Document Type:	Final Statistical Analysis Plan
Document Dates:	09 November 2020
Study Title:	Phase II Randomized, Multi-Center, Double-Blind, Global Study to Determine the Efficacy and Safety of Durvalumab plus Olaparib Combination Therapy Compared with Durvalumab Monotherapy as Maintenance Therapy in Patients whose Disease has not Progressed Following Standard of Care Platinum-Based Chemotherapy with Durvalumab in First-Line Stage IV Non-Small Cell Lung Cancer (ORION)
Protocol Reference Number:	D9102C00001
NCT Number:	NCT03775486

AstraZeneca

Statistical Analysis Plan

Study Code	D9102C00001
Edition Number	3.0
Date	09 November 2020

A Phase II Randomized, Multi-Center, Double-Blind, Global Study to Determine the Efficacy and Safety of Durvalumab plus Olaparib Combination Therapy Compared with Durvalumab Monotherapy as Maintenance Therapy in Patients whose Disease has not Progressed Following Standard of Care Platinum-Based Chemotherapy with Durvalumab in First-Line Stage IV Non-Small Cell Lung Cancer (ORION)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
1L	First line
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ALK	Anaplastic lymphoma kinase
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AML	Acute myeloid leukemia
AST	Aspartate aminotransferase
BICR	Blinded independent central review
BoR	Best overall response
BP	Blood pressure
CI	Confidence interval
COVID-19	Coronavirus Disease 2019
CR	Complete response

Abbreviation or special term	Explanation
CrCl	Creatinine clearance
CRO	Contract research organization
CSP	Clinical study protocol
CSR	Clinical study report
CT	Computerized tomography
CTCAE	Common Terminology Criteria for Adverse Event
CV	Coefficient of variation
CCI	[redacted]
DAE	Discontinuation of investigational product due to adverse events
DBL	Database lock
DBP	Diastolic blood pressure
DCO	Data cutoff
DNA	Deoxyribonucleic acid
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
CCI	[redacted]
FAS	Full analysis set
HR	Hazard ratio
HRQoL	Health-related quality of life
HRR	Homologous recombination repair
HRRm	Homologous recombination repair mutant
ICU	Intensive care unit
IDMC	Independent data monitoring committee
CCI	[redacted]
IHC	Immunohistochemistry
IP	Investigational product
IPD	Important protocol deviation
IRMC	Independent review master charter

Abbreviation or special term	Explanation
ITT	Intent-to-treat
IxRS	Interactive voice/web response system
KM	Kaplan-Meier
LD	Longest diameter
LIMS	Laboratory information management system
MDS	Myelodysplastic syndrome
MedDRA	Medical dictionary for regulatory activities
MMRM	Mixed-effect model for repeated measures
MRI	Magnetic resonance imaging
CCI	[redacted]
MTP	Multiple testing procedure
NA	Not applicable
nAb	Neutralizing antibody
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
NSCLC	Non-small cell lung cancer
NTL	Non-target lesion
OAE	Other significant adverse events
ORR	Objective response rate
OS	Overall survival
PD	Progressive disease
PD-L1	Programmed cell death ligand 1
PFS	Progression-free survival
CCI	[redacted]
PK	Pharmacokinetic(s)
PR	Partial response
PRO	Patient-reported outcome
q12w	Every 12 weeks
q8w	Every 8 weeks
QLQ-C30	30 item core quality of life questionnaire
QLQ-LC13	13-item lung cancer quality of life questionnaire

Abbreviation or special term	Explanation
QoL	Quality of life
QTc	QT interval corrected for heart rate
QTcF	QT interval corrected for heart rate using Fridericia's formula
RDI	Relative dose intensity
REML	Restricted maximum likelihood
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious adverse event
SAP	Statistical analysis plan
SBP	Systolic blood pressure
SD	Stable disease
sf	Significant figure
SoC	Standard of care
TC	Tumor cells
TEAE	Treatment-emergent adverse events
TL	Target lesion
CCI	[redacted]
TNM	Classification of malignant tumors (tumor, lymph nodes, metastasis)
TSH	Thyroid-stimulating hormone
ULN	Upper limit of normal
VAS	Visual analogue scale
WHO	World Health Organization

AMENDMENT HISTORY

Category*: Change refers to	Date	Description of change	In line with the CSP?	Rationale
Primary or secondary endpoints	18 March 2020	Definition of time to symptom deterioration and time to HRQoL/function deterioration was changed from an unconfirmed to a confirmed deterioration.	NA	Consistency with other studies
Statistical analysis method for the primary or secondary endpoints	18 March 2020	Changed endpoints subject to MMRM analysis from multi-item symptom scales, the 5 functional scales and global health status/QoL scale to the primary PRO symptoms of interest. Changed endpoints subject to time-to-event analysis from multi-item symptom scales, the 5 functional scales and global health status/QoL scale to all symptom scales, the 5 functional scales and global health status/QoL scale. Changed endpoints subject to improvement rate analysis from multi-item symptom scales, the 5 functional scales and global health status/QoL scale to all symptom scales, the 5 functional scales and global health status/QoL scale.	Yes	Consistency with updates in the protocol
Other	18 March 2020	Updated to specify that 2 interim analyses will be formed for the key secondary endpoint of OS, and to provide a formal alpha-spending plan for the analysis of OS.	Yes	Consistency with updates in the protocol
Other	18 March 2020	Global Product Statistician transition from PPD to PPD.	NA	NA
Other	09 November2020	Updated to remove the interim analysis of PFS at 104 events; updated to remove the first interim analysis of OS.	Yes	Consistency with updates in the protocol
Statistical analysis method for the primary or secondary endpoints	09 November2020	Updated subgroup analysis to use eCRF data instead of IXRS data for the subgroups based on strata.	Yes	Consistency with other studies

Statistical analysis method for the primary or secondary endpoints	09 November2020	Added analysis of covariate effect on HR estimate for the primary endpoint	Yes	Consistency with other studies
Other	09 November2020	Added a summary of agreement of investigator assessment of RECIST progression and BICR assessment	NA	Consistency with other studies
Statistical analysis method for the primary or secondary endpoints	09 November2020	Added subgroup analysis for OS.	Yes	Consistency with other studies
Statistical analysis method for the primary or secondary endpoints	09 November2020	Added summary statistics of follow up time for OS.	Yes	Consistency with other studies
Statistical analysis method for the primary or secondary endpoints	09 November2020	Added a summary of confirmed objective response for Maintenance Phase	Yes	Consistency with other studies
Other	09 November2020	Updated summary of overall tumor response as reported by investigator in the Initial therapy phase to be presented at the last visit	Yes	Consistency with protocol
Statistical analysis method for the primary or secondary endpoints	09 November2020	Added an additional restriction that PRO data within 12 month of randomization or PD to be included for the MMRM analysis. Added an interaction term of baseline by visit to the MMRM model and removed random subject effect from the model. Added specification for visit sequencing and windowing.	Yes	Consistency with other studies
Statistical analysis method for the primary or secondary endpoints	09 November2020	Removed some symptom scales from time to deterioration analysis and improvement rate analysis for the PRO parameters.	Yes	Consistency with other studies
Other	09 November2020	Added summaries of chemotherapy exposure in initial therapy phase for subjects treated in the initial phase and for subjects entered maintenance phase.	Yes	Additional summary

Other	09 November2020	Updated exposure calculation for olaparib/placebo to exclude duration of delays.	Yes	Consistency with other studies
Other	09 November2020	Updated relative dose intensity (RDI) to include exposure up to the actual last day of dosing.	Yes	Consistency with TA SAP
Other	09 November2020	Updated RDI for olaparib/placebo to use 300 mg BID as the planned dose.	Yes	Consistency with other studies
Other	09 November2020	Updated summaries of AESI; added summaries of AEPI.	Yes	Consistency with AZ program wise updates
Other	09 November2020	Added summaries of infection AEs	Yes	Consistency with other studies
Data presentations	09 November2020	Updated summaries of death.	Yes	Consistency with other studies
Other	09 November2020	Some inclusion exclusion criteria removed from the entry criteria IPD (IPD #2).	Yes	Consistency with updated protocol deviation
Other	09 November2020	Removed the specified changes to analysis in Section 6, as the analyses are consistent with the protocol amendment.	Yes	Consistency with updates in the protocol
Other	09 November2020	Updated subsequent anti-cancer therapy for safety definition to include radiotherapy, with the exception of palliative radiotherapy.	Yes	Consistency with other studies
Other	09 November2020	Added a section for COVID-19 data summaries.	Yes	Consistency with AZ program wise updates

* Pre-specified categories are
Primary or secondary endpoints; Statistical analysis method for the primary or secondary endpoints; Derivation of primary or secondary endpoints; Multiple Testing Procedure; Data presentations; Other

1 STUDY DETAILS

This statistical analysis plan (SAP) contains a more detailed description of the analyses in the clinical study protocol (CSP). This SAP is based on version 4.0 of the CSP.

1.1 Study objectives

Table 1 Study objectives

Primary Objective:	**Endpoint/Variable:**
To assess the efficacy of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of PFS (Investigator-assessed)	PFS: Time from date of randomization until the date of objective radiological disease progression according to Investigator assessment using RECIST 1.1 or death by any cause in the absence of progression
Secondary Objectives:	**Endpoint/Variable:**
<u>Key Secondary Objective:</u>	
To further assess the efficacy of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of OS	OS: Time from date of randomization until the date of death by any cause
<u>Additional Secondary Objectives:</u>	
To further assess the efficacy of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of ORR and DoR	ORR: Percentage of patients with an Investigator-assessed response of CR or PR after randomization DoR: Time from the date of first documented response following randomization until the first date of documented progression or death in the absence of disease progression
To further assess the efficacy of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of PFS (Investigator-assessed) in the HRRm population	PFS: Time from date of randomization until the date of objective radiological disease progression according to Investigator assessment in HRRm population using RECIST 1.1 or death (by any cause in the absence of progression)
To assess the PK of durvalumab in combination with olaparib	Concentration of durvalumab
To assess disease-related symptoms and HRQoL in patients treated with durvalumab plus olaparib combination therapy compared with durvalumab monotherapy	Change from baseline and time to deterioration (for maintenance phase) in EORTC QLQ-C30 and QLQ-LC13
To investigate the immunogenicity of durvalumab	Presence of ADAs for durvalumab

Table 1 Study objectives

Safety Objective:	Endpoint/Variable:
To assess the safety and tolerability profile of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy	AEs, physical examinations, laboratory findings, and vital signs
CCI	CCI

ADA Anti-drug antibody; AE Adverse event; CCI ; DoR Duration of response; EORTC European Organization for Research and Treatment of Cancer: CCI ; HRQoL Health-related quality of life; CCI HRRm Homologous recombination repair related gene mutation; CCI CCI ORR Objective response rate; OS Overall survival; CCI PFS Progression-free survival; CCI PK Pharmacokinetics; CCI QLQ-C30 30-item Core Quality of Life Questionnaire; QLQ-LC13 13-item Lung Cancer Quality of Life Questionnaire; RECIST 1.1 Response Evaluation Criteria in Solid Tumors version 1.1; CCI .

1.2 Study design

This is a Phase II randomized, multi-center, double-blind, global study to determine the efficacy and safety of durvalumab plus olaparib combination therapy compared with durvalumab monotherapy as maintenance therapy in subjects whose disease has not progressed following standard of care (SoC) platinum-based chemotherapy with durvalumab in first-line Stage IV NSCLC. There will be approximately 80 sites in the study. During the initial therapy phase, approximately 350 to 400 subjects will receive treatment with durvalumab, along with the Investigator's choice of platinum-based doublet therapy for squamous NSCLC (nanoparticle albumin-bound [nab]-paclitaxel plus carboplatin or gemcitabine plus carboplatin/cisplatin) and nonsquamous NSCLC (nab-paclitaxel plus carboplatin or pemetrexed plus carboplatin/cisplatin) for 4 cycles.

It is estimated that approximately 350 to 400 subjects will be enrolled in the initial therapy phase in order for approximately 250 subjects who have not progressed (i.e., maintained complete response [CR], partial response [PR], or stable disease [SD] throughout the initial therapy phase according to Investigator-assessed RECIST 1.1) to be randomized into the maintenance phase of the study (subjects completing the initial therapy phase who are not randomized cannot continue durvalumab). Subjects will be randomized 1:1 to receive either durvalumab plus placebo or durvalumab plus olaparib maintenance therapy. Randomization will be stratified by histologic subtype (squamous or nonsquamous) and objective response (CR/PR or SD; obtained at the last visit prior to randomization [Cycle 4 scan]) during the initial therapy phase.

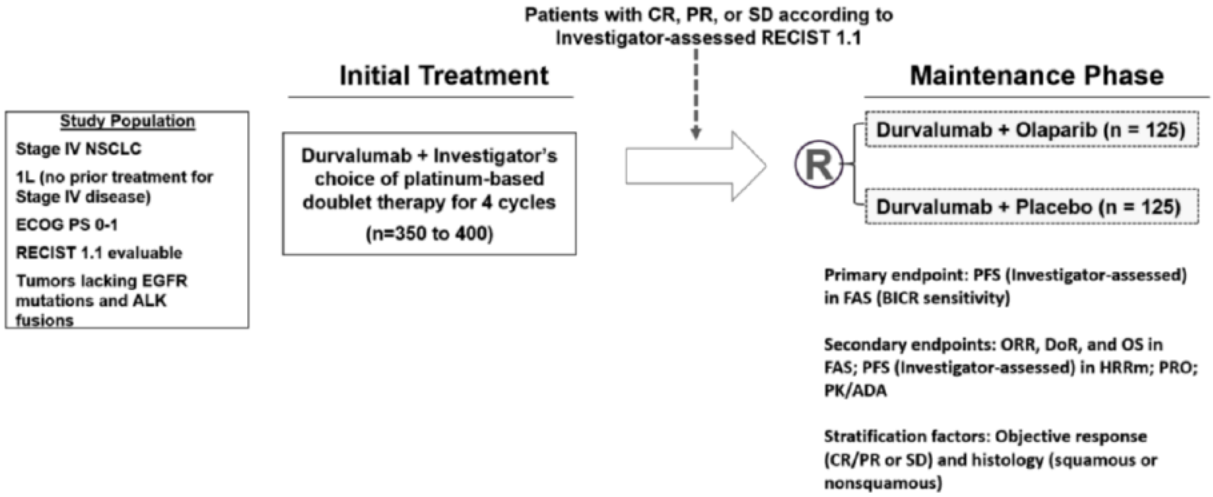
Confirmation of eligibility criteria for randomization (eligibility scan and other specific criteria; see Sections 5.1 and 5.2 of the Clinical Study Protocol [CSP] for criteria that must be met at randomization) will take place 14 to 28 days after Cycle 4 Day 1 of the initial therapy phase. Laboratory assessments for eligibility should be taken after the last dose of chemotherapy in the initial therapy phase. If determined eligible, subjects will be randomized within 5 weeks after Cycle 4 Day 1 of the initial therapy phase; every effort should be made to minimize the time between confirmation of eligibility, randomization, and starting maintenance treatment. Subjects will receive maintenance treatment until specific discontinuation criteria are met, including clinical disease progression (as assessed by the Investigator) or RECIST 1.1-defined radiological progressive disease (PD), unacceptable toxicity, and withdrawal of consent. Note that crossover within the study will not be permitted.

Tumor evaluation using RECIST 1.1 will be conducted at screening (within 28 days prior to the first dose of study medication administered during the initial therapy phase), 14 to 28 days after Cycle 2 Day 1 and Cycle 4 Day 1 of the initial therapy phase, and every 8 weeks (q8w) ±1 week during the maintenance phase (for the first 48 weeks, and then q12w ±1 week thereafter) until RECIST 1.1-defined radiological PD plus one or more additional follow-up scans, if clinically feasible.

After treatment discontinuation for any reason other than RECIST 1.1-defined radiological PD, scanning/tumor assessments will continue until RECIST 1.1-defined radiological PD plus one or more additional follow-up scans (if clinically feasible). If treatment is discontinued due to RECIST 1.1-defined radiological PD, one or more additional follow-up scans (if clinically feasible) will be performed.

For an overview of the study design, see Figure 1.

Figure 1 Study design



1L First-line; ADA Anti-drug antibodies; ALK Anaplastic lymphoma kinase; BICR Blinded independent central review; CR Complete response; ECOG Eastern Cooperative Oncology Group; EGFR Epidermal growth factor receptor; DoR duration of response; FAS Full analysis set; HRRm Homologous recombination repair related gene mutation; NSCLC Non-small cell lung cancer; ORR Objective response rate; OS Overall survival; PFS Progression-free survival; PK Pharmacokinetic; PR Partial response; PRO Patient-reported outcome; PS Performance status; R Randomization; RECIST 1.1 Response Evaluation Criteria in Solid Tumors version 1.1; SD Stable disease.

1.3 Number of subjects

The study is sized to characterize the PFS benefit of durvalumab in combination with olaparib versus durvalumab monotherapy in subjects with Stage IV NSCLC whose disease has not progressed after 4 cycles of durvalumab with platinum-based first-line chemotherapy.

Approximately 350 to 400 subjects will be enrolled in the initial therapy phase of the study. Approximately 250 subjects globally who have not progressed will be randomized in a 1:1 ratio to either the durvalumab plus olaparib treatment arm or the durvalumab plus placebo treatment arm, approximately 125 subjects per arm. The randomization will be stratified based on objective response to durvalumab plus chemotherapy (CR/PR or SD; obtained at the last visit prior to randomization [Cycle 4 scan]) and histology (squamous or nonsquamous).

The DCO for the primary analysis of PFS will occur when approximately 163 PFS events have occurred across the durvalumab plus olaparib treatment arm and durvalumab plus placebo treatment arm (approximately 65% maturity). CCI [redacted]

CCI [redacted]

CCI [redacted]

One interim analysis of OS for durvalumab + olaparib vs durvalumab monotherapy will be performed at the time of the primary analysis for PFS when approximately 109 death events (67% information fraction) will be available. The final OS analysis will occur when 163 death events have occurred across the two treatment arms. CCI [redacted] The alpha will be split between the two OS analyses using the Lan and DeMets (Lan and DeMets 1983) spending function that approximates the O'Brien Fleming approach, with the boundaries for the treatment comparison derived based upon the actual number of OS events observed at the time of analysis.

2 ANALYSIS SETS

2.1 Definition of analysis sets

Definitions of the analysis sets for each outcome variable are provided in Table 2.

2.1.1 Full analysis set (FAS) for the initial therapy phase

The FAS for the initial therapy phase will include all subjects who received at least 1 dose of durvalumab and/or chemotherapy in the initial therapy phase of the study. The FAS for initial therapy phase will be used to summarize subject disposition and demographic characteristics for all subjects receiving initial therapy treatment in the study. It will also be used to summarize the tumor response as reported by investigators in the initial therapy phase.

2.1.2 Full analysis set (FAS)

The FAS for the maintenance phase will include all randomized subjects with treatment groups assigned in accordance with the randomization, regardless of the treatment actually received. Subjects who were randomized but did not subsequently proceed to receive study treatment are included in the analysis in the treatment arm to which they were randomized. The FAS therefore follows the principles of intent-to-treat (ITT).

The stratified analyses performed for FAS will be based on subjects' objective response to durvalumab plus chemotherapy during the initial therapy phase (CR/PR or SD) and histology (squamous or nonsquamous) reported in the IxRS system that they were subsequently randomized on. Sensitivity analyses may be performed based on the eCRF data and other source data if a substantial number of subjects are mis-stratified.

The FAS will be used for the primary efficacy analysis of PFS and all secondary efficacy analyses (including PROs). The HRRm subgroup of the FAS will be used for the secondary efficacy analysis of PFS.

Summaries of demographic and subject characteristics will be reported for the FAS.

2.1.3 Safety analysis sets (SAF)

2.1.3.1 Safety analysis set (SAF) for the initial therapy phase

The SAF for the initial therapy phase will consist of all subjects who received at least 1 dose of durvalumab during the initial therapy phase of the study. Minimal exposure and adverse event safety data will be summarized for the initial therapy phase only (not including post-randomization data for randomized subjects), unless unexpected safety signals are observed**.**

2.1.3.2 Safety analysis set (SAF)

This is the safety analysis set for the maintenance phase and it will consist of all subjects who received at least 1 dose of any study treatment (durvalumab and/or olaparib/placebo) during the maintenance phase of the study. Subjects will be classified based on the treatment actually received, that is, erroneously treated subjects will be summarized according to the treatment they actually received (e.g., those randomized to durvalumab plus placebo who receive one or more doses of olaparib in error will be reported in the durvalumab plus olaparib arm). Subjects who only receive durvalumab will be summarized according to the arm that they were randomized to.

Safety and tolerability summaries will be produced using the SAF.

2.1.4 PK analysis sets

All subjects who receive at least 1 dose of durvalumab for whom any post-dose data are available and who do not violate or deviate from the protocol in ways that would significantly affect the PK analyses will be included in the PK analysis set. The population will be defined by the Study Physician, Pharmacokineticist, and Statistician prior to any analyses being performed. Durvalumab PK data will be summarized for the initial therapy phase and maintenance phase.

Table 2 Summary of outcome variables and analysis populations

Outcome variable	Populations
Efficacy data	
PFS	FAS for maintenance phase (all subjects for the primary analysis and the subset of subjects with HRRm for a secondary analysis)

Table 2 Summary of outcome variables and analysis populations

Outcome variable	Populations
OS, ORR, DoR, PROs	FAS for maintenance phase
	ORR will be based on the subset of subjects in FAS who have measurable disease at randomization
	DoR will be based on the subset of subjects in FAS who have measurable disease at randomization and achieve objective tumor response after randomization
Demography	FAS for initial therapy phase and FAS for maintenance phase
PK	PK analysis set
Safety data	
Exposure	SAF for initial therapy phase and SAF for maintenance phase
AEs	SAF for initial therapy phase and SAF for maintenance phase
Laboratory measurements	SAF for initial therapy phase and SAF for maintenance phase
WHO/ECOG performance status	SAF for maintenance phase
Vital signs	SAF for maintenance phase

AE Adverse event; DoR Duration of response; ECOG Eastern Cooperative Oncology Group; FAS Full analysis set; HRRm Homologous recombination repair related gene mutation; ORR Objective response rate; OS Overall survival; PFS Progression-free survival.; PK Pharmacokinetics; PRO Patient-reported outcome; WHO World Health Organization.

2.2 Violations and deviations

2.2.1 Important protocol deviations

In accordance with ICH E3 a protocol deviation is any change, divergence, or departure from the study design or procedures defined in the protocol. Important protocol deviations (IPDs) are a subset of protocol deviations that may significantly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety, or well-being.

All IPDs will be programmatically derived from the eCRF data where possible. The following general categories will be considered important protocol deviations and will be listed and discussed in the CSR as appropriate:

1. Received prohibited concomitant systemic anti-cancer medications (including other anti-cancer agents). Please refer to the CSP Section 6.4 for the concomitant medications that are detailed as being ‘excluded’ from permitted use during the study. This will be used as a guiding principle for the physician review prior to database lock.

2. Subjects who deviate from the following key entry criteria per the Clinical Study Protocol (CSP).

    a) Inclusion criteria:

    Inclusion Criteria #5, 6 and 8 apply to screening prior to receiving initial therapy.

    5: Histologically or cytologically documented Stage IV NSCLC not amenable to curative surgery or radiation.

    6: Patients must have tumors that lack activating EGFR mutations and ALK fusions. If a patient has squamous histology or is known to have a tumor with a KRAS mutation, then EGFR and ALK testing are not required.

    8: No prior chemotherapy or any other systematic therapy for Stage IV NSCLC. Patients who have received prior platinum-containing adjuvant, neo-adjuvant, or definitive chemoradiation are eligible, provided that progression has occurred >12 months from end of last therapy.

    Inclusion criteria #16 applies to maintenance therapy only.

    16: Patients must have documented radiographic evidence of a timepoint tumor response of CR, PR, or SD according to Investigator-assessed RECIST 1.1 guidelines following the 4 cycles of platinum-based chemotherapy. An objective response does not have to be confirmed in order for the patient to be randomized.

    b) Exclusion criteria:

    Exclusion Criteria #2 applies to screening prior to receiving initial therapy.

    2: Mixed small-cell lung cancer and sarcomatoid variant NSCLC histology.

    Exclusion criteria #29 applies to maintenance therapy only.

    29: Inability to complete 4 cycles of platinum-based chemotherapy for any reason or discontinuation of durvalumab during initial therapy treatment. Dose interruptions or delays are not exclusionary.

3. Subjects randomized to either treatment arm that did not receive any durvalumab.

4. Subjects randomized to either treatment arm that did not receive any olaparib/placebo.

5. Subjects who received an alternative study treatment to that which they were randomized.

6. No baseline RECIST 1.1 assessment on or before start of initial therapy or baseline RECIST scan > 42 days before start of initial therapy. Note that although the screening period for baseline RECIST assessment was 28 days, an additional 14-day window should be applied thus only baseline RECIST assessments of greater than 42 days will be deemed as constituting an important deviation.

7. Randomized subjects who are determined to be eligible for the maintenance phase, but have not been randomized within 5 weeks of Cycle 4 Day 1 of the initial therapy phase.

Subjects who receive at least 1 dose of durvalumab during the initial therapy phase will be included in the SAF for the initial therapy phase as described in CSP Section 9.3.2.1. Subjects who receive the wrong treatment at any time during the maintenance phase will be included in the SAF for the maintenance phase as described in CSP Section 9.3.2.2. During the study, decisions on how to handle errors in treatment dispensing (with regard to continuation/discontinuation of study treatment or, if applicable, analytically) will be made on an individual basis with written instruction from the study team leader and/or statistician.

The important protocol deviations will be listed and summarized for all subjects who were randomized by randomized treatment group, and listed for subjects in the initial therapy phase who weren't randomized.

In addition to the programmatic determination of the deviations above, other study deviations captured from the eCRF module for inclusion/exclusion criteria will be tabulated and listed. Any other deviations from monitoring notes or reports (as recorded in CTMS), will be reported in an appendix to the CSR.

If any deviation is considered to impact upon PK, a subject or particular data for a subject may be excluded from the PK analysis set. None of the other deviations beyond those in the analysis set definitions will lead to subjects being excluded from the analysis sets described in section 2.1.

A per-protocol analysis excluding subjects with specific important protocol deviations is not planned; however, a 'deviation bias' sensitivity analysis may be performed on the progression free survival endpoint excluding subjects with deviations that may affect the efficacy of the trial therapy if > 10% of subjects in either treatment group have 1 or more important protocol deviations.

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.

Errors in stratifications (based upon stratification information recorded in the IxRS and eCRF source data) will also be summarized separately to the important protocol deviations.

2.2.2 Monitoring of important protocol deviations

The IPDs will be programmatically identified within the clinical database by programmed edit checks or via manual validation checks. A programmatically derived IPD report will be created listing all identified IPDs and the data used to identify them. This report will be reviewed at regular IPD review meetings held on at least a monthly basis. At this meeting, programmatically-derived IPDs will be checked to ensure that they have been correctly classified.

On an ongoing basis throughout the study, monitoring notes or summaries will be reviewed to determine any important post-entry deviations that are not identifiable via programming.

The final classification of IPDs will be made prior to database lock or data cut-off.

3 PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST visit responses

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

Tumor evaluation using RECIST 1.1 will be conducted at screening (within 28 days prior to the first dose of study medication administered during the initial therapy phase), 14 to 28 days after Cycle 2 Day 1 and Cycle 4 Day 1 of the initial therapy phase. Confirmation of eligibility criteria for randomization will take place 14 to 28 days after Cycle 4 Day 1 of the initial therapy phase. If determined eligible, subjects will be randomized within 5 weeks after Cycle 4 Day 1 of the initial therapy phase; every effort should be made to minimize the time between confirmation of eligibility, randomization, and starting maintenance treatment. Tumor assessments are then performed every 8 weeks ±1 week following randomization for the first 48 weeks and every 12 weeks ±1 week thereafter until confirmed progressive disease (PD) as per Investigator assessment of RECIST 1.1 and Investigator determination that the subject is no longer benefiting from treatment with the IP.

If an unscheduled assessment is performed, and the subject 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 subjects being assessed at a different frequency than other subjects.

There will be 2 baseline assessments, the first for the initial therapy phase and the second for the maintenance phase. A first "Initial Therapy Baseline" scan should be collected during screening (Days -28 to -1) for disease staging and for use as a RECIST 1.1 baseline scan for the initial therapy phase on-study scans (collected 14 to 28 days after Cycle 2 Day 1 and Cycle 4 Day 1 of the initial therapy phase). A subject's diagnostic scan may be used as a baseline scan for the initial therapy phase only if taken within 28 days of first-dose administration and

in accordance with the scan acquisition requirements. The scan at 14 to 28 days after Cycle 4 Day 1 of the initial therapy phase will be compared with the initial therapy baseline scan to determine eligibility for the maintenance phase and will also be the "Maintenance Baseline" scan for the assessment of response during the maintenance phase (for those subjects eligible for randomization into the maintenance phase) with new RECIST 1.1 baseline assignment of TLs/NTLs.

For the initial therapy phase, the tumor responses per the investigator will be reported.

For the maintenance phase, the analyses of the primary endpoint (PFS in the FAS) and secondary endpoints (ORR and DoR in the FAS for subjects with measurable disease at randomization, and PFS in the HRRm subgroup of the FAS) will be based on Investigator assessments using RECIST 1.1.

From the investigator's review of the imaging scans, the RECIST tumor response data will be used to determine each subject's visit response according to RECIST version 1.1. At each visit, subjects will be programmatically assigned a RECIST 1.1 visit response of complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD), using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a subject has had a tumor assessment which cannot be evaluated, then the subject 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).

For subjects with no disease at baseline for the maintenance phase due to a complete response to initial therapy (i.e. no TLs and no NTLs), evaluation of overall visit responses will be based on absence/presence of new lesions. If no TLs and no NTLs are recorded at a visit, both the TL and NTL visit response will be recorded as NA and the overall visit response will be no evidence of disease (NED). If a new lesion is observed then the overall visit response will be PD.

Please refer to Section 3.1.1 for the definitions of CR, PR, SD and PD for TL applicable for subjects with measurable disease at baseline, Section 3.1.2 for NTL applicable for subjects with NTL identified at baseline, and Section 3.1.3 for overall response for all subjects regardless of baseline disease status.

RECIST outcomes (i.e. PFS, ORR, and DoR) will be calculated programmatically from the site investigator data (see Section 3.2) from the overall visit responses.

Sensitivity analyses for the primary endpoint will be performed, including analyzing PFS according to BICR in the FAS. The BICR will be performed on all radiological scans of all randomized subjects. Please see Section 3.1.4 for details.

3.1.1 Target lesions (TLs) – site investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (except lymph nodes which must have

short axis ≥ 15 mm) with CT or MRI and which is suitable for accurate repeated measurements.

A subject can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (representative of all lesions involved suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded for the initial therapy or maintenance phase, then the measurement from the one that is closest and prior to first dose of on-study treatment (initial therapy phase) and the one prior to randomization (maintenance phase) will be used to define the associated baseline sum of TLs. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

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

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

TL response is summarized in Table 3.

Table 3 TL visit responses

Visit Reponses	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 <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not Evaluable (NE)	Only relevant in certain situations (i.e. if any of the target lesions 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 target lesion response

Visit Reponses	Description
Not applicable (NA)	No TLs are recorded at baseline

CR Complete response; PR Partial response; PD Progression of disease; NE Not evaluable; SD Stable disease; TL Target lesion.

Rounding of TL data

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

Missing TL data

For a visit to be evaluable then all TL measurements should be recorded. However, a visit response of PD will 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 LD recorded.

If there is at least one TL measurement missing and a visit response of PD cannot be assigned, the visit response will be not evaluable (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 will 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 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 are also met.

- Step 3: If not all lesions meet the CR criteria (i.e. a pathological lymph node selected as TL has short axis > 10mm or the reappearance of previously disappeared lesion) or a new lesion appears, then response will be set to PD

- Step 4: If after steps 1 – 3 a response can still not be determined the response will be set to remain as CR

TL too big to measure

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

TL too small to measure

If a TL becomes too small to measure then this will be indicated as such on the case report form and a value of 5mm will be entered into the database and used in TL calculations. However a smaller value may be used if the radiologist has not indicated 'too small to measure' on the case report form and has entered a smaller value that can be reliably measured. If a TL response of PD results (at a subsequent visit) then this will be reviewed by the study team blinded to treatment assignment.

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation, palliative surgery, or embolization), will be handled as specified below. Once a lesion has had intervention then it will 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 subject would be assigned a TL response of PD.

- Step 3: If after both steps PD has not been assigned, then, if appropriate, a scaled sum of diameters will be calculated (as long as ≤ 1/3 of the TLs have missing measurements), treating the lesion with intervention as missing, and PR or SD then assigned as the visit response. Subjects with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10mm for lymph nodes) and the lesions that have been subject to intervention also have a value of 0 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 will be treated as missing and scaled up (as per step 2 above).

Scaling (applicable only for irradiated lesions/lesion intervention)

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

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

Example of scaling

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

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

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

68 x 74 / 62 = 81 mm

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

Lesions that split in two

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

Lesions that merge

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

Change in method of assessment of TLs

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

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

At each visit an overall assessment of the NTL response will 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.

NTL response will be derived based on the Investigator's overall assessment of NTLs as shown in Table 4.

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

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

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

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

New lesions will be identified via a Yes/No tick box. The 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.

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

Subjects with 'symptomatic progression' in the maintenance phase 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.

Table 4 NTL visit responses

Visit Responses	Description
Complete Response (CR)	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10 mm short axis).
Progressive Disease (PD)	Unequivocal progression of existing NTLs. Unequivocal progression may be due to an important progression in one lesion only or in several lesions. In all cases the progression MUST be clinically significant for the physician to consider changing (or stopping) therapy.
Non-CR/Non-PD	Persistence of one or more NTLs with no evidence of progression.
Not Evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: For subjects 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

CR Complete response; NA Not applicable; NE Not evaluable; NTL Non-Target lesion; PD Progressive disease.

3.1.3 Overall visit response – site investigator data

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

Table 5 Overall visit responses

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

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

CR Complete response; NA Not applicable; NE Not evaluable; NED No evidence of disease; PD Progressive disease; PR Partial response; SD Stable disease.

3.1.4 Independent Review

A planned BICR of all radiological imaging data will be carried out using RECIST version 1.1. All radiological scans for all randomized subjects (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. As described in Section 3.1 there will be 2 baseline assessments, the first for the initial therapy phase and the second for the maintenance phase. Each phase will be reviewed separately for randomized subjects. 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 subject, 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 subjects with TLs at baseline: CR, PR, SD, PD, NE; for subjects with NTLs only: CR, SD, PD, NE; for subjects with no disease identified at baseline: PD, no evidence of disease [NED], NE). If a subject has had a tumor assessment that cannot be evaluated then the subject will be assigned a visit response of NE (unless there is evidence of progression in which case the response will be assigned as PD). RECIST assessments/scans contributing towards a particular visit may be performed on different dates and for the central review the date of progression for each reviewer will be provided based on the earliest of the scan dates of the component that triggered the progression.

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

changes in target lesion dimensions. Endpoints (of ORR, PFS and DoR) will be derived programmatically from this information.

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

A BICR of all subjects will be performed for the database lock for PFS, which will cover all of the scans up to the DCO.

Further details of the BICR will be documented in the independent review master charter (IRMC).

3.2 Outcome variables

3.2.1 Progression free survival (PFS)

PFS is defined as the time from randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the subject withdraws from randomized therapy or receives another anti-cancer therapy prior to progression (i.e. date of PFS event or censoring – date of randomization + 1). Subjects who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST assessment. However, if the subject progresses or dies immediately after two or more missed visits, the subject will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the two missed visits. (Note: NE visit is not considered as missed visit).

Given the scheduled visit assessment scheme (i.e. eight-weekly for the first 48 weeks then twelve-weekly thereafter) the definition of 2 missed visits will change.

- If the previous RECIST assessment is less than study day 274 (i.e. week 39) then two missing visits will equate to 18 weeks since the previous RECIST assessment, allowing for early and late visits (i.e. 2 x 8 weeks + 1 week for an early assessment + 1 week for a late assessment = 18 weeks).

- 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 22 weeks (i.e. take the average of 8 and 12 weeks which gives 10 weeks and then apply same rationale, hence 2 x 10 weeks + 1 week for an early assessment + 1 week for a late assessment = 22 weeks). The time period for the previous RECIST assessment will be from study days 274 to 330 (i.e. week 39 to week 47).

- From week 48 onwards (when the scheduling changes to twelve-weekly assessments), two missing visits will equate to 26 weeks (i.e. 2 x 12 weeks + 1 week for an early assessment + 1 week for a late assessment = 26 weeks).

If the subject has no evaluable visits or does not have baseline data for the maintenance phase, they will be censored at Day 1 unless they die within 2 visits of baseline (16 weeks plus 1 week allowing for a late assessment within the visit window); then they will be treated as an event with date of death as the event date.

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

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

- For BICR assessments, the date of progression will be determined based on the **earliest** of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or of the reviewer who read baseline first if there is no adjudication for ICR data.

- For investigational assessments, the date of progression will be determined based on the **earliest** of the dates of the component that triggered the progression.

- For both BICR and investigational assessments, when censoring a subject for PFS the subject will be censored at the **latest** of the dates contributing to a particular overall visit assessment.

Note: for TLs only the latest scan date is recorded out of all scans performed at that assessment for the TLs and similarly for NTLs only the latest scan date is recorded out of all scans performed at that assessment for the NTLs.

3.2.2 Overall survival (OS)

Overall survival is defined as the time from the date of randomization until death due to any cause regardless of whether the subject withdraws from randomized therapy or receives another anti-cancer therapy (i.e. date of death or censoring – date of randomization + 1). Any subject not known to have died at the time of analysis will be censored based on the last recorded date on which the subject 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 both the PFS and OS analyses, and if subjects are confirmed to be alive or if the death date is post the DCO date these subjects will be censored at the date of DCO. The status of ongoing, withdrawn (from the study) and “lost to follow-up” subjects at the time of the final OS analysis should be obtained by the site personnel by checking the subject’s notes, hospital records, contacting the subject’s general practitioner and checking publicly-available death registries. In the event that the subject has actively withdrawn consent to the processing of their personal data, the vital status of the subject can be obtained by site personnel from publicly available resources where it is possible to do so under applicable local laws.

Note: For any OS analysis performed prior to the final OS analysis, in the absence of survival calls being made, it may be necessary to use all relevant CRF fields to determine the last

recorded date on which the patient was known to be alive for those patients still on treatment (since the SURVIVE module is only completed for patients off treatment). The last date for each individual patient is defined as the latest among the following dates recorded on the case report forms (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

3.2.3 Objective response rate (ORR)

ORR is defined as the percentage of subjects with at least one investigator-assessed visit response of CR or PR and will be based on a subset of all randomized subjects with measurable disease at randomization per the site Investigator.

Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Subjects who discontinue randomized treatment without progression, receive a subsequent anti-cancer therapy and then respond will not be included as responders in the ORR.

In addition to the unconfirmed ORR described above confirmed ORR will also be calculated. A confirmed response of CR/PR is defined as a response of CR/PR at 1 visit and confirmed by repeat imaging not less than 4 weeks after the visit when the response was first observed with no evidence of progression between the initial and CR/PR confirmation visit. Both visits contributing to a response must be prior to subsequent therapy for the subject to be considered as a responder.

In the case where a subject has two non-consecutive visit responses of PR, then, as long as the time between the 2 visits of PR is greater than 4 weeks and there is no PD between the PR visits, the subject will be defined as a responder. Similarly, if a subject has visit responses of CR, NE, CR, then, as long as the time between the 2 visits of CR is greater than 4 weeks, then a best response of CR will be assigned.

3.2.4 Duration of response (DoR)

DoR will be defined as the time from the date of first documented response after randomization until date of documented progression or death in the absence of disease progression (i.e. date of PFS event or censoring – date of first response + 1). The end of response should coincide with

the date of progression or death from any cause used for the PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR.

If a subject does not progress following a response, then their DoR will use the PFS censoring time. DoR will not be defined for those subjects who do not have documented response after randomization.

3.2.5 Best objective response (BoR)

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in Section 3.1.3. It is the best response a subject has had following randomization, but prior to starting any subsequent anti-cancer therapy and up to and including RECIST progression or the last evaluable 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 subjects with no disease at baseline for the maintenance phase), PD and NE.

For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. SD should be recorded at least 8 weeks minus 1 week, i.e. at least 49 days (to allow for an early assessment within the assessment window), after randomization. For CR/PR, the initial overall visit assessment that showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

BoR will be determined programmatically based on RECIST from the overall visit response using all site investigator data after randomization up until the first progression event or date of starting any subsequent anti-cancer therapy. The denominators for each case will be consistent with those used in the ORR analysis.

For subjects whose progression event is death, BoR will be calculated based upon all evaluable RECIST assessments after randomization and prior to death.

For subjects who die with no evaluable RECIST assessments, if the death occurs ≤9 weeks (i.e. 8 weeks + 1 week to allow for a late assessment within the assessment window) after randomization, then BoR will be assigned to the progression (PD) category. For subjects who die with no evaluable RECIST assessments, if the death occurs >9 weeks after randomization then BoR will be assigned to the NE category.

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

3.2.6 Patient-reported outcome variables

The following PROs will be administered in this study: European Organization for Research and Treatment of Cancer 30-item Core Quality of Life Questionnaire and 13-item Lung Cancer Quality of Life Questionnaire (EORTC QLQ-C30 and QLQ-LC13), CCI

CCI [redacted] CCI [redacted] see CSP for questionnaire specimens). All questionnaires will be scored according to published scoring guidelines. All PRO analyses will be based on the FAS, unless otherwise stated.

3.2.6.1 EORTC QLQ-C30 and QLQ-LC13

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce the following scales:

- 5 multi-item functional scales: physical, role, cognitive, emotional, and social
- 3 multi-item symptom scales: fatigue, pain, and nausea/vomiting
- 5 single item symptom scales: dyspnea, loss of appetite, insomnia, constipation, and diarrhea
- A 2-item Global health status/QoL scale
- 1 item on the financial impact of the disease.

The EORTC QLQ-LC13 is a lung cancer-specific module from the EORTC for lung cancer comprising 13 scales (cough, hemoptysis, dyspnea, site-specific pain, sore mouth, dysphagia, peripheral neuropathy, alopecia, and pain medication). With the exception of a 3-item scale for dyspnea, all are single items. The dyspnea scale will only be used for analysis if all 3 items have been scored; the component items will also be summarized as single-item measures.

The EORTC QLQ-C30 and EORTC QLQ-LC13 will be scored according to the EORTC QLQ-C30 scoring manual (Fayers et al 2001) and EORTC-QLQ-LC13 instructions respectively. An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales, each of the function scales, and the global health status/QoL scale in the EORTC QLQ-C30, and each of the symptom scales in the EORTC QLQ-LC13. Higher scores on the global health status and function scales indicate better health status/function, but higher scores on symptom scales represent greater symptom severity.

The global health status/HRQoL will be assessed using the EORTC-QLQ-C30 global QoL scale which includes 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 primary PRO measures will be subject-reported lung cancer symptoms assessed using the EORTC QLQ-LC13 and EORTC QLQ-C30, namely:

- QLQ-LC13: Dyspnea (multi-item scale based on three questions: “Were you short of breath when you rested; walked; climbed stairs?”)
- QLQ-LC13: Cough: one item (“How much did you cough?”)
- QLQ-LC13: Chest pain: one item (“Have you had pain in your chest?”)

- QLQ-C30: Fatigue (multi-item scale based on three questions: "Did you need rest; Have you felt weak; Were you tired?")
- QLQ-C30: Appetite loss: one item ("Have you lacked appetite?")

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change of ≥10 in the score from baseline for scales from the EORTC QLQ-C30 and QLQ-LC13 (Osoba et al 1998). For example, a clinically relevant deterioration or worsening in chest pain (as assessed by EORTC QLQ-LC13) is defined as an increase in the score from baseline of ≥10. A clinically relevant 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, the change in symptoms/ functioning from baseline will be categorized as improved, stable, or worsened as shown in Table 6.

Table 6 Visit response for symptoms and health-related quality of life

Score	Change from baseline	Visit response
EORTC QLQ-C30/QLQ-LC13 symptom scales	≥+10	Worsened
	≤-10	Improved
	Otherwise	Stable
EORTC QLQ-C30 functional and global health status/QoL scales	≥+10	Improved
	≤-10	Worsened
	Otherwise	Stable

EORTC QLQ-C30 European Organization for Research and Treatment of Cancer 30-Item core quality-of-life questionnaire; QLQ-LC13 13-Item lung cancer quality-of-life questionnaire; QoL Quality of life.

A subject's best overall response in symptoms, function, or global health status/QoL will be derived as the best response the subject achieved, based on evaluable PRO data collected during the study period. Best overall response will follow the same principles as the ORR analysis, and only responses prior to subsequent therapy will be included.

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.

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 a subsequent assessment, or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the subject withdraws from IP or receives another anticancer therapy prior to symptom deterioration. 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. Subjects with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Subjects 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. Also, if symptoms deteriorate after 2 or more missed PRO assessment visits or the subject dies after 2 or more missed PRO assessment visits, the subject will be censored at the time of the last PRO assessment where the symptom could be evaluated.

If the subject has no visits or does not have baseline data they will be censored at Day 1 unless they die within 2 visits of baseline (4 weeks plus 3 days allowing for a late assessment within the visit window).

The population for the analysis of time to symptom deterioration will include a subset of the FAS who have baseline scores of ≤90.

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 a subsequent visit or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the subject withdraws from IP or receives another anticancer therapy prior to HRQoL/function deterioration. 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. Subjects with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Subjects 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. Also, if HRQoL deteriorates after 2 or more missed PRO assessment visits or the subject dies after 2 or more missed PRO assessment visits, the subject will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated.

If the subject has no visits or does not have baseline data they will be censored at day 1 unless they die within 2 visits of baseline (4 weeks plus 3 days allowing for a late assessment within the visit window).

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.

Symptom, HRQoL and function improvement rate

The improvement rate (for symptom, function and HRQoL) will be defined as the number (%) of subjects with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement in that symptom from baseline, defined as

- a decrease from baseline score ≥10 for EORTC QLQ-C30 and QLQ-LC13 symptom scales
- an increase from baseline score of ≥10 for EORTC QLQ-C30 and QLQ-LC13 function scales and global health status/QoL

The denominator will consist of a subset of the FAS with evaluable data (i.e. who have a baseline symptom score ≥10 or a baseline global health status/QoL or functional score of ≤90 for the respective analyses).

CCI [redacted]

3.2.6.5 PRO compliance rates

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30 and QLQ-LC13, CCI [redacted] 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 e.g. a questionnaire from a subject who has not withdrawn from the study at the scheduled assessment time but excluding subjects in countries with no available translation. For subjects that have progressed, the latest of progression and safety follow-up will be used to assess whether the subject is still under HRQoL follow-up at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.
- Evaluable questionnaire = a questionnaire with a completion date and at least one subscale that is non-missing.
- Overall PRO compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.

Compliance over time will be calculated separately for each visit, including baseline, as the number of subjects with an evaluable questionnaire at the time point (as defined above), divided by number of subjects 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.

CCI [redacted]

CCI [redacted]

3.3 Safety variables

3.3.1 General considerations for safety assessments

Safety data will be analyzed separately for the initial therapy phase and the maintenance phase. For the initial therapy phase, safety analyses will be based on the SAF for the initial therapy phase; while for the maintenance phase, the SAF for the maintenance phase will be used.

For the initial therapy phase, any safety summaries by visit will only use scheduled nominal study visits. For the maintenance phase, time windows will be defined for any presentations that summarize values by visit. The following conventions will also apply:

- Baseline will generally be the last value obtained prior to the first dose of study medication in the maintenance phase. For the initial therapy phase, baseline will be the last value obtained prior to the first dose of study treatment in the initial therapy phase. Where safety data are summarized over time for each phase, study day will be calculated in relation to date of first study treatment in each phase respectively.
- The time windows will be exhaustive so that data recorded at any time point in the maintenance phase has the potential to be summarized. Inclusion within the time window will be based on the actual date and not the intended date of the visit.
- All unscheduled visit data in the maintenance phase will 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 there is an even number of days 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 (with 4 weeks between scheduled assessments) are:

  - Day 1, visit window N/A
  - Day 29, visit window 2 to 42
  - Day 57, visit window 43 to 70
  - Day 85, visit window 71 to 98
  - Day 113, visit window 99 to 126
  - Day 141, visit window 127 to 154
  - Day 169, visit window 155 to 182
  - Day 197, visit window 183 to 210
  - Day 225, visit window 211 to 238
  - Day 253, visit window 239 to 266

- Day 281, visit window 267 to 294
- Day 309, visit window 295 to 322
- Day 337, visit window 323 to 350

- For summaries based on the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- Listings will display all values contributing to a time point for a subject.
- For visit-based summaries:
    - If there is more than one value per subject within a time window then the closest value to the scheduled visit date will be summarized, or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the subject that contributed to the summary table, wherever feasible. (Note: for summaries of extreme values, all post-baseline values collected will be used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.)
    - To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group, visit data will only be summarized if the number of observations is greater than the minimum of 20 and > 1/3 of subjects dosed.
- For summaries at a subject level, all values will be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a subject level statistic such as a maximum.

3.3.1.1 Handling of missing data

The following considerations are made for missing safety data, diagnostic dates, AE dates and concomitant medications/procedures:

- 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.
- For missing diagnostic dates, if day and/or month are missing then 1st of the month and/or January will be used. If year is missing, leave the complete date as missing.
- For missing start dates for AEs or concomitant medications/procedures, the following will be applied:
    a. Missing day: Impute the 1st of the month unless month/year is same as month/year of first dose of study drug then impute first dose date (either initial therapy phase or maintenance phase depending on the month/year).
    b. Missing day and month: Impute 1st January unless year is the same as first dose date then impute first dose date (first dose date of maintenance will take the priority if the subject entered the maintenance phase).

c. Completely missing – impute first dose date (first dose date of maintenance phase will take the priority if the subject entered the maintenance phase) 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.

Note: When imputing a start date ensure that the new imputed date is sensible, i.e., it is prior to the end date of the AE or concomitant medication/procedure (otherwise set it to the end date).

- For missing stop dates of AEs or concomitant medications/procedures, the following will be applied:
  a. Missing day: Impute the last day of the month unless month/year is same as month/year of last dose of study drug then impute last dose date (either initial therapy phase or maintenance phase depending on the month/year).
  b. Missing day and month: Impute 31st December of the year unless year is the same as last dose of study drug then impute last dose date (last dose date of maintenance will take the priority if the subject entered the maintenance phase).
  c. Completely missing date for a medication: Check whether the medication is still ongoing and when it started in relation to study drug before imputing a date. If the ongoing flag is missing then assume that the medication is still being taken (i.e. do not impute a date). If the medication has stopped and its start date is prior to first dose date then impute the first dose date; if it started on or after first dose date then impute to the day after the last dose date (first dose date/last dose date of maintenance will take the priority).
  d. If an AE has a completely missing end date then it will be treated as ongoing.

  Note: When imputing a stop date ensure that the new imputed date is sensible, i.e., it is after the start date of the AE or concomitant medication/procedure (otherwise set it to the start date).

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

- Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.
- For partial subsequent anti-cancer therapy start dates, the following will be applied:
  a. Missing day: If the month/year is the same as treatment end date then impute to the day after treatment, otherwise first day of the month (last dose date of maintenance will take the priority).

b. Missing day and month: If year is the same as treatment end date then impute to the day after treatment, otherwise 1st January of the same year as anti-cancer therapy date (last dose date of maintenance will take the priority).

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

3.3.2 Exposure and dose interruptions

3.3.2.1 Maintenance phase

Exposure will be defined as follows:

Total (or intended) exposure of durvalumab

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

Total (or intended) exposure of olaparib/placebo

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

Actual exposure of durvalumab

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

Actual exposure of olaparib/placebo

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

To calculate actual exposure, dose interruptions will include those where a subject forgot to take a dose.

The duration of dose delays for durvalumab will be calculated as follows:

- Total duration of dose delays = sum of positive values of (date of the dose – date of previous dose – (28+3) days)
    - If there are no delays, the duration sums to 0, as infusions are performed every 4 weeks

The duration of dose interruptions for olaparib/placebo will be calculated as:

- Total duration of dose interruptions = sum of (end date of interruption – start date of interruption + 1)
    - If there are no interruptions, the duration sums to 0, as olaparib/placebo is administered daily

The actual exposure calculation makes no adjustment for any dose reductions that may have occurred for olaparib/placebo, however the number and proportion of subjects with dose reductions will be summarized.

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

Both the first date and last date of study medication administration will be determined from EX module where the actual dose received is greater than 0.

Safety follow-up will be defined as:

- Total safety follow-up = minimum of (90 days after the last dose of durvalumab or olaparib/placebo, whichever is later, date of withdrawal of consent, date of death, date of DCO) – first dose date in maintenance phase +1

Missed or forgotten doses

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

3.3.2.2 Initial therapy phase

Exposure in the initial therapy phase will be measured by total exposure, actual exposure, and the number of cycles received, for durvalumab and each of the chemotherapy agents. A cycle corresponds to a period of 21 days. If a cycle is prolonged due to toxicity, this will still be counted as one cycle. A cycle will be counted if treatment is started even if the full dose is not delivered.

Safety follow-up for subjects treated in initial therapy phase only will be defined as:

- Minimum of (90 days after the last dose of durvalumab, date of withdrawal of consent, date of death, date of DCO) – first dose date in initial therapy phase +1

For subjects treated in both initial therapy and maintenance phases, safety follow-up for the initial phase will be defined as:

- The day prior to first dose in maintenance phase – first dose date in initial therapy phase +1

3.3.3 Dose intensity – Maintenance Phase

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

Relative dose intensity (RDI) will be defined for durvalumab (for each treatment arm), olaparib, and placebo as follows:

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

For durvalumab, when deriving actual dose administered the volume before and after infusion will also be considered.

For olaparib, D = 300 mg x 2 x total (intended) exposure

3.3.4 Adverse events

AEs and SAEs will be collected throughout the study, from date of informed consent until the end of follow-up period, which is defined as 90 days after the last dose of durvalumab or 30 days after the last dose of olaparib/placebo, whichever is later for subjects who were treated in the maintenance phase, or 90 days after the last dose of durvalumab for subjects who were only treated in the initial therapy phase.

A treatment emergent adverse event (TEAE) for the maintenance phase is an AE with an onset date or a pre-existing AE worsening (by investigator report of a change in intensity) following the first dose of study treatment in the maintenance phase up to and including min (date of last dose of study treatment + 90 days, day before the first dose of subsequent anti-cancer therapy (including radiotherapy, with the exception of palliative radiotherapy)).

A TEAE for the initial therapy phase is an AE with an onset date or a pre-existing AE worsening following the first dose of study treatment in the initial therapy phase up to and including min (date of last dose of study treatment + 90 days, day before the first dose of subsequent anti-cancer therapy) for the initial therapy failures or prior to the first dose of the maintenance phase for the subjects treated in the maintenance phase.

Any AE occurring before any study treatment (i.e. before the administration of the first dose in the initial therapy phase) and without worsening after initial of study treatment will be referred to as ‘pre-treatment’.

The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 5.0).

For the durvalumab + olaparib arm and the durvalumab + placebo arm in the maintenance phase, in the event of the components being administered separately then date of first dose/last dose will be derived using the earliest/latest dosing date of the components.

Other significant adverse events (OAEs)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs and 'Discontinuation of Investigational Product due to Adverse Events' (DAEs). Based on the expert's judgement, significant adverse events of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered as other significant adverse events (OAEs) and reported as such in the CSR. A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs.

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

Infection Adverse events

Infection AEs will be summarized by pooled terms and PTs in two ways: (1) using MedDRA HLGT/HLT pooled terms (2) Custom pooled terms. The following summaries will be reported for both HLGT/HLT pooled terms and custom pooled terms and PTs:

- Infection AEs by CTCAE grade
- Serious Infection AEs (including event rate)

Overall Infection AE summaries will be presented, including the number and percentage of patients in each of these categories.

AEs of special interest and AEs of possible interest

Some clinical concepts (including some selected individual preferred terms and higher-level terms) have been considered "AEs of special interest" (AESI) and "AEs of possible interest" (AEPIs) to the durvalumab and AESI for the olaparib program. All AESIs are being closely monitored in clinical studies using durvalumab alone, and durvalumab in combination with other anti-cancer agents.

AEs of special interest and AEs of possible interest for durvalumab

AESIs 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). In addition, infusion-related reactions and hypersensitivity/anaphylactic reactions are also considered AESIs.

AEPIs 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 AEs not routinely arising from an inflammatory or immune-mediated mechanism of action – typically quite general clinical terms that usually present from a multitude of other causes – are classified as AEPIs.

These AESIs and AEPIs have been identified as Pneumonitis, Hepatic events, Diarrhea/Colitis, Intestinal perforations, Adrenal Insufficiency, Type 1 diabetes mellitus, Hyperthyroid events, Hypophysitis, Hypothyroid events, Thyroiditis, Renal events, Dermatitis/Rash, Pancreatic events, Myocarditis, Myasthenia gravis, Guillain-Barre syndrome, Myositis, Infusion/hypersensitivity reactions and Other rare/miscellaneous. Other categories may be added or existing terms may be merged as necessary following review by an AstraZeneca medically qualified expert. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which MedDRA preferred terms contribute to each AESI/AEPI. A further review will take place prior to Database lock (DBL) to ensure any further terms not already included are captured within the categories.

The AESIs for the study treatments can be found in Section 8.3.12 of the CSP.

Adverse events of special interest for olaparib

AESIs for olaparib are the important potential risks of:

- MDS/AML
- new primary malignancy (other than MDS/AML)
- pneumonitis

Immune-mediated Adverse Events (imAE) for durvalumab

imAEs will be identified from both AEs of special interest (AESIs) and AEs of possible interest (AEPIs) based on programmatic rules that consider interventions involving systemic steroid therapy, immunosuppressant use, and/or endocrine therapy (which, in the case of AEPIs, occurs after first considering an Investigator's causality assessment and/or an Investigator's designation of an event as immune-mediated). 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). Further details are provided in the imAE charter.

In addition, the Sponsor may perform medical review of those AESIs and AEPIs and classify them as imAEs or not imAEs via an independent manual adjudication process.

3.3.5 Laboratory data

Blood samples for determination of hematology, clinical chemistry, and TSH will be collected throughout the study, from screening to 90 days following the discontinuation of treatment. Urinalysis will be collected at screening and throughout the study as clinically indicated.

Post-baseline data obtained up until 90 days following discontinuation of durvalumab or olaparib/placebo, whichever is later, will be considered as "on-study".

Change from baseline in hematology and clinical chemistry variables will be calculated for each post-dose visit. For the definition of baseline and the derivation of post-baseline visit values considering visit windows and handling of multiple records within a visit window, derivation rules as described in Section 3.3.1 will be used.

CTC grades will be defined at each visit according to the CTC grade criteria using project ranges as required, after conversion of lab result to corresponding preferred units. The following parameters have CTC grades defined for both high and low values: Lymphocytes (absolute count) , Potassium, Sodium, Magnesium and Corrected calcium so high and low CTC grades will be calculated.

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

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

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-study value (depending on the direction of an adverse effect) will be defined for each laboratory parameter as the maximum (or minimum) post-dose value up to and including min (date of last dose of study treatment + 90 days, day before the first dose of subsequent anti-cancer therapy).

Project reference ranges will be used throughout for reporting purposes. The denominator used in laboratory summaries of CTC grades will only include evaluable subjects, i.e., those who had sufficient data for the assessment of abnormality to be performed.

For example:

- If a CTCAE criterion is based on change from baseline, then both the baseline and at least 1 post-baseline value are required for a subject to be evaluable.
- If a CTCAE criterion is not based on change from baseline,1 or more post-baseline values are required for a subject to be evaluable.

3.3.6 ECGs

ECG data will be obtained at screening and as clinically indicated throughout the study.

The following ECG variables will be collected in the eCRF: ECG mean heart rate, PR interval, QRS duration, QT interval, QTcF interval, RR interval and overall ECG evaluation of normal or abnormal.

Any clinically significant abnormalities detected require triplicate ECG results. If a QT interval corrected for heart rate using Fridericia's formula (QTcF) value >470 ms, 2 additional 12-lead ECGs should be obtained over a brief period (e.g., 30 minutes) to confirm the finding.

3.3.7 Vital signs

Vital signs (blood pressure [BP], pulse, respiratory rate, and temperature) will be collected every 4 weeks throughout the study, from screening to 90 days after the last dose of study treatment with the exception of body weight which will be collected from screening to 6 months after the last dose of study treatment.

Change from baseline in vital signs variables will be calculated for each post-baseline visit. For derivation of post-baseline visit values using visit windows and handling of multiple records within a visit window, derivation rules as described in Section 3.3.1 will be used.

3.3.8 WHO/ECOG performance status

WHO/ECOG performance status will be assessed during the study whilst subjects are receiving treatment, and also at time points that are consistent with tumor assessments post treatment discontinuation and at initiation of subsequent cancer therapy, using the following scale:

0: Fully active; able to carry out all usual activities without restrictions

1: Restricted in strenuous activity, but ambulatory and able to carry out light work or work of a sedentary nature (e.g., light housework or office work)

2: Ambulatory and capable of self-care, but unable to carry out any work activities; up and about more than 50% of waking hours

3: Capable of only limited self-care; confined to bed or chair more than 50% of waking hours

4: Completely disabled, unable to carry out any self-care, and totally confined to bed or chair

5: Dead

Any significant change from baseline for the initial therapy or maintenance phases will be reported as an AE.

3.3.9 Concomitant medication

Any medications taken by the subject at any time between the date of the first dose (including the date of the first dose) of study treatment up to the date of last dose of study treatment + 90 days in the study will be considered as concomitant medication. Any medication that started prior to the first dose of the study treatment and ended after the first dose or is ongoing will be considered as both prior and concomitant medication.

Allowed and disallowed concomitant medications will be presented by ATC classification and generic term.

3.4 Pharmacokinetic variables

The PK analyses will be performed at AstraZeneca or an appointed CRO. PK concentration data and summary statistics will be tabulated. Further exploratory analysis of PK data, if conducted, will be reported separately from the main CSR.

3.5 Immunogenicity variables

Serum samples for durvalumab antidrug antibodies ADA assessments will be conducted utilizing a tiered approach (screen, confirm, titer), and ADA data will be collected at scheduled visits shown in the CSP (Section 8.5.2). ADA result from each sample will be reported as either positive or negative. If the sample is positive, the ADA titer will be reported as well. In addition, the presence of neutralizing antibody (nAb) may be tested for all ADA-positive samples using a ligand-binding assay. The nAb results will be reported as positive or negative.

The baseline ADA result is defined as the reported result of the pre-dose initial therapy C1 sample. The number and percentage of ADA-evaluable subjects (those in the SAF for the initial therapy phase with non-missing baseline ADA result and at least 1 post-baseline ADA result in either initial therapy phase or maintenance phase) who fulfil the following criteria will be determined. A subject is defined as being ADA positive if a positive ADA result is available at any time, including baseline and all post-baseline measurements; otherwise ADA negative.

- ADA positive at any visit; the percentage of ADA-positive subjects in the ADA evaluable population is known as ADA prevalence
- Defined as either treatment-induced and treatment-boosted ADA; the percentage of subjects fulfilling this criterion in the ADA evaluable population is known as ADA incidence
- ADA positive post-baseline and positive at baseline

- ADA positive post-baseline and not detected at baseline (treatment-induced ADA)
- ADA not detected post-baseline and positive at baseline
- Treatment-boosted ADA, defined as a baseline positive ADA titer that was boosted to a 4-fold or higher level following drug administration
- Persistently positive ADA, defined as having at least 2 post-baseline ADA positive measurements with at least 16 weeks (112 days) between the first and last positive measurement or an ADA positive result at the last available assessment. The category may include subjects meeting these criteria who are ADA positive at baseline
- Transiently positive ADA, defined as having at least one post-baseline ADA positive measurement and not fulfilling the conditions for persistently positive. The category may include subjects meeting these criteria who are ADA positive at baseline
- nAb positive at any visit

3.6 Biomarker variables

3.6.1 HRR mutation status

Subject tumor samples (either newly acquired or archival [<3 years old]) will be provided at screening to determine HRR mutation status by a central laboratory using tissue-based HRR genes mutation testing.

The tumor sample will be tested for loss of function and alterations in specific HRR genes that have not yet been determined. If the test results indicate that the subject has at least 1 qualifying mutation in any of these genes, the subject will be considered HRRm.

3.6.2 PD-L1 expression status

The tumor samples will also be used to retrospectively establish the subject's pretreatment PD-L1 expression status in subjects who are enrolled into initial therapy (defined by the Ventana SP263 PD-L1 IHC assay). Subjects will be classified into 3 subgroups based on their PD-L1 expression status:

- ≥50% of tumor cells (TC) with membrane staining for PD-L1 at any intensity (PD-L1 TC ≥50%)
- 1% to 49% of TC with membrane staining for PD-L1 at any intensity (PD-L1 TC 1% to 49%)
- <1% of TC with membrane staining for PD-L1 at any intensity (PD-L1 TC <1%)

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

[redacted]

4 ANALYSIS METHODS

The primary objective is to assess the efficacy of the durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of PFS (Investigator-assessed) for subjects who are in the FAS of the maintenance phase. The formal statistical analysis will be performed to test the following main hypotheses:

- Null Hypothesis (H_{10}): No difference between durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of PFS for the FAS
- Alternative Hypothesis (H_{11}): There is a difference between durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of PFS for the FAS

If a statistically significant difference is observed in the study, i.e., reject the null hypothesis of no difference in favor of H_{11}, then the following hypotheses can also be tested:

- Null Hypothesis (H_{20}): No difference between durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of OS for the FAS
- Alternative Hypothesis (H_{21}): There is a difference between durvalumab plus olaparib combination therapy compared with durvalumab monotherapy in terms of OS for the FAS

CCI [redacted]

4.1 General principles

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

- Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, 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 based on 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.
- For PK data the geometric mean and CV will be presented to 4 significant figures (sf), minimum and maximum will be presented to 3 sf and n will be presented as an integer.
- The primary and secondary efficacy analyses (including PROs) will be performed on all subjects in the FAS for the maintenance phase, and additional secondary analyses will be performed on the HRRm subgroup of the FAS for the maintenance phase, for PFS. PK data will be summarized and analyzed based on the PK analysis set. Safety and treatment exposure data will be summarized for the SAF for both the initial therapy phase and maintenance phase. Study population and demography data will be summarized based upon the FAS for initial therapy phase and maintenance phase.
- Outputs will be summarized by treatment arm.
- SAS® version 9.4 or above will be used for all analyses.

Baseline

In general, for efficacy endpoints the last observed measurement prior to randomization will be considered the baseline measurement. However, 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. The PRO endpoints are scheduled to be collected on the first day of randomized treatment; these data will be used as baseline provided they are collected on or before the first day of study treatment in the maintenance phase.

For safety endpoints the last observation before the first dose of study treatment in the maintenance phase will be considered the baseline measurement unless otherwise specified. Similarly, the last observation before the first dose of study treatment in the initial therapy phase will be the baseline for the period. For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose. Assessments on the day of the first dose where neither time nor a nominal pre-dose indicator are captured will be considered prior to the first dose if such procedures are required by the protocol to be conducted before the first dose.

In all summaries change from baseline variables will be calculated as the post-treatment value minus the value at baseline. The percentage change from baseline will be calculated as:

= (post-baseline value - baseline value) / baseline value x 100

4.2 Analysis methods

Results of all statistical analysis will be presented using 95% CI and a 2 sided p-value, unless otherwise stated.

Table 7 details the endpoints that are subject to formal statistical analysis (including primary and secondary analysis) and pre-planned sensitivity analyses. Note: all endpoints compare durvalumab + olaparib versus durvalumab + placebo in all randomized subjects (FAS for the maintenance phase), unless otherwise indicated. The 2 stratification factors are histology (squamous or nonsquamous) and objective response to durvalumab plus chemotherapy obtained at the last visit prior to randomization (CR/PR or SD [Cycle 4 scan]). These factors will be covariates in the stratified log-rank test, logistic regression and mixed-effect model repeated measure models.

Prior to unblinding, the number of subjects across both treatment groups in each level of strata will be reviewed, and the planned stratification factors may be removed or levels may be combined if too few subjects are represented in any cell.

Table 7 Formal statistical analyses to be conducted and pre-planned sensitivity analyses

Endpoints analyzed	Notes
Progression free survival (PFS)	Primary analysis for all subjects in the FAS and a secondary analysis in the HRRm subgroup of the FAS. Analysis is conducted with a stratified log-rank test using Investigator assessment per RECIST 1.1. Sensitivity analyses for the primary analysis: 1) Analysis using BICR RECIST 1.1 assessments 2) Analyses using Investigator RECIST 1.1 assessments (i) Interval censored analysis – evaluation time bias (ii) Analysis using alternative censoring rules – attrition bias
Overall survival (OS)	Secondary analysis for all subjects in the FAS. Analyzed using a stratified log-rank test.
Objective response rate (ORR)	Secondary analysis for all subjects in the FAS. Logistic regression using Investigator assessment per RECIST 1.1.
Duration of response (DoR)	Secondary analysis for all subjects in the FAS. KM estimates using Investigator assessments per RECIST 1.1.
Time to deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Secondary analysis for all subjects in the FAS. Stratified log-rank test as per PFS analysis.
Symptom improvement rate (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Secondary analysis for all subjects in the FAS. Logistic regression.
Change from baseline in symptoms (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Secondary analysis for all subjects in the FAS. Mean change from baseline using a Mixed Model Repeated Measurements (MMRM) analysis.

Table 7 Formal statistical analyses to be conducted and pre-planned sensitivity analyses

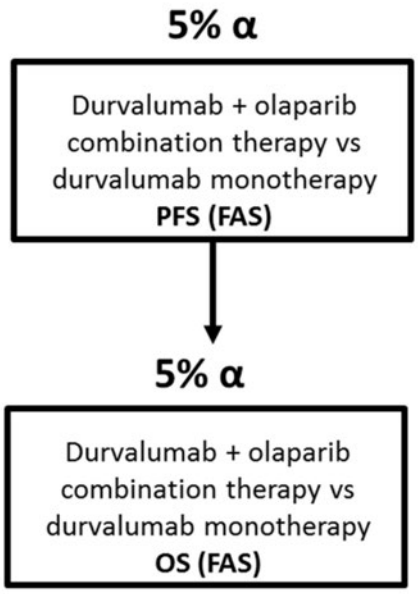
CCI [redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	

4.2.1 Multiple testing strategy

The overall 5% type I error rate will first be allocated to test the primary endpoint of PFS for durvalumab plus olaparib combination therapy versus durvalumab monotherapy (FAS). If the primary analysis of PFS is significant, 5% alpha will be recycled to the lower level in the hierarchy, where the 5% alpha will be used for the test of OS for durvalumab plus olaparib combination therapy versus durvalumab monotherapy (FAS) (5% alpha).

As shown in Figure 2, the primary endpoint of PFS (FAS) will be tested at 1 primary analysis of PFS. The key secondary endpoint, OS (FAS), will be tested at 2 time points: 1 interim analysis and 1 final analysis. The test at the primary analysis of PFS (FAS) will be considered as 1 test family, as will the tests for the interim analysis and final analysis of OS. As long as 1 test in the family can be rejected, the family is rejected. Thus, the assigned total alpha to the family will be recycled to the next MTP level. Details of the interim analyses are provided in Section 5.

Figure 2 Multiple Testing Procedure



FAS Full analysis set; OS Overall survival; PFS Progression free survival

4.2.2 Primary endpoint: Progression free survival

An analysis of the primary endpoint PFS will occur when it is expected that approximately 163 PFS events have occurred (65% maturity). PFS will be based on the programmatically derived RECIST 1.1 using investigator data.

The analysis will be performed for subjects in the FAS using a stratified log-rank test adjusting for objective response to durvalumab plus chemotherapy in the initial therapy phase (CR/PR versus SD) and histology (squamous or nonsquamous) for generation of the p-value, and using a method that corresponds to the Breslow approach for handling ties (Breslow 1974).

The model will include these effects regardless of whether the inclusion of effects significantly improves the fit of the model.

The effect of durvalumab + olaparib versus durvalumab + placebo will be estimated by the HR together with its 95% CI from a stratified Cox model (an HR less than 1 will favor durvalumab in combination with olaparib). The CI will be calculated using a profile likelihood approach. The stratified Cox model will be fitted using PROC PHREG (in SAS) with the Efron method to control for ties and the strata variables included in the strata statement.

The response to initial therapy and histology covariates used in the statistical modelling will be based on the values reported in the IxRS at randomization.

KM plots of PFS will be presented by treatment arm, by treatment arm and response to initial therapy, and by treatment arm and histology.

Summaries of the number and percentage of subjects experiencing a PFS event and type of event (RECIST 1.1 or death) will be provided along with the median PFS and 95% CI for each treatment.

The assumption of proportionality will be assessed. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) versus 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 will 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 using the approach of Gail and Simon (Gail and Simon 1985).

The PFS analysis described above will be repeated for a subset of subjects in the FAS with HRRm.

Sensitivity Analyses

The analysis of PFS as assessed by BICR in the FAS will be performed as a sensitivity analysis using the same methodology as specified for PFS as assessed by the site Investigator in the FAS.

Sensitivity analyses will also 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 evaluable RECIST assessment will be analyzed using a stratified log-rank test. For subjects whose death was treated as 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 even in highly asymmetric assessment schedules (Sun and Chen 2010). To support this analysis, the mean of subject-level average inter-assessment times will be tabulated for each treatment. This approach will use the Investigator RECIST assessments.

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of subjects who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, subjects who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a KM plot of the time to censoring where the censoring indicator of the PFS analysis is reversed. This approach will use the Investigator RECIST 1.1 assessments.

Disagreements between investigator and central reviews of RECIST 1.1 progression will be presented for each treatment group. The summary will include the early discrepancy rate which is the frequency of central review declared progressions before the investigator review as a proportion of all central review progressions and the late discrepancy rate which is the

frequency of central review declared progressions after the investigator review as a proportion of all discrepancies.

As stratification variables will be defined according to data from the IxRS, if there are a sufficient number of subjects who are mis-stratified, a sensitivity analysis may be carried out using the baseline data collected in the eCRF.

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.

4.2.2.1 Additional supportive summaries/graphs

The treatment status at progression of subjects at the time of analysis will be summarized for the FAS. This will include the number (%) of subjects who were on treatment at the time of progression, the number (%) of subjects who discontinued study treatment prior to progression, the number (%) of subjects 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 subjects who have discontinued treatment.

The number of subjects prematurely censored will be summarized by treatment arm. A subject is defined as prematurely censored if the subject had not progressed and the latest scan prior to DCO was more than 1 scheduled tumor assessment interval (+2 weeks) prior to the DCO date.

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

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

Additionally, summary statistics for the number of days between RECIST assessments will be presented by treatment groups. Summary statistics for the number of weeks between the time of progression and the last evaluable RECIST assessment prior to progression will be presented for each treatment group for subjects who progress.

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

In addition, a summary of new lesions (i.e. sites of new lesions) will be produced.

Subgroup Analyses

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1 using Investigator assessments) between durvalumab plus olaparib combination therapy versus durvalumab monotherapy in the following subgroups of the FAS (but not limited to):

- Sex (male or female)
- Age at study entry (<65 or ≥65 years of age) as recorded in DM module

- PD-L1 status (<1%, 1% to 49%, ≥50%, unknown)
- Histology (squamous or nonsquamous)
- Objective response to initial therapy (CR/PR or SD)
- Smoking (smoker [current or former] or non-smoker [never smoker])
- Race (Asian or non-Asian)
- HRRm status (yes, no, unknown)
- Investigator's choice of chemotherapy (cisplatin doublet versus carboplatin doublet; nab-paclitaxel doublet versus pemetrexed doublet versus gemcitabine doublet)
    - For subjects who switched between cisplatin and carboplatin, subjects will be included in the subgroup corresponding to the first therapy received

The subgroup analyses for the stratification factors (specifically, histology and objective response to initial therapy) will be based on the values reported in the eCRF. Other baseline variables may also be assessed if there is clinical 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 and/or predictive factors. If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events across both treatment groups in a subgroup), the relationship between that subgroup and PFS will not be formally analyzed. In this case, only descriptive summaries will be provided.

For each subgroup, the HR (durvalumab + olaparib vs. durvalumab + placebo) and 95% CI will be calculated from a Cox proportional hazards model with treatment as the only covariate. These will be presented on a forest plot including the HR and 95% CI from the overall population.

The primary interpretation in the HRRm positive subgroup will be based on the stratified log-rank test as specified for the primary analysis of PFS.

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.

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).

Effect of covariates on HR estimate

Cox proportional hazards modelling will be employed to assess the effect of pre-specified covariates on the HR estimate for the treatment comparisons of the primary endpoint based on the FAS. Before embarking on more detailed modelling, an initial model will be constructed

containing treatment and the stratification factors alone to ensure that any output from the Cox modelling is likely to be consistent with the results of the stratified log-rank test.

The results from the initial model and the model containing additional covariates will be presented.

Additional covariates for this model will be sex, age at study entry, PD-L1 status at study entry, histology (squamous vs nonsquamous) based on eCRF data, objective response to initial therapy based on eCRF data, smoking status, race, HRRm status and investigators choice of chemotherapy (cisplatin doublet vs carboplatin doublet). The model will include the effect regardless of whether the inclusion of effect significantly improves the fit of the model providing there is enough data to make them meaningful.

4.2.3 Overall survival

OS in the FAS will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. The effect of durvalumab plus olaparib combination therapy versus durvalumab monotherapy will be estimated by the HR together with its corresponding two-sided 95% CI from a stratified Cox model. KM plots will be presented by treatment arm.

Summaries of the number and percentage of subjects 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 and 95% CI for each treatment.

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

In addition, subgroup analyses will be conducted comparing OS between durvalumab plus olaparib combination therapy versus durvalumab monotherapy in the subgroups defined for the primary PFS analysis.

For each subgroup, the HR (durvalumab + olaparib vs. durvalumab + placebo) and 95% CI will be calculated from a Cox proportional hazards model with treatment as the only covariate. These will be presented on a forest plot including the HR and 95% CI from the overall population.

A summary of the duration of follow-up will be summarized for all subjects as well as for censored subjects only, presented by treatment group.

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

4.2.4 Objective response rate

The ORR will be based on the programmatically derived RECIST 1.1 assessment using the Investigator tumor data. The ORR will be compared between durvalumab + olaparib and durvalumab + placebo using logistic regression models adjusting for the same factors as the primary endpoint PFS (response to initial therapy and histology). The results of the analysis will be presented in terms of an odds ratio (an odds ratio of greater than 1 will favor the durvalumab and olaparib combination therapy over the durvalumab and placebo) together with its associated profile likelihood CI (using the option 'LRCI' in SAS PROC GENMOD) and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model).

Summaries will be produced that present the number and percentage of subjects with a tumor response (CR/PR). For each treatment arm, best overall response (BoR) will be summarized by n (%) for each category (CR, PR, SD, PD, NED, and NE). No formal statistical analyses are planned for BoR.

The ORR analysis described above will be repeated for a subset of subjects in the FAS with HRRm.

Summaries of the number and percentage of subjects achieving a confirmed objective response will be presented.

Overall tumor response as reported by investigators at the last visit in the initial therapy phase will be summarized descriptively as number of subjects and corresponding percentages for each response category (CR, PR, SD, PD, and NE).

4.2.5 Duration of response

KM estimates will be provided for the DoR in responding subjects (i.e., median DoR and 95% CIs) by treatment arm, including the associated KM curves (without any formal comparison of treatment arms or p-value attached).

The DoR analysis described above will be repeated for a subset of subjects in the FAS with HRRm.

4.2.6 Patient-reported outcomes

PRO analyses will be conducted for the FAS.

Compliance rates summarizing questionnaire completion at each visit will be tabulated. By visits summaries will use visits windows defined in Section 3.3.1.

4.2.6.1 EORTC QLQ-C30 and QLQ-LC13

Change from baseline

Summaries of absolute and change from baseline values for symptoms, function, and HRQoL will be reported for the all symptom scales, the 5 functional scales, and global health status/QoL scale by visit for each treatment arm.

The mean change from baseline will be analyzed for the primary PRO symptoms of interest (QLQ-LC13 Dyspnea, LC13 Cough, LC13 Chest pain, C30 Fatigue, and C30 Appetite loss).

Change from baseline in these pre-specified PRO symptom scores of (QLQ-LC13 Dyspnea, LC13 Cough, LC13 Chest pain, C30 Fatigue, and C30 Appetite loss) 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 or 20 subjects in a visit in any treatment arm). It is acknowledged that subjects 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 in order to include the discontinuation and follow up visits, a generic visit variable will be derived for each subject 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 in order to select only those visits occurring within the first 12 months of randomization or until PD. This will follow the rules for assigning visit windows in Section 3.3.1.

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

	Study Day	
Generic visit	Subject X	Subject Y
Baseline	Baseline	Baseline
1	29	28
2	57	50 (discontinuation)
3	85	85
4	113	113
5	130 (discontinuation)	141
6	169	169

The MMRM model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, response to initial therapy and histology 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 95% CIs will be presented along with an estimate of the treatment difference, 95% CI 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.

A plot will be produced of adjusted mean change from baseline against time, with treatment group identified within the plot. The corresponding 95% CIs for each time point will be overlaid.

Time to deterioration

Time to deterioration in symptom, function, and HRQoL will be analyzed for the following functional and symptom scales:

- EORTC QLQ-C30: Global health status/QoL, Physical, role, cognitive, emotional, and social functioning
- EORTC QLQ-C30: fatigue, nausea/vomiting, pain, dyspnea, insomnia, loss of appetite, constipation, and diarrhea
- EORTC QLQ-LC13: Dyspnea, Coughing, Hemoptysis, Pain in chest, Pain in arm or shoulder, Pain in other parts

For treatment comparison, the methods used will be as described in the primary PFS analysis and illustrated using a KM plot by treatment arm. For each of the analyses, time to deterioration will be presented using a Kaplan-Meier plot. Summaries of the number and percentage of subjects experiencing a clinically meaningful deterioration or death, as well as who were censored, and the median time to deterioration will also be provided for each treatment arm.

The hazard ratio, p-value, and 95% CI will be presented graphically on a forest plot, for all indicated subscales.

Summaries of the number and percentage of subjects in each response category at each visit for each ordinal item (in terms of the proportion of subjects in the categories of improvement, no change, and deterioration as defined in Section 3.2.6.1) will also be produced for each treatment arm.

Improvement rate

A summary of the improvement rate for symptoms, function, and HRQoL will be generated for the following functional and symptom scales:

- EORTC QLQ-C30: Global health status/QoL, Physical, role, cognitive, emotional, and social functioning
- EORTC QLQ-C30: fatigue, nausea/vomiting, pain, dyspnea, insomnia, loss of appetite, constipation, and diarrhea
- EORTC QLQ-LC13: Dyspnea, Coughing, Hemoptysis, Pain in chest, Pain in arm or shoulder, Pain in other parts

The improvement rate for the above scales will be analyzed by comparing the treatment arms using a logistic regression model as described for the analysis of ORR. The results of the analysis will be presented in terms of an odds ratio together with its associated profile likelihood 95% 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 odds ratio and 95% CI for each scale/item will be presented graphically on a forest plot. If there are very few responses in one treatment arm, a Fisher's exact test will be considered.

4.2.6.2CCI

CCI [redacted]

[redacted]

[redacted]

4.2.8 Safety analyses

Safety and tolerability will be presented overall for the initial therapy phase and by treatment arm for the maintenance phase.

A brief overview of safety data during the initial therapy phase will be summarized using the SAF for the initial therapy phase. This includes summaries and listings of AEs and exposure to durvalumab.

For the maintenance phase SAF, data from all cycles will be combined in the presentation of safety data. AEs will be summarized by number and percentage of subjects experiencing each AE and by rate of AEs per person-years at risk. Other safety data will be assessed in terms of physical examination, serum chemistry, hematology, urinalysis, vital signs, and ECGs. Exposure to each study treatment, time on study, dose delays or dose interruptions, and dose reductions will also be summarized.

Safety data for the initial therapy phase will be summarized for the period including the on-treatment + follow-up (follow-up defined as 90 days after the last dose of durvalumab) and prior to start of subsequent therapy for the initial therapy failures or on-treatment up to the first dose of the maintenance phase for the subjects treated in the maintenance phase.

Unless otherwise stated, safety data in the maintenance phase will be summarized for the 'on-study' period (on-treatment + follow-up), which is defined as from the first dose in the maintenance until 90 days after the last dose of durvalumab or olaparib/Placebo, whichever is later.

Safety data will be summarized only. No formal statistical analyses will be performed on the safety data.

The following sections describe the planned safety summaries. However, additional safety tables may be required to aid interpretation of the safety data.

4.2.8.1 Adverse Events

All AEs, both in terms of current MedDRA preferred term and CTCAE grade, will be listed and summarized descriptively by count (n) and percentage (%). The AE summaries, unless otherwise stated, will be based on treatment-emergent AEs in each treatment phase up until the

initiation of the first subsequent anti-cancer therapy following discontinuation of treatment or until the end of follow-up period (whichever occurs first). This will more accurately depict AEs attributable to study treatment only as a number of AEs up to the end of the follow-up period 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 the end of the follow-up period (i.e. without taking subsequent anti-cancer therapy into account).

Initial therapy phase

Summary information (the number and percent of subjects by system organ class and preferred term) will be tabulated for:

- All AEs
- All SAEs
- AEs leading to discontinuation of any study drug
- AEs with outcome of death (only include subjects who were not randomized into maintenance phase)

The summary will also be presented by maximum reported CTCAE grade, system organ class and preferred term.

Maintenance phase

Summary information (the number and percent of subjects by system organ class and preferred term separated by treatment group) will be tabulated for:

- All AEs (repeated separately to include AEs following any subsequent therapy)
- All AEs possibly related to durvalumab only (as determined by the reporting investigator)
- All AEs possibly related to olaparib/placebo only (as determined by the reporting investigator)
- All AEs possibly related to either study medication (as determined by the reporting investigator)
- Most common AEs (occurring in at least 5% of subjects)
- AEs with CTCAE grade 3 or 4
- AEs with CTCAE grade 3 or 4, possibly related to durvalumab only (as determined by the reporting investigator)
- AEs with CTCAE grade 3 or 4, possibly related to olaparib/placebo only (as determined by the reporting investigator)

- AEs with CTCAE grade 3 or 4, possibly related to either study medication (as determined by the reporting investigator)
- Most common AEs with CTCAE grade 3 or 4 (occurring in at least 1% of subjects)
- AEs with outcome of death
- AEs with outcome of death possibly related to durvalumab only (as determined by the reporting investigator)
- AEs with outcome of death possibly related to olaparib/placebo only (as determined by the reporting investigator)
- AEs with outcome of death possibly related to either study medication (as determined by the reporting investigator)
- All SAEs
- AEs leading to discontinuation of durvalumab only
- AEs leading to discontinuation of olaparib/placebo only
- AEs leading to discontinuation of either study medication
- AEs leading to discontinuation of durvalumab only, possibly related to durvalumab only (as determined by the reporting investigator)
- AEs leading to discontinuation of olaparib/placebo only, possibly related to olaparib/placebo only (as determined by the reporting investigator)
- AEs leading to discontinuation of either study medication, possibly related to either study medication (as determined by the reporting investigator)
- AEs leading to dose interruption of durvalumab only
- AEs leading to dose interruption of olaparib/placebo only
- AEs leading to dose interruption of either study medication
- Infusion reaction AEs (as determined by the reporting investigator)
- AEs ongoing from initial therapy

Summaries of other significant AEs may be produced.

Multiple events per subject will not be accounted for apart from on the episode level summaries.

An overall summary of the number and percentage of subjects in each of the above categories will be presented.

A truncated AE table of most common AEs and another table showing most common AEs with CTCAE grade 3 or 4, showing all events that occur in at least 5% and 1% of subjects overall, respectively, will be summarized by preferred term, by decreasing frequency. This cut-off may be modified after review of the data. When applying a cut-off (i.e., x %), the raw percentage

will be compared to the cut-off and no rounding will be applied first (i.e., an AE with frequency of 4.9% will not appear if a cut-off is 5%).

Each AE event rate (per 100 patient years) and SAE event rate will also be summarized by preferred term within each system organ class. For each preferred term, the event rate is defined as the number of subjects with at least 1 event during the treatment period plus the follow-up period (or until the initiation of the first subsequent therapy following discontinuation of treatment) divided by the total treatment duration only (excluding the follow up period, in days), summed over subjects and then multiplied by 365.25 x 100 to present in terms of per 100 patient years.

Summaries of the number and percentage of subjects with AEs will be provided by maximum reported CTCAE grade, system organ class, preferred term and treatment group.

In addition, all AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug for the initial therapy phase and the maintenance phase respectively. Pre-treatment AEs and AEs that occur after a subject has received further therapy for cancer (following discontinuation of IP) will be included in the AE listings. A separate data listing of AEs occurring after the end of follow-up period will also be produced.

4.2.8.2 Deaths

An overall summary of all deaths and all deaths on-treatment or within the follow-up period will be produced for subjects in the FAS for the maintenance phase.

For all deaths, the summary will include:

- Death related to disease under investigation only
- Death related to disease under investigation and an AE on treatment or within 90 days follow-up with an outcome of death
  - AE onset prior to subsequent therapy
  - AE onset after start of subsequent therapy
- AE on treatment or within 90 days follow-up with an outcome of death only
  - AE onset prior to subsequent therapy
  - AE onset after start of subsequent therapy
- Death after end of follow-up period and not due to disease under investigation or not a TEAE
- Other deaths

For all deaths on-treatment or within the follow-up period, the summary will include:

- Death related to disease under investigation only

- Death related to disease under investigation and an AE on treatment or within 90 days follow-up with an outcome of death
  - AE onset prior to subsequent therapy
  - AE onset after start of subsequent therapy
- AE on treatment or within 90 days follow-up with an outcome of death only
  - AE onset prior to subsequent therapy
  - AE onset after start of subsequent therapy
- Other deaths

A listing of all deaths will also be produced for subjects who were treated in the initial therapy phase only and for subjects in the FAS for the maintenance phase, respectively.

4.2.8.3 AEs of special interest and AEs of possible interest (AESI/AEPI)

Preferred terms will be used to identify AESIs/AEPIs for durvalumab and olaparib (see Section 3.3.4 for the list).

Preferred terms of AESI's or AEPI's will be identified before DBL and documented in the Study Master File. Summary tables of grouped MedDRA preferred terms will be produced. For each grouped term, the number (%) of subjects experiencing any of the specified terms will be presented by maximum CTCAE grade.

For the maintenance phase, summaries by grouped term and preferred term will be presented separately for durvalumab AESI/AEPI and olaparib AESI as below:

- All AESI
- Serious AESI

Summaries by grouped term and maximum reported CTCAE grade, by grouped term and outcome of events, will be presented respectively, for durvalumab and olaparib AESI/AEPI .

An overall summary of AEPI/AESI for durvalumab will be presented with the following categories:

- Any AESI/AEPI
- Any AESI/AEPI of CTCAE Grade 3 or 4
- Any serious AESI/AEPI (including events with outcome of death)
- Any AESI/AEPI with outcome of death
- Any AESI/AEPI, causally related to study treatment
- Any AESI/AEPI of CTCAE Grade 3 or 4, causally related to study treatment
- Any serious AESI/AEPI, causally related to study treatment

- Any AESI/AEPI with outcome of death, causally related to study treatment
- Any AESI/AEPI leading to concomitant medication use (system coricosteroids)
- Any AESI/AEPI leading to concomitant medication use (high dose steroids)
- Any AESI/AEPI leading to concomitant medication use (endocrine therapy)
- Any AESI/AEPI leading to concomitant medication use (other immunosuppressants)
- Any AESI/AEPI leading to discontinuation of study treatment

4.2.8.4 Immune mediated Adverse Events

Programmatically-generated immune mediated adverse events will be presented. Details of the programmatically generated immune mediated adverse event summaries will be confirmed before database lock.

The imAEs (as classified by the Sponsor) will also be summarized in the same manner as for the summaries for AESI/AEPI described above. See further details in the imAE Charter with respect to derivation rules.

4.2.8.5 Laboratory assessments

Specific outputs will be produced for Hy's Law, ALT, AST and total bilirubin as outlined in the Global Safety SAP and TA Standard SAS output.

Data collected "on study" or until the initiation of the first subsequent therapy following discontinuation of treatment or end of follow-up period (whichever occurs first) will be used for reporting. This will more accurately depict laboratory toxicities attributable to study treatment only as a number of toxicities up to 90 days following discontinuation of durvalumab or olaparib/placebo, whichever is later, are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, summaries of laboratory data may also be produced containing data collected up until the end of follow-up period (i.e. without taking subsequent therapy into account). Any data collected after the end of follow-up period will not be summarized.

Laboratory data (hematology, clinical chemistry, and TSH parameters) will be summarized over time in terms of absolute values and change from baseline at each scheduled measurement by actual treatment group. Data summaries will be provided in preferred units.

Additionally, a summary table of the CrCl level and change in CrCl from baseline will be presented by treatment arm and visit. A shift plot of the baseline CrCl versus the minimum observation on treatment will be presented by subject.

Shift tables for laboratory values by worst CTC grade will be produced, and for specific parameters separate shift tables indicating hyper- and hypo- directionality of change from

baseline will be produced. The laboratory parameters for which CTC grade shift outputs will be included but not limited are:

- Hematology: Hemoglobin (hypo), Leukocytes, Lymphocytes, absolute count-hypo and hyper, Neutrophils, absolute count, 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, Creatinine, amylase, lipase

For the parameters with no applicable CTCAE grading that are listed in the CSP, shift tables from baseline to worst value may be provided.

Only laboratory data summary over time for observed values and change from baseline at each scheduled measurement will be presented for the initial therapy phase.

4.2.8.6 Liver Enzyme Elevations and Potential Hy's law

The following summaries will include the number (%) of subjects 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.)

Narratives will be provided in the CSR for subjects who have ALT ≥ 3x ULN plus total bilirubin ≥ 2x ULN or AST ≥ 3x ULN plus total bilirubin ≥ 2x ULN at any visit.

Liver biochemistry test results over time for subjects with elevated ALT or AST (i.e. ≥ 3x ULN), and elevated total bilirubin (i.e. ≥ 2x ULN) (at any time) will be plotted. Individual subject 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 maximum ALT and AST vs. maximum total bilirubin by treatment group will also be produced with reference lines at 3xULN for ALT, AST, and 2xULN for total bilirubin. In each plot, total bilirubin will be in the vertical axis.

4.2.8.7 ECGs

Overall evaluation of ECG is collected at screening visits in terms of normal or abnormal, and the relevance of the abnormality is termed as "clinically significant" or "not clinically significant". A shift table of baseline evaluation to worst evaluation will be produced if there is sufficient data.

ECG data up to the date of last dose of study medication + 30 days will be included in the summary table.

4.2.8.8 Vital signs

Vital signs data up to the date of last dose of study medication + 30 days will be included in the summary tables.

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.

4.2.8.9 Time to Subsequent Therapy from discontinuation of study treatment

Descriptive summaries will be produced for time to subsequent therapy from discontinuation of study treatment (the latest of either durvalumab or olaparib/placebo) . These summaries are supportive of the adverse event and laboratory data outputs.

4.2.8.10 WHO/ECOG performance status

World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will be summarized over time for the SAF. A shift table of baseline evaluation to worst evaluation will be produced.

WHO/ECOG data up to the date of last dose of study medication + 90 days will be included in the summary table.

4.2.9 Pharmacokinetic analyses

Summaries of PK concentration data of durvalumab will be provided for all evaluable subjects in the PK analysis set.

4.2.10 Immunogenicity analyses

A summary of the number and percentage of subjects in the entire study who develop detectable ADA to durvalumab by ADA categories (Section 3.5) will be presented based on the ADA evaluable set. The summary will be repeated for the following groups of subjects:

- SAF subjects by treatment group
- SAF subjects for initial therapy phase only

Immunogenicity results will be listed for subjects in the SAF in each of the groups listed above regardless of ADA-evaluable status. ADA titer and nAb data will be presented for samples confirmed positive for the presence of ADA to durvalumab. AEs in ADA positive subjects by ADA positive category will be listed.

The effect of immunogenicity on PK, efficacy, and safety will be evaluated, if data allow.

4.2.11 Demographic and baseline characteristics data

The following will be summarized for all subjects in the FAS for the initial therapy phase:

- Subject disposition (reason for initial therapy treatment discontinuation; number of subjects randomized)
- Demographics (age at study entry, age group [<50, ≥50 - <65, ≥65 - <75 and ≥75 years] at study entry, sex, race, and ethnicity)
- Inclusion in analysis sets

The following will be summarized for all subjects in the FAS in the maintenance phase (unless otherwise specified), by treatment group:

- Subject disposition (reason for maintenance treatment discontinuation; reason for study discontinuation)
- Important protocol deviations
- Inclusion in analysis sets
- Stratification factors from IxRS versus source data (response to initial therapy and histology)
- PD-L1 expression status (<1%, 1% to 49%, ≥50%, unknown) at study entry
- HRRm status (yes, no, unknown) at study entry
- Investigator's choice of chemotherapy in the initial therapy phase (cisplatin doublet versus carboplatin doublet; nab-paclitaxel doublet versus pemetrexed doublet versus gemcitabine doublet)
- Demographics (age at study entry, age group [<50, ≥50 - <65, ≥65 - <75 and ≥75 years] at study entry, sex, race, and ethnicity)
- Subject characteristics at baseline (height [measured prior to initial therapy], weight [measured prior to randomization])
- Subject recruitment by region, country and center
- Previous disease-related treatment modalities
- Previous disease-related treatments prior to this study
- Disease characteristics at initial diagnosis (TNM classification, histology type and IASLC Stage) and baseline at randomization (WHO/ECOG performance status)
- Extent of disease at study entry
- Medical history (past and current)
- Surgical history

- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use, categorised (never, current, former) and number of pack years

The medications will be coded following AZ standard drug dictionary/WHO Drug dictionary as applicable.

4.2.12 Treatment exposure

For the initial therapy phase, total exposure and actual exposure, dose delays and interruptions, reasons for dose delays and interruptions, and total number of cycles received will be summarized for durvalumab. In addition, number and percentage of subjects who received any chemotherapy agents, cumulative doses and number of cycles received will be summarized.

For the maintenance phase, the following summaries related to study treatment will be produced by actual treatment group:

- Time from randomization to first dose of durvalumab, and to first dose of olaparib/placebo, by treatment group
- Total exposure to durvalumab and olaparib/placebo by treatment group
- Actual exposure of durvalumab and olaparib/placebo by treatment group
- Total number of durvalumab cycles received by each treatment group
- Number of, reasons for, and duration of dose delays for durvalumab (per treatment arm) and interruptions for olaparib/placebo, respectively. Dose interruptions will be based on investigator initiated dosing decisions. In addition, interruptions due to AEs and due to reasons other than AEs will be summarized separately. Number of and reason for dose reduction for olaparib/placebo will also be summarized.
- RDI (relative dose intensity) of durvalumab (per treatment arm) and olaparib/placebo
- Cumulative doses and number of cycles received for chemotherapy agents during the initial therapy phase

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

4.2.13 Subsequent Therapy

Subsequent therapies received after discontinuation of study treatment in the maintenance phase will be summarized by treatment group.

CCI [redacted]

CCI [redacted]

CCI [redacted]

4.2.16 Listings

In addition to listings mentioned within the analysis text, selected data will be listed, meeting the requirements of ICH E3, i.e.:

- Discontinued subjects (Appendix 12.2.1 in the AZ CSR)
- Protocol deviations (Appendix 12.2.2 in the AZ CSR)
- Subjects excluded from the efficacy analysis (Appendix 12.2.3 in the AZ CSR)
- Demographic data (Appendix 12.2.4 in the AZ CSR)
- Study drug administration and PK concentration data of durvalumab (Appendix 12.2.5 in the AZ CSR)
- Individual tumor assessment data per investigator (Appendix 12.2.6 in the AZ CSR)
- Individual efficacy response data including PFS by BICR (Appendix 12.2.6 in the AZ CSR)
- Overall survival status (Appendix 12.2.6 in the AZ CSR)
- Adverse events (AEs): all AEs and serious adverse events (SAEs) (Appendix 12.2.7 in the AZ CSR)
- Listing of individual laboratory measurements by subject

4.2.17 Coronavirus Disease 2019 (COVID-19)

Depending on the extent of any impact, summaries and listings of data relating to subjects diagnosed with COVID-19, and impact of COVID-19 on study conduct (in particular missed visits, delayed or discontinued IP, and other protocol deviations) may be generated. In addition, PFS and OS sensitivity analyses may be performed by repeating the summaries and analyses except that any patient who had a death with primary/secondary cause as being COVID-19 related (including infection) reported as fatal will be censored at their COVID related death date. For AE and Deaths summaries of COVID-19 related events including infections and deaths maybe produced.

5 INTERIM ANALYSES

5.1 Analysis Methods

A primary analysis for PFS will occur at approximately 163 events. CCI [redacted]

CCI [redacted]

There will be 1 interim analysis performed for OS. The OS interim will occur at the time of the primary analysis for PFS. It is anticipated that approximately 67% of the OS events will be available for this OS IA (approximately 109 of 163 OS events). CCI [redacted]

At the time of the primary analysis of PFS/interim analysis of OS the study will be unblinded and all endpoints will be reported. Subjects will continue to be followed up for survival until approximately 163 subjects have had an OS event at which time the final OS analysis would be performed.

5.2 Independent data monitoring committee

The safety of all AstraZeneca clinical studies is closely monitored on an ongoing basis by AstraZeneca representatives in consultation with Patient Safety. Issues identified will be addressed; for instance, this could involve amendments to the Clinical Study Protocol and letters to Investigators.

An IDMC composed of independent experts will be established to perform an assessment of the safety of durvalumab + olaparib combination therapy in this population on an ongoing basis. The committee will meet at a frequency outlined in the IDMC Charter. . Safety reviews will be carried out by the IDMC in an unblinded manner.

After review of the unblinded data, the IDMC will recommend whether the study should continue unchanged, be stopped, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to AstraZeneca/MedImmune. The report will include the recommendation and any potential protocol amendments and will not contain any unblinding information. The final decision to modify or stop the study will sit with the sponsor.

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

6 CHANGES OF ANALYSIS FROM PROTOCOL

There are no changes made to the analysis specified in the protocol.

7 REFERENCES

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

CCI [redacted]

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

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

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;16(1):139-44.

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;97-108.

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.

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: d9102c00001-sap-ed-3		
Document Title:	Statistical Analysis Plan Edition 3	
Document ID:	Doc ID-004193930	
Version Label:	3.0 CURRENT LATEST APPROVED	
Server Date **(dd-MMM-yyyy HH:mm 'UTC'Z)**	**Signed by**	**Meaning of Signature**
21-Dec-2020 14:25 UTC	PPD	
21-Dec-2020 17:28 UTC	PPD	

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