treatment value.

Additionally maximum increases from baseline (including scheduled and unscheduled assessments) will be summarized based on the following categories:

- Change >60 msec;
- Change >30 msec but ≤60 msec;
- Change ≤30 msec.

In the event that neither Fridericia's nor Bazett's correction adequately adjusts for heart rate, an additional correction such as a population or patient-specific baseline correction could be used and should be fully justified.

For PR and QRS maximum increases from baseline, the following categories will be applied:

- PR change ≥50% if absolute baseline value was <200 msec;
- PR change ≥25% if absolute baseline value was ≥200 msec;
- QRS change ≥50% if absolute baseline value was <100 msec,
- QRS change ≥25% if absolute baseline value was ≥100 msec.

If more than one ECG is measured at a nominal time post-dose (eg, triplicate ECGs within 2-4 minutes), the mean will be used to represent a single observation per patient and time post-dose. If any of the three individual ECGs results in a QTc ≥500 msec and the mean is not ≥500 msec, then that patient's data will be described in the safety section in the study report in order to place the ≥500 msec value in appropriate clinical context. On the other hand, such individual ≥500 msec value within a triplicate will not be included in the categorical analysis unless the average from that triplicate is also ≥500 msec. Data listings

will contain the means from a triplicate as well as the parameters from each of the three ECGs. Note that using the mean value may result in a patient having a measurement that is not represented by an actual ECG.

6.5.6. Physical Examination

Physical examination findings will only be listed.

6.5.7. Left Ventricular Ejection Fraction (LVEF)

LVEF% will be summarized as frequency (number and percentage) of patients with:

- A shift from baseline normal to at least one result below the institutional lower limit of normal during the on-treatment period;
- ≥20-point decrease from baseline in LVEF%.

A patient will be included in the categories above if any post-baseline assessment (including unscheduled assessments) meet the criteria; however only post-baseline assessments which use the same method of assessment (ECHO or MUGA) as baseline will be considered.

6.5.8. Performance Status

The ECOG shift from baseline to the highest score during the post-baseline period will be summarized by Cohort.

6.5.9. Mood and Suicidal Ideation and Behavior Analyses

- **The Beck Depression Inventory-II (BDI-II) - Assessment of Mood**

An assessment of mood will be administered to patients via the Beck Depression Inventory-II (BDI-II) scale at the timepoints described in the Schedule of Activities of the Study Protocol. This is a 21 item self-reported scale, with each item rated by patients on a 4 point scale (ranging from 0-3.). The scale includes items capturing mood, (loss of pleasure, sadness, and irritability), suicidal ideation, and cognitive signs (punitive thoughts, self- criticism, self-dislike pessimism, poor concentration) as well as somatic signs (appetite, sleep, fatigue, libido).

Scores are obtained by adding up the total points from the series of answers. Higher total scores indicate more severe depressive symptoms. The standardized cutoffs are as follows:

- 0–13: minimal depression;
- 14–19: mild depression;
- 20–28: moderate depression;
- 29–63: severe depression.

The following descriptive statistics will be provided: n, arithmetic mean and SD. Frequencies and percentages will be displayed on items capturing mood, (loss of pleasure, sadness, irritability), suicidal ideation, and cognitive signs (punitive thoughts, self- criticism, self-dislike, pessimism, poor concentration) as well as somatic signs (appetite, sleep, fatigue, libido) on the BDI-II scale will be summarized.

- **Columbia Suicide Severity Rating Scale (C- SSRS)**

Frequencies and percentages will be displayed for patients with suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent based on the C-SSRS during treatment.

7. INTERIM ANALYSES

No formal interim analysis will be conducted for this study. However, as this is an open-label study, the sponsor conduct unblinded reviews of the data during the course of the study for the purpose of safety assessment, facilitating pharmacokinetic/pharmacodynamic modeling, and/or to support clinical development.

8. CONSIDERATION ON COVID-19 IMPACT

During the study, COVID-19 emergency started and may have impacted study conduct. The following analysis will be conducted to assess the impact due to COVID-19.

Protocol deviations due to COVID-19 will be summaried.

A listing of patients impacted by COVID-19 will be provided.

If there are deaths due to COVID-19, sensitiviety analysis for PFS and OS may be considerd.

9. REFERENCES

1. Amit O, et al. Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis and recommendation from a PhRMA working group. European Journal of Cancer 47:1772-1778, 2011.

2. Brookmeyer R and Crowley J. A confidence interval for the median survival time. Biometrics 38:29-41, 1982.

3. Kalbfleisch JD, Prentice, RL. Statistical Analysis of Failure Time Data, 2nd Edition.Hoboken, Wiley Interscience.

Appendix 1. Derived Objective Response per mRECIST v1.1 based on ICR Data

The primary endpoint is Objective Response (OR) in patients whose disease has progressed after crizotinib (Cohort 1) by independent central radiology (ICR) assessment per mRECIST v1.1 specified in protocol Appendix 3.

Per protocol Appendix 3, up to 5 CNS lesions were permitted to be identified as target lesion in addition to 5 extracranial lesions (up to 2 lesions per organ) at baseline. Furthermore, it was expected same set of brain lesions were used for both overall and intracranial assessment per RECIST v1.1 (Protocol Appendix 3).

However, per ICR (PAREXEL) operational procedure, overall assessment and intracranial assessment are performed independently for study B7461024; Overall assessment was based on RECIST 1.1 with up to 2 CNS lesions and 5 lesions in total were permitted to be identified as target lesion at baseline. While intracranial assessment was based on modified RECIST 1.1 that allows up to 5 measurable CNS lesions. Hence the brain lesions in the overall assessment and intracranial assessment may be different. And the overall assessment was conducted in a different way as specified in protocol Appendix 3.

To mitigate the potential risk due to the difference between ICR's procedure and that defined in protocol Appendix 3, programming derived overall assessment per modified RECIST v1.1 based on ICR tumor lesion data will be conducted to support the efficacy analysis.

Given the ICR data provided by PAREXEL, for each patient, the extracranial lesions identified in the overall assessment will be combined with the intracranial lesion identified in the intracranial assessment. The tumor response will then be derived based on tumor measurement/assessment per RECIST v1.1.

As there will be 2 radiologist to review each patient, for the derived response, only the readings from the reader agreed by the adjudicator will be used for the derivation. If there's no adjudication, which means that there's no discrepancy in the two readers' assessment for each time point, the reader with more representative extracranial lesion selected at baseline will be used for the derivation. The selection will follow the below rule:

- If both the extracranial assessment and intracranial assessment have adjudicators, the adjudicator's assessment will be applied, no matter whether the extracranial assessment and intracranial assessment have chosen the same reviewer's readings.
- If only the extracranial assessment or only intracranial assessment have adjudicator but not both of them, both of the extracranial assessment and intracranial assessment based on the exist adjudicator's decision will be applied.
- If no adjudicator appeared for both of extracranial and intracranial assessment, the selection will follow the below rule:
    - The one with more extracranial target lesion at baseline in overall assessment will be used if there was no adjudication;
    - If the two readers identified the same number of extracranial target lesion, the one with the larger SOD of extracranial target lesion will be chosen.