Statistical Analysis Plan

NCT Number: NCT03596866

Title: A Phase 3 Randomized Open-label Study of Brigatinib (ALUNBRIG®) Versus Alectinib (ALECENSA®) in Advanced Anaplastic Lymphoma Kinase-Positive Non–Small-Cell Lung Cancer Patients Who Have Progressed on Crizotinib (XALKORI®)

Study Number: Brigatinib-3001

Document Version and Date: Version 2.0, 16 June 2021

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

STATISTICAL ANALYSIS PLAN

STUDY NUMBER: Brigatinib-3001

A Phase 3 Randomized Open-label Study of Brigatinib (ALUNBRIG®) Versus Alectinib (ALECENSA®) in Advanced Anaplastic Lymphoma Kinase-Positive Non–Small-Cell Lung Cancer Patients Who Have Progressed on Crizotinib (XALKORI®)

PHASE 3

Version: 2.0

Date: 16 June 2021

Prepared by: Takeda Development Center Americas, Inc (TDCA)

[redacted], PhD

[redacted], Oncology Statistics

Based on:

Protocol Version: Amendment 04

Protocol Date: 08 March 2021

1.1 Approval Signatures

Electronic signatures can be found on the last page of this document.

2.0 TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

3.0 LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ALK+	anaplastic lymphoma kinase-positive
AST	aspartate aminotransferase
BIRC	blinded Independent Review Committee
BMI	body mass index
CIF	cumulative incidence functions
CRF	case report form
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D-5L	EuroQoL–5 Dimensions–5 Levels
FA	Final analysis
FISH	fluorescence in situ hybridization
HR	hazard ratio
HRQoL	health-related quality of life
IA	interim analysis
iCNS	intracranial central nervous system
iDOR	intracranial duration of response
iORR	intracranial objective response rate
iPD	intracranial progression without prior systemic progression
ITT	intent-to-treat
IxRS	Interactive voice/web-based response system
KM	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NSCLC	non–small-cell lung cancer
OBF	O'Brien Fleming
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PT	Preferred Term
QD	once daily
QLQ-BN20	Quality of Life Brain Cancer Module
QLQ-C30	Quality of Life Questionnaire-Core 30
QLQ-LC13	Quality of Life Lung Cancer Module

RANO-BM	Response Assessment in Neuro-Oncology Brain Metastases
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SOC	System Organ Class
TLFs	tables, listings, and figures
TTR	time to response
ULN	upper limit of normal
WHO Drug	World Health Organization Drug Dictionary

4.0 OBJECTIVES

4.1 Primary Objective

The primary objective is to compare the efficacy of brigatinib to that of alectinib in patients with ALK+ locally advanced or metastatic NSCLC who have progressed on crizotinib as evidenced by PFS as assessed per Response Evaluation Criteria in Solid Tumors (RECIST) v1.1.

4.2 Secondary Objectives

The secondary objectives are:

1. To compare the efficacy of brigatinib with that of alectinib as evidenced by overall survival (OS), PFS as assessed by the investigator, ORR, DOR, and time to response (all as assessed per RECIST v1.1).
2. To compare the efficacy of brigatinib in the CNS to that of alectinib as evidenced by iORR, iDOR, and time to iPD as assessed per modified RECIST criteria.
3. To assess the safety and tolerability of brigatinib in comparison with alectinib.
4. To collect plasma concentration-time data for brigatinib to contribute to population pharmacokinetic (PK) analyses.
5. To assess patient-reported symptoms and health-related quality of life (HRQoL) with the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Core 30 (QLQ-C30) (v3.0) and its Quality of Life Lung Cancer Module (QLQ-LC13), in patients treated with brigatinib compared with those treated with alectinib.

4.3 Exploratory Objectives

The exploratory objectives are:

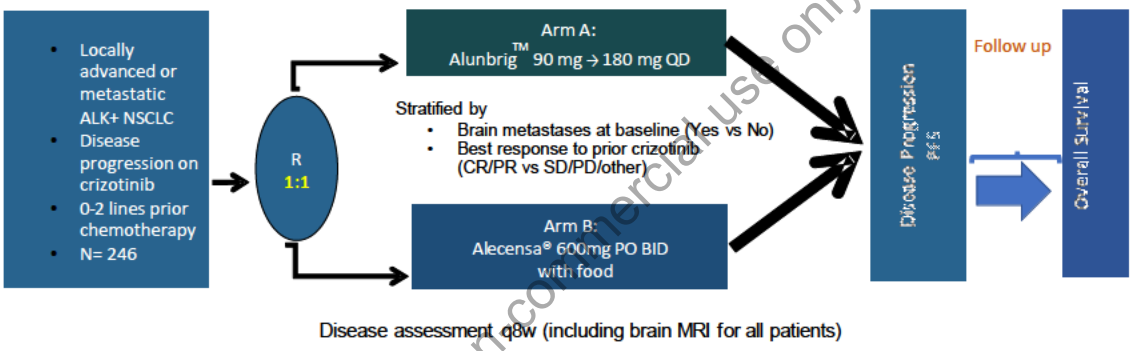
1. To compare the efficacy in the CNS of brigatinib to that of alectinib as evidenced by iORR, iDOR, and time to iPD, per the Response Assessment in Neuro-Oncology Brain Metastases (RANO-BM) criteria.
2. To explore the molecular determinants of efficacy and safety with brigatinib and alectinib.
3. To evaluate health resource utilization.
4. To use patient reported outcomes to assess morbidity related to CNS symptoms.

4.4 Study Design

This is a phase 3, randomized, open-label, comparative, multicenter, international study in which patients with ALK+ NSCLC who have progressed on crizotinib will be randomized in a 1:1 fashion to receive brigatinib in Arm A or alectinib in Arm B. Patients will be stratified by the presence of intracranial CNS metastases at baseline (Yes versus No) and best prior response to crizotinib therapy as assessed by the investigator (complete response [CR]/partial response [PR] vs any other response/status unknown).

The expected number of patients is 246 enrolled at 100 to 120 sites globally. Patients will be treated on each arm until they experience progressive disease (PD) assessed by the investigator or intolerable toxicity, or until any other discontinuation criterion is met. On each arm, treatment will be given continuously. For logistical convenience, every 28 days is defined as 1 cycle. Continuation of study drug beyond progression is permitted, if there is potential for continued clinical benefit (eg, absence of clinical symptoms or signs, indicating clinically significant disease progression requiring alternative systemic anticancer therapy; no decline in performance status; absence of rapid disease progression or threat to vital organs or critical anatomical sites [eg, respiratory failure due to tumor compression, spinal cord compression] requiring urgent use of alternative anticancer therapy; and no significant, unacceptable or irreversible toxicities related to study treatment).

Figure 4.1 Study Design



Abbreviations: ALK+, anaplastic lymphoma kinase-positive; BID, twice daily; CR, complete response; MRI, magnetic resonance imaging; NSCLC, non–small-cell lung cancer; PD, progressive disease; PFS, progression-free survival; PO, oral; PR, partial response: QD, once daily; q8w, every 8 weeks; R, randomization; SD, stable disease.

The primary endpoint of the study is to compare the efficacy of brigatinib with that of alectinib in patients with ALK+ locally advanced or metastatic NSCLC who have progressed on crizotinib as evidenced by BIRC-assessed PFS. An interim analysis (IA) to assess the primary endpoint will be conducted once 70% of PFS events (approximately 115 events) have been observed. The primary analysis for the primary endpoint will be conducted when approximately 164 PFS events have been observed, which is approximately 30 months after the first patient has been randomized. The final analysis for OS will occur when approximately 156 deaths occur, which is approximately 5 years after the first patient is randomized.

AEs will be assessed from the time a patient signs the informed consent form (ICF) until 30 days after last dose of the last study drug a patient receives. Patients' signs and symptoms, laboratory values, vital signs, ECGs, and any other relevant special examinations as clinically indicated will be obtained to evaluate the safety and tolerability of brigatinib. For patients randomized to the brigatinib arm of the study (Arm A), sparse PK samples will be collected during the study to

measure plasma concentrations of brigatinib. Before randomization, ALK status must be confirmed with documentation from a pathology report of ALK+ tumor tissue.

5.0 ANALYSIS ENDPOINTS

5.1 Primary Endpoint

- The primary endpoint is PFS as assessed by the blinded Independent Review Committee (BIRC) per RECIST v1.1.

5.2 Secondary Endpoints

Key Secondary Endpoint:

- OS.

Other Secondary Endpoints:

1. PFS, as assessed by the investigator per RECIST v1.1.
2. Confirmed ORR, as assessed by the investigator and BIRC per RECIST v1.1.
3. DOR, as assessed by the investigator and BIRC.
4. Time to response, as assessed by the investigator and BIRC.
5. Confirmed iORR, as assessed by BIRC per modified RECIST v1.1 (as described in protocol and BIRC charter).
6. iDOR, as assessed by the BIRC per modified RECIST v1.1.
7. Time to iPD without prior systemic progression, as assessed by the BIRC per modified RECIST v1.1.
8. HRQoL assessed with the global health status/quality of life and other function and symptom domains from EORTC QLQ-C30 (v3.0) and EORTC QLQ-LC13.

5.3 Safety Endpoints

The safety endpoint assessments will include physical and laboratory examinations, vital signs, and electrocardiograms (ECGs). AEs will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v4.03.

5.4 Exploratory Endpoints

1. CNS efficacy outcomes as assessed by the BIRC per RANO-BM criteria (iORR, iDOR, and time to iPD).
2. Molecular determinants of efficacy and safety with brigatinib and alectinib as uncovered by genetic alterations noted on tumor tissue DNA or plasma samples of circulating tumor DNA.
3. HRQoL measured by EuroQoL–5 Dimensions–5 Levels (EQ-5D-5L) questionnaires.
4. To evaluate healthcare resource utilization.
5. Items from the EORTC Quality of Life Brain Cancer Module (QLQ-BN20) used to assess morbidity related to CNS symptoms.

6.0 DETERMINATION OF SAMPLE SIZE

For the purposes of this sample size calculation, the median PFS for alectinib is estimated as 9 months and the median PFS for brigatinib is estimated to be 15 months on the basis of the outcomes observed in previous single-arm studies (AP26113-13-201 and AP26113-11-101). Approximately 246 patients will be randomized 1:1 to receive brigatinib or alectinib. A total of 164 PFS events (progression or death among the randomized patients) will provide approximately 90% power to detect a 6-month improvement in PFS (HR=0.60). This power projection is based on a 2-sided log-rank test and is controlled at the 2-sided 0.05 level, adjusting for the proposed interim analysis plan.

This study is not powered to demonstrate a statistically significant difference in OS with a total of approximately 246 patients and an assumption of median OS of 34 and 29 months for the brigatinib and alectinib arms, respectively, in this post-crizotinib setting (HR=0.853).

7.0 METHODS OF ANALYSIS AND PRESENTATION

7.1 General Principles

Statistical tests and confidence intervals will be reported as 2-sided and will be assessed at α=0.05 significance level, ie, 95% CI, unless otherwise stated. For IA and final analysis, the actual α level spent on each endpoint involved in formal testing will be clearly reported.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations (sd) will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate. P-values will be rounded to 3 decimal places.

Where appropriate, variables will be summarized descriptively by study visit. Unless otherwise specified, the baseline value is defined as the value collected at the time closest to, but prior to, the start of study drug administration for safety analysis and randomization date for efficacy analysis.

For the categorical variables, the count and proportion of each possible value will be tabulated by treatment group. For continuous variables, the number of subjects with non-missing values, mean, median, SD, minimum, and maximum values will be tabulated.

All statistical analyses will be conducted using SAS® Version 9.4 or higher.

7.1.1 Study Definitions

7.1.1.1 *Definition of Study Days*

Study Day 1 is defined as the date on which a subject is administered their first dose of study drug. Other study days are defined relative to the Study Day 1 with Day 1 being Study Day 1 and Day -1 being the day prior to Study Day 1.

7.1.1.2 *Definition of Study Visit Windows*

The screening period is defined as the 28 days prior to the date which a subject is administered their first dose of the study drug. Following visits may be arranged ±3 days from the Day 1 of each cycle, with the exception of disease assessment, which can be arranged with a ±7-day window. Once radiological disease progression is observed or patients have started a new systemic anticancer therapy, the survival follow-up shall be arranged with a ±14-day window.

7.1.1.3 *Conventions for Missing Adverse Event Dates*

In general, the imputation will be conservative such that onset dates will be imputed to be as early as possible and resolution dates will be imputed to be as late as possible. Resolution date will be imputed first and then used to impute incomplete onset date. If the AE resolution date is prior to the corresponding AE start date after applying the imputation rules below, it will be imputed as the AE start date.

Imputation for Resolution Date:

- If day is missing but month and year are non-missing (??-MMM-YYYY),
  - Min (last day of the month, data cutoff date, death date).
- If day and month are missing (??-???-YYYY),
  - Min (31-DEC-YYYY, data cutoff date, death date).
- If date is completely missing (e.g., AE is ongoing),
  - Min (data cutoff date, death date, last dose date+30 days).

  In such case, the imputed end date will not show in the listings. Instead "Ongoing", "Continuing" or similar wording will be used.

Imputation for Onset Date:

- If day is missing but month and year are non-missing (??-MMM-YYYY):
  - If year and month are the same as year and month of first dose date:
    - If resolution date or imputed resolution date ≥ first dose date, impute as first dose date.
    - If resolution date or imputed resolution date < first dose date, impute as: Max (the first day of the month, informed consent date).
  - If year is the same as year of first dose date and month is **after** month of first dose date, impute as the first date of the month.
  - If year is the same as year of first dose date and month is **before** month of first dose date, impute as:

    Max (the first day of the month, informed consent date).
  - If year is **after** year of first dose date, impute as the first date of the month

    If year is **before** year of first dose date, impute as:

    Max (the first day of the month, informed consent date).
- If day and month are missing and year is non-missing (??-???-YYYY), impute as follows:
  - If year is the same as year of first dose date:
    - If resolution date or imputed resolution date is on or after first dose date, impute as first dose date.

      If resolution date (or imputed resolution date) is prior to first dose date, impute as: Max (the first day of the year, informed consent date).
  - If year is **after** year of first dose date, impute as 01-JAN-YYYY.
  - If year is **before** year of first dose date, impute as:

Max (01-JAN-YYYY, informed consent date).

- If date is completely missing:
  - If resolution date (or imputed resolution date) is on or after first dose date, impute as first dose date.
  - If resolution date (or imputed resolution date) is prior to first dose date, impute as informed consent date.

7.1.1.4 *Conventions for Missing Anti-Cancer Therapies and Prior/Concomitant Medication Dates*

Prior Anti-Cancer Therapies Start Date

- If day is missing but month and year are non-missing (YYYY-MM-UU), impute as YYYY-MM-01
- If day and month are missing (YYYY-UU-UU), impute as YYYY-01-01
- If month is missing (YYYY-UU-DD), but day and year are non-missing, ignore the day and impute as YYYY-01-01

If after applying above rules, any prior treatment start dates are **before** diagnosis date, impute as diagnosis date.

Prior Anti-Cancer Therapies Stop Date

- If day is missing but month and year are non-missing (YYYY-MM-UU), impute as the earliest of:
  - The last day of the month
  - Randomization date
- If day and month are missing (YYYY-UU-UU), impute as the earliest of:
  - YYYY-12-31
  - Randomization date
- If date is entirely missing, impute as Randomization date.

If after applying above rules, any prior treatment stop dates are prior to corresponding prior treatment start date, impute as start date.

Prior/Concomitant Medication Start Date

- If day is missing but month and year are non-missing (YYYY-MM-UU), impute as YYYY-MM- 01.
- If day and month are missing (YYYY-UU-UU), impute as YYYY-01-01.
- If month is missing (YYYY-UU-DD), but day and year are non-missing, ignore the day and impute as YYYY-01-01.
- If date is entirely missing, impute as the first does date.

Prior/Concomitant Medication End Date

- If day is missing but month and year are non-missing (YYYY-MM-UU), impute as the last day of the month
- If day and month are missing (YYYY-UU-UU), impute as YYYY-12-31.
- If date is entirely missing and not ongoing, impute as the last date of study treatment.

If imputed date is before start date set as start date. Do not impute for ongoing records.

Subsequent Anti-Cancer Treatment Start Date

- If the month and year are present but the day is missing:
    - If the onset month and year are the same as the month and year of the last dose of study drug, the day of last dose + 1 will be imputed.
    - If the onset month and year are not the same as the month and year of the last dose of study drug, the first day of the month is imputed.
- If only a year is present:
    - If the onset year is the same as the year of the last dose of study drug, the date of last dose + 1 will be imputed.
    - If the onset year is not the same as the year of the last dose of study drug, the first day of the year is imputed.
- If no components of the onset date are present, the date of the last dose of study drug+ 1 will be imputed.

7.2 Analysis Sets

7.2.1 Full Analysis Set

The full analysis set is based on the intent-to-treat (ITT) principle and includes all patients randomized to each regimen regardless of whether they are ALK+ by an FDA approved test (Vysis ALK Break Apart FISH Probe Kit, Ventana ALK (D5F3) CDx Assay, Foundation Medicine's FoundationOneCDx) or a local test other than FISH and immunohistochemistry, or whether they receive study drug or adhere to the assigned dose. The primary analyses of efficacy will be based on the full analysis set according to the treatment they were randomized to receive, regardless of errors.

7.2.2 Safety Analysis Set

The safety analysis set for each regimen includes all patients receiving at least 1 dose of study drug. Safety-related endpoints will be analyzed using the safety analysis set according to the actual treatment received.

7.2.3 Per-Protocol Analysis Set

The per-protocol analysis set will exclude all patients in the safety analysis set who do not meet key entry criteria, have no measurable disease at baseline, or have no adequate postbaseline response assessment by BIRC unless the reason is death or early discontinuation due to disease progression per investigator. All decisions to exclude patients and the reasons will be documented prior to database lock.

To specify, the per-protocol population will fulfill the following criteria:

- Have histologically or cytologically confirmed stage IIIB (locally advanced or recurrent) or stage IV NSCLC
- Have documentation of ALK rearrangement
- Had progressive disease while on crizotinib, as assessed by the investigator or treating physician. (Note: crizotinib does not need to be the last therapy a patient received. The patient may have received chemotherapy as his/her last systemic anticancer therapy.)
- Have had no other ALK inhibitor other than crizotinib.
- Treatment with crizotinib for at least 4 weeks before progression.
- Have had no more than 2 prior regimens of systemic anticancer therapy (other than crizotinib) in the locally advanced or metastatic setting*.

  Note: a systemic anticancer therapy regimen will be counted if it is administered for at least 1 complete cycle. A new anticancer agent used as maintenance therapy will be counted as a new regimen. Neoadjuvant or adjuvant systemic anticancer therapy will be counted as a prior regimen if disease progression/recurrence occurred within 12 months upon completion of this neoadjuvant or adjuvant therapy.

  *Systemic therapy followed by maintenance therapy will be considered as one regimen if the maintenance therapy consists of a drug or drugs that were used in the regimen that immediately preceded maintenance.
- Have at least one measurable target lesion per RECIST v1.1 according to the investigator
- At least two adequate post-baseline radiographic response assessments unless the reason for no post-baseline radiographic response assessment is one of the following:
  - Death
  - Discontinuation due to documented disease progression per RECIST v1.1
  - Discontinuation due to AE

7.2.4 PRO Analysis Set

The PRO analysis set includes all patients with baseline and at least 1 post-baseline PRO measurement in the full analysis set.

7.3 Disposition of Subjects

Study information, including date first subject signed ICF, date of first subject randomized, date of last subject's last visit/contact, date of last subject's last procedure for collection of data for primary endpoint, MedDRA version, WHO Drug version, and SAS version will be presented in listings.

Screen failures will be summarized and listed by the reasons for failure to be enrolled in the study.

A table and listing of significant protocol deviations will be presented. The criteria may include, but are not limited to:

- Major inclusion/exclusion violations
- Patients meeting criteria of dose discontinuation due to toxicity but continued treatment
- Missed pregnancy testing
- Taken excluded concomitant medications

Patient disposition will be summarized by treatment arm for the full analysis set and will include the following:

- Patients randomized
- Patient treatment status as ongoing, discontinued, or never treated
- Patients who have discontinued from treatment will be further categorized by primary reason for discontinuation from treatment
    - Adverse event
    - Protocol deviation
    - Lost to follow-up
    - Withdrawal by subject
    - Investigator discretion
    - Intolerable toxicity
    - Clinical progression without radiological PD per RECIST
    - Progressive disease per RECIST
    - Other

- Patients who have discontinued from study will be further categorized by reason for discontinuation from study
    - Death
    - Adverse event
    - Protocol deviation
    - Lost to Follow-up
    - Withdrawal by subject
    - Study terminated by sponsor
    - Pregnancy
    - Other
- Total follow-up time in months, defined as (last contact date – randomization date + 1)/30.4375 days; in categories of <6, 6-12, 12-24, and ≥24 months as appropriate.

7.4 Demographic and Other Baseline Characteristics

Demographics and baseline characteristics will be summarized by treatment arm using descriptive statistics for full analysis set. The variables include but are not limited to age, gender, race, ethnicity, height, weight, BMI, ECOG, ALK+ (FDA approved tests or other), smoking history, disease stage at initial diagnosis (stage I/II vs III vs IV) and study entry (stage III vs IV), time since initial diagnosis, brain metastases at baseline by investigator assessment (Yes vs no), best response to prior crizotinib (CR/PR vs other), prior chemotherapy for systemic disease (other than crizotinib) yes or no, and number of prior chemotherapies for systemic disease (other than crizotinib) (0, 1 or 2).

The two stratification factors, brain metastases at baseline and best response to prior crizotinib will be tabulated by stratum using IxRS and eCRF data separately. The concordance rate between IxRS and eCRF will be calculated.

Continuous variables will be summarized by mean, median, standard deviation, and range; categorical variables will be summarized by count and percentage.

7.5 Medical History and Concurrent Medical Conditions

Medical history will be coded using the most recent version of MedDRA available at the time of coding. General medical history will be summarized for each treatment group by System Organ Class (SOC) and Preferred Term (PT) using counts and percentages. By-patient listing will also be presented for medical history for the full analysis set. Ongoing conditions, including related to the underlying NSCLC and symptoms entered in eCRF will be flagged in the listing.

7.6 Medication History and Concomitant Medications

7.6.1 Prior Therapy

Prior anticancer medications will be coded by preferred term using the World Health Organization (WHO) Drug Dictionary. Summary tables will show the number and percentage of patients who had received prior radiotherapy, prior radiotherapy to the brain, time from last prior radiotherapy to the brain to Day 1, prior systemic therapy for NSCLC, the most recent prior systemic therapy (crizotinib or other), best response to most recent prior therapy (CR/PR vs other), best response to crizotinib (CR/PR vs other), number of prior chemotherapies for systemic disease (other than crizotinib) (0, 1 or 2), time from last dose of prior systemic anti-cancer therapy to Day 1, and prior anticancer surgery. The full analysis set will be used for analysis of prior therapy.

7.6.2 Concomitant Medications

Concomitant medications are medications ongoing at the time of the first dose of study drug or medications that started after first dose and within 30 days of the last dose of study drug. Concomitant medications will be coded by preferred term using the WHO Drug Dictionary. The number and percentage of patients taking concomitant medications from the first dose of study drug through the end of treatment will be tabulated by Anatomical Therapeutic Chemical (ATC) classification pharmacological subgroup and WHO drug generic term for each treatment group. By-patient listing will also be presented for concomitant medications.

Concomitant procedures will not be coded but will be presented in a data listing. The safety analysis set will be used for analysis of concomitant medications.

7.6.3 Subsequent Anticancer Therapy

The number and percentage of patients receiving each type systemic (ALK inhibitors, immunotherapy, chemotherapy, investigational drugs) and line of systemic anticancer therapy as well as radiotherapy or anticancer surgery subsequent to the randomized drug (1^{st}, 2^{nd}, etc. following brigatinib or alectinib), as well as the best response to these systemic therapies will be tabulated by treatment arm for the full analysis set.

Subsequent therapies will be also presented in a by-patient listing by arm.

7.7 Study Drug Exposure and Compliance

Study drug exposure will be summarized using the following measures by actual treatment arm:

- Time on study treatment (days and months)
- Total cumulative dose administered (mg)
- Dose intensity (mg/day)
- Relative dose intensity (%)

- Dose interruption of at least 3 days
    - Number of patients with at least one occurrence
    - Number of patients who returned to target dose after interruption
    - Duration of longest dose interruption ≥3 days
- Dose reduction of at least 3 days
    - Number of patients with at least one occurrence
    - Number of patients who returned to target dose after reduction
    - Duration of longest dose reduction ≥3 days

Time on treatment will be defined as the time interval from the first dose date to the last dosing date and computed with the following formula: Time (days) on treatment= last non-zero dose date – first dose date + 1

Dose intensity will be calculated with the following formula: Dose intensity (mg/day) = Total cumulative dose (mg) / Time on treatment (days). Relative dose intensity will be defined as the proportion of the planned dose received by patients:

100%*(Total cumulative dose administered / Total dose planned.

In the brigatinib arm, daily planned dose will be 90 mg in the 7-day lead-in period and 180 mg from day 8 onward; in the alectinib arm, daily planned dose will be 1,200 mg (600 mg BID).

Dose modifications will be summarized by dose interruption and dose reduction. A patient will be identified as having dose interruption if this patient had no exposure to study drug (0 mg) for at least 3 consecutive days. A patient will be identified as having dose reduction if this patient had a period of reduced dosage of at least 3 consecutive days, as long as the dose received was less than the target dose but greater than 0 mg on some of the days in this period. Periods of time in which a patient alternates between reduced dosing and dose interruptions will be handled in the following manner:

- The entire period between the last receipt of the target dose and either resumption at the target dose or discontinuation of treatment will be considered a single dose reduction period.
- Any period of 3 or more days with no receipt of study drug within that dose reduction period will also be treated as a dose interruption.

7.8 Efficacy Analysis

7.8.1 Primary Efficacy Endpoint

7.8.1.1 *Data Handling Rules for the Primary Analysis of the Primary Endpoint*

The primary endpoint, PFS assessed by BIRC, is defined as the time interval from the date of randomization until the first date at which disease progression is objectively documented via

RECIST v1.1, or death due to any cause, whichever occurs first. It will be censored for subjects who have not had a PFS event at the data cutoff. Follow-up for PFS will also be censored in the primary efficacy analyses upon documentation of certain other events (such as initiation of new anticancer treatment prior to a PFS event).

The detailed scheme of progression and censoring for the primary analysis of PFS is specified in Table 7.1.

Table 7.1 Scheme of Progression and Censoring for the Primary Analysis of PFS

Rule	Situation	Analysis Date	Outcome
1	No baseline and/or no post-baseline assessment, no new anticancer treatment*, and no death	Randomization	Censored
2	No progression or death, and new anticancer treatment* is not initiated	Date of last adequate progression-free radiographic assessment	Censored
3	New anticancer treatment* is initiated prior to PD or death	Date of last adequate progression-free radiographic assessment before new anticancer treatment*	Censored
4	PD or death after ≤1 consecutively missed radiographic assessment	Date of documented PD or death, whichever is earlier	Progressed
5	PD or death after ≥2 consecutively missed radiographic assessments	Date of last adequate progression-free radiographic assessment prior to the missed interval	Censored

* Any radiotherapy, systemic anticancer therapy, and any surgical removal of remaining tumor intended for complete resection.

PFS will be calculated for each patient in months and will be defined as (PFS Analysis date – randomization date + 1)/30.4375. The reason for an event or censoring will be tabulated alongside the Kaplan-Meier (KM) estimates for all time-to-event endpoints. The reason for a PFS event will be either PD per RECIST criteria or death due to any cause; the reason for censoring will be one of the following: new anticancer treatment started prior to documented PD per RECIST criteria, PFS event observed after ≥2 consecutively missed radiographic assessment, no valid post-baseline assessment or missing/incomplete baseline assessment, and no progression or death.

7.8.1.2 *Evaluation of the Primary Efficacy Endpoint*

The primary analysis of the primary endpoint will be performed using a 2-sided stratified log-rank test using the stratification factors (at randomization): presence of intracranial CNS metastases at baseline [Yes vs No], and best prior response to crizotinib therapy as assessed by the investigator [CR/PR vs any other response/status unknown] to compare the BIRC-assessed PFS of patients randomized to brigatinib with the BIRC-assessed PFS of patients randomized to alectinib. The overall (2-sided) type I error rate will be controlled at 0.05. The primary analysis will be based on the full analysis set. Median PFS and 95% CI will be estimated for each

treatment arm using the KM method. The HR and its 95% CI will be estimated using the stratified Cox regression model with the same stratification factors.

7.8.1.3 *Interim Analysis of the Primary Endpoint*

Refer to Section **7.12 Interim Analysis** for greater detail. One interim analysis is planned for this study after approximately 115 PFS events (70% of the total 164 PFS events) have been observed approximately 22 months after the first patient is randomized. An O'Brien Fleming (OBF) Lan DeMets (DeMets and Lan, 1994) alpha spending function will be used to control the overall alpha level at 0.05 (2-sided). A gamma spending function (Hwang, Shih and De Cani, 1990) with parameter -6 will be used for the non-binding futility stopping boundary to control the overall power at 90%. If exactly 115 PFS events occurred at the time of IA of PFS, the 2-sided alpha to be spent would be 0.0149 and type II error to be spent would be 0.017. If the p-value doesn't meet the prespecified stopping rule for efficacy or futility at the IA then the primary analysis will be conducted after 164 PFS events are observed. The efficacy and futility stopping boundaries used in the analysis will be adjusted based on the actual number of events observed at each analysis using the O'Brien-Fleming Lan-DeMets (DeMets and Lan, 1994) alpha spending function and gamma spending function (Hwang, Shih and De Cani, 1990).

The first occurrence of a successful evaluation of the primary endpoint, defined as surpassing the critical value at a pre-planned analysis, will complete the inferential statistical evaluation of the primary endpoint for this study. All subsequent analyses after the conclusion of the primary analysis will be non-inferential, even for analyses conducted as specified in the protocol for subsequent analyses of the primary endpoint.

7.8.1.4 *Sensitivity Analyses of the Primary Endpoint*

Sensitivity analyses of the primary endpoint of PFS per BIRC will also be performed in the following large categories:

- Alternative analysis populations
  - Per-protocol analysis set
  - ALK+ confirmed by an FDA approved test (Vysis ALK Break-Apart FISH Probe Kit, Ventana ALK (D5F3) CDx Assay; Foundation One CDx)
- Alternative censoring rules. Table 7.2 only describes the distinct rules for each sensitivity analysis while omitting those otherwise identical rules.

Table 7.2 Scheme of Progression and Censoring for the Sensitivity Analyses of PFS

Rule	Situation	Analysis Date	Outcome
S1: EMA Guideline			
3	New anticancer treatment* is initiated prior to PD or death	Ignored, neither a reason for event nor censoring per se	NA

Rule	Situation	Analysis Date	Outcome
5	PD or death after ≥2 consecutively missed radiographic assessments	Date of documented PD or death, whichever is earlier	Progressed
S2: Table 7.1 + treatment discontinuation			
6	Treatment discontinuation for undocumented PD (clinical progressive disease) without radiographic assessment	Date of last adequate progression-free radiographic assessment prior to treatment discontinuation	Censored
S3: S1 + treatment discontinuation			
7	Treatment discontinuation for undocumented PD (clinical progressive disease) without radiographic assessment	Date of last adequate progression-free radiographic assessment prior to treatment discontinuation	Progressed
S4: S3 + subsequent therapy			
3	New anticancer treatment* is initiated prior to PD or death	Date of start of new anticancer therapy*	Progressed

* Any radiotherapy, systemic anticancer therapy, and any surgical removal of remaining tumor intended for complete resection.

Subgroup analyses will be performed by baseline prognostic factors to compare the PFS within subgroups (Table 7.3). Unstratified Cox regression models will be used to compute the HR and 95% CI, due to limited number of events within subgroups. Forest plots will be created to facilitate visual comparison.

Table 7.3 Key Subgroups

Variable	Definition of subgroups
Age	< 65 years; ≥ 65 years
Sex	Male; female
Race	Asian; non-Asian
Cancer stage at study entry	IV; < IV
Baseline ECOG	0-1; 2
Smoking history	Never smoker; former or current smoker
Brain metastases at baseline	Yes; no
Best response to prior crizotinib	CR or PR; Any other response or status unknown
Prior chemotherapy for systemic disease (other than crizotinib)	Yes; no
• Number of prior chemotherapies for systemic disease (other than crizotinib)	0; 1 or 2

7.8.2 Secondary Efficacy Endpoints

OS is only key secondary endpoint. Formal statistical tests will be performed only when PFS per BIRC is statistically significant.

The other secondary efficacy endpoints are defined in Subsections 7.8.2.2 -7.8.2.9. Unless otherwise stated, secondary efficacy endpoints of response will use BIRC assessments with sensitivity analyses performed per investigator assessment.

7.8.2.1 *Overall Survival (OS)*

OS is defined as the time interval from randomization until death due to any cause in full analysis set. It will be censored on the date of last contact for those patients who are alive at the data cutoff date. The primary analysis of OS will be performed using a 2-sided stratified log-rank test to compare the OS of patients randomized to brigatinib with the OS of patients randomized to alectinib, following a significant BIRC-assessed PFS. OS could potentially be tested on a maximum of three occasions, ie, concurrent with the IA and FA of BIRC-assessed PFS, as well as when an estimated 156 deaths occur, approximately 5 years after the first patient is randomized, assuming the median OS of 34 and 29 months for the brigatinib and alectinib arms, respectively (HR=0.853). The statistical significance of OS will be determined with the O'Brien Fleming Lan DeMets alpha spending function based on the information fraction available at each analysis time point, relative to an estimated total of 156 deaths. This study is not powered to demonstrate a statistically significant difference in OS with a total of approximately 246 patients. Additional details on the testing of OS can be found in **Section 7.12**.

Median OS and 95% CI will be estimated by treatment arm using the KM method. The HR and its 95% CI will be estimated using the stratified Cox regression model. In addition, to account for the confounding effects of subsequent therapy, alternative methods such as the inverse probability of censoring weighted (IPCW [Robins and Finkelstein, 2000]) and marginal structural model (MSM [Robins, Hernan et al. 2000]) may be used as sensitivity analyses for OS.

In IPCW and MSM, covariates that would affect disease progression and subsequent therapy will be individually evaluated and included alongside OS in the model in deriving the weights adjusting for both time-fixed and time-varying confounding effects associated with the alternative therapies. Covariates will be dropped from the weighting calculation and final OS model if >5% missing data. The adjusted K-M curves will also be presented along the HR, 95% CI, and adjusted p-value based on IPCW and MSM methods. SAS PROC PHREG procedure with counting process type of data input, taking multiple observations per subject, will be used as the final Cox model for OS.

Subgroup analyses of OS will be presented to assess consistency of OS in key subgroups (Asian vs. non-Asian; brain metastases at baseline, best response to prior crizotinib). The unstratified Cox regression model will be used to compute the HRs, due to limited number of events within subgroups. Forest plots will be created to facilitate visual comparison.

7.8.2.2 *PFS, as assessed by the investigator per RECIST v1.1*

Analogous to PFS assessed by BIRC, the investigator-assessed PFS is defined as the time interval from the date of randomization until the first date at which disease progression is objectively documented via RECIST v1.1 per investigator interpretation, or death due to any cause, whichever occurs first. It will follow the same censoring rules and analysis timing as that for the BIRC-assessed PFS.

7.8.2.3 *Confirmed ORR*

Confirmed ORR is defined as the proportion of the patients who have achieved a confirmed response of CR or PR using RECIST v1.1 after the initiation of study treatment. Confirmed responses are responses that persist on repeat imaging ≥4 weeks (allowing a minus 3-day window) after initial response. The primary analysis of ORR will be analyzed based on the full analysis set; a sensitivity analysis of ORR will be based on the per-protocol analysis set. Patients without at least one post-baseline assessment will be considered non-responders. The exact 2-sided 95% binomial confidence intervals will be calculated using the Clopper-Pearson method. The difference in ORR between the two treatment arms will be calculated with a two-sided 95% CI. ORR per BIRC and per investigator will be analyzed separately.

Best target lesion responses will be graphed in waterfall plots to illustrate the best percentage change in target lesions from baseline by arm.

7.8.2.4 *Duration of Response (DOR)*

Duration of response (DOR) is defined as the time interval from the time that the measurement criteria are first met for CR or PR (whichever is first recorded) until the first date that the PD is

objectively documented or death in patients with confirmed response. For the responders who have not progressed or died, they will be censored at the last tumor assessment date prior to receiving subsequent anticancer therapy. DOR will be analyzed among the confirmed responders only. DOR will be estimated for each treatment arm using the KM method. DOR per BIRC and per investigator will be analyzed separately.

7.8.2.5 *Time to Response (TTR)*

Time to response is defined as the time interval from randomization until the initial observation of CR or PR. Time to response will be summarized by arm using descriptive statistics among the confirmed responders only. TTR per BIRC and per investigator will be analyzed separately.

7.8.2.6 *Confirmed Intracranial Objective Response Rate*

Confirmed iORR, as assessed by the BIRC, is defined as the proportion of the patients who have achieved a confirmed response of CR or PR in the CNS per a modified RECIST v1.1 after the initiation of study treatment. Confirmed responses are responses that persist on repeat imaging ≥4 weeks (allowing a minus 3-day window) after initial response. The exact 2-sided 95% binomial confidence intervals will be calculated using the Clopper-Pearson method. The difference in iORR between the two treatment arms will be calculate with a two-sided 95% CI.

Analyses of this endpoint will be performed in the following populations:

- Patients with brain metastases at enrollment per BIRC
- Patients with measurable brain metastases at enrollment per BIRC defined as having at least one target lesion in the brain
- Patients with only non-measurable brain metastases at enrollment per BIRC

7.8.2.7 *Intracranial Duration of Response (iDOR)*

iDOR is defined as the time interval from the time that the measurement criteria are first met for CR or PR in the CNS (whichever is first recorded) until the first date that the PD in the CNS (iPD) is objectively documented or death in patients with confirmed intracranial response. iDOR will be estimated by treatment arm using the KM method among patients with an intracranial objective response with respect to the following populations:

- Patients with brain metastases at enrollment per BIRC
- Patients with measurable brain metastases at enrollment per BIRC defined as having at least one target lesion in the brain
- Patients with only non-measurable brain metastases at enrollment per BIRC

7.8.2.8 *Time to iPD without prior systemic progression*

Time to iPD without prior systemic progression, as assessed by the BIRC, is defined as the time interval from randomization until the first date at which intracranial disease progression (iPD) is

objectively documented via a modification of RECIST v1.1 without prior systemic progression. iPD is defined as progression due to newly developed iCNS lesions and/or progression of preexisting baseline iCNS lesions. All patients in full analysis set will be included in the analysis regardless of their baseline status of brain metastases.

The time to a cause-specific event is defined as the time from randomization to the first cause-specific event, ie, the earliest of iPD without prior systemic PD, systemic PD without prior iPD, or death without prior iPD or systemic PD. To account for the competing risks among three types of events, a stratified log-rank test will be computed on the basis of the cause-specific hazard functions. The cause-specific HRs and 95% CIs will be estimated in a stratified Cox regression model. Patients experiencing one cause-specific event will be censored simultaneously for the other cause-specific events in the context of competing risks analysis. Should more than one type of the competing events occur on the same day, iPD will take precedence over systemic (extracranial) PD and systemic (extracranial) PD over death.

The probabilities of iPD, systemic PD, and death by treatment arm with the 95% CIs will be estimated and plotted using the cumulative incidence functions (CIF). A comparison of the incidence of progression in the CNS as the first site of disease progression between treatment arms, eg, at 6 and 12 months or other meaningful landmarks after randomization, will be based on the CIF estimates.

7.8.2.9 *HRQoL assessed with the global health status/quality of life and other function and symptom domains from EORTC QLQ-C30 (v3.0) and EORTC QLQ-LC13.*

Analyses of PRO data will be performed using the PRO analysis set, which will be defined as patients with baseline and at least 1 post-baseline measurement in the full analysis set. The global health status/QoL and functions from EORTC QLQ-C30 and symptoms measured by LC13 will be of special interest.

The descriptive statistics of the actual value and change from baseline of the EORTC QLQ-C30 and QLQ-LC13 scores will be summarized by treatment group over time. The change from baseline scores will be analyzed using linear mixed-effects models to compare the 2 treatment arms. Covariates will include treatment group, visit, the interaction between treatment group and visit, baseline score and stratification factors (brain metastases at baseline and best response to prior crizotinib). The estimated means with 95% CIs will be provided by treatment groups at each time point and overall across different time points. The mean differences between treatment groups along with 95% CIs and p-values will also be presented at each time point and overall. *All items for a PRO instrument must come from the same visit date. If a PRO instrument has >50% missing data at the baseline visit, the screening values will be used instead for the baseline values (if screening instrument has <50% missing data). For example, if a patient answered only 5 questions on EORTC QLQ-C30 at the baseline visit, but answered all 30 questions at the screening visit, all 30 questions from the screening visit will be identified as the baseline item scores for EORTC QLQ-C30.*

Time to deterioration, time to improvement and duration of improvement will be analyzed for each of EORTC QLQ-C30 and LC13 scores and the composite score of cough, dyspnea, and

pain in the chest from LC13. For this composite score, the deterioration (or improvement) will be defined as any of cough, dyspnea, and pain in the chest scores that had ≥ 10 points deterioration (or improvement) from baseline.

Time to deterioration will be defined as time from the date of randomization to the earliest date, at which, the patient's score had a ≥ 10 points deterioration from baseline.

Time to improvement will be defined as time from the date of randomization to the earliest date, at which, the patient's score had a ≥ 10 points improvement from baseline.

Duration of improvement will be defined as time from the date of first improvement to the date of first occurrence of deterioration after the improvement.

Patients without an event will be censored at the date of last QoL measurement. A 2-sided stratified log-rank test will be used to compare the treatment groups. In addition, an unadjusted, stratified Cox model will be used to estimate the hazard ratio and its 95% CI. The Kaplan-Meier survival curves will also be provided for each treatment group.

The number and percentage of patients with improved and worsened EORTC QLQ-C30 and LC13 scores will also be summarized by treatment group over time.

In addition to the estimates at each time point, the overall and average percentage of improvement and percentage of worsening and corresponding estimates across all time points will be provided. Overall will be defined as patients with at least one event across all time points. Average will be defined based on the average change from baseline across all time points.

For percentage of improvement and percentage of worsening, odds ratios with 95% CI and p-values from a Cochran-Mantel-Haenszel tests stratified by the stratification factors (brain metastases at baseline and best response to prior crizotinib) will be presented to compare two treatment groups. The risk ratios and risk difference with corresponding 95% CIs will also be presented.

For the percentage calculation, the number of patients in the ITT-PRO population will be used as in the denominator, which assumes patients with missing data will be counted as stable (not improved or worsened). Details of scoring and initial handling of missing item scores are included in the EORTC QLQ-C30 and QLQ-LC13 scoring guidelines. Further investigation on patterns of missing data and subsequent sensitivity analysis may be conducted. EORTC QLQ-C30 and QLQ-LC13 questionnaires compliance will be summarized by treatment arm at each time point (and overall).

Details of scoring and initial handling of missing item scores are included in the EORTC QLQ-C30 and QLQ-LC13 scoring guidelines. Further investigation on patterns of missing data and subsequent sensitivity analysis may be conducted. EORTC QLQ-C30 and QLQ-LC13 questionnaires compliance will be summarized by treatment arm at each time point and overall.

7.8.3 Exploratory Endpoints

7.8.3.1 *CNS efficacy outcomes as assessed by the BIRC per RANO-BM criteria (iORR, iDOR, and time to iPD).*

Statistical methods same as those described in subsections 7.8.2.6, 7.8.2.7, and 7.8.2.8 will be used with substitution of BIRC assessments per RANO-BM criteria.

7.8.3.2 *Molecular determinants of efficacy and safety with brigatinib and alectinib as uncovered by genetic alterations noted on tumor tissue DNA or plasma samples of circulating tumor DNA.*

The assessments of relationships between molecular genetic information from tumor and plasma samples with efficacy and safety are out the scope of the Clinical Study Report and may be explored in a standalone document.

7.8.3.3 *HRQoL measured by EuroQoL–5 Dimensions–5 Levels (EQ-5D-5L) questionnaires.*

EQ-5D scores and change from baseline will be summarized in descriptive statistics by arm and over time. In addition, there is a general question about the overall health (EQ-VAS) with range 0-100 (larger number indicating better health). The EQ-VAS values and change from baseline will be summarized by visit.

EQ-5D-5L questionnaires compliance will be summarized by treatment arm at each time point (and overall).

7.8.3.4 *Evaluation of healthcare resource utilization.*

Health resource utilization data will be summarized in descriptive statistics of medical encounters (length of stay, inpatient/outpatient admissions, and reasons), number of missing days from work or other activities by patient and caregiver by treatment arm.

7.8.3.5 *Items from the EORTC Quality of Life Brain Cancer Module (QLQ-BN20) used to assess morbidity related to CNS symptoms.*

To reduce the clinical site and respondent burden, 5 items from the brain cancer module QLQ BN20 will be used to assess morbidity related to CNS symptoms, including headaches (item 4, Q34), coordination (item 15, Q45), and communication deficit (items 11-13, Q41-43). The descriptive statistics of the actual value and change from baseline of these EORTC QLQ-BN20 scores will be summarized by treatment group over time.

In addition, the change from baseline of QLQ-BN20 scores will be analyzed using linear mixed models. Time to deterioration, time to improvement, duration of improvement and the number and percentage of patients with improved and worsened QLQ-BN20 scores will also be summarized by treatment group, following the details provided in section 7.8.2.8 for the analysis of EORTC QLQ-C30 and LC13 scores.

EORTC QLQ-BN20 questionnaires compliance will be summarized by treatment arm at each time point (and overall).

7.9 Pharmacokinetic/Pharmacodynamic Analysis

7.9.1 Pharmacokinetic Analysis

Plasma concentrations of brigatinib will be listed and summarized by time point using descriptive statistics (eg, mean, sd, geometric mean, %CV, median, and range). PK data collected in this study will contribute to population PK and exposure/response (safety and efficacy) analyses. These analyses may include data from other brigatinib clinical studies. The analysis plans for these analyses will be developed and the results reported separately.

7.9.2 Pharmacodynamic Analysis

Not applicable.

7.10 Other Outcomes

Not applicable.

7.11 Safety Analysis

Safety assessments will include evaluations of incidence of AEs, severity and type of AEs, physical and laboratory examinations, vital signs, and electrocardiograms (ECGs) according to the Schedule of Events. AE verbatim terms will be coded using MedDRA Dictionary v21.0 or higher with the severity graded by the NCI CTCAE v4.03. Subjects will remain on treatment until they meet one or more criteria for withdrawal.

All safety analyses will be performed using the safety analysis set, which includes all subjects who have received at least one dose of randomized study treatment. Unless otherwise specified, all safety analyses will be reported according to actual study treatment received.

7.11.1 Adverse Events

All adverse events (AEs) starting/worsening on or after the first dose of study treatment and no later than 30 days after the last dose date will be considered as treatment-emergent adverse events (TEAEs).

All AEs entered in the clinical database will be listed in by-subject listings or available for review in appropriate datasets.

The incidence rates of TEAEs, as well as the frequency of occurrence of overall toxicity categorized by maximum toxicity grades (severity), will be tabulated by MedDRA primary system organ class (SOC) and preferred terms (PT). Patients with the same AE more than once will have that event counted only once within each SOC and once within each PT. In addition, TEAEs will be summarized by causal relationship to study treatment (in the investigator's opinion) and action taken on study treatment, including dose modifications, interruptions and

discontinuation. Grade 3 or higher TEAEs and Grade 3 or higher drug-related TEAEs (with G5 reported separately) will be summarized, both overall and restricted to those drug-related. Serious treatment-emergent AEs (SAEs), both overall and restricted to those drug-related, will also be summarized. Most commonly reported (≥10% of patients in any arm or ≥5% absolute difference between arms) TEAEs and SAEs will be tabulated by PT. Patients with the same AE more than once will have that event counted only once within each PT.

All deaths occurring during a period of exposure to randomized therapy or within a period of up to 30 days following discontinuation from randomized therapy will be summarized by AEs leading to death and causal relationship to study treatment. All deaths occurring later than 30 days but resulting from treatment related adverse event(s) will be summarized by AEs leading to death and causal relationship to study treatment.

Table 7.4 Summary of AE Analysis Tables

TEAEs	by SOC & PT, by PT, by SOC & PT & Max Grade
Drug-related TEAEs	by SOC & PT, by PT, by SOC & PT & Max Grade
Grade 3 or higher TEAEs (G5 separately)	by SOC & PT
Grade 3 or higher drug-related TEAEs (G5 separately)	by SOC & PT
SAEs	by SOC & PT, by PT, by SOC & PT & Max Grade
Drug-related SAEs	by SOC & PT, by PT, by SOC & PT & Max Grade
Grade 3 or higher SAEs (G5 separately)	by SOC & PT
Grade 3 or higher drug-related SAEs (G5 separately)	by SOC & PT
TEAEs resulting in death	by SOC & PT
TEAEs resulting in treatment discontinuation	by SOC & PT
TEAEs resulting in treatment modification	by SOC & PT
Most common (≥10% in any arm or ≥5% absolute difference between arms) TEAEs	by PT
Most common (≥2% in any arm) SAEs	by PT
Most common (≥2% in any arm) Gr3+ TEAEs	by PT
Most Frequent (≥5% in any arm) Non-Serious TEAEs	by SOC & PT

. AEs in special categories will also be summarized in terms of number of subjects with at least one event in the category and by number of subjects with at least one event for each constituent preferred term. The special categories to be analyzed will be aligned with the risk profile of brigatinib and alectinib.

Exploratory analyses of the incidence of treatment-emergent AEs may also be summarized in subgroups based on selected demographic and baseline disease characteristics, randomization stratification factors, and ALK mutation.

7.11.2 Clinical Laboratory Evaluations

For the purposes of summarization in the tables and listings, all laboratory values will be converted to standardized units. If a lab value is reported using a non-numeric qualifier (eg, less than (<) a certain value, or greater than (>) a certain value), the given numeric value will be used in the summary statistics, ignoring the non-numeric qualifier. If a subject has repeated laboratory values for a given time point, the value from the last evaluation will be used.

The parameters to be analyzed are as follows:

- Hematology: hematocrit, hemoglobin, leukocytes with differential, neutrophils (absolute neutrophil count [ANC]), platelet count.
- Serum chemistry: albumin, alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), amylase, bilirubin (total), blood urea nitrogen (BUN), calcium, creatine kinase, creatinine, chloride, glucose, insulin, lactate dehydrogenase (LDH), lipase, magnesium, phosphate, potassium, sodium.

Summary statistics (mean, standard deviation, median, and range) will be calculated for the raw data and for their changes from baseline at each time point of assessment and for the changes from baseline to the last value. Individual values outside the normal ranges will be identified (by "H" for high and "L" for low) in the data listings displaying the absolute values for each subject.

Shift tables from baseline to worst value on study, i.e., worst grade by NCI CTCAE v4.03 or categorized as low, normal, high if no grades are defined will be provided for the following laboratory parameters: hemoglobin, ANC, platelets, albumin, ALT, AST, amylase, total bilirubin, calcium, creatine kinase, creatinine, glucose, insulin, lipase, magnesium, phosphate, potassium, and sodium.

Mean laboratory values with standard deviations over time for key lab parameters will be plotted by arm, including but not limited to ANC, platelets, and liver function tests (ALT, AST, and total bilirubin).

7.11.3 Vital Signs

The actual values of vital sign parameters including temperature, blood pressures, heart rate, pulse rate, respiratory rate, and body weight, will be summarized over time for each treatment arm, including change from baseline. A by-patient listing will also be presented.

7.11.4 12-Lead ECGs

ECG (QT, PR, QRS, and RR) intervals, and ventricular rate will be summarized at each scheduled time point, along with mean change from baseline to each posttreatment time point. In addition, the overall interpretation of ECG results (normal, not clinically significant abnormal,

clinically significant abnormal) is collected by eCRF according to the scheduled measurements. Shifts in ECG interpretation will be presented as cross-tabulations of numbers of patients with normal, not clinically significant abnormal and clinically significant abnormal ECG interpretation results, including categories of missing and totals.

7.11.5 Other Observations Related to Safety

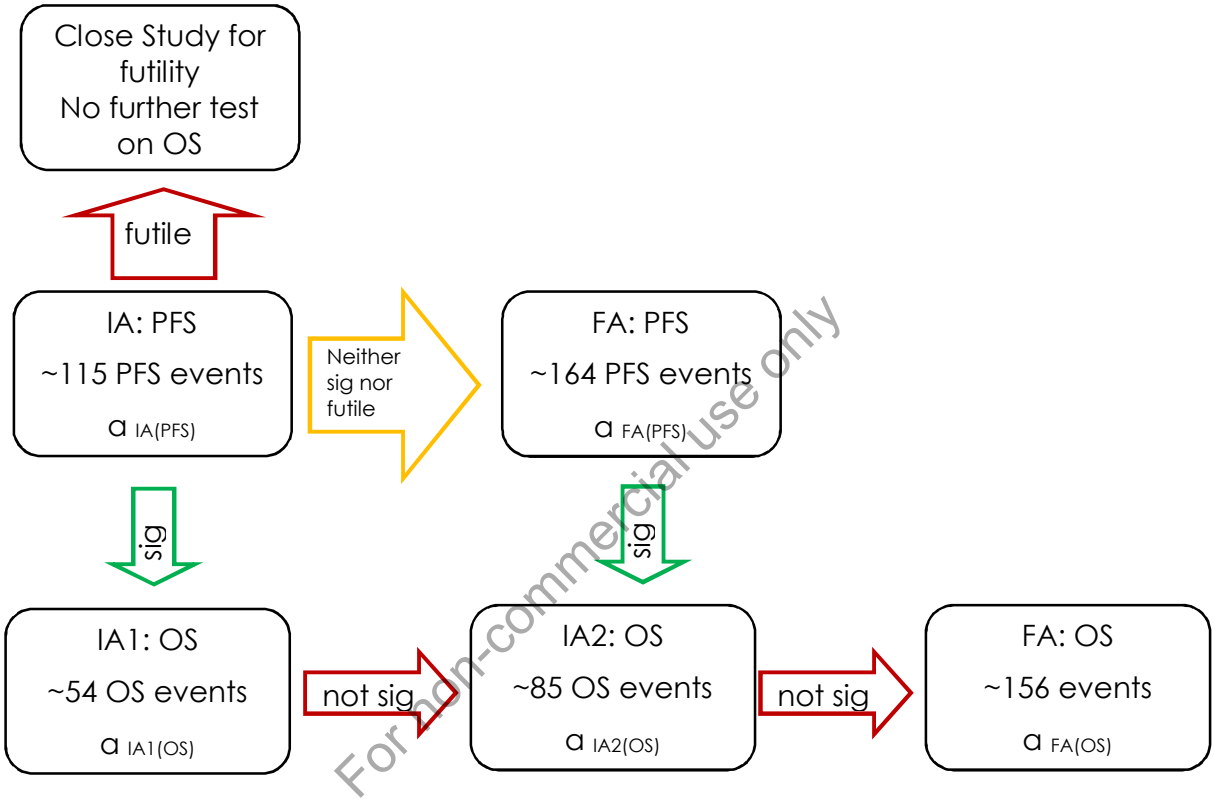
Pregnancy test results will be provided in a by-patient listing.

ECOG performance status scores from baseline to the worst post-baseline value will be summarized in a shift table. By-patient listing of ECOG performance scores will be presented by time.

7.12 Interim Analysis

The family-wise error rate (FWER) will be strongly controlled for the testing of BIRC-assessed PFS and OS accounting for IA by applying the fixed sequence testing procedure with separate O'Brien Fleming (OBF) Lan DeMets alpha spending functions for efficacy stopping boundary with respect to BIRC-assessed PFS and OS. A gamma spending function with parameter -6 will be used for the non-binding futility stopping boundary of BIRC-assessed PFS to control the overall power at 90%. If exactly 115 PFS events occurred at the time of IA for PFS, which was estimated to occur approximately 22 months after first patient randomized, PFS would be significant if its p-value from the stratified log-rank test < 0.0149 and the study will be stopped for futility if the p-value >0.538. Both efficacy and futility boundaries will need to be recalculated if the observed number of PFS events deviates from the expected 115 events. The α levels for OS will be determined using the OBF function based on the information fraction available at each analysis time point, relative to an estimated total of 156 deaths. Assuming BIRC-assessed PFS is significant at IA and exactly 54 OS events have occurred concurrently, OS would be significant at this IA1 of OS if p-value < 0.000278; or significant at the IA2 of OS with 85 deaths if p-value < 0.0047; or finally significant at FA of OS with 156 deaths if p < 0.048. The significance boundaries will be recalculated based on the actual number of observed deaths at IA2 and FA of OS, respectively (Figure 7.1).

Figure 7.1 Multiple testing procedures for controlling FWER among the primary and key secondary endpoints



7.13 Changes in the Statistical Analysis Plan

SAP version 1.0 Summary of Changes

Clarified the primary analysis of the following secondary efficacy endpoints will be based on the confirmed responses (CR or PR): ORR, DOR, TTR, iORR, and iDOR.

Clarified the endpoint of Time to iPD in the protocol is restricted to Time to iPD without prior systemic progression, as assessed by the BIRC, defined as the time interval from randomization until the first date at which intracranial disease progression (iPD) is objectively documented via a modification of RECIST v1.1 without prior systemic progression (Section 7.8.2.8).

SAP version 2.0 Summary of Changes

Details of futility analysis to be performed with the planned interim analysis following protocol amendment 4 are added.

Minor changes e.g. imputation rules, summary of disposition, AE and lab are also included.

8.0 REFERENCES

1. FDA Guidance: "Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics Guidance for Industry" (April 2015)
2. CHMP Guideline on the evaluation of anticancer medicinal products in man Appendix 1 (EMA/CHMP/27994/2008/Rev.1)
3. DeMets DL, Lan KK. Interim analysis: The alpha spending function approach. Stat Med. 1994 Jul 15-30;13(13-14):1341-52.
4. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eu J Cancer 2009;45:228-247.
5. Robins JM, Hernan M, Brumback B. Marginal strucural models and causal inference in epidemiology. Epidemiology 2000; 11(5): 550-560.
6. Robins JM and Finkelstein DM. Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests. Biometrics 2000; 56: 779-788.
7. Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. Stat Med 1990;9(12):1439-45.

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
[redacted]	Biostatistics Approval	18-Jun-2021 21:38 UTC