Official Title: Randomized, Multicenter, Phase III, Open-Label Study of Alectinib Versus Crizotinib in Treatment-Naive Anaplastic Lymphoma Kinase-Positive Advanced Non-Small Cell Lung Cancer

NCT Number: NCT02075840

Document Date: Statistical Analysis Plan Version 4: 03-Mar-2017

STATISTICAL ANALYSIS PLAN

TITLE: **RANDOMIZED, MULTICENTER, PHASE III, OPEN-LABEL STUDY OF ALECTINIB VERSUS CRIZOTINIB IN TREATMENT-NAIVE ANAPLASTIC LYMPHOMA KINASE–POSITIVE ADVANCED NON–SMALL CELL LUNG CANCER**

PROTOCOL NUMBER: BO28984

STUDY DRUG: Alectinib

VERSION NUMBER: 4

IND NUMBER: 111723

EUDRACT NUMBER: 2013-004133-33

SPONSOR: F. Hoffmann—La Roche Ltd

PLAN PREPARED BY: [redacted]

DATE FINAL: 30 November 2015

DATE(S) AMENDED: Version 2: 1 November 2016
Version 3: 9 December 2016
Version 4: See electronic date stamp below.

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory (Clinical)	03-Mar-2017 08:06:53

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Definition
ALK	anaplastic lymphoma kinase
CDOR	CNS duration of response
CORR	CNS objective response rate
CR	complete response
CSR	Clinical Study Report
CT	computed tomography
DOR	duration of response
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
EORTC	European Organization for Research and Treatment of Cancer
FDA	U.S. Food and Drug Administration
FISH	fluorescence in situ hybridization
FPP	FISH positive population
HR	hazard ratio
iDCC	independent Data Coordinating Cente
iDMC	independent Data Monitoring Committee
IHC	immunohistochemistry
IRC	Independent Review Committee
ITT	intent-to-treat
IxRS	interactive voice or Web-based response system
MRI	magnetic resonance imaging
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	not evaluable
NSCLC	non–small cell lung cancer
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PK	pharmacokinetic
PR	partial response
PRO	patient-reported outcome
PS	Performance Status
QLQ-C30	Core Quality of Life Questionnaire
QLQ-LC13	Quality of Life Questionnaire Lung Cancer Module
QoL	quality of life

Abbreviation or Term	Definition
RANO	Response Assessment in Neuro-Oncology
RECIST	Response Evaluation Criteria in Solid Tumors
SAP	Statistical Analysis Plan
SD	stable disease
SDF	survival distribution function
TTD	time to deterioration
ULN	upper limit of normal

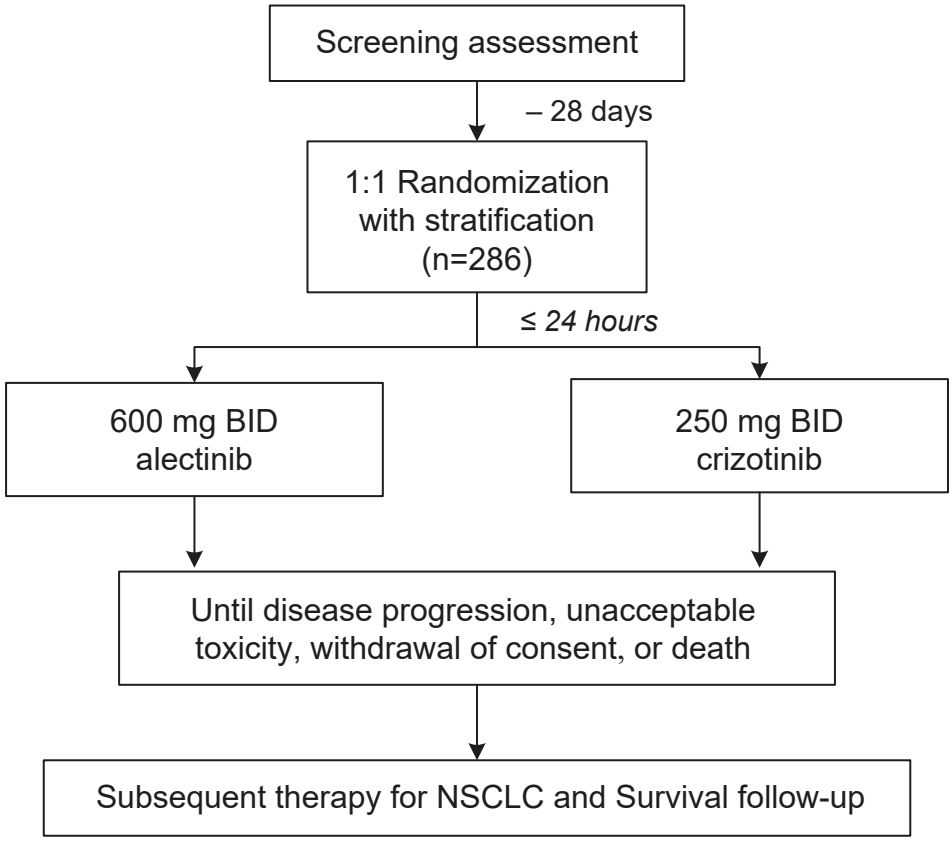
1. BACKGROUND

Alectinib is a small molecule that is highly selective and a potent inhibitor of anaplastic lymphoma kinase (ALK). Alectinib is currently being developed for the treatment of patients who have ALK-positive non–small cell lung cancer (NSCLC), as detected by a U.S. Food and Drug Administration (FDA)-approved test.

2. STUDY DESIGN

This is a randomized, active-controlled, multicenter, Phase III, open-label study in patients with treatment-naive ALK-positive advanced NSCLC. All patients are required to provide pretreatment tumor tissue to confirm the presence of ALK rearrangement (by immunohistochemistry [IHC] test). Patients will be randomized in a 1:1 ratio into one of the two treatment arms to receive either alectinib or crizotinib (Figure 1).

Figure 1 Summary of Study Design



BID = twice daily; n = number of patients; NSCLC = non-small cell lung cancer.

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. The outcome measures are listed there.

2.2 DETERMINATION OF SAMPLE SIZE

The focus of this clinical trial is hypothesis testing, and the primary endpoint of progression-free survival (PFS) was used to determine the sample size of the study.

The Phase III PROFILE 1014 study of crizotinib versus standard pemetrexed–platinum-based chemotherapy in previously untreated patients with ALK-positive non-squamous NSCLC reported a median PFS of 10.9 months for crizotinib (Solomon et al. 2014). A hazard ratio (HR) of 0.65 for alectinib versus crizotinib (i.e., an increase in median PFS from 10.9 months to 16.8 months) will be targeted.

In this study, 286 patients will be enrolled in a 1:1 randomization allocation. Enrollment will take approximately 24 months on the basis of an assumption of non-linear recruitment as follows:

- Month 1: 1 patient per month
- Month 2: 2 patients per month
- Month 3: 4 patients per month
- Month 4: 6 patients per month
- Month 5: 8 patients per month
- Month 6: 10 patients per month
- Month 7: 12 patients per month
- Months 8–12: 13 patients per month
- Months 13–14: 14 patients per month
- Month 15 onwards: 15 patients per month

Approximately 170 PFS events are required to achieve 80% power of the log-rank test at a two-sided alpha level of 5%. This number of events is estimated to occur approximately 33 months after the first patient has been enrolled.

To illustrate sensitivity, if only 160 PFS events are observed by the clinical cutoff, the study has 78.1% power; if only 165 PFS events are observed, the study has 79.3% power with the assumed HR of 0.65 and an alpha of 0.05.

No interim analysis for efficacy or futility is planned.

An analysis of overall survival (OS) will be performed at the time of the final analysis of the primary endpoint of PFS. A survival follow-up analysis will be performed once approximately 50% of patients (i.e., 143 patients) have died. The median OS in the crizotinib arm is assumed to be 24 months and the expected median OS in the alectinib treatment arm is 30 months, equating to an HR of 0.83. On the basis of the sample size ($n = 286$), the trial will not be powered to demonstrate a statistically significant difference

in OS of this magnitude. At the time of the final analysis of the primary endpoint of PFS, on the basis of the above assumptions, 106 OS events are expected to have occurred. The events required for the survival follow-up analysis are expected to occur approximately 42 months after the first patient has been enrolled.

2.3 ANALYSIS TIMING

This study is event driven with a recruitment period of approximately 24 months. The required number of events for the primary analysis of the primary endpoint (investigator-assessed PFS) is expected approximately 33 months after the first patient has been enrolled. If the alectinib median PFS by investigator is not estimable at the time of the primary analysis due to an insufficient number of events in this treatment arm, an exploratory update of PFS analysis will be performed when approximately 50% patients in the alectinib arm will have experienced a PFS event. This analysis will be conduced purely for the purpose of obtaining a median estimate of PFS in the alectinib treatment arm, no formal treatment comparisons will be performed for this endpoint after the primary analysis.

A survival follow-up analysis will be performed once approximately 50% of patients have died, which is estimated to occur approximately 42 months after the first patient has been enrolled.

The study will formally end once the survival follow-up analysis is complete.

3. STUDY CONDUCT

3.1 RANDOMIZATION ISSUES

Patients will be randomly assigned in a 1:1 allocation ratio to the two treatment arms via a block stratified randomization procedure.

Randomization will guard against systematic selection bias and should ensure the comparability of treatment groups. To assist balance in important prognostic factors, randomization will be stratified by Eastern Cooperative Oncology Group Performance Status (ECOG PS; 0/1 vs. 2), race (Asian vs. non-Asian), and CNS metastases at baseline (yes vs. no).

The randomization is checked three times during the study conduct by the independent Data Coordinating Center (iDCC).

3.2 INDEPENDENT REVIEW FACILITY

An independent imaging group will be used to evaluate tumor assessments for determination of progression and response rate according to Response Evaluation Criteria in Solid Tumors (RECIST), Version 1.1 for the primary analysis. Imaging studies (computed tomography [CT]/magnetic resonance imaging [MRI]/bone scans) will be acquired according to a standard protocol and will be forwarded to the independent

reviewers. In addition, relevant cytologic and medical data (e.g., relevant cytology reports documenting malignant pleural effusions, bone marrow aspirations, cerebral spinal fluid, CNS symptoms, corticosteroid use, etc.) may be forwarded, if available, to the independent reviewers to aid with assessment of progressive disease (PD) and response. Investigator tumor assessments will not be reconciled with the Independent Review Committee (IRC) tumor assessments. Further details are included in the IRC Charter. Details of imaging handling procedures are also described in a separate laboratory manual.

For evaluation of the CNS endpoints, the IRC will perform an assessment of scans based on RECIST v1.1 and Response Assessment in Neuro-Oncology (RANO) criteria. More details are given in the IRC charter.

3.3 DATA MONITORING

Although the study is open label, the Sponsor will not analyze aggregated data by treatment group before analysis of the primary endpoint of PFS.

An independent Data Monitoring Committee (iDMC) has been incorporated into the trial design to objectively and carefully review the accumulating safety data for alectinib. Regular safety interim analyses to be reviewed by the iDMC have been planned to evaluate safety information from the study. Outputs for these reviews are prepared by an external iDCC.

4. STATISTICAL METHODS

The analyses outlined in this Statistical Analysis Plan (SAP) supersede those specified in the protocol for the purpose of a regulatory filing.

4.1 ANALYSIS POPULATIONS

4.1.1 Intent-To-Treat Population

The primary analysis population for efficacy is the intent-to-treat (ITT) population, defined as all randomized patients. Patients will be assigned to the treatment group to which they were randomized.

4.1.2 Safety Population

The primary analysis population for safety is the Safety Population, defined as all patients who received at least one dose of study medication. Patients will be assigned to treatment groups as treated, and all patients who received any dose of alectinib will be included in the alectinib treatment arm.

4.1.3 FISH Positive Population

A secondary analysis population for efficacy is the fluorescence in situ hybridization (FISH) Positive Population (FPP), defined as all patients in the ITT population who were ALK-positive as assessed using the Vysis FISH assay. Patients will be assigned to the

treatment group to which they were randomized. A supportive efficacy analysis will be performed based on the FISH assay rather than the IHC.

An analysis of the primary endpoint of investigator-assessed PFS based on the FPP will be performed at the time of the primary analysis. T-tests to compare baseline characteristics by treatment group will be performed to assess any imbalance between treatment arms in the FPP. In the event that the randomization is not maintained in the FPP (t-test $p<0.01$), log-rank analyses of investigator-assessed PFS will be stratified by randomization strata and the significant baseline characteristics will be included in the Cox model. A major discrepancy between the Ventana IHC assay and the Vysis FISH assay may necessitate a follow-up analysis of the primary endpoint based on approximately 170 PFS events observed in the FPP to ensure 80% power of the log-rank test under the assumption of an HR of 0.65 in this secondary population.

4.1.4 Pharmacokinetic-Evaluable Population

The Pharmacokinetic-Evaluable Population is defined as all patients who received any dose of study medication and who have at least one post-baseline pharmacokinetic (PK) sample available.

4.2 ANALYSIS OF STUDY CONDUCT

Study enrollment, patient disposition, reasons for discontinuation from the study treatment, and reason for study termination will be summarized for all patients in the ITT population.

Protocol deviations, including violations of inclusion and/or exclusion criteria and deviations during study conduct will be reported and summarized.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic, baseline disease characteristics, and lung cancer history, including prior brain radiation, will be compared descriptively between both treatment arms for the ITT population. Descriptive baseline summaries of continuous data will present the group mean, standard deviation, median, minimum, and maximum. Descriptive baseline summaries of discrete data will present the category counts as frequencies and percentages.

A summary of concordance of stratification factors determined by electronic Case Report Form (eCRF) versus interactive voice or Web-based response system (IxRS) will also be reported.

The baseline value of any variable will be defined as the last available value recorded prior to the first administration of study medication.

Previous and concomitant cancer therapy will also be summarized, including radiotherapy and surgery, as well as subsequent anti-cancer therapy. Previous and concurrent diseases and medications will also be summarized.

The Abbott Vysis FISH test will be used as an exploratory assay after patients have been enrolled in this study. A descriptive summary of baseline characteristics, including randomization strata by treatment arm, will be produced for the FPP to assess treatment group comparability.

4.4 EFFICACY ANALYSIS

4.4.1 Primary Efficacy Endpoint

PFS is defined as the time from date of randomization to the date of first documented disease progression or death, whichever occurs first. The primary endpoint of PFS will be determined on the basis of investigator assessment of progression using RECIST v1.1. Patients who have not experienced disease progression or death at the time of analysis will be censored at the last tumor assessment date either during study treatment or during follow-up. Patients with no post-baseline tumor assessment will be censored at the date of randomization.

Patients who discontinue treatment prior to disease progression (e.g., due to toxicity) will continue in the study and will be followed until disease progression and for OS regardless of whether they subsequently receive anti-cancer therapy.

The treatment comparison of PFS will be based on a stratified log-rank test at the 5% level of significance (two-sided). The randomization stratification factors are ECOG PS (0/1 vs. 2) and race (Asian vs. non-Asian), as recorded on the eCRF, and CNS metastases at baseline (yes vs. no). These factors will be included in the stratified log-rank analysis as long as an individual stratum includes $>10\%$ of the ITT population. For analysis purposes, stratification according to CNS metastases at baseline will be performed on the basis of the IRC assessment rather than the investigator assessment, given that the independent assessment by neuroradiologists is deemed to be the most reliable and will correspond to the populations used to assess the CNS efficacy endpoints. Results from an unstratified log-rank test will also be presented as a supportive analysis.

Because patients were stratified on the basis of CNS metastases at baseline by investigator assessment, baseline characteristics grouped by CNS metastases at baseline by IRC will be summarized by treatment arm. Concordance of CNS metastases at baseline between investigator and IRC will also be reported.

Additional supportive analyses include Kaplan-Meier and Cox modelling approaches. The Kaplan-Meier method will be used to estimate the median PFS for each treatment arm with 95% confidence limits, and a Kaplan-Meier curve will be constructed to provide a visual description of the difference between the treatment arms. A stratified Cox

proportional hazard regression model will be used including treatment in order to provide an estimate of the treatment effect expressed as an HR (alectinib vs. crizotinib), as well as a 95% CI. The proportional hazards assumption will be assessed both graphically from the Kaplan-Meier plot as well as by adding a treatment by time interaction term to the Cox regression model. If the proportional hazards assumption is not met, alternative appropriate methods will be used.

The difference between the two treatment groups for the primary objective will be assessed and tested for the following hypothesis: the survival distribution function (SDF) of the alectinib treatment group is the same as for the crizotinib treatment group versus the alternative that the two distributions are different:

H_0: SDF (alectinib) = SDF (crizotinib)

versus

H_A: SDF (alectinib) $\neq$ SDF (crizotinib),

where SDF denotes the survival distribution function of the parameter PFS.

4.4.2 Secondary Efficacy Endpoints

All secondary efficacy endpoints will be analyzed in the ITT population unless otherwise specified.

If the primary endpoint of investigator-assessed PFS is statistically significant at a two-sided 5% significance level based on the stratified log-rank test, the following secondary endpoints will be tested in the following sequential order (O'Neill 1997), each at a two-sided 5% significance level:

- PFS by IRC
- Time to CNS progression by IRC RECIST criteria
- Objective response rate (ORR) by investigator assessment
- OS

All tests in the sequence will be based on a stratified log-rank test at the 5% level of significance (two-sided). The stratification factors included and the analysis population will be the same as for the primary hypothesis test. Results from unstratified tests will also be presented as supportive analyses.

4.4.2.1 PFS by IRC

An analysis of PFS on the basis of the IRC assessments will be performed using the same methodology as specified for PFS on the basis of investigator assessment.

A concordance analysis between the IRC-determined and the investigator-determined PD status (yes vs. no) will be provided by treatment arm together with a summary of agreement between the IRC-determined and the investigator-determined PD dates. The early discrepancy rate (investigator assessed prior to IRC) and the late discrepancy rate

(investigator assessed after IRC) will be calculated for each treatment arm and the differential discordance calculated as the rate on the alectinib arm minus the rate on the crizotinib arm (Amit et al. 2011). In addition, a table will be created displaying the number of visits with investigator tumor assessment but missing IRC assessment.

4.4.2.2 Time to CNS Progression by IRC RECIST Criteria

Time to CNS progression is defined as the time from randomization until first radiographic evidence of CNS progression by independent review. An independent central radiological review will be performed for all patients, and the analysis of CNS progression will be based on the data from the independent review. All patients in the ITT population will be included in the analysis regardless of their baseline status of CNS metastases. CNS progression is defined as progression due to newly developed CNS lesions and/or progression of preexisting baseline CNS lesions. On the basis of RECIST v1.1, this is defined as a new post-baseline CNS/brain lesion(s) and/or an increase of $\geq$20% in the sum of longest diameters of the measurable baseline CNS lesions compared with nadir and/or unequivocal progression of non-measurable baseline CNS lesions.

In order to account for the competing risks inherent in the comparison of CNS progression between alectinib and crizotinib, a stratified two-sided log-rank test will be computed on the basis of a cause-specific hazard function. Results from unstratified tests will also be presented as supportive analyses.

The probability of CNS progression, non-CNS progression, and death by treatment group with 95% CIs will each be estimated using cumulative incidence functions. Gray's test to compare the risk of CNS progression between alectinib and crizotinib will also be performed as a supportive analysis.

Similar analyses of CNS progression based on IRC RANO criteria will be performed.

4.4.2.3 Objective Response Rate

ORR, on the basis of investigator assessment, is defined as the percentage of patients in the ITT population with measurable disease at baseline who attain a complete response (CR) or partial response (PR). Per RECIST v1.1, confirmation of objective response is not required for this secondary endpoint. Patients without a post-baseline tumor assessment will be considered non-responders, as will patients with a best overall response of stable disease (SD), PD, or not evaluable (NE).

An estimate of ORR and its two-sided 95% CI will be calculated using the Clopper-Pearson method for each treatment arm. Response rates in the treatment groups will be compared using a stratified Mantel-Haenszel test, including the randomization stratification factors. Results from unstratified tests will also be presented as supportive analyses. The difference in ORR between the two treatment arms will be

presented together with a two-sided 95% CI on the basis of a normal approximation to the binomial distribution.

ORR by IRC will be analyzed similarly as a supportive analysis in the population of patients identified to have measureable disease at baseline according to the IRC.

4.4.2.4 Duration of Response

For patients with measurable disease at baseline who have experienced an objective response (CR or PR) during the study as assessed by the investigator, duration of response (DOR) is defined as the duration from the first tumor assessment that supports the patient's objective response (CR or PR, whichever is first recorded) to first documented disease progression or death due to any causes, whichever occurred first. Patients who have not progressed or died at the time of analysis will be censored at the last tumor assessment date. Because the determination of DOR is based on a non-randomized subset of patients, formal hypothesis testing will not be performed. DOR will be estimated using Kaplan-Meier methodology and an HR and its 95% CI on the basis of a Cox proportional regression model with stratification factors will be calculated. The proportional hazards assumption will be assessed graphically.

4.4.2.5 Overall Survival

OS is defined as the time from the date of randomization to the date of death due to any cause. Patients who are not reported as having died at the time of analysis will be censored at the date when they were last known to be alive. Patients who do not have post-baseline information will be censored at the date of randomization. OS will be analyzed using the same methodology as specified for the primary endpoint. A survival follow-up analysis will be performed based on more mature data.

4.4.2.6 CNS Objective Response Rate according to RECIST v1.1 and RANO Criteria by IRC

CNS objective response rate (CORR) will be summarized by treatment group in the subgroup of patients in the ITT population who have measurable CNS lesions at baseline as determined by the IRC. These summaries will be provided for both RECIST v1.1 and RANO. CNS responses according to RECIST v1.1 do not have to be confirmed, whereas confirmation is incorporated into the RANO criteria.

CORR will be summarized by treatment group in the subgroup of patients in the ITT population who have measurable and non-measurable CNS lesions at baseline as determined by the IRC. For this analysis, patients with non-measurable CNS disease at baseline by IRC who experience a CR in CNS lesions will be included as responders.

In addition, a subgroup analysis of CORR will be performed according to prior brain radiation status (yes vs. no) for the two analysis populations above.

4.4.2.7 CNS Duration of Response according to RECIST v1.1 and RANO Criteria by IRC

This analysis will be performed including all patients in the ITT population who have measurable CNS lesions at baseline as determined by the IRC who had a CNS response (CR or PR). CNS duration of response (CDOR) is the time from CNS response to CNS PD by IRC.

4.4.3 Exploratory Efficacy Endpoints

The Abbott Vysis FISH test will be used as an exploratory assay after patients have been enrolled in this study. As part of the rationale for the biomarker assessment, results from these analyses will be used to quantify the degree of correlation between FISH and IHC for the detection of ALK-positive NSCLC. The proportion of patients in the ITT population who have ALK-positive NSCLC by the FISH assay will be summarized. The FPP will be a secondary analysis population for efficacy. This analysis population will be used to perform a supportive analysis of the study data based on the FISH assay rather than the IHC.

4.4.4 Sensitivity Analyses

A sensitivity analysis will be performed on the primary endpoint of PFS with the following changes from the primary analysis:

- Censor patients at the last adequate tumor assessment prior to the start of non–protocol-specified anti-cancer therapy received prior to observing progression.
- Censor patients for whom documentation of disease progression or death occurs after ≥ 2 missed tumor assessments. These patients will be censored at the last tumor assessment prior to the missed assessments.
- Censor patients who discontinue study treatment (due to personal preference or toxicity) and/or withdraw or are lost to follow-up prior to observing progression.

Two additional sensitivity analyses for PFS include:

- The effect of missing tumor assessments will be assessed if the number of missing assessments in either arm is $> 5\%$. For patients with progression determined following one or more missing tumor assessments, the progression will be backdated to the first missing tumor assessment.
- The effect of loss to follow-up will be assessed depending on the number of patients who are lost to follow-up. If $> 5\%$ of patients are lost to follow-up for PFS in either treatment arm, a “worst-case” analysis will be performed in which patients who are lost to follow-up will be considered to have progressed at the last date they were known to be progression-free.

Summary statistics on timing of scans (regardless of outcome) will be provided by treatment arm. Nelson-Aalen methods may be used to estimate the cumulative hazard rate.

A sensitivity analysis for PFS and OS will be performed based on the stratification factors entered at randomization in the IxRS system.

4.4.5 Subgroup Analyses

PFS by investigator and IRC assessments will be presented separately for important subgroups including age (<65, ≥ 65), sex, race (Asian, non-Asian), and smoking status and baseline prognostic characteristics including baseline ECOG PS, CNS metastases at baseline as determined by IRC, and prior brain radiation (in patients with CNS metastases at baseline). The HR including a 95% CI will be presented separately for each level of the categorical variables.

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Standard non-compartmental analysis may be conducted for PK data collected from patients participating in serial and/or intensive PK collections for relevant analytes, as data allow, as appropriate, and if needed. PK parameters, including but not limited to area under the concentration-time curve, maximum concentration, and time to maximum concentration, will be calculated on the basis of the available data as appropriate and where data allow. Additional PK parameters may be calculated as deemed appropriate.

Individual and mean plasma concentrations at each sampling timepoint and/or PK parameters for alectinib and metabolite(s) will be listed, as appropriate.

Summary statistics (e.g., means, standard deviation, coefficient of variation percent, geometric means, medians, and ranges) for plasma concentrations and/or PK parameters for alectinib and metabolite(s) will be presented by nominal collection times (plasma concentrations only), as appropriate. Additional plots or summary statistics may be constructed or calculated, as appropriate.

Results of PK and/or any PK/pharmacodynamic analyses may be reported outside the Clinical Study Report (CSR).

Non-linear mixed-effects modeling (with software NONMEM) will be used to analyze the sparse and/or serial/intensive plasma concentration-time data for alectinib. The PK data from this study may be pooled with data from other studies. Population and individual PK parameters will be estimated and the influence of various covariates (such as age, sex, and body weight) on these parameters will be investigated.

Exploratory analyses will be conducted to investigate the relationship between alectinib PK exposure and efficacy/safety parameters.

Results of the mixed-effects modeling and exploratory analyses will be reported in a document separate from the CSR.

4.6 SAFETY ANALYSES

All safety analyses will be performed on the Safety Population; that is, all patients who receive any dose of study medication (see Section 4.1.2) by treatment arm with all patients who received any dose of alectinib will be included in the alectinib treatment arm.

4.6.1 Exposure of Study Medication

Study treatment (alectinib and crizotinib) exposure, including treatment duration, dose intensity, number of doses, percentage of planned dose, and total cumulative dose will be summarized with descriptive statistics. The number of missed doses will also be displayed.

4.6.2 Adverse Events

Verbatim description of adverse events will be mapped to MedDRA thesaurus terms and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.0 (NCI CTCAE v4.0). All adverse events occurring during or after the first study treatment will be listed and summarized by mapped term, appropriate thesaurus level and NCI CTCAE grade.

Summary tables of the following will be provided:

- All adverse events
- Serious adverse events
- Severe adverse events (Grade 3 or higher)
- Adverse events leading to study treatment discontinuation
- Adverse events leading to dose reduction
- Adverse events leading to dose interruption
- Treatment-related adverse events
- Adverse events leading to death
- Adverse events by highest NCI CTCAE grade
- Selected adverse events relating to ALK inhibitors and/or the tyrosine kinase inhibitor class and alectinib data

A summary table of common adverse events, that is, those occurring in at least 10% of patients, will be provided.

Selected adverse events are defined as follows:

- Hepatotoxicity
- Interstitial lung disease
- Vision disorders
- Skin disorders (e.g., photosensitivity, rash)
- Anemia
- Gastrointestinal disorders (e.g., nausea, vomiting, diarrhea)
- Abnormal renal function (e.g., serum creatinine increase, renal impairment, renal failure)
- Severe myalgia and CPK elevations
- Edema
- Bradycardia

Rates of each selected adverse event group will be summarized including incidence of Grade 3 or higher events, serious adverse events, and treatment-related events.

Multiple occurrences of the same event will be counted once at the maximum severity. All deaths and causes of death will be summarized.

Adverse events by sex, age (<65 vs. ≥ 65 years), race (non-Asian vs. Asian), and CNS metastases at baseline (yes vs. no) will be presented.

4.6.3 Laboratory Data

Laboratory data will be classified according to NCI CTCAE v4.0 and will be summarized descriptively over time, including change from baseline. Highest NCI CTCAE grade after baseline will also be reported and shift tables from baseline to worst value during the study post-baseline will be presented. An additional shift table for serum creatinine will include only elevations above the upper limit of normal (ULN) (i.e., elevations above baseline, which are included in the full NCI CTCAE v4.0 definition, will be excluded from this shift table).

A Hy's law analysis will be provided. The potential Hy's law quadrant is defined as ALT or AST increases above 3-fold the ULN with concomitant total bilirubin increases above 2-fold the ULN.

In order to evaluate and compare potential hypogonadism in the two treatment arms, total and free testosterone, sex hormone-binding globulin (if available), follicle-stimulating hormone, and luteinizing hormone levels in the blood will be summarized for males by treatment arm over time.

4.6.4 Vital Signs and ECOG Performance Status

Vital signs will be summarized descriptively by treatment group over time, including change from baseline. ECOG PS will also be tabulated over time by treatment group.

4.6.5 Electrocardiograms

The following parameters are captured on the eCRF and will be listed and summarized by treatment group over time: heart rate; RR (time from one R wave to next R wave); PQ; QRS and QT duration; and QT interval corrected using Fridericia's formula.

4.7 PATIENT-REPORTED OUTCOME ANALYSES

Through the use of the European Organization for Research and Treatment of Cancer (EORTC) Core Quality of Life Questionnaire (QLQ-C30) and EORTC Quality of Life Questionnaire Lung Cancer Module (QLQ-LC13), lung cancer symptoms, symptoms commonly associated with cancer treatments, and disease and treatment's impact on patients' functioning and health-related quality of life will be collected at every study visit (Week 4, Week 8, and then after every 4 weeks) until disease progression and during post-progression treatment in case of isolated, asymptomatic CNS progression; at post-treatment visit (4 weeks after permanent treatment discontinuation); and at subsequent 8-weekly survival follow-up visits for 6 months. For patients who discontinue treatment for reasons other than disease progression and who progress within the first 6 months of survival follow-up, patient-reported outcome (PRO) measures will be administered every 4 weeks until disease progression and then decreased to every 8 weeks until 6 months post-treatment. For patients who discontinue treatment for reasons other than disease progression and have not yet progressed at 6 months post-treatment, PRO measures will be administered every 4 weeks until disease progression and will no longer be required thereafter.

The baseline PRO assessment will be the first PRO assessment completed within 7 days of randomization. For the following PRO analyses, there may be multiple PRO assessments completed within the same calendar day; the earliest assessment will be used for analysis. In the cases that there are multiple PRO assessments completed within the same time window scheduled for PRO assessments, the PRO assessment closest to the visit will be chosen for analysis.

The EORTC QLQ-C30 and EORTC QLQ-LC13 will be scored according to the EORTC scoring manual 3rd edition (Fayers et al. 2001). The QLQ-C30 and QLQ-LC13 are composed of both multi-item scales and single-item measures including functional scales, symptom scales, and a global health status/quality of life (QoL) scale.

For multi-item subscales, if $\leq$50% of items within the multi-item subscale are missing at a given timepoint, the multi-item score will be calculated on the basis of the non-missing items. If $>$50% of items are missing or if a single-item measure is missing, the subscale is missing.

All of the scales and single-item measures will be linearly transformed so that each score will range from 0 to 100. A high score for a functional scale represents a high/healthy level of functioning, a high score for the global health status/QoL represents a high QoL; however, a high score for a symptom scale/item represents a high level of symptomatology/problems.

The EuroQoL 5 Dimension questioannire analyses are out of scope of this SAP.

4.7.1 Time to Deterioration of Patient-Reported Lung Cancer Symptoms, Global Health Status, and Cognitive Function

Time to deterioration (TTD) analyses will be performed on lung cancer symptom scores, global health status scores, and cognitive function scale scores in the ITT population and will include all data collected through disease progression and survival follow-up.

TTD is defined as the time from randomization until the first confirmed clinically meaningful deterioration. Confirmed clinically meaningful deterioration in lung cancer symptoms is defined as a ≥10-point increase from baseline in a symptom score that must be held for at least two consecutive assessments or an initial ≥10-point increase above baseline followed by death within 5 weeks from the last assessment. Conversely, confirmed clinically meaningful deterioration in global health status or function is defined as a ≥10-point decrease from baseline in a symptom score that must be held for at least two consecutive assessments or an initial ≥10-point decrease from baseline followed by death within 5 weeks from the last assessment. A ≥10-point change in the score is perceived by patients as being clinically significant (Osoba et al. 1998).

TTD will be documented for each of the following symptom, global health status, or cognitive function scale scores:

- Cough (Question 31 on the EORTC QLQ-LC13)
- Dyspnea single item (Question 8 on the QLQ-C30)
- Dyspnea multi-item subscale (Questions 33-35 on the QLQ-LC13)
- Chest pain (Question 40 on the QLQ-LC13)
- Arm and/or shoulder pain (Question 41 on the QLQ-LC13)
- Fatigue multi-item subscale (Questions 10, 12, and 18 on the QLQ-C30)
- Global Health Status scale (Questions 29 and 30 on the QLQ-C30)
- Cognitive Function Scale (Questions 20 and 25 on the QLQ-C30)

In addition to those single symptom endpoints, a 3-symptom composite endpoint will also be assessed. In this instance, symptom deterioration will be determined as a ≥10-point increase above baseline in any of the following lung cancer symptom scores, whichever occurs first (cough, chest pain, and dyspnea multi-item subscale). In the same manner as the single symptom endpoints, symptom deterioration will need to be confirmed for the original symptom; a ≥10-point increase from baseline in a symptom

score (i.e., cough) that must be held for at least two consecutive assessments (i.e., cough for two assessments) or an initial ≥10-point increase above baseline followed by death within 5 weeks from the last assessment.

Patients without baseline or post-baseline EORTC symptom or scale scores will be censored at the date of randomization. Patients without deterioration at the time of analysis will be censored at the last time they were known to have not deteriorated.

There will be no imputation for missing data for the TTD analysis. If there is a substantial amount of missingness, patient baseline and disease characteristics will be compared between patients with and without missing baseline assessments using descriptive statistics.

TTD of the prespecified symptoms will be summarized using the Kaplan-Meier method. The estimated Kaplan-Meier plots will be provided for each score previously described (including the 3-symptom composite score). A stratified log-rank test will be the primary method to compare the time to first deterioration between the treatment groups. The median time and two-sided 95% CI for the median will also be provided. Estimates of the treatment effect will be expressed as HRs with use of a stratified Cox model, including 95% CIs for the ITT population.

A sensitivity analysis will be performed to account for multiple PRO assessments completed within 7 days of randomization, the same day, or multiple assessments completed within the same time window scheduled for PRO assessments. The sensitivity analyses will use the following data:

- In the case that there are multiple baseline PRO assessments completed within 7 days of randomization, the assessment with the best score (lowest score for symptoms, highest score for global health status or function scales) for any of the pre-specified scores will be selected.
- In case of multiple assessment within the same calendar day, the worst score (highest score for symptoms, lowest score for global health status or function scales) for any of the prespecified scores will be selected.
- In the case of multiple assessments within a single time window scheduled for PRO assessments, the assessment with the worst score (highest score for symptoms, lowest score for global health status or function scales) for any of the prespecified scores will be selected.

There is risk that PFS endpoints (or death) may confound differences in CNS events and in PRO cognitive function, therefore, a competing risks analysis may be done on the TTD for Cognitive Function.

4.7.2 Additional Patient-Reported Outcomes Analyses

Only patients in ITT population with a baseline assessment and at least one post-baseline assessment (PRO-evaluable subset) will be included in the following

analyses. Missing post-baseline items and subscales will not be imputed and will be summarized as observed.

Compliance rates will be summarized by listing the number and proportion of patients in the PRO-evaluable subset who completed the PRO assessments at each timepoint by treatment arm. Reasons for non-completion will be summarized, if available, in the CRF.

All of the following analyses will be performed on all timepoints during the treatement period as well as at the time of disease progression per RECIST v1.1 (PRO assessment completed within ±7 days of date of radiographic pharmacodynamic) at the last dose of treatment received before treatment discontinuation of any cause, at the 4-week post-treatment discontinuation visit, and at the 8-weekly survival follow-up visits through 6 months.

- EORTC scores and change from baseline will be descriptively analyzed using means, standard deviations, medians, and range, by treatment arm at baseline and each of the timepoints previously defined.
- The number and proportion of patients at each of the 4-response category reflecting no symptom/impairment to severe symptom/impairment will be documented at baseline.
- Graphs of the mean changes and standard errors over time from the baseline assessment of items and subscales will be provided for each treatment arm.
- The number and proportion of patients who improved, worsened, or remained stable compared with baseline score will be summarized for all item-score and subscales score of the EORTC QLQ-C30 questionnaire and the QLQ-LC13 by treatment arm at baseline and each of the timepoints previously defined. This analysis should be performed for the PRO-evaluable population as well as the subgroups of the patients with/without CNS metastases at baseline. A patient will be deemed improved at a timepoint if there is a ≥10-point decrease from baseline for any symptom or a ≥10-point increase from baseline for any functioning scales or global health status scale. A patient will be deemed worsened at that timepoint if they have a ≥10-point increase from baseline for any symptom or a ≥10-point decrease from baseline for any functioning scales or global health status. If patients do not fit the improved or worsened criteria, they will be defined as stable for that specific item/scale at that timepoint.

Additional analyses will be conducted specifically for the cognitive function scale per the EORTC QLQ-C30:

- The number and proportion (%) of patients with ≥10-point decrease from baseline in the Cognitive Function Scale score per the EORTC QLQ-C30 in each treatment arm at (1) any point during the treatment or (2) at date of CNS Progression (PRO assessment completed within ±7 days of date of CNS PD). This analysis will not include PRO data from the 4-week post-treatment discontinuation visit or the PRO data collected at the 8-weekly survival follow-up visits through 6 months.

- A histogram of worst change scores from baseline (decline of 16.66 vs. 33.33 vs. 50 vs. 66.66 vs. 83.33 vs. 100 points from baseline) in the Cognitive Function Scale score in each treatment arm at any point on treatment. This analysis will not include PRO data from the 4-week post-treatment discontinuation visit or the PRO data collected at the 8-weekly survival follow-up visits through 6 months.

4.8 EXPLORATORY ANALYSES

Exploratory outcome measures defined in the protocol other than the Abbot Vysis FISH subgroup analysis (see Section 4.4.3) and onset of hypogonadism in adult men are outside of the scope of this SAP and will not be included in the CSR.

The objective to determine the correlation between ALK status as assessed by plasma ALK polymerase chain reaction and/or plasma ALK sequencing tests with ALK status obtained using the Ventana ALK IHC and FISH Vysis ALK Break Apart FISH Probe Kit (Abbott) is out of scope of this SAP and will be conducted outside of the CSR analyses.

4.9 INTERIM ANALYSES

No interim analysis for efficacy or futility is planned.

5. REFERENCES

Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. Eur J Cancer 2011:1772–8.

Fayers PM, Aaronson N, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: EORTC, 2001.

Osoba D, Rodriguez G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16:139–44.

O'Neill RT. Secondary endpoints cannot be validly analyzed if the primary endpoint does not demonstrate clear statistical significance. Control Clin Trials 1997;18:550–6.

Solomon BJ, Mok T, Kim DW, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N Engl J Med 2014;371:2167–77.

Appendix 1
Protocol Synopsis

TITLE:	**RANDOMIZED, MULTICENTER, PHASE III, OPEN-LABEL STUDY OF ALECTINIB VERSUS CRIZOTINIB IN TREATMENT-NAIVE ANAPLASTIC LYMPHOMA KINASE-POSITIVE ADVANCED NON–SMALL CELL LUNG CANCER**
PROTOCOL NUMBER:	BO28984
VERSION NUMBER:	4
EUDRACT NUMBER:	2013-004133-33
IND NUMBER:	111723
UTN NUMBER:	U1111-1160-7882
TEST PRODUCT:	Alectinib (RO5424802)
PHASE:	Phase III
INDICATION:	Anaplastic lymphoma kinase–positive (ALK-positive) non–small cell lung cancer (NSCLC)
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives

Efficacy Objectives

The primary efficacy objective for this study is as follows:

- To evaluate and compare the efficacy of alectinib compared to crizotinib in patients with treatment-naive anaplastic lymphoma kinase (ALK)-positive advanced non-small cell lung cancer (NSCLC), as measured by investigator-assessed progression-free survival (PFS)

The secondary efficacy objectives for this study are as follows:

- To evaluate and compare the Objective Response Rate (ORR) and Duration of Response (DOR)
- To evaluate and compare the time to progression in the CNS on the basis of Independent Review Committee (IRC) review of radiographs by Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 and Revised Assessment in Neuro Oncology (RANO) criteria, as well as:
    - To evaluate CNS objective response rate (C-ORR) in patients with CNS metastases who have measurable disease in the CNS at baseline
    - To assess CNS duration of response (C-DOR) in patients who have a CNS Objective Response
    - To assess CNS progression rates (C-PR) at 6, 12, 18, and 24 months on the basis of cumulative incidence
- To evaluate and compare the PFS assessment by the IRC
- To evaluate and compare the overall survival (OS)

Appendix 1
Protocol Synopsis (cont.)

Safety Objective

The safety objective for this study is as follows:

- To evaluate the safety and tolerability of alectinib compared to crizotinib

Pharmacokinetic Objective

The secondary pharmacokinetic (PK) objective for this study is as follows:

- To characterize the pharmacokinetics of alectinib and metabolite(s)

Patient-Reported Outcome Objectives

The secondary patient-reported outcome (PRO) objectives for this study are as follows:

- To evaluate and compare time to deterioration (TTD) in patient-reported lung cancer symptoms of cough, dyspnea (single item and multi-item subscales), chest pain, arm and shoulder pain, and fatigue as measured by the European Organization for the Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire – Core (QLQ-C30) and the supplemental lung cancer module (QLQ-LC13) as well as a composite of three symptoms (cough, dyspnea, chest pain)
- To evaluate and compare PROs of health-related quality of life (HRQoL), patient functioning, and side effects of treatment as measured by the EORTC QLQ-C30 and EORTC QLQ-LC13

Exploratory Objectives

The exploratory objectives for this study are as follows:

- To evaluate and compare patient's health status as assessed by the EuroQoL 5 Dimension (EQ-5D-3L) questionnaire to generate utility scores for use in economic models for reimbursement
- To evaluate and compare an onset of hypogonadism in adult men by measuring total testosterone and free testosterone (either by direct measurement or by calculation using albumin and sex hormone-binding globulin [SHBG]), follicle-stimulating hormone (FSH), and luteinizing hormone (LH) levels in blood
- To evaluate and compare efficacy in patients with treatment-naive ALK-positive NSCLC as assessed by the FISH Vysis® ALK Break Apart FISH Probe Kit (Abbott).
- To evaluate and compare efficacy and safety in patients having treatment-naive ALK positive NSCLC as assessed by plasma ALK assays (polymerase chain reaction [PCR] and/or sequencing)
- To determine the correlation between ALK status as assessed by plasma ALK PCR and/or plasma ALK sequencing tests, with ALK status obtained using the Ventana ALK IHC and FISH Vysis ALK Break Apart FISH Probe Kit (Abbott)
- To investigate molecular mechanisms of resistance to ALK inhibitors
- To investigate detection of ALK mutations/fusions in plasma

Study Design

Description of Study

This is a randomized, active controlled, multicenter Phase III open-label study in patients with treatment-naive ALK-positive advanced NSCLC. All patients are required to provide pretreatment tumor tissue to confirm the presence of ALK rearrangement (by Ventana immunohistochemistry [IHC] testing). Patients will be randomized 1:1 into one of the two treatment arms to receive either alectinib or crizotinib.

This study will comprise approximately 180 centers, in around 30 countries worldwide.

The primary endpoint of the study is investigator-assessed PFS.

Appendix 1
Protocol Synopsis (cont.)

Central randomization will be performed via an interactive voice or web-based response system (IxRS) using the following stratification factors: Eastern Cooperative Oncology Group Performance Status (ECOG PS) (0/1 vs. 2), race (Asian vs. non-Asian), and CNS metastases at baseline (yes vs. no). An IxRS manual containing relevant information will be provided to each study site.

The experimental arm will receive alectinib at 600 mg orally twice daily (BID), taken with food. The control arm will receive crizotinib at 250 mg orally BID, taken with or without food. The first dose of the study drug should be administered as soon as possible after randomization, preferably within 24 hours, and no later than 48 hours after randomization.

Patients will be treated until disease progression, unacceptable toxicity, withdrawal of consent, or death. After disease progression (as per RECIST v1.1), patients should discontinue the study medication. After disease progression, patients will be treated at the discretion of the investigator according to local practice. Information regarding the nature and the duration of subsequent therapies will be collected.

In case of isolated asymptomatic CNS progression (e.g., new CNS oligometastases), a local therapy can be given (e.g., stereotactic radiotherapy or surgery) followed by continuation of either alectinib (in alectinib arm) or crizotinib (in crizotinib arm) until systemic disease progression and/or symptomatic CNS progression. The decision to continue the treatment beyond isolated, asymptomatic CNS progression is at the investigator's discretion for patients who can continue to benefit from respective treatment.

Patients who discontinue treatment prior to disease progression (e.g., due to unacceptable toxicity or withdrawal of consent) will continue to be followed until disease progression and for OS regardless of whether they subsequently receive non-study anti-cancer therapy. Data for subsequent therapy will be collected for the analysis of OS.

Number of Patients

Approximately 286 patients will be randomly assigned in a 1:1 allocation ratio to the two treatment arms via a block stratified randomization procedure and over a planned recruitment period of 24 months.

Randomization will guard against systematic selection bias and should ensure the comparability of treatment groups. To assist balance in important prognostic factors, randomization will be stratified by ECOG PS (0/1 vs. 2), race (Asian vs. non-Asian), and CNS metastases at baseline (yes vs. no).

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Histologically or cytologically confirmed diagnosis of advanced or recurrent (Stage IIIB not amenable for multimodality treatment) or metastatic (Stage IV) NSCLC that is ALK-positive as assessed by the Ventana IHC test. Sufficient tumor tissue to perform ALK IHC and ALK FISH is required. Both tests will be performed at designated central laboratories.
- Age $\geq$ 18 years old
- Life expectancy of at least 12 weeks
- ECOG PS of 0–2
- No prior systemic treatment for advanced or recurrent NSCLC (Stage IIIB not amenable for multimodality treatment) or metastatic (Stage IV) NSCLC
- Adequate hematologic function:

  Platelet count $\geq 100 \times 10^9$/L

  ANC $\geq$ 1500 cells/μL

  Hemoglobin $\geq$ 9.0 g/dL

Appendix 1
Protocol Synopsis (cont.)

- Adequate renal function:

  An estimated glomerular filtration rate (eGFR) calculated using the Modification of Diet in Renal Disease equation of at least 45 mL/min/1.73 m^2
- Patients must have recovered from effects of any major surgery or significant traumatic injury at least 28 days before the first dose of study treatment.
- Measurable disease (by RECIST v1.1) prior to the administration of study treatment
- Prior brain or leptomeningeal metastases allowed if asymptomatic (e.g., diagnosed incidentally at study baseline). Asymptomatic CNS lesions might be treated at the discretion of the investigator as per local clinical practice. If patients have neurological symptoms or signs due to CNS metastasis, patients need to complete whole brain radiation or gamma knife irradiation treatment. In all cases, radiation treatment must be completed at least 14 days before enrollment and patients must be clinically stable.
- For all females of childbearing potential, a negative pregnancy test must be obtained within 3 days before starting study treatment.
- For women who are not postmenopausal ($\geq$ 12 months of non-therapy-induced amenorrhea) or surgically sterile (absence of ovaries and/or uterus), agreement to remain abstinent or use single or combined contraceptive methods that result in a failure rate of $<$ 1% per year during the treatment period and for at least 3 months after the last dose of study drug. Abstinence is only acceptable if it is in line with the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception. Examples of contraceptive methods with a failure rate of $<$ 1% per year include tubal ligation, male sterilization, hormonal implants, established and proper use of combined oral or injected hormonal contraceptives, and certain intrauterine devices. Alternatively, two methods (e.g., two barrier methods such as condom and cervical cap use) may be combined to achieve a failure rate $<$ 1% per year. Barrier methods must always be supplemented with the use of a spermicide.
- For men, agreement to remain abstinent or use a condom plus an additional contraceptive method that together result in a failure rate of $<$ 1% per year during the treatment period and for at least 3 months after the last dose of study drug. Abstinence is only acceptable if it is in line with the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.
- Able and willing to provide written informed consent prior to performing any study-related procedures and to comply with the study protocol, including patients must be willing and able to use the electronic patient-reported outcome (ePRO) device.

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study entry:

- Patients with a previous malignancy within the past 3 years are excluded (other than curatively treated basal cell carcinoma of the skin, early gastrointestinal (GI) cancer by endoscopic resection, in situ carcinoma of the cervix, or any cured cancer that is considered to have no impact in PFS and OS for the current NSCLC).
- Any GI disorder that may affect absorption of oral medications, such as mal-absorption syndrome or status post-major bowel resection

Appendix 1
Protocol Synopsis (cont.)

- Liver disease characterized by:

  ALT or AST $>3\times$ULN ($\geq 5\times$ULN for patients with concurrent liver metastasis) confirmed on two consecutive measurements

  OR

  Impaired excretory function (e.g., hyperbilirubinemia) or synthetic function or other conditions of decompensated liver disease such as coagulopathy, hepatic encephalopathy, hypoalbuminemia, ascites, and bleeding from esophageal varices

  OR

  Acute viral or active autoimmune, alcoholic, or other types of acute hepatitis
- National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.0) Grade 3 or higher toxicities due to any prior therapy such as radiotherapy (excluding alopecia), which have not shown improvement and are strictly considered to interfere with current study medication
- History of organ transplant
- Co-administration of anti-cancer therapies other than those administered in this study
- Patients with baseline QTc$>$470 ms or symptomatic bradycardia
- Administration of strong/potent cytochrome P450 (CYP) 3A inhibitors or inducers within 14 days prior to the first dose of study treatment and while on treatment with alectinib or crizotinib
- Administration of agents with potential QT interval prolonging effects within 14 days prior to the first administration of study drug for all patients and while on treatment through the end of the study for crizotinib-treated patients only (see protocol for details)
- History of hypersensitivity to any of the additives in the alectinib drug formulation (see protocol for details)
- History of hypersensitivity to any of the additives in the crizotinib drug formulation (see protocol for details)
- Pregnant or lactating women
- Known HIV positivity or AIDS-related illness
- Any clinically significant concomitant disease or condition that could interfere with, or for which the treatment might interfere with, the conduct of the study or the absorption of oral medications or that would, in the opinion of the Principal Investigator, pose an unacceptable risk to the patient in this study
- Any psychological, familial, sociological, or geographical condition potentially hampering compliance with the study protocol requirements and/or follow-up procedures; those conditions should be discussed with the patient before trial entry

Length of Study

The time from first patient screened to end of study, defined below, will be approximately 43 months.

End of Study

This study is event driven, with a recruitment period of approximately 24 months. The required number of events for the final analysis of the primary endpoint is expected approximately 33 months after the first patient has been enrolled. Patients are to be treated until disease progression, unacceptable toxicity, withdrawal of consent, or death, whichever occurs first. Follow-up for survival will continue until the survival follow-up analysis or the Sponsor decides to end the trial, whichever occurs first. The study will formally end once the survival follow-up analysis is complete.

Appendix 1
Protocol Synopsis (cont.)

A survival follow-up analysis will be performed once approximately 50% of patients (i.e., 143 patients) have died, which is estimated to occur approximately 42 months after the first patient has been enrolled.

Outcome Measures

Efficacy Outcome Measures

The efficacy outcome measures for this study are as follows:

- PFS, which is defined as the time from randomization to the first documented disease progression, as determined by the investigators (primary endpoint) or IRC (secondary endpoint) using RECIST v1.1 or death from any cause, whichever occurs first. Patients without an event will be censored at the last tumor assessment either during follow-up or during study treatment. Patients with no post-baseline assessments will be censored at the date of randomization.
- ORR, which is defined as the percentage of patients who attain complete response (CR) or partial response (PR); response, as determined by the investigators using RECIST v1.1. Patients without any assessments will be regarded as non-responders.
- Time to CNS progression, which is defined as the time from randomization to the first occurrence of disease progression in the CNS as determined by IRC using RECIST v1.1 and RANO (separate assessments and analyses), as well as C-ORR in patients with CNS metastases who have measurable disease in the CNS at baseline, C-DOR in patients who have a CNS Objective Response, and C-PR at 6, 12, 18, and 24 months
- DOR, which is defined as the time from when response (CR or PR) was first documented to first documented disease progression or death (whichever occurs first). This will only be calculated for patients who have a best overall response of CR or PR. Patients who do not progress or die after they have had a response are censored at the date of their last tumor measurement.
- OS, which is defined as the time from randomization to death from any cause. Patients without an event will be censored at the last date known to be alive. Patients without any follow-up information will be censored at the date of randomization.

Safety Outcome Measures

The secondary safety outcome measures for this study are as follows:

- Serious and non-serious adverse events
- Safety laboratory tests values
- Vital signs (blood pressure, heart rate), ECG
- Physical examination

Pharmacokinetic Outcome Measures

The PK outcome measures for this study are as follows:

- Sparse (pre-dose) PK samples for measurement of alectinib and its major metabolite(s) will be collected in all study patients receiving alectinib treatment
- Serial/intensive PK sampling will be collected in a subset of consenting patients enrolled to receive alectinib treatment (approximately 10%−15%, at least approximately n = 20)
- PK parameters will be determined as appropriate and where data allow:

  The pharmacokinetics of alectinib (and metabolite[s], if appropriate) will be described, and the between-patient variability will be estimated using a population PK approach. The potential influence of covariates that contribute significantly to the between-patient differences in PK parameters of alectinib will also be explored and quantified.

  Non-compartmental analysis may be conducted in patients undergoing serial/intensive PK sample collection, as appropriate and where data allow.

Appendix 1
Protocol Synopsis (cont.)

Patient-Reported Outcome Measures

The PRO measures for this study are as follows:

- EORTC QLQ-C30 and the EORTC QLQ-LC13 to determine the impact of alectinib compared with crizotinib as measured by TTD in patient-reported lung cancer symptoms (e.g., cough, dyspnea [single item and multi-item scales], pain in chest, pain in arm/shoulder, fatigue)
- The EORTC QLQ-C30 and EORTC QLQ-LC13 to measure PROs of HRQoL, patient functioning, and side effects of therapy compared between patients treated with alectinib and those treated with crizotinib

Exploratory Outcome Measures

The exploratory outcome measures for this study are as follows:

- EQ-5D-3L to generate utility scores for use in economic models for the purpose of reimbursement
- Total testosterone and free testosterone levels (either by direct measurement or by calculation using albumin and SHBG), FSH, and LH in blood to measure an onset of hypogonadism in adult men
- Results from the FISH Vysis® ALK Break Apart FISH Probe Kit (Abbott) to evaluate and compare efficacy in patients with treatment-naive NSCLC that is ALK-positive by FISH test
- ALK fusion status in circulating tumor nucleic acids from plasma to evaluate and compare efficacy and safety in patients with treatment-naive NSCLC that is ALK-positive by plasma ALK tests (PCR and/or sequencing) for diagnostic purposes
- Post-progression tumor mutation status to study molecular mechanisms of resistance to ALK inhibitors
- ALK mutation status in plasma DNA to monitor efficacy and disease progression

Investigational Medicinal Products

Test Product

Alectinib comes in a capsule dosage form containing the following active ingredient:

[Chemical name] 9-Ethyl-6, 6-dimethyl-8-[4-(morpholin-4-yl) piperidin-1-yl]-11-oxo-6, 11-dihydro-5H-benzo[b]carbazole-3-carbonitrile hydrochloride

Each capsule contains 150 mg of alectinib (as free base) along with lactose monohydrate, carmellose calcium, hydroxypropyl cellulose, SLS, and magnesium stearate.

Alectinib capsules should be stored in accordance with the storage instructions on the label. Alectinib capsules should be administered orally BID with food in the morning and evening.

Comparator

Crizotinib comes in a hard capsule dosage form. Each capsule contains 250 mg or 200 mg crizotinib. Crizotinib hard capsules should be stored in accordance with the storage instructions on the label. Crizotinib capsules should be administered orally BID.

For further details, see the local prescribing information for crizotinib (Xalkori®).

Non-Investigational Medicinal Products

Standard-of-care therapy administered to all patients in addition to test product or comparator.

Statistical Methods

Primary Analysis

PFS is defined as the time from date of randomization to the date of first documented disease progression or death, whichever occurs first. The primary endpoint of PFS will be determined on the basis of investigator assessment of progression using RECIST v1.1. Patients who have not experienced disease progression or death at the time of analysis will be censored at the last

tumor assessment date either during study treatment or during follow-up. Patients with no post-baseline tumor assessment will be censored at the date of randomization.

Patients who discontinue treatment prior to disease progression (e.g., due to toxicity) will continue on study and will be followed until disease progression and for OS regardless of whether they subsequently receive anti-cancer therapy.

The treatment comparison of PFS will be based on a stratified log-rank test at the 5% level of significance (two-sided). The stratification factors are the randomization stratification factors: ECOG PS (0/1 vs. 2), race (Asian vs. non-Asian), and CNS metastases at baseline (yes vs. no), as recorded on the eCRF. Results from an unstratified log-rank test will also be presented.

The Kaplan-Meier method will be used to estimate the median PFS for each treatment arm with 95% confidence limits, and a Kaplan-Meier curve will be constructed to provide a visual description of the difference between the treatment arms. A stratified Cox proportional regression model will be used including treatment in order to provide an estimate of the treatment effect expressed as a hazard ratio (HR) (alectinib vs. crizotinib), as well as a 95% CI.

The difference between the two treatment groups will be assessed and tested for the following hypothesis: the survival distribution function (SDF) of the alectinib treatment group is the same as for the crizotinib treatment group versus the alternative that the two distributions are different:

H0: SDF (alectinib) = SDF (crizotinib)
versus
H1: SDF (alectinib) ≠ SDF (crizotinib),

where SDF denotes the survival distribution function of the parameter PFS.

Determination of Sample Size

The focus of this clinical trial is hypothesis testing, and the primary endpoint of PFS was used to determine the sample size of the study.

At the time of initially writing the protocol, there were no Phase III data available on crizotinib in the first-line setting in ALK-positive patients. The median PFS for crizotinib administered as second-line therapy from the global randomized Phase III PROFILE 1007 study was 7.7 months (95% CI: 6.0%, 8.8%). From the Phase II single arm study of patients who had received ≥ 1 line of chemotherapy (with 85% patients having received ≥ 2 prior chemotherapy regimens), the median PFS was 8.5 months (95% CI: 6.2%, 9.9%). The Phase III PROFILE 1014 study of crizotinib versus standard pemetrexed–platinum-based chemotherapy in previously untreated patients with ALK-positive non-squamous NSCLC reported a median PFS of 10.9 months for crizotinib.

Thus, an initial assumption of median PFS of 9.8 months for the crizotinib arm has been reassessed based on data from the Phase III study PROFILE 1014 to 10.9 months. An HR of 0.65 for alectinib versus crizotinib (i.e., an increase from 10.9 months median PFS to 16.8 months) will be targeted.

In this study, 286 patients will be enrolled in a 1:1 randomization allocation. Enrollment will take approximately 24 months on the basis of an assumption of non-linear recruitment as follows:

- Month 1: 1 patient per month
- Month 2: 2 patients per month
- Month 3: 4 patients per month
- Month 4: 6 patients per month
- Month 5: 8 patients per month
- Month 6: 10 patients per month
- Month 7: 12 patients per month
- Months 8–12: 13 patients per month
- Months 13–14: 14 patients per month
- Month 15 onwards: 15 patients per month

Appendix 1
Protocol Synopsis (cont.)

A total of 170 PFS events are required to achieve 80% power at a two-sided alpha level of 5%. This number of PFS events is estimated to occur approximately 33 months after the first patient has been enrolled.

Interim Analyses

No interim analysis for efficacy or futility is planned.