AstraZeneca

Statistical Analysis Plan

Study Code	D933KC00001
Edition Number	4.0
Date	27 April 2023

A Phase III, Randomized, Placebo-controlled, Double-blind, Multi-center, International Study of Durvalumab Given Concurrently with Platinum-based Chemoradiation Therapy in Patients with Locally Advanced, Unresectable Non-small Cell Lung Cancer (Stage III) (PACIFIC2)

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
[redacted]	[redacted]
AE	Adverse Event
AEPI	Adverse Event of Possible Interest
AESI	Adverse Event of Special Interest
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
APF12	Alive and Progression Free at 12 months
APF18	Alive and Progression Free at 18 months
AST	Aspartate Aminotransferase
AUC	Area Under the Concentration versus Time Curve
BICR	Blinded Independent Central Review
BMI	Body Mass Index
BoR	Best Objective Response
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
CR	Complete Response
CrCl	Creatinine Clearance
CRF	Case Report Form
CRT	Chemoradiation Therapy
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Event
CCI [redacted]	
CV	Coefficient of Variation
DBL	Database Lock
DBP	Diastolic Blood Pressure
DCO	Data Cutoff
DCR	Disease Control Rate

Abbreviation or special term	Explanation
DoR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EGFR	Epidermal Growth Factor Receptor
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQoL 5 dimension, 5-level health state utility index
FAS	Full Analysis Set
HLT	Higher Level Term
HR	Hazard Ratio
HRQoL	Health-Related Quality of Life
IA	Interim Analysis
IDMC	Independent Data Monitoring Committee
IP	Investigational Product
IPCW	Inverse Probability of Censoring Weighting
ITT	Intent-to-Treat
IV	Intravenous
IVRS	Interactive Voice Response System
LD	Longest Diameter
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MRI	Magnetic Resonance Imaging
NA	Not Applicable
nAb	Neutralizing Antibody
NE	Not Evaluable
NED	No Evidence of Disease
NL	New Lesion
NSCLC	Non-Small Cell Lung Cancer
NTL(s)	Non-Target Lesion(s)
ORR	Objective Response Rate
OS	Overall Survival
OS24	Proportion of Patients Alive at 24 Months from Randomization

Abbreviation or special term	Explanation
PD	Progressive Disease
[redacted]	[redacted]
PFS	Progression-Free Survival
[redacted]	[redacted]
PGIS	Patients' Global Impression of Severity
[redacted]	
PR	Partial Response
PRO	Patient-Reported Outcome
PT	Preferred Term
PTV	Planning Target Volume
[redacted]	
QLQ-C30	30-Item Core Quality of Life Questionnaire
QLQ-LC13	13-Item Lung Cancer Quality of Life Questionnaire
QoL	Quality of Life
RDI	Relative Dose Intensity
RECIST	Response Evaluation Criteria In Solid Tumors
REML	Restricted Maximum Likelihood
RPSFT	Rank Preserving Structural Failure Time
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SD	Stable Disease
SoC	Standard of Care
SOC	System Organ Class
TFST	Time to First Subsequent Therapy or Death
TL(s)	Target Lesion(s)
TSH	Thyroid Stimulating Hormone
TSST	Time to Second Subsequent Therapy or Death
TTDM	Time to Death or Distant Metastasis
TTR	Time to Response
ULN	Upper Limit of Normal
WHO	World Health Organization

AMENDMENT HISTORY

CATEGORY Change refers to:	**Date**	**Description of change**	**In line with CSP?**	**Rationale**
Version 4				
Secondary objectives table	05/26/2023	Section 6: Specified that the secondary objectives table includes duplicated rows for [redacted] and that only the investigation of [redacted] of durvalumab when in combination with CRT will be carried out.	No	To avoid confusion on which analysis will be carried out.
Reference update	05/26/2023	Section 3.3.1: Updated the reference for the EORTC QLQ-C30 Scoring Manual from 2nd Edition (Fayers et al 1999) to 3rd Edition (Fayers et al 2001)	Yes	To have the same reference as included in the CSP
Immunogenicity analysis	05/26/2023	Section 3.5.2: Made the definition of CCI [redacted] CCI [redacted] CCI [redacted]	Yes	To avoid any confusion regarding the definitions. The revised language is consistent with the existing programming for CCI [redacted]
Analysis set definition	04/27/2023	Section 2.1.3: Providing detailed rules for defining CCI [redacted] CCI	Yes	Obtain detailed CCI [redacted] and provide clarification for programming team to select patients/visits CCI [redacted]
Secondary endpoint OS interim analysis	04/27/2023	Section 5.1.2: Clarify that, CCI [redacted] CCI [redacted] CCI [redacted] of the	Yes	Adding details and clarifications to the OS interim and final analysis.

		planned number of OS events for the OS final analysis (approximately CCI OS events at the PFS final analysis), [redacted], and will use [redacted] the CCI or CCI alpha, depending on whether ORR is statistically significant.		
Handling changes to the analysis in the protocol	4/27/2023	Section 6: added clarification that analyses in the SAP supersede those in the protocol.	Yes	To add clarification on situations when SAP and protocol have different analysis approaches, as this is a general principal for the SAP.
Version 3				
Multiple testing procedure	8/12/2022	Section 4.1.2 and section 5: Clarification to the alpha spend for ORR and OS24	Yes	Ensure strong control of the overall type 1 error
	8/12/2022	Figure 2: Added a line going from OS24 back to ORR	Yes	To fully illustrate the CCI CCI CCI [redacted] Section 4.1.2.
IDMC Review	8/12/2022	Section 1.2 and Section 5.1.3: Clarified that the CCI CCI [redacted]	No	CCI [redacted] CCI [redacted] CCI [redacted] CCI [redacted] [redacted] [redacted]
Interim analyses	8/12/2022	Section 1.3 and Section 5.1.2: Clarified that CCI [redacted] CCI [redacted] instead of	No	Based on blinded review of the OS data the number of

		the [redacted] with alpha adjustment for the OS interim [redacted] information fraction		OS events at the PFS final analysis is expected to reach [redacted]
Statistical analysis method for secondary endpoint(s)	8/12/2022	Section 4.2.2.1: Change the estimate of [redacted]	Yes	To use an approach based on the difference in proportions which is accepted by FDA
	8/12/2022	Section 4.2.2.3: Remove reference to the treatment comparison at 24 months being a Hazard Ratio	Yes	OS24 Hazard Ratio removed as the formal statistical test output (p value) will be reported by the method described in Klein 2007 and the difference between the treatment groups in OS24 will be described as outlined in section 4.2.2.3.
Safety reporting	8/12/2022	Section 4.1.3: Remove the rule that only produces safety summaries if number of observations is >20 and >1/3 patients dosed for the ECG summary table only	Yes	To enable the ECG summary by visit table to be populated in the TFLs even if there are only a small number of observations
	8/12/2022	Section 4.1.4: Update the imputation of partial stop dates of adverse events and medications/therapies to have the last dose date as the latest date, not the first dose date, in	Yes	Ensure imputed dates are conservative

		the situation the stop date is in the same month as last dose		
	8/12/2022	Section 4.2.4.1: Add summaries of AEs of maximum CTCAE Grade 3 or 4	Yes	Outputs needed for regulatory submissions
	8/12/2022	Section 4.2.4.1: Remove section on Infection Adverse Events	Yes	No longer required from a safety perspective
Data presentation	8/12/2022	Section 4.2.9: Add a cross-tabulation comparison of stratification factors recorded in the IVRS vs eCRF	Yes	To summarize mis-stratifications
	8/12/2022	Section 2.2.1: Included COVID-19 related IPDs	No	To summarize impact of COVID-19
	8/12/2022	Section 4.2.11: Included COVID-19 listing based on Clinical Trial Management System	No	To summarize impact of COVID-19
Derivation of secondary endpoint(s)	8/12/2022	Section 3.3.1.1 and Section 3.3.1.2: Updated approach taken for deterioration for EORTC QLQ-C30 and QLQ-LC13 to use a 'look-forward' approach	No	Conservative approach taken to flag deterioration
	8/12/2022	Section 3.3.1.3 and Section 3.3.1.4: Updated approach taken for improvement rate for EORTC QLQ-C30 and QLQ-LC13 so that missed visits are not allowed in between consecutive assessments of improvement.	No	Conservative approach taken to flag improvement
	8/12/2022	Appendix A: Added section to clarify confirmation and censoring of visits for time to deterioration	No	Clarification and further detail provided
Version 2				
Primary endpoint(s)	9/25/2020	Updated to reflect the changes in protocol amendments #3 and #4 (i.e., CCI [redacted] CCI)	Yes	Align to the changes made to the primary endpoint definition in the CSP

		CCI [redacted] These changes are reflected in: • Section 1.1.1 - 1.1.2 • Section 3.2.1 • Section 3.2.1.2 moved to Section 3.2.2.1 - Sections 3.2.2.[x] renumbered accordingly • Section 4.2.1.1 • Section 4.2.1.2 moved to Section 4.2.2.1 - Sections 4.2.2.[x] renumbered accordingly		
Multiple testing procedure	9/25/2020	Updated text to reflect CCI [redacted] CCI [redacted] CCI [redacted] These changes are reflected in: • Section 1.2 - 1.3 • Section 4 - 4.1.2 • Section 5.1.1 - 5.1.2 - 5.1.3	Yes	Align to changes made in the CSP to CCI [redacted] CCI [redacted] CCI [redacted] CCI [redacted]
Derivation of primary endpoint(s)	9/25/2020	Section 3.2.1.1: Clarified (according to TA SAP) that the censoring date for PFS is last evaluable RECIST1.1 assessment.	Yes	Clarification that the last RECIST 1.1 assessment used for PFS censoring must have been an evaluable RECIST 1.1 assessment
Derivation of secondary endpoint(s)	9/25/2020	Section 3.2.1.2: Removed visit windowing rule from the derivation of the supportive endpoint “change in tumor size”.	Yes	Align with other efficacy endpoints ensuring all RECIST 1.1 assessments are used in the analysis
	9/25/2020	Section 3.2.2.7: Removed the 2 missed visits rule for [redacted] and	Yes	As [redacted] is collected per local standard clinical practice,

		clarified the censoring rule for patient who died as [redacted]		visit intervals vary among patients
	9/25/2020	Section 4.2.1.1: Added clarifications for the [redacted].	Yes	To ensure sufficient patient numbers are within each subgroup for analysis
Statistical analysis method for secondary endpoint(s)	9/25/2020	Section 3.2.2.5: Removed supportive analysis for time to response (TTR).	Yes	Deemed to be no longer required
	9/25/2020	Section 4.2.1: Removed the details on the management of ties for the stratified log-rank test.	Yes	Text was redundant as it is not necessary to define the ties handling method for a log rank test
	9/25/2020	Section 4.2.2.1: Clarified that summary of BoR is required on both BICR and site investigator data.	Yes	To support the ORR analysis which is also produced for BICR and for investigator data
	9/25/2020	Section 4.2.2.4: Removed formal testing for rate of complete response.	Yes	Clarify that referring to the same method as for testing ORR does not include a formal test for Rate of Complete Response
	9/25/2020	Section 4.2.2.5: Added landmark analysis for DoR.	No	To support the DoR endpoint
	9/25/2020	Section 4.2.2.7: Added landmark analysis for TTDM.	No	To support the TTDM endpoint
	9/25/2020	Section 4.2.3: Added clarifications on the Bonferroni procedure. Added clarifications on the PGIS analysis.		Clarify that the Bonferroni procedure is used to control the overall type 1 error at 5% 2-sided;

				Clarified how the PGIS data will be summarized in further detail
Safety reporting	9/25/2020	Section 3.4.1: Modified AE of interest section to include the definition of AE of possible interest (AEPI). Removed definition of other significant AE (OAE).	Yes	To align with safety reporting of AEs of both special and possible interest. Clarified that AEs classed as OAEs no longer need to be summarized
	9/25/2020	Section 3.4.5: Removed derivation of actual exposure for all chemotherapy regimens, since actual exposure and RDI will be summarized for durvalumab or matching placebo only.	Yes	Actual exposure calculations are not required for chemotherapy regimens
	9/25/2020	Section 4.1.4: Added details of the imputation rules for missing/incomplete dates of adverse events and medications/therapies. Added details of the imputation rules for incomplete dates of death and start of subsequent cancer therapies.	Yes	To align with the Oncology TA SAP guidance to ensure consistency with how missing/incomplete AE/medication start and stop dates are imputed
	9/25/2020	Section 4.2.4.1: Added details for the summaries of adverse events of special interest, infection adverse events and adverse events of pneumonitis and radiation pneumonitis. Added subsection related to summary of long-term tolerability. This subsection describes the graphical summary of prevalence plots and cumulative incidence plots. Modified summary of AE of interest to include AEPI.	Yes	To further support the analysis of safety data
	9/25/2020	Section 4.2.4.2: Added summary of reversibility of creatinine clearance. Clarified summary for	Yes	To further support the analysis of safety data

		TSH abnormalities in subsection "Thyroid Function Test Results".		
Data presentation	9/25/2020	Section 3.6.2: Added clarification on the EQ-5D-5L questionnaire.	Yes	Additional details on the EQ-5D-5L added for readability
	9/25/2020	Section 4.1.3: Clarified different display rules for safety and PRO assessments for visit based summaries.	Yes	Distinguish between PRO data and safety data for clarity on the minimum amount of data required to create summaries per visit
	9/25/2020	Section 4.2.11: Added new section to investigate and report the impact of COVID-19 on the study.	No	To summarize impact of Covid-19
	9/25/2020	Section 6: Clarified removal of OAE. Updated wording.	No	Align with changes of analysis from CSP
N/A	10/19/2018	Initial approved SAP	N/A	N/A

1. STUDY DETAILS

1.1 Study objectives

1.1.1 Primary objective

Primary objective:	Endpoint/variables:
To assess the efficacy of durvalumab + SoC CRT compared with placebo + SoC CRT in terms of PFS	PFS using BICR assessments according to RECIST 1.1 [a]

[a] PFS will be based on programmatically derived PFS by BICR assessment according to RECIST 1.1. See statistical methods section for further details.

1.1.2 Secondary objectives

Secondary objectives:	Endpoint/variables:
To further assess the efficacy of durvalumab + SoC CRT compared with placebo + SoC CRT in terms of ORR, OS, OS24, rate of CR, DoR, DCR, [redacted], and TTDM	ORR using BICR assessments according to RECIST 1.1 OS and OS24 Rate of CR, DoR, DCR, and TTDM using BICR assessments according to RECIST 1.1 [redacted] as defined by local standard clinical practice
CCI [redacted]	CCI [redacted]
CCI [redacted]	CCI [redacted]
CCI [redacted]	CCI [redacted]
To assess symptoms and health-related QoL in patients treated with durvalumab + SoC CRT compared with placebo + SoC CRT using EORTC QLQ-C30 v3 and QLQ-LC13	EORTC QLQ-C30 and QLQ-LC13: Change in symptoms, functioning, and global health status/QoL

1.1.3 Safety objectives

Safety objective:	Endpoint/variables:
To assess the safety and tolerability profile of durvalumab + SoC CRT compared with placebo + SoC CRT	AEs, physical examinations, vital signs including blood pressure, pulse, electrocardiograms, and laboratory findings including clinical chemistry, hematology, and urinalysis

1.1.4 Exploratory objectives [a]

Exploratory objectives:	Endpoint/variables:
CCI [redacted]	[redacted]

CCI	CCI

[a] Exploratory objectives maybe presented outside the CSR

1.2 Study design

This is a Phase III, randomized, double-blind, placebo-controlled, multi-center, international study assessing the efficacy and safety of durvalumab given concurrently with platinum-based CRT (durvalumab + standard of care [SoC] CRT) in patients with locally advanced, unresectable NSCLC (Stage III).

Study period:

Number of patients:

Approximately CCI patients with locally advanced, unresectable NSCLC (Stage III) will be recruited and 300 patients randomized in a 2:1 ratio to durvalumab + SoC CRT or placebo + SoC CRT. Patients will be stratified by age CCI and stage CCI CC CCI

Treatments and treatment duration:

All patients will receive 1 of the following platinum-based SoC chemotherapy options, based on Investigator's discretion, in addition to radiation therapy: cisplatin/etoposide, carboplatin/paclitaxel, pemetrexed/cisplatin, or pemetrexed/carboplatin. Chemotherapy treatment regimens are outlined in Table 6.

Patients will also receive durvalumab CCI mg or placebo every CCI via intravenous infusion concurrent with SoC CRT (i.e., starting on Cycle 1 Day 1 [±3 days]). Patients with complete response (CR), partial response (PR), or stable disease (SD) based on Investigator assessment at the 16-week tumor evaluation following completion of SoC CRT will continue to receive durvalumab/placebo as consolidation treatment, unless there is unacceptable toxicity, withdrawal of consent, or another discontinuation criterion is met. Patients with RECIST 1.1-defined radiological progressive disease at the 16-week tumor evaluation following completion of SoC CRT will proceed to follow-up.

Independent Data Monitoring Committee:

An independent data monitoring committee (IDMC) composed of independent experts will be convened to confirm the safety and tolerability of the proposed dose and schedule of durvalumab + SoC CRT. The first safety review will take place CCI have completed SoC CRT and have had at least CCI The second safety review will take place when CCI have completed SoC CRT and have had at least CC days of follow-up. Safety reviews will be carried out by the IDMC in an unblinded manner. After review of the

unblinded data, the IDMC will make a recommendation on whether the study should continue recruitment as planned or hold recruitment.

An additional safety review for CCI patients will take place CCI This review will be carried out by the IDMC in an unblinded manner. After review of the unblinded data, the IDMC will make a recommendation on whether the study should continue recruitment as planned or hold recruitment in CCI

The IDMC will meet approximately every 6 months thereafter to continue safety monitoring. In addition, the IDMC will review:

- CCI

Note: At the PFS final analysis, AstraZeneca will be unblinded to the data (regardless of the outcome). Hence the IDMC will not be required to review the efficacy data for the OS IA.

Full details of the IDMC procedures, processes, and IAs can be found in the IDMC Charter.

Figure 1

CR Complete response; CRT Chemoradiation therapy; ECOG Eastern Cooperative Oncology Group; n Number of patients; NSCLC Non-small cell lung cancer; ORR Objective response rate; OS Overall survival; PD Progressive disease; PFS Progression-free survival; PR Partial response; SD Stable disease; SoC Standard of care; WHO World Health Organization.

1.3 Number of patients

The primary objective of this study is to assess the efficacy of durvalumab + SoC CRT compared with placebo + SoC CRT in terms of PFS per Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1) as assessed by Blinded Independent Central Review (BICR). The key secondary endpoints (i.e., those included in the multiple testing procedure) are ORR per RECIST 1.1 as assessed by BICR, OS and OS24.

A multiple testing procedure will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoints of PFS and the key secondary endpoints, ORR, OS and OS24. The family-wise error rate is strongly controlled at 5% for these endpoints.

Approximately 300 patients will be randomized 2:1 to 2 arms: durvalumab + SoC CRT or placebo + SoC CRT. A recruitment period of approximately [redacted] is anticipated. [redacted]

The study is powered for the [redacted]. There are up to [redacted] analysis timepoints [redacted] approximately [redacted] of [redacted] PFS events (information fraction of [redacted] have occurred. It is estimated that this DCO will occur [redacted] months after the last patient has been randomized. The DCO for the PFS final analysis will occur when approximately [redacted] events have occurred. It is estimated that this DCO will occur [redacted] after the last patient has been randomized.

If the true PFS HR is [redacted] the study will provide greater than [redacted] power to demonstrate a statistically significant PFS effect with a [redacted] this translates to a [redacted] month benefit in median PFS over [redacted] months on placebo if PFS is exponentially distributed. The smallest treatment effect that would be statistically significant is an HR of [redacted] at final analysis.

Per protocol there are [redacted] for formal statistical testing of OS. [redacted] when approximately [redacted] PFS events have occurred) where it was predicted per protocol, that approximately [redacted] of [redacted] OS events (information fraction of [redacted] would be reached. The DCO for the [redacted] planned [redacted] approximately [redacted] of [redacted] OS events (information fraction of [redacted] are reached. However, [redacted] Hence, [redacted] will now actually be performed using the [redacted] target number of OS events (see Section 5.1 for further details).
The DCO for the final OS analysis will occur when approximately [redacted] events are reached. [redacted] Any subsequent protocolled analysis timepoints may still be assessed but will be considered as an updated exploratory analysis outside of the multiple testing procedure.

If the true OS HR is [redacted] the study will have greater than [redacted] overall power to demonstrate a statistically significant OS effect with a [redacted] this translates to a [redacted] month benefit in median OS over [redacted] months on placebo if OS is exponentially distributed. The smallest treatment difference that would be statistically significant is a HR of [redacted] at final analysis. The study has [redacted] power to detect a statistically significant difference in ORR of

[redacted] with a 2-sided significance level of [redacted] This assumes the ORR for the patients randomized to placebo + SoC CRT is [redacted] The smallest treatment effect that would be statistically significant is a difference in ORR of [redacted]

2. ANALYSIS SETS

2.1 Definition of analysis sets

Summaries of the analysis sets for each outcome variable are provided in Table 1.

Table 1 Summary of outcome variables and analysis populations

Outcome variable	Populations
Efficacy Data	
PFS	FAS (ITT population)
ORR, OS, OS24, CR rate, DoR, DCR at [redacted], PRO endpoints, and TTDM	FAS (ITT population)
Study Population/Demography Data	
Demography	FAS (ITT population)
Baseline and disease characteristics	FAS (ITT population)
Important deviations	FAS (ITT population)
Medical/surgical history	FAS (ITT population)
Previous anti-cancer therapy	FAS (ITT population)
Concomitant medications/procedures	FAS (ITT population)
[redacted]	
Safety Data	
Exposure	Safety analysis set
AEs	Safety analysis set
Laboratory measurements	Safety analysis set
Vital Signs	Safety analysis set

[redacted] CR Complete response; DCR Disease control rate; DoR Duration of response; FAS Full analysis set; ITT Intent-to-Treat; ORR Objective response rate; OS Overall survival; OS24 Overall Survival at 24 months; PFS Progression-free survival; [redacted] PRO Patient-reported outcomes; TTDM Time to death or distant metastasis

2.1.1 Full analysis set (ITT population)

The full analysis set (FAS) will include all randomized patients. The FAS will be used for all efficacy analyses (including PROs). Treatment arms will be compared based on randomized

study treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment will be included in the analysis in the treatment arm to which they were randomized. The analysis of data using the FAS therefore follows the principles of ITT.

2.1.2 Safety analysis set

The safety analysis set will consist of all patients who received at least 1 dose of randomized treatment (durvalumab or matching placebo). Safety data will be summarized using the safety analysis set according to treatment received, that is, erroneously treated patients (e.g., those randomized to treatment A but actually given treatment B) will be summarized according to the treatment they actually received.

[redacted]

2.2 Violations and deviations

The following general categories will be considered important deviations and will be listed and discussed in the CSR as appropriate for the study:

- Deviation 1: Patients randomized but who did not receive durvalumab/matching placebo
- Deviation 2: Patients randomized but who did not receive protocol defined CRT
- Deviation 3: Patients who deviate from the following key entry criteria in CSP:
    a) Inclusion criteria: 4, 5, 6;
    b) Exclusion criteria: 8, 11, 12, 14, 17, 20, 21, 22.
- Deviation 4: Baseline RECIST scan > 42 days before randomization.
- Deviation 5: No baseline RECIST 1.1 assessment on or before date of randomization.

- Deviation 6: Received prohibited concomitant medications.
  Refer to the CSP section 6.5 for those medications that are detailed as being 'prohibited' from permitted use during the study. This will be used as a guiding principle for the physician review of all medications prior to DBL.

- Deviation 7: Patients randomized who received treatment other than that to which they were randomized.

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

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 regards 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 summarized by randomized treatment group, including COVID-19 related IPDs. Deviation 1 will lead to exclusion from the Safety analysis set. None of the other deviations will lead to patients being excluded from the analysis sets described in Section 2.1 CCI [redacted] A per-protocol analysis excluding patients with specific important protocol deviations is not planned; however, a 'deviation bias' sensitivity analysis may be performed excluding patients with deviations that may affect the efficacy of the trial therapy if CCI [redacted] of patients in either treatment group:

- did not have the intended disease or indication or

- did not receive any randomized therapy.

The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock and will be documented prior to the primary analysis being conducted.

In addition to the programmatic determination of the deviations above, other study deviations captured from the 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.

In addition, all COVID-19 related non-important PDs and issues will be summarized and listed and included in the CSR.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST visit responses

For all patients, the RECIST tumor 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 tumor assessments are to be performed no more than [CCI] before randomization and ideally as close as possible to and prior to the date of randomization . Tumor assessments will begin at [CCI] ± 1 week after randomization, then every [CCI] ± 1 week through [CCI] (relative to the date of randomization), and then every [CCI] ± 1 week thereafter (relative to the date of randomization) until 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 minimize 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 tumor 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 CR, PR, SD or 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 tumor assessment which cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response will be assigned as PD).

RECIST outcomes (i.e., PFS and ORR etc.) will be calculated programmatically for the BICR and site investigator data from overall visit responses.

3.1.1 Site investigator assessments using RECIST 1.1

3.1.1.1 Target Lesions (TLs)

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 computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. A patient can have a maximum of five target lesions recorded at baseline with a maximum of two target lesions per organ (representative of all lesions involved and suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to randomization 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 non-target lesions (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 3.1.2.2 for further details). If a patient does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 2 TL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all TLs since baseline. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10 mm.
Partial response (PR)	At least a 30% decrease in the sum of the diameters of TL, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.
Progression of 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.
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 1 decimal place (d.p.) before assigning a target lesion 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 target lesion 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;
- A NTL visit response of PD is recorded;
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 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 lesion diameter 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

A CR response can only be followed by CR, PD or NE. 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 is also met.
- Step 3: If not all lesions meet the CR criteria and the sum of lesions meets the criteria for PD 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 target lesion 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 target lesion 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 target lesion 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 then this will be reviewed by the study team blinded to treatment assignment.

Irradiated lesions/lesion intervention

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

The current study includes radiotherapy as study treatment, and this will not be considered as a TL intervention in the eCRF. Any TL (including lymph nodes), which has had intervention in addition to study treatment during the study (e.g., irradiation / palliative surgery / embolization), should be handled in the following way and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumors:

- 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 during the study [note: the radiotherapy component of study treatment is not considered an intervention])

If > 1/3 of target lesion measurements are missing (because of intervention) then target lesion response will be NE, unless the sum of diameters of non-missing target lesion 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 target lesions has increased by ≥5mm from nadir).

If ≤ 1/3 of the target lesion 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; it had a BL measure of 29.3cm.

The sum of lesions 1-4 at the follow-up is 26 cm. The sum of the corresponding lesions at the nadir visit is 26.8 cm.

Scale up as follows to give an estimated TL sum of 28.4cm:

$$\frac{26}{26.8} \times 29.3 = 28.4cm$$

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 target lesions 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 and MRI are the only methods of assessment that can be used within this trial, with CT and MRI being the preferred methods. 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.

Note, 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.

3.1.1.2 Non-Target Lesions (NTLs) and new lesions.

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

Non-target lesion response will be derived based on the Investigator's overall assessment of NTLs as follows:

Table 3 NTL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all NTLs since baseline. All lymph nodes must be non-pathological in size (<10 mm short axis).
Non CR/non PD	Persistence of one or more NTLs with no evidence of progression.
Progression (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.
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' based on 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 tumor burden has increased sufficiently to merit a discontinuation of therapy. A modest 'increase' in the size of one or more NTLs is usually not sufficient to qualify for unequivocal progression status.

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

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

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

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

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

3.1.1.3 Overall visit response

Table 4 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 4 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**

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
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**

3.1.2 Blinded Independent Central Review (BICR)

A planned 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 Organization (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 tumor 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, DoR, etc..) 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 DCO.

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

BICR according to RECIST 1.1 will be regarded as primary in terms of the efficacy analyses.

3.2 Outcome variables

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

RECIST 1.1 outcomes (i.e., PFS, ORR, etc.) will be derived using the overall visit responses and relevant dates from the BICR. This will be repeated using the programmatically derived overall visit responses from investigator RECIST 1.1 assessments.

3.2.1 Primary efficacy outcome endpoints

3.2.1.1 Progression-free survival (PFS)

PFS (per RECIST 1.1 as assessed by the BICR) is defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from therapy or receives another anticancer therapy prior to progression (i.e., date of event or censoring – date of randomization + 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 1.1 assessment. [redacted]

Given the scheduled visit assessment scheme (i.e., week [redacted] after randomization, then [redacted] weekly through to week [redacted] and then [redacted] weekly thereafter) the definition of 2 missed visits will change over time. If the previous RECIST assessment is between study day 106 (i.e., week 15) and less than study day [redacted] then two missed visits will equate to [redacted] since the previous RECIST assessment, allowing for early and late visits (i.e., [redacted] + [redacted] for an early assessment + [redacted] for a late assessment = [redacted] If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from eight-weekly to twelve-weekly this will equate to [redacted] (i.e., take the average of 8 and [redacted] which gives [redacted] and then apply same rationale, hence [redacted] + [redacted] for an early assessment + [redacted] for a late assessment = [redacted] The time period for the previous RECIST assessment will be from study days 274 to 344 (i.e., week 39 to week 49). From week 49 onwards (when the scheduling changes to twelve-weekly assessments), two missing visits will equate to [redacted] (i.e., [redacted] + [redacted] for an early assessment + [redacted] for a late assessment = [redacted]

If the patient has no visits or does not have baseline data they will be [redacted] unless they die within 2 visits of baseline [redacted] plus [redacted] allowing for a [redacted] within the visit window).

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

RECIST 1.1 assessments/scans contributing toward 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** 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 Investigator assessments, the date of progression will be determined based on the **earliest** of the RECIST 1.1 assessment/scan dates of the component that indicates progression.
- When censoring a patient for PFS, the patient will be censored at the **latest** of the scan dates contributing to a particular overall visit assessment.

Note: For TLs, only the latest scan date within an imaging visit window is recorded in the RECIST 1.1 eCRF 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.

Time to first subsequent therapy or death (TFST)

As a supportive summary to PFS, time to first subsequent therapy or death (TFST) is defined as the time from the date of randomization to the earlier of start date of the first subsequent anti-cancer therapy after discontinuation of randomized treatment, or death (i.e., date of first subsequent cancer therapy/death or censoring – date of randomization + 1). Any patient not known to have had a first subsequent anti-cancer therapy will be censored at the last date that the patient was known not to have received a first subsequent anti-cancer therapy (obtained from the TTSCAPRX form). If a patient terminated the study for reason other than death before first subsequent therapy, these patients will be censored at the earliest of their last known to be alive and termination dates. Patients not receiving randomized treatment would have TFST calculated in the same way, i.e., time from date of randomization to the subsequent therapy or death.

3.2.2 Secondary efficacy outcome endpoints

3.2.2.1 Objective response rate (ORR)

ORR (per RECIST 1.1 as assessed by BICR) is defined as the number (%) of patients with a confirmed response of CR or PR and will be based on all patients in the ITT population (FAS).

A confirmed response of CR or PR means that a response of CR or PR is recorded at 1 visit and confirmed by repeat imaging not less than [redacted] after the visit when the response was first observed, with no evidence of progression between the initial and CR/PR confirmation visit. Therefore, data obtained up until progression, or the last assessment in the absence of progression, will be included in the assessment of ORR. Patients who discontinue treatment without progression, receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is considered a subsequent anti-cancer therapy), and then respond will not be included as responders in the ORR (i.e., both visits contributing to a response must be prior to subsequent therapy for the patient to be considered as a responder).

In the case where a patient has two non-consecutive visit responses of PR, then, as long as the time between the 2 visits of PR is greater than [redacted] and there is no PD between the PR visits, the patient will be defined as a confirmed responder. Similarly, if a patient has visit responses of CR, NE, CR, then, as long as the time between the 2 visits of CR is greater than [redacted] then a best response of CR will be assigned.

ORR will also be obtained using the algorithm described above for the RECIST 1.1 site Investigator tumor data.

Best objective response (BoR)

Best objective response (BoR) is calculated based on the overall visit response from each RECIST assessment. It is the best response a patient has had following randomization but prior to starting any subsequent cancer therapy and up to and including RECIST progression or the last assessment in the absence of RECIST progression.

Categorization 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.

CR or PR must be confirmed. 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 [redacted] minus [redacted] i.e. [redacted] (to allow for an early assessment within the assessment window), after randomization. For CR/PR, the initial overall visit assessment which showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

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 PFS event is death, BoR will be calculated based upon all RECIST assessments prior to death.

For patients who die with no RECIST assessments, if the death occurs [redacted] (i.e., [redacted] for first assessment + [redacted] + [redacted] for a late assessment) after randomization, then BoR will be assigned to the progression (PD) category. For patients who die with no RECIST assessments, if the death occurs [redacted] after randomization 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 defined in the section above are satisfied at any time following randomization, prior to RECIST progression and prior to starting any subsequent cancer therapy.

Change in tumor size

For supportive purposes, percentage change from baseline in tumor size will be derived at each scheduled tumor assessment visit (i.e., week 16, week 24, week 32, etc. hereafter referred to as week X for convenience) using BICR data and site investigator data. Best percentage change from baseline in tumor size will also be derived as the biggest decrease or the smallest increase in tumor size from baseline.

This is based on RECIST TL measurements taken at baseline and at the timepoint of interest. Tumor size is defined as the sum of the longest diameters of the TLs based upon RECIST assessments. TLs are measurable tumor lesions. Baseline for RECIST is defined to be the last assessment prior to randomization. The change in TL tumor size at week X will be obtained for each patient by taking the difference between the sum of the TLs at week X and the sum of the TLs at baseline. To obtain the percentage change in TL tumor size at week X the change in TL tumor size is divided by the sum of the TLs at baseline and multiplied by 100 (i.e., (week X - baseline) / baseline * 100).

The above derivations will be programmed for the BICR based upon RECIST 1.1 assessments. Measurements from the reviewer selected by the adjudicator will be used when adjudication for overall visit response has occurred, but in the case where no adjudication was required, the measurements from the reviewer who reviewed the baseline scan first will be used for this analysis.

3.2.2.2 Overall survival (OS)

Overall survival (OS) is defined as the time from the date of randomization until death due to any cause regardless of whether the patient withdraws from randomized therapy or receives another anti-cancer therapy (i.e., date of death or censoring – date of randomization + 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 that 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

3.2.2.3 Proportion of patients alive at 24 months (OS24)

The proportion of patients alive at 24 months (i.e., OS24) will be defined as the Kaplan-Meier estimate of OS at 24 months. In addition, the proportion of patients alive at 12 months (i.e., OS12) will be presented. This will be defined as the Kaplan-Meier estimate of OS at 12 months.

3.2.2.4 Rate of complete response

The rate of CR (per RECIST 1.1 as assessed by BICR) is defined as the number (%) of patients with a confirmed response of CR.

A confirmed response of CR means that a response of CR is recorded at 1 visit and confirmed by repeat imaging not less than [CCI] after the visit when the response was first observed, with no evidence of progression between the initial and CR confirmation visit. Therefore,

data obtained up until progression, or the last assessment in the absence of progression, will be included in the assessment of the rate of CR. Patients who discontinue treatment without progression, receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is considered a subsequent anti-cancer therapy), and then have a CR will not be included as responders in the rate of CR (i.e., both visits contributing to a confirmed response of CR must be prior to subsequent therapy for the patient to be considered as a responder in the rate of CR).

In the case where a patient has two non-consecutive visit responses of CR, then, as long as the time between the 2 visits of CR is greater than [redacted] and there is no PD between the CR visits, the patient will be assigned to have a confirmed CR.

The rate of CR will also be obtained using the algorithm described above for the RECIST 1.1 site Investigator tumor data.

3.2.2.5 Duration of response (DoR)

DoR (per RECIST 1.1 as assessed by BICR) will be defined as the time from the date of first documented response (which is subsequently confirmed) until the first 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 RECIST 1.1 PFS endpoint.

The time of the initial response will be defined as the latest of the dates contributing toward the first visit response of CR or PR that was subsequently confirmed. If a patient does not progress following a response, then his or her DoR will be censored at the PFS censoring time. DoR will not be defined for those patients who do not have documented confirmed response.

3.2.2.6 Disease control rate (DCR)

DCR [redacted] (per RECIST 1.1 as assessed by BICR) is defined as the percentage of patients who have a best objective response (BoR) of CR or PR [redacted] or who have SD [redacted] after randomization [redacted]
Data obtained up until progression, or the last assessment in the absence of progression, will be included in the assessment of DCR.

3.2.2.7 Time to death or distant metastasis (TTDM)

TTDM (per RECIST 1.1 as assessed by BICR) will be defined as the time from the date of randomization until the first date of distant metastasis or death in the absence of distant metastasis. Distant metastasis is defined as any NL that is outside of the radiation field according to RECIST 1.1 or proven by biopsy. For this reason, the TTDM endpoint will be determined from recurrent disease that occurs outside of the structures contained within the thorax, including lymph nodes, pulmonary, pleural, and mediastinal metastatic sites and

excluding the heart. The locations of distant metastases will be defined and documented prior to database lock/unblinding.

Patients who have not developed distant metastasis or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment. However, if the patient has distant metastasis or dies after 2 or more missed visits, the patient will be censored at the time of the latest RECIST 1.1 assessment prior to the 2 missed visits. If the patient has no visits or does not have baseline data, he/she will be censored at Day 1 unless they die within 2 visits of baseline.

[redacted] The date of the [redacted] will be programmatically determined from investigator-assessed data (see section 3.1.1 for details). The date of [redacted] will be recorded by the investigator and defined according to local standard clinical practice and may involve any of; objective radiological, symptomatic progression or death. The date of the [redacted] assessment and investigator opinion of [redacted] at each assessment will be recorded in the electronic case report form (eCRF). [redacted] should be reviewed at regular assessments following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and [redacted] if the patient has not had [redacted]

Time to second subsequent therapy or death (TSST)

[redacted] time to second subsequent therapy or death (TSST) is defined as the time from the date of randomization to the earlier of start date of the second subsequent anti-cancer therapy after discontinuation of first subsequent treatment, or death (i.e., date of second subsequent cancer therapy/death or censoring – date of randomization + 1). Any patient not known to have had a second anti-cancer subsequent therapy will be censored at the last date that the patient was known not to have received a second subsequent anti-cancer therapy (obtained from the TTSCAPRX form). If a patient terminated the study for reason other than death before second subsequent therapy, these patients will be censored at the earliest of their last known to be alive and termination dates. Patients not receiving randomized treatment would have TSST calculated in the same way, i.e. time from date of randomization to the subsequent therapy or death.

3.3 Patient reported outcome (PRO) variables

Symptoms and overall quality of life will be assessed using the PRO questionnaires, EORTC QLQ-C30 and QLQ-LC13 (secondary endpoints). All questionnaires will be scored according

to published guidelines or the developer's guidelines, if published guidelines are not available. All PRO analyses will be based on the FAS, unless stated otherwise.

3.3.1 EORTC QLQ-C30 and QLQ-LC13

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea/vomiting), 5 individual items (dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a global measure of health status. The global health status/HRQoL will be assessed using 2 items from the QLQ-C30: "How would you rate your overall health during the past week? (Item 29) and "How would you rate your overall QoL during the past week? (Item 30).

The QLQ-LC13 is a lung cancer specific module from the EORTC for lung cancer comprising 13 questions to assess lung cancer symptoms (cough, hemoptysis, dyspnea, and site-specific pain), treatment-related symptoms (sore mouth, dysphagia, peripheral neuropathy, and alopecia), and pain medication. With the exception of a multi-item scale for dyspnea, all are single items. The dyspnea scale will only be used if all 3 items have been scored; otherwise, the items are treated as single-item measures.

An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, the functional scales and the global health status scale in the QLQ-C30 and for each of the symptom scales/items in the QLQ-LC13 according to the EORTC QLQ-C30 Scoring Manual and EORTC QLQ-LC13 instructions.

Higher scores on the global health status/QoL and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity. For each subscale, if <50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded. If there is evidence that the missing data are systematic, missing values will be handled to ensure that any possible bias is minimized.

Definition of clinically meaningful changes

Changes in score compared to baseline will be evaluated. A minimum clinically meaningful change is defined as a change in the score from baseline of ≥10 for scales/items from the QLQ-C30 and the QLQ-LC13 (Osoba et al 1998). For example, a clinically meaningful deterioration or worsening in chest pain (as assessed by QLQ-LC13) is defined as an increase in the score from baseline of ≥10. A clinically meaningful improvement in fatigue (as assessed by QLQ-C30) is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, change in symptoms/functioning from baseline will be categorized as improved, stable or worsening as shown in Table 5.

Table 5 Visit responses for symptoms and HRQoL

Score	Change from baseline	Visit response
QLQ-C30/QLQ-LC13 Symptom scales/items	≥+10	Worsened/Deterioration
	≤-10	Improved
	Otherwise	Stable/No change
QLQ-C30 functional scales and global health status/QoL	≥+10	Improved
	≤-10	Worsened/Deterioration
	Otherwise	Stable/No change

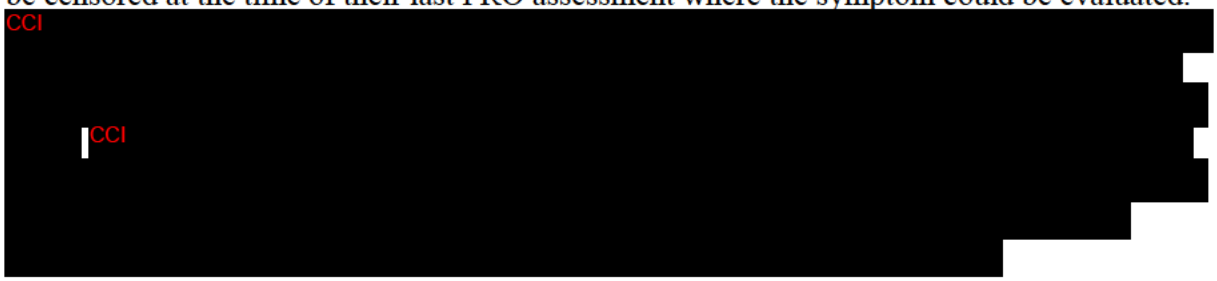
EORTC European Organisation for Research and Treatment of Cancer; QLQ C30 30 item core quality of life questionnaire; QLQ-LC13 13-item Lung Cancer Quality of Life Questionnaire; QoL Quality of life.

For the visit level summaries of Improvement/Deterioration/No change then all patients with a baseline and post-baseline score will be included, thus the denominator may differ from the time to deterioration and improvement rate endpoints derived below.

3.3.1.1 Time to symptom deterioration (QLQ-C30 and QLQ-LC13)

For each of the symptoms scales/items in the EORTC QLQ-C30 and QLQ-LC13, time to symptom deterioration will be defined as the time from randomization until the date of the first clinically meaningful symptom deterioration (an increase in the score from baseline of ≥10) that is confirmed at the next available subsequent assessment at least 14 days apart, or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the patient withdraws from IP or receives another anticancer therapy prior to symptom deterioration. Missed visits are allowed in between assessments confirming deterioration. This is considered a conservative approach whereby a deterioration is considered a 'negative' outcome and therefore should be assigned as such, regardless of missed visits. Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the symptom change could be evaluated. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Patients whose symptoms (as measured by EORTC QLQ-C30 and QLQ-LC13) have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the symptom could be evaluated.



See Appendix A for further details on the derivation of the confirmation of deterioration 2 missed visit rule for ePRO and length of 2 missed visit window.

If the patient has no visits or does not have baseline data they will be censored at day 1 unless they die within 2 visits of baseline CCI plus 3 days CCI

The population for the analysis of time to symptom deterioration will include a subset of the FAS who have baseline scores of CCI

3.3.1.2 Time to HRQoL/function deterioration (QLQ-C30)

For HRQoL and function (as measured by EORTC QLQ-C30), time to deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful deterioration (a decrease in the score from baseline of ≥10) that is confirmed at the next available subsequent assessment at least CCI or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient withdraws from IP or receives another anticancer therapy prior to HRQoL/function deterioration. Missed visits are allowed in between assessments confirming deterioration. This is considered a conservative approach whereby a deterioration is considered a 'negative' outcome and therefore should be assigned as such, regardless of missed visits. Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the HRQoL/function change could be evaluated. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Patients whose HRQoL or function have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the HRQoL/function could be evaluated, prior to the 2 missed visits. Also, if HRQoL deteriorates after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated. The 2 missed visit rule for ePRO will take a 'look-forward' approach i.e. if there are 2 consecutive missed visits at any time prior to the confirmed deterioration event, the event will be censored at the last available assessment prior to the 2 missed visits. Confirmation of deterioration will first be determined, then the censoring rules will be applied.

See Appendix A for further details on the derivation of the confirmation of deterioration, 2-missed visit rule and length of 2 missed visit window.

If the patient has no visits or does not have baseline data they will be censored at day 1 unless they die within 2 visits of baseline CCI plus CCI

The population for the analysis of time to QoL/function deterioration will include a subset of the FAS population who have baseline scores of ≥10.

3.3.1.3 Symptom improvement rate (QLQ-C30 and QLQ-LC13)

The symptom improvement rate will be defined as the number (%) of patients with a minimum of 2 consecutive assessments at least 14 days apart that show a clinically

meaningful improvement (a decrease in score of ≥10) in that symptom from baseline. Missed visits are not allowed in between consecutive assessments of improvement. This is considered a conservative approach whereby an improvement is considered a 'positive' outcome and therefore for any avoidance of doubt in the assessment of an improvement, missed visits are not allowed.

The denominator will consist of a subset of the FAS population who have a baseline symptom score of ≥10.

3.3.1.4 HRQoL/function improvement rate (QLQ-C30)

The HRQoL/function improvement rate will be defined as the number (%) of patients with a minimum of 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (an increase in score of ≥10) in that scale from baseline. Missed visits are not allowed in between consecutive assessments of improvement. This is considered a conservative approach whereby an improvement is considered a 'positive' outcome and therefore for any avoidance of doubt in the assessment of an improvement, missed visits are not allowed.

The denominator will consist of a subset of the FAS population who have a baseline QoL/function score of ≤90.

3.3.1.5 Change from baseline

Change from baseline in the pre-specified PRO symptom scores of dyspnea (LC13), cough (LC13), pain in chest (LC13), fatigue (C30), appetite loss (C30), physical function (C30) and global health status/QoL (C30) will be analyzed making use of all data from baseline.

3.3.1.6 Compliance rate

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30 and EORTC-QLQ-LC13 respectively. These will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least one individual item completed.

- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time up until PFS2 e.g., a questionnaire from a patient who has not withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. Date of study discontinuation and death date (whichever occurs earliest) will be mapped to the nearest visit date to define the number of expected forms.

- Evaluable questionnaire = a questionnaire with a completion date and at least one subscale that is non-missing.

- Overall PRO compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.

- Overall patient compliance rate is defined for each randomized treatment group as: Total number of patients with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above), divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

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

3.4 Safety

Safety and tolerability will be assessed in terms of adverse events (AEs) (including serious adverse events [SAEs]), deaths, laboratory data, vital signs, electrocardiograms (ECGs) and exposure. These will be collected for all patients.

'On treatment' will be defined as assessments between date of the first dose and CCI following last dose of randomized treatment (durvalumab/placebo) or up to and including the date of initiation of the first subsequent therapy (whichever occurs first).

The Safety analysis set will be used for reporting of safety data.

3.4.1 Adverse events

AEs and SAEs will be collected throughout the study, from date of informed consent until CC CCI after the last dose of randomized treatment. A treatment emergent adverse event (TEAE) is an AE with an onset date or a pre-existing AE worsening following the first dose of randomized treatment through to CCI after the last dose of randomized treatment (i.e., the last dose of durvalumab/placebo). The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03).
Adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

AEs of special interest and AEs of possible interest

Some clinical concepts (including some selected individual preferred terms [PTs] and higher level terms [HLTs]) have been considered "AEs of special interest" (AESI) to the durvalumab

program. AESIs represent pre-specified risks which are of importance to a clinical development program.

The AESIs reported in the AstraZeneca-sponsored durvalumab studies are defined as AEs with a likely inflammatory or immune-mediated pathophysiological basis resulting from the mechanism of action of durvalumab and requiring more frequent monitoring and/or interventions such as corticosteroids, immunosuppressants, and/or endocrine therapy. Endocrine therapies include standard endocrine supplementation, as well as treatment of symptoms resulting from endocrine disorders (for example, therapies for hyperthyroidism include beta blockers [e.g., propranolol], calcium channel blockers [e.g., verapamil, diltiazem], methimazole, propylthiouracil, and sodium perchlorate).

The Adverse Events of Possible Interest (AEPIs) reported in the AstraZeneca-sponsored durvalumab studies are defined as AEs that could have a potential inflammatory or immune-mediated pathophysiological basis resulting from the mechanism of action of durvalumab but are more likely to have occurred due to other pathophysiological mechanisms, thus, the likelihood of the event being inflammatory or immune-mediated in nature is not high and/or is most often or usually explained by the other causes. These AESIs and AEPIs are identified as a list of categories provided by the clinical 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 preferred terms contribute to each AESI/AEPI. A further review will take place prior to DBL to ensure any further terms not already included are captured within the categories.

3.4.2 Laboratory data

Laboratory data will be collected throughout the study, from screening to follow-up visit as described in the CSP. Blood and urine samples for determination of hematology, clinical chemistry, and urinalysis will be collected as described in Section 8.2.1 of the CSP. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 4.1.3 below will be used.

Change from baseline in hematology and clinical chemistry variables will be calculated for each post-dose visit on treatment. CTCAE grades will be defined at each visit according to the CTCAE grade criteria using local or project ranges as required, after conversion of lab result to corresponding preferred units. The following parameters have CTCAE grades defined for both high and low values: Potassium, Sodium, Magnesium, Glucose and Corrected calcium.

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

Corrected calcium (mmol/L) = Total Calcium (mmol/L) + ([40 – Albumin (g/L)] x 0.02)

Creatinine clearance will be derived according to the Cockroft-Gault formula (Cockcroft and Gault 1976).
Males:

$$\text{Creatinine CL} = \frac{\text{Weight (kg) x (140 - Age)}}{\text{72 x serum creatinine (mg/dL)}}$$

Females:

$$\text{Creatinine CL} = \frac{\text{Weight (kg) x (140 - Age)}}{\text{72 x serum creatinine (mg/dL)}} \text{ x } 0.85$$

Absolute values will be compared to the project reference range and classified as low (below range), normal (within range or on limits of range) and high (above range).

The maximum or minimum on-treatment value (depending on the direction of an adverse effect) will be defined for each laboratory parameter as the maximum (or minimum) post-baseline value at any time.

Local reference ranges will be used for the primary interpretation of laboratory data at the local laboratory. Project reference ranges will be used throughout for reporting purposes. The denominator used in laboratory summaries of CTCAE grades will only include evaluable patients, in other words those who had sufficient data to have the possibility of an abnormality.

For example:

- If a CTCAE criterion involves a change from baseline, evaluable patients would have both a baseline and at least 1 post-dose value recorded.
- If a CTCAE criterion does not consider changes from baseline, to be evaluable the patient need only have 1 post-baseline value recorded.

3.4.3 ECGs

Resting 12-lead electrocardiograms (ECGs) are recorded at screening and as clinically indicated thereafter. Categorical summaries of change from baseline in overall ECG assessments (recorded as “abnormal” and “normal”) will be created if a sufficient number of ECG assessments are recorded. The decision to create summaries of ECG data will be made at the blind data review meeting (BDRM).

3.4.4 Vital signs

Vital signs data obtained up until the 30 days from date of last dose of randomized treatment will be used for reporting. Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 4.1.3 below will be used. The denominator in vital signs data should include only those patients with recorded data.

3.4.5 Study treatments

Study treatment in this study refers to chemotherapy and radiation (CRT), durvalumab, and placebo. See Table 6 for further details on the IPs. Exposure will be defined for durvalumab or placebo and the SoC CRTs outlined in Table 6.

Table 6 Study Treatments

	Study treatment name	Route of administration	Dosing instructions
Durvalumab	Durvalumab (MEDI4736)	IV	CCI mg IV CCI [b]
Placebo	Saline solution	IV	Saline volume matching durvalumab volume
Standard of care	Cisplatin/ Etoposide	IV	Cisplatin CCI on Days 1 and 8 Etoposide CCI on Days 1–5 CCI days × 2 cycles (+1 additional induction cycle optional) Concurrent thoracic radiotherapy
	Carboplatin/ Paclitaxel	IV	Carboplatin AUC 2 and Paclitaxel CCI mg/m^2 on Day CCI basis for CCI - Concurrent thoracic radiotherapy **Optional**: paclitaxel CCI and carboplatin CCI given either as 1 induction cycle prior to initiation of radiotherapy **OR** as 1-2 consolidation cycles after radiotherapy is completed
	Pemetrexed/ Cisplatin [a]	IV	Pemetrexed CCI and cisplatin CCI on (+1 additional induction cycle optional)
	Pemetrexed/ Carboplatin [a]	IV	Pemetrexed CCI and carboplatin CCI on
	Radiation	External beam radiation	5 fractions/ week for CCI (±3 days) (Total CCI

Note: Cycles of durvalumab/placebo are cycles of chemotherapy are per local prescribing guidelines.

[a] Patients with CCI NSCLC only. Administer vitamin B12 and folic acid as per pemetrexed prescribing instructions.

[b] If a patient's weight falls to 30 kg or below, the patient should receive CCI dosing equivalent to CCI of durvalumab CCI until the weight improves to >30 kg, at which point the patient should start receiving the fixed dosing of durvalumab CCI mg CCI

AUC Area under the curve; IV Intravenous; CCI Every CCI CCI Every 28; SoC Standard of care.

3.4.5.1 Treatment exposure for durvalumab or placebo

Exposure will be defined as follows:

Total (or intended) exposure of durvalumab (MEDI4736) or placebo

- Total (or intended) exposure = min (last dose date where dose > 0 mg + 27 days, date of death, date of DCO) – first dose date + 1

Actual exposure of durvalumab (MEDI4736) or placebo

- Actual exposure = intended exposure – total duration of dose delays, where intended exposure will be calculated as above.

Calculation of duration of dose delays (for actual exposure):

- Since patients will receive CCI mg via IV infusion CCI the duration of dose delays will be calculated as follows:

  For all dosing dates:

  Total duration of dose delays= Sum of (Date of the dose - Date of previous dose – 28 days)

  Thus, if no delays were encountered, the duration would sum up to 0, since infusions were done every four weeks.

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

3.4.5.2 Treatment exposure for SoCs

Exposure will be defined as follows:

Cisplatin and Etoposide

Total (or intended) exposure of cisplatin

- Minimum of (last infusion/dose date of the last cycle + 6 days (if last infusion/dose date was Day 1 of cycle) or last infusion/dose date of the last cycle + 20 days (if last infusion/dose date was Day 8 of cycle), date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of etoposide

- Minimum of (last infusion/dose date of the last cycle + 23 days (if last infusion/dose date was Day 5 of cycle) or last infusion/dose date of the last cycle (if last

infusion/dose date was Day 1-4 of cycle), date of death, date of DCO) – first infusion/dose date of first cycle + 1

Carboplatin and Paclitaxel

Total (or intended) exposure of carboplatin

- Minimum of (infusion/dose date of the last cycle + 6 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of paclitaxel

- Minimum of (infusion/dose date of the last cycle + 6 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of carboplatin and paclitaxel with optional induction cycle

- Minimum of (last infusion/dose date of the last cycle + 20 days (if last infusion/dose date was Day 1 of induction cycle) or last infusion/dose date of the last cycle + 6 days (if last infusion/dose date was after induction cycle), date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of carboplatin and paclitaxel with optional consolidation cycle

- Minimum of (last infusion/dose date of the last cycle + 6 days (if last infusion/dose date was not from consolidation cycle) or last infusion/dose date of the last cycle + 20 days (if last infusion/dose date was from consolidation cycle), date of death, date of DCO) – first infusion/dose date of first cycle + 1

Pemetrexed and Cisplatin

Total (or intended) exposure of pemetrexed

- Minimum of (infusion/dose date of the last cycle + 20 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of cisplatin

- Minimum of (infusion/dose date of the last cycle + 20 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Pemetrexed and Carboplatin

Total (or intended) exposure of pemetrexed

- Minimum of (infusion/dose date of the last cycle + 20 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Total (or intended) exposure of carboplatin

- Minimum of (infusion/dose date of the last cycle + 20 days, date of death, date of DCO) – first infusion/dose date of first cycle + 1

Radiation

Total Dose of Radiotherapy in Grays will be calculated Fraction Dose multiplied by Number of Fraction Doses.

Total (or intended) exposure

- Minimum of (radiotherapy stop date, date of death, date of DCO) – radiotherapy start date + 1

Patients who permanently discontinue during a dose delay

If a decision is made to permanently discontinue study treatment in-between cycles or during a dose delay, then the date of last administration of study medication recorded will be used in the programming.

3.4.6 Dose intensity

Dose intensity will be derived for study treatment durvalumab or placebo. Relative dose intensity (RDI) is the percentage of the actual dose intensity delivered relative to the intended dose intensity through to treatment discontinuation.

RDI will be defined as follows:

- RDI = 100% * d/D, where d is the actual cumulative dose delivered up 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. When accounting for the calculation of intended cumulative dose 3 days should be added to the date of last dose to reflect the protocol allowed window for dosing.

When deriving actual dose administered the volume before and after infusion will also be considered.

Example of dose intensity for durvalumab (MEDI4736)

Table 7 Dose Intensity scenarios

		Study Day									
RDI	Patient	1	29	57	85	113	141	169	197	225	
100%	1	X	X	X	X	X	X	X	X	X	PD
100%	2	X	X	X	X	X	X	X	X[D]		PD
56%	3	X	X			X	O	X		X	PD

X: Dose of CCI taken; O: Dose missed; [D]: Dose discontinued; PD: Progressive Disease

Patients 1-3 progressed on Day 230, so the intended dose through to progression was CCI of durvalumab = CCI

Patient 1 received a total of 13.5g of durvalumab, whereas other patients received less treatment due to:

- Early stopping prior to PD (Patient 2)
- Dosing delays (Patient 3)

The Patient 2 example illustrates that for RDI, the end of actual dosing period is calculated based on the smallest recovery period after the last non-zero dose.

Patient 1: RDI = (9 * 1.5g)/ CCI = 100%

Patient 2: RDI = (8 * 1.5g)/ CCI = 100%

Patient 3: RDI = (5 * 1.5g)/ CCI = 56%

Analyses to evaluate the of durvalumab will be performed by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

3.5.1 CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

CCI CCI CCI

[redacted]

3.6 Exploratory variables

3.6.1 Calculation or derivation of patient-reported outcome variables–PGIS

Patients' overall impression of the severity of their cancer symptoms will be assessed using the Patient Global Impression of Severity (PGIS) questionnaire. The lung cancer overall status can be rated as (1) no symptoms, (2) very mild, (3) mild, (4) moderate, (5) severe and (6) very severe. The data will be presented using summaries and descriptive statistics.

3.6.2 Calculation or derivation of patient-reported health state utility (EQ-5D-5L)

The health state utility will be assessed using the EQ-5D-5L (exploratory). The EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal. Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care.

The index comprises of 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and extreme problems). A unique EQ-5D health state is referred to by a 5-digit code, allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses). Where value sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied (Oemar and Janseen 2013). In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analog scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

4. ANALYSIS METHODS

There will be 1 treatment comparison of interest:

- Durvalumab (MEDI4736) CCI + SoC CRT vs placebo + SoC CRT

and the formal statistical analyses will be performed to test the main hypothesis:

- H0: No difference between durvalumab + SoC CRT and placebo + SoC CRT
- H1: Difference between durvalumab + SoC CRT and placebo + SoC CRT

The primary objective of this study is to assess the efficacy of durvalumab + SoC CRT compared with placebo + SoC CRT in terms of PFS (per RECIST 1.1 as assessed by BICR).

The key secondary endpoints (i.e., those included in the multiple testing procedure) are ORR (per RECIST 1.1 as assessed by BICR), OS and OS24. The study is powered for the primary endpoint (PFS) and if PFS is statistically significant then the multiple testing procedure allows for statistical testing of the key secondary endpoints of ORR, OS and OS24.

Results of all statistical analysis will be presented using a CCI CI and 2-sided p-value, unless otherwise stated.

There will be . Based on review of blinded OS data, at the time of is now expected to reach the CCI IA target number CCI rather than the Hence, only C will now actually be performed using the CCI target number of

1. CCI will occur when approximately events have occurred CCI of the target number of PFS events). It is estimated that this DCO will occur CC months after the last patient has been randomized.
2. CCI occur when approximately CCI events have occurred CCI maturity). It is estimated that this DCO will occur CCI after the last patient has been randomized. Approximately CCI events CCI maturity) are expected to have occurred at this DCO.
3. CCI occur when approximately CCI events CCI maturity) have occurred.

4.1 General principles

Efficacy and PRO data will be summarized and analyzed on the FAS analysis set. Safety and treatment exposure data will be summarized based upon the safety analysis set. Study population and demography data will be summarized based upon the FAS analysis set.

Data will be presented in data listings by treatment group and subject number. All summaries will be presented by treatment group, unless otherwise specified.

A month is operationally defined to be 30.4375 days. Six months is operationally defined to be 183 days.

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 summarized 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 summarized by frequency counts and percentages for each category.

- Unless otherwise stated, percentages will be calculated out of the population total and for each treatment group.

- 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.

- P-values will be rounded to 3 decimal places. P-values less than 0.0005 (e.g., 0.0002) will not be rounded to 3 decimal places (e.g., 0.000) but instead be displayed as <0.001. P-values output as <0.0001 by statistical software will not be rounded and displayed in the same way ('<0.0001').

- SAS® version 9.4 will be used for all analyses.

4.1.1 Baseline

In general, for efficacy and PRO endpoints the last observed measurement prior to randomization will be considered the baseline measurement. However, for PRO endpoints if an evaluable assessment is only available after randomization but before the first dose of randomized treatment then this assessment will be used as baseline. For safety endpoints the last observation before the first dose of randomized treatment will be considered the baseline measurement unless otherwise specified. For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose.

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

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

4.1.2 Methods for multiplicity control

The multiple testing procedure (Figure 2) will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoints of PFS and the key secondary endpoints, [redacted]

[redacted]

analysis, an alpha level of CCI will be allocated to the OS analysis and an alpha level [redacted]

Formal statistical testing of ORR will occur when PFS is statistically significant. In order to make an allowance for the CCI where this could be analyzed [redacted] a Haybittle Peto spending function will be applied with the assumption of a CCI Information Fraction at the first potential timepoint (as most responses are expected to occur early in the study). If additional alpha is recycled to ORR at a later timepoint, the formal testing of ORR will be based on the data at the time when PFS is statistically significant.

Formal statistical testing of OS24 will occur after 24 month of LSI and when OS is statistically significant. In order to make an allowance for the 3 possible datacuts where this could be analyzed an alpha adjustment will be applied by using the exact same alpha that is being used at the timepoint for OS (O'Brien Fleming spending function).

For the PFS endpoint, CCI and the alpha level will be controlled at the timepoints by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach.

For the OS endpoint, CCI and the alpha level will be controlled at the timepoints by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach. The O'Brien Fleming boundaries for the will be adjusted depending on the alpha used for the endpoint.

Figure 2 Multiple testing procedures for controlling the type 1 error rate

For PRO symptoms, the overall type I error (5% 2 sided) will be controlled across the five primary PRO measures of cough, dyspnea and chest pain as assessed by the EORTC QLQ-LC13 and fatigue and appetite loss as assessed by the EORTC QLQ-C30 using the Bonferroni procedure. These are not part of the main multiple testing procedure. CCI

4.1.3 Visit window for safety and PRO assessments

Time windows will need defining for any presentations that summarize values by visit. The following conventions should also apply:

- The time windows should be exhaustive so that data recorded at any time point has the potential to be summarized. Inclusion within the time window should be based on the actual date and not the intended date of the visit.
- All unscheduled visit data should 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 half way 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 29, visit window 2 – 42
- Day 57, visit window 43 – 70
- Day 85, visit window 71 – 98
- Day 113, visit window 99 – 126

- 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 should be used, or the earlier in the event the values are equidistant from the nominal visit date. If there are two values recorded on the same day and the parameter is CTCAE gradable then the record with the highest toxicity grade should be used. Alternatively, if there are two records recorded on the same day and the toxicity grade is the same (or is not calculated for the parameter) then the average of the two records should be used. The listings should highlight the value for that patient that went into the summary table, wherever feasible. Note: in summaries of extreme values all on-treatment values collected are used including those collected at unscheduled visits.
  - To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group visit data should only be summarized:
    - (safety data except ECG) if the number of observations is greater than the minimum of 20 and > 1/3 of patients dosed;
    - (PRO) if the number of observations is greater than 20.
- For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.
- Baseline will be defined as the last non-missing measurement prior to the first dose of randomized treatment. For laboratory data, any assessments made on day 1 will be considered pre-dose. If there are two visits 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) with assessment time missing, 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 summarized over time, study day will be calculated in relation to date of first treatment.

4.1.4 Imputation rules

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. Missing CTCAE grades will not be imputed.

Patients with a partial date of birth (i.e., for those countries where year of birth only is given) will have an assumed date of birth of 1st Jan [given year]) for calculation of age at randomization.

Partial dates for the following modules will be imputed: prior cancer therapy, previous radiotherapy and prior and concomitant medications and adverse events, as well as for other modules, where required.

The original incomplete or missing dates will be presented in the listings.

<u>Adverse events and medications</u>

Adverse events: all AEs will be considered as treatment-emergent unless the opposite can be clearly stated. Imputation will be done only in the context of identifying TEAEs.

Medications/therapies: all medications will be considered as concomitant unless the opposite can be clearly stated.

In practice, for adverse events and medications, original incomplete or missing start dates will be imputed as below:

- Missing day: impute the 1st of the month unless month is same as month of first dose of study drug then impute first dose date;
- Missing day and month: impute 1st January unless year is the same as first dose date then impute first dose date;
- 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 i.e., is prior to the end date of the AE or medication.

Original incomplete or missing stop dates for adverse events and medications/therapies will be imputed as below:

- Missing day: impute the last day of the month unless month is same as month of last dose of study drug then impute last dose date;
- Missing day and month: impute 31st December unless year is the same as last dose date then impute last dose date;
- Completely missing:
    - AE: since there is no ongoing flag recorded in eCRF, then assume that AE is still present (i.e., do not impute a date);
    - Medication: if the ongoing flag is missing then assume that medication is still being taken (i.e., do not impute a date). If the medication has stopped and start date of medication is prior to first dose date then impute the first dose date, if the medication started on or after first dose date then impute a date that is after the last dose date.

When imputing a stop date, ensure that the new imputed date is sensible i.e., is after the start date of the AE or medication.

Duration of AE/medications will not be derived using imputed dates.

Imputation of partial death dates

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

- For missing day only – using the 1st of the month
- For missing day and month – using the 1st of January

Imputation of partial start dates of subsequent anti-cancer therapy

- For missing day only – using the 1st of the month
- For missing day and month – using the 1st of January

4.2 Analysis methods

Efficacy data will be summarized and analyzed using the FAS.

The following table (Table 8) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint.

Table 8 Formal statistical analyses to be conducted and pre-planned sensitivity analyses

Endpoints analyzed	Notes
Progression-free survival	Primary analysis using stratified log-rank test using BICR assessments (RECIST 1.1) Sensitivity analyses using BICR assessments (RECIST 1.1) – Interval censored analysis – evaluation time bias – Analysis using alternative censoring rules – attrition bias [redacted] stratified log-rank test using site Investigator assessments (RECIST 1.1) – ascertainment bias Additional analysis using Cox proportional hazards models to determine the effect of covariates on the HR estimate Additional analysis using Cox proportional hazards models to determine the consistency of treatment effect between stratification factors via the approach of Gail and Simon 1985. Subgroup analysis using Cox proportional hazards model
Objective response rate	Primary analysis using CMH test, stratified by age [redacted] and stage [redacted] Sensitivity analysis using a CMH test repeated using the site Investigator data based on RECIST 1.1 As a sensitivity analysis the ORR analyzed using logistic regression adjusting for the same factors as for PFS
Overall survival	Stratified log-rank test Sensitivity analysis using a Kaplan-Meier plot of time to censoring where the censoring indicator of the primary analysis is reversed – attrition bias Additional analysis using Cox proportional hazards models to determine the effect of covariates on the HR estimate Additional analysis using Cox proportional hazards models to determine the consistency of treatment effect between stratification factors via the approach of Gail and Simon 1985. Subgroup analysis using Cox proportional hazards model
Proportion of patients alive at 24 months	Kaplan-Meier estimates of survival at 24 months and p-value (following the method described by Klein et al 2007)
Rate of CR	Analysis using CMH test using BICR assessment RECIST 1.1 data Sensitivity analysis using the CMH test using site Investigator tumor data (RECIST 1.1)
Duration of response	Kaplan-Meier estimates

Endpoints analyzed	Notes
Disease control rate	Summarized by descriptive statistics using BICR assessment RECIST 1.1 data and site Investigator tumor data (RECIST 1.1)
Time to death or distant metastasis	Stratified log-rank test using BICR tumor data (RECIST 1.1)
[redacted]	Stratified log-rank test
Time to first subsequent therapy [a]	Stratified log-rank test
Time to second subsequent therapy [a]	Stratified log-rank test
Change from baseline in key symptoms (EORTC QLQ-C30 and QLQ-LC13)	Mixed model repeated measures analysis
HRQoL/Function improvement rate (EORTC QLQ-C30 endpoints)	Logistic regression
Symptom improvement rate (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Logistic regression
Time to HRQoL/Function deterioration (EORTC QLQ-C30 endpoints)	Stratified log-rank test
Time to symptom deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Stratified log-rank test

[a] Endpoints to be analyzed at the time of overall survival analysis
BICR Blinded Independent Central Review; CMH Cochran-Mantel-Haenszel; EORTC European Organisation for Research and Treatment of Cancer; HRQoL Health-related quality of life; ORR Objective response rate; PFS Progression-free survival; [redacted] QLQ-C30 v3 30-item Core Quality of Life Questionnaire, version 3; QLQ-LC13 13-item Lung Cancer Quality of Life Questionnaire; RECIST 1.1 Response Evaluation Criteria in Solid Tumors, version 1.1.

4.2.1 Primary efficacy endpoints

4.2.1.1 Progression-free survival

PFS based on the BICR data will be analyzed using a stratified log-rank test adjusting for age [redacted] and stage [redacted] for generation of the p-value. The HR and its CI will be estimated from a stratified Cox proportional hazards model (Cox, 1972) (with ties = Efron and the stratification variables included in the strata statement) and the CI calculated using a profile likelihood approach.

The effect of treatment will be estimated by the hazard ratio (HR) together with its corresponding CCI % CI and p-value for the ITT population.

The covariates in the statistical modelling will be based on the values entered into IVRS at randomization, even if it is subsequently discovered that these values were incorrect.

Supportive summaries/graphs

A Kaplan-Meier (KM) plot of PFS will be presented by treatment group. The total number of events, median PFS (calculated from the Kaplan-Meier plot, with CCI CIs) and the PFS rate at CCI months will be summarized by treatment arm.

The assumption of proportionality will be assessed. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) vs log (time) and, if these raise concerns, by fitting a time-dependent covariate to assess the extent to which this represents random variation. If a lack of proportionality is evident, the variation in treatment effect can be described by presenting piecewise HR calculated over distinct time-periods. In such circumstances, the HR can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. If lack of proportionality is found, this may be a result of treatment-by-covariate interactions, which will be investigated.

The treatment status at progression of patients at the time of analysis will be summarized. 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.

The number of patients prematurely censored will be summarized by treatment arm. A patient would be defined as prematurely censored if they had not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than one scheduled tumor assessment interval plus CCI if time period between randomization and DCO for that patient is CCI or less; CCI otherwise) prior to the DCO date.

Additionally, summary statistics will be given for the number of days from censoring to DCO for all censored patients.

A summary of the duration of follow-up will be summarized using median time from randomization to date of censoring (date last known to be non-progressor) in censored (not progressed) patients only, presented by treatment group.

Additionally, summary statistics for the number of weeks between the time of progression and the last RECIST assessment prior to progression will be presented for each treatment group.

Summaries of the number and percentage of patients who miss two or more consecutive RECIST assessments will be presented for each treatment group.

All of the collected RECIST 1.1 data will be listed for all randomized patients. In addition, a summary of new lesions (i.e., sites of new lesions) will be produced.

Sensitivity analyses

The following sensitivity analyses will be performed:

1. Evaluation-time bias

A sensitivity analysis will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time points. The midpoint between the time of progression and the previous RECIST assessment (using the final date of the assessment) will be analyzed using a stratified log-rank test, as described for the primary analysis of PFS. Note that midpoint values resulting in non-integer values should be rounded down. For patients whose death was treated as a PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust to even highly asymmetric assessment schedules (Sun and Chen 2010). To support this analysis, the mean of patient-level average inter-assessment times will be tabulated for each treatment. This approach will use the BICR data.

2. Attrition bias

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following two, or more, non-missing tumor assessments will be included. In addition, and within the same sensitivity analysis, patients who take subsequent therapy (note that for this analysis radiotherapy is considered a subsequent anti-cancer therapy) prior to their last RECIST assessment or progression or death will be censored at their last assessment prior to taking the subsequent therapy. This analysis will be supported by a Kaplan-Meier plot of the time to censoring using the PFS data from the primary analysis and where the censoring indicator of the PFS analysis is reversed. This approach will use the BICR data.

3. Ascertainment bias

Ascertainment bias will be assessed by analyzing site Investigator data. The stratified log-rank test will be repeated on the programmatically derived PFS using the site Investigator data based upon RECIST. The HR and CI will be presented.

If there is an important discrepancy between the primary analysis using the BICR and this sensitivity analysis using site Investigator data assessments, then the proportion of patients with site but no central progression will be summarized; such patients have the potential to introduce bias in the central review due to informative censoring. An approach that imputes an event at the next visit in the central review analysis may help inform the most likely HR value (Fleischer et al 2011), but only if an important discrepancy exists.

Disagreements between investigator and central reviews of RECIST progression will be presented for each treatment group. The summary will include the early discrepancy rate which is the frequency of investigator review declared progressions before the central review as a proportion of all investigator review progressions and the late discrepancy rate which is the frequency of investigator review declared progressions after the central review as a proportion of all discrepancies.

A forest plot illustrating the hazard ratio and 95% confidence interval will be provided to compare the primary and sensitivity analyses of progression free survival.

Subgroup analysis

Subgroup analyses will be conducted comparing PFS between treatments in the following subgroups of the FAS (but not limited to):

- Planned chemotherapy treatment regimen (identified prior to randomization)
  - Carboplatin vs Cisplatin
- Planned radiation therapy (identified prior to randomization)
  - Intensity-modulated radiation therapy vs 3-dimensional conformal radiation therapy)

- Region
  - Asia vs. Europe vs. South America vs. North America.
    This will be determined from the center number (CENTRE). If less than 20 patients will be randomized within North America, then this region will be pooled together with South America.

- Race/ethnicity
  - White, Black/African-American, Asian, Other [Native Hawaiian/Pacific Islander or American Indian/Alaska Native or Others]).
    This will be determined from the response to “Race” (DEM module) on the eCRF at screening

- Sex
  - Female vs. Male

- Age at randomization
  - CCI [redacted]

- Smoking status
  - Smoker vs. Non-smoker
    Patient is categorized as smoker if there exists a record in SU_NIC with any of the following options for “What type of substance was used?” 'Cigarettes', 'Cigarillos', 'Cigars', 'Pipe Tobacco', 'Tobacco for Smoking'. Else patient is a non-smoker. Non-smoker is a patient that has never smoked.

- Stage of disease at study entry
  - CCI CCI CCI
- Histology
  - CCI CCI
- WHO performance status at baseline
  - Normal activity [PSTAT=0] vs. restricted activity [PSTAT=1]
    Patients with a missing performance status will be included in the 'restricted activity' category.
- CCI

- EGFR mutation
  - Positive vs. Negative vs. unknown
    Receptor status values of "mutation not detected" or "negative" will be considered "negative". Receptor status values of "mutation detected" or "positive" will be considered "positive".
- Planning target volume (PTV)
  - $\geq 450cm^3$ vs. $<450cm^3$

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

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

No adjustment to the significance level for testing will be made since all these subgroup analyses will be considered exploratory and may only be supportive of the primary analysis of PFS. For each subgroup, the HR (for the treatment comparisons of interest) and CCI CI will be calculated from a Cox proportional hazards model that contains treatment. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties,

using the by statement to obtain HR and CCI CI for each subgroup level separately. These will be presented on a forest plot including the HR and CCI CI, along with the results of the overall primary analysis.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 10 events (5 events in each group/arm)), the relationship between that subgroup and PFS will not be formally analyzed. In this case, only descriptive summaries will be provided.

Unless there is a marked difference between the results of the statistical analyses of the PFS from the BICR tumor data and that of the site Investigator, the subgroup analyses will only be performed upon the PFS endpoint using BICR data.

Effect of covariates on the HR estimate

Cox proportional hazards modelling will be employed to assess the effect of pre-specified covariates on the HR estimate. The result from the model in the primary analysis and the model containing additional covariates will be presented.

Additional covariates for this model will be sex, smoking status, histology, region and race.

The model will include the effect regardless of whether the inclusion of effect significantly improves the fit of the model provided there is enough data to make them meaningful.

Consistency of treatment effect between stratification factors

Interactions between treatment and stratification factors will also be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

Time to first subsequent therapy or death (TFST)

The time to the start of subsequent therapy will be analyzed using the same methodology and model as that used for the primary analysis of PFS. The HR for the treatment effect together with its CCI CI will be presented. In addition, medians and a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment arm and the time between progression and starting subsequent therapy will be assessed based upon the investigator data defined date of progression. This will be summarized per treatment arm, but no formal comparisons will be made. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

In patients who received a subsequent cancer therapy, a summary table of first subsequent cancer therapies by treatment arm will be provided.

4.2.2 Secondary efficacy endpoints

4.2.2.1 Objective response rate

The ORR will be based on BICR data. The ORR will be compared between treatment arms using a Cochran-Mantel-Haenszel (CMH) test, stratified by age [redacted] and stage ([redacted]).

The results of the analysis will be presented in terms of a difference in proportions together with the (1 – allocated alpha)% CI and p-value. The confidence intervals for the difference in proportions between groups will be estimated using Miettinen and Nurminen's (MN) confidence limits.

The covariates in the statistical modelling will be based on the values entered into IVRS at randomization, even if it is subsequently discovered that these values were incorrect.

As a sensitivity analysis, the ORR will be analyzed using logistic regression adjusting for the same factors as for PFS. The results of the analysis will be presented in terms of an odds ratio together with its associated profile likelihood [redacted] CI (e.g., using the option 'LRCI' in SAS procedure GENMOD) and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model).

The analysis of ORR using a CMH test will be repeated using the site investigator data based on RECIST 1.1 as a sensitivity analysis to confirm the results of the primary analysis.

Summaries will be produced that present the number and percentage of patients with a tumor response (CR/PR). Overall visit response data will be listed for all patients (i.e., the FAS). For each treatment arm, the BoR will be summarized on both BICR and site investigator data by n (%) for each category (CR, PR, NED, SD, PD, and NE). No formal statistical analyses are planned for BoR.

Subgroup analysis

Subgroup analyses will be conducted comparing the difference in proportions between treatment arms in the same subgroups as specified for the PFS subgroup analysis with treatment as the only factor in the model.

For each subgroup, the difference in proportions between treatment groups and [redacted] CI will be calculated. These will be presented on a forest plot including the difference in proportions and [redacted] CI.

Consistency of treatment effect between stratification factors

Interactions between treatment and stratification factor will also be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

4.2.2.2 Overall survival

OS will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. The effect of treatment will be estimated by the HR together with its corresponding (1 – allocated adjusted alpha) % CI. Kaplan Meier plots of OS will be presented by treatment arm.

Summaries of the number and percentage of patients who have died, those still in survival follow-up, those lost to follow-up, and those who have withdrawn consent will be provided along with the median OS for each treatment.

Sensitivity analyses

A sensitivity analysis for OS will examine the censoring patterns to rule out attrition bias, achieved by a Kaplan-Meier plot of time to censoring where the censoring indicator of OS is reversed.

The number of patients prematurely censored will be summarized by treatment arm. A patient would be defined as prematurely censored if their survival status was not defined at the DCO.

In addition, duration of follow-up will be summarized using medians:

- In censored patients who are alive at DCO only: Time from randomization to date of censoring (date last known to be alive) for each arm.

- In all patients: Time from randomization to the date of death or to the date of censoring for censored patients, regardless of treatment arm.

Subgroup analyses maybe performed if there are a sufficient number of OS events (≥10 events [≥5 events in each group/arm]).

Impact of switching (crossover outside of this study) to immunotherapies (or other potentially active investigational agents) on OS analyses

Exploratory analyses of overall survival adjusting for the impact of treatment switching will be performed to inform decision-makers, including payers. Methods such as the rank preserving structural failure time (RPSFT), inverse probability of censoring weighting (IPCW), and two steps methods (Latimer 2014) will be explored. The final choice of methods will be based on numerous factors including, but not limited to, the completeness of data, the degree of treatment switching, maturity of data, whether switching occurs very early or later in the trial, and the plausibility of the underlying assumptions including the constant treatment effect for the rank preserving method and the exchangeability assumption for IPCW and two step methods. If the described methods are deemed to be insufficient to describe the disease and treatment process, other methods may need to be explored. Further detail will be provided in the Payer Analysis Plan.

Subsequent therapies received after discontinuation of treatment will be summarized and listed by treatment group. Patients who subsequently received an immunotherapy agent or entered an immunotherapy trial will be summarized and listed by treatment group according to line of subsequent therapy, i.e., immediately after immunotherapy or as a later line.

4.2.2.3 Proportion of patients alive at 24 months

The proportion of patients alive at 24 months (i.e., OS24) will be summarized (using the Kaplan-Meier curve) and presented by treatment arm. For each treatment arm, the survival rate at 24 months based on Kaplan-Meier method will be presented, along with its (1 – allocated alpha)% CI. The computation of the CI will be based on a log-log transformation.

For the comparison between treatments, the test will be based on the method described in Klein et al 2007. The test statistic and its variance estimate are as follows:

- $\text{test statistic} = \ln \frac{\ln \hat{S}_1(t)}{\ln \hat{S}_2(t)}$

- $\text{variance estimate} = \frac{\hat{\sigma}_1(t)^2}{\ln^2 S_1(t)} + \frac{\hat{\sigma}_2(t)^2}{\ln^2 S_2(t)}$

where $\hat{\sigma}_i(t)^2 = \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$ is the variance for ln S(t) derived from Greenwood's formula for the variance of $S(t)$ and can be estimated from standard software packages, where d_i and n_i refer to the number of deaths and patients at risk for each risk set.

- The z-statistic is then calculated as: $\frac{\text{test statistic}}{\sqrt{\text{variance estimate}}}$

For the stratified analysis, the test statistic and its variance estimate in each strata will be estimated and combined by weighting inversely proportionately according to each within stratum variance (Whitehead and Whitehead 1991). A Z-test will be performed, and the p-value from the test will be presented.

The proportion of patients alive at 12 months (OS12) will be defined as the Kaplan-Meier estimate of OS at 12 months. For each treatment arm, the survival rate at 12 will be presented, along with its 95% confidence interval.

4.2.2.4 Rate of complete response

Rate of CR will be analyzed in the same way as for the ORR analysis without formal testing. Rate of CR will be assessed using investigator and BICR data.

4.2.2.5 Duration of response (DoR)

Kaplan-Meier estimates will be provided for the DoR in confirmed responded patients (i.e., median DoR and CCI CIs) by treatment arm, including the associated Kaplan-Meier curves (without any formal comparison of treatment arms or p-value attached). Landmarks DoR rates at 6, 12, 18 months will also be summarized.

4.2.2.6 Disease control rate (DCR)

DCR will be summarized using descriptive statistics by treatment arm using investigator and BICR data.

4.2.2.7 Time to death or distant metastasis (TTDM)

TTDM will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. Kaplan-Meier plots, medians and landmarks such as 12, 18 and 24 months will be summarized to support the analysis. TTDM will be based on BICR data.

[redacted]

Time to second subsequent therapy or death (TSST)

For supportive purposes, the time to the start of second subsequent therapy will be analyzed using the same methodology and model as that used for the analysis of PFS2. The HR for the treatment effect together with its CCI CI will be presented. In addition, a Kaplan-Meier plot of the time to the start of second subsequent therapy will be presented by treatment arm. This will be summarized per treatment arm, but no formal comparisons will be made. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

4.2.3 Secondary patient-reported outcomes

The PRO endpoints identified as primary are:

- Dyspnea: multi-item scale based on 3 questions ("Were you short of breath when you rested; walked; climbed stairs" – QLQ-LC13),
- Cough: 1 item ("How much did you cough?" – QLQ-LC13),
- Chest pain: 1 item ("Have you had pain in your chest" – QLQ-LC13).

- Fatigue: multi-item based on 3 questions ("Did you need rest; Have you felt weak; Were you tired" – QLQ-C30)

- Appetite loss: 1 item ("Have you lacked appetite" – QLQ-C30)

The physical functioning and overall health status domains of the EORTC QLQ-C30 are furthermore pre-specified endpoints of interest.

The analysis will be performed using a linear mixed model for repeated measures (MMRM) analysis of change from baseline in the scores for each assessment timepoint (see next section for further details) and the Bonferroni procedure for adjusting the significance level will be used to aid interpretation. The 5 key endpoints will be tested at the 1% significance level adjusted according to the Bonferroni procedure, in order to control the overall type I error (5% 2 sided).

Mixed models repeated measures (MMRM) analysis

Change from baseline in the 5 pre-specified symptoms as well as physical functioning and overall health status will be analyzed using a mixed model for repeated measures (MMRM) analysis making use of all data from baseline up to 12 months. The analysis will be to compare the average treatment effect from the point of randomization until PD or 12 months (whichever is earlier) unless there is excessive missing data (defined as >75% missing data).

It is acknowledged that patients will discontinue treatment at different timepoints during the study and that this is an important time with regards to symptoms and HRQoL data collection. To account for this, and to include the discontinuation and follow up visits, a generic visit variable will be derived for each patient in order that the average treatment effect can be analyzed using the above method. Each visit will be assigned a sequential number. The time from randomization to each of these will be derived to select only those visits occurring within the first 12 months of randomization or until PD.

As an example, say a patient X attends the first 4 scheduled visits of a 2-weekly schedule and then discontinues treatment, whilst patient Y discontinues treatment after the first scheduled visit, the first 6 generic visits would be as follows:

Generic visit	Study day (week)	
	Patient X	Patient Y
Baseline	Baseline	Baseline
1	15	15
2	29	26 (discontinuation)

Generic visit	Study day (week)	
	Patient X	Patient Y
3	43	43
4	57	57
5	82 (discontinuation)	85
6	113	113

The MMRM model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, age (<65 vs ≥65 years) and stage CCI vs CCI as well as the continuous fixed covariate of baseline score and the baseline score-by-visit interaction. Restricted maximum likelihood (REML) estimation will be used. An overall adjusted mean estimate will be derived that will estimate the average treatment effect over visits giving each visit equal weight. For this overall treatment comparison, adjusted mean estimates per treatment group and corresponding CCI CIs will be presented along with an estimate of the treatment difference, CCI CI, 99% CI (for the 5 key symptoms only) and p-value.

An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be tried in order until convergence is reached: Toeplitz with heterogeneity, autoregressive with heterogeneity, Toeplitz, autoregressive and compound symmetry.

Multiple imputation techniques for missing values may be considered to explore the robustness of any treatment effect.

Time to deterioration

Time to symptom and function/HRQoL deterioration will be analyzed for each of the symptom scales/items, function scales, and global health status/QoL in EORTC QLQ-C30 and QLQ-LC13. This will be achieved by comparing between treatment arms using a stratified log-rank test as described for the primary analysis of PFS. The HR and CCI CI for each scale/item will be presented graphically on a forest plot.

For each of the symptom scales/items, functional scales, and global health status/QoL, time to deterioration will be presented using a Kaplan-Meier plot. Summaries of the number and percentage of patients experiencing a clinically meaningful deterioration or death and the median time to deterioration will also be provided for each treatment arm.

Symptom and function/HRQoL improvement rate

A summary of the symptom improvement rate for all symptom scales/items in EORTC QLQ-C30 and QLQ-LC13 will be produced. Similarly, a summary of function/HRQoL improvement rate for each of the 5 function scales (physical, role, emotional, cognitive, and social) and global health status/QoL will be produced.

Symptom and HRQoL/function improvement rates will be analyzed by comparing between treatment arms using a logistic regression model adjusting for the same factors as PFS. The odds ratio (an odds ratio greater than 1 will favor durvalumab) together with its associated profile likelihood CCI CI (e.g., using the option 'LRCI' in SAS procedure GENMOD) for each scale/item will be presented graphically on a forest plot. If there are very few responses in 1 treatment arm, a Fisher's exact test will be considered.

Change from baseline

Summaries of original and change from baseline values of each symptom scale/item, the global HRQoL score, and each functional domain will be reported by assessment timepoint for each treatment arm. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each assessment timepoint for each ordinal item (in terms of the proportion of patients in the categories of improvement, stable, and deterioration as defined in Table 5) will also be produced for each treatment arm.

EuroQol-5-Dimension 5-Level questionnaire (EQ-5D-5L)

Descriptive statistics will be reported for the health state domain (e.g., proportion in each domain) and the visual analog scale by visit, as well as the change in the visual analog scale value and the derived utility index value from baseline. To support future economic evaluations, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment and pre- and post-progression.

Patients' Global Impression of Severity (PGIS)

Summaries and descriptive statistics of PGIS, considered as categorical data, will be presented by assessment timepoint and overall for each treatment arm, based on the FAS.

4.2.3.1 Health care resource use

The potential impact the disease and treatment has on health care resource use will be analyzed for the purposes of submissions to payers. Descriptive statistics (as appropriate, including means, median, ranges or frequencies and percentages) will be provided for each arm on the different types of hospital admissions, the length of stay of people admitted in to hospital for at least one overnight stay and length of stay of people admitted to intensive care / high dependency units, as well as the primary sign or symptom the patient presents with. To

support submissions to payers, additional analyses may be undertaken and these will be outlined in a separate Payer Analysis Plan.

4.2.4 Safety data

Safety and tolerability data from all cycles of treatment will be combined and will be presented by treatment arm using the safety population. Safety summaries will be descriptive only. No formal statistical analyses will be performed on the safety variables.

The following sections describe the planned safety summaries for AEs, vital signs, laboratory parameters, ECG. However, additional safety summaries (not specified in this SAP) may need to be produced to aid interpretation of the safety data.

4.2.4.1 Adverse events

All AEs, both in terms of current MedDRA preferred term and CTCAE grade, will be summarized descriptively by count (n) and percentage (%) for each treatment group. Any AE occurring before randomized treatment (i.e., before the administration of the first infusion on Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as 'pre-treatment'. However, any AE occurring before the administration of the first dose on Study Day 1 that increases in severity after the first dose will be regarded as treatment emergent and thus will be included in the majority of summary tables.

AEs observed up until CCI following discontinuation of the study treatment (i.e., the last dose of randomized treatment) or until the initiation of the first subsequent anti-cancer therapy (including radiotherapy, except for palliative radiotherapy) following discontinuation of study treatment (whichever occurs first) will be used for reporting of all the AE summary tables. This will more accurately depict AEs attributable to study treatment only as some of AEs up to CCI following discontinuation of the study treatment are likely to be attributable to subsequent therapy.

However, to assess the longer-term toxicity profile, some of the AE summaries may also be produced containing AEs observed up until CCI following discontinuation of the study treatment (i.e., without taking subsequent therapy into account).

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

Summary information (the number and percent of patients by system organ class (SOC) and PT separated by treatment group) will be tabulated for:

- All AEs

- All AEs causally related to durvalumab/placebo only (as determined by the reporting investigator)
- All AEs causally related to SoC CRT only (as determined by the reporting investigator)
- All AEs causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator)
- AEs with CTCAE grade 3 or 4
- AEs causally related to durvalumab/placebo only (as determined by the reporting investigator) with CTCAE grade 3 or 4
- AEs causally related to SoC CRT only (as determined by the reporting investigator) with CTCAE grade 3 or 4
- AEs causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator) with CTCAE grade 3 or 4
- AEs with maximum CTCAE grade 3 or 4
- AEs causally related to durvalumab/placebo only (as determined by the reporting investigator) with maximum CTCAE grade 3 or 4
- AEs causally related to SoC CRT only (as determined by the reporting investigator) with maximum CTCAE grade 3 or 4
- AEs causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator) with maximum CTCAE grade 3 or 4
- Most common AEs
- Most common AEs with CTCAE grade 3 or 4
- Most common AEs with maximum CTCAE grade 3 or 4
- AEs with outcome of death
- AEs with outcome of death causally related to durvalumab/placebo only (as determined by the reporting investigator)
- AEs with outcome of death causally related to SoC CRT only (as determined by the reporting investigator)
- AEs with outcome of death causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator)

- All SAEs
- All SAEs causally related to durvalumab/placebo only (as determined by the reporting investigator)
- All SAEs causally related to SoC CRT only (as determined by the reporting investigator)
- All SAEs causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator)
- AEs leading to discontinuation of durvalumab/placebo only
- AEs leading to discontinuation of SoC CRT only
- AEs leading to discontinuation of durvalumab/placebo or SoC CRT
- AEs leading to discontinuation of durvalumab/placebo only, causally related to durvalumab/placebo only (as determined by the reporting investigator)
- AEs leading to discontinuation of SoC CRT only, causally related to CRT only (as determined by the reporting investigator)
- AEs leading to discontinuation of SoC CRT only, causally related to durvalumab/placebo only (as determined by the reporting investigator)
- AEs leading to discontinuation of durvalumab/placebo or SoC CRT, causally related to durvalumab/placebo or SoC CRT (as determined by the reporting investigator)
- AEs leading to dose interruption of durvalumab/placebo only
- AEs leading to dose interruption of SoC CRT only
- AEs leading to dose delay of durvalumab/placebo or SoC CRT

An overall summary of the number and percentage of patients in each category will be presented, as will an overall summary of the number of episodes in each category. In addition, a truncated AE table of most common AEs and other tables showing most common AEs with CTCAE grade 3 or 4, and most common AEs with maximum CTCAE grade 3 or 4, showing all events that occur in at least 5% of patients overall will be summarized by preferred term, by decreasing frequency in the total column (the total column will not be displayed in the AE tables). This cut-off may be modified after review of the data. When applying a cut-off (i.e., x %), the raw percentage should be compared to the cut-off, no rounding should be applied first (i.e., an AE with frequency of 4.9% will not appear if a cut-off is 5%). Summary statistics showing the time to onset and the duration of the first AE may also be presented as appropriate.

Each AE event rate (per 100 patient years) will also be summarized by preferred term within each SOC for the output summarizing all AEs. For each preferred term, the event rate is defined as the number of patients with that AE divided by the total treatment duration (days) of randomized treatment summed over patients and then multiplied by 365.25 x 100 to present in terms of per 100 patient years.

Summaries of the number and percentage of patients with AEs will be provided by maximum reported CTCAE grade, SOC, PT and treatment group.

Fluctuations observed in CTCAE grades during study will be listed for those AEs which are CTCAE ≥ 3.

In addition, all AEs will be listed.

Deaths

Two summaries of all deaths will be provided with number and percentage of patients by treatment group, categorized as:

- Total number of deaths (regardless of date of death)
- Death related to disease under investigation only as determined by the investigator
- Death related to disease under investigation and an AE with outcome of death
    a. AE onset prior to subsequent therapy. Which includes AEs with an onset date (or pre-treatment AEs that increase in severity) on or after the date of first dose and up to and including CCI following the last dose of study medication (durvalumab/placebo), or AE start date <= the date of initiation of the first subsequent therapy (whichever occurs first).
    b. AE onset after start of subsequent therapy. Which includes AEs with start date >CCI following the last dose of study medication (durvalumab/placebo) and AE start date > the date of initiation of the first subsequent therapy (whichever occurs first).
- AE with outcome of death only
    a. AE onset prior to subsequent therapy. Which includes AEs with an onset date (or pre-treatment AEs that increase in severity) on or after the date of first dose and up to and including CCI following the last dose of study medication (durvalumab/placebo), or AE start date <= the date of initiation of the first subsequent therapy (whichever occurs first).
    b. AE onset after start of subsequent therapy. Which includes AEs with start date >CCI following the last dose of study medication (durvalumab/placebo)

and AE start date > the date of initiation of the first subsequent therapy (whichever occurs first).

- Death after end of safety follow up period (last dose of study medication (durvalumab/placebo) + CCI and not due to disease under investigation
- Unknown reason for death
- Other deaths

This summary will be repeated for all deaths on-treatment or within CCI of last dose of durvalumab/placebo.

Adverse events of special interest and possible interest

PTs used to identify adverse events of special interest (AESI) and adverse event of possible interest (AEPI), as defined in Section 3.4.1, will be listed before DBL and documented in the Study Master File.

Grouped summary tables of certain MedDRA PTs will be produced. For each 'grouped' term, the number (%) of patients experiencing any of the specified terms will be presented by maximum CTCAE grade. Additional summaries will include duration of AE and time to onset of first AE for each grouped term and preferred term within it. 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.

Additional summaries of the above-mentioned grouped AE categories will include number (%) of patients who have:

- At least one AESI/AEPI presented by outcome
- At least one AESI/AEPI causally related to study treatment
- At least one AESI/AEPI leading to discontinuation of study treatment

Additionally, there will be several summaries of AESIs/AEPIs requiring concomitant treatment, and particularly the relationship of AESIs/AEPIs to the use of immunosuppressive agents (i.e., depicting which AESI/AEPI triggered immunosuppressive use) and, separately, to the use of immunosuppressive agents at high doses.

Similarly to AESI/AEPI, summaries will be provided for the number (%) of patients experiencing adverse events of pneumonitis and radiation pneumonitis.

Summary of long term tolerability

To assess long term tolerability, provided that there are a sufficient number of patients with events to warrant it, prevalence plots and cumulative incidence plots will be presented for

each of the AESI grouped terms and any other events considered important after review of the safety data, provided there are ≥ 10 events.

A prevalence plot provides information on the extent to which the events may be an ongoing burden to patients. The prevalence at time t after first dose of study treatment is calculated as the number of patients experiencing the event divided by the number of patients receiving study treatment or in safety follow-up at time t; generally, t is categorized by each day after dosing. The prevalence is plotted over time split by treatment arm. Multiple occurrences of the same event are considered for each patient but a patient is only counted in the numerator whilst they are experiencing one of the occurrences of the event. These plots will be presented for each of the AESI grouped terms and will only be produced for AESIs that have ≥10 events.

A cumulative incidence plot is a plot of the raw cumulative incidence and cumulative incidence function over time with the treatment groups presented on separate plots. The raw cumulative incidence is the actual probability that a patient will have experienced their first occurrence of the event by a given time point. The cumulative incidence function estimates the cumulative incidence if the data cut-off had not been imposed and all patients had completed safety follow-up (Pintilie M.). This plot will be presented for "Pneumonitis" AESI and will only be produced if the AESI have ≥10 events.

4.2.4.2 Laboratory assessments

Data obtained up until the CCI following discontinuation of study treatment (durvalumab/placebo) or until the initiation of the first subsequent anti-cancer therapy (including radiotherapy, except for palliative radiotherapy) following discontinuation of study treatment (whichever occurs first) will be used for reporting. This will more accurately depict laboratory toxicities attributable to study treatment only as some toxicities up to CCI following discontinuation of the study treatment are likely to be attributable to subsequent therapy.

Any data post CCI after the last dose of the study treatment will not be summarized.

Data summaries will be provided in preferred units.

Scatter plots (shift plots) of baseline to maximum value/minimum value (as appropriate) on treatment (i.e., on-treatment is defined as data collected between the start of treatment and the relevant follow-up period following the last dose of study treatment) may be produced for certain parameters if warranted after data review.

Box-plots of absolute values by week, and box-plots of change from baseline by week, may be presented for certain parameters if warranted after data review. For continuous laboratory assessments absolute value and change from baseline will be summarized using descriptive statistics at each scheduled assessment time by actual treatment group.

Shift tables for laboratory values by worst CTCAE grade will be produced, and for specific parameters separate shift tables indicating hyper- and hypo- directionality of change will be produced. The laboratory parameters for which CTCAE grade shift outputs will be produced are:

- Hematology: Hemoglobin; Leukocytes; Lymphocytes (count, absolute); Neutrophils (count, absolute); Platelets

- Clinical chemistry: ALT, AST, ALP, Total Bilirubin, Albumin, Magnesium – hypo and – hyper, Sodium – hypo and – hyper, Potassium – hypo and – hyper, Corrected Calcium – hypo and – hyper, Glucose – hypo and – hyper, Creatinine.

Additional summaries will include a shift table for urinalysis (Bilirubin, Blood, Glucose, Ketones, Protein) comparing baseline value to maximum on-treatment value if a sufficient number of urinalysis assessments are recorded.

Liver enzyme elevations and Hy's law

The following summaries will include the number (%) of patients who have:

- Elevated ALT, AST, and Total bilirubin during the study

  - ALT ≥ 3x –≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x, and >20x Upper Limit of Normal (ULN) during the study

  - AST ≥ 3x–≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x, and >20x ULN during the study

  - Total bilirubin ≥2x-≤3x, >3x-≤5x, >5x ULN during the study

  - ALT or AST ≥3x-≤5x, >5x - ≤8x, >8x - ≤ 10x, >10x - ≤ 20x, >20x ULN during the study

  - ALT or AST ≥3x ULN and Total bilirubin ≥2x ULN during the study (Potential Hy's law): The onset date of ALT or AST elevation should be prior to or on the date of Total Bilirubin elevation

Liver biochemistry test results over time for patients with elevated ALT or AST (i.e., ≥ 3x ULN), and elevated Total Bilirubin (i.e., ≥ 2x ULN) (at any time) will be plotted. Individual patient data where ALT or AST (i.e., ≥ 3x ULN) plus Total Bilirubin (i.e., ≥ 2x ULN) are elevated at any time will be listed also.

Plots of ALT and AST vs. Total Bilirubin by treatment group will also be produced with reference lines at 3×ULN for ALT, AST, and 2×ULN for Total Bilirubin. In each plot, Total Bilirubin will be on the vertical axis.

Assessment of Thyroid Function Test Results

The following summaries will include the number and percentage of patients who have elevated or low thyroid stimulating hormone (TSH).

- TSH > ULN
- TSH > ULN with TSH ≤ ULN at baseline
- TSH > 3 X ULN
- TSH > 3 X ULN with TSH ≤ ULN at baseline
- TSH > 10 X ULN
- TSH > 10 X ULN with TSH ≤ ULN at baseline
- TSH < LLN
- TSH < LLN with TSH ≥ LLN at baseline

A separate summary will present:

- Number of subjects with at least one post-baseline TSH result
    - On-treatment elevated TSH > ULN
    - On-treatment elevated TSH > ULN with TSH≤ ULN at baseline
    - On-treatment elevated TSH > ULN with at least one T3 free/T4 free < LLN
    - On-treatment elevated TSH > ULN with all other T3 free/T4 free ≥ LLN
    - On-treatment elevated TSH > ULN with T3 free/T4 free missing
    - On-treatment decreased TSH < LLN
    - On-treatment decreased TSH < LLN with TSH≥ LLN at baseline
    - On-treatment decreased TSH < LLN with at least one T3 free/T4 free > ULN
    - On-treatment decreased TSH < LLN with all T3 free/T4 free ≤ ULN
    - On-treatment decreased TSH < LLN with T3 free/T4 free missing

Assessment of Renal Function Test Abnormalities

In addition to the analysis for serum creatinine, the number and percentage of patients with creatinine clearance (CrCl) rate during treatment period meeting the following categories will be presented:

- Normal: CrCl ≥ 90 mL/min
- Mild Impairment: CrCl ≥ 60 - < 90 mL/min
- Moderate Impairment: CrCl ≥ 30 - < 60 mL/min
- Severe Impairment: CrCl ≥ 15 - < 30 mL/min
- Kidney Failure: CrCl < 15 mL/min

A summary of the reversibility of creatinine clearance will be provided as the number and percentage of patients who shifted to a worse on-treatment renal impairment category from baseline and the number and percentage of patients whose worsened renal impairment was reversible and transient, defined as subsequent creatinine clearance value that is higher than the worst on-treatment creatinine clearance value and in a better impairment category.

Creatinine clearance rate will be calculated using serum Creatinine and the Cockcroft-Gault formula (Cockcroft and Gault 1976).

4.2.4.3 ECGs

ECG data obtained up until the safety follow-up will be included in the summary tables.

Overall evaluation of ECG is collected in terms of normal or abnormal, and the relevance of the abnormality is termed as "clinically significant" or "not clinically significant". ECG evaluations will be summarized using a shift table of baseline evaluation to worst evaluation on-treatment, if a sufficient number of ECG assessments are recorded.

4.2.4.4 Vital signs

Vital signs data obtained up until the final visit will be included in the summary tables.

Box plots for absolute values and change from baseline by week may be presented for certain vital signs parameters, if warranted after data review.

Vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, temperature, respiratory rate and weight) will be summarized over time in terms of absolute values and change from baseline at each scheduled measurement by actual treatment group

4.2.4.5 Physical examination

Individual physical examination data will not be summarized.

4.2.4.6 Other safety data

Data from positive pregnancy tests will not be summarized.

4.2.5 WHO performance status

All WHO performance status data will be summarized over time for the FAS population.

CCI [redacted]

CCI [redacted]

4.2.9 Demographic and baseline characteristics data

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

- Patient disposition (All patients)
- Important protocol deviations
- Inclusion in analysis populations
- Demographics (age, age group [redacted] sex, race and ethnicity)
- Patient characteristics at baseline (height, weight, weight group, body mass index (BMI) and BMI group)
- Patient recruitment by region, country and center
- Previous disease-related treatment modalities
- Time from initial diagnosis to randomization
- Disease characteristics at baseline (WHO performance status, primary tumor location, histology type, tumor grade and overall disease classification)
- Extent of disease at study entry
- TNM classification at screening and at diagnosis
- Medical history (past and current)

- Relevant surgical history (as appropriate)
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use, categorized (never, current, former)
- Stratification factors as per IVRS and eCRF data, as well as discrepancies between IVRS vs eCRF stratification

4.2.10 Treatment exposure and intensity

The following summaries related to study treatment will be produced for the safety analysis set by treatment group:

- Total exposure
- Actual exposure (durvalumab or matching placebo only)
- Reasons for dose delays/interruptions. Dose interruptions will be based on investigator initiated dosing decisions
- Reasons for premature termination of radiation therapy
- Number of infusions received
- [redacted]

For patients on study treatment at the time of the interim analysis, the DCO date will be used to calculate exposure.

4.2.11 Coronavirus Disease 2019 (COVID-19)

A listing of all patients affected by a COVID-19 related study disruption by unique subject number identifier and investigational site will be generated along with the description of how the individual's participation was altered. A listing of patients with reported issues in the Clinical Trial Management System due to COVID-19 pandemic will be generated. Additional analyses might be conducted to investigate the impact of COVID-19 on study endpoints.

5. INTERIM ANALYSES

5.1 Analysis methods

Interim safety monitoring will be conducted by an IDMC. Interim analyses will be performed for efficacy as described below:

5.1.1 PFS interim analysis

CCI [redacted]

The alpha level allocated to PFS will be controlled at the [redacted] by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, [redacted].

The significance level for the PFS analyses will be calculated using the statistical software package EAST by specifying the information fraction for each analysis. The information fraction is calculated as the number of PFS events at the analysis time-point divided by the total number of events at the final analysis time-point. For example, with the alpha level of 5%, CCI [redacted] of PFS events required at the time of the primary PFS analysis are available at the time of the interim CCI [redacted] the 2-sided significance level to be applied for the PFS IA would be CCI [redacted] and the 2-sided significance level to be applied for the [redacted] would be CCI [redacted]

It is expected that recruitment will have completed prior to the results of the IA being available.

5.1.2 OS interim analyses

Per protocol, CCI [redacted]

However, CCI [redacted] at the time CCI [redacted] the number of CCI [redacted] is now estimated to reach the CCI [redacted] rather than the CCI [redacted] Hence, if at the time of the CCI [redacted] is within approximately CCI [redacted]

The alpha level allocated to OS will depend on the results of the PFS and ORR analysis (see Section 4.1.2). It will be controlled at the [redacted] analysis by using the Lan DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, where the significance level applied at the IA depends upon the proportion of information available.

The significance level for the OS analyses will be calculated using the statistical software package EAST by specifying the information fraction for each analysis. The information fraction is calculated as the number of OS events at the analysis time-point divided by the total number of events at the final analysis time-point. For example, if the alpha level is 5% and [redacted] of OS events required at the time of the primary OS analysis are available at the time of the IA [redacted] the 2-sided significance level to be applied for the OS IA would be [redacted] respectively, and the 2-sided significance level to be applied for the primary OS analysis would be [redacted]

5.1.3 ORR interim analysis

Formal statistical testing of ORR will occur when PFS is statistically significant.

In order to make allowance for the 2 possible datacuts where this could be analyzed (i.e., the interim or final PFS), the alpha level allocated to ORR will be controlled at the interim and final timepoints by using the Haybittle Peto spending function. Assuming an information fraction of [redacted] (as most responses are expected to occur early in the study), and applying an alpha of [redacted]

5.1.4 OS24 interim analysis

5.1.5 The statistical testing of OS24 will be performed when OS is statistically significant.Independent Data Monitoring Committee

An IDMC composed of independent experts will be convened to confirm the safety and tolerability of the proposed dose and schedule of durvalumab + SoC CRT. The first safety review will take place when [redacted] have completed SoC CRT and have had at least [redacted] The second safety review will take place when the [redacted] have completed SoC CRT and have had at least [redacted] Safety reviews will be carried out by the IDMC in an unblinded manner. After review of the unblinded data, the IDMC will make a recommendation on whether the study should continue recruitment as planned or hold recruitment.

An additional safety review for [redacted] will take place [redacted] This review will be carried out by the IDMC in an unblinded manner. After review of the unblinded data, the IDMC will make a recommendation on whether the study should continue recruitment as planned or hold recruitment [redacted]

The IDMC will also meet approximately every 6 months thereafter to continue safety monitoring.

In addition:

- CCI [redacted]

Note: At the final PFS analysis, AstraZeneca will be unblinded to the data (regardless of the outcome). Hence the IDMC will not be required to review the efficacy data for the OS IA.

Full details of the IDMC procedures, processes, and IAs can be found in the IDMC Charter.

6. CHANGES OF ANALYSIS FROM PROTOCOL

- Clarified the criteria for all RECIST endpoints (i.e., ORR, PFS, TTDM, etc..) of how non-evaluable assessments should be handled at missed visits and in censoring for time to event endpoints.
- Added landmark analyses for both DoR and TTDM endpoints.
- Removed section and corresponding analysis for "Other significant adverse events (OAEs)".
- Added listings and analyses to investigate and assess the impact of COVID-19 on study endpoints.
- Section 1.1.2 listing secondary objectives includes duplicated rows for immunogenicity analysis. To clarify, only investigation of immunogenicity of durvalumab when in combination with CRT will be carried out. This change is not reflected in Section 1.1.2 since the table is kept consistent with the one in CSP.

As a general principal, analyses outlined in the SAP supersede those in the protocol.

7. REFERENCES

Berry et al 1991
Berry G, Kitchin RM, Mock PA. A comparison of two simple hazard ratio estimators based on the logrank test. Stat Med 1991 May;10(5):749-55.

Breslow, 1974
Breslow, NE. Covariance Analysis of Censored Survival Data. Biometrics 1974; 30:89–99.

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009 Feb 28;28(5):739-61.

Butts et al 2014
Butts C, Socinski MA, Mitchell PL, Thatcher N, Havel L, Krzakowski M, et al, on behalf of the START trial team. Tecemotide (L-BLP25) vs placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. Lancet Oncol 2014;15(1):59-68.

Cockcroft and Gault 1976
Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976; 16(1):31-41.

Collett 2003
Collett D. Modelling survival data in medical research: 2nd ed. Chapman and Hall/CRC. 2003. Available on request.

Ellis et al 2008
Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008 Jul;29(4):456-65.

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.

Fleischer et al 2011
Fleischer F, Gaschler-Markefski B, Bluhmki E. How is retrospective independent review influenced by investigator-introduced informative censoring: a quantitative approach. Stat Med 2011 Dec 20;30(29):3373-86.

Gail and Simon 1985
Gail M, Simon R. Tests for qualitative interactions between treatment effects and patient subsets. Biometrics 1985 Jun;41(2):361–72.

Glimm et al 2009
Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine 2009; 29:219–28.

Hanna et al 2004
Hanna N, Shepherd FA, Fossella FV, Pereira JR, De Marinis F, von Pawel J, et al. Randomized phase III trial of pemetrexed vs docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J Clin Oncol 2004 May 1;22(9):1589-97.

Klein et al 2007
Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007 Oct 30; 26(24): 4505–19.

Lan and DeMets 1983
Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70:659-63.

Nishino et al 2013
Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH and Hodi FS. Developing a common language for tumor response to immunotherapy: Immune-related response criteria using unidimensional measurements. Clin Cancer Res 2013;19:3936-43

Oemar and Janseen 2013
Oemar M, Oppe M. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument V5.0 (October 2013). Available from URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

Osoba et al 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998 Jan;16(1):139-44.

Pintilie M.
Competing risks: A practical perspective. Wiley.

Robins and Tsiatis 1991
Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics-Theory and Methods 1991; 20(8):2609-31.

Robins 1993
Robins JM. Information Recovery and Bias Adjustment in Proportional Hazards Regression Analysis of Randomized Trials Using Surrogate Markers. Proceedings of the Biopharmaceutical Section, American Statistical Association 1993; 24-33.

Sellke and Siegmund 1983
Sellke T, Siegmund D. Sequential analysis of the proportional hazards model. Biometrika 1983;70:315-26.

Stone et al 2015
Stone A, Macpherson E, Smith A, Jennison C. Model free audit methodology for bias evaluation of tumour progression in oncology. Submitted.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method with the Conventional Approach for Interval-Censored Data Analysis. Stat Biopharm Res 2010;2(1):97-108.

Whitehead and Whitehead 1991
Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Statistics In Medicine 1991;10:1665-77.

Wolchok et al 2009
Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbé C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009 Dec 1;15(23):7412-20.

Ye et al 2012
Ye Y, Li A, Lui L, Yao B. A group sequential Holm procedure with multiple primary endpoints. Statistics in Medicine 2012; 32:1112-1124.

8. APPENDIX

APPENDIX A. 2 MISSED VISITS

ePRO Assessments

As per Sections 3.3.1.1 and 3.3.1.2, confirmation of deterioration will first be determined, then the censoring rules will be applied. For ePRO assessments, we will apply a 'look-forward' approach to the 2 missed visit rule.

For example:

Baseline	V1	V2	V3	V4	V5	V6	V7	V8	Observed confirmed deterioration visit*	Time to Deterioration derivation
	NC	DT	DT	NC	MV	NC	IMP	NC	Event at V2	Event at V2
	MV	MV	NC	DT	MV	DT	DT	NC	Event at V4	No event, censor at baseline
	MV	DT	NC	MV	MV	DT	DT	NC	Event at V6	No event, censor at V3
	NC	DT	NC	MV	IMP	DT	NC	DT	Event not confirmed, but if last DT then confirmed	Event at V8
	NC	DT	NC	MV	IMP	DT	MV	DT	Event at V6	Event at V6
	NC	IMP	NC	IMP	MV	MV	IMP	DT	Event at V8, last DT	No event, censor at V4
	MV	MV	NC	MV	MV	DT	DT	NC	Event at V6	No event, censor at baseline
	NC	DT	MV	MV	DT	NC	IMP	DT	Event at V2	Event at V2

DT Deterioration; IMP Improvement; MV Missed visit; NC No change (/stable).

*Missed visits are allowed in between assessments confirming deterioration, and confirmation of deterioration requires at least 14 days between assessments.

Length of 2 missed visit window for EORTC QLQ-C30 and QLQ-LC13

The scheduled visit assessments and visit windows for the QLQ-C30 and QLQ-LC13 are every CCI (±3 days) for the first CCI from randomization, and then every CCI (±3 days) up to CCI relative to date of randomization, and then every CCI (±1 week) through CCI relative to date of randomization, and then every CCI (±1 week) thereafter until PFS2.

Missing baseline or within 2 visits of baseline:

- A patient without baseline will be censored at day 1 unless they die ≤ day 31.
- Any assessment ≤ day 31 is within 2 visits CCI + 3 days) of baseline and therefore cannot be censored for missing visits.

Schedule: CCI ±3 days:

- For any assessment > day 1 then visit schedule is CCI ±3 days and thus the maximum time between 2 consecutive assessments is 3+14+14+3 = 34 days.

Schedule: CCI ±3 days:

- For any assessents where the visit schedule changes from CCI ±3 days to CCI ±3 days, the maximum time between 2 consecutive assessments is 3+14+28+3 = 48 CCI
- For any assessments where the visit scheulde is CCI ±3 week, the maximum time between two consecutive assessments is 3+28+28+3 = 62 days.

Schedule: CCI ±1 week:

- For any assessents where the visit schedule changes from CCI ±3 days to CCI ±1 week, the maximum time between 2 consecutive assessments is 3+28+56+7 = 94 days.
- For any assessments where the visit scheulde is CCI ±1 week, the maximum time between two consecutive assessments is 7+56+56+7 = 126 days.

Schedule: CCI ±1 week:

- For any assessents where the visit schedule changes from CCI ±1 week to CCI ±1 week, the maximum time between 2 consecutive assessments is 7+56+84+7 = 154 days.
- For any assessments where the visit scheulde is CCI ±1 week, the maximum time between two consecutive assessments is 7+84+84+7 = 182 days.

APPENDIX B. [redacted]

Rule #	Category	Programming rules
	General	[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]

		[redacted]
4	[redacted]	[redacted]

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: d933kc00001-sap-ed-4		
Document Title:	d933kc00001 sap ed 4	
Document ID:	Doc ID-004953503	
Version Label:	2.0 CURRENT LATEST APPROVED	
Server Date **(dd-MMM-yyyy HH:mm 'UTC'Z)**	**Signed by**	**Meaning of Signature**
05-Jun-2023 14:20 UTC	PPD	Content Approval
05-Jun-2023 10:24 UTC	PPD	Content Approval

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