AstraZeneca

Statistical Analysis Plan

Study Code	D5084C00009
Edition Number	3.0
Date	16 August 2022

A Multi-centre Phase II, Double-Blind, Randomised Study of Savolitinib in Combination with Osimertinib vs Savolitinib in Combination with Placebo in Patients with EGFRm+ and MET Amplified Locally Advanced or Metastatic Non-Small Cell Lung Cancer who have Progressed Following Treatment with Osimertinib

Phastar Study Statistician	PPD
AstraZeneca Study Statistician	PPD
Global Product Statistician	PPD

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AESI	Adverse event of special interest
ANCOVA	Analysis of covariance
ATC	Anatomical therapeutic chemical
BICR	Blinded independent central review
BoR	Best objective response
BSR	Baseline scaled ratio
CI	Confidence interval
COVID-19	Coronavirus 2019 disease
CR	Complete response
CRF	Case report form
CRO	Contract research organisation
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed tomography
CTCAE	Common terminology criteria for adverse events
ctDNA	Circulating tumour DNA
CV	Coefficient of variation
DAE	Discontinuation of investigational product due to adverse events
DBL	Database lock
DCO	Data cut-off
CCI	[redacted]
DoR	Duration of response
d.p.	Decimal place
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG	Eastern cooperative oncology group
eCRF	Electronic case report form
EGFR	Epidermal growth factor receptor
EGFRm+	Epidermal growth factor receptor mutation positive

Abbreviation or special term	Explanation
FAS	Full Analysis Set
FDA	Food and Drug Administration
FISH	Fluorescence in situ hybridisation
HLA	Human leukocyte antigen
HR	Hazard ratio
IDMC	Independent Data Monitoring Committee
CCI	[redacted]
IPD	Important protocol deviation
ITT	Intention-to-treat
IWRS	Interactive Web Response System
LD	Longest diameter
LFTs	Liver function test
Ln	Natural logarithm or logarithm to the base e
LSI	Last subject in
MET	Hepatocyte growth factor receptor
MRI	Magnetic resonance imaging
NA	Not applicable
NE	Not evaluable
NED	No evidence of disease
NSCLC	Non-small cell lung cancer
NTL	Non-target lesion
OAE	Other significant adverse events
ORR	Objective response rate
OS	Overall survival
PD	Progressive disease
PFS	Progression free survival
PK	Pharmacokinetics
PP	Per-protocol
PR	Partial response
PSI	Pulmonary symptom index
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours

Abbreviation or special term	Explanation
CCI	[redacted]
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAF	Safety analysis set
SAS®	A commercially available integrated system of software products, commonly used for reporting and analysis of Clinical Studies
SD	Stable disease
SoA	Schedule of activities
TL	Target lesion
ULN	Upper limit of normal
WHO	World health organisation

1 STUDY DETAILS

This statistical analysis plan (SAP) contains a more detailed description of the analyses in the clinical study protocol (CSP). This SAP is based on version 3.0 of the CSP.

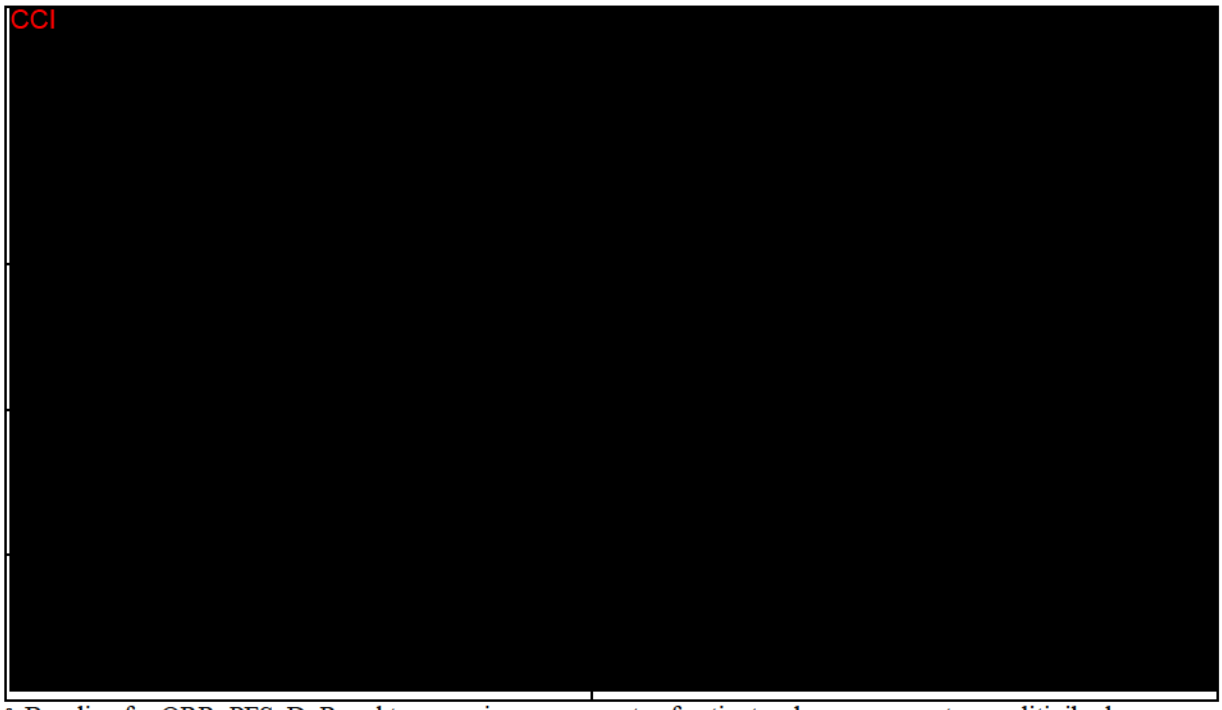
1.1 Study Objectives

The study objectives and the corresponding endpoints/variables are shown in Table 1.

Table 1 Study objectives and corresponding endpoints/variables

Primary objective:	**Primary endpoints/variables:**
To assess the ORR of savolitinib in combination with osimertinib versus savolitinib in combination with placebo in patients with EGFRm+, MET amplified locally advanced or metastatic NSCLC who have progressed on previous osimertinib therapy.	Objective response rate is defined as the proportion of patients with measurable disease who have a CR or PR as determined by the investigator at the local site per RECIST 1.1.
Secondary objectives:	**Secondary endpoints/variables:**
To determine the efficacy of savolitinib in combination with osimertinib versus savolitinib in combination with placebo in patients with EGFRm+, MET amplified locally advanced or metastatic NSCLC who have progressed on previous osimertinib therapy.	• Progression-free survival is defined as time from randomisation until progression per RECIST 1.1 as assessed by the investigator or death due to any cause. • Duration of response is defined as the time from the date of first documented response until date of documented progression per RECIST 1.1 as assessed by the investigator or death in the absence of disease progression. • Tumour size assessment is defined as the percentage change from baseline in TLs at 12 weeks per RECIST 1.1 as assessed by the investigator. • Overall survival is defined as time from randomisation until the date of death due to any cause.

To evaluate the efficacy of savolitinib plus osimertinib in patients who cross-over after progression on savolitinib plus placebo. [a]	• Objective response rate is defined as the proportion of patients with measurable disease who have a CR or PR as determined by the investigator at the local site per RECIST 1.1. • Progression-free survival is defined as time from the first dose in the cross-over period until progression per RECIST 1.1 as assessed by the investigator or death due to any cause. • Duration of response is defined as the time from the date of first documented response during the cross-over period until date of documented progression per RECIST 1.1 as assessed by the investigator or death in the absence of disease progression. • Tumour size assessment is defined as the percentage change from baseline in TLs at 12 weeks per RECIST 1.1 as assessed by the investigator.
To evaluate the PK of savolitinib and osimertinib.	Plasma concentrations of savolitinib, osimertinib, and their metabolites.
To determine the prevalence of ctDNA clearance after savolitinib plus osimertinib or savolitinib plus placebo treatment in this patient population.	Total clearance in EGFR mutations at 6-weeks after therapy initiation (percentage and absolute change from baseline in EGFR mutation allele frequencies).
Safety objective:	**Safety endpoints/variables:**
To evaluate the safety and tolerability of savolitinib plus osimertinib or savolitinib plus placebo.	Safety and tolerability will be evaluated in terms of AEs, SAEs and discontinuation rate due to AEs. Clinical chemistry/haematology including LFTs, ECHOs, ECGs and vital signs including blood pressure and heart rate.
Tertiary/exploratory objectives:	**Tertiary/exploratory endpoints/variables:**



[a] Baseline for ORR, PFS, DoR and tumour size assessments of patients who cross-over to savolitinib plus osimertinib from savolitinib plus placebo will be the progression scan on savolitinib plus placebo acquired within 28 days of the start of treatment in the cross-over period.

1.2 Study Design

This is a multi-centre, Phase II, double-blind, randomised study designed to determine the efficacy of savolitinib administered orally in combination with osimertinib versus savolitinib in combination with placebo in patients with EGFRm+, MET amplified, locally advanced or metastatic NSCLC who have progressed on previous osimertinib treatment.

Prior to CSP Version 3.0, approximately 56 patients were planned to be randomised in a ratio of 1:1 to receive treatment with savolitinib once daily plus osimertinib once daily, or savolitinib once daily plus placebo. At least half of the patients randomised into the study will be second line patients who were treated with osimertinib as first line therapy. Randomisation will be stratified according to the number of prior lines of therapy (i.e., osimertinib monotherapy as first line or ≥ second line [which includes patients who received osimertinib monotherapy before or after chemotherapy]).

Prospective testing of tumour specimens to determine MET amplification by central fluorescence in situ hybridisation (FISH) using tumour specimens collected after progression on prior treatment with osimertinib should be performed at Pre-screening in advance of Screening. Patients determined to be MET amplified in Pre-screening will undergo screening during the 28 days prior to randomisation to confirm eligibility.

All patients confirmed as eligible will begin treatment on Day 1 with savolitinib plus osimertinib or savolitinib plus placebo. Treatment will continue once daily in 28-day cycles until either

objective PD by RECIST 1.1 is assessed, unacceptable toxicity occurs, consent is withdrawn, or another discontinuation criterion is met. Patients randomised to the savolitinib plus placebo arm may cross-over to open-label savolitinib plus osimertinib following investigator-assessed objective PD to ensure that all patients enrolled may have the opportunity to receive the combination of savolitinib plus osimertinib.

A non-comparative interim futility analysis for the savolitinib plus placebo arm was planned to occur after 20 patients overall (10 per arm) had the opportunity of being treated for 2 RECIST post-baseline scans (12 weeks). Recruitment was to continue during this interim futility analysis.

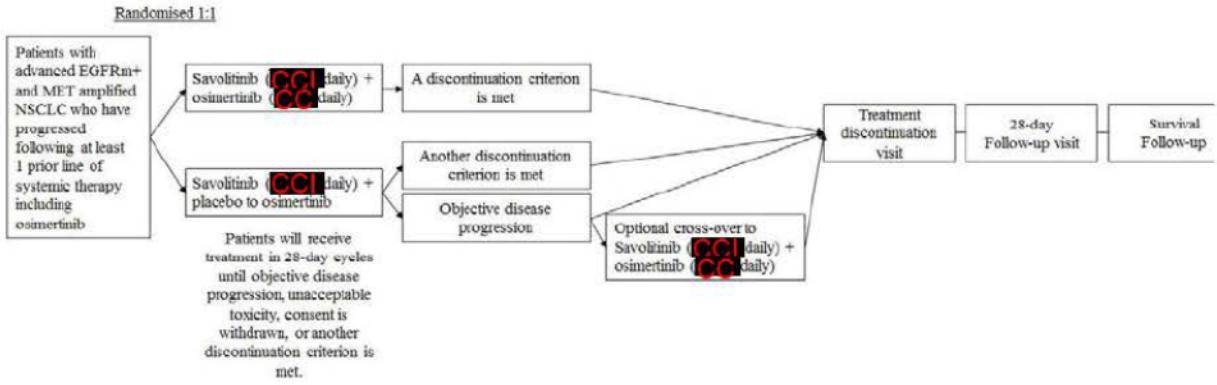
Randomised study design, cross-over study design and end of study design are described in the CSP Section 4.1.

Under CSP V3.0, as a result of AstraZeneca's decision to terminate study recruitment early, the planned interim futility analysis and primary analysis will not be performed, and there will be no requirement for data review by the IDMC. Alternatively, an initial data cut-off (DCO) will occur following the early termination of study recruitment to allow an early review of the data by AstraZeneca. The study will be unblinded after completion of data cleaning activities on the initial DCO data, and patients on savolitinib monotherapy (savolitinib plus placebo) will be given the opportunity to cross-over to the savolitinib plus osimertinib arm after approval by AstraZeneca.

If it is decided that a patient will switch to savolitinib plus osimertinib before final database lock, then they will follow the scheduled procedures described in CSP Table 2. Patients may otherwise choose to remain on savolitinib monotherapy (as long as, in the opinion of the investigator, they are still receiving clinical benefit and have not reported PD). All patients will be followed until the final DCO for the final analysis, which will occur after all randomised patients have had a 9-month follow-up. CCI [redacted]

For an overview of the study design see Figure 1.

Figure 1 Study Design



EGFRm+ = Epidermal growth factor receptor mutation positive; MET = Hepatocyte growth factor receptor; NSCLC = Non-small cell lung cancer.

1.3 Number of patients

Prior to CSP V3.0, the treatment of approximately 56 patients (28 per arm) with EGFRm+ and MET amplified locally advanced or metastatic NSCLC who have progressed following treatment with osimertinib was considered adequate for this study CCI [redacted]

At least half of the patients planned to be randomised into the study will be second-line patients who were treated with osimertinib as first-line therapy.

Under CSP V3.0, the planned number of randomised patients will not be met. This is a result of the decision by AstraZeneca to terminate study recruitment early.

2 ANALYSIS SETS

2.1 Definition of Analysis Sets

Patients will be assigned to the following analysis sets prior to any analyses being performed.

Full Analysis Set (FAS)

The FAS will include all randomised patients with treatment groups assigned in accordance with the randomisation, regardless of the treatment actually received. Patients who were randomised but did not subsequently receive treatment are included in the FAS. The analysis of data using the FAS therefore follows the principles of intention to treat.

The FAS will be the primary population for reporting efficacy data and to summarise baseline characteristics.

Safety analysis set (SAF)

The Safety analysis set will consist of all randomised patients who received any amount of study treatment. Safety data will not be formally analysed but summarised using the safety analysis set according to the treatment received. If a patient receives any amount of an experimental therapy, they will be summarised in the treatment group corresponding to the first experimental treatment they received.

Cross-over analysis set

The Cross-over analysis set will consist of all patients who were randomised to savolitinib in combination with placebo and received at least one dose of savolitinib plus osimertinib in the cross-over period of the study.

It will be used as the population for reporting the efficacy and safety data for patients who crossed over to treatment with savolitinib plus osimertinib following objective progression on savolitinib plus placebo.

PK analysis set

All patients who receive at least one dose of savolitinib or osimertinib as per the protocol, for whom there are at least one reportable PK concentration without any protocol deviation impacting PK, will be included in the PK analysis set. If a patient receives any amount of an experimental therapy, they will be summarised in the treatment group corresponding to the first experimental treatment they received.

Monotherapy analysis set

Under CSP V3.0, the Interim futility analysis set will no longer be required, since the planned interim futility analysis will not be performed. Prior to CSP V3.0, it would have included the first 20 enrolled patients (10 per arm) who had the opportunity to be treated for 2 RECIST post-baseline scans (12 weeks). Therefore, a Monotherapy analysis set is defined including the first 10 randomised patients from the savolitinib plus placebo arm only, to be selected after any randomised patients who were accidentally unblinded are first excluded.

Details of the analysis sets are presented in Table 2.

Table 2 Summary of outcome variables and analysis sets

Outcome variable	***Analysis set***	***Summarised by biomarker groups[a]***
Efficacy data		
ORR, PFS, DoR, OS	*FAS*	Yes
Change in tumour size	*FAS*	Yes
Study Population/Demography Data		

Table 2 Summary of outcome variables and analysis sets

Outcome variable	*Analysis set*	*Summarised by biomarker groups*[a]
Demography characteristics (e.g. age, sex etc.)	*FAS*	Yes
Baseline and disease characteristics	*FAS*	Yes
Important deviations	*FAS*	Yes
Medical/surgical history	*FAS*	Yes
Previous anti-cancer therapy	*FAS*	Yes
Concomitant medications/procedures	*FAS*	Yes
Subsequent anti-cancer therapy	*FAS*	Yes
PK Data		
PK data	*PK*	No
Monotherapy analysis		
ORR	*Monotherapy*	No
Safety data		
Exposure	*Safety*	Yes
Adverse events	*Safety*	No
Laboratory measurements	*Safety*	No
Vital signs	*Safety*	No
ECGs	*Safety*	No
ctDNA data	*Safety*	No
Cross-over analysis		
Baseline characteristics	*Cross-over analysis set*	No
Efficacy endpoints	*Cross-over analysis set*	No
Safety data	*Cross-over analysis set*	No

[a] Biomarker groups are defined as subsets of randomised patients who are FISH10+, CCI , and FISH10+ and/or CCI , as described in Section 3.5.

2.2 Protocol Deviations

The general categories listed in Appendix 9.1 will be considered important protocol deviations (IPDs) and will be either observable and provided by AstraZeneca or programmatically derived from the electronic case report form (eCRF) data.

Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1. During the study, decisions on how to handle errors in treatment dispensing (with regard to continuation/discontinuation of study treatment or, if applicable, analytically) will be made on an individual basis with written instruction from the study team leader and/or statistician.

The important protocol deviations will be listed and summarised by randomised treatment group and discussed in the clinical study report (CSR) as appropriate. None of the deviations will lead to patients being excluded from 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).

In addition to the IPDs, other study deviations captured from the case report form (CRF) module for inclusion/exclusion criteria will be tabulated and listed. Any other deviations from monitoring notes or reports will be reported in an appendix to the CSR.

Depending on the extent of any impact, summaries of data relating to patients diagnosed with COVID-19, and impact of COVID-19 on study conduct may be generated including:

- Disposition (discontinued IP due to COVID-19 and withdrew study due to COVID-19)
- Deviations (overall deviations plus if due to COVID-19 and not due to COVID-19)
- Summary of COVID-19 disruption (visit impact, drug impacted)
- Listing for patients affected by the COVID-19 pandemic.

3 PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST Visit Responses

For all patients, the RECIST tumour response data will be used to determine each patient's visit response according to RECIST version 1.1. It will also be used to determine if and when a patient has progressed in accordance with RECIST and their best objective response to study treatment.

Baseline radiological tumour assessments are to be performed no more than 28 days before the start of randomised treatment and ideally as close as possible to the start of study treatment.

For patients who cross-over to savolitinib plus osimertinib based on confirmed objective disease progression during the initial randomised treatment period (including patients who have already crossed over based on confirmed PD prior to CSP V3.0) the baseline scan for response endpoints during the cross-over period will be the progression scan on savolitinib plus placebo, which should be acquired within 28 days of the start of treatment in the cross-over period. Scheduled RECIST 1.1 tumour assessments are not required for these patients, and the subsequent assessment of disease progression will be based on standard-of-care scan or investigator determined clinical progression.

For patients who cross-over after study unblinding and who have not objectively progressed during the initial randomised treatment period, scheduled RECIST 1.1 tumour assessments are

required using initial randomisation as baseline. Patients who then discontinue study intervention after cross-over for reasons other than investigator-confirmed disease progression will continue scheduled RECIST 1.1 tumour assessments until objective disease progression per RECIST 1.1 as assessed by investigator, unless they withdraw consent to the entire study. These tumour assessments are then performed every 6 weeks following the start of study treatment up to 24 weeks relative to initial randomisation, then every 8 weeks until 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. This schedule is to be followed in order to minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

From the investigator's review of the imaging scans, the RECIST tumour response data will be used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD), using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a patient has had a tumour assessment that 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 Section 3.1.1 for the definitions of CR, PR, SD and PD.

RECIST outcomes (i.e. progression free survival [PFS], objective response rate [ORR] etc.) will be calculated programmatically for the site investigator data (see Section 3.2) from the overall visit responses.

Patients with measurable or non-measurable disease assessed at baseline by computed tomography (CT) / magnetic resonance imaging (MRI) will be entered in this study.

3.1.1 Target lesions (TLs) – site investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD), (except lymph nodes which must have short axis ≥ 15 mm) with CT or MRI and which is suitable for accurate repeated measurements. A patient can have a maximum of five measurable lesions recorded at baseline with a maximum of two lesions per organ (representative of all lesions involved and 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 and prior 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 NTLs at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits.

Note: For patients who do not have measurable disease at entry (i.e. no TLs) but have non-measurable disease, evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions (see Section 3.1.3 for further details). If a patient does not have measurable disease at baseline, then the TL visit response will be not applicable (NA).

Table 3 TL Visit Responses (RECIST 1.1)

Visit Responses	Description
Complete response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to < 10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.
Not evaluable (NE)	Only relevant in certain situations (i.e. if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit; and scaling up could not be performed for lesions with interventions). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response.
Not applicable (NA)	No TLs are recorded at baseline.

Rounding of TL data

For calculation of PD and PR for TLs percentage changes from baseline and previous minimum should be rounded to one d.p. 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 then all TL measurements should be recorded. However, a visit response of PD should still be assigned if any of the following occurred:

- A new lesion is recorded
- An 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 ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared.

Note: the nadir can only be taken from assessments where all the TLs had a LD recorded.

If there is at least one TL measurement missing and a visit response of PD cannot be assigned, the visit response is NE.

If all TL measurements are missing then the TL visit response is NE. Overall visit response will also be NE, unless there is a progression of non-TLs or new lesions, in which case the response will be PD.

Lymph nodes

For lymph nodes, if the size reduces to < 10mm then 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 < 10mm and all other TLs are 0mm then although the sum may be > 0mm the calculation of TL response should be over-written as a CR.

TL visit responses subsequent to CR

Only CR, PD or NE can follow a CR. If a CR has occurred, then the following rules at the subsequent visits must be applied:

- Step 1: If all lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then 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 < 10mm.
- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to NE irrespective of whether, when referencing the sum of TL diameters, the criteria for PD are also met.
- Step 3: If not all lesions meet the CR criteria (i.e. a pathological lymph node selected as TL has short axis > 10mm or the reappearance of previously disappeared lesion) or a new lesion appears, then response will be set to PD
- Step 4: If after steps 1 – 3 a response can still not 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, then this will be indicated as such on the case report form and a value of 5mm will be entered into the database and used in TL calculations. However, a smaller value may be used if the radiologist has not indicated 'too small to measure' on the case report form and has entered a smaller value that can be reliably measured. If a TL response of PD results (at a subsequent visit) 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. 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 tumours:

- 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, treat the lesion diameter (for those lesions with intervention) as missing and if ≤ 1/3 of the TLs have missing measurements then scale up as described in the 'Scaling' section below. If the scaling results in a visit response of PD then the patient would be assigned a TL response of PD.

- Step 3: If, after both steps, PD has not been assigned, then, if appropriate (i.e. if ≤ 1/3 of the TLs have missing measurements), the scaled sum of diameters calculated in step 2 should be used, and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or < 10mm for lymph nodes) and the lesions that have been subject to intervention have a value of 0 (or < 10mm for lymph nodes) recorded. If scaling up is not appropriate due to too few non-missing measurements, then the visit response will be set as NE.

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).

Scaling (applicable only for irradiated lesions/lesion intervention)

If > 1/3 of TL measurements are missing (because of intervention) then the TL response will be NE, unless the sum of diameters of non-missing TL would result in PD (i.e. if using a value of 0 for missing lesions, the sum of diameters has still increased by 20% or more compared to nadir and the sum of TLs has increased by ≥ 5mm from nadir).

If ≤ 1/3 of the TL measurements are missing (because of intervention) then the results will be scaled up (based on the sizes at the nadir visit to give an estimated sum of diameters) and this will be used in calculations; this is equivalent to comparing the visit sum of diameters of the non-missing lesions to the nadir sum of diameters excluding the lesions with missing measurements.

Example of scaling

Lesion 5 is missing at the follow-up visit; the nadir TL sum including lesions 1 – 5 was 74 mm.

The sum of lesions 1 – 4 at the follow-up is 68 mm. The sum of the corresponding lesions at the nadir visit is 62 mm.

Scale up as follows to give an estimated TL sum of 81 mm:

$68 \times 74 / 62 = 81$ mm

CR will not be allowed as a TL response for visits where there is missing data. Only PR, SD or PD (or NE) could be assigned as the TL visit response in these cases. However, for visits with ≤ 1/3 lesion assessments not recorded, the scaled up sum of TLs diameters will be included when defining the nadir value for the assessment of progression.

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 0cm.

Change in method of assessment of TLs

CT, MRI and clinical examination are the only methods of assessment that can be used within a trial, with CT and MRI being the preferred methods and clinical examination only used in special cases. If a change in method of assessment occurs, between CT and MRI this will be considered acceptable and no adjustment within the programming is needed.

If a change in method involves clinical examination (e.g. CT changes to clinical examination or vice versa), any affected lesions should be treated as missing. The TL visit response may still

be evaluable if the number of missing TL measurements at a visit is ≤ 1/3 of the total number of TLs.

3.1.2 Non-target lesions (NTLs) and new lesions – site investigator data.

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

NTL response will be derived based on the investigator's overall assessment of NTLs as follows:

Table 4 NTL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10 mm short axis).
Progressive disease (PD)	Unequivocal progression of existing NTLs. Unequivocal progression may be due to an important progression in one lesion only or in several lesions. In all cases, the progression MUST be clinically significant for the physician to consider changing (or stopping) therapy.
Non-CR/Non-PD	Persistence of one or more NTLs with no evidence of progression.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: For patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met.
Not applicable (NA)	Only relevant if there are no NTLs at baseline.

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 determination of disease progression. 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, but should not overtly affect the derivation.

Symptomatic progression 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 progression' requiring discontinuation of treatment without objective evidence of disease progression at that time should continue to undergo tumour assessments where possible until objective disease progression is observed.

3.1.3 Overall visit response – site investigator data

Table 5 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 5 Overall visit responses

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
CR	CR or NA	No (or NE)	**CR**
CR	Non-CR/Non-PD or NE	No (or NE)	**PR**
PR	Non-PD or NE or NA	No (or NE)	**PR**
SD	Non-PD or NE or NA	No (or NE)	**SD**
PD	Any	Any	**PD**
Any	PD	Any	**PD**
Any	Any	Yes	**PD**
NE	Non-PD or NE or NA	No (or NE)	**NE**
NA	CR	No (or NE)	**CR**
NA	Non-CR/Non-PD	No (or NE)	**SD**
NA	NE	No (or NE)	**NE**
NA	NA	No (or NE)	**NED**

3.1.4 Independent review

A planned blinded independent central review (BICR) of all radiological imaging data will be carried out using RECIST version 1.1. All radiological scans for all patients (including those at unscheduled visits, or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca appointed Contract Research Organisation (CRO) for central analysis. The imaging scans will be reviewed by two independent radiologists using RECIST 1.1 and will be adjudicated, if required (i.e. two reviewers' review the scans and adjudication is performed by a separate reviewer in case of a disagreement). For each patient, the BICR will define the overall visit response (i.e. the response obtained overall at each visit by assessing TLs, NTLs and new lesions) data and no programmatic derivation of visit response is necessary. (For patients with TLs at baseline: CR, PR, SD, PD, NE; for patients with NTLs only: CR, SD, PD, NE; for patients with no disease identified at baseline: PD, no evidence of disease [NED], NE). If a patient has had a tumour assessment that cannot be evaluated then the patient will be assigned a visit response of NE (unless there is evidence of progression in which case the response will be assigned as PD). RECIST assessments/scans contributing towards a particular visit may be performed on different dates and for the central review the date of progression for each reviewer will be provided based on the earliest of the scan dates of the component that triggered the progression.

If adjudication is performed, the reviewer that the adjudicator agreed with will be selected as a single reviewer (note in the case of more than one review period, the latest adjudicator decision will be used). In the absence of adjudication, the records for all visits for a single reviewer will be used. The reviewer selected in the absence of adjudication will be the reviewer who read the baseline scan first. The records from the single selected reviewer will be used to report all BICR RECIST information including dates of progression, visit response, censoring and changes in target lesion dimensions. Endpoints (of ORR, PFS and duration of response [DoR]) will be derived programmatically from this information.

Results of this independent review will not be communicated to investigators and the management of patients will be based solely upon the results of the RECIST 1.1 assessment conducted by the investigator.

A BICR of all patients will be performed for the final database lock, which will cover all of the scans up to the data cut-off (DCO).

Further details of the BICR will be documented in the BICR Charter.

3.2 Efficacy Variables

3.2.1 Objective response rate (ORR)

ORR is defined as the percentage of patients with at least one investigator-assessed visit response of CR or PR and will be based on a subset of all randomised patients with measurable disease at baseline per the site investigator. ORR will also be defined using the BICR data to

define a visit response of CR or PR, with the denominator defined as subset of all randomised patients with measurable disease at baseline per BICR.

Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Patients who discontinue randomised treatment without progression, receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy) and then respond will not be included as responders in the ORR.

3.2.2 Progression free survival (PFS)

PFS is defined as the time from randomisation until the date of 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 analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST assessment. However, if the patient progresses or dies immediately after two or more consecutive missed visits, the patient will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the two missed visits (Note: NE visit is not considered as missed visit).

Given the scheduled visit assessment scheme (i.e. six-weekly for the first 24 weeks then eight-weekly thereafter) the definition of 2 missed visits will change, as follows and summarised in Table 6.

- If the previous RECIST assessment is less than study day 134 (i.e. week 19) then two missing visits will equate to 14 weeks since the previous RECIST assessment, allowing for early and late visits (i.e. 2 x 6 weeks + 1 week for an early assessment + 1 week for a late assessment = 14 weeks).
- If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from six-weekly to eight-weekly this will equate to 16 weeks (i.e. 6 weeks + 8 weeks + 1 week for an early assessment + 1 week for a late assessment = 16 weeks). The time period for the previous RECIST assessment will be from study days 134 to 168 (i.e. week 20 to week 24).
- From week 25 onwards (when the scheduling changes to eight-weekly assessments), two missing visits will equate to 18 weeks (i.e. 2 x 8 weeks + 1 week for an early assessment + 1 week for a late assessment = 18 weeks).

Table 6 Definition of Two Missed Visits

Scheduled assessment	Previous assessment scan	Two missed visits window
Week 7	No evaluable scan or no screening scan	2 x 6 weeks + 2 weeks = 14 weeks (98 days) (relative to randomisation)
Week 7	Screening scan	2 x 6 weeks + 2 weeks = 14 weeks (98 days) (relative to randomisation)
Q6W up to Week 25	> Day 1 – Day 133 (up to Week 19)	2 x 6 weeks + 2 weeks = 14 weeks (98 days)
	> Day 133 – Day 168 (Week 20 – Week 24) (change period from Q6W to Q8W)	6 weeks + 8 weeks + 2 weeks = 16 weeks (112 days)
Q8W thereafter	Day 169 onwards (Week 25 onwards)	2 x 8 weeks + 2 weeks = 18 weeks (126 days)

If the patient has no evaluable visits or does not have baseline data, they will be censored at Day 1 unless they die within two visits of baseline (12 weeks plus 1 week allowing for a late assessment within the visit window).

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

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

- For BICR assessments, the date of progression will be determined based on the **earliest** of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or of the reviewer who read baseline first if there is no adjudication for BICR data.
- For investigational assessments, the date of progression will be determined based on the **earliest** of the dates of the component that triggered the progression.
- For both BICR and investigational assessments, when censoring a patient for PFS the patient will be censored at the **latest** of the dates contributing to a particular overall visit assessment.

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.

3.2.3 Duration of response (DoR)

DoR will be defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression (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 PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR.

If a patient does not progress following a response, then their DoR will use the PFS censoring time.

3.2.4 Overall survival (OS)

Overall survival 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). 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 (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made in the week following the date of data cut-off (DCO) for the analysis, and 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 at the time of the final OS analysis 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, the vital status of the patient can be obtained by site personnel from publicly available resources where it is possible to do so under applicable local laws.

Note: For any OS analysis performed prior to the final OS analysis, in the absence of survival calls being made, it may be necessary to use all relevant CRF fields to determine the last recorded date on which the patient was known to be alive for those patients still on treatment (since the SURVIVE module is only completed for patients off treatment if a survival sweep is not performed). The last date for each individual patient is defined as the latest among the following dates recorded on the CRFs:

- AE start and stop dates
- Admission and discharge dates of hospitalization
- Study treatment date
- End of treatment date
- Laboratory test dates

- Date of vital signs
- Disease assessment dates on RECIST CRF
- Start and stop dates of alternative anticancer treatment
- Date last known alive on survival status CRF
- End of study date.

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.

3.2.5 Best objective response (BoR)

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in Section 3.1.3. It is the best response a patient has had following randomisation, but prior to starting any subsequent cancer therapy and up to and including RECIST progression or the last evaluable assessment in the absence of RECIST progression. Categorisation of BoR will be based on RECIST using the following response categories: CR, PR, SD, NED (applies only to those patients entering the study with no disease at baseline), PD and NE.

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 6 weeks minus 1 week, i.e. at least 35 days (to allow for an early assessment within the assessment window), after randomisation. For CR/PR, the initial overall visit assessment that showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

BoR will be determined programmatically based on RECIST from the overall visit response using all BICR data up until the first progression event. It will also be determined programmatically based on RECIST using all site investigator data up until the first progression event. The denominators for each case will be consistent with those used in the ORR analysis.

For patients whose progression 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 $\leq$ 91 days (i.e. [2 × 6 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 > 91 days after randomisation then BoR will be assigned to the NE category.

A patient will be classified as a responder if the RECIST criteria for a CR or PR are satisfied at any time following randomisation, prior to RECIST progression and prior to starting any subsequent cancer therapy.

3.2.6 Change in TL tumour size

The percentage change from baseline in TL tumour size at 12 weeks is based on RECIST TL measurements taken at baseline and at week 12. Tumour size is the sum of the longest diameters of the TLs. TLs are measurable tumour lesions. Baseline for RECIST is defined to be the last evaluable assessment prior to randomisation. The percentage change in TL tumour size at week 12 will be obtained for each patient taking the difference between the sum of the TLs at week 12 and the sum of the target lesions at baseline divided by the sum of the TLs at baseline times 100 (i.e. [(week 12 - baseline) / baseline] × 100).

Patients who progress before week 12 and enter the cross-over period of the study should have had a tumour assessment performed at the time of progression. The tumour size from their latest progression assessment will be used instead of a week 12 assessment.

No imputation will be used for summaries of the percentage change in tumour size from baseline by visit, but for summaries of the percentage change at week 12, imputation will be used.

TL imputation

For patients who have less than or equal to one-third of TLs missing (because of intervention) at week 12, assessment data from missing lesions will be scaled up proportionally to the sum of the corresponding lesions at baseline to give an estimated sum of diameters as described in Section 3.1.1.

Apply a window around the week 12 visit

Whenever TL tumour size data for the week 12 visit is available then this should be used in the analysis (Note: or the visit at which progression was documented if before week 12). The windowing rules detailed in CSP Section 1.3 will be used to assign the RECIST scan results to the visit; therefore, any RECIST scan performed within ± 1 week of the protocol scheduled visit will be used for that visit.

Best percentage change

The absolute change and percentage change from baseline in the sum of tumour size at each assessment will be calculated. The best change in tumour size (i.e. depth of response) is the largest decrease from baseline or the smallest increase from baseline in the absence of a reduction and will include all assessments prior to the earliest of death in the absence of progression, any evidence of progression, the start of subsequent anti-cancer therapy or the last evaluable RECIST assessment if the patient has not died, progressed or started subsequent anti-cancer therapy.

If best percentage change cannot be calculated due to missing data (including if the patient has no TLs at baseline), 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 patient has no post-baseline assessment and has died
- If a patient has new lesions or progression of NTLs or TLs
- If a patient has withdrawn due to PD and has no evaluable TL data before or at PD.

3.3 Safety Variables

3.3.1 Exposure and dose interruptions

Exposure (i.e. duration of treatment) will be defined for both savolitinib and osimertinib/placebo as follows:

Total (or intended) exposure of study treatment

- Total (or intended) exposure = min(last dose date where dose > 0 mg, date of death, date of DCO) – first dose date + 1

Actual exposure of study treatment

- Actual exposure = intended exposure – total duration of dose interruptions, where intended exposure will be calculated as above, and a dose interruption is defined as any length of time where the patient has not taken any of the planned daily dose.

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

Number of treatment cycles received

Exposure will also be measured by the number of cycles received. A cycle corresponds to a period of 28 days. If a cycle is prolonged due to toxicity, this should still be counted as one cycle. A cycle will be counted if treatment is started even if the full dose is not delivered.

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.

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

3.3.2 Dose intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation. RDI will be defined as follows:

- RDI = 100% × d/D, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the actual last day of dosing. D is the total dose that would be delivered, if there were no modification to dose or schedule.

3.3.3 Adverse events (AEs)

Adverse events and serious adverse events (SAEs) will be collected throughout the study, from date of informed consent until 28 days after the last dose of study treatment (or end of follow-up period). Events will be defined as treatment emergent if they onset or worsen (by investigator report of a change in intensity), during the treatment period as defined in the protocol. 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 (using the CTCAE version referenced in the Clinical Study Protocol).

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 'Discontinuation of Investigational Product due to Adverse Events' (DAEs). Based on the expert's judgement, significant adverse events of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered other significant adverse events (OAEs) and reported as such in the CSR. A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs.

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

AEs of special interest

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

3.3.4 Analysis of Total Calcium per NCI CTCAE criteria

As applicable, values will be converted to standard units and will be graded using CTCAE version referenced in the Clinical Study Protocol. Corrected calcium records will be programmatically derived from Total Calcium and Albumin and appended to the lab dataset for grading.

3.4 Pharmacokinetic Variables

The following PK concentrations will be summarised for savolitinib, osimertinib and their metabolites:

- After single dosing: C_{1h} and C_{3h} post-dose.
- After multiple dosing: $C_{pre\text{-}dose}$, and at C_{1h}, and C_{3h}, post-dose at Cycle 2 and at pre-dose, 1, 3, 4 and 6 hours post-dose at Cycle 3 Day 1 and pre-dose at Cycle 6 Day 1.

Appropriate PK parameters such as AUC_{ss}, C_{ssmax}, T_{ssmax}, CL_{ss}/F at Cycle 3, Day 1 may also be calculated for savolitinib, osimertinib and their metabolites using non-compartmental analysis.

Pharmacokinetic concentration data will be collected and summarised as per the protocol using nominal sampling times. If calculated, appropriate PK parameters may be reported in the CSR.

3.5 Biomarker Variables

3.5.1 Secondary Variables

Circulating-tumour DNA changes will be evaluated as a surrogate marker of clinical efficacy.

The percentage change in EGFR mutation ctDNA allele frequencies at Week 6 and at other timepoints will be obtained for each patient taking the difference between the allele frequency of EGFR sensitising mutations in ctDNA at Week 6 and the ctDNA at baseline divided by the allele frequency of EGFR sensitising mutations in ctDNA at baseline multiplied by one hundred (i.e., [Week 6 - baseline]/baseline × 100). Patients with undetectable EGFR mutations at

baseline will be excluded from this analysis. Samples with undetectable EGFR mutations at Week 6 will be interpreted as 0% allele frequency so that the percentage change is 100%.

The absolute change in ctDNA at Week 6 and over time will be obtained for each patient by the difference between the EGFR mutation ctDNA allele frequencies at timepoint and at baseline.

Patients with detectable EGFR mutations at baseline and available data at 6 weeks will also be categorised into the following groups of ctDNA status: (1) ctDNA clearance: defined by a 100% percent change in EGFR mutations in ctDNA (i.e., undetectable at Week 6) and (2) ctDNA non-clearance: defined by a percent change in EGFR mutations in ctDNA less than 100% (i.e., detectable at Week 6). Patients with undetectable EGFR mutations at baseline will be categorised as "uninformative".

3.5.2 CCI

CCI

[redacted]

4 ANALYSIS METHODS

4.1 General Principles

The below mentioned general principles will be followed throughout the study:

- Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarised by the number of observations, mean, standard

deviation, median, upper and lower quartiles minimum, and maximum. For log-transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarised by frequency counts and percentages for each category.

- Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment group. Overall totals will be calculated for baseline summaries only.
- For continuous data, the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.
- For categorical data, percentages will be rounded to 1 decimal place.
- SAS® version 9.4 or higher will be used for all analyses.
- There are three MET amplification and/ or overexpression biomarker groups of analytical interest as follows:
    - All patients: This population includes all randomised patients.
    - FISH10+ patients: This population includes patients who meet the high biomarker cut-off for FISH as defined in Section 3.5.
    - CCI patients: This population includes patients who meet the high biomarker cut-off for CCI as defined in Section 3.5.

  These groups are not mutually exclusive, i.e. a patient may be a member of all three groups.
- Unless otherwise specified, analyses will be presented by treatment group and biomarker group (FISH10+, CCI, FISH10+ and/or CCI, and all patients).
- Additional combinations of line of therapy and biomarker group may be further investigated.

In general, for efficacy endpoints the last observed measurement prior to randomisation will be considered the baseline measurement. However, if an evaluable assessment is only available after randomisation but before the first dose of randomised treatment then this assessment will be used as baseline. 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 percentage change from baseline will be calculated as (post-baseline value - baseline value) / baseline value × 100. For any variable subjected to log transformation, the change from baseline calculated and summarised on the log scale will be back-transformed and presented as a 'baseline scaled ratio' (BSR). Percentage change will then be calculated as (BSR - 1) × 100.

Efficacy data will be summarised and analysed based upon the FAS. Safety and treatment exposure data will be summarised based upon the safety analysis set. Study population and demography data will be summarised based upon the FAS. PK data will be summarised based upon the PK analysis set. Summaries will be repeated for the cross-over analysis set as applicable.

The stratification in the statistical analyses will be based on the values entered in the interactive web response system (IWRS) at randomisation, even if it is subsequently discovered that these values were incorrect.

4.2 Analysis Methods

4.2.1 Multiplicity

There will be no adjustment for multiplicity as CCI [redacted], and no inferential analyses are planned.

4.2.2 Primary Efficacy Endpoint

4.2.2.1 Objective response rate (ORR)

The ORR will be based on the site investigator RECIST data, and using all scans regardless of whether they were scheduled or not. The ORR will be summarised by n (%) in each treatment arm.

Summaries by treatment group will be produced that present the number and percentage of patients with a tumour response (CR/PR) based upon the number of patients with measurable disease at baseline per the site investigator. The response rates by treatment group will be presented with a 2-sided 95% CI using the Clopper-Pearson method.

This analysis of ORR will be repeated for second line patients (i.e. patients with one prior line of therapy before the study treatment).

These summaries will also be produced for ORR per BICR (based upon the number of patients with measurable disease at baseline per BICR). If there is an important discrepancy between the primary analysis using the site investigator data and this sensitivity analysis using BICR data then the proportion of patients with site but no central confirmation of progression will be

summarised; such patients have the potential to induce bias in the central review due to informative censoring. Disagreements between investigator and central reviews of RECIST progression will be presented for each treatment group.

Additionally, a summary of ORR in the first 10 monotherapy arm patients who have had the opportunity to be treated for two RECIST post-baseline scans (12 weeks, the Monotherapy Analysis Set) will be performed. This summary will have no impact on the CoC study and will only be used for exploratory purposes. The first 10 randomised patients will be selected from the savolitinib plus placebo arm, after first excluding any randomised patients who were accidentally unblinded.

For each treatment group best objective response (BoR) will be summarised by n (%) for each category (CR, PR, SD, PD and NE) based upon the FAS. This analysis will be repeated for second line patients (i.e. patients with one prior line of therapy before the study treatment).

4.2.2.1.1 Subgroup analyses

Subgroup analyses will be conducted, summarising ORR in the treatment groups in the following subgroups of the FAS:

- Prior lines of therapy (osimertinib monotherapy as first line versus ≥ second line)
- FISH allocation status (amplified versus polysomy, or missing).

The subgroups for prior lines of therapy will be determined by the values entered into the IWRS at randomisation, even if it is subsequently discovered that these values were incorrect.

All patients randomised in the study will be MET amplified (FISH+). The FISH allocation status of amplification is considered a ratio of 'Mean MET/CEP7 RATIO' ≥ 2, and polysomy where this ratio < 2. If the ratio variable is missing, patients will be presented in a missing category.

The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors. These subgroup analyses will be repeated for the FISH10+ biomarker group and for the CCI biomarker group.

4.2.3 Secondary Efficacy Endpoints

4.2.3.1 Progression free survival (PFS)

The PFS will be based on the site investigator RECIST data, and using all scans regardless of whether they were scheduled or not. The PFS will be summarised in both treatment arms.

If there are at least 20 events, the effect of treatment will be estimated by the hazard ratio (HR) (an HR less than 1 will favour savolitinib plus osimertinib) together with its corresponding 95% CI for the FAS. The HR and its CI will be estimated from a stratified Cox Proportional Hazards model (with ties = Efron and baseline stratification), adjusting for the number of prior lines of therapy (osimertinib monotherapy as first line or ≥ second line [which includes patients who received osimertinib monotherapy before or after chemotherapy]), and the CI calculated using a profile likelihood approach.

Kaplan-Meier plots of PFS will be presented by treatment arm. Summaries of the number and percentage of patients experiencing a PFS event, and the type of event (RECIST 1.1 progression or death) will be provided along with summaries of PFS (n, events, medians, quartiles, proportion progression free at 3, 6, 9, and 12 months and corresponding 95% CIs) by treatment group. The number and percentage of patients censored in the analysis as well as by reason for censoring will be provided for each treatment group.

This analysis of PFS will be repeated for second line patients (i.e. patients with one prior line of therapy before the study treatment).

These analyses will also be produced for PFS per BICR. If there is an important discrepancy between the primary analysis using the site investigator data and this sensitivity analysis using BICR data then the proportion of patients with site but no central confirmation of progression will be summarised; such patients have the potential to induce bias in the central review due to informative censoring. Disagreements between investigator and central reviews of RECIST progression will be presented for each treatment group.

The effect of treatment will be estimated by the HR for second line patients and per BICR only if there are at least 20 events within those analysis sets.

The treatment status at progression of patients at the time of analysis will be summarised based upon the FAS. This will include the number (%) of patients who were on treatment at the time of progression, the number (%) of patients who discontinued study treatment prior to progression, the number (%) of patients who have not progressed and were on treatment or discontinued treatment. This will also provide distribution of number of days prior to progression for the patients who have discontinued treatment.

4.2.3.2 Duration of response (DoR)

Descriptive data (n, number of responses that have progressed, median, quartile, minimum and maximum DoR) will be provided by treatment group for the duration of response in responding patients, including the associated Kaplan-Meier curves (without any formal comparison or p-value attached). These analyses and plots will be repeated for second line patients.

Descriptive summaries for the DoR will be based upon the FAS.

4.2.3.3 Overall survival (OS)

Overall survival will be analysed as described for the investigator-assessed PFS endpoint in Section 4.2.3.1, with the exception that the analyses will not be repeated in second line patients.

Summaries of OS by treatment arm (n, events, medians, quartiles, proportion alive at 3, 6, 9 and 12 months and corresponding 95% CIs) based upon the FAS and a Kaplan-Meier plot will be provided.

In addition, duration of follow-up will be summarised using medians in all patients based upon the FAS, calculated as the time from randomisation to the date of death (i.e. overall survival) or

to the date of censoring (date last known to be alive) for censored patients, regardless of treatment arm.

4.2.3.4 Change in TL tumour size

The absolute values, absolute change in TL tumour size from baseline and percentage change in TL tumour size from baseline will be summarised using descriptive statistics and presented at each timepoint and by randomised treatment group, based upon the FAS.

The best change in TL tumour size from baseline (where best change in TL size is the maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction) will also be summarised and presented by randomised treatment group. The number and percentage of patients in each treatment group whose best percentage change data is imputed (detailed in Section 3.2.6) will also be presented.

The percentage change in week 12 TL tumour size from baseline will also be summarised and presented by randomised treatment group. The number and percentage of patients in each treatment group whose week 12 data is imputed will also be presented.

Tumour size will also be presented graphically using waterfall plots for each treatment group, to present each patient's week 12 percentage change in TL tumour size as a separate bar, with the bars ordered from the largest increase to the largest decrease. Reference lines at the -30% and +20% change in TL tumour size level will be added to the plots, which correspond with the definition of 'partial response' and 'progressive disease' respectively. All progressions will be marked with a '●' or designated with patterns or colours for ORR categories. The scale in these plots will be fixed to be from -100 to 100 to avoid presenting extreme values. Values that are capped as a result of this restriction to the scale will be marked with '#'. Values will be ordered in descending order, with the imputations due to death appearing first, followed by a gap, followed by all other patients. Imputed values will be clearly marked with '*' and patients with imputation where there is a death or evidence of progression will have different shading to each other and the other patients to make it clear that these are different. These plots will be repeated for second line patients and for ≥ third line patients.

Additionally, 'spider' plots will be produced for each treatment group. These will depict each patient's actual change and percentage change from baseline in TL tumour size as a line over time, and progression due to non-target and/or new lesions will be indicated. These plots will be repeated for second line patients and for ≥ third line patients.

4.2.4 Data cut-offs

The data cut-off for futility analysis was planned to take place after 20 patients overall (10 per arm) have had the opportunity to be treated for 2 RECIST post-baseline scans (12 weeks). In order to ensure that 10 patients in the savolitinib plus placebo group have been randomised for inclusion into the futility analysis set, the IWRS was to be used to determine when 22 patients have been randomised into the study.

The data cut-off for primary analysis was planned to take place 6 months after last subject in (LSI). The data cut-off for final analysis was planned to take place at the earlier of 18 months after LSI or when 70% of patients have progressed or died.

Under CSP v3.0, an initial DCO will occur to allow an early review of data by AstraZeneca. A final DCO will occur 9 months after the last patient has been randomised, after which the final analysis will be performed. CCI [redacted] These changes are a result of the decision by AstraZeneca to terminate study recruitment early.

4.2.5 Safety

4.2.5.1 General considerations for safety assessments

Time windows will be defined for any presentations that summarise values by visit. The following conventions will apply:

- The time windows will be exhaustive so that data recorded at any time point has the potential to be summarised. Inclusion within the time window will be based on the actual date and not the intended date of the visit.
- All unscheduled visit data have the potential to be included in the summaries.
- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls halfway between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits, then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for vital signs data are:
  - Day 8, visit window 2 – 11
  - Day 15, visit window 12 – 18
  - Day 22, visit window 19 – 25
  - Day 29, visit window 26 – 43
  - Day 57, visit window 44 – 71
  - Day 85, visit window 72 – 99
  - Day 113, visit window 100 – 127.
- 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.
- For visit based summaries, if there is more than one value per patient within a time window then the closest value to the scheduled visit date will be summarised, or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the patient that contributed to the summary table, wherever feasible. Note: in summaries of extreme values all post-baseline values

collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.

- Baseline for safety assessments will generally be the last value obtained prior to the first dose of study medication). Alternatively, if two visits are equally eligible to assess patient status at baseline (e.g., screening and baseline assessments both on the same date prior to first dose with no washout or other intervention in the screening period), the average can be taken as a baseline value. For non-numeric laboratory tests (i.e. some of the urinalysis parameters) where taking an average is not possible then the best value would be taken as baseline as this is the most conservative. In the scenario where there are two assessments on day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. Where safety data are summarised over time, study day will be calculated in relation to date of first treatment.

- Per Section 9.4.3 of the CSP, adverse events occurring in the savolitinib in combination with placebo group will be split into two presentations based on whether they occurred before or after cross-over.

Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e. below the lower limit of quantification) or > x (i.e. above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

Furthermore:

- 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.

- For missing AE start dates, the following will be applied:

    a. Missing day: impute the 1st of the month unless month is the same as month of the first dose of study drug then impute first dose date

    b. Missing day and month: impute 1st January unless year is the same as first dose date then impute first dose date

    c. Completely missing: impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date

When imputing a start date ensure that the new imputed date is sensible e.g. is prior to the end date of the AE.

- For missing AE end dates, the following will be applied:

a. Missing day: impute the last day of the month unless month is the same as month of the last dose of study drug then impute last dose date.

b. Missing day and month: impute 31st December unless year is the same as last dose date then impute last dose date.

Flags will be retained in the database indicating where any programmatic imputation has been applied, and in such cases, any durations would not be calculated.

- 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:

c. For Missing day only: using the 1st of the month

d. For Missing day and Month: using the 1st of January

With the exception of exposure data, safety data will not be summarised by biomarker group.

4.2.5.2 Adverse events (AEs)

Adverse events will be coded using the most recent version of MedDRA that have been released for execution at AstraZeneca. Adverse events will be graded according to the National Cancer Institute CTCAE (Version 5.0).

Summary information (the number and percent of patients per treatment group) by system organ class and preferred term will be tabulated for savolitinib plus placebo and savolitinib plus osimertinib for:

- All AEs
- All AEs possibly related to study medication (as determined by the reporting investigator)
- AEs with maximum reported CTCAE grade
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, possibly related to study medication (as determined by the reporting investigator)
- Most common AEs (occurring in at least 10% of patients)
- Most common non-serious AEs (occurring in at least 5% of patients)
- AEs with outcome of death
- AEs with outcome of death possibly related to study medication (as determined by the reporting investigator)
- All SAEs
- All SAEs possibly related to study medication (as determined by the reporting investigator)

- AEs leading to discontinuation of study medication
- AEs leading to discontinuation of study medication, possibly related to study medication (as determined by the reporting investigator)
- AEs leading to dose reduction of study medication
- AEs leading to dose reduction of study medication, possibly related to study medication (as determined by the reporting investigator)
- AEs leading to dose interruption of study medication
- AEs leading to dose interruption of study medication, possibly related to study medication (as determined by the reporting investigator).

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

Any AE occurring before the first dose of study intervention will be included in the data listings but will not be included in the summary tables of AEs. Adverse event summary tables will include only treatment-emergent AEs. The following events are considered treatment emergent:

- Adverse events with an onset date on or after the first dose of randomised study intervention and within the safety follow-up period (28 days [± 7 days] after last dose of all study intervention). Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of all study intervention) will be flagged in the data listings.
- Worsening of pre-existing events on or after the first dose of study intervention and within the safety follow-up period.

Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of all study intervention) will be flagged in the data listings, but not included in the summaries.

4.2.5.3 Adverse events of special interest (AESI)

Preferred terms used to identify AESI will be listed before DBL and documented in the Trial Master File. Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI grouping. Groupings will be based on preferred terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

Summaries of the above-mentioned grouped AE categories (the number and percent of patients per treatment group) by system organ class and preferred term will be tabulated for:

- All AESIs
- All AESIs possibly related to study medication (savolitinib only, osimertinib/placebo only, or both savolitinib and osimertinib/placebo only)

- All AESIs by maximum reported CTCAE grade.

4.2.5.4 Vital signs

Summaries of vital signs data will include all data obtained up until 28 days after the last dose of study treatment. Absolute values and change from baseline for diastolic and systolic BP, heart rate, body temperature, height and weight will be summarised at each visit by treatment. The mean, standard deviation, median, minimum value, maximum value, and lower and upper quartiles at each time point will be presented.

4.2.5.5 Laboratory summaries

For clinical chemistry and haematology, numerical summaries by treatment group of absolute values and change from baseline will be presented which include the mean, standard deviation, median, minimum value, maximum value, and lower and upper quartiles at each visit in which laboratory values are taken. Shift tables will be produced for CTCAE grade changes, and for baseline to minimum/ maximum values on treatment, in haematology and clinical chemistry parameters, and for baseline to maximum value on treatment for urinalysis parameters. Plots for laboratory parameters of baseline versus minimum/ maximum post-baseline values will be presented.

The liver function tests will be summarised for each patient with potential Hy's Law, who has ALT or AST > 3x ULN and total bilirubin > 2x ULN at any time during the study by time point, and with elevation in transaminases preceding or coinciding with (on the same day as) the elevation in BILI.

4.2.5.6 Electrocardiograms (ECGs)

QT interval corrected for heart rate using Fridericia's correction (QTcF) evaluations will be done based on triplicate 12-lead electrocardiograms (ECGs) as indicated in the schedule of activities (SoA) of the CSP. The following ECG variables will be collected: PR interval, QRS duration, QT interval, RR interval, and overall ECG evaluation.

The overall evaluation of an ECG will be either "normal" or "abnormal" with abnormalities categorised as either "clinically significant" or "not clinically significant".

Summaries of ECG data will include all data obtained up to and including 28 days after the date of the last dose of study treatment. The ECG assessment at baseline versus end of treatment will be presented by treatment group. Absolute values and change from baseline will be summarised at each visit by treatment group.

For QTc, there are 3 recordings at each visit and therefore for the analysis the average of the non-missing values will be taken as the result at that visit. The QTcF will be calculated in the eCRF as follows (where QT and RR are in seconds):

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

The number and percentage of patients with a maximum post-baseline QTcF result exceeding ICH boundaries will be presented by treatment group, both as absolute values and as increases from baseline.

4.2.5.7 Multi-gated Acquisition (MUGA)/Echocardiogram (ECHO)

An ECHO or MUGA scan to assess left ventricle ejection fraction will be conducted during the main study screening period, Q12W (± 2 weeks) relative to randomisation, and at the end of treatment as indicated in the SoAs of the CSP. Absolute values at baseline and end of treatment visit and change from baseline to end of treatment visit for LVEF results will be summarised.

4.2.5.8 Exposure

Duration of exposure will be summarised for the SAF. The following summaries will be produced:

- Total exposure and actual exposure of savolitinib and osimertinib
- Exposure plot over time showing a line for each treatment arm
- Summary statistics (mean, standard deviation, median, quartiles, minimum, maximum) of RDI of savolitinib and osimertinib
- Summary of dose interruptions and reductions, with reasons, of savolitinib, osimertinib and placebo separately for the first 6 weeks and at any time (see Section 3.3.1 for details of analysis of missed or forgotten doses)
- Cumulative exposure over time, showing the number and percentage of patients treated for at least 0, 6 and 12 months
- Total number of treatment cycles received.

The calculation of exposure is given in Section 3.3.1.

Analysis methods for relative dose intensity

Summary statistics (mean, standard deviation, median, quartiles, minimum, maximum) will be presented for dose intensity.

4.2.6 Demographics and baseline characteristics

The following will be summarised for all patients in the FAS (unless otherwise specified) by treatment group and biomarker group:

- Patient disposition (including screening failures and reason for screening failure)
- Important protocol deviations
- Inclusion in analysis sets
- Demographics (age, age group [< 40, ≥ 40 - < 65, ≥ 65 - < 75 and ≥ 75 years], sex, race, ethnicity, and country)

- Patient characteristics at baseline (height, weight)
- Previous disease-related treatment modalities
- Number of prior regimens of previous chemotherapy prior to this study
- Disease characteristics at baseline (ECOG performance status, primary tumour location, histology type, tumour grade, overall disease classification, FISH allocation status, prior lines of therapy, and immediate prior line of therapy)
- Extent of disease at baseline
- Disease related medical history (past and current)
- Relevant surgical history
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use, categorised (never, current, former).

The medications will be coded following WHO Drug dictionary.

4.2.7 Concomitant and other treatments

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

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

Medications received prior to, concomitantly, or post-treatment will be coded using the AstraZeneca 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 mediation or therapy summaries, incomplete medication or radiotherapy start and stop dates will be imputed as detailed in Section 4.2.5.1.

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

- Prior medications are those taken prior to study treatment 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.

In addition, all post-treatment anti-cancer medications and surgical procedures will be summarised for the full analysis set.

The following summaries will be produced:

- Summary of prior medications
- Summary of concomitant medications
- Summary of post study treatment cancer therapies.

All concomitant and other treatment data will be listed.

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

4.2.8 Pharmacokinetic data

Plasma concentrations of savolitinib, osimertinib and their metabolites (M2 and M3 for savolitinib; AZ5104 for osimertinib) will be summarised by nominal sample time using standard summary statistics for PK concentrations (geometric mean, geometric coefficient of variation, geometric mean ± standard deviation, arithmetic mean, standard deviation, median, minimum, maximum and n) for the PK parameters listed in Section 3.4.

The time dependency of the PK on multiple dosing will be assessed by the calculation of the ratios of:

- C_{3h} Cycle 2 Day 1 / C_{3h} Cycle 1 Day 1
- $C_{pre\text{-}dose}$ Cycle 3 Day 1 / $C_{pre\text{-}dose}$ Cycle 2 Day 1
- $C_{pre\text{-}dose}$ Cycle 6 Day 1 / $C_{pre\text{-}dose}$ Cycle 2 Day 1

All plasma concentrations will be listed.

At Cycle 3 Day 1, PK parameters of C_{ssmax}, AUC_{ss}, T_{ssmax}, CL_{ss}/F for osimertinib, savolitinib and if appropriate for its metabolites, will be determined using non-compartment analysis by Covance PK alliance on behalf of AstraZeneca. Actual sample times will be used for the analysis.

Pharmacokinetic data will not be summarised by biomarker group.

4.2.9 Biomarkers

Biomarker endpoints will be analysed for the SAF.

The absolute change and the percentage change in EGFR mutation ctDNA allele frequencies at Week 6 will be summarised with descriptive statistics (n, mean, SD, median, min, max). Individual-level changes from baseline to all post-baseline timepoints will be summarised for

all patients. Waterfall plots of percentage change ctDNA allele frequencies at Week 6 will be included and spider plots of absolute and percentage changes in ctDNA allele frequencies over time may also be included.

The number and percentage of patients with detectable EGFR mutations at baseline and available data at 6-weeks, split by ctDNA clearance and non-clearance will be presented.

CCI [redacted]

Other summaries and analyses for exploratory biomarkers may be documented in a separate analysis plan and will be reported in a separate report outside the CSR.

4.2.10 Analyses for cross-over patients

The assessment schedule for patients who cross-over will be determined by the point of cross-over, as described in Section 3.1.

Selected analyses will be repeated for the cross-over analysis set, including baseline characteristics, efficacy endpoints and safety analyses where sufficient data allow.

Any AE summary tables for cross-over patients will include all AEs that occurred after the start of cross-over treatment up until the end of the 28-day follow-up period. The 28-day follow-up period will be defined as the 28 days following discontinuation of treatment.

Cross-over analyses will not be presented by biomarker group.

5 INTERIM ANALYSES

Prior to CSP V3.0, a non-comparative interim futility analysis for the savolitinib plus placebo arm was planned to occur after 20 patients overall (10 per arm) had the opportunity to be treated for 2 RECIST post-baseline scans (12 weeks).

Under CSP V3.0, as a result of AstraZeneca's decision to terminate study recruitment early, the planned interim futility analysis will not be performed and there will be no requirement for data review by an external independent data monitoring committee (IDMC). Alternatively, following the termination of study recruitment, an initial DCO will occur to allow an early review of the data by AstraZeneca. The study will be unblinded after completion of data cleaning activities on the initial DCO data, and patients on savolitinib monotherapy (savolitinib plus placebo) will be given the opportunity to cross-over to the savolitinib plus osimertinib arm. If the investigator believes patients are gaining clinical benefit and patients have not reported PD, patients may continue to receive savolitinib monotherapy. All patients will be followed until the final analysis.

6 CHANGES OF ANALYSIS FROM PROTOCOL

The formula given in Section 3.5 for the percentage change in EGFR mutation ctDNA allele frequencies at week 6 and other time points will be [Week 6 – baseline]/baseline × 100, contrary to Section 9.4.5 of the CSP V3 which states it will be [Week 6 baseline]/baseline × 100.

Previous versions of the SAP included comparative inferential analyses; however, since the SAP v2.0 update, it has been decided to remove all such analyses due to the lower than planned sample size following the CSP v3.0 amendment, with the exception of PFS and OS if there are at least 20 events. Where a comparative analysis was planned, summaries by treatment group will still be presented.

7 COUNTRY SPECIFIC ANALYSES

Not applicable.

8 REFERENCES

Alosh, M et al. Statistical considerations on subgroup analysis in clinical trials. Stat Bio Res. 2015; 7 (4):286-303

Collett D. Modelling survival data in medical research. Chapman & Hall, 1994.

[Not referenced]Dodd LE, Korn EL, Freidlin B, Gray R, Bhattacharyra S. An audit strategy for progression free survival. Biometrics, 2011; 67(3): 1092-9

Eisenhauer EA, Therasse P, Bogaerts J et el. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). E J Cancer. 2009; 45:228-247

Ellis S, Carroll K J, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials. 2008; 29:456-465

FDA endpoint guidance. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics.

Grizzle JE. A note on stratifying versus complete random assignment in clinical trials. Control Clin Trials. 1982; 3:365-368

Klein JP, Logan B, Harhoff M et al. Analyzing survival curves at a fixed point in time. Stat Med. 2007; 26:4505-4519

Lin DY, Dai L, Cheng G, Sailer MO. On confidence intervals for the hazard ratio in randomized clinical trials. Biometrics 2016; 72(4):1098-1102

9 APPENDICES

9.1 Important Protocol Deviations

Protocol deviations are collected, reviewed and reconciled throughout the study. Important protocol deviations (IPDs) are identified from the complete set of protocol deviations. IPDs are those which may significantly impact the reliability of the study data or that may significantly affect a patient's rights, safety, or wellbeing.

A set of pre-determined IPDs are listed in the protocol deviations plan. The protocol deviations plan also indicates which IPDs are identified by programmatic checks.

The IPDs are grouped into the following important protocol deviation (IPD) categories, where full details of the individual IPDs within each IPD category are provided in the protocol deviations plan:

- Inclusion and exclusion criteria deviations
- Discontinuation criteria for study product met but participant not withdrawn from study treatment
- Discontinuation criteria for overall study withdrawal met but participant not withdrawn from study
- IP deviation
- Excluded medications taken
- Deviations to study procedure
- Other important deviations.

The following general categories will be considered important protocol deviations (IPDs) and will be programmatically derived from Veeva Clinical Vault (VCV) data. These will be listed and summarised by randomised treatment group and discussed in the CSR as appropriate. Refer to the CSP for full details of the inclusion/exclusion criteria.

- Patients who deviate from key entry criteria per the CSP (Deviation 1)
    - Inclusion criteria 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 18
    - Exclusion criteria 3, 7, 11, 12, 13, 16, 17, 21, 23
- Discontinuation criteria for study product met but participant not withdrawn from study treatment (Deviation 2)
- Discontinuation criteria for overall study withdrawal met but participant not withdrawn from study (Deviation 3)
- Investigational product deviation (Deviation 4)
- Excluded medications taken (Deviation 5)
- Baseline RECIST > 42 days before date of randomisation (Deviation 6)
- No baseline RECIST assessment on or before date of randomisation (Deviation 7)
- Other important deviations (Deviation 8)

- Patients randomised but who did not receive study treatment (Deviation 9).

10 DOCUMENT REVISION HISTORY

Version	Description of Update
1.0	First approved SAP version.
1.3	The following changes have been made to version 1: (i) Signature pages have been removed, since signatures will be collected electronically. (ii) List of abbreviations has been updated, and abbreviations have been spelled out in full on first mention throughout. (iii) Section 2.1 Definition of Analysis Sets – More detail has been added to the definition of the PK analysis set. (iv) Section 3.1.1 Target lesions (TLs) – Site investigator data has been updated to add text related to RECIST in non-registration studies about scaling up missing sum of TL when there are ≤ 1/3 missing. (v) Section 3.5 Biomarker variables – The formula for percentage change in EGFR ctDNA allele frequencies at week 6 and other time points was corrected, and will be [Week 6 – baseline]/baseline×100. (vi) Section 4.2 Analysis Methods – In Table 6, clarification has been added that the profile likelihood confidence interval will be calculated for PFS. (vii) Section 4.2.2.1 Objective response rate (ORR) – More detail has been added about the use of the Cochran-Mantel-Haenszel test, and on summaries of the ORR presented by treatment group. (viii) Section 4.2.2.1.2 Subgroup analyses – has been updated to include subgroup analysis by FISH allocation status. (ix) Section 4.2.3.4 Change in TL tumour size – In the description of the ANCOVA models for absolute and percentage change in TL tumour size, the term "the absolute (percentage) change in week 12 value" has been removed, since this would not be a covariate in the models. (x) Section 4.2.5.1 General considerations for safety assessments – has been updated with the following: a. The visit windows for vital signs data: Day 15 and Day 22 had incorrect upper and lower limits. b. Details regarding vital signs schedule in brackets were removed.

Version	Description of Update
	c. A bullet point describing summaries at patient levels was removed, since this was partially a repeat of a previous point, and did not make sense as we would not present a summary at patient level. d. A point was added to clarify that adverse events occurring in the savolitinib + placebo group will be split into two presentations, depending on whether they occurred before or after cross-over. (xi) Section 4.2.5.2 Adverse events (AEs) – has been updated with the following: a. Summaries of AEs with CTCAE grade 3 or 4 have been changed to grade 3 or higher. b. Most common AEs are defined as occurring in at least 10% of patients, changed from at least 5%. (xii) Section 4.2.5.3 Adverse events of special interest (AESI) – This section has been added. (xiii) Section 4.2.5.4 Vital signs – This section has been added. (xiv) Section 4.2.5.5 Laboratory summaries – This section has been added. (xv) Section 4.2.5.6 Electrocardiograms (ECGs) – This section has been added. (xvi) Section 4.2.5.7 Multi-gated Acquisition (MUGA)/Echocardiogram (ECHO) – This section has been added. (xvii) Section 4.2.5.8 Exposure – This has been updated as follows: a. The bullet "Summary of duration of exposure of savolitinib and osimertinib" has been removed since this is a duplicate, and the first bullet has been changed to "Total exposure and actual exposure of savolitinib and osimertinib" for clarification. b. Summary of dose delays have been removed, as these are not applicable to this study. c. Cumulative exposure over time showing the number and percentage of patients treated for at least 0, 6 and 12 months was added. (xviii) Section 4.2.8 Pharmacokinetic data – Median has been added to the list of summary statistics for PK parameters. (xix) Section 4.2.10 Analyses for cross-over patients – This has been updated as follows: a. To remove ambiguity about analysis of OS, which may only be repeated if there are sufficient events.

Version	Description of Update
	b. Added summaries of exposure and ECHO data. (xx) Clarified that ctDNA and PK parameter analyses may only be repeated if there are sufficient data. Section 6 Changes of Analysis from Protocol – This has been updated to reflect changes in Sections 3.5 and 4.2.3.4 that are inconsistent with the protocol. (xxi) Section 9.1 Important Protocol Deviations – This has been updated to provide a summary and refer to the protocol deviations plan.
2.0	The following changes have been made following early termination of study recruitment and under CSP v3.0: (i) Section 1.2 Study Design – Updated to reflect early termination of study recruitment and its impact on crossover procedures and removal of futility analysis. (ii) Section 1.3 Number of patients – Updated to clarify the planned sample size will not be met. (iii) Section 2.1 Definition of Analysis Sets: a. Added that patients will be assigned to analysis sets prior to analyses being performed b. Changed Safety analysis set abbreviation from SAS to SAF c. Added Monotherapy analysis set, and explained the interim futility analysis set is no longer required. d. Table 2: Additional column flagging variables to be summarised by biomarker groups, and removal of interim futility analysis. Simplified cross-over analysis variables. (iv) Section 3.1 Derivation of RECIST Visit Responses – Amended scheduled assessments details for patients who cross-over. (v) Section 3.1.4 Independent review – Amended to state BICR will no longer be performed. (vi) Section 3.2.1 Objective response rate – Removed ORR defined using BICR data. (vii) Section 3.2.2 Progression free survival: a. Amended definition of two missed visits and added Table 6. b. Removed RECIST rules applying to BICR assessments.

Version	Description of Update
	(viii) Section 3.2.5 Best objective response – Removed derivation using BICR data. (ix) Section 3.2.6 Change in TL tumour size – Added details of imputation for percentage change at week 12. (x) Section 3.4 Pharmacokinetic variables – Removed Cycle 11 Day 1 from summarised data. (xi) Section 3.5 Biomarker variables – Added sub-headings for secondary and exploratory variables, and definitions of biomarker groups (FISH10+ and CCI). (xii) Section 4.1 General Principles: a. Added definitions of biomarker groups and presentation of analyses by both biomarker and treatment group. b. Removed paragraph about stratification, since this applied to comparative analyses which will no longer be done. (xiii) Section 4.2 Analysis Methods – Removed Table 7 since comparative/ formal statistical analysis will no longer be done. (xiv) Section 4.2.1 Multiplicity – Amended to state there will be no adjustment for multiplicity. (xv) Section 4.2.2.1 Objective response rate: a. Removed sentence about summaries per BICR. b. Removed all comparative/inferential analyses. c. Added summary of ORR in monotherapy analysis set. d. Added repeat of ORR analysis and BoR analysis for patients in second line patients. (xvi) Section 4.2.2.1.1 ORR sensitivity analyses – Section removed. (xvii) Section 4.2.2.1.1 Subgroup analyses: a. Clarified that values for prior lines of therapy will come from IWRS b. Added explanation of FISH allocation status (moved from section 3.5) c. Removed logistic regression modelling and all mention of comparative analyses.

Version	Description of Update
	d. Added repeat of subgroup analysis for biomarker groups. (xviii) Section 4.2.3.1 Progression free survival: a. Removed all comparative/inferential analyses. b. Added repeat of PFS analysis and KM plots of PFS for patients in second line patients. (xix) Section 4.2.3.1.1 PFS sensitivity analyses – Section removed. (xx) Section 4.2.3.2 Duration of response – Added repeat of DoR analysis and KM plot for patients in second line patients. (xxi) Section 4.2.3.3 Overall survival – Removed bullet point and combined with paragraph above. (xxii) Section 4.2.3.4 Change in TL tumour size: a. Removed all comparative/inferential analyses. b. Added repeats of waterfall and spider plots in second line patients. (xxiii) Section 4.2.4 Data cut-offs – Amended so originally planned futility, primary and final analyses will no longer take place, and instead that under CSP v3.0 there will be an initial DCO for early data review and a final DCO after all patients have had a 9 month follow-up. (xxiv) Section 4.2.5.2 Adverse events: a. Amended so summary information will be tabulated by treatment group (not savolitinib or osimertinib or the combination). b. Added definition of treatment emergent. (xxv) Section 4.2.6 Demographic and baseline characteristics – Amended age groups, and added to disease characteristics at baseline. (xxvi) Section 4.2.8 Pharmacokinetic data – Removed calculation of ratio of $C_{pre\text{-}dose}$ Cycle 11 Day 1 / $C_{pre\text{-}dose}$ Cycle 2 Day 1 since Cycle 11 due to removal of requirement for PK sample collection at Cycle 11. Clarified this will not be summarised by biomarker group. (xxvii) Section 4.2.10 Analyses for cross-over patients: a. Referred to Section 3.1 for assessment schedule

Version	Description of Update
	b. Simplified to say selected analyses will be repeated for these patients, including baseline, efficacy and safety. c. Clarification that cross-over analysis will not be repeated by biomarker groups. (xxviii) Section 5 Interim Analyses – Removal of interim analysis with explanation that it is no longer required under CSP V3.0. (xxix) Section 6 Changes of Analysis from Protocol: a. Noted the removal of comparative analyses from the study due to low sample size. b. Noted the removal of BICR. c. Removed the change previously noted relating to ANCOVA models, since this is now redundant. (xxx) Section 9.1 Important Protocol Deviations – List amended in line with updated non-compliance handling plan v3, and TA SAP.
3.0	The following changes have been made following the decision to reintroduce the BICR and associated analyses, and some comparative analyses: (i) Section 3.1.4 Independent review – Reverted back to text from v1.3 which states a planned blinded independent central review (BICR) will be carried out. (ii) Section 3.2.1 Objective response rate – Reverted back to text from v1.3 which includes sentence about BICR data. (iii) Section 3.2.2 Progression free survival – Reverted back to text from v1.3 regarding rules applying to BICR assessments. (iv) Section 3.2.5 Best objective response – Reverted back to text from v1.3 which includes derivation using BICR data. (v) Section 4.1 General Principles – Re-inserted sentence on stratification variables source to be used. (vi) Section 4.2.2.1 Objective response rate – Added back sentence from v1.3 about summaries per BICR data, and text originally from an ORR sensitivity analysis section describing how disagreements will be presented (vii) Section 4.2.3.1. Progression free survival (PFS): a. Re-inserted text from v1.3 describing estimation of HR and its corresponding CI, with added clarification that this will only be done if there are $\geq$ 20 events. Similarly, this will be repeated for second line patients and per BICR if there are $\geq$ 20 events. b. Added back sentence from v1.3 about summaries per BICR data, and text originally from a PFS sensitivity analysis section describing how disagreements will be presented.

Version	Description of Update
	(viii) Section 4.2.3.3 Overall survival (OS) – Added clarification that analyses will not be repeated for second line patients. (ix) Section 6 Changes of Analysis from Protocol – Removed mention of removal of BICR, and clarified comparative analyses may still be performed for PFS and OS if there are at least 20 events.

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: d5084c00009-sap-ed-3.0		
Document Title:	Statistical Analysis Plan Edition 3.0	
Document ID:	CCI [redacted]	
Version Label:	3.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
16-Aug-2022 14:20 UTC	PPD	Author Approval
16-Aug-2022 15:24 UTC	PPD	Content Approval
16-Aug-2022 13:09 UTC	PPD	Content Approval

CCI [redacted].