Statistical Analysis Plan

Study Code D5164C00001

Edition Number 4.0

A Phase III, Double-blind, Randomised, Placebo-Controlled Multi-centre, study to assess the efficacy and safety of AZD9291 versus Placebo, in Patients with Epidermal Growth Factor Receptor Mutation Positive Stage IB-IIIA Non-small Cell Lung Carcinoma, following Complete Tumour Resection With or Without Adjuvant Chemotherapy (ADAURA)

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE(s)	Adverse event(s)
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
B	Blood
BP	Bodily pain
BPr	Blood pressure
BUN	Blood urea nitrogen
CI(s)	Confidence interval(s)
cMET	Proto-oncogene encoding Hepatocyte Growth Factor Receptor
CNS	Central Nervous System
CSR	Clinical Study Report
CT	Computer tomography
CTCAE	Common Terminology Criteria for Adverse Event
ctDNA	Circulating tumour deoxyribonucleic acid
CV	Coefficient of variation
DCO	Data cut off
DFS	Disease free survival
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
eCRF	Electronic case report form
EGFR	Epidermal Growth Factor Receptor
Ex19del	Deletions in exon 19
FAS	Full Analysis Set
GH	General health perceptions
Gmean	Geometric mean
HbA1C	Haemoglobin A1C
HDU	High dependency unit

Abbreviation or special term	Explanation
HER2	Human Epidermal Growth Factor Receptor 2
HR	Hazard ratio
HRQoL	Health Related Quality of Life
IP	Investigational Product
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
IPCW	Inverse probability of censoring weighting
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
L858R	Exon 21
LDH	Lactate dehydrogenase
LLN	Lower limit of normal
LLoQ	Lower limit of quantification
LOQ	Limit of quantification
LVEF	Left Ventricular Ejection Fraction
MCS	Mental component summary
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mental health
MRI	Magnetic resonance imaging
NC	Not calculable
NPV	Negative predictive value
NQ	Non-quantifiable
NSCLC	Non-small Cell Lung Cancer
OAE(s)	Other significant adverse event(s)
OS	Overall Survival
P	Plasma
PCS	Physical component summary
PF	Physical functioning
PFS	Progression-free survival

Abbreviation or special term	Explanation
PH	Proportional Hazards
PK	Pharmacokinetics
PPV	Positive predictive value
PRO(s)	Patient reported outcome(s)
PT(s)	Preferred term(s)
QRS complex	Name for the combination of three of the graphical deflections seen on the electrocardiogram
QT	Interval on the electrocardiogram representing the duration of depolarization and repolarization of the heart
QTc	The QT interval corrected for heart rate
QTcF	Fridericia QTc correction
RBC	Red blood cells
RE	Role limitations-emotional
RP	Role limitations-physical
RPSFT	Rank preserving structural failure time
RR	Interval on the electrocardiogram representing the time from the onset of one QRS complex to the onset of the next QRS complex
S	Serum
SAE(s)	Serious adverse event(s)
SF	Social function
SF-36 v2	Short form 36 health survey version 2
SOC	System organ class
StdDev	Standard deviation
T790M	An amino acid substitution at position 790 in EGFR, from a threonine to a methionine
TFST	Time to first subsequent therapy
TKI	Tyrosine kinase inhibitor
t_{max}	The time after administration of a drug when the maximum plasma concentration is reached; when the rate of absorption equals the rate of elimination
TSST	Time to second subsequent therapy or death

Abbreviation or special term	Explanation
U	Urine
ULoQ	Upper limit of quantification
VT	Vitality
WHO	World Health Organisation

AMENDMENT HISTORY

Date	Brief description of change
23 June 2020	Section 7: revision to section to reflect new multiple testing procedure
28 April 2020	Section 3.1.5: added text to provide clarity on handling missing visits regarding the SF-36V2 Health Survey Section 4.2.1: added text to allow for data collected on the day of randomisation to be used as baseline Section 4.3.7.1: clarified text regarding sensitivity analysis to support variable derivation Section 4.3.7.3: inclusion of additional analysis of time to new brain lesion or death Section 7: new section added to reflect new multiple testing procedure
18 December 2018	Minor grammatical updates, correction of typos and formatting issues such as missing page numbers have been updated. Updates mainly for clarity or to provide missing information. Specific changes: Following abbreviations and special terms added to the list of abbreviations: CNS, IP, PH, RR, QRS complex. Section 1.1 Outcome measures clarified to ensure alignment between the SAP and the clinical study protocol version 2. Section 1.2 and Section 1.3 Updates to ensure alignment with the protocol including Figure 1 replaced with the Figure presented in the clinical study protocol. If more mature data on Overall Survival is required, then it may be analysed at a later timepoint. Any exploratory objectives not described in this SAP will be reported separately to the CSR. Section 2.1 Analysis set definitions clarified for full, safety, and PK analysis sets. PK analysis set is defined using safety analysis set instead of FAS originally specified in the protocol. Section 2.2 Protocol deviations expanded to define the protocol deviations considered important for CSR. Section 3.1.1 Study day definitions updated to clarify that all full analysis set summaries are relative to randomisation date, and all safety analysis set summaries are relative to first dose.

Section 3.1.3 Information added on handling partial dates for deaths and diagnostics.

Section 3.2.1 Further information on censoring added.

Section 3.2.3 Further information on survival calls added and eCRF pages to use. Information on handling partial dates added.

Section 3.2.4 Further detail of the derivation of PFS (exploratory) added.

Section 3.2.5 Further detail of the derivation of TFST added.

Section 3.2.6 Further detail of the derivation of TSST added.

Section 3.2.7 Details provided on how brain metastases are identified in the CRF.

Section 3.2.8.1 Clarification that SF-36 v2 (standard) is used as stated in the protocol. Time to deterioration definition clarified as confirmed worsening. Added in description of SF-36 compliance.

Section 3.2.9 Information on how health care resource use is captured and how the length of stay and missing discharge dates are calculated has been clarified.

Section 3.3.4 Additional section added to clarify how the adverse events of special interest are to be identified.

Section 3.3.5 Clarification on how missed or forgotten doses should be recorded in eCRF and what dates to use for patients who permanently discontinued during a dose interruption has been added. Added total safety follow-up calculation.

Section 3.3.6 Laboratory safety variables corrected within Table 4. Derivation of corrected calcium and creatinine clearance added.

Section 3.3.7 Detail added about how to handle replicate ECG measurements. RR within the ECG QTcF formula defined.

Section 4 For the primary and any subsequent analyses, all safety data will only be summarised until start of subsequent cancer therapy.

Section 4.1 Updated in line with other changes.

Section 4.2.1 Baseline measurements clarified for Full and Safety analysis sets.

Section 4.2.2 Definition of significantly less than 70 DFS events added.

Section 4.3.1 Further detail added on patient disposition.

Section 4.3.3 Following categorical variables were added to the list of demographic and other baseline characteristics: recruitment by country and centre, disease characteristics at baseline. Following categorical variables were removed as not captured at baseline: Extent of disease upon entry to study (metastatic, locally advanced, both), Site of local/metastatic disease, brain metastases and Visceral metastases. Additional information was added to clarify that stratification factors

will be summarised as recorded in IVRS (as well as eCRF) at the time of randomisation
Section 4.3.5 WHO drug dictionary is used in place of AstraZeneca drug dictionary. Prior, Concomitant and Post IP medications and anti-cancer therapies definitions clarified. Details of adjuvant disease-related treatment modalities and adjuvant chemotherapy by stage (IB, II, IIIA) summaries added. Additional information added to clarify that both disallowed and allowed concomitant medications will be summarised.
Section 4.3.6 Analysis set to be used clarified.
Section 4.3.7.1 As supportive information, summary tables added of treatment status at disease recurrence or death, censoring relative to DCO and disease recurrence status at time of DCO. Additional details provided for the proportionality assessments and subgroup analyses for DFS analyses. Subgroup analysis for age updated to show it will be performed using age at screening, not baseline. Region will not be included in quantitative interactions. 96% CIs will also be produced. A summary of DFS by stage and adjuvant therapy removed as included in subgroup analyses.
Section 4.3.7.2 DFS rate detail moved here. DFS rate will also be summarised at 6, 12, 18 and 48 months. As supportive information, summary table added of survival status at DCO. Timepoints of OS analyses added to estimate OS rate at 2, 3, as well as 5 years. OS rate will be estimated for non-IBs and overall. An OS analysis in certain centres will not be performed. Compliance with SF-36 will be summarised by visit.
Section 4.3.7.3 Clarification provided for healthcare resource use. Information on analysis of PFS has been moved here. Information added to provide clarification how PFS, TFST and TSST will be summarised. Analysis of Plasma-Derived ctDNA EGFR Mutation Status at Baseline Compared to Status at Disease Recurrence will not be produced for all patients in the FAS.
Section 4.3.8.1 Text from Section 3.3 related to AE summaries moved here.
Section 4.3.8.2 Life table and prevalence plot will also be produced for stomatitis.
Section 4.3.8.3 Summary of deaths clarified that it will be produced for the FAS.
Section 4.3.8.4 Information on how exposure will be summarised has been moved here. Detail added about duration of exposure summaries.
Section 4.3.8.9 Clarification provided that percentages will only be calculated where there is at least one baseline and one post-baseline measurement.
Section 4.3.8.10 Details of WHO performance status descriptive summary clarified.

	Section 4.3.9 Text on 'Analysis of PK Concentration Data' from section 4.3.7.2 moved here. Clarifications provided for the statistical parameters used in the analysis of PK concentration calculations.

1. STUDY DETAILS

1.1. Study Objectives

Primary Objective:	**Outcome Measure:**
To assess the efficacy of AZD9291 compared to placebo as measured by disease free survival (DFS).	DFS by investigator assessment.

Secondary Objectives:	**Outcome Measure:**
To further assess the efficacy of AZD9291 compared with placebo.	- DFS rate at 2, 3 and 5 years. - Overall Survival (OS). - OS rate at 5 years.
To assess the effect of AZD9291 compared with placebo on health-related quality of life (HRQoL).	Changes in generic HRQoL as measured by the SF-36 (version 2, standard).
To characterise the pharmacokinetics (PK) of AZD9291 and its metabolites (AZ5104 and AZ7550).	PK plasma concentrations of AZD9291, and metabolites AZ5104 and AZ7550; and ratio of metabolite to AZD9291 for each PK sample (included in the clinical study report [CSR]). PK data from this study will be analysed using a population PK approach and reported separately to the Clinical Study Report (CSR). Data from this study may form part of a pooled analysis with data from other studies

Safety Objective:	**Outcome Measure:**
To assess the safety and tolerability profile of AZD9291 compared with placebo.	- Adverse events (graded by Common Terminology Criteria for Adverse Event [CTCAE] v4) - Clinical chemistry, haematology and urinalysis - Vital signs, Physical Examination, Weight - Digital electrocardiogram (ECG) - Left ventricular ejection fraction (LVEF) - World Health Organisation (WHO) Performance Status - Ophthalmologic assessment

Exploratory Objective*:	**Outcome Measure:**
To compare health resource use associated with AZD9291 treatment versus placebo	Health Resource Use Module
To compare the effects of AZD9291 with placebo on post recurrence outcomes.	- Time to next treatment(s) - Type of recurrence (local/regional or distant) - Site(s) of recurrence - Type of next treatment(s) (including procedures, radiotherapy and anticancer agents) - Progression-free survival (PFS) as determined by investigator assessment
To assess the efficacy of AZD9291 in patients with confirmed baseline amino acid substitution at position 790 in epidermal growth factor receptor (EGFR), from a threonine to a methionine (T790M) status (positive/negative) using a high sensitivity method yet to be determined (retrospective)	- DFS by investigator assessment - OS

* Results from such exploratory analyses may be reported separately from the CSR.

Exploratory Objective*:	**Outcome Measure:**
To collect and store biopsy material (multiple cores where possible) from all screened patients for exploratory analysis of molecular mechanisms associated with development of non-small cell lung cancer (NSCLC) and response to treatment.	Key genetic and proteomic markers to include, but not limited to, EGFR mutations, human epidermal growth factor receptor 2 (HER2), and proto-oncogene encoding Hepatocyte Growth Factor Receptor (cMET) expression and/or amplification. Data generated may be reported separately and may also form part of a pooled analysis with other AZD9291 studies.
To collect and store deoxyribonucleic acid (DNA) for future exploratory research into genes/genetic variation that may influence response to AZD9291 (safety and efficacy) and/or susceptibility to/development of cancers.	Correlation of polymorphisms with variation in pharmacodynamics, safety or response observed in patients treated with AZD9291 or comparator.
To compare the baseline tumour EGFR mutation status in all randomised patients with evaluable results from baseline plasma.	Comparison of EGFR mutation status between tumour deoxyribonucleic acid (DNA) and plasma derived ctDNA.
To compare plasma-derived ctDNA EGFR mutation status at baseline and at disease recurrence.	Comparison of EGFR mutation status in plasma samples at baseline and at disease recurrence.
To assess any changes in EGFR mutation status (including T790M) using the mandated serial plasma samples coupled with a high sensitivity method yet to be determined (retrospective).	Assessment of EGFR mutation status in serial plasma samples.
To explore the relationship between PK and selected endpoints (which may include efficacy, safety, and/or patient reported outcomes [PROs]), where deemed appropriate.	Correlation of PK with other primary/secondary/exploratory endpoints in patients treated with AZD9291.

1.2. Study Design

This is a phase III, double-blind, randomised, placebo-controlled study to assess the efficacy and safety of AZD9291 versus placebo in patients with stage IB-IIIA non-squamous, NSCLC with a centrally confirmed, common sensitising EGFR mutations (deletions in exon 19

[Ex19del] and exon 21 [L858R] either alone or in combination with other EGFR mutations) as confirmed by a central test, who have had complete tumour resection, with or without post-operative adjuvant chemotherapy.

Patients will be randomised 1:1 to receive either AZD9291 or placebo. Patients must have sufficiently recovered from surgery and completed any standard of care adjuvant chemotherapy, if applicable, prior to randomisation. Patients must be randomised within 10 weeks of complete surgical resection if adjuvant chemotherapy was not administered and within 26 weeks if adjuvant chemotherapy was administered.

Accounting for patients who are found to have wild type EGFR, sample attrition and 10% screen fail rate for other reasons, it is estimated that 3200 patients will be screened to randomise approximately 700 patients. It is assumed that approximately 60% of patients will be recruited from Asia and 40% from non-Asian countries. All patients must be staged according to the 7th edition of AJCC Cancer Staging Manual, 2010. The proportion of patients randomised with stage IB cancer will be approximately 30% and the proportion of patients randomised with Stage II-IIIA cancer will be approximately 70%. Patients will be stratified by stage (IB vs II vs IIIa), mutation type status as confirmed by a central laboratory using a tissue based test (Ex19Del/L858R either alone or in combination with other EGFR mutations as confirmed by a central test), and race (Asian/non-Asian).

Following complete resection and prior to treatment initiation, all patients will be required to have a baseline computer tomography (CT) scan (chest and abdomen including liver and adrenal glands) within 28 days of study treatment initiation to confirm that disease is not present. Patients will continue on randomised treatment until recurrence of disease, treatment discontinuation criterion is met, or treatment is completed. The treatment duration period is three years. Patients being on study treatment at the time of DFS primary analysis may continue treatment until discontinuation criterion is met or three years treatment is completed. The study database will be locked at this point. A further analysis will be performed if there are significantly less than 70 DFS events in the IB population at the time of the primary analysis. If more mature data on Overall Survival is required, then it may be analysed at a later timepoint. Once the study is closed serious adverse events (SAEs) will only be reported as per Section 6.4 of the protocol.

Patients will undergo safety assessments at baseline, two weeks, four weeks, 12 weeks and every 12 weeks until treatment is completed or discontinued. All study patients must have a 28 day follow up visit after treatment is stopped.

Patients will undergo regular radiological assessments for disease recurrence at 12 weeks, 24 weeks and then every 24 weeks until five years, then yearly thereafter. DFS shall be measured from the day of randomisation until date of recurrence or death (by any cause) in the absence of recurrence by investigational site assessment.

Patients who discontinue treatment prior to disease recurrence will continue to be followed for DFS according to study plan. Following disease recurrence, patients will undergo radiological imaging for subsequent progression in accordance with local clinical practice and will be

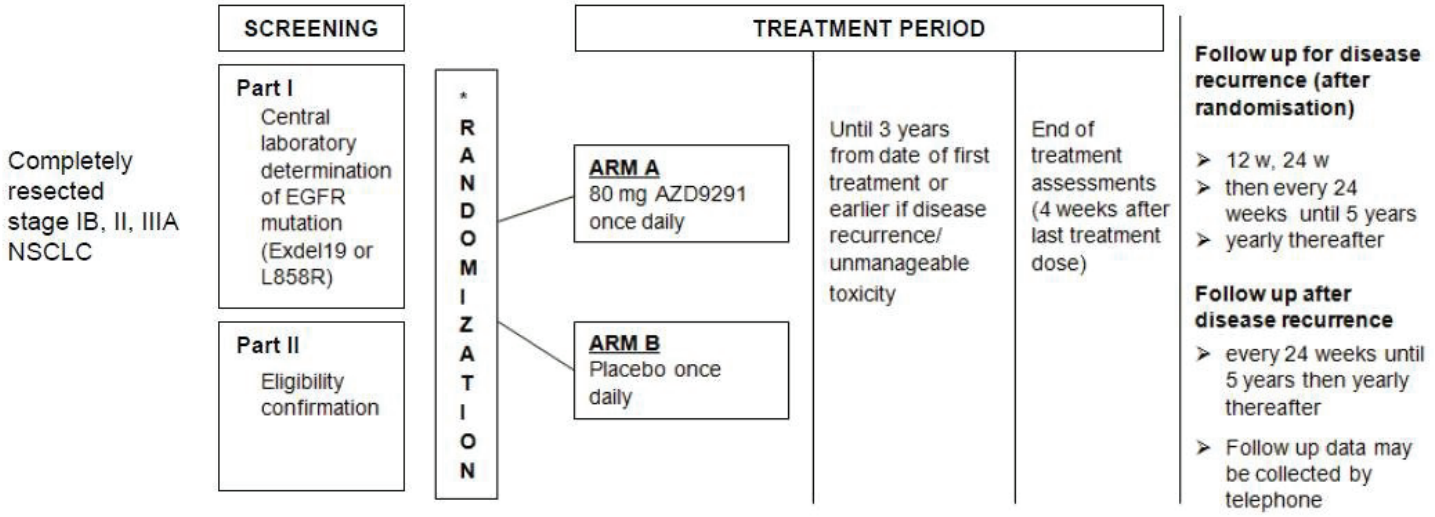
followed for survival every six months until five years post randomisation, then yearly thereafter. Follow up will continue until the study is closed.

At disease recurrence, patients will be restaged according to the 7th edition of AJCC Cancer Staging Manual, 2010, and all sites of NSCLC relapse will be recorded. Treatments received by the patient after relapse will be determined by the physician. Post recurrence cancer treatments and procedures will be recorded.

Tumour and blood sampling for biomarker, translational science and PK will also be performed. Any exploratory objectives not described in this SAP will be reported separately to the CSR.

The overall study design is shown in Figure 1 below. The study schedule is detailed in Section 4 of the protocol.

Figure 1 Study Design



***Stratification:**
– Stage (IB vs. II vs IIIA)
– EGFR mutation (Exdel19 or L858R)
– Race (Asian/non-Asian)

Randomization will be 1:1

1.3. Number of Patients

Approximately 700 patients will be randomised globally in a 1:1 ratio (AZD9291: placebo) to this study. The primary endpoint of the study is DFS based on investigator assessment. DFS will first be tested in the subset of patients with Stage IIA-IIIA cancer. If statistical significance is shown for DFS in the subset of patients with Stage IIA-IIIA cancer, then it will be tested for the overall population with the test mass split between first and second analyses.

The primary analysis of DFS will occur when approximately 247 disease recurrence events have been observed in approximately 490 patients who are in Stage IIA-IIIA (i.e. non-IB). If the true DFS hazard ratio (HR) for the comparison of AZD9291 versus placebo in this patient population is 0.70, 247 disease recurrence events will provide 80% power to demonstrate a statistically significant difference in DFS at a 5% two-sided significance level (translating to an improvement in median DFS from 40 months to 57 months, assuming DFS is exponentially distributed). The minimum DFS HR that would be statistically significant (i.e. $p<0.05$, two-sided) is 0.78.

If the true DFS HR for the comparison of AZD9291 versus placebo in the overall population is 0.70, 314 disease recurrence events will provide approximately 90% power to demonstrate a statistically significant difference in DFS at a 4% two-sided significance level (translating to an improvement in median DFS from 46 months to 66 months, assuming DFS is exponentially distributed). The minimum DFS HR that would be statistically significant (i.e. $p<0.04$, two-sided) is 0.79.

At the time of the primary analysis it is anticipated that approximately 195 OS events (28% maturity) will have occurred (assuming a median OS of 96 months for the placebo arm) in the overall population. If the true OS HR for the comparison of AZD9291 versus placebo is 0.66, 195 death events provide approximately 80% power to demonstrate a statistically significant difference at the 4% (two-sided) significance level. With 195 death events and assuming a true OS HR of 0.85 there is approximately 90% chance of observing an HR < 1.02.

Assuming 28 months non-linear recruitment the data cut-off for the primary analysis is estimated to occur approximately 68-70 months after first patient randomised.

It is estimated that 3200 patients will be screened in order to randomise 700 patients.

2. ANALYSIS SETS

2.1. Definition of Analysis Sets

2.1.1. Full Analysis Set

The full analysis set (FAS) will include all randomised patients.

The FAS will be used for all demographic summaries and efficacy analyses, and treatment groups will be compared on the basis of randomised study treatment, regardless of the treatment actually received.

2.1.2. Safety Analysis Set

The safety analysis set will consist of all patients who received at least one dose of study treatment.

Safety data will not be formally analysed but summarised using the safety analysis set, according to the treatment received; i.e. erroneously treated patients (e.g. those randomised to treatment A but actually given treatment B) will be summarised according to the treatment they actually received. If a patient received both treatments, then they will be summarised according to the active treatment.

2.1.3. Pharmacokinetic Analysis Set

The PK analysis set is defined as patients in the safety analysis set who received AZD9291 and have at least one measurable PK concentration, supported by the relevant date and time of this sample. For each time a PK sample was taken from a patient, the dosing data for that day and for multiple dosing, the dose date for the two days prior to the sample days must be available. For any individual sample from a patient to be included in the PK analysis set, the full sample data and dosing data need to be present for that sample/patient.

The pharmacokineticist will agree to the strategy for dealing with data affected by protocol deviations before any formal statistical analysis is performed. Protocol deviations may include changes to the procedures that may impact the quality of the data or any circumstances that can alter the evaluation of the PK. Examples may include but are not limited to: vomiting following oral dosing occurring within the timeframe of 2 times the median time after administration of a drug when the maximum plasma concentration is reached when the rate of absorption equals the rate of elimination (t_{max}); sample processing errors that lead to inaccurate bioanalytical results; incomplete dose administered; incomplete PK profile collected; and/or use of disallowed concomitant medication. In the case of a relevant protocol deviation or event, affected PK data collected will be excluded from the summaries and statistical analyses, but will still be reported in the study result listings. Any exclusions will be detailed in the CSR.

2.2. Protocol Violations and Deviations

Important protocol deviations are defined as those deviations from the protocol likely to have an impact on the perceived efficacy and/or safety of study treatments.

A list of important protocol deviations and any action to be taken regarding the exclusion of patients or affected data from specific analyses are defined in Table 1 below and a list of deviations that are regarded as important are in Table 2 below.

Note that the contents of these tables are not an exhaustive list. A complete list of anticipated protocol deviations (including important protocol deviations) will be compiled separately and finalised prior to unblinding. Note a longer list of important protocol deviations may be used for the IDMCs.

Table 1 Protocol Deviations with Action to be Taken for Analysis

Protocol Deviation	Action to be Taken for Analysis
Patient randomised and did not receive any study medication	Exclude from the Safety analysis set
Patient was given incorrect study medication	Analyse "As-randomised" for the FAS. Analyse "As-treated" for the Safety analysis set. If only the incorrect study medication was taken for the whole study, then analyse according to treatment taken. If patient received both AZD9291 and Placebo then analyse according to active treatment.

Table 2 Important Protocol Deviations

Criteria Type	Important Deviation Description
Inclusion Criteria	• Lack of Histologically confirmed diagnosis of primary non-small lung cancer (NSCLC) on predominantly non-squamous histology. • MRI or CT scan of the Brain was not done before randomization. • Non-Japanese Sites' patients with pathology at diagnosis: Post Surgery the AJCC Stage (7th edition) for the patient is not Stage IB, II or IIIA. • Japanese Sites' patients with pathology at Diagnosis: Post Surgery the AJCC Stage for the patient is not Stage II or IIIA. • No confirmation that the tumour harbours an EGFR mutation known to be associated with EGFR TKI (tyrosine kinase inhibitor) sensitivity (including Ex19del and L858R). • Not complete surgical resection of the primary NSCLC, and all surgical margins of resection not negative for tumour.

Exclusion Criteria	• Any anti-cancer therapy as follows:
	  • Prior treatment with neoadjuvant or adjuvant EGFR-TKI.
	  • Pre-operative (neo-adjuvant) platinum based or other chemotherapy.

Table 2 Important Protocol Deviations

Criteria Type	Important Deviation Description
	• Pre-operative or post-operative or planned radiation therapy for the current lung cancer. • Any other prior anticancer therapy, including investigational therapy, for treatment of NSCLC other than standard platinum based doublet post-operative adjuvant chemotherapy. • Treatment with an investigational drug within five half-lives of the compound or any of its related material, if known. • Past medical history of ILD, drug-induced ILD, radiation pneumonitis which required steroid treatment, or any evidence of clinically active ILD.
IP Administration/ Study Treatment	• Patient was randomized but did not receive any treatment • Patient received medication in error e.g. patient dosed but not fully eligible for study treatment or patient received the incorrect treatment e.g. did not receive the treatment they were randomised to. • Patients with corneal ulceration and/or ILD who resume study treatment.
Procedures/Tests	• Baseline disease assessment performed more than 28 days prior randomization i.e. not performed during screening window. • Disease Recurrence scans not performed in accordance with the protocol time schedule RELATIVE to RANDOMISATON date or not at all.
Disallowed Medications	• Patient receiving anti-cancer therapies (e.g. anticancer drugs, radiotherapy) or other concomitant strong CYP3A inducers or herbal medications which are strong CYP3A inducers during the study, other than IP.

The following summary will be provided:

- A summary of the number and percentage of patients with an important protocol deviation by treatment group and overall and by type of deviation (Analysis set: FAS)

A by-patient listing of important protocol deviations will be provided in the CSR.

3. PRIMARY, SECONDARY AND EXPLORATORY VARIABLES

3.1. General Variables

3.1.1. Study Day Definitions

For the purpose of data summaries based on the full analysis set, for example efficacy, Study Day 1 is defined as the date of randomisation to study treatment. For visits (or events) that occur on or after randomisation, study day is defined as:

$$\text{date of visit (or event)} - \text{date of randomisation} + 1.$$

For visits (or events) that occur prior to randomisation, study day is defined as:

$$\text{date of visit (or event)} - \text{date of randomisation.}$$

There is no Study Day 0.

For the purpose of the safety data summaries, and any summaries based on the safety analysis set, Dose Day 1 is defined as the date of first dose of study treatment (referred to in the protocol as Week 1 Day 1). For visits (or events) that occur on or after first dose, dose day is defined as:

$$\text{date of visit (or event)} - \text{date of first dose of study treatment} + 1.$$

For visits (or events) that occur prior to first dose, dose day is defined as:

$$\text{date of visit (or event)} - \text{date of first dose of study treatment.}$$

There is no Dose Day 0.

For listings (such as for adverse events [AEs]) that include the derivation of "days since last dose", this is defined as:

$$\text{event date} - \text{date of last dose.}$$

Events that occur on the same day as the last dose of study treatment will therefore be described as occurring zero days from the last dose of study treatment.

3.1.2. Visit Windows

For summaries of vital signs, laboratory data, digital ECG, and HRQoL, etc., assessments will be assigned to calculated visit windows (using study day).

The time windows should be exhaustive so that data recorded at any time point has the potential to be summarised. Inclusion within the visit window should be based on the actual date and not the intended date of the visit. For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.

The window for the visits following baseline (including unscheduled visits) will be constructed in such a way that the upper limit of the interval falls half way between the two visits.

For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval). Listings should display all values contributing to a time point for a patient; they should also highlight the value for that patient that was used in the summary table, wherever feasible.

For visit-based summaries:

- If there is more than one value per patient within a visit window then the closest to the planned study day value should be summarised, or the earlier in the event the values are equidistant from the planned study day. The visit will be missing if no assessment was reported within the specified visit window around the planned study day.
- To prevent very large tables or plots being produced that contain many cells with meaningless data, summary statistics will be presented where at least 10 patients in either treatment group have data recorded at a particular visit. If there are assessments which fall in to a window for a visit that is not scheduled, summary statistics will be presented if at least 10 patients in either treatment group have data recorded for that visit.

3.1.3. Handling Missing Data

In general, other than for partial dates, missing data will not be imputed and will be treated as missing, unless otherwise specified.

3.1.3.1. Imputation of partial dates

Concomitant Medication and Adverse Events Start Dates

- If year is missing (or completely missing), do not impute.
- If (year is present and month and day are missing) or (year and day are present and month is missing), impute as January 1st.
- If year and month are present and day is missing, impute day as first day of the month.

Concomitant Medication and Adverse Events End Dates

- If year is missing (or completely missing), do not impute.
- If (year is present and month and day are missing) or (year and day are present and month is missing), impute as December 31st.
- If year and month are present and day is missing, impute day as last day of the month.

In addition for AEs, if for a partial start date the AE start date could (when also considering the AE end date) potentially be on the first study treatment date, the AE start date will be imputed with the first study treatment date to assume a "worst case" scenario; e.g. AE from UNK-Feb-2014 to 23-Mar-2014 with first study treatment date 21-Feb-2014, then the AE start date will be imputed to 21-Feb-2014.

Death Dates

- If a patient is known to have died where only a partial death date is available then the date of death will be imputed as the latest of the last date known to be alive +1 from the database and the death date using the available information provided:
    a. For Missing day only – using the 1st of the month

    b. For Missing day and Month – using the 1st of January

If there is evidence of death but the date is entirely missing, it will be treated as missing, i.e. censored at the last known alive date.

Diagnostic Dates

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

3.1.4. Imputation Rules for Lab Values Outside of Quantification Range

Lab values below the lower limit of quantification (LLoQ) that are reported as "<LLoQ" or "≤LLoQ" in the database will be imputed by LLoQ × 0.99 for analysis purposes. The original value will be listed.

Lab values above the upper level of quantification (ULoQ) that are reported as ">ULoQ" or "≥ULoQ" in the database will be imputed by ULoQ × 1.01 for analysis purposes. The original value will be listed.

3.1.5. Imputation rules for SF-36v2

Imputation rules will be implemented per user's manual for SF-36V2 Health Survey, third edition.

3.2. Efficacy Assessments

3.2.1. Disease Free Survival

DFS is defined as the time (in days) from the date of randomisation until the date of disease recurrence or death (by any cause in the absence of recurrence) regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to recurrence (i.e. date of DFS event or censoring – date of randomisation + 1).

Patients who are disease-free and alive at the time of analysis will be censored at the date of their last assessment for disease recurrence. However, if the patient recurs or dies immediately after two or more consecutive missed visits, the patient will be censored at the time of the latest evaluable assessment for disease recurrence prior to the two missed visits. Given the scheduled visit assessment scheme (i.e. twelve-weekly for the first 24 weeks and then twenty-four-weekly thereafter) the definition of 2 missed visits will change. If the previous assessment is less than study day 78 (i.e. week 11) then two missing visits will equate to 25 weeks since the previous assessment, allowing for late visit (i.e. 2x12 weeks + 1 week for a late assessment =25 weeks). If the two missed visits occur over the period when the scheduled frequency of assessments changes from twelve-weekly to twenty-four-weekly this will equate to 38 weeks (i.e. take the average 12 and 24 weeks which gives 18 weeks and then 2 x 18 weeks + 1 week of an early assessment + 1 week for a late assessment = 38 weeks). The time period for the previous assessment will be from study days 78 to 161 (i.e. week 11 up to week 23). If the previous assessment is from week 23 (day 162) onwards (when the scheduling

changes to twenty-four weekly assessments), two missing visits will equate to 50 weeks (i.e. 2 x 24 weeks + 1 week for an early assessment + 1 week for a late assessment = 50 weeks).

If the patient has no evaluable visits or does not have baseline data (chest, abdomen & brain scans), they will be censored at Day 1 unless they die within two visits of baseline (24 weeks plus 1 week allowing for a late assessment within the visit window) in which case the event will be used.

The DFS time will always be derived based on scan/assessment dates and not visit dates. If assessments contributing towards a particular visit are performed on different dates, for example, a biopsy confirming disease recurrence following the scan where recurrence was suspected, the date of recurrence will be the earliest of the dates of the assessment that triggered the recurrence. When censoring a patient for DFS the patient will be censored at the latest of the dates contributing to a particular overall visit assessment.

DFS rate at 6 months (i.e. 182.625 days), 12 months (i.e. 365.25 days), 18 months (i.e. 547.875 days), 24 months (i.e. 730.5 days), 36 months (i.e. 1095.75 days), 48 months (i.e. 1461 days) and 60 months (i.e. 1826.25 days) is defined as the proportion of patients alive and disease free at the timepoint, estimated from Kaplan-Meier (KM) plots of the primary endpoint of DFS at the time of the primary analysis.

3.2.2. Evidence of Disease Recurrence

Recurrence will be categorised as local/regional or distant. When recurrence is first documented at any site, complete restaging is required to identify all sites of recurrence per CRF.

3.2.2.1. Local or Regional Recurrence

Local or regional recurrence is defined as recurrence in the area of the tumour bed, hilum or mediastinal lymph nodes.

3.2.2.2. Distant Recurrence

Distant recurrence is defined as spread of disease beyond the area of the tumour bed, hilum or mediastinal lymph nodes.

3.2.3. Overall Survival

OS is defined as the time from the date of randomisation until death due to any cause regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy (i.e. date of death or censoring – date of randomisation + 1). The SURVIVE and DEATH modules of the eCRF will be used. Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive (recorded within the SURVIVE module of the eCRF). The date of contact will not be used. Other eCRF pages will not be considered.

For the futility analysis, survival calls will be made strictly after the date of the Data Cut Off (DCO) for the analysis. For the primary analysis and any subsequent analyses, survival calls will be made two weeks following the date of DCO for the analysis. If patients are confirmed to be alive or if the death date is post the DCO date these patients will be censored at the date of DCO. The status of ongoing, withdrawn (from the study) and "lost to follow-up" patients should be obtained by the site personnel by checking the patient's notes, hospital records, contacting the patient's general practitioner and checking publicly-available death registries. In the event that the patient has actively withdrawn consent to the processing of their personal data, death dates may be found by checking publicly available death registries where it is possible to do so under applicable local laws.

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

a. For Missing day only – using the 1st of the month

b. For Missing day and Month – using the 1st of January

If there is evidence of death but the date is entirely missing, it will be treated as missing, i.e. censored at the last known alive date.

OS rate at two years (i.e. 730.5 days), three years (i.e. 1095.75 days) and five years (i.e. 1826.25 days) is defined as the proportion of patients alive at the timepoint, estimated from a KM plot of OS at the time of the primary analysis.

3.2.4. Progression-Free Survival (exploratory)

PFS is defined as the time from the date of randomisation to the date of disease progression (following disease recurrence) or death (from any cause in the absence of progression) (i.e. date of PFS event or censoring – date of randomisation + 1). The date of disease progression will be recorded by the investigator and defined according to local standard clinical practice including: objective radiological, symptomatic or other progression. Patients alive and for whom a disease progression has not been observed will be censored at the last time known to be alive and without a disease progression; i.e. censored at the latest progression assessment date or disease recurrence assessment date if the patient has not had a recurrence, progression or death. However, if the patient experiences a progression or dies immediately after two or more consecutive missed visits, the patient will be censored at the time of the last progression assessment prior to the two missed visits.

3.2.5. Time to First Subsequent Therapy or Death (exploratory)

Time to first subsequent therapy or death (TFST) is defined as the time from the date of randomisation to the earlier of the date of first subsequent anti-cancer therapy or procedure start date following study treatment discontinuation, or death (i.e. date of first subsequent anti-cancer therapy/procedure/death or censoring – date of randomisation + 1). Any patient not known to have had a first subsequent therapy/procedure or not known to have died at the time

of the analysis will be censored at the last known time to have not received first subsequent therapy; i.e. the last follow-up visit where this was confirmed. If a patient terminated the study for a reason other than death before first subsequent therapy, such patients will be censored at the earliest of their last known to be alive and termination dates. Patients not receiving randomised treatment would have TFST calculated as time from date of randomisation to the initial therapy or death.

3.2.6. Time to Second Subsequent Therapy or Death (exploratory)

Time to second subsequent therapy or death (TSST) is defined as the time from the date of randomisation to the earlier of the date of second subsequent anti-cancer therapy or procedure start date following study treatment discontinuation, or death (i.e. date of second subsequent anti-cancer therapy/procedure/death or censoring – date of randomisation + 1). Any patient not known to have died at the time of the analysis and not known to have had a second subsequent therapy/procedure will be censored at the last known time to have not received second subsequent therapy; i.e. the last follow-up visit where this was confirmed. If a patient terminated the study for a reason other than death before second subsequent therapy, such patients will be censored at the earliest of their last known to be alive and termination dates. Patients not receiving randomised treatment would have TSST calculated in the same way, i.e. time from date of randomisation to the subsequent therapy or death.

3.2.7. Brain Metastases

ADAURA population consists of patients with EGFR Mutation Positive Stage IB-IIIA NSCLC and brain/CNS metastasis present at baseline is an exclusion criterion. On study, the appearance of brain cancer metastases will be those recorded on the CRF (e.g. DISREC) as Brain/CNS at disease recurrence (not including those recorded as Other CNS).

3.2.8. Health-Related Quality of Life

Patient-reported HRQoL will be assessed using the SF-36 questionnaire.

3.2.8.1. SF-36 v2 (standard)

The SF-36 includes eight domains: Physical Functioning (PF); Role Limitations-Physical (RP), Vitality (VT), General Health Perceptions (GH), Bodily Pain (BP), Social Function (SF), Role Limitations-Emotional (RE), and Mental Health (MH). These can be summarised into two summary scores: The Physical Component Summary (PCS) and Mental Component Summary (MCS). Final scores for each scale range from 0-100 with higher scores indicating better health. The general population has a mean score of 50 with a standard deviation (StdDev) of 10.

The absolute values and change from baseline will be calculated for each domain and summary scale at each scheduled post-baseline assessment. The visit response to the SF-36 at each assessment will also be categorised as improved, worsened and stable based on the

changes from baseline using the criteria for a minimum clinically important difference as shown in Table 3.

Table 3 SF-36 Visit Response Categories

	Visit response		
Score	**Improved**	**Worsened**	**Stable**
PCS	≥ + 3.1	≤ - 3.1	Otherwise
MCS	≥ + 3.8	≤ - 3.8	Otherwise
PF	≥ + 3.5	≤ - 3.5	Otherwise
RP	≥ + 3.2	≤ - 3.2	Otherwise
BP	≥ + 4.5	≤ - 4.5	Otherwise
GH	≥ + 5.7	≤ - 5.7	Otherwise
VT	≥ + 5.5	≤ - 5.5	Otherwise
SF	≥ + 5.0	≤ - 5.0	Otherwise
RE	≥ + 3.8	≤ - 3.8	Otherwise
MH	≥ + 5.5	≤ - 5.5	Otherwise

Time to deterioration of HRQoL is defined as time from date of randomisation to the date of first clinically important worsening confirmed at the subsequent assessment, or death (by any cause) in the absence of a clinically important worsening, provided death occurs within two assessment visits of the last assessment where HRQoL could be evaluated and regardless of whether the patients withdraws from randomised therapy or receives another anticancer therapy prior to symptom deterioration.

3.2.8.2. Compliance

Summary measures of overall compliance and compliance over time will be derived for patient-reported HRQoL. This will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least one individual item completed.
- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time e.g. a questionnaire from a patient who has not died, is not lost to follow-up or withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.

- Evaluable questionnaire = a questionnaire with a completion date and at least one subscale that is non-missing.
- Overall PRO compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.
- Overall patient compliance rate is defined for each randomised treatment group as: Total number of patients with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above), divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

Compliance over time will be calculated separately for each visit, including baseline as the number of patients with an evaluable questionnaire at the time point (as defined above), divided by number of patients still expected to complete questionnaires. Similarly, the evaluable rate over time will be calculated separately for each visit, including baseline, as the number of evaluable questionnaires (per definition above), divided by the number of received questionnaires.

3.2.9. Health Care Resource Use (exploratory)

The Healthcare Resource Use Module will be completed by the investigational site for any healthcare resource use between visits. The site will ask patients for any health resource use between visits (i.e., excluding routine follow-up clinic visits associated with the clinical trial but including both planned and unplanned admissions) during treatment period.

To investigate the impact of treatment and disease on health care resource, the following variables will be captured:

- Planned and unplanned hospital attendances beyond trial protocol mandated visits (including physician visits, emergency room visits, day cases and admissions)
- Primary sign or symptom the patient presents with
- Length of hospital stay
- Length of any time spent in an intensive care unit (ICU)

Where admitted overnight, the length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation (length of hospital stay = end date of hospitalisation – start date of hospitalisation + 1). Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation. The length of ICU stay will be calculated using the same method.

3.3. Safety and Tolerability

Safety and tolerability will be assessed in terms of AEs (including SAEs), exposure, deaths, laboratory data, vital signs (pulse and blood pressure [BPr]), ECGs, physical examination, WHO performance status and ophthalmologic assessments. These will be collected as detailed in the CSP.

3.3.1. Adverse Events

AEs and SAEs will be collected throughout the study, from date of main informed consent until 28 days after the last dose of study treatment. The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03).

3.3.2. Serious Adverse Events

An SAE is an AE occurring during any study phase (i.e. run-in, treatment, washout, follow-up), that fulfils one or more of the following criteria:

- Results in death.
- Is immediately life-threatening.
- Requires in-patient hospitalisation or prolongation of existing hospitalisation.
- Results in persistent or significant disability/incapacity or substantial disruption of the ability to conduct normal life functions.
- Is a congenital abnormality or birth defect.
- Is an important medical event that may jeopardise the patient or may require medical intervention to prevent one of the outcomes listed above.

3.3.3. Other Significant Adverse Events (OAEs)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs and AEs leading to discontinuation. Based on the expert's judgment, significant AEs of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered OAEs and reported as such in the CSR. A similar review of laboratory/vital signs (pulse and BPr)/ECG data will be performed for identification of OAEs.

Examples of these could be marked haematological and other laboratory abnormalities, and certain events that lead to intervention (other than those already classified as serious), dose reduction or significant additional treatment.

3.3.4. Adverse Events of Special Interest (AESIs)

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

represent pre-specified risks that are considered to be of importance to a clinical development program.

These AESIs have been identified as a list of categories provided by the patient safety team. Other categories may be added as necessary or existing terms may be merged. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which higher-level terms and which preferred terms contribute to each AESI. Further reviews may take place prior to database lock (DBL) to ensure any further terms not already included are captured within the categories.

Preferred terms used to identify adverse events of special interest will be listed before database lock and documented in the Study Master File.

3.3.5. Exposure and Dose Interruptions

Total exposure time (months) will be calculated from the first dose date to the last dose date as

$$\text{Total exposure} = \frac{(\text{last dose date where dose} > 0\text{mg} - \text{first dose date}) + 1.}{30.4375}$$

Actual exposure time (months) will be calculated from the first dose date to the last dose date, taking account of dose interruptions as

$$\text{Actual exposure} = \frac{((\text{last dose date where dose} > 0\text{mg} - \text{first dose date}) + 1) - \text{total duration of dose interruption (i. e. number of days with dose} = 0\text{mg})}{30.4375}$$

The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

Missed or forgotten doses
Missed and forgotten doses should be recorded on the DOSE module as a dose interruption with the reason recorded as “Subject forgot to take dose”. These missed or forgotten doses will not be included as dose interruptions in the summary tables but the information will appear in the listing for dosing. However, these missed and forgotten doses will be considered in the derivation of actual exposure.

Patients who permanently discontinue during a dose interruption
If a patient permanently discontinues study treatment during a dose interruption, then the date of last administration of study medication recorded on DOSDISC will be used in the programming.

Example of the dosing pattern for a patient who permanently discontinued during a dose interruption:

Dose	Start date	Stop date	Reason for change
300mg	23/01/11	14/02/11	NA
0mg	15/02/11	28/02/11	AE (dose interrupted)
200mg	01/03/11	13/03/11	AE dose restarted and reduced)
0mg	14/03/11	15/03/11	AE (dose interrupted)
0mg	16/03/11		disease recurrence (permanent discontinued)

The data will be recorded as above on the DOSE module. The date of last dose for this patient used in the programming and recorded on the DOSDISC module is 13/03/11 and the reason for permanent discontinuation recorded on DOSDISC should be AE. This patient will have one dose interruption according to the summary tables as the second dose interruption will not be included as an interruption.

Safety Follow-up

- Total Safety Follow-up = min((last dose date + 28 days), date of withdrawal of consent, date of death, date of DCO) – first dose date +1

3.3.6. Laboratory Safety Variables

Table 4 below shows the laboratory safety variables that will be summarised.

Table 4 Laboratory safety variables

Clinical chemistry (Serum [S] / Plasma [P])	Hematology (Blood [B])	Urinalysis (Urine [U])
S/P-Albumin	B-Haemoglobin	U-Glucose
S/P-ALT	B-Leukocyte	U-Protein
S/P-AST	B-Haematocrit	U-Blood
S/P-Alkaline phosphatise	B-Red blood cells (RBC) count	
S/P-Bilirubin, total	B-Absolute leukocyte differential count:	
S/P-Calcium, total	Neutrophils	
S/P-Creatinine	Lymphocytes	
Creatinine Clearance†	Monocytes	
S/P-Glucose (fasting, on PK days only)‡	Basophils	
S/P-Lactate dehydrogenase (LDH)§	Eosinophils	
S/P-Magnesium	B-Platelet count	

† Creatinine clearance will be derived using the method of Cockcroft and Gault (Cockcroft & Gault, 1976).
‡ Patients will be required to fast (water only) for at least eight hours prior to the collection of a fasting glucose sample required on PK days. Random glucose sample will be collected on non-PK days.
§ LDH is an additional variable collected at Screening only.

Clinical chemistry (Serum [S] / Plasma [P])	Hematology (Blood [B])	Urinalysis (Urine [U])
S/P-Potassium	B-Reticulocytes	
S/P-Sodium	Haemoglobin A1C (HbA1C)	
S/P-Urea nitrogen/Blood urea nitrogen (BUN)		

Corrected calcium will be derived during creation of the reporting database using the following formula:

$$\text{Corrected calcium (mmol/L)} = \text{Total calcium (mmol/L)} + ([40 - \text{albumin (g/L)}] * 0.02)$$

Creatinine clearance will be derived during creation of the reporting database according to the Cockroft-Gault formula:

$$\text{Creatinine clearance (mL/min)} = ([140 - \text{age at randomization}] * \text{weight (kg)} [* 0.85 \text{ if subject is female}]) / (72 * \text{serum creatinine (mg/dL)})$$

3.3.7. ECG

If there are replicate ECG measurements for a patient on a particular date for the same visit then the mean of the replicate measurements will be taken. For baseline and for the patient's first year on study then only digital ECG measurements will be used in the analysis. From the second year onwards, standard ECG and digital ECG measurements will be used.

Fridericia QTc correction (QTcF) will be calculated programmatically using the reported ECG values (RR and QT) as:

$$\text{QTcF} = \frac{\text{QT}}{\sqrt[3]{}}$$

where RR is the time interval in seconds between the QRS complexes.

3.4. Pharmacokinetic Variables

Plasma concentrations of AZD9291 and its metabolites A7550 and AZ5104 will be determined using validated bioanalytical method by Covance on behalf of AstraZeneca. The ratio of metabolite to AZD9291 will be calculated for each PK sample collected.

The plasma concentration data for AZD9291 and metabolites will also be analysed using a population PK approach, which may include exploring the influence of covariates on PK, if the data allows. The data collected in this study may also be combined with similar data from other studies and explored using population PK and/or PK-pharmacodynamic methods. A

separate analysis plan will be written to describe these analyses. The results of any such analyses will be reported separately from the CSR.

4. ANALYSIS METHODS

4.1. Planned Analyses

The following table (Table 5) summarises the planned summaries.

Table 5 Planned Summaries

Demographics and patient characteristics
Demography and other baseline characteristics
Medical history
Prior and concomitant medications
Efficacy, Health-related Quality of Life (HRQoL) and PK
DFS
DFS rate, OS, OS rate, PFS
SF-36
PK (plasma concentrations and ratio of metabolite to AZD9291)
Safety
Exposure
Adverse Events
Serious adverse events
Adverse events of special interest
Deaths
Laboratory evaluations
Vital signs
Physical examination
ECG
LVEF
WHO performance status
Ophthalmologic assessments

4.2. General Principles

Continuous data will be summarised using descriptive statistics (number of observations, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum). Frequencies and percentages will be used for summarising categorical (discrete) data.

Confidence intervals (CIs), when presented, will generally be constructed at the 95% confidence level. For binomial variables, the normal approximation methods will be employed unless otherwise specified.

A month is defined to be 30.4375 days.

Data will be presented in data listings by patient identifier and treatment arm.

All summaries will be presented by treatment group unless otherwise specified.

4.2.1. Baseline Measurements and Change from Baseline Variables

In general, for summaries using the full analysis set, the last observed measurement prior to randomisation will be considered the baseline measurement. However, if an evaluable assessment is only available on the day of randomisation then this assessment will be used as baseline and assumed to have occurred prior to first dose. For summaries using the safety analysis set the last observation prior to the first dose of study treatment will be considered the baseline measurement unless otherwise specified.

For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose.

Assessments on the day of first dose where neither time nor a nominal pre-dose indicator are captured will be considered prior to first dose if such procedures are required by the protocol to be conducted before first dose.

In all summaries, change from baseline variables will be calculated as:

$$\text{post-treatment value} - \text{baseline value}\,.$$

The percentage change from baseline will be calculated as:

$$\frac{(\text{post-baseline value} - \text{baseline value})}{\text{baseline value}} \times 100.$$

4.2.2. Multiple Testing Strategy

A single, primary analysis is planned for this study, with a further analysis performed only if there are significantly less than 70 DFS events (i.e. 63 DFS events or less) in the IB population at the time of the primary analysis. The primary endpoint of DFS and secondary endpoint of OS will be tested in the subset of patients with stage IIA-IIIA cancer as well as in the overall population. In order to describe the nature of the benefits of AZD9291 treatment, primary and secondary endpoints will be tested at a two-sided significance level of 5%. However, in order to strongly control the type I error at 5% two-sided level, a hierarchical testing procedure will also be employed across the primary endpoint and secondary endpoint of OS, intended for key label claims.

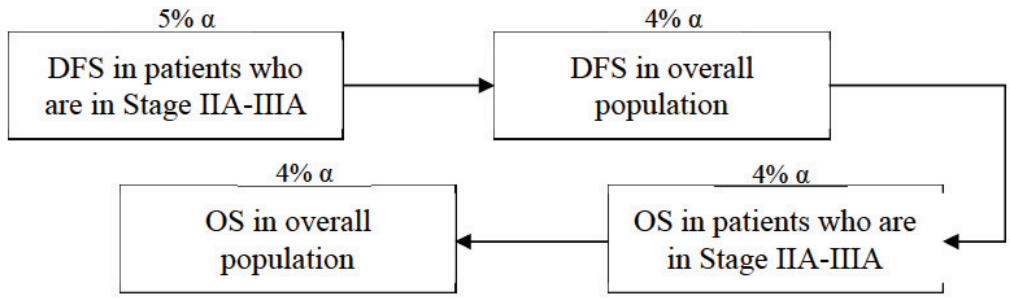
DFS in patients who are in Stage IIA-IIIA (i.e. non-IB) is tested first using the full test mass (full test mass = alpha).

DFS in the overall population will only be tested, with the test mass split between first and second analyses, if statistical significance is shown for DFS in patients who are Stage IIA-IIIA. OS (in both populations) will only be tested, with the test mass split between first and second analyses, if statistical significance is shown for DFS in the overall population.

Both DFS in the overall population and OS in both populations will be tested at the time of the primary analyses of DFS in patients who are Stage IIA-IIIA and may be tested again when at least 70 DFS events in the IB population have been observed (if there were significantly less than 70 DFS events at the primary analysis). Only a proportion of alpha will be spent at the first analyses for both endpoints to control for multiple testing.

The hierarchical testing procedure at the time of the primary analysis is shown in Figure 2. At the time of the primary analysis, DFS in patients who are in Stage IIA-IIIA is tested first using the full test mass of 5%. If this result is statistically significant, DFS in the overall population will be tested at 4%. If this result is statistically significant, OS in patients who are in Stage IIA-IIIA will be tested at 4%. If this result is statistically significant, OS in the overall population will be tested at 4%.

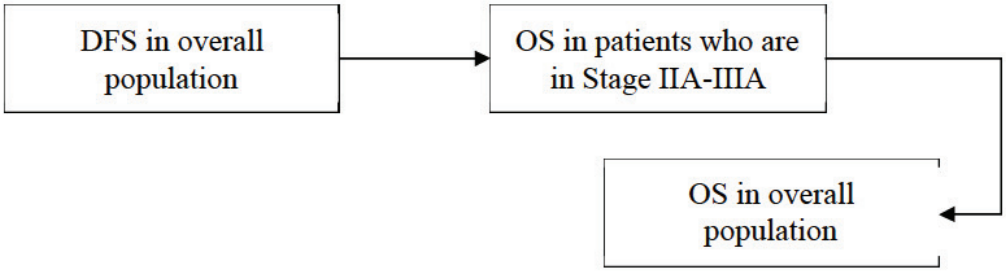
Figure 2 Hierarchical Testing Procedure at time of Primary Analysis



If there are significantly less than 70 DFS events in the IB population at the time of the primary analysis, further follow up of all patients will be performed and a second analysis of DFS in the overall population and OS in both populations will be conducted. The hierarchical testing procedure at this follow-up is shown in Figure 3. DFS in the overall population will be testing using the remaining alpha taking account of correlation (Stone, 2010). If this result is statistically significant, OS in patients who are in Stage IIA-IIIA will be tested using the remaining alpha (taking account of correlation). If this result is statistically significant, OS in the overall population will be tested using the remaining alpha (taking account of correlation). The alpha level taking account of correlation for the second analyses of DFS in the overall population and the OS analyses will be calculated from the observed number of events during the study using the software package EAST by selecting boundary family: Haybittle-Peto (p-value) and specifying the information fraction and fixed significance level used at the first analysis. The information fraction is calculated as the number of events observed at the first analysis divided by the total number of events observed at the second analysis.

For example, assuming that 287 DFS events are observed in the overall population at the first analysis and 378 events are observed at the second DFS analysis of the overall population. The information fraction would be 0.759 (287/378 events). This would result in a significance level for the second analysis of DFS in the overall population of 2.6%.

Figure 3 Hierarchical Testing Procedure at Follow-up Analysis (if performed)



4.3. Analysis Methods

4.3.1. Patient Disposition and Data Sets Analysed

Patient disposition will be listed and summarised for the FAS. Summaries will include the number and percentage of patients:

- Randomised
- Treated
- Patients ongoing study treatment at the data cut-off
- Included in each analysis set (FAS, Safety, and PK).

In addition, the number and percentage of patients who discontinued treatment and who discontinued the study, including a breakdown of the main reason for discontinuation will be presented for all patients. This will also include the number and percentage of patients who completed their three years treatment.

The total number of patients enrolled will also be summarised.

4.3.2. Protocol Deviations

All important protocol deviations will be listed and summarised for the FAS. All protocol deviations will be defined by the study team before the database lock.

4.3.3. Demographic and Other Baseline Characteristics

Demographic, patient, and disease characteristics at baseline will be listed and summarised for the FAS.

Standard descriptive statistics (n, mean, SD, median, minimum, maximum) will be presented for the continuous variables of:

- Age (years)
- Weight (kg)
- Height (cm)
- Body mass index (kg/m^2), calculated as: $\frac{\text{weight}}{\text{height}^2}$
- Nicotine consumption (number of pack years)

The total frequency counts and percentages of patients will be presented for the categorical variables of:

- Age group (years) (grouped as <50, ≥50-<65, ≥65-<75, ≥75)
- Sex
- Race
- Ethnic group
- Recruitment by country and centre
- AJCC staging at diagnosis (IB, IIA, IIB, IIIA)
- Disease characteristics at baseline (Primary tumour location and laterality, histology type, lung cancer resection type, primary tumour grade, and overall classification)
- Smoking status (former, current, never)
- WHO performance status (Normal activity, restricted activity)
- EGFR mutation type (including Ex19del, L858R)

In addition, stratification factors will be summarised as recorded in IVRS at the time of randomisation (a summary table of stratification factors as recorded in the eCRF will also be presented for information). A listing of all patients showing all the EGFR mutations identified by the cobas® central test will be presented.

4.3.4. Medical History

Disease related medical history and relevant surgical history will be coded using MedDRA. All disease related medical history will be listed and the number and percentage of patients with any disease related medical history will be summarised for the FAS by system organ class (SOC) and preferred term (PT).

All relevant surgical history will be listed and summarised similarly.

4.3.5. Concomitant and Other Treatments

Information on any treatment within the four weeks prior to initiation of study treatment and all concomitant treatments given up to 28 days after discontinuation of study treatment, or disease recurrence (whichever is later), with reasons for the treatment, will be recorded in the electronic case report form (eCRF). Thereafter, only subsequent regimens of anti-cancer therapy will be recorded in the eCRF.

Other anti-cancer therapies, investigational agents, and radiotherapy should not be given while the patient is on study treatment.

Medications received prior to, concomitantly, or post-study treatment will be coded using the WHO Drug Dictionary Anatomical Therapeutic Chemical (ATC) Classification codes. Concomitant medications will be summarised for the FAS by ATC classification codes.

For the purpose of inclusion in prior and/or concomitant medication or therapy summaries, incomplete medication or radiotherapy start and stop dates will be imputed as detailed in Section 3.1.3.1.

Prior medications, prior anti-cancer therapies, concomitant medications (disallowed and allowed), post IP medications, and post IP anti-cancer therapies are defined based on start and stop dates as follows:

- Prior medications and anti-cancer therapies are those received prior to screening with a stop date prior to the first dose of study treatment.
- Concomitant medications are those with a stop date on or after the first dose date of study treatment (and could have started prior to or during treatment).
- Post IP medications and post IP anti-cancer therapies are those with a start date after the last dose date of study treatment.

If it is not possible to strictly determine whether a medication is prior or concomitant due to partial or missing dates the medication will be assumed to be concomitant.

The following summaries will be produced for the FAS

- Summary of prior medications
- Summary of concomitant medications
- Summary of post IP medications
- Summary of prior anti-cancer therapies
- Summary of post IP anti-cancer therapies

All treatment data will be listed.

Missing coding terms should be listed and summarised as "Not coded".

In addition, the summaries of the number and percentage of patients with adjuvant disease-related treatment modalities and with the adjuvant chemotherapy by stage (IB, II, IIIA) will be presented for the FAS.

4.3.6. Brain Metastases

The proportion of patients who develop brain metastases during the trial will be summarised by treatment group for the FAS.

4.3.7. Efficacy

All efficacy analyses will be performed on the FAS. Results of all statistical analyses will be presented using a 95% CI and two-sided p-value.

4.3.7.1. Primary Outcome: Disease Free Survival (DFS)

DFS in the subset of patients with stage IIA-IIIA cancer will be analysed using a log-rank test stratified by stage (II versus IIIA), mutation type (Ex19Del versus L858R either alone or in combination with other EGFR mutations) and race (Asian versus Non-Asian) for the generation of the p-value and using the Breslow approach for handling ties. The covariates in the statistical modelling will be based on the values entered into IVRS at randomisation, even if it is subsequently discovered that these values were incorrect.

DFS in the overall population will be analysed using a log-rank test stratified by stage (IB versus II versus IIIA), mutation status (Ex19Del versus L858R either alone or in combination with other EGFR mutations as confirmed by a central test) and race (Asian versus Non-Asian) for the generation of the p-value and using the Breslow approach for handling ties. The HR and CI will be obtained directly from the U and V statistics as follows (Berry, et al., 1991), (Selke & Siegmund, 1983):

$$H = \exp\left(\frac{U}{V}\right)$$

$$95\%\ \text{CI for } H = \left(\exp\left[\frac{U}{V} - \frac{1.96}{\sqrt{V}}\right], \exp\left[\frac{U}{V} + \frac{1.96}{\sqrt{V}}\right]\right)$$

Where $U = \sum_k U_k = \sum_k \sum_i (d_{1ki} - e_{1ki})$ is the stratified log-rank test statistic (with d_{1ki} and e_{1ki} the observed and expected events in group 1, stratum *k*) and $\sqrt{V} = \sqrt{\sum_k v_k}$ is the standard deviation obtained from the LIFETEST procedure with a STRATA term for the stratification variables.

The assumption of proportionality will be assessed. In the event of non-proportionality, the HR will be interpreted as an average HR over the observed extent of follow-up unless there is extensive crossing of the survival curves. Proportionality will be tested firstly by examining the plots of complementary log-log (event times) versus log (time) and, if judged necessary from examination of these plots, a time dependent covariate will be fitted to assess the extent to which the lack of proportionality represents random variation. If a lack of proportionality is evident, the variation in treatment effect can be described by presenting piecewise HR calculated over distinct time-periods. If lack of proportionality is found this may be a result of a treatment-by-covariate interaction, which will be investigated.

A Kaplan-Meier (KM) plot of DFS will be presented by treatment group. The total number of events and median DFS (calculated from the KM plot, with two-sided 95% CIs and with two-sided 96% CIs) will be summarised.

The type of DFS event will be broken down at the time of data cut off, whereby an event includes both disease recurrence and death in the absence of disease recurrence. Various reasons may be provided for patient censoring, for example, lost to follow up or a disease recurrence occurring after the last evaluable assessment. Additionally, those censored patients will be summarised according to whether more than one scheduled assessment interval (+ 2 Weeks) has elapsed prior to the data cut-off, i.e. prematurely censored.

Further tables will summarise the treatment status of patients at the time of disease recurrence or censoring in terms of having no randomised treatment, completed or discontinued treatment prior to disease recurrence or censoring, or ongoing randomised treatment at the time of disease recurrence or censoring.

DFS Sensitivity Analyses

a) Quantitative Interactions

The presence of quantitative interactions will be assessed by means of an overall global interaction test. This will be performed by comparing the fit of a Cox proportional-hazards (PH) model including treatment, covariates for race, stage, and mutation status, and all covariate-by-treatment interaction terms, with one that excludes the interaction terms and will be assessed at the two-sided 10% significance level. If the fit of the model is not significantly improved, then it will be concluded that overall the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant, and all excluded interactions are non-significant. Throughout this process, all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

Any quantitative interactions identified using this procedure will then be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail & Simon, 1985).

b) Evaluation-time bias

In order to assess possible evaluation-time bias that could occur if scans are not performed at the protocol-scheduled time points, the midpoint between the time of recurrence and the previous evaluable assessment will be analysed using a log rank test stratified by stage, mutation status and race. Note that midpoint values resulting in non-integer values should be rounded down.

For patients who die in the absence of recurrence, the date of death will be used to derive the DFS time used in the analysis. Patients with no events (i.e. patients censored for DFS) will be censored at day 1 for this analysis.

c) Attrition bias

Possible attrition bias will be assessed by repeating the primary DFS analysis, except that the actual DFS times rather than the censored times of patients who recurred or died in the absence of recurrence immediately following two or more non-evaluable assessments, will be included. In addition, and within the same sensitivity analysis, patients who take subsequent (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy) therapy prior to their last evaluable assessment or recurrence or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. A KM plot of the time to censoring, where the censoring indicator of the primary DFS analysis is reversed, will be presented to assess the number of patients being followed over time.

Subgroup Analyses

Subgroup analyses will be conducted comparing DFS between the treatments in the following subgroups of the FAS:

- Stage (IB / II / IIIA)
- EGFR mutation type (Ex19Del / L858R either alone or in combination with other EGFR mutations)
- EGFR mutation status detectable in plasma-derived ctDNA (Ex19Del / L858R / unknown)
- Pre-treatment T790M mutation status (Positive / Negative / unknown)
- Race (Asian / Non-Asian)
- Adjuvant chemotherapy (Yes / No)
- Gender (Male / Female)
- Age at screening (<65 / ≥65)
- Smoking history (Yes/No)

The subgroup analyses for the stratification factors will be based on the values entered into the IVRS, all other factors will be based on values recorded on the eCRF.

Other baseline variables may also be assessed if there is clinical or biological justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic factors.

No adjustment to the significance level for testing will be made since the subgroup analysis may only be supportive of the primary analysis of DFS. For each subgroup level, the HR and 95% CI will be calculated from a single Cox PH model that contains a term for treatment, the subgroup covariate of interest, and the treatment by subgroup interaction term. The HR will be

obtained for each level of the subgroup from this model. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties.

These HRs and the associated 95% CIs will be summarised and presented on a forest plot, along with the results of the overall population and non-IBs.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events per level in a subgroup), the relationship between that subgroup and DFS will not be formally analysed. In this case, only descriptive summaries will be provided.

4.3.7.2. Secondary Outcomes

Analysis of DFS Rate

DFS rate data will be analysed using the same model as for the primary analysis of DFS.

KM plots will be produced to assess if recurrence rates have reached an apparent plateau. Estimates of DFS rate at 6 months, 12 months, 18 months, 24 months, 36 months, 48 months and 60 months obtained from the plot will be presented for each arm. If the shape of the DFS curve suggests a plateau, a cure rate model will also be fitted.

Analysis of Overall Survival

Survival Status will be summarised in terms of those who have died or censored whereby patients may be alive and ongoing in the study or discontinued the study prior to death. OS data will be analysed using the same methodology and model as for the analysis of DFS (with the exception that the sensitivity, subgroup and exploratory analyses will not be performed), provided there are sufficient events available for a meaningful analysis (i.e. >20 deaths [if not, descriptive summaries will be provided]).

OS rate will be estimated at 2, 3 and 5 years will be estimated for each arm from a KM plot of OS at the time of the primary analysis. This will be estimated for non-IBs and overall.

If more mature data on OS is required, then it may be analysed at a later timepoint.

Additional exploratory analysis of OS adjusting for the impact of treatment switching may be performed[**] if patients randomised to placebo subsequently receive a T790M directed EGFR-TKI. Methods such as Rank Preserving Structural Failure Time (RPSFT) (Robins & Tsiatis, 1991), Inverse Probability of Censoring Weighting (IPCW) (Robins & Finkelstein, 2000) and other methods in development may be explored. The decision to adjust and final choice of methods will depend on the observed data and the plausibility of the underlying assumptions and the results from these analyses will be reported separately from the CSR.

[**] internally by AstraZeneca.

Analysis of Health-Related Quality of Life

The compliance with SF-36 will be summarised by visit.

The scores for each of the eight health domains scores and for each of the PCS and MCS measures from the SF-36 will be summarised in terms of mean score and change from baseline values at each post-baseline assessment. The responses to each of the health domain scores and summary measures will also be categorised in terms of improved, worsened, and stable at each post-baseline assessment.

The primary HRQoL outcome measures of interest are time to deterioration of the two aggregated summary scores (MCS and PCS). The probability of making a type I error (5% two-sided) will be split equally between these two analyses. Time to deterioration in the subset of patients with stage IIA-IIIA cancer will be analysed using a log-rank test stratified by stage (II, IIIA), mutation type (Ex19Del, L858R either alone or in combination with other EGFR mutations) and race (Asian, Non-Asian) using the Breslow approach for handling ties.

For two missed visits rule please refer to user's manual for SF-36V2 Health Survey, third edition.

4.3.7.3. Exploratory Outcomes

Healthcare Care Resource Use

Health Care Resource Use data (i.e. the hospital admission details including type of attendance and primary symptom for admission) will be listed and presented using frequencies and percentages. The duration of admission/attendance and duration of intensive care unit (ICU)/high dependency unit (HDU) stay will be summarised using descriptive statistics, based on the FAS. It will summarised from randomisation and up to and including treatment discontinuation.

Analysis of Progression-Free Survival

Progression status at the time of analysis by treatment arm will be categorised into radiological, symptomatic, other progression or death, with reasons for censoring also summarised accordingly. PFS data will be analysed for the overall population using the same methodology and model as for the analysis of DFS with the exception that the sensitivity, subgroup and exploratory analyses will not be performed. The HR for the treatment effect together with its 95% CI will be presented. In addition, medians and Kaplan-Meier plot will be presented by treatment arm to support the analysis. It will be summarised per treatment arm, but no formal comparisons will be made. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint. Time from randomisation to progression will be summarised by treatment arm.

Post-Relapse Outcomes

TFST and TSST will be summarised for the overall population using the same methodology and model as for the analysis of DFS with the exception that the sensitivity, subgroup and

exploratory analyses will not be performed. The HR for the treatment effect together with its 95% CI will be presented. In addition, medians and a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment arm and the time between recurrence and starting subsequent therapy will be assessed. This will be summarised per treatment arm, but no formal comparisons will be made. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

In patients who received a subsequent anti-cancer therapy, a summary table of first (and second) subsequent anti-cancer therapies by treatment arm will be provided, as well as response to first subsequent anti-cancer therapy by treatment arm (if available).

A summary of the number of patients prematurely censored will also be produced.

Analysis of EGFR Mutation Status Between Tumour DNA And Plasma-Derived CtDNA

Cross-tabulation summaries of the mutation status from EGFR mutation status from tumour DNA (Positive, Negative, Total) versus the mutation status from plasma-derived ctDNA (Positive, Negative, Unknown, Total) will be produced for:

- all screened patients who are evaluable (Positive, Negative) for baseline tumour mutation status
- all screened patients who are evaluable for both the baseline tumour mutation status and the recurrence mutation status

The concordance (i.e. both positive or both negative), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) rates for the EGFR mutation status will be produced for subgroups of patients as mentioned above. The 95% CIs using Wilson score intervals will be presented for each of these, however if the number of patients in the subgroup is <30, then the Clopper-Pearson method (Clopper & Pearson, 1934) will be used to derive the 95% CI.

Sensitivity is defined as:

$$100 \times \frac{\text{number of patients who are EGFR postive from both tumour sample and plasma-derived}}{\text{number of patients who are EGFR postive from tumour sample}}$$

Specificity is defined as:

$$100 \times \frac{\text{number of patients who are EGFR negative from both tumour sample and plasma-derived}}{\text{number of patients who are EGFR negative from tumour sample}}$$

PPV is defined as:

$$100 \times \frac{\text{number of patients who are EGFR positive from both tumour sample and plasma-derived}}{\text{number of patients who are EGFR positive from plasma-derived}}$$

NPV is defined as:

$$100 \times \frac{\text{number of patients who are EGFR negative from both tumour sample and plasma-derived}}{\text{number of patients who are EGFR negative from plasma-derived}}$$

Analysis of Plasma-Derived ctDNA EGFR Mutation Status at Baseline Compared to Status at Disease Recurrence

Cross-tabulation summaries of the mutation status from the first plasma sample at baseline (Positive, Negative, Unknown, Total) versus the mutation status from the plasma sample taken at disease recurrence (Positive, Negative, Unknown, Total) will be presented:

- in patients from the FAS population who are evaluable for both the first plasma sample and the recurrence plasma sample

The concordance rate and its 95% CI for the EGFR mutation status of the first plasma sample versus the progression plasma sample will also be presented for groups described above.

Analysis of Time to New Brain Lesion or Death

Time to new brain lesion or death will be summarised for the stage II/IIIA and overall population using the same methodology and model as for the analysis of DFS with the exception that the sensitivity, subgroup and exploratory analyses will not be performed. An event is a recurrence with evidence of a CNS lesion or death by any cause. Patients without an event will be censored at the last evaluable disease assessment. The HR for the treatment effect together with its 95% CI will be presented. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

4.3.8. Safety and Tolerability

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

4.3.8.1. Adverse Events and Serious Adverse Events

All AEs, both in terms of MedDRA PT and CTCAE grade, will be summarised descriptively by count (n) and percentage (%) for each treatment arm. The latest MedDRA dictionary version will be used for coding. Missing coding terms should be listed and summarised as "Not coded".

The summary tables will include all AEs that occurred after the start of treatment up until the end of the 28-day follow-up period. For the IDMCs, the 28-day follow-up period will be defined as 28 days following discontinuation of study treatment. For the primary analysis and any subsequent analysis, the 28-day follow-up period will be defined as 28 days following discontinuation of study treatment and before starting subsequent cancer therapy. Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of study treatment) will be flagged in the data listings for all analyses.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of CTCAE grade and relationship to study treatment. Frequencies and percentages of patients reporting each PT will be presented (i.e. multiple events per patient will not be accounted for apart from on the episode level summaries).

Summary information (the number and percent of patients by treatment) by SOC and PT will be tabulated for:

- All AEs
- All AEs causally related to study treatment
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to treatment
- AEs with outcome of death
- AEs with outcome of death causally related to treatment
- AEs leading to dose modification (interruptions and/or reductions)
- All SAEs
- All SAEs causally related to study treatment
- AEs leading to discontinuation of treatment
- AEs leading to discontinuation of treatment, causally related to treatment
- OAEs
- OAEs causally related to treatment

In addition, a truncated AE table of most common AEs, showing all events that occur in at least 5% of patients overall will be summarised by PT, by decreasing frequency. This cut-off may be modified after review of the data.

AEs will be assigned CTCAE grades and summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, SOC, PT and actual treatment group. Fluctuations observed in CTCAE grades during the study will be listed (where collected).

AEs with outcome of death, SAEs, AEs leading to discontinuation of treatment, AEs causally related to treatment and OAEs will be listed.

Any AE occurring before study treatment will be included in the data listings but will not be included in the summary tables of AEs.

A separate summary of AEs occurring more than 28 days after discontinuation of AZD9291 (where reported) as well as those occurring prior to treatment will be produced. These events will not be included in other AE summaries.

4.3.8.2. Adverse Events of Special Interest

Summary tables of AEs of special interest will be produced. The number (%) of patients experiencing any of the specified terms will be presented overall and by maximum CTCAE grade. Additional summaries of time to onset of first AE for each grouped term and each PT within it; time to onset of first CTCAE grade three or higher and duration of AEs of special interest will be produced. In addition, the summary tables in the AEs section listed above will be repeated for grouped AEs of special interest.

Life tables and prevalence plots for rash, diarrhoea, fatigue and stomatitis will also be produced.

The management of AEs of diarrhoea and rash will be summarised in terms of medication taken, dose modification and proportion of time on treatment.

Data collected on the specific eCRF for skin reactions will also be listed and summarised to further assess tolerability of specific events. Summaries will be by PT group for patient level and event level, overall, by effect and by CTCAE grade. The number of skin reactions per patient will also be summarised.

4.3.8.3. Deaths

A summary of deaths for the FAS will be provided with number and percentage of patients, categorised as:

- Related to disease under investigation,
- AE outcome=death,
- Both related to disease under investigation and with AE outcome=death,
- AE with outcome=death >28days after last treatment dose,
- Patients with unknown reason for death.

A corresponding listing will also be produced.

4.3.8.4. Exposure

Exposure will be summarised for safety analysis set. The following summaries will be produced:

- Summary of duration of exposure of study treatment including total exposure, actual exposure, and cumulative exposure over time (number and percentage for: ≥1 day, ≥6 months, ≥12 months, ≥18 months, ≥24 months, ≥30 months, ≥36 months)
- Summary of interruptions and reductions of study treatment

4.3.8.5. Laboratory Evaluations

All laboratory data recorded in the eCRF will be listed. If any additional analytes to those in Table 4 are also recorded, then these will be listed only.

All values will be classified as low (below range), normal (within range) and high (above range) based on project-specific reference ranges. As applicable, values will be converted to standard units and will be graded using CTCAE v4.0.

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised.

For clinical chemistry and haematology, shift tables will present movements in and out of reference range or, if applicable, CTCAE grade changes from baseline to the maximum/minimum value on treatment will be provided. Corresponding shift tables ("Negative", "Trace", "Positive", "0", "+", "++", "+++") will be produced for urinalysis.

Plots of both the maximum post-baseline alanine aminotransferase (ALT) and aspartate aminotransferase (AST) versus the maximum post-baseline total bilirubin, expressed as multiples of their upper limit of reference range will be produced. Box plots of absolute values and change from baseline for all haematology and clinical chemistry parameters will also be presented.

A pregnancy test will also be performed at screening only for women of child-bearing potential.

4.3.8.6. Vital Signs (Pulse and BPr) and Weight

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised.

Percentage change from baseline for pulse, BPr and weight will be summarised by treatment group and visit.

4.3.8.7. Physical Examination

Abnormalities identified from physical examination will be listed.

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised. For each physical examination body system, the number and percentage of patients with abnormalities at baseline and post-baseline will be summarised.

4.3.8.8. ECG

All ECG data received will be presented in data listings.

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised.

QTc summaries will be presented for patients in the full analysis set.

The following ECG parameters will be summarised (absolute values and change from baseline) by visit: QTcF, RR, PR, QRS and QT.

Plots of mean observed QTc parameters and change from baseline in QTc parameters versus time will be presented. Shift plots of the maximum change from baseline versus the baseline value for QTcF, with reference lines for 450 ms, ±30 ms and ±60 ms change will be presented.

QTc outliers are defined as QTcF values following dosing that are greater than 450 ms or are increases from baseline greater than 30 ms. QTcF outliers will be highlighted in the data listings and summarised using the following categories:

- Values >450 ms, >480 ms., >500 ms.,
- Increase from baseline of >30 ms., Increase from baseline of >60 ms., Increase from baseline of >90 ms.,
- Values >450 ms and increases of >30 ms. Values >500 ms and increases of >60 ms.

The number and percentage of patients who meet the ECG outlier criteria at any assessment post-date of first dose will be summarised.

4.3.8.9. Left Ventricular Ejection Fraction

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised.

LVEF outliers are defined as LVEF values following dosing that are:

- ≥10 percentage points decrease from baseline and <50%, or
- ≥15 percentage points decrease from baseline and ≥50%.
- LVEF outliers and abnormal values will be highlighted in the data listings and summarised.

Percentages will be calculated for patients with a baseline and at least one post baseline measurement.

Patients with outliers that are associated with an AE (related or unrelated to study treatment) will be listed and summarised as appropriate, along with the duration of the abnormality.

4.3.8.10. WHO Performance Status

For the primary analysis and any subsequent analysis only data collected before starting further cancer therapy will be summarised.

WHO performance status will be listed and summarised as the frequency and percentage of patients by treatment and visit for the safety analysis set.

4.3.8.11. Ophthalmologic Assessment

Ophthalmologic assessment will be listed by treatment and visit for the safety analysis set.

4.3.9. PK Concentration Data

All plasma concentrations of AZD9291 and metabolites (AZ5104 and AZ7550) will be listed for the PK analysis set.

The following summary statistics will be presented for AZD9291, AZ5104 and AZ7550 plasma concentrations and metabolite(s) to AZD9291 ratio by treatment day and nominal sample time window for the PK analysis set:

- Geometric mean (gmean) calculated as e^{μ}, where μ is the mean of the data on a logarithmic scale. The corresponding error bars will be calculated as $\exp(\mu \pm s)$ and referred to as Gmean ± StdDev in the table
- Geometric standard deviation (GeoSD) calculated as e^{s}, where s is the standard deviation of the log transformed data
- Coefficient of variation (CV) calculated as $100 \times \sqrt{(\exp(s^2) - 1)}$
- Arithmetic mean calculated using untransformed data
- Standard deviation (StdDev) calculated using untransformed data
- Minimum
- Median
- Maximum
- Number of observations (n)
- n≤ LLOQ

Non-quantifiable (NQ) values of plasma concentrations will be handled as follows:

- If, at a given time point, 50% or less of the plasma concentrations are NQ, the gmean, CV, gmean ± StdDev, arithmetic mean and StdDev will be calculated by substituting the limit of quantification (LOQ) for values which are NQ.
- If more than 50%, but not all, of the concentrations are NQ, the gmean, CV, gmean ± StdDev, arithmetic mean and StdDev will be reported as not calculable (NC).
- If all the concentrations are NQ, the gmean and arithmetic mean will be reported as NQ and the CV, gmean ± StdDev and StdDev as NC.
- The number of values above LLOQ will be reported for each time point along with the total number of collected values.

If data are available for less than three patients, no summary statistics other than minimum, maximum and n will be presented.

The concentration data for Cycle 1 Day 1 and Cycle 3 Day 1 may be displayed graphically for the PK analysis set, as appropriate. Displays may include plasma concentration profiles versus time and gmean concentrations (± StdDev) both on the linear and on the log-scale versus time, as appropriate.

5. INTERIM ANALYSES

An Independent Data Monitoring Committee (IDMC) will be convened and will meet every six months for the first two years from the first patient randomised, and yearly thereafter. Further meetings for review of safety data may be convened at the discretion of IDMC. The IDMC will review safety assessments and make recommendations to continue, amend, or stop the study based on safety findings. Serious adverse events, adverse events, and other safety data will be reviewed, and individual and aggregated safety data will be evaluated by the IDMC.

The IDMC will conduct a futility analysis. Further details will be documented in the IDMC Charter prior to the first DMC safety review meeting. The IDMC will review the futility outcomes and provide a recommendation for whether the study should continue, stop or be modified in some way.

6. CHANGES OF ANALYSIS FROM PROTOCOL

The study protocol states for the subgroup analysis of DFS by age (see “Subgroup Analysis” in Section 4.3.7.1 above) that the age should be as collected at randomisation. However, age is only captured at the screening visit, as such the subgroup analysis will be performed using age at screening.

PK analysis set is defined using safety analysis set instead of FAS originally specified in the protocol to be able to interpret the PK according to the treatment the patients actually received.

In Table 4, Haemoglobin A1C (HbA1C) is listed under Haematology (column 2) instead of Clinical chemistry (column 1) as originally presented in the clinical study protocol. This is to correct an error from the Protocol.

For the primary and any subsequent analyses, all safety data will only be summarised until start of subsequent cancer therapy to aid interpretation.

If more mature data on Overall Survival is required, then it may be analysed at a later timepoint.

7. REVISION TO MULTIPLE TESTING STRATEGY

This section supersedes earlier text regarding alpha allocation.

At the time of the scheduled IDMC meeting for safety in April 2020, the IDMC requested efficacy data. Based on the review of this data, the IDMC made the recommendation to unblind the study and complete primary reporting. Given these unplanned analyses of efficacy, the alpha allocation needs to be revised to control the type I error.

No changes will be made to the order of the hypothesis being tested.

Primary Endpoint: DFS in stage II/IIIA

The primary analysis was planned to be conducted when approximately 247 DFS events were observed in the stage II/IIIA population. This represents 50% maturity based on the planned sample size of 490 subjects.

Two unplanned interim analysis of DFS in the stage II/IIIA population were conducted at the time of observing 86 DFS events and 156 DFS events respectively. The corresponding information fractions were 0.35 and 0.63 where the final number of events would have been 247. The Lan DeMets approach that approximates the O' Brien and Fleming spending function will be used to maintain an overall 2-sided 5% type I error. Using statistical software package EAST, the following stopping boundaries are obtained and shown below in Table 6.

The exact number of events, associated information fractions and 2-sided p-values will be confirmed after database lock and unblinding.

Table 6 Alpha allocation under Lan-DeMets with O'Brien-Fleming type spending function, stage II/IIIA population

Timepoint	Number of events/ information fraction/maturity	Critical value (HR)	2-sided p-value
IDMC6 (April 2019) - safety and futility review	86/0.35/18%	0.4590	0.00030
IDMC7 (April 2020) - safety review	156/0.63/33%	0.6588	0.009384
Primary planned analysis per protocol (expected Q1 2022)	247/1.0/53%	0.7763	0.04701

An exploratory DFS analysis in the stage II/IIIA population will be conducted once approximately 247 DFS events have occurred. This analysis will not be formally tested and will be considered exploratory. At this time, an exploratory analysis of DFS in the overall population will also be conducted.

DFS in the overall population

If testing of the DFS in stage II/IIIA population is statistically significant, the full 2-sided 5% alpha can be recycled forward into testing endpoints pre-specified in the hierarchal testing procedure. The next test in the hierarchal procedure is to test the DFS in the overall population.

Two unplanned interim analysis of DFS in the overall population were conducted at the time of observing 109 DFS events and 196 DFS events respectively. This equates to an information fraction of 0.34 and 0.62, where the final number of events would have been 317. The Lan DeMets approach that approximates the O'Brien and Fleming spending function will be used to maintain an overall 2-sided 5% type I error. Using statistical software package EAST, the following stopping boundaries are obtained and shown below in Table 7.

The exact number of events, associated information fractions and 2-sided p-values will be confirmed after database lock and unblinding.

Table 7 Alpha allocation under Lan-DeMets with O'Brien-Fleming type spending function, overall population

Timepoint	Number of events/ information fraction/maturity	Critical value (HR)	2-sided p-value
IDMC6 (April 2019) - safety and futility review	109/0.34/16%	0.4938	0.00025
IDMC7 (April 2020) - safety review	196/0.62/29%	0.6886	0.00885
Primary planned analysis per protocol (expected Q1 2022)	317/1.0/47%	0.8002	0.04718

An exploratory DFS analysis in the overall population will be conducted once approximately 247 DFS events have occurred in the stage II/IIIA population and approximately 70 DFS events in the stage IB, if there is less than 63 events at the time of the exploratory DFS analysis in the stage II/IIIA population. This analysis will not be formally tested and will be considered exploratory.

Overall Survival in Stage II/IIIA population

If the test of DFS in overall population is statistically significant, OS in stage II/IIIA population will be tested using the Haybittle-Peto boundary with alpha allocation of 0.0002 (two-sided) for each of the interim analysis and overall 2-sided alpha of 5%. The final analysis of OS will be conducted when approximately 94 deaths have been observed (approximately 20% maturity). The exact 2-sided alpha will be calculated based the exact information fraction at the time of the analysis. Alpha will be fully exhausted.

Depending on the length of time to reach approximately 94 OS events, an additional exploratory analysis reporting the 3, 4 and 5 year OS landmarks may be conducted after the final OS analysis. This analysis will not be formally tested.

Overall Survival in overall population

If the test of OS in the stage II/IIIA population is statistically significant, OS in the overall population will be tested using the Haybittle-Peto boundary with alpha allocation of 0.0002 (two-sided) for each of the interim analysis and overall 2-sided alpha of 5%. This final analysis of OS will be conducted at the same time as the OS analysis in the stage II/IIIA population. The exact 2-sided alpha will be calculated based the exact information fraction at the time of the analysis. Alpha will be fully exhausted.

Depending on the length of time to reach approximately 94 OS events in the stage II/IIIA population, an additional exploratory analysis reporting the 3, 4 and 5 year OS landmarks in the overall population may be conducted after the final OS analysis. This analysis will not be formally tested.

Current and Future Analysis

Given the recommendation by the IDMC in April 2020, all analysis specified to be conducted at the time of primary reporting will be conducted at this time.

No further statistical testing of DFS will be conducted after this unplanned interim analysis (IDMC7, DCO January 2020) and any further analysis will be considered as exploratory only. Safety data may also be reported at this time.

One further analysis with statistical testing of OS will be conducted, exploratory updates may be considered after this final OS analysis. Safety data may also be considered at this time.

8. REFERENCES

Berry, G., Kitchin, R. & Mock, P., 1991. A comparison of two simple hazard ratio estimators based on the logrank test. *Statistics in Medicine,* Volume 10, pp. 749-55.

Clopper, C. J. & Pearson, E. S., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika,* Volume 26, pp. 404-413.

Cockcroft, D. W. & Gault, M. H., 1976. Prediction of creatinine clearance from serum creatinine. *Nephron,* 16(1), pp. 31-41.

Gail, M. & Simon, R., 1985. Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics,* Volume 41, pp. 361-72.

Robins, J. M. & Finkelstein, D. M., 2000. Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests. *Biometrics,* 56(3), pp. 779-788.

Robins, J. M. & Tsiatis, A. A., 1991. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics - Theory and Methods,* 20(8).

Selke, T. & Siegmund, D., 1983. Sequential analysis of the proportional hazards model. *Biometrika,* Volume 70, pp. 315-26.

Stone, A., 2010. The application of bespoke spending functions in group-sequential designs and the effect of delayed treatment switching in survival trials. *Pharm Stat,* 9(2), pp. 151-61.

SIGNATURE PAGE

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature

Document Name: d5164c00001-sap-ed-4		
Document Title:	Statistical Analysis Plan Edition 4.0	
Document ID:	Doc ID-002968078	
Version Label:	5.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
[redacted]	[redacted]	Author Approval
[redacted]	[redacted]	Content Approval

Notes: (1) Document details as stored in ANGEL, an AstraZeneca document management system.