AstraZeneca

Statistical Analysis Plan

Study Code	D5165C00001/CAURAL
Edition Number	2.0 final
Date	21st August 2017

A Phase III, Multi-Centre, Open Label, Randomized Study to Assess the Efficacy and Safety of AZD9291 in Combination with MEDI4736 versus AZD9291 Monotherapy in Patients with Locally Advanced or Metastatic Epidermal Growth Factor Receptor T790M mutation-positive Non-Small Cell Lung Cancer who have received Prior Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Therapy (CAURAL)

A Phase III, Multi-Centre, Open Label, Randomized Study to Assess the Efficacy and Safety of AZD9291 in Combination with MEDI4736 versus AZD9291 Monotherapy in Patients with Locally Advanced or Metastatic Epidermal Growth Factor Receptor T790M mutation-positive Non-Small Cell Lung Cancer who have received Prior Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Therapy (CAURAL)

Study Statistician PPD _____ 31st August 2017
Date

A Phase III, Multi-Centre, Open Label, Randomized Study to Assess the Efficacy and Safety of AZD9291 in Combination with MEDI4736 versus AZD9291 Monotherapy in Patients with Locally Advanced or Metastatic Epidermal Growth Factor Receptor T790M mutation-positive Non-Small Cell Lung Cancer who have received Prior Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Therapy (CAURAL)

Global Product Statistician PPD 31/08/2017.

Date

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine transaminase
AST	Aspartate transaminase
ATC	Anatomical Therapeutic Chemical
BP	Blood pressure
CI	Confidence interval
CR	Complete response
CSP	Clinical study protocol
CSR	Clinical study report
CT	Computerized tomography scan
CTCAE	Common Terminology Criteria for Adverse Events
DCO	Data cut-off
DCR	Disease control rate
DNA	Deoxyribonucleic acid
DoR	Duration of response
ECG	Electrocardiogram
eCRF	Electronic case report form
EGFR	Epidermal growth factor receptor
EGFRm+	EGFR mutation positive
FAS	Full Analysis Set
HIV	Human immunodeficiency virus
IP	Investigational product
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
L858R	An amino acid substitution at position 858 in EGFR, from a Leucine (L) to an Arginine (R)

Abbreviation or special term	Explanation
LD	Longest diameter
LDH	Lactate dehydrogenase
LLoQ	Lower limit of quantification
LVEF	Left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic resonance imaging
MUGA	Multi-gated acquisition scan
N/A	Not applicable
NE	Not evaluable
NSCLC	Non-small cell lung cancer
NTL	Non-target lesion
OAE	Other significant adverse event
ORR	Objective response rate
OS	Overall survival
PD	Progression of disease
PD-L1	Programmed Death Ligand 1
PFS	Progression Free Survival
PID	Percentage intended dose
PK	Pharmacokinetics
PR	Partial response
PT	Preferred term
QD	Once daily
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious adverse event
SD	Stable disease
SLD	Sum of longest diameter
SOC	System organ class
T790M	An amino acid substitution at position 790 in EGFR, from a Threonine (T) to a Methionine (M)

Abbreviation or special term	Explanation
T790M-	T790M mutation negative
T790M+	T790M mutation positive
TKI	Tyrosine kinase inhibitor
TL	Target lesion
ULoQ	Upper limit of quantification
WHO	World Health Organization

AMENDMENT HISTORY

Date	Brief description of change
21st August 2017	To reflect changes documented in Protocol Version 2.0 02-Mar-2017.

1. STUDY DETAILS

1.1 Study objectives

Primary objective

Objective:	Outcome Measure:
To investigate the safety and tolerability profile of AZD9291 in combination with MEDI4736.	Adverse events (graded by Common Terminology Criteria for Adverse Events (CTCAE v4)) Clinical chemistry, haematology and urinalysis Vital signs (pulse and blood pressure), Physical Examination, Weight Digital Electrocardiogram (ECG) Echocardiogram/Multi Gated Acquisition Scan (MUGA) (for Left Ventricular Ejection Fraction) WHO Performance Status

Exploratory objectives

Objective:	Outcome Measure:
To assess the safety and tolerability of AZD9291 as a single agent	Adverse events (graded by CTCAEv4) Clinical chemistry, haematology and urinalysis Vital signs (pulse and blood pressure), Physical examination, Weight Centrally reviewed digital Electrocardiogram (ECG) Echocardiogram/Multi gated Acquisition Scan (MUGA) (for left ventricular Ejection fraction) WHO performance status Serious Adverse events (graded by CTCAEv4)

Objective:	**Outcome Measure:**
To obtain a preliminary assessment of the efficacy of AZD9291 in combination with MEDI4736 and AZD9291 monotherapy	Objective Response Rate (ORR) Progression Free Survival (PFS) landmark at 6 and 12 months Overall Survival (OS) landmark at 12 months Duration of Response (DoR) Disease Control Rate (DCR) Tumour shrinkage using Investigator assessments according to RECIST 1.1
To assess the PK of AZD9291 as a single agent and in combination with MEDI4736.	Plasma concentrations of AZD9291 and metabolites AZ5104 and AZ7550.
To characterise the PK and immunogenicity of MEDI4736 after single dosing and at steady state after multiple dosing when given intravenously to patients with EGFRm non small-cell lung cancer (NSCLC) in combination with AZD9291	Plasma concentrations of MEDI-4736 to characterise PK Blood samples for anti-drug antibodies (ADA) to characterise immunogenicity.

The following exploratory objectives may be reported separately from the main Clinical Study Report (CSR) and will not be covered in this Statistical Analysis Plan.	
To characterise the pharmacodynamics of MEDI4736 after single dosing and at steady state after multiple dosing when given intravenously to patients with EGFRm non small-cell lung cancer (NSCLC) in combination with AZD9291	Blood samples for soluble PD-L1 to characterise pharmacodynamics .
To investigate the relationship of response and PFS in patients who are PD-L1 positive and those that are PD-L1 negative.	Retrospective evaluation of baseline PD-L1 expression in tumour (based on immunohistrochemistry assessment).
To collect and store deoxyribonucleic acid (DNA) for future exploratory research into genes/genetic variation that may influence PK or response AZD9291 as a single agent and in combination with MEDI4736 (i.e. absorption, distribution, metabolism, excretion, safety and efficacy) and/or susceptibility to/development of cancers.	Correlation of polymorphisms with variation in Pharmacokinetics (PK), pharmacodynamics, safety or response observed in patients treated with AZD9291 as a single agent and in combination with MEDI4736.

Objective:	Outcome Measure:
To collect and store tumour samples and blood-based (plasma and serum) samples for potential exploratory research into factors that may influence susceptibility to/development of NSCLC/cancer and/or AZD9291 as a single agent and in combination with MEDI4736 (where response is defined broadly to include efficacy, tolerability or safety).	Collection of tumour and blood-based samples to include, but not be limited to, investigation of biomarkers such as Epidermal growth factor receptor (EGFR) mutations (T790M, L858R, etc). These samples may be shared with a diagnostic partner for development of a companion diagnostic if needed. The samples may also be used to investigate the relationship between PK and blood-borne biomarkers. Samples may be analysed retrospectively. Any biomarker data generated may be reported separately and may also form part of a pooled analysis with other AZD9291 studies.

1.2 Study design

This trial was initially designed as a phase III, multi-centre, open label, randomized study to assess the efficacy and safety of AZD9291 (80 mg, orally, once daily) in combination with MEDI4736 (10 mg/kg, IV infusion every 2 weeks [q2w]) versus AZD9291 monotherapy (80 mg, orally, once daily) in patients with a confirmed diagnosis of EGFR T790M mutation positive (T790M+) locally advanced or metastatic NSCLC (Stage IIIB-IV), who have progressed following prior therapy with an approved EGFR-TKI agent.

All patients must have documented radiological progression on EGFR-TKI treatment and on the last treatment administered prior to enrolling in the study.

A mandatory biopsy was needed for central testing of T790M mutation status following confirmed disease progression on previous treatment.

Three hundred fifty patients were planned to be randomized, however recruitment was stopped after enrolling 29 patients. Early termination of enrolment into the study was decided in the light of a higher than anticipated incidence of interstitial lung disease-like events in patients receiving the combination of AZD9291 with MEDI4736 in a separate multi-arm Phase 1b open label study D5160C00006 (TATTON).

Two summaries of the study data are planned:

Primary Summary
A primary summary will be performed with a data cut-off date representative of when the last country with active patients receives ethics and regulatory approval for protocol version 2.0. This summary will be used for the CSR and will report data according to the randomised or actual treatment groups as specified in Section 2.

Final Summary

Patients had been initially randomized to receive either AZD9291 in combination with MEDI4736 or AZD9291 monotherapy. Depending on the treatment being received, the patients will be categorized as follows for the purpose of delineating the required study plan and assessments:

GROUP A - Patients on AZD9291 monotherapy

Group A includes patients continuing to receive AZD9291 monotherapy following data cut-off (DCO) for the primary analysis, who were either randomised to receive AZD9291 monotherapy or were randomised to receive AZD9291 in combination with MEDI4736 but who have discontinued MEDI4736 and been followed up for 90 days post discontinuation of MEDI4736

Patients may continue dosing with the drug according to the protocol as long as they are continuing to show clinical benefit, as judged by the Investigator, and in the absence of a discontinuation criterion. These patients have no pre-scheduled visits until treatment discontinuation (apart those needed for the drug accountability and dispensing) and will follow the normal clinical site approach. Following approval of clinical study protocol (CSP) version 2.0 at the site only SAE, outcomes of pregnancy and drug dispensing/ accountability data will be collected for Group A patients. This may result in sparse data in some of the summaries in the primary summary.

GROUP B - Patients on MEDI4736 in combination with AZD9291 or MEDI4736 monotherapy

Group B includes patients randomised to receive MEDI4736 and AZD9291 and who are continuing to receive MEDI4736 either in combination with AZD9291 or as a monotherapy

Patients may continue treatment with MEDI4736 monotherapy or in combination with AZD9291 as long as they are continuing to show clinical benefit, as judged by the Investigator, and in the absence of a discontinuation criterion. Group B patients should follow the schedule / assessments specified in the CSP. Group B patients that discontinue MEDI4736 during the study can be switched to Group A and follow the AZD9291 Monotherapy assessment schedule following completion of the 90 day safety follow up. Following approval of CSP version 2.0

The overall study design is shown in Figure 1 below. The study schedule for Group A and Group B subjects is detailed in Tables 1 and 2 of the CSP, respectively.

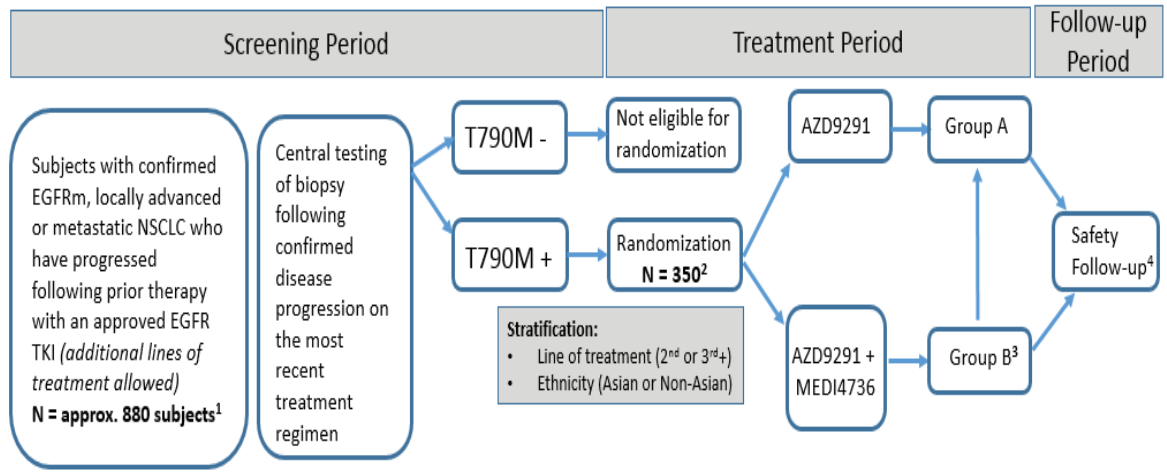
A final addendum summary will be performed when the last patient discontinues MEDI4736 and completes their required safety follow up. All outputs from the primary summary will be repeated at the final summary, except those relating to change in tumour size at 6 weeks, which will have been fully reported during the primary analysis. Only data collected between

the primary summary data cut-off and the time when last patient receiving MEDI4736 reaches the end of safety follow up, will be summarised. If three or fewer patients are included in this summary, then only listings will be produced, with no summary tables.

If all patients have discontinued from MEDI4736 treatment at the time of DCO for the primary study summary, the primary summary will be the only summary of data for the study.

Final summary will show data summarized by Group A and Group B as defined above. For demography, exposure and SAEs data will be summarized for both Groups A and B. Other data will be reported for Group B only. Patients that do not meet criteria for either Group A or Group B definition will be listed along with any relevant data that has been collected.

Figure 1 Study flow chart



1 Due to termination of recruitment, final screened patients = 60

2 Due to termination of recruitment, final randomized patients = 29

3 Patients receiving the combination treatment who discontinue MEDI4736 may be switched to Group A following the completion of the required safety follow up period

4 Patients receiving only AZD9291 at the time of treatment discontinuation should be contacted 30 days after their last dose to follow up on any existing SAEs and monitor for new SAEs that may be related to the IP

Patients who discontinue the combination treatment (AZD9291 + MEDI4736) or discontinue AZD9291 <90 days after discontinuing MEDI4736 will complete a 90 day safety follow up period

1.3 Number of subjects

The original plan was to recruit approximately 350 subjects, stratified by current line of treatment ($2^{nd}/3^{rd}$ line plus) and ethnicity (Asian/non-Asian) , randomized in a 1:1 ratio to the treatments as specified below:

- AZD9291 80mg QD
- AZD9291 80mg QD and MEDI4736 10mg/kg q2w

The study was terminated early with only 29 patients randomized. This number is not considered sufficient for statistical comparison of the treatment groups, and all data reporting will be using summaries and listings.

2. ANALYSIS SETS

2.1 Definition of analysis sets

These analysis sets apply only to the primary summary. The final summary will include only data collected between the primary summary data cut-off and the time when last patient receiving MEDI4736 reaches the end of safety follow up.

2.1.1 Full Analysis Set

The Full Analysis Set (FAS) will include all randomized subjects. The FAS will be used for all efficacy summaries and treatment groups will be presented on the basis of randomized treatment, regardless of the treatment actually received.

2.1.2 Safety Analysis Set

The Safety Analysis Set will consist of all subjects who received at least one dose of randomized treatment. Safety data will be summarized using the Safety Analysis Set, according to the treatment that patients received in their first cycle. Patients who received at least one dose of MEDI4736 in the first treatment cycle (15 +/-3 days) will be regarded as being in the combination treatment. Additional summaries might be provided for the patients who were randomised and received the combination but later discontinued MEDI4736 as appropriate.

2.1.3 PK Analysis Set

Patients in the FAS who have at least one measurable PK concentration, supported by the relevant date and time of this sample, and for each time a PK sample was taken the dosing data for that day, and for samples taken after multiple dosing the dosing data for the 2 days prior to the sample day as well as the sample day. For any individual sample to be included in the PK Analysis Set the full sample data and dosing data needs to be present for that sample.

Table 1 below shows the analysis set to be used for each outcome variable.

Table 1 Summary of outcome variables and analysis populations

Outcome variable		Populations
Demography		Safety
Efficacy Data	PFS	FAS
	ORR	
	DoR	
	DCR	
	Tumor shrinkage	
	OS	
PK data	Plasma concentrations	PK
	Metabolites	
Pharmacodynamics and immunogenicity	ADA	Safety
Safety Data	Adverse Event (AE)	Safety
	Laboratory data	
	Vital signs	
	ECG	
	Left ventricular ejection fraction (LVEF)	
	World Health Organization (WHO) performance status	

2.2 Violations and deviations

All important deviations related to the study inclusion or exclusion criteria and study conduct will be listed and summarised by randomized treatment group. None of the deviations will lead to any subjects being excluded from any of the analysis sets described in Section 2.1 (with the exception of the PK Analysis Set, if the deviation is considered to impact upon PK respectively).

The following general categories will be considered important deviations. This list is not exhaustive and additional important deviations may be added prior to database lock. A full list is provided within the study Protocol Deviations documentation.

- Informed consent procedure deviation (e.g., no informed consent signed prior to any screening procedure)
- Eligibility criteria deviation (e.g., any inclusion criteria not met or exclusion criteria met)
- Prohibited medication deviation (e.g., subject received other anticancer agents, investigational agents, or radiotherapy while on study treatment)

The categorization of these as important deviations is not automatic and will depend on duration and the perceived effect on efficacy.

In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed by the clinical team to determine any important post-entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made prior to database lock.

Incorrect stratifications will be summarized and listed, separately to the important protocol deviations. For this, a comparison will be made between the stratification factors entered into Interactive Voice Response System (IVRS) vs. the electronic case report form (eCRF) data.

A summary will be produced showing any subjects who received a treatment different to that assigned by the randomization for any duration of treatment (for example if a subject randomized to AZD9291 monotherapy but did receive at least one dose of MEDI4736). Note, this is not due to tolerability issues where subjects experience a drug interruption and/or drug delay.

Subjects who receive the wrong treatment at any time will be included in the Safety Analysis Set as described in Section 2.1.2. During the study, decisions on how to handle these situations will be made on an individual basis with written instruction from the study team leader/physician and/or statistician.

3. PRIMARY AND EXPLORATORY VARIABLES

3.1 Derivation of RECIST visit responses

All RECIST assessments, whether scheduled or unscheduled, will be included in the calculations. This is also regardless of whether a patient discontinues study drug.

For all patients, the RECIST version 1.1 (see further Appendix C of the CSP) tumour response data will be used to determine each patient's visit response. It will also be used to determine if and when a patient has progressed and also their best overall response.

The baseline tumour assessment is part of the screening procedures and should be performed within 28 days before the start of study treatment.

Prior to approval of CSP version 2.0, efficacy for all patients will be assessed by objective tumour assessments every 8 weeks (relative to the first dose) for the first 48 weeks, then every 12 weeks thereafter, until confirmed objective disease progression as defined by RECIST 1.1 (irrespective of the reason for stopping study drug and/or subsequent therapy). For patients who discontinue study drug due to toxicity or a reason other than confirmed progression of disease (PD), objective tumour assessments should be continued every 8 weeks for 48 weeks (relative to first dose) then every 12 weeks thereafter until confirmed objective disease

progression. If an unscheduled assessment is performed, and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits. Patients on monotherapy AZD9291 will be assessed according to local practice until objective disease progression.

Disease progression requires confirmation, the confirmatory scan should occur preferably at the next scheduled visit and no earlier than 4 weeks after the initial assessment of PD in the absence of clinically significant deterioration. Study treatment will continue between the initial assessment of progression and confirmation for progression.

Progression would be considered confirmed if the following criteria are met:

- ≥20% increase in the sum diameters of target lesions (TLs) compared with the nadir at 2 consecutive visits with an absolute increase of at least 5 mm.

  The assessment of progression of ≥20% increase in the sum diameters of TLs compared with the nadir is at the first progression time point relative to the nadir (the smallest sum of diameters and this may be at baseline or subsequent follow-up visit). The confirmed scan confirms the persistence of the ≥20% increase relative to the nadir.

- And/or significant progression (worsening) of non target lesions (NTLs) or new lesions at the confirmatory PD time point compared with the first time point where progression of NTLs or new lesions identified.

- And/or additional new unequivocal lesions at the confirmatory PD time point compared with the first time point new lesions identified.

Patients who discontinue treatment for reasons other than progression would undergo a 3-month follow-up. Progressions occurring during this follow-up will not necessarily have a second confirmatory scan.

Following approval of CSP version 2.0, objective tumour assessments every 8 or 12 weeks will only be carried out for Group B patients up to disease progression or up to 3 months following treatment discontinuation, and confirmation of progression will only be required for Group B patients who continue to receive MEDI4736 following progression , while patients discontinuing treatment may not have scans to confirm progression. Following approval of CSP version 2.0, patients in Group A will not have tumour assessment scans.

RECIST 1.1 without confirmation of progression will be regarded as primary in terms of the efficacy reporting. A sensitivity summary requiring confirmation will be conducted. The confirmation should occur at least 28 days from the first scan documenting radiological progression.

At each visit for the site investigator data, an overall visit response will be programmatically determined - using the information from TLs, NTLs and new lesions.

RECIST outcomes (i.e., PFS and ORR etc) will be calculated using a computer program for site investigator data.

RECIST results based on blinded independent central review were planned in the original protocol, but were not included in version 2.0 of the protocol, and no such review was performed.

3.1.1 Investigator assessment using RECIST 1.1

3.1.1.1 Target lesions

Measurable disease is defined as having at least one measurable lesion which is ≥10 mm in the longest diameter (LD) (except lymph nodes which must have short axis ≥15 mm) with CT or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements.

A patient can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (representative of all lesions involved suitable for accurate repeated measurement) and these are referred to as TLs. If more than one baseline scan is recorded then measurements from the one that is closest to randomisation will be used to define the baseline sum of TLs. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

All other lesions (or sites of disease) not recorded as TL should be identified as NTL at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits.

If a patient has had a tumour assessment which cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response will be assigned as PD).

Please refer to CSP Appendix C for the definitions of complete response (CR), partial response (PR), stable disease (SD) and PD and the derivations of overall visit response using the information from target lesions (TL), non-target lesions (NTL) and new lesions.

Rounding of TL data

For calculation of PD and PR for TLs, percentage changes from baseline and previous minimum should be rounded to 1 decimal place before assigning a TL response. For example, 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%.

Missing TL data

For a visit to be evaluable, all TL measurements should be recorded. However, a visit response of PD should be assigned if any of the following occurred:

- A new lesion is recorded;
- A NTL visit response of PD is recorded;
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥5 mm from nadir even assuming the non-recorded TLs have disappeared.

The nadir can only be taken from assessments where all the TLs had an LD recorded.

Lymph nodes

For lymph nodes, if the size reduces to <10 mm, these are considered non-pathological.

However a size will still be given and this size should still be used to determine the TL visit response as normal. In the special case where all lymph nodes are <10 mm and all other TLs are 0 mm, although the sum may be >0 mm, the calculation of TL response should be over-written as a CR.

TL visit responses subsequent to CR

A CR response can only be followed by CR, PD or NE. If a CR has occurred, the following rules at the subsequent visits must be applied:

- Step 1: If all lesions meet the CR criteria (i.e., 0 mm or <10 mm for lymph nodes), the response will be set to CR irrespective of whether the criteria for PD of TL is also met (i.e., if a lymph node LD increases by 20% but remains <10 mm).
- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e., 0 mm or <10 mm for lymph nodes), the response will be set to NE irrespective of whether when referencing the sum of TL diameters the criteria for PD is also met.
- Step 3: If not all lesions meet the CR criteria, and the sum of lesions meets the criteria for PD, the response will be set to PD.
- Step 4: If, after steps 1 through 3, a response cannot be determined, the response will be set to remain as CR.

TL too big to measure

If a TL becomes too big to measure, this should be indicated in the database and a size ('*x*') above which it cannot be accurately measured should be recorded. If using a value of *x* in the calculation of TL response would not give an overall visit response of PD, then this will be flagged and reviewed by the study team blinded to treatment assignment. It is expected that a visit response of PD will remain in the vast majority of cases.

TL too small to measure

If a TL becomes too small to measure a value of 5 mm will be entered into the database and used in TL calculations, unless the radiologist has indicated and entered a smaller value that can be measured reliably. If a TL response of PD results then this will be reviewed by the study team blinded to treatment assignment.

Irradiated lesions/lesion intervention

Previously irradiated lesions (i.e., lesion irradiated prior to entry into the study) should be recorded as NTLs and should not form part of the TL assessment.

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation/palliative surgery/embolisation), should be handled in the following way:

- Step 1: the diameters of the TLs (including the lesions that have had intervention) will be summed and the calculation will be performed in the usual manner. If the visit response is PD this will remain as a valid response category.

- Step 2: If there was no evidence of progression after step 1, the lesion diameter (for those lesions with intervention) will be treated as missing and scaled up as described previously as long as there remain ≤1/3 of the TLs with missing measurements. If the scaling results in a visit response of PD then the subject will be assigned a TL response of PD.

- Step 3: If after both steps PD has not been assigned, then a scaled sum of diameters will be calculated, treating the lesion with intervention as missing, and PR or SD then assigned as the visit response. Subjects with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10 mm for lymph nodes) and the lesions that have been subject to intervention also have a value of 0 recorded.

At subsequent visits the above steps will be repeated to determine the TL and overall visit response. When calculating the previous minimum, lesions with intervention should be treated as missing and scaled up (as per step 2 above).

Once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumors.

Lesions that split in two

If a TL splits in two, then the LDs of the split lesions should be summed and reported as the LD for the lesion that split.

Lesions that merge

If two TLs merge, then the LD of the merged lesion should be recorded for one of the TL sizes and the other TL size should be recorded as 0mm.

Change in method of assessment of TLs

Computerized tomography (CT) scan and MRI are the only methods of assessment that can be used within this trial. If a change in method of assessment occurs between CT scan and MRI this will be considered acceptable and no adjustment within the programming is needed.

3.1.1.2 NTLs and new lesions

At each visit an overall assessment of the NTL response should be recorded by the investigator. This section provides the definitions of the criteria used to determine and record overall response for NTL at the investigational site at each visit.

Please refer to CSP Appendix C for the definitions of CR, PR, SD and PD and the derivations of overall visit response using the information from target lesions (TL), non-target lesions (NTL) and new lesions.

To achieve 'unequivocal progression' on the basis of NTLs, there must be an overall level of substantial worsening in non-target disease such that, even in the presence of SD or PR in TLs, the overall tumour burden has increased sufficiently to merit a discontinuation of therapy. A modest 'increase' in the size of one or more NTLs is usually not sufficient to qualify for unequivocal progression status.

Details of any new lesions will also be recorded with the date of assessment. The presence of one or more new lesions is assessed as progression.

A lesion identified at a follow up assessment in an anatomical location that was not scanned at baseline is considered a new lesion and will indicate disease progression.

The finding of a new lesion should be unequivocal: i.e. not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumour.

New lesions will be identified via a Yes/No tick box. The absence and presence of new lesions at each visit should be listed alongside the TL and NTL visit responses.

A new lesion indicates progression so the overall visit response will be PD irrespective of the TL and NTL response.

If the question 'Any new lesions since baseline' has not been answered with Yes or No and the new lesion details are blank this is not evidence that no new lesions are present and should be treated as NE in the derivation of overall visit response.

Symptomatic deterioration is not a descriptor for progression of NTLs: it is a reason for stopping study therapy and will not be included in any assessment of NTLs.

Patients with 'symptomatic deterioration' requiring discontinuation of treatment without objective evidence of disease progression at that time should continue to undergo RECIST 1.1 tumour assessments where possible until objective disease progression is observed.

3.1.1.3 Overall visit response

Table 2 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 2 Overall visit response

Target Lesion	Non-target Lesion	New lesion(s)	Overall response
CR	CR	No	CR
CR	NA	No	CR
NA	CR	No	CR
CR	Non CR/Non PD	No	PR
CR	NE	No	PR
PR	Non PD or NE	No	PR
SD	Non PD or NE	No	SD
NA	Non CR/Non PD	No	SD (Non CR/Non PD)
NE	Non PD or NE	No	NE
NA	NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD
NA	NA	No	NED

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable, NED = no evidence of disease, NA = not applicable (only relevant if there were no TL/NTL at baseline).

3.2 Efficacy outcome variables

Following approval of CSP version 2.0 tumour based efficacy assessments will not be collected for Group A patients.

3.2.1 Progression free survival landmark summary

The endpoint PFS, as assessed by the investigator, using RECIST 1.1 will be reported both with and without confirmation of progression. will be defined as the time from the date of randomisation until the date of the first objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to progression (i.e., date of PFS event or censoring – date of randomisation + 1).

Patients who have not progressed or died at the time of summary will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment. However, if

the patient progresses or dies after 2 or more missed visits, the patient will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the two missed visits with the exception of 2 missed visits from baseline (i.e. if the subject dies within 119 days of the last evaluable visit), the event of death will be used in the PFS summary).

Given the scheduled visit assessment scheme (i.e. eight-weekly for the first 48 weeks then twelve-weekly thereafter) the definition of 2 missed visits will change. If the previous RECIST assessment is less than study day 232 (i.e. week 32 + 7 day visit window) then two missing visits will equate to 17 weeks (2 x 8 week scans + 7 day visit window) since the previous RECIST assessment, allowing for late visits.

If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from eight-weekly to twelve-weekly this will equate to 21 weeks. The time period for the previous RECIST assessment will be from study day 287. From week 49 (week 48 + 7 days) onwards, two missing visits will equate to 25 weeks (2 x 12 week scans + 7 day visit window).

If the patient has no evaluable visits or does not have baseline data they will be censored at baseline unless they die within 112 days of baseline (2 x 8 week scans).

The PFS time will always be derived based on scan/assessment dates not visit dates.

RECIST 1.1 assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- For investigational assessments, the date of progression will be determined based on the earliest of the RECIST assessment/scan dates of the component that indicates progression

Note: For TLs, only the latest scan date is recorded out of all scans performed at that assessment for the TLs and similarly for NTLs only the latest scan date is recorded out of all scans performed at that assessment for the NTLs. Additionally, PFS will be obtained using the algorithm described above for the RECIST site investigator tumour data. An additional sensitivity summary using a modified version of the algorithm will be used, whereby any objective disease progression must be confirmed by the next scheduled scan. The confirmatory scan must be no sooner than 4 weeks after the initial suspected progression. If disease progression is confirmed (or disease progression occurs and no further scans are recorded) then the date of progression will be when it was originally observed. Patients with a single disease progression and no further tumour assessment scans will be treated as PD in the summary.

The proportion of patients alive and progression free at 6 months (i.e., APF6) will be defined as the Kaplan-Meier (KM) estimate of PFS at 6 months.

The proportion of patients alive and progression free at 12 months (i.e., APF12) will be defined as the KM estimate of PFS at 12 months.

3.2.2 Objective response rate (ORR)

ORR (per RECIST 1.1 as assessed by the investigator) is defined as the number and percentage of patients with measurable disease with at least 1 visit response of CR or PR. Summaries will be produced based on both unconfirmed and confirmed response, where the response is confirmed at least 4 weeks later. ORR will be based on all randomised patients. Therefore, data obtained up until progression, or the last evaluable assessment in the absence of progression, will be included in the assessment of ORR.

A subject who goes off treatment without progression to receive a subsequent anti-cancer therapy, and then responds (CR or PR), will not be included as a responder in the ORR.

3.2.3 Duration of response (DoR)

DoR (per RECIST 1.1 as assessed by the investigator) will be defined as the time from the date of first documented response (CR or PR) until the first date of documented progression or death in the absence of disease progression, whichever is earlier (i.e., date of PFS event or censoring – date of first response + 1). The end of response should coincide with the date of progression or death from any cause used for the RECIST 1.1 PFS endpoint).

The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of CR or PR. If a patient does not progress following a response, then the corresponding DoR will be censored at the PFS censoring time. Summaries will be produced based on both confirmed and unconfirmed progression.

DoR will not be defined for those patients who do not have documented response.

3.2.4 Disease control rate (DCR)

DCR is defined as the percentage of subjects who have a best overall response of CR or PR or SD at ≥4 weeks, prior to any PD event. Summaries will be produced based on both confirmed and unconfirmed response.

A subject's best overall response is calculated as defined below.

3.2.5 Best overall response

Best overall response (BOR) is calculated based on the overall visit response from each RECIST assessment. It is the best response a patient has had following randomisation but prior to starting any subsequent anti-cancer therapy and prior to RECIST progression or the last evaluable assessment in the absence of RECIST progression and subsequent anti-cancer therapy.

Categorisation of BOR will be based on RECIST using the following response categories: CR, PR, SD, PD and NE.

BOR will be determined programmatically based on RECIST from the overall visit response at each visit using investigator assessment data including all data up until the first progression,

the start of any subsequent cancer therapy or the last evaluable assessment in the absence of progression.

For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. SD should be recorded at least 16 weeks +/- 1 week, i.e. at least 119 days (to allow for the assessment window), after randomisation (i.e. study day 120). For CR/PR, the initial overall visit assessment which showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

For patients whose PFS event is death, BOR will be calculated based upon all evaluable RECIST assessments prior to death.

For patients who die with no evaluable RECIST assessments, if the death occurs ≤17 weeks (i.e., 16 weeks + 1 week to allow for a late assessment within the assessment window) after randomisation, then BOR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST assessments, if the death occurs >17 weeks (i.e., 16 weeks + 1 week) after the date of first dose then BOR will be assigned to the NE category.

Progression events that have been censored due to them being more than two missed visits after the last evaluable assessment will not contribute to the BOR derivation.

Summaries will be produced based on both confirmed and unconfirmed response.

3.2.6 Tumour shrinkage

Tumour shrinkage will be assessed using RECIST 1.1 tumour response. The absolute change and percentage change from baseline in sum of tumour size at each assessment will be calculated. Tumour size is the sum of the longest diameters (SLD) of the TLs. The percentage change in SLD at each week for which data are available will be obtained for each subject taking the difference between the SLD at each week and the SLD at baseline divided by the SLD at baseline multiplied by 100 (i.e., [week n – baseline] / baseline x 100). The change from baseline will be obtained for each subject taking the difference between the SLD at each week and the SLD at baseline (i.e., week n - baseline).

The best percentage change in SLD from baseline or the minimum increase from baseline in the absence of a reduction from baseline based on all post-baseline assessments prior to progression or start of subsequent anti-cancer therapy.

If best percentage change cannot be calculated due to missing data, a value of +20% will be imputed as the best percentage change from baseline in the following situations (otherwise best percentage change will be left as missing):

- If a subject has no post-baseline assessment and has died;
- If a subject has new lesions or progression of NTLs;
- If a subject has withdrawn due to PD and has no evaluable TL data before or at PD.

3.2.7 Overall survival landmark summary

OS is defined as the time from the date of randomisation until death due to any cause (i.e., date of death or censoring – date of randomisation + 1). Any patient not known to have died at the time of the summary will be censored based on the last recorded date on which the patient was known to be alive (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls which were originally planned for the 2 weeks following the date of data cut-off (DCO), will no longer be made following approval of CSP version 2.0.

The proportion of patients alive at 12 months (i.e., OS12) will be defined as the KM estimate of OS at 12 months

3.3 Safety variables

Following approval of CSP version 2.0 safety assessments will not be collected for Group A patients, except for SAEs.

3.3.1 Adverse Events

All AEs will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA) and graded according to the National Cancer Institute CTCAE (version 4.0.3 or later).

Preferred terms used to identify adverse events of special interest (AESI) will be listed before database lock and documented in the Study Master File. Groupings of certain MedDRA preferred terms will be based on preferred terms provided by the medical team and a listing of the preferred terms in each grouping will be produced.

The grouped preferred terms for AESI for AZD9291 are shown below. A separate list will be produced for MEDI 4736. The individual preferred terms will be available in the programming specification documentation:

- Ocular surface effects
- Dry eye/conjunctivitis
- Interstitial lung disease and acute interstitial pneumonitis
- Nail and nail bed conditions
- Skin disorders including dermatitis
- Upper GI tract inflammatory events
- Cardiotoxicity
- Diarrhoea

- Colitis/Enterocolitis (Gastrointestinal Disorder)
- Hepatitis/ Hepatic Toxicity (Hepatic Function Abnormality)
- Infusion-related Reactions
- Anaphylaxis and Serious Allergic Reactions (Hypersensitivity)
- Neuromuscular Toxicity/Neuropathy
- Endocrinopathy

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

Any AE occurring within 30 days of discontinuation of AZD9291 and within 90 days for MEDI4736 will be included in the AE summaries. Any events in this period that occur after a subject has received further therapy for cancer (following discontinuation of AZD9291) will be flagged in the data listings. If AZD9291 is discontinued less than 60 days after MEDI4736, subjects should be followed up to 90 days from the date of MEDI4736 discontinuation.

3.3.2 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 treatment 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, blood pressure [BP] and weight), and/or 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.

3.3.3 Laboratory data

Laboratory data will be collected throughout the study, from screening to follow-up visit as described in the CSP. Blood and urine samples for determination of haematology, clinical chemistry, and urinalysis will be collected as described in Section 5.2.1 of the CSP. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 4 below will be used. Values recorded as "<*x*" will be reported as the maximum number of this range, i.e. *x*.

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

Local reference ranges will be used for the primary interpretation of laboratory data at the local laboratory. Project reference ranges will be used throughout for reporting purposes.

3.3.4 Vital signs

Vital signs data obtained up until the 30 days from date of last dose of study treatment will be used for reporting. Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 4 below will be used.

3.3.5 ECGs

Baseline is the pre-dose time point on the first dosing day (Day 1 Cycle 1). For each time point three ECG recordings should be taken within an approximate 5 minute period. The variables to be reported from the continuous ECG measurements are RR, PQ, QRS, and QT intervals. The QT interval will be corrected for RR (the duration of a heart beat) to obtain corrected (QTc) variables.

QTcF will be calculated using the formula:

$$QTcF=QT/RR^{b}$$

where b=1/3.

To obtain a single value of QTcF, RR, PR, QRS and QT at each specified time point, the mean of the triplicate values at that time point will be used. For each subject, the change-from baseline in an ECG variable at each time point will be calculated as the difference between the mean of the replicate value at each post-dose time point and the mean of pre-dose baseline replicate value. All ECG data will be obtained from the vendor, eRT, and there will be no site interpretation.

QTc outliers are defined as QTcF values following dosing that are greater than 450 ms or are increases from baseline greater than 30 ms.

3.3.6 Left ventricular ejection fraction

Echocardiogram/MUGA data will be collected every 16 weeks throughout the study, from screening to treatment discontinuation as described in the CSP.

Change from baseline in LVEF will be calculated for each post-dose visit on treatment.

Abnormal LVEF are defined as values less than the lower limit of normal (project reference).

Absolute values will also be flagged if they meet either of the following criteria (considered as outliers):

- ≥10 percentage points decrease from baseline and <50% LVEF, or

- ≥15 percentage points decrease from baseline and ≥50% LVEF.

3.3.7 Duration of exposure and duration of treatment

Total and actual exposure will be calculated separately for AZD9291 or MEDI4736. Total exposure to each drug (AZD9291 or MEDI4736) will be time (days) from the first dose to the last dose:

Total exposure = (last dose date where dose >0 mg - first dose date) + 1

Actual exposure to each drug (AZD9291 or MEDI4736) will be time (days) from first dose to the last dose, taking account of dose interruptions.

Actual exposure = ((last dose date where dose >0 mg - first dose date) + 1) - total duration of dose interruption (i.e., number of days with dose = 0mg).

Dose reductions are not permitted per the CSP for MEDI4736. The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

The duration of treatment will be the time (days) from the first dose to the last dose of either treatment:

The last dose date for the combination arm is defined as the latest of the last non-zero dose of either agent. For the AZD9291 monotherapy arm, this will be the last non-dose dose of AZD9291.

Duration of treatment = (last dose date where dose >0 mg - first dose date) + 1

3.3.8 Dose intensity

Relative dose intensity (RDI) is the percentage of the actual dose intensity delivered relative to the intended dose intensity through to treatment discontinuation. Percentage intended dose (PID) is the percentage of the actual dose delivered relative to the intended dose through to progression. Both will be derived using the date of objective disease progression as defined by RECIST v1.1 using the investigator site assessments. If the investigator considered that it was in the patient's best interest to continue study treatment past this time, this was not included in the derivation of RDI and PID.

RDI and PID will be defined as follows:

- RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the earlier of progression (or a censoring event) or the actual last day of dosing and D is the intended cumulative dose up to the earlier of progression (or a censoring event) or the actual last day of dosing.

- PID = 100% * d/D, where d is the actual cumulative dose delivered up to progression (or a censoring event) and D is the intended cumulative dose up to progression (or a

censoring event). D is the total dose that would be delivered, if there were no modifications to dose or schedule.

RDI and PID for AZD9291 and MEDI4736 will be calculated for the entire intended treatment period (censored at data cut-off).

3.4 PK and pharmacodynamics

3.4.1 Plasma concentrations of AZD9291 and metabolites AZ5104 and AZ7550

All plasma concentrations of AZD9291, AZ5104 and AZ7550 will be listed. Summaries of plasma concentrations by nominal sample window will be presented for the PK Analysis Set. Following approval of CSP version 2.0, PK samples were only collected in the combination treatment group.

The ratio of metabolite to AZD9291 in each PK sample will be listed. Summaries of ratio of metabolite to AZD9291 by nominal sample window will be presented for the PK Analysis Set.

The following summary statistics will be presented:

- The geometric mean (gmean, calculated as exp [μ], where μ is the mean of the data on a logarithmic scale).
- Coefficient of variation (CV, calculated as 100 √ [exp(s2)-1], where s is the standard deviation of the data on a log scale).
- Gmean ± standard deviation (calculated as exp[μ±s]).
- Geometric Standard Deviation (GeoSD) calculated using transformed data.
- Arithmetic mean calculated using untransformed data.
- Standard Deviation calculated using untransformed data.
- Minimum, median, maximum and number of observations (n).

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, geoSD, arithmetic mean and standard deviation 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, geoSD, arithmetic mean and standard deviation will be reported as not calculable (NC). The max value will be reported from the individual data, and the min and median will be set as NQ.· If all the concentrations are NQ, the gmean and arithmetic mean will be reported as NQ and the CV, geoSD and standard deviation as NC.

The number of values above lower limit of quantification (LLOQ) will be reported for each time-point along with the total number of collected values. If data are available for less than

3 patients, no summary statistics other than minimum, maximum and n will be presented. The plasma concentration data for AZD9291 and its metabolites will also be analysed using a population PK approach, which may include exploring the influence of covariates on PK, if the data allows.

A pharmacokinetic-pharmacodynamic approach may be used to investigate the relationship between PK and selected primary, secondary and/or exploratory endpoints, where deemed appropriate. These results of these analyses will be reported separately from the study CSR. A separate data analysis plan will be prepared to describe such analyses.

3.4.2 Serum concentrations of MEDI4736

PK analysis of the serum concentration data for MEDI4736 will be performed by Clinical Pharmacology, AstraZeneca/MedImmune or delegate on behalf of Clinical Pharmacology.

The actual sampling times will be used in the PK calculations. Pharmacokinetic concentration data and summary statistics will be tabulated. Individual and mean blood MEDI4736 concentration-time profiles will be generated. PK parameters will not be reported as the data collected will not be sufficient for their determination.

3.4.3 Pharmacodynamics and immunogenicity for ADA

Analyses to evaluate the pharmacokinetics and immunogenicity of MEDI4736 will be performed by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

Immunogenicity results will be analysed descriptively by summarizing the number and percentage of patients who develop detectable anti-MEDI4736 antibodies. The immunogenicity titre will be reported for samples confirmed positive for the presence of anti-MEDI4736 antibodies. Summaries will be based upon all patients from the safety population who have at least one measurable ADA blood sample.

4. SUMMARY METHODS

All reporting of data will be descriptive, using summary tables and listings. No statistical hypothesis testing will be performed.

Due to the small numbers of patients, summaries will not be grouped by the stratification factors involved in the randomization.

4.1 General principles

Continuous data will be summarized using descriptive statistics (number of observations, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum). Frequencies and percentages will be used for summarizing categorical (discrete) data. Confidence intervals (CIs), when presented, will generally be 2-sided and constructed at the 95% level. For binomial variables, exact methods will be employed unless otherwise specified.

A month is operationally defined to be 30.4375 days.

For the primary summary, data will be presented in data listings by the randomized treatment group and subject number. All summaries will be presented by randomized treatment group unless otherwise specified.

For the final summary, data will be presented in data listings by Group A / Group B and subject number. All summaries will be presented by Group A / Group B unless otherwise specified.

4.1.1 Baseline measurements and change from baseline variables

In general, for efficacy endpoints the last observed measurement prior to randomization will be considered the baseline measurement. For safety endpoints the last observation before 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 the first dose where neither time nor a nominal pre-dose indicator are captured will be considered prior to the first dose if such procedures are required by the protocol to be conducted before the first dose.

In all summaries change from baseline variables will be calculated as the post-treatment value minus the value at baseline. The % change from baseline will be calculated as (post-baseline value - baseline value) / baseline value x 100.

4.1.2 Study day definitions

For the purpose of efficacy data summary, Study Day 1 is defined as the date of randomization to study treatment. For visits (or events) that occur on or after randomization, study day is defined as (date of visit [event] - date of randomization + 1). For visits (or events) that occur prior to randomization, study day is defined as (date of visit [event] - date of randomization). There is no Study Day 0.

For the purpose of safety data summary, 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 (date of visit [event] - date of first dose of study treatment + 1). For visits (or events) that occur prior to first dose, dose day is defined as (date of visit [event] - date of first dose of study treatment). There is no Dose Day 0.

For listings (such as for AEs) that include the derivation of “days since last dose,” this is defined as (event date - date of last dose). Events that occur on the same day as the last dose of study drug will therefore be described as occurring zero days from the last dose of study drug.

4.1.3 Visit windows

For summaries of vital signs, laboratory data, ECG and target lesion size, 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 have the potential to be summarized. Inclusion within the visit window should be based on the actual date and not the intended date of the visit. For summaries at a subject level, all values should be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a subject 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 subject; they should also highlight the value for that subject that was used in the summary table, wherever feasible.

For visit based summaries:

- If there is more than one value per subject within a visit window then the closest to the planned study day value should be summarized, 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 subjects in either treatment group have data recorded at a particular visit.

4.1.4 Handling missing data

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

4.1.4.1 Imputation of partial dates

Initial diagnosis date:

- If year is missing (or completely missing), do not impute.

- If only day is missing, impute day as 15th of the month.

- If day and month are missing, impute as July 1st.

Concomitant medication start date

- 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 end date

- 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.

4.1.5 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 x 0.99 for reporting 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 x 1.01 for reporting purposes. The original value will be listed.

4.1.6 Rounding rules for reported percentages

For percentages ≥10%:

- Values ≥X.5 or above round to X+1.
- Values >X but <X.5 round to X.

For percentages <10%:

- Values ≥X.Y5 or above round to X.Y+0.1.
- Values >X.Y but <X.Y5 round to X.Y.

4.2 Summary methods

4.2.1 Subject disposition and data sets analyzed

Subject disposition will be listed and summarized for all patients. Summaries will include the number and percentage of subjects:

- Randomized

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

In addition, the number and percentage of subjects who discontinued treatment and who discontinued the study, including a breakdown of the main reason for discontinuation will be presented for all subjects.

4.2.2 Protocol deviations

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

4.2.3 Demographic and other baseline characteristics

Demographic and baseline subject characteristics will be listed and summarized for the Safety Analysis Set. Standard descriptive statistics will be presented for the continuous variables of:

- Age (years).
- Weight (kg).
- Height (cm).
- Body mass index (kg/m2) (calculated as [weight/height^2] where weight is in kg and height is in m).

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

- Age group (years) (grouped as <50, ≥50 to <65, ≥65 to <75 and ≥75).
- Sex.
- Race.
- Ethnicity.
- Overall disease classification (metastatic, locally advanced, both).
- Site of disease.
- Baseline TL size (mean and categories: <40, 40-79, 80-119, and ≥120mm).
- Histology.
- Smoking status.

- WHO performance status (0/1).

Receptor status at baseline will be summarised.

4.2.4 Medical history

Disease related medical history and relevant surgical history will be coded using MedDRA. All disease related medical and surgical history will be listed.

4.2.5 Anti-cancer medications and radiotherapy (prior, during, post-treatment with investigational product)

Medications received prior to, concomitantly, or post-treatment will be coded using the AstraZeneca Drug Dictionary Anatomical Therapeutic Chemical (ATC) Classification codes. Concomitant medications/radiotherapy will be listed and summarized for the Safety Analysis Set by ATC classification codes.

For the purpose of inclusion in prior and/or concomitant mediation or therapy summaries, incomplete medication or radiotherapy start and stop dates will be imputed as detailed in Section 4.1.5.1.

Prior medications, concomitant and post-treatment medications are defined based on imputed start and stop dates as follows:

- Prior medications are those taken 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-treatment medications are those with a start date after the last dose date of study treatment.

EGFR-TKI therapies will be defined as approved EGFR-TKI therapies and will include gefitinib, erlotinib and afatinib.

4.2.5.1 Anti-cancer therapy

All anti-cancer therapies at baseline and post discontinuation will be summarized for the Safety Analysis Set. The number of anti-cancer therapy regimens and number of cycles will be summarized using descriptive statistics. The anti-cancer therapy agent, therapy class, best response, and treatment status will be summarized using frequencies and percentages. Any anti-cancer therapies will also be listed.

4.2.5.2 Radiotherapy

All radiotherapies will be summarized for the Safety Analysis Set. Any radiotherapy will also be listed.

4.2.6 Exposure

Exposure will be listed and summarized for Safety Analysis Set. The following summaries will be produced:

- Summary of duration of exposure of AZD9291
- Summary of duration of exposure of MEDI4736
- Summary of duration of treatment for each arm
- RDI and PID of AZD9291
- RDI and PID of MEDI4736
- Summary of interruptions and reductions of AZD9291
- Summary of interruptions and delays of MEDI4736

4.2.7 Efficacy

All efficacy summaries will be produced using the FAS.

4.2.7.1 Progression free survival landmark summary

Progression status assessed by investigator at time of landmark summary (at Month 6 and Month 12 from the last subject first dose) will be summarised.

Kaplan-Meier plots and estimates will be produced at time of landmark summary along with other descriptive summaries based on the FAS.

4.2.7.2 Sensitivity summary

The following sensitivity summary will be conducted:

(a) With confirmation of progression: RECIST 1.1 criteria will be used to assess PFS with the requirement for confirmation of progression (by investigator). Progressions will only be used in the summary where there is a scan no earlier than 4 weeks later that documents progression or where there is no following scan. The first scan documenting radiological progression will be used regardless.

4.2.7.3 Subgroup summary

No subgroup summaries are planned.

4.2.7.4 ORR

The ORR will be based on the programmatically derived RECIST outcome using the investigator data.

ORR will be analyzed in terms of the proportion of responders out of the number of available patients per treatment arm. The associated 95% confidence intervals will also be provided using the Clopper-Pearson exact method for binomial proportions.

4.2.7.5 OS

The proportion of patients alive at 12 months will be reported. Kaplan-Meier plots and estimates will be produced at the time of summary.

4.2.7.6 DoR

DoR will be reported using summary statistics (n, median and quartiles) and KM plots to describe the duration of response in those patients with an initial response on study.

4.2.7.7 DCR

DCR will be presented as the proportion of patients with disease control, out of the number of patients available per treatment arm. The associated 95% confidence intervals will also be provided using the Clopper-Pearson exact method for binomial proportions.

4.2.7.8 Tumour shrinkage

The best absolute change in target lesion tumour size from baseline, and percentage change in in target lesion tumour size from baseline will be summarised using descriptive statistics and presented at each time point by randomized treatment group.

4.2.8 Safety outcomes

The Safety Analysis Set will be used for all safety and tolerability tables, figures and listings except were expressly noted.

4.2.8.1 Adverse events

All AEs, both in terms of MedDRA PT and CTCAE grade, will be listed and summarized descriptively by count (n) and percentage (%) for each treatment group. MedDRA will be used for coding. Any AE occurring before study treatment (i.e., before Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as ‘pre-treatment’. The summary tables will include all AEs that occurred after the start of treatment up until the end of the follow-up period.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator’s assessment of severity and relationship to study drug. Frequencies and percentages of subjects reporting each preferred term will be presented (i.e., multiple events per subject will not be accounted for apart from on the episode level summaries).

Summary information (the number and percent of subjects by treatment) 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 study treatment
- Fatal AEs
- All SAEs
- AEs leading to discontinuation
- OAEs

An overall summary of the number and percentage of subjects in each category will be presented, as will an overall summary of the number of episodes in each category. In addition, a truncated AE table of the most common AEs, showing all events that occur in at least 5% of subjects overall will be summarized by preferred term, 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 subjects will be provided by maximum reported CTCAE grade, SOC, PT and actual treatment group.

In addition, AEs with an outcome of death, SAEs, and OAEs will be listed.

4.2.8.2 AEs of special interest

Summary tables of AESIs will be produced. The number (%) of patients experiencing any of the specified terms will be presented overall, by maximum CTCAE grade, and by outcome.

4.2.8.3 Deaths

A summary of deaths will be provided with the number and percentage of subjects, categorized as:

- Related to disease under investigation,
- Fatal AE,
- Both related to disease under investigation and fatal AE,
- Fatal AE >30 days after last treatment dose,
- Deaths >30 days after last treatment dose, unrelated to AE or disease under investigation, and
- Subjects with unknown reason for death.

A corresponding listing will also be produced.

4.2.8.4 Laboratory evaluations

All laboratory data recorded in the eCRF will be listed. If any additional analytes to those in Table 5 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 reference ranges provided by the local or central laboratory. As applicable, values will be graded using CTCAE v4.0.

Plots of both the maximum post-baseline alanine transaminase (ALT) and aspartate transaminase (AST) versus the maximum post-baseline total bilirubin, expressed as multiples of their upper limit of reference range will be produced.

The laboratory variables to be summarised are shown in Table 3 below.

Table 3 Laboratory safety variables

Clinical chemistry (Serum/Plasma)	**Hematology (Blood)**	**Urinalysis (Urine)**
Albumin	Hemoglobin	Glucose
Alanine transaminase (ALT)	Leukocytes	Protein
Aspartate transaminase (AST)	Hematocrit	Blood
Alkaline phosphatase	Red blood cell (RBC) count	
Bilirubin, total	Absolute leukocyte differential count:	
Calcium, total	Neutrophils	
Creatinine	Lymphocytes	
Glucose	Monocytes	
Magnesium	Basophils	
Phosphate	Eosinophils	
Potassium	Platelet count	
Sodium	Reticulocytes	
Urea nitrogen/Blood urea nitrogen (BUN)		
TSH[1]		
Lactate dehydrogenase (LDH)[2]		
Hepatitis B surface antigen[2]		
Hepatitis C antibody[2]		
Human immunodeficiency virus (HIV) antibody[2]		
Activated partial thromboplastin time[2]		
International normalised ratio[2]		

[1] Free T3 and Free T4 to be measured only if TSH is abnormal.

[2] measured at the screening visit only

4.2.8.5 Vital signs (pulse and BP) and weight

All vital signs data will be listed. Absolute and percentage change from baseline for pulse, BP and weight will be summarized by treatment group, overall and by visit.

4.2.8.6 Physical examination

Abnormalities identified from physical examination will be listed.

4.2.8.7 ECG

All ECG data received will be presented in data listings. QTc summaries will be presented for subjects in the Safety Analysis Set. The ECG parameters that will be summarized (absolute values, change and percentage change from baseline) are: QTcF, RR, PR, QRS and QT.

The observed values and change from baseline in QTcF will be summarized by visit using summary statistics.

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 summarized 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 subjects who meet the ECG outlier criteria at any assessment post-date of first dose will be summarized.

4.2.8.8 Left ventricular ejection fraction (LVEF)

Abnormal LVEF are defined as values less than the lower limit of normal. LVEF outliers and abnormal values will be highlighted in the data listings and LVEF will be summarized.

Further, subjects with outliers that are associated with an AE (related or unrelated to AZD9291) will be listed and summarized as appropriate, along with the duration of the abnormality.

4.2.8.9 WHO performance status

WHO performance status will be listed.

4.3 PK, pharmacodynamics and immunogenicity summary

AZD9291, MEDI4736 and metabolites AZ5104 and AZ7550 PK concentrations will be summarised. AZD9291 and MEDI4736 concentrations will be plotted using box-blots. Individual concentrations and ratios of metabolites will be listed.

ADA concentrations will be summarised and listed.

5. INTERIM ANALYSES

No interim analysis is planned.

6. CHANGES FROM PROTOCOL

OS is defined as the time from the date of randomisation until death due to any cause, whereas in the CSP it is defined as the time from the date of first dose until death due to any cause. As this is a randomised study the principle that efficacy analyses are measured from the date of randomisation is adhered to.

7. REFERENCES

None

8. APPENDIX

N/A.