A Randomised, Multicentre, Double-blind, Placebo-controlled, Phase III Study of First-line Carboplatin and Paclitaxel in Combination with Durvalumab, Followed by Maintenance Durvalumab with or without Olaparib in Patients with Newly Diagnosed Advanced or Recurrent Endometrial Cancer (DUO-E)

PPD **(AstraZeneca)**	PPD	Date

TABLE OF CONTENTS **PAGE**

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AESI	Adverse event of special interest
AML	Acute myeloid leukaemia
APTT	Activated partial thromboplastin time
BICR	Blinded independent central review
BoR	Best objective response
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CR	Complete response
CRF	Case report form
CRO	Contract research organisation
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Event;
ctDNA	Circulating tumour DNA
DBL	Database lock
DCO	Data cut-off
DoR	Duration of response
eCRF	Electronic case report form
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer - Quality of Life Questionnaire Core 30
EORTC QLQ-EN24	European Organisation for Research and Treatment of Cancer - Quality of Life Questionnaire – Endometrial Cancer Module
EQ-5D-5L	EuroQoL five dimensions, five level
FAS	Full Analysis Set
FIGO	Federation Internationale de Gynecologie et d'Obstetrique
HLA-LOH	human leukocyte antigen – loss of heterozygosity
HR	Hazard ratio
HRD	homologous recombination deficiency

Abbreviation or special term	Explanation
HRR	homologous recombination repair
HRRm	homologous recombination repair mutation
HRQoL	Health related quality of life
ICH	International Conference on Harmonisation
ICR	Independent central review
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
INR	International normalised ratio
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
LD	Longest diameter
lsmean	Least squares mean
LSR	Last subject randomised
MDS	Myelodysplastic syndrome
MMR	Mismatch repair
MMRM	Mixed-effect model for repeated measures
MRI	Magnetic resonance imaging
MSI	Microsatellite instability
NA	Not applicable
NE	Not evaluable
NTL	Non-target lesion
ORR	Objective response rate
OS	Overall survival
QAPFS	Quality adjusted progression free survival
QTWiST	Quality-adjusted time without significant symptoms of toxicity
PD	Progressive disease
PD-L1	programmed death ligand 1
PGI-BR	patient global impression of benefit/risk
PGIC	patient global impression of change

Abbreviation or special term	Explanation
PGIS	patient global impression of severity (in symptoms)
PGI-TT	patient global impression of treatment tolerability
PFS	Progression free survival
PFS2	Second progression free survival
PK	Pharmacokinetics
PP	Per-protocol
PR	Partial response
PRO	Patient reported outcome
PRO-CTCAE	Patient reported outcomes version of Common Terminology Criteria for Adverse Event
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
REML	Restricted maximum likelihood
SAE	Serious adverse event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SAS®	A commercially available integrated system of software products, commonly used for reporting and analysis of Clinical Studies
SD	Stable disease
TCR	T-cell receptor
TDT	Time to study treatment discontinuation or death
TFST	Time to first subsequent therapy or death
TIL	Tumour-infiltrating lymphocyte
TMB	Tumour mutational burden
TL	Target lesion
TNM	Tumour, Node, Metastasis
TSST	Time to second subsequent therapy or death
ULN	Upper limit of normal
VAS	Visual analogue scale

AMENDMENT HISTORY

Version 4.0, 24 Nov 2022

Category: Change refers to	Date	Description of change	In line with CSP? Y (version)/ N/ NA	Rationale
Derivation of primary or secondary endpoints	24 Nov 2022	The planned tumour predictive biomarker subgroup analysis has been updated to include HRRm status (HRRm vs non-HRRm vs Unknown), replacing the previously specified HRD status subgroup analysis (Sections 4.2.2.2 and 4.2.11)	Y (v5)	Within the endometrial cancer setting, cut-off(s) for tumour HRD status measure of genomic instability relevant tests are not analytically or clinically defined. Instead, this subgroup analysis will focus on mutations in genes involved in the HRR pathway (termed HRR genes).
Statistical analysis method for the primary or secondary endpoints	24 Nov 2022	CA-125 and History of Debulking Surgery have been removed from the planned subgroup analysis and interaction test. FIGO stage has been removed from the interaction test (Section 4.2.2.2)	NA	Removed subgroup analysis and interaction test by CA-125 and History of Debulking Surgery status given these are not key prognostic factors for this patient population. Removed FIGO stage from the interaction test as this is only summarised in Newly Diagnosed patients and not the entire ITT population on which the Global interaction test is performed.
Other	24 Nov 2022	Correction made to the estimated timing of the second interim analysis for OS, which is predicted to occur 51 months after *first* patient randomised (but was incorrectly stated to occur 51 months after *last* patient randomised) (Section 1.3)	N	Typographical error

	24 Nov 2022	Clarification that summaries of the number and percentage of patients with AEs leading to dose modifications will be presented by treatment arm (Section 4.2.10.2)	Y (v5)	To align with CSP
	24 Nov 2022	Clarification that for EORTC QLQ-EN24, there are a total of 13 scales/scores (and not 11).	Y (v5)	Typographical error

Version 3.0, 29 June 2022

Category: Change refers to	**Date**	**Description of change**	**In line with CSP? Y (version)/ N/ NA**	**Rationale**
Derivation of primary or secondary endpoints	29 June 2022	Objectives updated to include the PFS comparison of Arm C (durvalumab in combination with platinum-based chemotherapy followed by maintenance durvalumab in combination with olaparib) versus Arm A (platinum-based chemotherapy) as primary. (Sections 1.1, 4.2, 4.2.2, 4.2.3 and 4.2.4)	Y (v5)	To enable the two key PFS comparison of interest (ie, Arm B versus Arm A and Arm C versus Arm A) to be tested independently.
Statistical analysis method for the primary or secondary endpoints	29 June 2022	The formal statistical analysis method, sample size assumptions, and timing of primary analysis of PFS was revised. The multiplicity testing strategy has been amended to enable the two key PFS comparisons of interest (ie, Arm B versus Arm A, and Arm C versus Arm A) to be tested independently with an equal alpha split of 2.5% assigned to each comparison of interest. In addition, based on external data, an assumed 3 month lag-effect for durvalumab, when	Y (v5)	To incorporate the impact of the lag effect due to evolving data on durvalumab (and other immunotherapies agents) from external trials in multiple tumour settings, including CASPIAN (NCT03043872; Paz Ares et al 2019), POSEIDON (NCT03164616; Johnson 2021) and TOPAZ-1 (NCT03875235; Oh et al 2022), which have demonstrated the potential for a lag/delay

		administered in combination with chemotherapy, has been taken into account. (Sections 1.3, 4.0 and 4.2.1)		in the treatment effect of immunotherapies when administered in combination with chemotherapy.
	29 June 2022	PFS futility boundary and analysis timepoint updated. (Section 5.1)	Y (v5)	The futility boundary and timing are updated to ensure sufficient data maturity and follow up after recruitment completion, to mitigate against the increased risk of incorrectly stopping the study for futility, given the potential for a lag/delay in the treatment effect.

Version 2.0, 01 February 2022

Category: Change refers to	**Date**	**Description of change**	**In line with CSP? Y (version)/ N/ NA**	**Rationale**
Derivation of primary or secondary endpoints	01 Feb 2022	Clarified that safety analysis set summaries in patients who were dosed with investigational product will be summarised by randomised treatment group. (Section 2.1)	Y (v3)	To clarify reporting of safety data by treatment groups and to clarify analysis set definitions to be applied to support reporting.
	01 Feb 2022	CT and MRI are the only acceptable methods for RECIST assessment of target lesions within this study (Section 3.1.1)	Y (v1)	Clinical examination is not a valid method of assessment for this study
	01 Feb 2022	Specify how OS censor date should be determined in the absence of survival calls (Section 3.2.2)	NA	To clarify which CRF fields will be used to determine last recorded date on which patients were known to be alive
	01 Feb 2022	Included definition of olaparib AESIs and durvalumab AEPI/AESIs and imAEs (Sections 3.5.3 and 4.2.10.3)	NA	To clarify TEAE, AESI/AEPI and imAE definitions.

	01 Feb 2022	Subgroup analysis categories updated to include "Unknown" where applicable (Section 4.2.2.2)	NA	To update subgroup analysis categories to include "Unknown" where applicable
	01 Feb 2022	PD-L1 cut points defined as a tumour area positivity score ≥1% (Section 4.2.2.2)	NA	To specify PD-L1 cut points for the subgroup analysis
	01 Feb 2022	Definition of intended exposure of carboplatin and paclitaxel updated to include substitute carboplatin/paclitaxel (Section 3.5.1)	NA	To include exposure of carboplatin substitute and paclitaxel substitute within the exposure summaries
Statistical analysis method for the primary or secondary endpoints	01 Feb 2022	Treatment discontinuation is defined as all investigational products being discontinued (Section 3.2.8)	NA	Clarification to the definition of treatment discontinuation for the purpose of analysis of TDT
	01 Feb 2022	Clarification that stratification factors will be incorporated within the Cox PH model as strata rather than as covariates (Sections 4.2.2 and 4.2.3)		To align with AZ standards
Data presentations	01 Feb 2022	Updated the exposure summaries of chemotherapy agents and durvalumab/placebo to present number of infusions instead of number of cycles (Section 3.5.1)	NA	To specify that number of infusions will be summarised for chemotherapy agents and durvalumab/placebo
	01 Feb 2022	Family history of cancer removed from the list of data to be summarised (Section 4.2.11)	NA	This data will be utilised for creating patient narratives. Aggregate summaries are not required
	01 Feb 2022	Description of patient disposition table updated to align with AZ standard outputs(Section 4.2.11)	NA	Updated to align with how data is collected and presented
	01 Feb 2022	Safety follow-up visit added to list of visit windows (Section 4.2.10.1)	NA	To clarify the visit window for safety assessment conducted after treatment discontinuation
Other	01 Feb 2022	Defined analyses and timings specifically for the China cohort (Section 2 and Section 7)	Y (v4)	To support regulatory submissions in China.

	01 Feb 2022	Updated wording of important protocol deviations to clarify requirements and updated criteria (Section 2.2)	NA	To clarify wording and expand upon the list of important PDs that could affect the safety of patients or efficacy of IP.
	01 Feb 2022	Updated to use Cycle 7 instead of Week 18 as the starting point of durvalumab dose frequency change (Section 3.5.1)	NA	To clarify that durvalumab dose frequency change starts at Cycle 7, not necessarily at week 18 because protocol allows delay of dosing
	01 Feb 2022	Outlined additional analyses to be performed as a result of the COVID-19 pandemic (Sections4.1 and 9.5)	NA	To comply with AstraZeneca requirements for clinical studies ongoing during the COVID-19 pandemic
	01 Feb 2022	Update definition of safety follow-up period throughout to reference 'last dose' rather than 'treatment discontinuation' (Sections 4.2.10.2, 4.2.10.4 and 4.2.12)	NA	To clarify using the last dose date, not treatment discontinuation date, to determine safety follow-up period

1. STUDY DETAILS

This statistical analysis plan (SAP) contains a more detailed description of the analyses in the clinical study protocol (CSP) for the DUO-E study (D9311C00001). This SAP is based on version 5.0 of the CSP.

1.1 Study Objectives

The primary objective of this study is to determine the efficacy of:

- Arm C (durvalumab+olaparib) versus Arm A (control)
- Arm B (durvalumab+placebo) versus Arm A (control)

in patients with newly diagnosed advanced or recurrent endometrial cancer. The primary objectives will be assessed by analysis of progression-free survival (PFS), and supported by additional time to event related endpoints.

The pharmacokinetics and immunogenicity of durvalumab will also be investigated.

A summary of all study objectives is given below:

Primary objective	**Endpoint/variable**
To demonstrate the efficacy of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab (Arm B) or durvalumab with olaparib (Arm C) when compared to platinum-based chemotherapy alone (Arm A) by assessment of progression-free survival (PFS), in patients with newly diagnosed advanced or recurrent endometrial cancer	PFS (per RECIST 1.1 as assessed by investigator), defined as the time from randomisation until the date of objective disease progression or death (by any cause in the absence of progression). This will be assessed via determining the efficacy of: • Durvalumab in combination with platinum based chemotherapy and continued as maintenance in combination with olaparib versus standard of care (SoC) platinum-based chemotherapy. • Durvalumab in combination with platinum based chemotherapy and continued as maintenance versus SoC platinum-based chemotherapy.
Secondary objectives	**Endpoint/variable**
To determine the efficacy of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab (Arm B) or durvalumab with olaparib (Arm C) when compared to platinum-based chemotherapy alone (Arm A) in newly diagnosed advanced or recurrent endometrial cancer patients by assessment of: PFS2, OS, ORR, DoR, TFST, TSST, and TDT	PFS2: Second progression-free survival is defined as the time from randomisation to the earliest of progression event subsequent to first subsequent therapy (assessed by the investigator per local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression), or death due to any cause. OS: Overall survival is defined as the time from the date of randomisation until death due to any cause. ORR: Objective response rate is the proportion of patients with measurable disease at baseline who have complete response (CR) or partial response (PR), as determined by the investigator at local site. DoR: Duration of response is time from the date of

	first documented response until date of documented progression or death in the absence of disease progression. TFST: Time to first subsequent therapy or death is time from randomisation to the earlier of start date of the first subsequent anti-cancer therapy after discontinuation of randomised treatment or death due to any cause. TSST: Time to second subsequent therapy or death is time from randomisation to the earlier of start date of the second subsequent anti-cancer therapy after discontinuation of first subsequent treatment or death due to any cause. TDT: Time to study treatment discontinuation or death is time from randomisation to the earlier of the date of study treatment discontinuation or death
To characterise the PK and immunogenicity of durvalumab and durvalumab in combination with olaparib	Serum concentrations of durvalumab Anti-drug antibodies (ADA) to durvalumab
To determine effects on symptoms, functioning and overall health-related quality of life (HRQoL) of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab (Arm B) or durvalumab with olaparib (Arm C) when compared to platinum-based chemotherapy alone (Arm A) in newly diagnosed advanced or recurrent endometrial cancer patients	Change from baseline in: • Physical functioning score of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 30 (EORTC QLQ-C30) • Role functioning score of the EORTC QLQ-C30 • Global health status/quality of life (QoL) score of the EORTC QLQ-C30 • All other functioning and symptom subscale scores of the EORTC QLQ-C30 (excluding the financial subscale) Time to deterioration in: • Physical functioning score of the EORTC QLQ-C30 • Role functioning score of the EORTC QLQ-C30 • Back/pelvic pain of the EORTC QLQ-EN24 • Gastrointestinal (GI) symptoms of the EORTC QLQ-EN24 • Urological symptoms of the EORTC QLQ-EN24
Safety objectives	**Endpoint/variable**
To evaluate the safety and tolerability of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab (Arm B) or durvalumab with olaparib (Arm C) compared to platinum-based chemotherapy alone (Arm A) in newly diagnosed advanced or recurrent endometrial cancer patients	Safety and tolerability will be evaluated in terms of AEs/serious AEs (SAEs), physical examination, vital signs including blood pressure, pulse, clinical laboratory including clinical chemistry/haematology parameters, and ECG Assessments related to AEs cover: • Occurrence/frequency • Relationship to investigational product (IP) as assessed by investigator • Common Terminology Criteria for Adverse Event (CTCAE) grade • Seriousness

	• Death • Discontinuation of IP • Dose modifications during the chemotherapy phase and the maintenance phase • AEs of special interest (AESIs) • Other significant AEs • Exposure • Immune-mediated adverse events (imAEs) – given the intended mechanisms of action of durvalumab, particular attention will be given to AEs that may follow enhanced T cell activation, or other imAE
Exploratory objectives	**Endpoint/variable**
To determine the efficacy of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab in combination with olaparib (Arm C) when compared to durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab (Arm B) in patients with newly diagnosed advanced or recurrent endometrial cancer	Will include, but is not limited to: • PFS (per RECIST 1.1 as assessed by investigator) • OS
To evaluate tumour predictive biomarkers of durvalumab and olaparib in advanced endometrial cancer patients [a]	Will include, but is not limited to the following measurements within the tumour: • Tumour tissue mismatch repair (MMR), microsatellite instability (MSI), tumour mutational burden (TMB) and PD-L1 status. • Mutation status of homologous recombination repair (HRR) genes and homologous recombination deficiency (HRD) score or other genomic scar of homologous recombination deficiency.
To evaluate additional tumour candidate predictive biomarkers of durvalumab and olaparib in advanced endometrial cancer patients [a]	May include, but is not limited to: • CD3+/CD8+ tumour-infiltrating lymphocyte (TIL) densities, Human leukocyte antigen – loss of heterozygosity (HLA-LOH), immune gene expression profiling and other exploratory biomarkers.
To further assess the efficacy of treatment through longitudinal analysis of blood samples collected at regular intervals on study [a]	May include but is not limited to: • Circulating tumour DNA (ctDNA) response to treatment. • Peripheral gene expression profiling, assessment of peripheral chemokines and cytokines, T-cell receptor (TCR) profiling.
To explore whether resistance mechanisms to treatment can be identified through analysis of tumour and blood samples – archival tumour sample and blood samples at baseline and on progression (tumour sample optional on progression) [a]	Analysis and outcome variables yet to be defined but may include molecular analysis of ctDNA.
Future exploratory research into factors that may influence development of cancer and/or response to study treatment (where response is defined broadly to include efficacy, tolerability or safety) may be	Analysis and outcome variables yet to be defined.

performed on the collected and stored blood or archival tumour samples that were mandatory for entry onto the study or on optional blood or tumour biopsy samples collected during the course of the study. [a]	
To collect and store DNA (according to each country's local and ethical procedures) for future exploratory research into genes/genetic variation that may influence response (i.e., distribution, safety, tolerability and efficacy) to study treatments and or susceptibility to disease (optional) [a]	To identify pharmacogenetic correlates for the response to treatment through the retrospective analysis of DNA extracted from an optional blood sample.
To explore health status of patients with durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab with or without olaparib in patients with advanced or recurrent endometrial cancer	Evaluation of health status by the assessment of • Health state utility derived from the EuroQoL five dimensions, five level health state utility index (EQ-5D-5L) • Quality-adjusted time without symptoms of disease or toxicity (Q-TWiST) [a] • Quality-adjusted PFS (QAPFS) [a]
To explore patient-reported treatment tolerability with durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab with or without olaparib in patients with advanced or recurrent endometrial cancer	Evaluation of selected symptoms from the patient-reported outcomes version of the CTCAE (PRO-CTCAE) and overall treatment tolerability using the patient global impression of treatment tolerability (PGI-TT).
To explore patient-reported severity of cancer symptoms, change in overall health condition, and overall benefit/risk evaluation for durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab with or without olaparib in patients with advanced or recurrent endometrial cancer	Evaluation of patient global impression of severity of cancer symptoms (PGIS), patient global impression of change in health condition (PGIC), and overall perception of benefit/risk (PGI-BR).
To explore healthcare resource associated with durvalumab and olaparib in advanced endometrial cancer patients	Key healthcare resource use will be collected using HOSPAD

[a] Results of these exploratory objectives may be presented outside the clinical study report.

1.2 Study Design

This is a randomised, double-blind, placebo controlled, multicentre Phase III study to evaluate the efficacy and safety of durvalumab in combination with platinum-based chemotherapy (carboplatin and paclitaxel) followed by maintenance durvalumab with or without olaparib in patients with advanced or recurrent endometrial cancer compared to platinum-based chemotherapy. Patients will undergo tumour biomarker assessment to determine the MMR status prior to main screening. After tissue for MMR testing has been shipped, patients can enter into the main screening period. Patients who signed the informed consent and who meet all of the inclusion criteria and none of the exclusion criteria will be eligible for randomisation.

Eligible patients will be randomised in a 1:1:1 ratio on Cycle 1 Day 1 to the following 3 treatment arms:

Arm A (control): Platinum-based chemotherapy (paclitaxel and carboplatin) with durvalumab placebo (IV) during the chemotherapy phase. Following completion of the chemotherapy phase, patients without objective disease progression will receive durvalumab placebo (IV) and olaparib placebo (tablets) in the maintenance phase until disease progression.

Arm B (Durvalumab+Placebo): Platinum-based chemotherapy (paclitaxel and carboplatin) with durvalumab (IV) during the chemotherapy phase. Following completion of the chemotherapy phase, patients without objective disease progression will receive durvalumab (IV) and olaparib placebo (tablets) in the maintenance phase until disease progression.

Arm C (Durvalumab+Olaparib): Platinum-based chemotherapy (paclitaxel and carboplatin) with durvalumab (IV) during the chemotherapy phase. Following completion of the chemotherapy phase, patients without objective disease progression will receive durvalumab (IV) and olaparib (tablets) in the maintenance phase until disease progression.

The randomisation scheme will be stratified according to:

Tumour tissue's mismatch repair (MMR) status: Patients with MMR deficient tumours versus those with proficient tumours.

Disease status: Patients with recurrent disease versus those newly diagnosed.

Geographic region: Asia versus rest of the world (RoW).

The overall schedule of the study is as follows:

1. Screening period
2. On- treatment period. The on-treatment period includes the following 2 phases:
    a. Chemotherapy phase.
    b. Maintenance phase.
3. Follow-up period.

1.3 Number of Subjects

Approximately 699 eligible endometrial cancer patients will be randomised globally at a 1:1:1 ratio to the study treatments.

The sample size was derived using the validated statistical software for the design, simulation and monitoring of clinical trials, EAST™ v6 and a validated non-proportional hazards-based AstraZeneca R-package. The sample size calculations were based on the following assumptions:

- The assumed median PFS of 12 months for the control arm is in line with data reported for carboplatin/paclitaxel in first-line endometrial cancer from the GOG-209 study (Miller et al 2012), as reported in the GOG-86P manuscript (Aghajanian et al 2018).

- The sample size has been derived on the assumption of a 3-month delay in separation of the PFS curves between Arm B versus Arm A and between Arm C versus Arm A. The assumed true average hazard ratio for the Durvalumab+Placebo arm is 0.70 (corresponding to an improvement in median PFS of 5.5 months over the assumed median PFS of 12 months in the control arm) and for the Durvalumab+Olaparib arm is 0.55 (corresponding to an improvement in median PFS of 11.2 months).

The data cut off for the primary analysis of PFS for the two comparisons of interest (Arm B versus Arm A and Arm C versus Arm A) will be undertaken at the same calendar time when approximately 299 PFS events have occurred (64% maturity) for the comparison of the Durvalumab+Placebo arm versus the control arm and approximately 281 PFS events have occurred (60% maturity) for the comparison of the Durvalumab+Olaparib arm versus the control (approximately 43 months after the first patient is randomised).

If the average true PFS HR is 0.70 for the Durvalumab+Placebo arm versus the control arm, the study will provide 80% power to demonstrate a statistically significant difference for PFS with overall 2-sided significance level of 2.5%; this translates to a 5.5 month benefit in median PFS over 12 months on the control arm. The smallest treatment difference that would be statistically significant is an HR of 0.77.

If the average true PFS HR is 0.55 for Durvalumab+Olaparib versus Control, this will provide >99% power to demonstrate a statistically significant difference for PFS with an overall 2-sided significance level of 2.5%; this translates to a 11.2 month benefit in median PFS over 12 months on the control arm. The smallest treatment difference that would be statistically significant is an HR of 0.76.

In addition, the sample size has been derived on the following assumptions:

- 27-month period of recruitment
- Approximately 10% uniform dropout rate over the study period.

The power calculations for OS were based on the following assumptions:

- Median OS of 22.7 months for the control arm
- The sample size has been derived on the assumption of a 3-month delay in separation of the OS curves between Arm B versus Arm A and between Arm C versus Arm A, hence the use of an average hazard ratio for OS. The assumed true average OS hazard ratio is 0.75 for the Durvalumab+Placebo arm versus Control arm and Durvalumab+Olaparib arm versus Control arm comparisons corresponding to an improvement in median OS of approximately 7.9 months over the assumed median OS of 22.7 months in the control arm.

The first interim analysis of OS will be performed at the time of the primary PFS analysis, based on the same date of data cut-off (DCO). For the comparison of the Durvalumab+Placebo arm versus the control arm, as well as Durvalumab+Olaparib versus

Control, it is anticipated that 74% of the target number of OS events will have occurred at this time (i.e., approximately 208 of 280 OS events per comparison).

A further analysis of OS may be performed at the same calendar time when approximately 244 OS events (87% of the target number of OS events) have occurred for the comparison of Durvalumab+Placebo vs the control arm, as well as Durvalumab+Olaparib vs the control arm, approximately 51 months after first patient is randomised.

A final analysis of OS may be performed at the same calendar time when approximately 280 OS events have occurred (60% maturity) for the comparison of the Durvalumab+Placebo arm versus the control arm, as well as the Durvalumab+Olaparib arm versus the control arm, approximately 63 months after the first patient is randomised. If the average true OS HR is 0.75 for the comparison of the experimental arm versus Control, the study will provide 55% power to demonstrate a statistically significant difference for OS with overall 2-sided significance level of 2.5%; this translates to a 7.9-month benefit in median OS over 22.7 months on the control arm . The smallest treatment difference that would be statistically significant is an HR of 0.76. Note that these estimates are based on the assumption that no confounding will occur.

For the OS comparisons, the alpha allocation for the secondary OS endpoints will be controlled at the interim and/or the final analysis timepoints separately for each treatment comparison by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates the O'Brien-Fleming approach, where the significance level applied at the interim analysis depends upon the proportion of information (i.e., information fraction) available.

2. ANALYSIS SETS

Note, Global recruitment to the study will close when approximately 699 patients are randomised. If necessary, enrolment in China will continue after global recruitment is closed (i.e. last subject randomised from a non-Chinese site) to allow inclusion of a China cohort consisting of approximately 129 randomised patients. The China cohort will support standalone safety and efficacy analyses of the patients from sites in China (please see Section 7 or details).

All populations and planned analyses described, relate to the Global population unless otherwise stated. A patient randomised in China prior to global recruitment closure will be included in both the Global ITT population and the China cohort ITT population. A patient randomised in China after the global recruitment closure will be included only in the China cohort ITT population.

2.1 Definition of Analysis Sets

The All Patients population refers to all enrolled patients, defined as those who signed the informed consent form.

In addition, four analysis sets will be defined:

- Full Analysis Set (FAS)
- Safety Analysis Set (SAF)
- Pharmacokinetic (PK) Analysis Set
- ADA Analysis Set

Full Analysis Set (FAS)

The full analysis set will include all randomised patients with treatment groups assigned in accordance with the randomisation, regardless of the treatment actually received. Patients who are randomised but do not subsequently receive treatment are included in the FAS. The analysis of data using the FAS therefore follows the principles of intent to treat. The FAS will be used to analyse efficacy data (including PROs), baseline characteristics data as well as Biomarker data, and patients will be summarised based on the treatment arm they are randomised to regardless of the treatment they actually receive.

Safety Analysis Set (SAF)

The safety analysis set will consist of all randomised patients who received any amount of study treatment (durvalumab/placebo or olaparib/placebo). Safety data will not be formally analysed but summarised using the SAF.

Patients who initially received a dose of durvalumab/placebo will be summarised according to the arm to which they were randomised. This is in order to provide a summary of the underlying safety profile that patients should expect when initially prescribed treatment (ie, SoC, SoC + durvalumab, or SoC + durvalumab + olaparib).

Maintenance Phase: Number of SAF patients who entered the maintenance phase (defined as receiving at least one dose of olaparib/olaparib placebo) will be summarised and additional summaries of safety by treatment phase may also be generated, as required.

PK Analysis Set

All patients who receive at least 1 dose of durvalumab per the protocol 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.

ADA Analysis Set

The ADA evaluable subjects are patients in the SAF who received at least 1 dose of durvalumab and have non-missing baseline ADA and at least 1 post-baseline ADA result.

Table 1 presents the summary of outcome variables and the analysis sets for these variables.

Table 1 Summary of outcome variables and analysis sets

Outcome variable	Analysis set
Efficacy data	
PFS	FAS
PFS2, OS, ORR*, DoR*, TFST, TSST, TDT	FAS
PROs, and symptom/HRQoL endpoints (including EORTC QLQ-C30 and EORTC QLQ-EN24, EQ-5D-5L, PGI-TT, PGIC, PGIS, PGI-BR and PRO-CTCAE)	FAS
Study Population/Demography Data	
Disposition	All patients set
Demography characteristics (e.g. age, sex etc.)	FAS
Baseline and disease characteristics	FAS
Protocol deviations (protocol deviations and important deviations)	FAS
Medical/surgical history	FAS
Previous anti-cancer therapy	FAS
Concomitant/previous medications/procedures	FAS
Subsequent anti-cancer therapy	FAS
PK Data (durvalumab only)	
PK data	PK
Immunogenicity Data (durvalumab only)	
Immunogenicity (ADA) data	ADA
Safety data	
Exposure	SAF
Adverse events	SAF
Laboratory measurements	SAF
Vital signs	SAF
ECGs	SAF
Biomarker data [a]	
Biomarkers	FAS

*Patients who are evaluable for the analysis of ORR are those with measurable disease at baseline. Patients who are evaluable for the analysis of DoR are those who responded in the ORR analysis.

[a] Biomarker data may be reported separately outside the CSR.

2.2 Protocol Deviations

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

- Patients randomised but who did not receive any study treatment.
- Patients who deviate from key entry criteria per the Clinical Study Protocol (CSP). (Inclusion criteria 1, 2b (main ICF), 3, 4, 5, 6, 7, 11 and 12; Exclusion criteria 5, 6, 12, 13, 15, 16 and 17)
- Baseline RECIST scan > 42 days before randomisation.
- No baseline RECIST 1.1 assessment on or before date of randomisation.
- Received prohibited concomitant medications (including other anti-cancer therapy or chronic use of immunosuppressive medications). Please refer to the CSP section 6.5.2 Table 7 for those 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 of all medications prior to database lock.
- Patients who received the investigational treatment at an incorrect dose or received an alternative investigational treatment to that which they were randomised.
- Patients who met discontinuation criteria but continued to be on study treatment

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

The important protocol deviations will be listed and summarised using FAS. The first above deviation will lead to exclusion from the SAF. None of the other deviations will lead to patients being excluded from the analysis sets described in section 2.1 (with the exception of the PK analysis set, a patient will be excluded from PK analysis if there is any deviation that is considered to impact PK analysis). A per-protocol analysis excluding patients with specific important protocol deviations is not planned; however, a 'deviation bias' sensitivity analysis may be performed on the progression free survival endpoint excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients in either of treatment group:

- Did not have the intended disease or indication or
- Did not receive any randomised therapy

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

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

Further details on the classification can be found in the study-specific protocol deviations list.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST Visit Responses

This section details the implementation of protocol-specific requirements in the assessment of Response Evaluation Criteria In Solid Tumours (RECIST) version 1.1 guidelines (Eisenhauer et al 2009) with regards to the per visit and overall Investigator assessment of tumour burden for this study.

For all patients, the RECIST tumour response data will be used to determine each patient's visit response according to RECIST 1.1 with minor modifications. It will also be used to determine if and when a patient has progressed in accordance with RECIST and their best objective response to study treatment. The modifications of the published RECIST guidelines include the removal of clinical examination and ultrasound as valid modalities to evaluate Target Lesions (TL), Non-Target Lesions (NTL) or new lesions, and the rule to identify new lesions on FDG-PET scans.

Baseline radiological tumour assessments are to be performed no more than 28 days before the start of randomised treatment and ideally as close as possible to the start of study treatment. Tumour assessments are then performed every 9 weeks for the first 18 weeks post randomisation and then every 12 weeks until disease progression.

If an unscheduled assessment is performed, and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits. This schedule is to be followed in order to minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

From the investigator's review of the imaging scans, the RECIST tumour response data will be used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD or PD, using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a patient has had a tumour assessment that cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response will be assigned as PD).

Please refer to Table 2 and Table 3 for the definitions of CR, PR, SD and PD for target and non-target lesions, respectively.

RECIST outcomes (i.e., PFS, ORR, etc.) will be calculated programmatically for the site investigator assessments (see section 3.2) from the overall visit responses.

Patients with measurable or non-measurable disease or no evidence of disease assessed at baseline by computed tomography (CT) / magnetic resonance imaging (MRI) will be entered in this study.

3.1.1 Target lesions (TLs) – site investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD), (except lymph nodes which must have short axis ≥ 15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. A patient can have a maximum of five measurable lesions recorded at baseline with a maximum of two lesions per organ (representative of all lesions involved and suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded, then measurements from the one that is closest and prior to randomisation will be used to define the baseline sum of TLs. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement. In such circumstance the next largest lesion, which can be measured reproducibly, should be selected.

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

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

For patients with no disease at baseline (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.

Table 2 TL Visit Responses (RECIST 1.1)

Visit Responses	Description
Complete response (CR)	Disappearance of all TLs since baseline. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <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.

Visit Responses	Description
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5 mm, taking as reference the smallest previous sum of diameters (nadir) since treatment started. It includes the baseline sum of diameters if that is the smallest on study.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.
Not evaluable (NE)	Only relevant in certain situations (i.e. if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit; and scaling up could not be performed for lesions with interventions). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response.
Not applicable (NA)	No TLs are recorded at baseline.

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 one decimal place before assigning a TL response. For example, 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%

Missing TL data

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

- A new lesion is recorded
- 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 longest diameter recorded.

Lymph nodes

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

TL visit responses subsequent to CR

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

- Step 1: If all lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node LD increases by 20% but remains < 10mm.
- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to NE irrespective of whether, when referencing the sum of TL diameters, the criteria for PD are also met.
- Step 3: If not all lesions meet the CR criteria (i.e. a pathological lymph node selected as TL has short axis > 10mm or the reappearance of previously disappeared lesion) or a new lesion appears, then response will be set to PD
- Step 4: If after steps 1 – 3 a response can still not be determined the response will be set to remain as CR

TL too big to measure

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

TL too small to measure

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

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolization), should be handled in the following way. Once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumours:

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

- Step 2: If there was no evidence of progression after step 1, treat the lesion diameter (for those lesions with intervention) as missing and if ≤ 1/3 of the TLs have missing measurements then scale up as described in the 'Scaling' section below. If the scaling results in a visit response of PD then the patient would be assigned a TL response of PD.
- Step 3: If, after both steps, PD has not been assigned, then, if appropriate (i.e. if ≤ 1/3 of the TLs have missing measurements), the scaled sum of diameters calculated in step 2 should be used, and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10 mm for lymph nodes) and the lesions that have been subject to intervention have a value of 0 (or <10 mm for lymph nodes) recorded. If scaling up is not appropriate due to too few non-missing measurements then the visit response will be set as NE.

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

Scaling (applicable only for irradiated lesions/lesion intervention)

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

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

Example of scaling

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

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

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

68 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 should be summed and reported on the eCRF as the LD for the lesion that split.

Lesions that merge

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

Change in method of assessment of TLs

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

If a change in method involves clinical examination (e.g. CT or MRI changes to clinical examination), any affected lesions should be treated as missing.

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

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

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

Table 3 NTL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all NTLs since baseline. All lymph nodes must be non-pathological in size (<10 mm short axis).
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.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: For patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met.
Not applicable (NA)	Only relevant if there are no NTLs at baseline.

CR = Complete response; PD = Progression of disease; NE = Not evaluable; NTL = Non-target lesion; TL = Target lesion.

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

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

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

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

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

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

If the question 'Any new lesions since baseline' has not been answered with Yes or No and the new lesion details are blank this is not evidence that no new lesions are present, but should not overtly affect the derivation.

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

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

3.1.3 Overall visit response – site investigator data

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

Table 4 Overall visit responses

Target lesions	Non-target lesions	New Lesions	Overall Visit response
CR	CR or NA	No (or NE)	CR
CR	Non-CR/Non-PD or NE	No (or NE)	PR

Target lesions	Non-target lesions	New Lesions	Overall Visit response
PR	Non-PD or NE or NA	No (or NE)	PR
SD	Non-PD or NE or NA	No (or NE)	SD
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD
NE	Non-PD or NE or NA	No (or NE)	NE
NA	CR	No (or NE)	CR
NA	Non-CR/Non-PD	No (or NE)	SD
NA	NE	No (or NE)	NE
NA	NA	No (or NE)	NED

CR = Complete response; PR = Partial response; SD = Stable disease; PD = Progression of disease; NE = Not evaluable; NA = Not applicable (only relevant if there were no target and/or non-target lesions at baseline); NED = No evidence of disease.

3.1.4 Independent review

A planned BICR of all radiological imaging data will be carried out using RECIST v1.1. All radiological scans for all patients (including those at unscheduled visits, or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca appointed Contract Research Organisation (CRO) for central analysis. The imaging scans will be reviewed by two independent radiologists using RECIST v1.1 and will be adjudicated, if required (i.e. two reviewers' review the scans and adjudication is performed by a separate reviewer in case of a disagreement). For each patient, the BICR will define the overall visit response (i.e. the response obtained overall at each visit by assessing TLs, NTLs and new lesions) and no programmatic derivation of visit response is necessary (For patients with TLs at baseline: CR, PR, SD, PD, NE; for patients with NTLs only: CR, SD, PD, NE; for patients with no disease identified at baseline: PD, no evidence of disease [NED], NE). If a patient has had a tumour assessment that cannot be evaluated, then the patient will be assigned a visit response of NE (unless there is evidence of progression in which case the response will be assigned as PD). RECIST assessments/scans contributing towards a particular visit may be performed on different dates and for the BICR, 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. The endpoint of PFS will be derived programmatically from this information.

Results of this independent review will not be communicated to investigators and the management of patients will be based solely upon the results of the RECIST v1.1 assessment conducted by the investigator.

A BICR of all patients will be performed for the final database lock for PFS, which will cover all the scans up to the DCO. After the primary PFS analysis, BICR review of scans will no longer be required.

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

3.2 Efficacy Variables

3.2.1 Progression free survival (PFS)

PFS (per RECIST 1.1 as assessed by the site investigator) is defined as the time from the date of randomisation until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from therapy or receives another anti-cancer therapy prior to progression (i.e. date of PFS event or censoring – date of randomisation + 1).

Patients who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment. However, if the patient progresses or dies after 2 or more consecutive missed assessments, the patient will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the 2 missed visits (Note: NE visit is not considered as missed visit).

Given the scheduled visit assessment scheme (i.e. nine-weekly for the first 18 weeks then twelve-weekly thereafter) the definition of 2 missed visits will change. If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from nine-weekly to twelve-weekly this will equate to 23 weeks (i.e. take the average of 9 and 12 weeks which gives 10.5 weeks and then apply 2 x 10.5 weeks + 1 week for an early assessment + 1 week for a late assessment = 23 weeks). The time period for the previous RECIST assessment will be from study days 57 to 120 (i.e. week 8 to week 17). From week 17 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 patient has no evaluable visits or does not have baseline data they will be censored at Day 1 unless they die within two visits (19 weeks) of baseline (2*9=18 weeks plus 1 week allowing for a late assessment within the visit window).

The PFS time will always be derived based on scan/assessment dates and not visit dates. RECIST 1.1 assessments/scans contributing toward a particular visit may be performed on different dates. The following rules will be applied:

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

3.2.2 Overall survival (OS)

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

Note: Survival calls will be made in the week following the date of each DCO, and if patients are confirmed to be alive or if the death date is post the DCO date these patients will be censored at the date of DCO. The status of ongoing, withdrawn (from the study) and "lost to follow-up" patients at the time of the OS analyses should be obtained by the site personnel by checking the patient's notes, hospital records, contacting the patient's general practitioner and checking publicly-available death registries. In the event that the patient has actively withdrawn consent to the processing of their personal data, the date of death of the patient can be obtained by site personnel from publicly available resources where it is possible to do so under applicable local laws.

Note that for any OS analysis performed prior to the final OS analysis, in the absence of survival calls being made, it may be necessary to use all relevant CRF fields to determine the last recorded date on which the patient was known to be alive for those patients still on treatment. 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 subsequent anticancer treatment

- Date last known alive on survival status CRF
- Completion or discontinuation date

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

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

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

3.2.3 Objective response rate (ORR)

ORR is defined as the percentage of patients with a confirmed investigator-assessed response of CR or PR. ORR will be based on a subset of FAS with measurable disease at baseline per the site investigator. A confirmed response of CR/PR means that a response of CR/PR is recorded 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.

Data obtained up until progression, or last evaluable assessment in the absence of progression will be included in the assessment of ORR. Patients who discontinue treatment without progression, receive a subsequent anti-cancer therapy and then respond will not be included as responders in the ORR (i.e. both visits contributing to a response must be prior to subsequent therapy for the patient to be considered as a responder).

In the case where a patient has two non-consecutive visit responses of PR, then, as long as the time between the 2 visits of PR is greater than 4 weeks and there is no PD between the PR visits, the patient will be defined as a responder. Similarly, if a patient has visit responses of CR, NE, CR, then, as long as the time between the 2 visits of CR is greater than 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 (which is subsequently confirmed) 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 that was CR or PR that was subsequently confirmed.

If a patient does not progress following a response, then their DoR will use the PFS censoring time. DoR will only be defined for patients who are responders in the ORR analysis.

3.2.5 Time from randomisation to second progression or death (PFS2)

Time from randomisation to second progression or death (PFS2) will be defined as the time from the date of randomisation to the earliest of the progression event subsequent to first subsequent therapy or death. The date of second progression will be recorded by the investigator in the eCRF and defined according to local standard clinical practice and may involve any of the following: progression by disease specific biomarker CA-125, symptomatic progression, objective radiological progression, other, or death. Second progression status will be reviewed every 12 weeks following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at date last known alive and without a second disease progression (i.e., censored at the latest of the PFS or PFS2 assessment date if the patient has not had a second progression or death).

3.2.6 Time to first subsequent therapy or death (TFST)

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

3.2.7 Time to second subsequent therapy or death (TSST)

As a supportive summary to PFS2, time to second subsequent therapy or death (TSST) is defined as the time from the date of randomisation to the earlier of start date of the second subsequent anti-cancer therapy after discontinuation of first subsequent treatment, or death (i.e. date of second subsequent cancer therapy/death or censoring – date of randomisation + 1). Any patient not known to have had a second anti-cancer subsequent therapy will be censored at the last date that the patient was known not to have received a second subsequent anti-cancer therapy (obtained from the TTSCAPRX form). If a patient terminated the study for reason other than death before second subsequent therapy, these patients will be censored at the earliest of their last known to be alive and termination dates. Patients not receiving randomised treatment would have TSST calculated in the same way, i.e. time from date of randomisation to the subsequent therapy (the second alternative cancer therapy or death).

3.2.8 Time to study treatment discontinuation or death (TDT)

Time to study treatment discontinuation or death (TDT) is defined as the time from randomisation to the earlier of the date of study treatment discontinuation or death (i.e. date of study treatment discontinuation/death or censoring – date of randomisation + 1). Any patient not known to have died at the time of analysis and not known to have discontinued study treatment will be censored based on the last recorded date on which the patient was known to be alive.

Note, for the purposes of the analysis of TDT, study treatment is defined as investigational treatment: durvalumab/placebo or olaparib/placebo. Treatment discontinuation is defined as all investigational products being discontinued. For example:

- If a patient receives both durvalumab/placebo and olaparib/placebo, the date of treatment discontinuation will be the later of the date of discontinuation of each drug respectively.
- If a patient only receives one of the two investigational products, treatment discontinuation will be the date of discontinuation for the investigational product administered.
- If a patient does not receive any investigational product, they will be censored at day 1.

3.2.9 Best objective response (BoR)

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

CR or PR must be confirmed. For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. SD should be recorded at least 9 weeks minus 1 week, i.e. at least 57 days (to allow for an early assessment within the assessment window), after randomisation. For CR/PR, the initial overall visit assessment that showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

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

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

For patients who die with no evaluable RECIST assessments, if the death occurs ≤19 weeks (i.e. 18 weeks + 1 week to allow for a late assessment within the assessment window) after randomisation, then BoR will be assigned to the progression (PD) category. For patients who

die with no evaluable RECIST assessments, if the death occurs > 19 weeks after randomisation then BoR will be assigned to the NE category.

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

3.3 Patient Reported Outcome (PRO) Variables

The following Patient-reported outcome measures will be used in this study:

- European Organisation for Research and Treatment of Cancer (EORTC) Core Quality of Life Questionnaire (EORTC QLQ-C30)
- EORTC Quality of Life Questionnaire – Endometrial Cancer Module (EORTC QLQ-EN24)
- EuroQoL five dimensions, five level (EQ-5D-5L) health state utility index
- Patient reported outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE).
- Patient global impression of severity of cancer symptoms (PGIS)
- Patient global impression of change (PGIC)
- Patient global impression of treatment tolerability (PGI-TT)
- Patient global impression of benefit/risk (PGI-BR)

EORTC QLQ-C30, EORTC QLQ-EN24, EQ-5D-5L, and PGIS will be assessed at baseline (prior to dosing on Day 1 of Cycle 1), then Q3W (±3 days) for the first 18 weeks, and then Q4W (±3 days) until second disease progression or death. Patients will also be assessed at the study treatment discontinuation visit (+2 days) unless completed within 3 days prior to the visit, and for those who discontinue for reasons other than PD, also at the PD visit (+2 days) unless completed within 3 days prior to the visit. PGIC will follow the same schedule except will not be completed at baseline.

PGI-TT and PRO-CTCAE will be assessed at baseline (prior to dosing on Day 1 of Cycle 1), then Q3W (±3 days) for the first 18 weeks, and then Q4W (±3 days) until treatment discontinuation visit. Patients will also be assessed at the study treatment discontinuation visit (+2 days) unless completed within 3 days prior to the visit, and 4 weeks (±3 days) after the study treatment discontinuation visit.

PGI-BR will be assessed at weeks 12, 15, and 18 (each ±3 days) and then Q8W (±3 days). PGI-BR will also be collected at the study treatment discontinuation visit (+2 days) unless completed within 3 days prior to the visit, and for those who discontinue for reasons other than PD, also at the PD visit (+2 days) unless completed within 3 days prior to the visit.

For all these PRO assessments, patients should have an assessment at the start of maintenance therapy visit unless completed within 3 days prior to the visit.

3.3.1 EORTC QLQ-C30 and EORTC QLQ-EN24

3.3.1.1 EORTC QLQ-C30

The EORTC QLQ-C30 (Aaronson et al 1993) is a questionnaire developed to assess cancer patients' measurement of global aspects of health-related quality of life (HRQoL). The QLQ-C30 is a 30-item self-administered questionnaire, with items grouped into five functional scales (physical, role, cognitive, emotional, and social), three symptom scales (fatigue, pain, and nausea/vomiting), and a 2-item global health related quality-of-life scale (global health status/QoL). Five single-item symptom measures are also included providing measurements on additional symptoms commonly reported by cancer patients (dyspnea, loss of appetite, insomnia, constipation, diarrhea) and 1 item on the financial impact of the disease. See scoring algorithm for the EORTC-QLQ-C30 in Appendix 9.1. Higher scores on the global health status/QoL and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent a higher level of symptoms. A clinically meaningful change or difference in EORTC QLQ-C30 subscales/scores is defined as an absolute change in score of ≥10 points (Osoba et al 1998).

3.3.1.2 EORTC QLQ-EN24

The EORTC QLQ endometrial cancer module EORTC QLQ-EN24 is a supplement to the EORTC QLQ-C30 and is used to assess endometrial cancer-specific symptoms and functioning (Greimel et al 2011) for patients with all stages of endometrial cancer. The module consists of 24 questions grouped into 5 multi-item scales to assess lymphoedema (2 items), urological symptoms (4 items), gastrointestinal symptoms (5 items), body image problems (2 items), and sexual/vaginal problems (3 items) as well as single items to assess back/pelvic pain, tingling/numbness, muscular/joint pain, hair loss, taste change, sexual interest, sexual activity, and sexual enjoyment. All questions are rated on a 4-point verbal rating scale: "Not at all," "A little," "Quite a bit," and "Very much." See scoring algorithm for the EORTC-QLQ-EN24 in Appendix 9.2. A higher score for the functional items represents a higher level of functioning, whereas a higher score for the symptom scales/items represents a higher level of symptoms. A clinically meaningful change or difference in EORTC QLQ-EN24 subscales/scores is defined as an absolute change in score of ≥10 points

3.3.2 EQ-5D-5L

The EQ-5D-5L will be used to explore the impact of treatment and disease state on patient-reported health state utility. The EQ-5D-5L, developed by the EuroQol Group, is a generic questionnaire consisting of the EQ-5D-5L descriptive system and the EQ Visual Analogue scale (EQ VAS).

3.3.2.1 EQ-5D-5L descriptive system

The EQ-5D-5L descriptive system questionnaire comprises of 5 dimensions of health covering mobility, self-care, usual activities, pain/discomfort and anxiety/depression. For each dimension, respondents select which statement best describes their health on that day from

any of the following five options of increasing levels of severity: 1 = “No problems”; 2 = “Slight problems”; 3 = “Moderate problems”; 4 = “Severe problems”; 5 = “Unable to/ extreme problems”.

3.3.2.2 Health state utility

Whereas a 1-digit number express the level selected for each individual dimension, a unique health state, termed the EQ-5D-5L profile, is reported as a five-digit code with a possible 3,125 health states from responses to the 5 dimensions. For example, state 11111 indicates no problems on any of the five dimensions, while state 12345 indicates no problems with mobility, slight problems with washing or dressing, moderate problems with doing usual activities, severe pain or discomfort and extreme anxiety or depression. The EQ-5D profile will be converted into a weighted health state utility value, termed the EQ-5D index, by applying a country-specific equation to the EQ-5D-5L profile that represents the comparative value of health states. This equation is based on national valuation sets elicited from the general population and the base case will be the UK perspective. Where a valuation set has not been published, the EQ-5D-5L profile will be converted to the EQ-5D index using a crosswalk algorithm (van Hout et al. 2012). The EQ-VAS is reported separately.

3.3.2.3 EQ VAS

Respondents also assess their health today using the EQ VAS. The EQ VAS is a 0-100 scale that records the respondent’s self-rated health status on a 10 cm vertical, visual analogue scale with endpoints labelled ‘the best health you can imagine’ corresponding to higher values on the VAS score and ‘the worst health you can imagine’ corresponding to lower values on the VAS score. This information can be used as a quantitative measure of health as judged by the individual respondents.

3.3.3 PRO-CTCAE

The PRO version of the CTCAE is an item library of symptomatic AEs experienced by patients while undergoing treatment of their cancer. It was developed to measure the relevant attributes of a symptom directly from patients for the purposes of understanding symptomatic toxicity from the patients' perspective. For this study, the following 3 items are considered relevant for assessment and are not captured by other PRO measures included in the study: 1 item for itching and 2 items for shivering/shaking chills. The PRO-CTCAE will only be administered in those countries where a linguistically validated version exists. Details of scoring algorithm for the PRO-CTCAE are included in Appendix 9.3.

3.3.4 PGIS

The PGIS (Patient global impression of severity of cancer symptoms) is included to assess how a patient perceives the overall current severity of cancer symptoms. This is a single item with six response options: “no symptoms”, “very mild”, “mild”, “moderate”, “severe” and “very severe”.

3.3.5 PGIC

The PGIC instrument captures the patient's overall valuation of response to treatment. The patient is asked to report the degree to which they have changed since entering the treatment period using a 7-point scale (1='Much better', 2=' Moderately better', 3=' A little better', 4='About the same', 5='A little worse', 6='Moderately worse', 7='Much worse').

3.3.6 PGI-TT

The PGI-TT is a single-item questionnaire assessing the overall bother associated with symptomatic AEs. The item is rated using a 6-point verbal scale including "Not at All", "A little bit", "Somewhat", "Quite a bit" and "Very Much".

3.3.7 PGI-BR

The PGI-BR is a 5-item questionnaire to assess the patient's perception of the overall benefits and risks of treatment. The 5 items include overall trial experience, efficacy, side effects, convenience and overall assessment of the benefits and harms of treatment. PGI-BR will only be administered in those countries where a linguistically validated version exists for a country. Details for the PGI-BR are included in Appendix 9.4.

3.3.8 Compliance

Summary measures of overall compliance and compliance over time will be derived for EORTC QLQ-C30, EORTC QLQ-EN24, and EQ-5D-5L for each randomised treatment arm separately. The compliances will be based upon the following definitions:

- Received: Number of patients that have completed questionnaire with completion date and at least one individual item completed.
- Expected: Number of patients who has not withdrawn from the study and expected to complete a questionnaire at a scheduled assessment time. Patients in countries with no available translation will be excluded.
- Evaluable: Number of patients for whom at least one subscale can be determined.
- Compliance rate (visit): Compliance rate for each specific visit (including baseline) is defined as the total number of evaluable subjects divided by total number expected for the visit, i.e. (Evaluable ÷ Expected) * 100.
- Completion rate (visit): Completion rate for each specific visit (including baseline) is defined as the total number of evaluable subjects for the visit divided by total number of randomised patients, i.e. (Evaluable ÷ Number of randomised patients) * 100.
- Evaluable rate (visit): Evaluable rate for each specific visit (including baseline) is defined as the total number of evaluable subjects divided by total number of patients that received a completed questionnaire for the visit, i.e. (Evaluable ÷ Received) * 100.
- Overall compliance rate is defined as the total number of patients with both an evaluable baseline and at least one evaluable postbaseline questionnaire divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

3.4 Health Care Resource Use Variables

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

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

Where admitted overnight, the length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation. That is;

Length of hospital stay = end date of hospitalisation – start date of hospitalisation + 1

Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation. The length of ICU stay will be calculated using the same method.

3.5 Safety Variables

Safety and tolerability of durvalumab in combination with platinum-based chemotherapy (paclitaxel and carboplatin) followed by maintenance durvalumab with or without olaparib will be assessed in terms of AEs (including SAEs, AESIs and immune-mediated AEs), physical examination, ECOG performance status, vital signs, clinical laboratory data, ECG and exposure.

3.5.1 Exposure and dose interruptions

Exposure will be determined separately for chemotherapy agents (paclitaxel and carboplatin), durvalumab/placebo and olaparib/placebo.

For chemotherapy agents (paclitaxel and carboplatin):

Total (or intended) exposure (days) of paclitaxel or paclitaxel substitute alternative treatment:

- Minimum of (last paclitaxel/substitute dose date where dose >0 [ml] + 20 days, date of death, date of DCO) –first paclitaxel/substitute dose date +1

Total (or intended) exposure (days) of carboplatin or carboplatin substitute alternative treatment:

- Minimum of (last carboplatin/substitute dose date where dose >0 [ml] + 20 days, date of death, date of DCO) –first carboplatin/substitute dose date +1

For durvalumab/placebo:

Total (or intended) exposure (days) of durvalumab/placebo:

- Total (or intended) exposure = min (last durvalumab/placebo dose date where dose > 0 [mg] + C days, date of death, date of DCO) – first durvalumab/placebo dose date + 1

Actual exposure (days) of durvalumab/placebo:

- Actual exposure = intended exposure – total duration of dose delays, where intended exposure will be calculated as above.
  Calculation of duration of dose delays (for actual exposure) is:
  For all dosing dates: Total duration of dose delays (= Sum of (Date of the dose – Date of previous dose – C days)). Thus, if no delays were encountered, the duration would sum up to 0, since infusions were done every C days.

C is equal to the scheduled number of days between doses minus one. C is equal to '20' if the last dose date falls prior to Cycle 7 where dosing is every 3 weeks and '27' if the last dose date falls Cycle 7 onwards where dosing is every 4 weeks.

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

For olaparib/placebo:

Total (or intended) exposure (days) of:

- Total (or intended) exposure = min (last olaparib/placebo dose date > 0 [mg], date of death, date of DCO) – first olaparib/placebo dose date +1

Actual exposure of study treatment

- Actual exposure = intended exposure – total duration of dose interruptions, where intended exposure will be calculated as above and a dose interruption is defined as any length of time where the patient has not taken any of the planned daily dose [i.e. sum of (end date of each interruption-start date of the interruption+1)]. To calculate actual exposure, dose interruptions will include those where a patient forgot to take a dose.

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

Number of treatment infusions received

Exposure of chemotherapy agents and durvalumab/placebo will also be measured by the number of infusions received. A cycle corresponds to a period of 21 days during the chemotherapy phase (cycles 1 to up to 6) and of 28 days during the maintenance phase for durvalumab/placebo. One infusion per cycle is planned for chemotherapy agents and

durvalumab/placebo. An infusion will be counted if treatment is started even if the full dose is not delivered.

Missed or forgotten doses

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

Patients who permanently discontinue during a dose interruption

If a patient permanently discontinues study treatment during a dose interruption, then the date of last administration of study medication recorded on DOSDISC will be used in the programming as the last dose date. The dose interruption that happens immediately before patients permanently discontinue the study treatment will not be included in the summary table as interruption (i.e. an interruption will only be included if the drug was restarted).

Safety Follow-up

- Total Safety Follow-up = min[(last dose date + xx days), date of withdrawal of consent, date of death, date of DCO)]– first dose date +1, where the 'last dose date + xx days' will be the 'last dose date of olaparib/placebo + 30 days' or 'last dose date of durvalumab/placebo + 90 days', whichever is later.

3.5.2 Dose intensity

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

- RDI = 100% 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. D is the total dose that would be delivered, if there were no modification to dose or schedule.

For Durvalumab/placebo

In deriving dose intensity parameters for durvalumab, the following should be considered:

- Dose reductions are not permitted.
- There might be dose interruptions and dose delays in which case a dose cycle may be prolonged. This cycle should be still counted as one cycle. A cycle is counted if treatment is started even if the full dose is not delivered.
- If a decision is made to permanently discontinue study treatment in-between cycles or during a cycle delay, then the date of the last administration of durvalumab/placebo recorded will be used in the determination of dose intensity.

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

For olaparib/placebo

- d is determined by adding up the received dose for each day from the date of first olaparib dose to the date of last olaparib dose.
- D = 300 mg x 2 x total (intended) exposure

3.5.3 Adverse events

SAEs will be collected right after the patient signed the informed consent and throughout the treatment period and the safety follow-up periods. All other AEs will be collected from Cycle 1 Day 1 throughout the treatment periods and the safety follow-up periods until the later date of:

- 30 days after the last dose of olaparib/placebo (the follow-up period for olaparib)
- 90 days after last dose of durvalumab/placebo (the follow-up period for durvalumab)

Any AE (including SAE) that is ongoing at the end of safety follow-up period needs to be followed up until it is resolved or the event is confirmed by the investigator to be unlikely to resolve, or the patient is lost to follow-up.

An adverse events will be defined as treatment emergent if it onsets, or worsens (by investigator report of a change in intensity) at or after the start of the first dose of any of the study treatments, including durvalumab/placebo or olaparib/placebo, throughout the treatment period and including the safety follow-up periods until the later date of 30 days after the last dose of olaparib/placebo and 90 days after last dose of durvalumab/placebo.

The latest version of Medical Dictionary for Regulatory Activities (MedDRA) 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).

Other significant adverse events (OAE)

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

Olaparib AEs of special interest

Olaparib adverse events of special interest (AESI) are events of scientific and medical interest specific to the further understanding of the safety profile of olaparib and may require close monitoring and rapid communication by the investigators to AstraZeneca. An AESI may be serious or non-serious.

These AESIs are identified as a list of categories provided by the AstraZeneca patient safety team. Other categories may be added as necessary or existing terms may be merged.

An AstraZeneca medically qualified expert, after consultation with the Global Patient Safety Physician, will review the AEs of interest and identify which preferred terms contribute to each AESI. A further review will take place prior to DBL to ensure any further terms not already included are captured within the categories.

Durvalumab 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" (AEPI) to the durvalumab program.

AESI are defined as AEs that include, but are not limited to, events with a potential inflammatory or immune mediated mechanism that may require more frequent monitoring and/or interventions such as corticosteroids, immunosuppressants, and/or endocrine therapy. Endocrine therapies include standard endocrine supplementation, as well as treatment of symptoms resulting from endocrine disorders (for example, therapies for hyperthyroidism include beta blockers [e.g., propranolol], calcium channel blockers [e.g., verapamil, diltiazem], methimazole, propylthiouracil, and sodium perchlorate).

The AEPIs reported in the AstraZeneca-sponsored durvalumab studies are defined as AEs that could have a potential inflammatory or immune-mediated pathophysiological basis resulting from the mechanism of action of durvalumab but are more likely to have occurred due to other pathophysiological mechanisms, thus, the likelihood of the event being inflammatory or immune-mediated in nature is not high and/or is most often or usually explained by the other causes.

The AESIs and AEPIs have been categorized into the following AESI/AEPI categories: Pneumonitis, Hepatic events, Diarrhoea/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.

Immune-mediated Adverse Events (imAE)

imAE will be identified from both the Durvalumab 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 a separate imAE charter. In addition, the Sponsor may perform medical review of those AESIs and classify them as imAEs or not imAEs via an independent manual adjudication process.

3.5.4 Laboratory data

Assessments of safety laboratory tests will be performed at a local laboratory at or near to the investigator site. Parameters related to laboratory data including haematology, clinical chemistry, urinalysis (see CSP, Tables 12 to 14 for a list of laboratory tests) will be assessed based on samples that will be collected throughout the study, from screening to end of safety follow-up visit. Absolute values will be compared to the laboratory reference range and classified as low (below range), normal (within range or on limits of range) and high (above range). All values classified as high or low will be flagged on the listings.

Additionally, assessment on coagulation (APTT and INR), disease specific tumour marker Cancer antigen 125 (CA125), bone marrow or blood cytogenetic analysis for patients with prolonged haematological toxicities, and pregnancy test will be conducted.

Derivation rules for post baseline visit values determination including visit window and multiple records will be described in Section4.2.10.1.

As applicable, values will be converted to standard units and will be graded using CTCAE version 5.0. Maximum post-baseline CTCAE grade will also be calculated.

3.5.5 Analysis of Total Calcium per NCI CTCAE criteria

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

Corrected calcium product will be derived during the creation of the analysis datasets (ADaM) using the following formula:

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

3.5.6 ECG

Resting 12-lead ECGs will be performed at screening or within 7 days prior to starting study treatment and when clinically indicated. Measurements should be taken after the subject has been rested in a supine position for at least 5 minutes. All ECGs will be assessed locally to determine whether they are clinically significantly abnormal/not clinically significantly abnormal. In case of clinically significant ECG abnormalities, including a QTcF value >470 ms, 2 additional 12-lead ECGs should be obtained over a brief period (e.g. 30 minutes) to confirm the finding. If there is a clinically significant abnormal finding, it will be recorded as an AE by the investigator if applicable.

ECG data obtained from screening up until the 30 days from date of last dose of olaparib/placebo treatment or 90 days from last infusion of durvalumab will be used for reporting. ECG changes from baseline will be assessed as 'worst change from baseline'.

3.5.7 Vital signs

Vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse, temperature and height [baseline only]) will be evaluated at the timepoints specified in the CSP and as clinically indicated at any other time, up until the end of safety follow-up (30 days post last dose for subjects receiving olaparib/placebo or 90 days after last dose of durvalumab/placebo, whichever is later). Generally, blood pressure, pulse and other vital signs should be measured prior to the start of the infusion. On the first infusion day BP and pulse will be collected before (approximately 30 minutes to 0 minutes before the beginning of the infusion), during (halfway through the infusion, approximately 30 minutes during the infusion) and at the end of the durvalumab/placebo infusion (approximately 60 minutes ±5 minutes). Weight will be assessed at screening and as clinically indicated at any other time.

For derivation of post baseline visit values, visit window and multiple records handling will be described in Section 4.2.10.1. Any clinically significant changes in vital signs should be recorded as an AE, if applicable (see CSP Sections 8.3.9).

3.5.8 Physical examination

Physical examination assessments including the following: general appearance, respiratory, cardiovascular, skin, head and neck (including ears, eyes, nose and throat), lymph nodes, thyroid, abdomen, pelvis, musculo-skeletal (including spine and extremities) and neurological systems will be performed at screening, Day 1 and subsequently as specified according to the study schedule (see CSP Tables 1 to 3). Any clinically significant changes in physical examination should be recorded as an AE, if applicable (see CSP Sections 8.3.8).

3.5.9 ECOG performance status

ECOG performance status will be assessed at screening (within 7 days prior to first dose of study treatment) and on treatment according to the study schedule (see CSP Tables 1 to 3). During follow up ECOG performance status might be collected at other site visits that the

patient attends, and in addition when information on subsequent anticancer therapy is provided.

Patients must have an ECOG performance status 0 or 1 within 7 days of starting study treatment to be eligible for enrolment. Any significant change from baseline or screening must be reported as an AE.

3.6 Pharmacokinetic variables (durvalumab only)

Pharmacokinetics will be assessed for durvalumab.

Serum samples will be analysed by a designated third party on behalf of AstraZeneca, using an appropriate bioanalytical method. Full details of the analytical method will be provided in a separate Bioanalytical Validation Report. Pharmacokinetic concentration data will be collected as per the protocol.

Due to sparse sampling in this study, pharmacokinetic data analyses will only comprise of summaries of serum concentrations for durvalumab. The following PK parameters will be determined after the steady-state doses: peak and trough concentration (as data allow).

3.7 Immunogenicity variables

Serum samples for anti-drug antibody (ADA) and ADA-neutralising antibodies for durvalumab will be collected at scheduled visits as per the Clinical Study Protocol. Placebo samples will not be analysed for ADA.

Samples will be measured for the presence (positive) or absence (negative) of ADAs. Tiered analysis will be performed to include screening, confirmatory, and titer assay components. Positive negative cut points that were previously statistically determined from drug-naïve validation samples will be employed to determine ADA positive or negative. If the sample is positive, the ADA titer will be reported as well. In addition, the presence of neutralizing antibody (nAb) will 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 sample. ADA-evaluable subjects are patients in the SAF who received at least 1 dose of durvalumab and have non-missing baseline ADA result and at least 1 post-baseline ADA. The following responses variables will be evaluated:

- ADA positive at any timepoint including baseline and all post-baseline. The percentage of ADA positive patients in the ADA evaluable population is known as ADA prevalence.
- ADA incidence (treatment-emergent ADA), defined as the percentage of ADA evaluable population with either treatment-induced or treatment-boosted ADA.
- 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 (greater than the analytical variance of the assay) following drug administration.
- Persistently positive ADA, defined as having at least 2 post-baseline ADA positive measurements within 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 includes patients 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 includes patients meeting these criteria who are ADA positive at baseline.
- nAb positive at any visit.

3.8 Exploratory Variables

3.8.1 Biomarker variables

Blood and tissue samples for exploratory biomarker analyses will be obtained according to the schedules presented in Section 1.1 of the CSP. Results of the exploratory biomarker data may be reported separately outside the CSR.

4. ANALYSIS METHODS

The primary objective of the study is to determine the efficacy by PFS (per RECIST 1.1 as assessed by investigator) of Durvalumab+Placebo and Durvalumab+Olaparib. This will be assessed by determining the efficacy of:

- Arm C (durvalumab in combination with SoC platinum-based chemotherapy and continued as maintenance in combination with olaparib) versus Arm A (SoC platinum based chemotherapy).
- Arm B (durvalumab in combination with SoC platinum-based chemotherapy and continued as maintenance) versus Arm A (SoC platinum-based chemotherapy).

The formal statistical analysis will be performed to test the hypotheses of interest:

- H_{0CA}: Arm C = Arm A versus H_{1CA}: Arm C ≠ Arm A
  and
- H_{0BA}: Arm B = Arm A versus H_{1BA}: Arm B ≠ Arm A

Where H_0 = the null hypothesis; H_1 = the alternate hypothesis.

The study will be considered positive (a success) if either of the above null hypotheses are rejected based on the primary analysis of PFS in the FAS.

OS will also be assessed for the Durvalumab+Placebo arm versus the control arm and for the Durvalumab+Olaparib arm versus the control arm. Details of type I error control for multiple efficacy tests will be described in section 4.2.1.

Pooling strategy

For the analysis of primary endpoint of PFS, the following pooling strategy will be applied across all three arms. If the number of events in the individual stratum are too small for a meaningful analysis (less than 5 events per stratum; a stratum is defined as strata1*strata2*…strataX*treatment; so with 3 stratification factors of each 2 levels and 3 treatments we have 2*2*2*3=24 stratum) stratification factors will be removed in the following order until there are at least 5 events in each stratum: Region (Asia / Rest of World); MMR status (Proficient / Deficient); Disease status (Recurrent / Newly diagnosed). The same pooling will be applied for both primary comparisons (Arm B vs Arm A and Arm C vs Arm A). All analyses will then be conducted with the same stratification factors as the primary analysis of PFS . If there are secondary endpoints or sensitivity analyses that still do not conform to the 5 event rule per stratum, an unstratified analysis will be conducted. This will be supported by an unstratified analysis of the primary endpoints.

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

Timing of efficacy analysis

The timing of efficacy analyses is summarised in Table 5 and is described below.

Table 5: Timing of efficacy analyses

Analysis	**Timepoint**
Primary PFS/First interim OS analysis	The same calendar time when there are approx. 299 PFS events across Durvalumab+Placebo arm and control arm and approx. 281 events across Durvalumab+Olaparib arm and control arm
Second interim OS Analysis	The same calendar time when there are approx. 244 deaths across Durvalumab+Placebo arm and control arm and approx. 244 deaths across Durvalumab+Olaparib arm and control arm
Final OS analysis	The same calendar time when there are approx. 280 deaths across Durvalumab+Placebo arm and control arm and approx.. 280 deaths across Durvalumab+Olaparib arm and control arm

4.1 General principles

Efficacy and PRO data will be summarised and analysed on the FAS analysis set. Safety and treatment exposure data will be summarised based upon the SAF. Study population and demography data will be summarised based upon the FAS analysis set.

Data will be presented in data listings by treatment group and subject number. All summaries will be presented by treatment group, unless otherwise specified. A month is operationally defined to be 30.4375 days. Six months is operationally defined to be 183 days.

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

- Descriptive statistics will be used for all variables, as appropriate, and will be presented by treatment group. Continuous variables will be summarised 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 summarised by frequency counts and percentages for each category.
- Unless otherwise stated, percentages will be calculated out of the total population for the corresponding treatment arm.
- 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.
- Where analysis models are stratified by the randomisation stratification factors; MMR status: MMR proficient tumours vs MMR deficient tumours, Disease status: Recurrent endometrial cancer vs Newly diagnosed endometrial cancer, Geographic region: Asia vs RoW, the strata obtained from the randomisation code will be used, not the values recorded in the electronic case report form (eCRF).
- For safety endpoints the last observation before the first dose of study treatment will be considered the baseline measurement unless otherwise specified. For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, will serve as sufficient evidence that the assessment occurred prior to first dose. Assessments on the day of the first dose where neither time nor a nominal pre-dose indicator are captured will be considered prior to the first dose if such procedures are required by the protocol to be conducted before the first dose.
- In general, for efficacy and PRO endpoints, baseline is defined as the last assessment prior to randomisation. However, if an evaluable assessment is not available before randomisation but available before first dose of study treatment then this assessment will be used as baseline.
- 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

  Percent change = (post-baseline value - baseline value) / baseline value x 100.
- P-values will be rounded to 3 decimal places. P-values less than 0.0005 (e.g. 0.0002) will not be rounded to 3 decimal places (e.g. 0.000) but instead be displayed as <0.001. P-values output as <0.0001 by statistical software will not be rounded and displayed in the same way (‘<0.0001’).
- SAS® version 9.4 (as a minimum) will be used for all analyses.

Depending on the extent of any impact, summaries of data relating to patients diagnosed with COVID-19, and impact of COVID-19 on study conduct (in particular missed visits, delayed or discontinued study treatments, and other protocol deviations) may be generated, see Section 9.5 for further information.

4.2 Analysis Methods

Table 6 provides the summary of all formal statistical analyses planned for this study together with pre-planned sensitivity analyses.

Table 6: Formal statistical analyses to be conducted and pre-planned sensitivity analyses

Endpoints analysed	**Notes**
Progression-free survival	• Primary analysis of Durvalumab+Placebo vs Control and Durvalumab+Olaparib vs Control using stratified log-rank test using investigator assessments (RECIST 1.1) • Exploratory analysis of investigator assessments (RECIST 1.1) (same method as primary analysis) for the comparison of Durvalumab+Olaparib vs Durvalumab+Placebo • Sensitivity analyses (durvalumab ± olaparib + vs control) using investigator assessments (RECIST 1.1) • Interval censored analysis – evaluation time bias • Analysis using alternative censoring rules – attrition bias • Sensitivity analysis via stratified log-rank test using BICR assessments (RECIST 1.1) – ascertainment bias - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control - Durvalumab+Olaparib vs Durvalumab+Placebo •Additional analysis using Cox proportional hazards models to determine the consistency of treatment effect between subgroups via the approach of Gail and Simon 1985. •Subgroup analysis using an unstratified Cox proportional hazards model (same comparisons as for the sensitivity analysis above)
Objective response rate	Descriptive summary of the number and percentage of patients with a confirmed tumour response according to the investigator (CR/PR) based on the number of patients with measurable disease at baseline Adjusted logistic regression (durvalumab ± olaparib + vs control)
Duration of response	Descriptive summary of the duration of response according to the investigator assessment in responding patients
PFS2, TFST, TSST, TDT	Stratified log-rank test for: - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control
Overall survival	• Stratified log-rank test for: - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control - Durvalumab+Olaparib vs Durvalumab+Placebo

Endpoints analysed	Notes
	• Sensitivity analysis using a Kaplan-Meier plot of time to censoring where the censoring indicator of the primary analysis is reversed – attrition bias - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control - Durvalumab+Olaparib vs Durvalumab+Placebo • Subgroup analysis using unstratified Cox proportional hazards model (same comparisons as for the sensitivity analysis above) at time of the final OS analysis
Change from baseline (EORTC QLQ-C30 and EN24 scores)	MMRM analysis: - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control
Time to deterioration (EORTC QLQ-C30 and QLQ-EN24 scores)	Stratified log-rank test for: - Durvalumab+Placebo vs Control - Durvalumab+Olaparib vs Control
EQ-5D-5L	Summary statistics
PK analysis (durvalumab only)	Descriptive statistics

4.2.1 Multiplicity

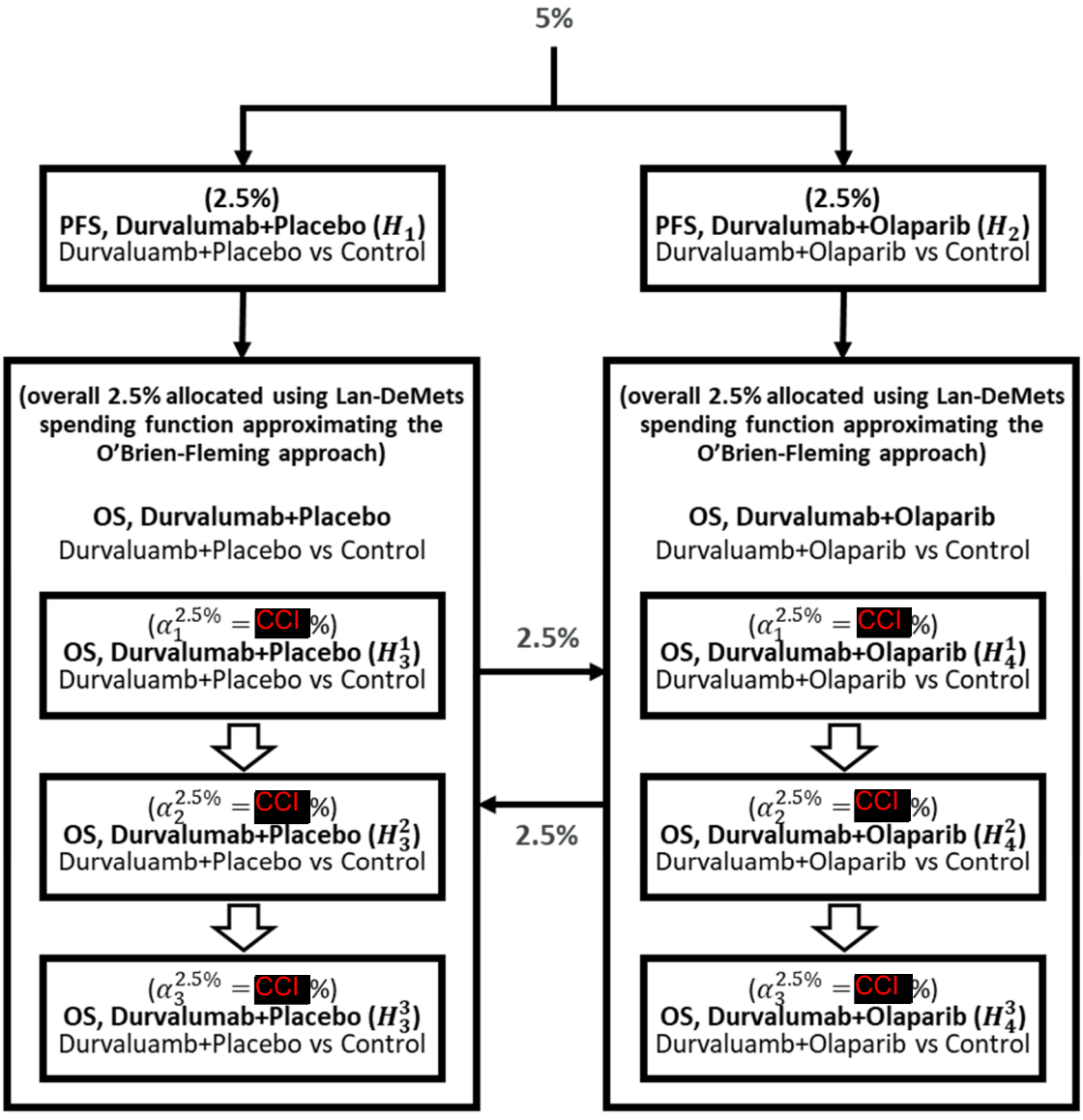
In order to strongly control the type I error at 5% (2-sided), a multiple testing procedure (MTP) with a gatekeeping strategy will be used across the key endpoints (PFS and OS) and treatment comparisons of interest (Durvalumab+Placebo arm versus control and Durvalumab+Olaparib arm versus control). If the higher level null hypothesis in the MTP is rejected for superiority, the following hypothesis will then be tested as shown in Figure 1.

Hypotheses will be tested using a multiple testing procedure with an alpha-exhaustive recycling strategy (Burman et al 2009). With this approach, hypotheses will be tested in a pre-defined order as outlined in Figure 1. According to alpha (test mass) splitting and alpha recycling, the test mass that becomes available after each rejected null hypothesis is recycled to secondary hypotheses not yet rejected. Since OS is tested at multiple timepoints (ie, 2 interim analyses and final analysis), the OS tests that for the same comparison (ie, shown in one box in the MTP) will be considered as one test family. As long as one test in the family can be rejected, the family is rejected thus the assigned total alpha to the family can be recycled to the next MTP level. This testing procedure stops when the entire test mass is allocated to non-rejected null hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses. Figure 1 shows the multiple testing framework.

Overall survival is tested at 2 interims and a final timepoint (see Section 5 for details). The alpha level allocated to OS will be controlled at the interim and primary time points by using the Lan DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, where the alpha level applied at the interim depends upon the proportion of information available. Note: If any interim analysis or primary analysis is statistically

significant, the overall alpha (two-sided) will be allocated to the next level. If the interim results do not meet the criterion of stopping for superiority for a given hypothesis, then follow-up will continue until the final target number of OS events for that comparison has been observed, following which the hypothesis will be retested. If the null hypothesis is then rejected, subsequent testing will continue hierarchically. The above testing procedure will ensure strong control of the family-wise error rate (Glimm et al 2009).

Figure 1: Illustration of Multiplicity strategy



Note: Alpha levels presented in this figure are 2-sided.

The overall 5% type I error rate will be controlled by initially assigning 2.5% alpha (2-sided) to each of the primary PFS comparisons of interest (i.e. Durvalumab+Placebo arm versus

Control ($PFS_{Durvalumab+Placebo}$) and Durvalumab+Olaparib versus Control ($PFS_{Durvalumab+Olaparib}$) (see Figure 1)).

If the PFS analysis in either arm is statistically significant, the respective 2.5% alpha (2-sided) will be allocated to the next level in a pre-defined order as outlined in Figure 1. For example, at the second level of the hierarchy, if the $PFS_{Durvalumab+Placebo}$ comparison is significant, 2.5% alpha (2-sided) will be assigned to the comparison of OS of the Durvalumab+Placebo arm versus the control arm ($OS_{Durvalumab+Placebo}$) and, if the $PFS_{Durvalumab+Olaparib}$ comparison is significant, 2.5% alpha (2-sided) will be assigned to the comparison of OS of the Durvalumab+Olaparib arm versus the Control arm ($OS_{Durvalumab+Olaparib}$).

If the $OS_{Durvalumab+Placebo}$ or $OS_{Durvalumab+Olaparib}$ analysis is statistically significant, the alpha (2-sided) will be allocated to the next level in a pre-defined order as outlined in Figure 1. For example, in this procedure, if the testing for $OS_{Durvalumab+Placebo}$ is significant at the alpha level specified in Figure 1 at either interim or final analysis, the full 2.5% alpha level for $OS_{Durvalumab+Placebo}$ can be propagated to the testing of $OS_{Durvalumab+Olaparib}$, which means that the $OS_{Durvalumab+Olaparib}$ will be tested at an overall alpha level of 5%. In case of alpha propagation from statistically significant testing of $OS_{Durvalumab+Placebo}$, the significance level for all interim and final analyses of $OS_{Durvalumab+Olaparib}$ will be recalculated based on a 5% alpha level overall. CCI [redacted]

Similarly, if the testing for $PFS_{Durvalumab+Olaparib}$ is significant and $OS_{Durvalumab+Olaparib}$ is significant at the 2.5% alpha level specified in Figure 1, then the full 2.5% alpha level for $OS_{Durvalumab+Olaparib}$ can be propagated to the testing of $OS_{Durvalumab+Placebo}$, which means that $OS_{Durvalumab+Placebo}$ will be tested at an overall alpha level of 5%. The significance level for all interim and final analyses of $OS_{Durvalumab+Placebo}$ will be recalculated based on 5% alpha level overall. CCI [redacted]
CCI [redacted]

Note, the interim/final OS analysis boundaries will ultimately be derived based on the actual number of events observed in the study; those referenced above and in Figure 1 are provided as examples only.

4.2.2 Progression free survival (PFS)

The primary PFS analysis of Durvalumab+Placebo versus Control and Durvalumab+Olaparib versus Control will be performed separately using a log-rank test stratified in accordance with the pre-defined pooling strategy (section 4.1) for generation of the p-value. The stratification variables in the statistical modelling will be based on the values entered into IVRS at randomisation, even if it is subsequently discovered that these values were incorrect.

The HR and its confidence interval will be estimated from a stratified Cox Proportional Hazards model (with ties = Efron and the stratification factors as strata) and the CI calculated using a profile likelihood approach (RISKLIMITS=PL).

An exploratory analysis to compare Durvalumab+Olaparib versus Durvalumab+Placebo will be performed. This analyses will be performed using the same methodology as for the primary endpoint described above.

Kaplan-Meier (KM) plots of PFS will be presented by treatment arm. Summaries of the number and percentage of subjects experiencing a PFS event, and the type of event (RECIST progression or death) will be provided along with median PFS for each treatment. Also, the proportion of patients alive and progression free at 6 monthly intervals from randomisation will be summarised (using the KM analysis) and presented by treatment group.

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

Supportive summaries/graphs

The number of patients prematurely censored will be summarised by treatment arm. A patient is defined as prematurely censored if they had not progressed and the latest scan prior to DCO was two or more scheduled tumour assessment interval (+ 2 weeks for protocol allowed window) prior to the DCO date. A KM plot, with tick marks to identify censored observations, of PFS will be presented by treatment group. In addition, duration of follow-up will be summarised using median time from randomisation to date of censoring (date last known to be non-progressor) in censored (not progressed) patients only, presented by treatment group.

Proportionality assumption

The assumption of proportionality will be assessed. Proportional hazard will be tested firstly by examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, Cox model with a time dependent covariate would be fitted to assess the extent to which this represents random variation. If a lack of proportionality is evident, the variation in treatment effect can be described by presenting piecewise HR calculated over distinct time periods. In such circumstances, the HR from the primary analysis can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. The treatment effect will also be described by presenting piecewise HRs calculated over distinct time-periods of 0-6, 6-12, 12-18, 18-24, >24 months. The piecewise model will be implemented by the addition of a time varying covariate/factor (based on the periods in the previous sentence) as per Collet 2003.

4.2.2.1 PFS sensitivity analyses

Sensitivity analyses will be conducted for the comparison of the Durvalumab+Placebo arm versus Control arm and Durvalumab+Olaparib versus Control arm.

(a) Evaluation time bias

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

(b) Attrition bias

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following two, or more, non-evaluable tumour assessments will be included. In addition, and within the same sensitivity analysis, patients who take subsequent therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy) prior to their last evaluable RECIST assessment or progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy.

This analysis will be supported by a Kaplan-Meier plot of the time to censoring using the PFS data from the primary analysis and where the censoring indicator of the PFS analysis is reversed.

(c) Ascertainment bias

Ascertainment bias will be assessed by analysing the BICR data. The stratified log rank test will be repeated on PFS using the BICR data based upon RECIST. The HR and 95% CI will be presented using stratified Cox Proportional Hazard model. This sensitivity analysis will also be conducted for the analysis of Durvalumab+Olaparib vs Durvalumab+Placebo.

If there is an important discrepancy between the primary analysis using the site investigator data and this sensitivity analysis using BICR data then the proportion of patients with site but no central confirmation of progression will be summarised; such patients have the potential to induce bias in the central review due to informative censoring. An approach of imputing an event at the next visit in the central review analysis may help inform the most likely HR value

(Fleischer et al 2011), but only if an important discrepancy is determined to exist by the study team.

Disagreements between investigator and central reviews of RECIST progression will be presented for each treatment group. The summary will include the early discrepancy rate which is the frequency of BICR progressions declared before the investigator review progressions (≥2 weeks earlier and including progressions declared by BICR but not investigator) as a proportion of all BICR review progressions, and the late discrepancy rate which is the frequency of BICR review progressions declared after the investigator review progressions (≥2 weeks later and including progressions declared by investigator review but not BICR) as a proportion of all discrepancies (including early and late discrepancies) (Amit et al 2010).

In the case where the distribution of discrepancy in progression assessment between BICR and local investigator across treatment groups is not similar, the PFS analysis may be biased due to informative censoring. The potential impact of informative censoring on parameter estimate will be assessed through sensitivity analysis, using either the methods of Jackson et al or Hsu and Taylor (Jackson et al 2014, Hsu and Taylor 2009) when considering time dependent covariates. This work will be presented separately and will not form part of the CSR.

(d) Deviation bias (if meaningful to do)

As a sensitivity analysis to the primary PFS endpoint, an analysis excluding patients with deviations that may affect the efficacy of the trial therapy will be performed if > 10% of patients:

- Did not have the intended disease or indication or
- Did not receive any randomised therapy

A stratified log-rank test will be repeated using the investigator RECIST data, using the same ties and stratification factors as described for the primary analysis of PFS. The HR and 95% CI will be presented using stratified Cox Proportional Hazard model.

4.2.2.2 Subgroup Analysis

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1 using investigator assessments) in the Durvalumab+Placebo arm versus Control arm, the Durvalumab+Olaparib arm versus Control arm, and the Durvalumab+Olaparib arm versus Durvalumab+Placebo arm to assess consistency of treatment effects across potential or expected prognostic factors.

The following subgroups of the full analysis set will be analysed for PFS:

Stratification factors:

- Disease status (Recurrent vs. Newly diagnosed)
- MMR status (Proficient vs. Deficient)
- Region (Asia vs. RoW)

Additional subgroups of interest include:

- Age at randomisation (<65 vs. ≥65 years of age)
  - This will be determined from "Age" (DM module) on the eCRF at screening
- Race (White vs. Black/African-American vs. Asian vs. Other [Native Hawaiian/Pacific Islander or American Indian/Alaska Native or Others]).
  - This will be determined from the response to "Race" (DM module) on the eCRF at screening.
- Homologous recombination repair mutation (HRRm) status (HRRm vs. non-HRRm vs. Unknown)
  - HRRm: sample with a pathogenic mutation in any of the pre-specified genes associated with HRR; non-HRRm: sample with no pathogenic mutations in any of the pre-specified genes associated with HRR; Unknown: sample where HRRm status was not available either due to test fail (unevaluable sample or assay failure)
  - HRRm status will be defined based upon the following genes: ATM, BRCA1, BRCA2, BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2, RAD51B, RAD51C, RAD51D and RAD54L
  - Note: This subgroup analysis will not include patients randomised in China.
- PD-L1 expression (Positive vs. Negative vs. Unknown)
  - Positive: sample with PD-L1 expression at any intensity above or equal to predefined cut-offs; Negative: sample with PD-L1 expression at any intensity below predefined cut-offs. Unknown: sample where PD-L1 expression was not available either due to a test fail (unevaluable sample or assay failure) or sample slide out of cut-slide stability.
  - The following cut-off will be assessed:
    - a tumour area positivity score and ≥1% cut off (TAP1)
  - PD-L1 expression will be evaluated applied to sections stained using Ventana SP263 immunohistochemistry assay. Other exploratory cut-offs may also be assessed as required.
- Histology (Endometroid vs. Serous vs. Others)
  - This will be determined from "Histology type" (Pathology: At Time of Diagnosis of Disease under Investigation module) on the eCRF at screening.
- Histological grade (High grade (G3) vs. Low grade (G1+G2))
  - This will be determined from "Tumour grade" (Pathology: At Time of Diagnosis of Disease under Investigation module) on the eCRF at screening.
- ECOG Performance Status (0 vs. 1)
  - This will be determined from "Performance status" (ECOG Performance Status module) on the eCRF at screening.
- FIGO stage at initial diagnosis in newly diagnosed patients (III vs. IV)
  - This will be determined from "FIGO stage" (Pathology: At Time of Diagnosis of Disease under Investigation module) on the eCRF at screening.

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

Other baseline variables may also be assessed if there is clinical or biological justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic factors. If a baseline imbalance is observed between treatment arms, ad-hoc subgroup analysis may be used to investigate any potential for impact on the main results.

No adjustment to the significance level for testing will be made since all these subgroup analyses will be considered exploratory and may only be supportive of the primary analysis of PFS.

For each subgroup level of a factor, the HR and 95% CI will be calculated from a Cox proportional hazards model that only contains a term for treatment. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties, and using a BY statement for the subgroup factor.

These HRs and associated two-sided 95% profile likelihood CIs will be summarised and presented on a forest plot, along with the results of the overall primary analysis.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events across both treatment groups in a subgroup), the HR and CI will not be produced for that subgroup. In this case, only descriptive summaries will be provided.

The presence of quantitative interactions will be assessed by means of an overall global interaction test for plausible subgroups:

The plausible subgroups with biological rationale for an interaction with treatment consist of the following subgroup covariates: disease status, MMR status, region, age group, race, HRRm status, PD-L1 expression, histology, histological grade and ECOG status.

This is performed by comparing the fit of a Cox proportional hazards model including treatment, all covariates, and all covariate-by treatment interaction terms, with one that excludes the interaction terms, and will be assessed at the 2-sided 10% significance level. If there are not more than 10 events per stratum for any covariate (i.e., within each stratum of a treatment*covariate interaction (2 treatments * 2 levels of the covariate = 4 stratum) a pre-defined pooling strategy should be applied to the covariate. If the pooling strategy does not meet the event criteria then the covariate-by-treatment interaction term should be omitted from the model. Moreover, if the covariate does not have more than 10 events per level of covariate then the main effect of the covariate will also be excluded. If the fit of the model is not significantly improved then it will be concluded that overall the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included

until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. Throughout this process all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

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

4.2.3 Overall survival (OS)

OS of Durvalumab+Placebo versus Control and Durvalumab+Placebo versus control will be analysed using the same methodology and model as that used for the analysis of PFS. The number and proportion of patients alive at 6 monthly intervals from randomisation (6, 12, 18, 24, etc) will be summarised (using the KM analysis) and presented by treatment group. A KM plot of OS will be presented by treatment arm.

The HR and its confidence interval will be estimated using a stratified Cox Proportional Hazards model (with ties = Efron and the stratification factors as strata) and the CI calculated using a profile likelihood approach.

In addition, an analysis of OS will be performed to compare Durvalumab+Olaparib versus Durvalumab+Placebo. This analysis will be performed using the same methodology as described above.

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

Subgroup analyses will be conducted comparing OS at the time of the final OS analysis, in the Durvalumab+Placebo arm versus Control arm, the Durvalumab+Olaparib arm versus Control arm, and the Durvalumab+Olaparib arm versus Durvalumab+Placebo arm to assess consistency of treatment effects across potential or expected prognostic factors. This will use the same methodology as the subgroup analyses for PFS (section 4.2.2.2).

4.2.3.1 OS sensitivity analysis

Sensitivity analyses for OS will examine the censoring patterns to rule out attrition bias with regard to the treatment comparisons and will be achieved by a Kaplan-Meier plot of time to censoring where the censoring indicator of OS is reversed.

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

In addition, duration of follow-up which is defined as time from randomisation to the date of death (i.e. overall survival) or to the date of censoring (date last known to be alive) for

censored patients regardless of treatment arm will be summarised using medians in all patients.

4.2.4 Objective response rate (ORR)

The ORR will be based on the site investigator RECIST data, and using all scans regardless of whether they were scheduled or not.

The ORR will be compared between Durvalumab+Placebo versus Control and Durvalumab+Olaparib vs Control using separate logistic regression models adjusting for the same stratification factors as the primary endpoint as covariates in the model. The results of the analyses will be presented in terms of an odds ratio (an odds ratio greater than 1 will favour investigational treatment arm) 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).

If there are not enough responses for a meaningful analysis using logistic regression then a CMH test will be presented.

Summaries will be produced that present the number and percentage of patients with a confirmed tumour response (CR/PR) based upon the number of patients with measurable disease at baseline per the site investigator.

For each treatment arm, best objective response (BoR) will be summarised by n (%) for each category (CR, PR, SD, PD and NE). No formal statistical analyses are planned for BoR.

4.2.5 Duration of response (DoR)

Descriptive data will be provided for the duration of response in responding patients, including the associated Kaplan-Meier curves (without any formal comparison or p-value attached).

Kaplan Meier plots of DoR based on the investigator assessment of RECIST will be presented. Median DoR will also be summarised calculated from the KM curve. Only patients who have a confirmed response will be included in this summary table.

4.2.6 Time from randomisation to second progression or death (PFS2)

Time from randomisation to second progression or death (PFS2) will be analysed using identical methods as outlined for PFS and adjusting for the same set of covariates, but no subgroup analysis will be performed. The HR for the treatment effect together with its 95% CI will be presented. Medians and Kaplan-Meier plots will be presented to support the analysis. The sensitivity analysis outlined in section 4.2.2.1 will not be repeated for PFS2 with the exception of a Kaplan-Meier plot of the time to censoring where the censoring indicator of PFS2 is reversed.

The number and percentage of patients experiencing a PFS2 event and the type of progression (Progression by disease specific biomarker CA-125, Symptomatic progression, Objective radiological progression, Other) will also be summarised by treatment arm, as well as summaries of deaths in the absence of second progression, and categories of PFS2 censoring. Time from randomisation to second progression will be summarised by treatment arm.

4.2.7 Time to first and second subsequent therapy or death and time to study treatment discontinuation or death

For supportive purposes, the time from the date of randomisation to the start of first subsequent therapy or death (TFST), time to second subsequent therapy (TSST) , as well as time to study treatment discontinuation or death (TDT) will be analysed using the same methodology and model as that used for the analysis of PFS. The HR for the treatment effect together with its 95% CI will be presented. In addition, medians and a Kaplan-Meier plot of the time to event will be presented by treatment arm.

Time between progression and starting subsequent therapy will be assessed. This will be summarised descriptively by treatment arm but no formal comparisons will be made. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

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

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

4.2.8 Patient reported outcomes (PROs)

For PRO summary by timepoint, all data including unscheduled assessment will be included according to the time window defined as described in section 4.2.10.1.

4.2.8.1 EORTC QLQ-C30 and EORTC QLQ EN24

Change from baseline

Scores from 0-100 will be derived for all multi-item or single item scales of the EORTC QLQ-C30 (except the single item scale financial difficulties) or QLQ-EN24.

For EORTC QLQ-C30, there are a total of 15 scales/scores:

- 5 functional scales in EORTC QLQ-C30: physical, role, cognitive, emotional, and social
- 3 multi-item symptom scales in EORTC QLQ-C30: fatigue, pain, and nausea/vomiting;
- 6 single item symptom scales in EORTC QLQ-C30: dyspnea, insomnia, appetite loss, constipation, diarrhea, financial difficulties

- Global health status/QoL scale in EORTC QLQ-30.

For EORTC QLQ-EN24, there are a total of 13 scales/scores:

- 3 functional scales in EORTC QLQ-EN24: sexual interest, sexual activity, sexual enjoyment
- 5 multi-item symptom scales in EORTC QLQ-EN24: lymphoedema, urological symptoms, gastrointestinal symptoms, poor body image, sexual/vaginal problems;
- 5 single item symptom scales in EORTC QLQ-EN24: pain in back and pelvis, tingling/numbness, muscular pain, hair loss, taste change

A summary of absolute and change from baseline for each EORTC QLQ-C30 (except for single item scale financial difficulties) and EORTC QLQ-EN24 scale will be reported by visit for each treatment group.

Change from baseline for each scale (14 scales of EORTC QLQ-C30 and 13 scales of EORTC QLQ-EN24) will be analysed using a mixed model for repeated measures (MMRM) of the change from baseline (defined as prior to first dose) in each scale for each visit for Durvalumab+Placebo versus Control as well as Durvalumab+Olaparib versus Control. The primary analysis will be to compare the average treatment effect from the point of randomisation for the first 12 months (which will include analysis visits obtained within the first 12 months, i.e. baseline, day 22 (week 3), weeks 6, 9, 12, 15, 18, 22, 26, 30, etc.) unless there is excessive missing data (defined as > 75% missing data).

The MMRM model will include patient, treatment, visit (analysis) and treatment by visit interaction as explanatory variables, the baseline score as a covariate along with the baseline score by visit interaction. Treatment, visit and treatment by visit interaction will be fixed effects in the model and patient will be included as an random effect. 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. The treatment by visit interaction will remain in the model regardless of significance.

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. The following provides sample code for implementing the MMRM analysis:

```
proc mixed data=scale_x method = reml;
class TRT VISIT SUBJECT;
model CBL = TRT VISIT TRT*VISIT BL BL*VISIT / s ddfm=kr;
repeated VISIT / type=UN subject=SUBJECT;
random intercept / subject= SUBJECT;
lsmeans TRT / at means pdiff diff alpha=0.05 cl;
```

where TRT is the randomised treatment, VISIT is the visit, CBL is the change from baseline score in the scale of interest, and BL is the baseline score of the scale of interest.

For the estimation of TRT*VISIT means an additional model will be run using all visits and the following lsmeans statement:

lsmeans TRT*VISIT / slice=VISIT pdiff diff alpha=0.05 cl;

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, and autoregressive. If there are still issues with the fit of the model or estimation of the treatment effects, subject will be treated as a fixed effect. For each treatment and visit, the adjusted (least squares) mean estimates, corresponding 95% CIs, estimates of the treatment difference, corresponding 95% CIs and p-values will be presented. 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.

Response (improvement, no change, deterioration)

At each post-baseline assessment, the change from baseline in each EORTC QLQ-C30 (except the single item scale financial difficulties) or EORTC QLQ-EN24 scale/item score will be categorized as improvement, no change, or deterioration as defined in Table 7. A summary table of the number (%) of subjects in each response category (improvement, no change, deterioration) for each scale will be presented for each visit by treatment group.

Table 7. Clinically meaningful change and visit response - EORTC QLQ-C30 and EORTC QLQ-EN24

Score	Change from baseline	Visit response
Health status/QoL scale and functional scales	≥10-point increase from baseline	Improvement
	≥10-point decrease from baseline	Deterioration
	<10-point increase from baseline	No change
	<10-point decrease from baseline	
	Missing visit response or missing baseline	Non-Evaluable [a]
Symptom scales	≥10-point increase from baseline	Deterioration
	≥10-point decrease from baseline	Improvement
	<10-point increase from baseline	No change
	<10-point decrease from baseline	
	Missing visit response or missing baseline	Non-Evaluable [a]

[a] Reason for non-evaluable visit response (i.e. due to missing baseline data, missing post-baseline data, missing due to withdrawal, or missing due to death) may be summarised, where required.

Time to deterioration

Time to deterioration (TTD) will be analysed for the following scale/score:

- Physical functioning score of the EORTC-QLQ-C30
- Role functioning score of the EORTC QLQ-C30
- Pain in back and pelvis of the EORTC QLQ-EN24
- Gastrointestinal (GI) symptoms of the EORTC QLQ-EN24
- Urological symptoms of the EORTC QLQ-EN24

Time to deterioration will be defined as:

Time to deterioration = date of event or censoring – date of randomisation + 1

where the event is the first clinically meaningful deterioration, i.e. an increase of ≥10 in the score from baseline for each symptom score and a decrease of ≥10 in the function scales or the global health status/QoL as defined in Table 6, that is confirmed at a subsequent visit (except if it was the patient's last available assessment) at least 14 days apart or death.

Patient who shows no clinical deterioration prior to the end of study will be censored at the last evaluable PRO assessment. If patient deteriorates or patient dies after 2 or more missed PRO assessment visit, then the patient will be censored at the last evaluable PRO assessment prior to the 2 missed visits.

Separate analyses will be conducted for each scale. In addition, sensitivity analyses will be performed whereby death is excluded as an event for TTD.

TTD will be analysed using a stratified log-rank test separately for Durvalumab+Placebo vs control and Durvalumab+Olaparib vs control as described for the primary PFS endpoint.

The effect of Durvalumab+Placebo vs control and Durvalumab+Olaparib vs control will be estimated by the HR together with its corresponding 95% CI and p value using the same methodology and model as that used for the analysis of PFS.

Kaplan-Meier plots will be presented by treatment group. A summary of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment group.

Only the subset of the FAS who have baseline scores of ≥10 will be included for the analysis of time to deterioration in functioning and a subset of the FAS who have baseline scores of ≤90 will be included for symptom deterioration analysis.

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

4.2.8.2 EQ-5D-5L

Descriptive statistics will be calculated for each scheduled visit/time point in the study, for each treatment group. The number of patients, the number of EQ-5D questionnaires

completed at each visit, the number and proportion responding to each dimension of the EQ-5D-5L will be summarised.

Additionally, summary statistics (e.g. n, mean, median, SD, min, max) will be reported for the EQ-5D index score and the EQ-VAS score, and the change from baseline for the EQ-5D index score and the EQ-VAS score.

Summary statistics will be supported by plots of mean and mean change from baseline in EQ-5D index score and EQ-VAS score, and associated 95% CIs, by time/scheduled visit and treatment arm. To support submissions to payers, additional analyses may be undertaken and these will be outlined in a separate Payer Analysis Plan.

4.2.8.3 PRO-CTCAE

PRO-CTCAE data will be summarised descriptively as the number and percent of subjects with each level of response for each PRO-CTCAE item at each visit by treatment group. A shift table of baseline evaluation to the worst post-baseline evaluation will also be produced.

4.2.8.4 PGIS

Responses on the PGIS will be summarised descriptively as the number and percent of subjects with each response at each visit by treatment group.

4.2.8.5 PGIC

Responses on the PGIC will be summarised descriptively as the number and percent of subjects with each response at each visit by treatment group.

4.2.8.6 PGI-TT

Responses on the PGI-TT will be summarised descriptively as the number and percent of subjects with each response at each visit by treatment group.

4.2.8.7 PGI-BR

Responses on each of the five questions (1 for the trial and 4 for the study medication) from the PGI-BR questionnaire will be summarised separately as the number and percent of subjects with each response at each visit by treatment group.

4.2.8.8 Compliance

For EORTC QLQ-C30, EORTC QLQ-EN24, and EQ-5D-5L, overall compliance rate and compliance or completion over time will be summarised descriptively for each treatment. For each summary table, number of patients received, expected and evaluable will also be included.

4.2.9 Health care resource use

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

4.2.10 Safety

Safety data will be presented using descriptive statistics unless otherwise specified. Safety and tolerability data will be presented by actual treatment group combining data from all cycles of treatment.

4.2.10.1 General considerations for safety and PRO assessments

Visit windows

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

- The time windows will be exhaustive so that data recorded at any time point has the potential to be summarised. Inclusion within the time window will be based on the actual date and not the intended date of the visit.

- All unscheduled visit data have the potential to be included in the summaries.

- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for vital signs data (with 3 and 4 weeks between scheduled assessments for chemotherapy phase and maintenance phase, respectively) are:

    - Day 1, visit window NA
    - Day 22, visit window 2 – 32
    - Day 43, visit window 33 – 53
    - Day 64, visit window 54 – 74
    - Day 85, visit window 75 – 95
    - Day 106, visit window 96 – 116
    - Day 127, visit window 117 – 140
    - Day 155, visit window 141 – 168
    - Day 183 visit window 169 – 196
    - Day 211, visit window 197 – 224

- In addition, for safety assessments only, the safety follow-up visit is defined as the latest of either 30 days following last dose of olaparib/placebo or 90 days following last dose of durvalumab/placebo, or until the initiation of the first subsequent therapy (including non-palliative radiotherapy), whichever occurs first.

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- Listings should display all values contributing to a time point for a patient.
- For visit based summaries
    - If there is more than one value per patient within a time window then the closest value to the scheduled visit date will be summarised, or the earlier, in the event the values are equidistant from the nominal visit date. The listings will highlight the value for the patient that contributed to the summary table, wherever feasible. Note: in summaries of extreme values all post baseline values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.
    - 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 summarised or plotted if the number of observations is greater than the minimum of 20 and > 1/3 of patients dosed.

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

Handling of missing data

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. Note that 0 should not be used as an imputed value in case the endpoint requires a log transformation. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

Generally, the imputation of dates is used to decide if an observation is treatment emergent for adverse events (AE) or concomitant medications (CM). The imputed dates are not advised to be used to calculate durations where the results would be less accurate.

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

- Partial or missing AE or CM start dates will be imputed as follows:

    a. Missing day - Impute the 1st of the month unless month is same as month of first dose of study drug then impute first dose date
    b. Missing day and month – impute 1st January unless year is the same as first dose date then impute first dose date
    c. Completely missing – impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date.

- Partial or missing AE or CM end dates will be imputed as follows:

    a. Missing day - impute the last day of the month. If the subject died in the same month, then set the imputed date as the death date
    b. Missing day and month – impute 31st December. If the subject died in the same year, then set the imputed date as the death date
    c. Completely Missing – No imputation

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

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

4.2.10.2 Adverse events (AEs)

All AEs, both in terms of MedDRA preferred term and CTCAE grade, will be listed and summarised descriptively by count (n) and percentage (%). The AE summaries, unless otherwise stated, will be based on treatment-emergent AEs up until the initiation of the first subsequent anti-cancer therapy following discontinuation of study treatment or until the end of safety follow-up period (latest of either 30 days following last dose of olaparib/placebo or 90 days following last dose of durvalumab/placebo), whichever occurs first. This will more accurately depict AEs attributable to only the study treatment as a number of AEs following discontinuation of the study treatment are likely to be attributable to subsequent therapy. In order to assess the longer-term toxicity profile, AE summaries will also be produced containing AEs observed up until the end of follow-up period (i.e. without taking subsequent therapy into account).

Any events in this period that occur after a patient has received further therapy for cancer will be flagged in the data listings.

A separate listing of AEs occurring more than 30 days after last dose of olaparib/placebo or 90 days after last dose of durvalumab/placebo (whichever is longer) will be produced. These events will not be included in the AE summaries.

Also, AEs occurring prior to start of study treatment will be listed separately, but will not be included in the summary tables (unless otherwise stated).

Key guidelines for counting incidence proportions of adverse events are as follows:

- When a patient has the same AE reported multiple times during an analysis period based on SOC and PT, the patient will only be counted once within a level of MedDRA in an AE incidence table.

- When assessing investigator-reported relationship to study drug of the AEs, if an AE changes in causal relationship during an analysis period for a patient, the related event will be chosen. Causally related AEs will include those reported as related by the investigator and those with a missing relationship. AE with a missing relationship will be presented as related in summary tables but will be presented in the data listing with a missing relationship.
- When summarising intensity of the AEs, if an AE changes in CTCAE grade during an analysis period for a patient, the AE with the maximum CTCAE grade will be chosen. In case the AE term (SOC and PT) is reported more than once, one of them with missing grade, and at least another with non-missing grade, the maximum CTCAE grade will be chosen from the non-missing grade values and the missing grade can be ignored. If all are of missing grade, then the AE severity will be summarised in an additional "Unknown" intensity category.

Treatment emergent adverse events will be presented for each treatment arm by MedDRA system organ class (SOC) and MedDRA preferred term (PT) covering number of patients reporting at least one event and number of events, where appropriate.

AEs will be assigned CTCAE grades (National Cancer Institute (NCI) CTCAE version 5.0) and summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, system organ class, preferred term and actual treatment arm. Fluctuations observed in CTCAE grades during study will be listed for those AEs which are CTCAE ≥ 3.

The number and percent of patients with TEAEs, as classified by SOC and PT, will be summarised by actual treatment arm for the safety analysis set. In addition, an adverse event of most common TEAEs, events that occur in at least 5% of patients in SAF will be summarised by preferred term, by decreasing frequency in the total column. The cut-off value of 5% may be modified after review of the data. When applying the cut-off value (i.e., 5%), the raw percentage will be compared to the cut-off value and no rounding will be applied (i.e., an AE with frequency of 4.9% will not be included in this table if a cut-off is 5%).

Summary information will be tabulated for:

- All TEAEs
- TEAEs by maximum CTCAE grade
- TEAEs with CTCAE grade 3 or 4
- TEAEs with maximum CTCAE grade 3 or higher
- TEAEs causally related to durvalumab/placebo*
- TEAEs causally related to olaparib/placebo*
- TEAEs causally related to carboplatin/paclitaxel*
- TEAEs causally related to any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)*
- TEAEs with maximum CTCAE grade 3 or 4, causally related to durvalumab/placebo*
- TEAEs with maximum CTCAE grade 3 or 4, causally related to olaparib/placebo*
- TEAEs with maximum CTCAE grade 3 or 4, causally related to carboplatin/paclitaxel*
- TEAEs with maximum CTCAE grade 3 or 4, causally related any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)*
- Most common TEAEs (occurring in at least 5% of subjects)
- TEAEs with outcome of death
- TEAEs with outcome of death causally related to durvalumab/placebo*
- TEAEs with outcome of death causally related to olaparib/placebo*
- TEAEs with outcome of death causally related to carboplatin/paclitaxel*
- TEAEs with outcome of death causally related to any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)*
- All serious treatment emergent adverse events (SAEs)
- All SAE causally related to durvalumab/placebo*
- All SAE causally related to olaparib/placebo*
- All SAE causally related to carboplatin/paclitaxel*
- All SAEs causally related to any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)*
- TEAEs leading to discontinuation of olaparib/placebo
- TEAEs leading to discontinuation of durvalumab/placebo
- TEAEs leading to discontinuation of carboplatin/paclitaxel
- TEAEs leading to discontinuation of any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)
- TEAEs leading to discontinuation of durvalumab/placebo, causally related to durvalumab/placebo*
- TEAEs leading to discontinuation of olaparib/placebo, causally related to olaparib/placebo*
- TEAEs leading to discontinuation of carboplatin/paclitaxel, causally related to carboplatin/paclitaxel*
- TEAEs leading to discontinuation of any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel), causally related to either durvalumab/placebo or olaparib/placebo or carboplatin/paclitaxel*
- Other significant AEs (OAEs)

- Immune mediated AEs (as determined by the reporting investigator)
- Infusion reaction AEs (as determined by the reporting investigator)

*Causality as determined by the reporting investigator

Summaries of the number and percentage of patients with AEs leading to dose modification of study medication by treatment arm will be presented for the following:

- AEs leading to a dose interruption of durvalumab/placebo
- AEs leading to a dose reduction of olaparib/placebo
- AEs leading to a dose interruption of olaparib/placebo
- AEs leading to a dose modification of olaparib/placebo, defined as a dose interruption or dose reduction of olaparib/placebo
- AEs leading to a dose interruption of carboplatin/paclitaxel
- AEs leading to a dose interruption of any treatment (durvalumab/placebo, olaparib/placebo, carboplatin or paclitaxel)

All AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and causal relationship to study drug. Adverse events with outcome of death, SAEs, AEs leading to discontinuation of treatment, AEs causally related to any study medication, and other significant adverse events (OAEs) will be listed separately.

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

Death

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

- Total number of death (regardless of date of death)
- Death related to disease under investigation only as determined by the investigator
- AE with outcome of death only
- Death related to disease under investigation and an AE with outcome of death
- Deaths after end of safety follow up period (as defined above), and not due to disease under investigation
- Unknown reason for death
- Other deaths

A list of all deaths will also be produced.

4.2.10.3 Adverse events of special interest (AESI) and adverse events of possible interest (AEPI)

Preferred terms used to identify AESI and adverse events of possible interest (AEPIs are for durvalumab only) will be listed before DBL and documented in the Study Master File.

For Durvalumab: Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI/AEPI grouping. Groupings will be based on preferred terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

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

- Any AESI/AEPI
- Any AESI/AEPI by PT and maximum CTCAE grade
- Any AESI/AEPI of maximum CTCAE grade 3 or 4
- Any serious AESI/AEPI
- Any AESI/AEPI with outcome of death
- Any AESI/AEPI causally related to study medication
- At least one AESI/AEPI leading to discontinuation of study medication
- Any AESI/AEPI leading to concomitant medication use (steroids)
- 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)

An overall AESI/AEPI summary will be presented, including number and percentage of patients in each of these categories.

For Olaparib: A separate summary table will also be produced capturing any toxicities of interest. Note the summary table capturing the toxicities of interest of MDS/AML and new primary malignancy includes events from first dose of study drug until the end of the study (i.e. not restricted to treatment emergent events).

The toxicities of interest are:

- Myelodysplastic syndrome/Acute Myeloid Leukemia
- New Primary Malignancy
- Pneumonitis

Immune-mediated Adverse events (imAEs)

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

4.2.10.4 Laboratory assessments

Data obtained up until the end of safety follow-up period (30 days following last dose of olaparib/placebo or 90 days following last dose of durvalumab/placebo) or until the initiation of the first subsequent therapy following discontinuation of treatment (whichever occurs first) will be used for reporting. This will more accurately depict laboratory toxicities attributable to only the study treatment as a number of toxicities in the follow-up period 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 90 days following last dose of durvalumab or up until 30 days following last dose of olaparib (i.e., without taking subsequent therapy into account).

Any data collected after the end of safety follow-up period will not be summarised.

A small selection of summaries of laboratory data may also be produced containing data from initiation of the first subsequent therapy following discontinuation of study treatment until 90 days following last dose of durvalumab or until 30 days following last dose of olaparib (i.e. summarising the laboratory data collected on patients taking subsequent therapy during the safety collection follow-up window post discontinuation of study treatment). These outputs will only be produced if the number of laboratory toxicities observed warrant the inclusion of such outputs for interpretational purposes.

Data summaries will be provided in International System (SI) of units.

Box-plots of absolute values and change from baseline for a selection of continuous laboratory assessments will be presented.

For all continuous laboratory assessments, absolute value, change from baseline and percentage change from baseline will be summarised using descriptive statistics at each scheduled assessment time by actual treatment arm. For categorical laboratory assessments, shift from baseline will be summarised using frequency and proportion at each scheduled assessment time by actual treatment arm.

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

- Haematology: Haemoglobin, Total white cell count (Leukocytes), Absolute lymphocyte count, Absolute neutrophil count, Platelet count
- Clinical chemistry: ALT, AST, ALP, LDH, Amylase, TSH, Total bilirubin, Gamma glutamyl transferase, Lactate Dehydrogenase, Lipase, Albumin– hypo and – hyper, Sodium – hypo and – hyper, Potassium – hypo and – hyper, Corrected calcium – hypo and – hyper, Glucose – hypo and – hyper, Magnesium- hypo and- hyper, Creatinine.

The denominator used in shift tables will only include evaluable patients, i.e., those who had sufficient data to have the possibility of an abnormality.

For example:

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

Urinalysis results (categorical data collected at baseline and only if clinically indicated post baseline) will be listed only.

Clinically significant laboratory results will be flagged and listed. Reference ranges will also be listed. All laboratory summaries and listings will be presented by actual treatment arm.

Liver Function

For liver biochemistry, the following summaries will include the number and percent of patients who have:

- Elevated ALT, AST, and Total bilirubin during the study
  - ALT ≥ 3x –≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x and >20x Upper Limit of Normal (ULN) during the study
  - AST ≥ 3x- –≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x,and >20x ULN during the study
  - Total bilirubin ≥2x-≤3x, >3x-≤5x, >5x ULN during the study
  - ALT or AST ≥3x - ≤5x, >5x - ≤8x, >8x - ≤ 10x, >10x - ≤ 20x, and >20x ULN during the study
  - ALT or AST ≥3x ULN and Total bilirubin ≥2x ULN during the study (Potential Hy's law: The onset date of ALT or AST elevation should be prior to or on the date of Total Bilirubin elevation)

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

Plots of ALT and AST vs. Total bilirubin all expressed as multiples of ULN by treatment group will also be produced with reference lines at 3×ULN for ALT, AST, and 2×ULN for Total bilirubin. In each plot, Total bilirubin will be in the vertical axis.

Assessment of Thyroid Function Test Results

The following summaries will include the number and percentage of patients who have elevated or low TSH.

- TSH > ULN
- TSH > ULN with TSH ≤ ULN at baseline

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

A separate summary will present:

- Number of subjects with at least one baseline and post-baseline TSH result
    - On-treatment elevated TSH > ULN and above baseline
    - On-treatment decreased TSH < LLN and below baseline
- Grade change from baseline to on treatment minimum and maximum

Absolute value and change from baseline of TSH, free T3 and free T4 will be summarised using descriptive statistics at each scheduled assessment time.

Assessment of Renal Function Test Abnormalities

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

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

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

Pregnancy

A listing of pregnancy results and, if applicable, a pregnancy report will be provided for safety analysis set.

4.2.10.5 Vital signs

Vital signs (pulse rate, systolic blood pressure (SBP), diastolic blood pressure (DBP), temperature, weight) will be summarised over time in terms of absolute values and changes from baseline at each scheduled visit.

For assessments related to the first infusion of durvalumab/placebo, data for the 3 timepoints (pre-dose, during and after infusion) will also be summarised. Change from pre-dose will be also summarised for during and after infusion vital signs.

Vital signs data will be listed by actual treatment arm.

4.2.10.6 ECG

ECG data obtained before study treatment and up until the 30-day (for olaparib) or 90-day (for durvalumab) safety follow-up visit will be listed for SAF. The listing will include the ECG interpretation (Normal, Borderline, Abnormal Not Clinically Significant, and Abnormal Clinically Significant).

4.2.10.7 Physical examination

A listing of physical examination results will be provided for SAF.

4.2.10.8 ECOG performance status

ECOG performance status data will be listed for SAF. A shift table of baseline evaluation to worst evaluation will be produced separately for each treatment arm.

4.2.11 Demographics and baseline characteristics

The following will be summarised for all patients in the FAS by randomised treatment arm except for patient disposition, in which case all patients set will be used:

- Patient disposition (including screening failures, randomised, received treatment in chemotherapy phase, received treatment in maintenance phase, discontinued treatment, reasons for treatment discontinuation, terminated from study and reasons for study termination).
- Important protocol deviations
- Inclusion in analysis sets
- Demographics (age, age group [< 65, vs ≥ 65 years], sex, race and ethnicity).
- Patient characteristics at baseline (height, weight, BMI (weight (kg)/[height (m)]^2))
- Stratification factors as per eCRF and randomisation
- Patient recruitment by country and centre
- Previous disease-related treatment modalities
- Prior cancer therapy
- Prior radiotherapy
- History of blood transfusions
- Post-discontinuation cancer therapy
- Disease characteristics at baseline (ECOG performance status, primary tumour location, histology type, tumour grade and overall disease classification, Baseline CA125 value)
- Extent of disease at baseline
- FIGO stage at baseline
- Disease related medical history (past and current)
- Relevant surgical history
- Allowed concomitant medications
- Disallowed concomitant medications

- MMR expression (proficient vs deficient), PD-L1 expression (percent of PD-L1 positive cells), HRRm status or any other biomarker data deemed important for this study.

Medical history will be coded using the latest version of MedDRA. The number and percentage of patients with medical history will be summarised by SOC and PT and presented for each treatment arm.

4.2.12 Concomitant and other treatments

Information on any treatment within the four weeks prior to initiation of study drug and all concomitant treatments given up to end of the clinical treatment phase of the study including the follow-up period following the last dose of study drug (30 days after last dose for olaparib/placebo, or 90 days for durvalumab/placebo) will be recorded in the eCRF. Thereafter, only subsequent regimens of anti-cancer therapy will be recorded in eCRF. Other anti-cancer therapies, investigational agents, and radiotherapy should not be given while the patient is on study drug. For a detailed list of restricted and prohibited medications please refer to CSP Section 6.5.2.

Medications received prior to, concomitantly, or post-treatment will be coded using the latest version of WHODrug. Concomitant medications will be summarised for the FAS by Anatomical Therapeutic Chemical category name (ATC level 4) and generic term (preferred term) by randomised treatment arm. Missing coding terms should be listed and summarised as "Not coded".

For the purpose of inclusion in prior and/or concomitant medication summaries, incomplete medication start and stop dates will be imputed as detailed in Section 4.2.10.1.

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

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

The following summaries will be produced:

- Summary of prior medications
- Summary of concomitant medications
- Summary of post treatment medications

4.2.13 Exposure

Exposure will be summarised for the safety analysis set by actual treatment arm. The following summaries will be produced:

- Total exposure duration of durvalumab/placebo, olaparib/placebo, carboplatin and paclitaxel
- Actual exposure duration of durvalumab/placebo and olaparib/placebo (regardless of dose reduction)
- RDI of durvalumab/placebo and olaparib/placebo.
- For Durvalumab/Placebo, summary of dose delay will be produced.
- For olaparib/placebo, summary of dose reductions and dose interruptions together with reasons for each dose modification will be produced.

Dose reductions and dose interruptions will be based on investigator initiated dosing decisions. Dose interruptions of olaparib due to "Subject Forgot to Take Dose" will be omitted from these summaries.

Overdose, defined as any dose or frequency of dosing that exceeds the dosing regimen as specified in the CSP. Possible overdose will be listed.

Subsequent Therapy

Subsequent cancer therapies received and together with the best response after discontinuation of study treatment will be summarised by randomised treatment arm.

4.2.14 Pharmacokinetic data (durvalumab only)

Plasma concentrations of durvalumab will be summarised by actual treatment group and per nominal sample time using standard summary statistics for PK concentrations (geometric mean, geometric coefficient of variation, geometric mean ± geometric standard deviation, arithmetric mean, standard deviation, coefficient of variation, minimum, maximum and n). All plasma concentrations will be listed.

- Data from patients excluded from the PK analysis set will be included in the data listings, but not in the summaries.
- Extra measurements (such as unscheduled or repeat assessments) will also not be included in summary tables but will be included in patient listings.
- For all data, descriptive statistics will follow the rounding convention of the individual data. Coefficients of variation (%CV and %GCV), where reported, will always be reported to 1 decimal place.

For descriptive statistics of serum concentrations, non-quantifiable (NQ) values of plasma concentrations will be handled as follows:

- If, at a given time point, 50% or less of the serum concentrations are NQ, the mean, standard deviation, geometric mean, geometric standard deviation, %CV, and %GCV will be calculated by substituting the lower limit of quantification (LLOQ) for values which are NQ.
- If more than 50%, but not all, of the concentrations are NQ, the mean, geometric mean, standard deviation, geometric standard deviation, %CV, and %GCV will be reported as not calculable (NC). The maximum value is reported from the individual data, and the minimum and median are set as NQ.
- If all the concentrations are NQ, the geometric mean and mean will be reported as NQ and the standard deviation, geometric standard deviation, %CV, and %GCV as NC.
- Three observations > LLOQ are required as a minimum for a serum concentration to be summarised. Two observations > LLOQ are presented as minimum and maximum with the other summary statistics as NC. The lower limit of quantitation will be reported in the CSR and in serum concentration tables, figures and listings.
- If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modelling approach.

4.2.15 Immunogenicity data (durvalumab only)

Immunogenicity results of all patients in the ADA evaluable population will be listed. The number and percentage of patients who develop detectable ADA to durvalumab within each ADA response category listed in section 3.7 will be summarised based on the ADA-evaluable population. The immunogenicity titer and neutralising ADA data will be listed for samples confirmed positive for the presence of anti-durvalumab antibodies. Details for the presentation and derivation of ADA data is provided in section 3.7.

The effect of immunogenicity as well as the effect of its neutralising properties on PK, pharmacodynamics, efficacy, and safety will be evaluated, if the data allow. A detailed plan will be written by the AstraZeneca Clinical Pharmacology group or designee.

4.2.16 Exploratory analyses

Summary and analyses for exploratory biomarkers and genetic data will be documented in a separate analysis plan and may be reported outside the CSR.

5. INTERIM ANALYSES

5.1 Futility and Interim Analyses

A futility analysis of PFS for the comparison of the Durvalumab+Placebo arm versus Control and the Durvalumab+Olaparib arm versus Control will be performed approximately 2-months post-LSR, and when a minimum of 50% of the target number of PFS events for each comparison has occurred (150 of 299 target events across the Durvalumab+Placebo and control arms, and 141 of 281 target events across the Durvalumab+Olaparib and control arms)

(approximately 25 months after the first patient has been randomised). The boundary for declaring futility and dropping an experimental arm will be observing a HR>1.15. The futility analyses will be performed by an IDMC (see Section 5.2).

In addition to the futility analysis, 2 interim OS analyses are planned. The first interim OS analysis will be performed at the time of the primary analysis of PFS when approximately 70% of the target number of OS events would have occurred (i.e., 208 of 280 OS events per comparison [Durvalumab+Placebo versus Control and Durvalumab+Olaparib versus Control]).

A second interim analysis of OS may be performed at the same calendar time when approximately 244 OS events (87% of the target number of OS events) have occurred for the comparison of the Durvalumab+Placebo arm versus the control arm, as well as the Durvalumab+Olaparib arm versus Control (approximately 51 months after the first patient is randomised). Multiplicity adjustments for these interim analyses and the stopping rule are discussed in Section 4.2.1.

Note: If the OS interim results do not meet the criterion of stopping for superiority for a given hypothesis, then follow-up will continue until the final target number of OS events for that comparison has been observed, following which, the hypothesis will be re-tested.

5.2 Independent Data Monitoring Committee

This study will use an external IDMC to assess ongoing safety analyses as well as for the PFS futility efficacy analysis. The committee will meet after the first 30 patients complete 1 cycle of study treatment to review the safety data from the study. The IDMC will meet at least every 6 months thereafter until the primary PFS analysis. Following each meeting, the IDMC will report to the sponsor and may recommend changes in the conduct of the study.

This committee will be composed of therapeutic area experts and biostatisticians, who are not employed by AstraZeneca and are free from conflict of interest.

Following the reviews, 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. 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. The sponsor or IDMC may call additional meetings if at any time there is concern about the safety of the study.

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

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

6. CHANGES OF ANALYSIS FROM PROTOCOL

Not applicable.

7. CHINA COHORT

The global recruitment into this study will close to all sites apart from China when approximately 699 patients have been randomised. Any patient from China, randomised before the global recruitment is closed (ie, last subject randomised from a non-Chinese site) will be included in both the global ITT population and the China cohort ITT population. A patient randomised in China after the global recruitment closure will be included only in the China cohort ITT population.

Approximately 129 patients from sites in China will be recruited and randomised in a 1:1:1 ratio to the study treatments and will follow the same study plan and procedures as patients recruited to the global study. The safety and efficacy data collected will be summarised and analysed separately to the global study safety and ITT analysis sets (as defined in Section 2.1).

The primary analysis of efficacy for the China cohort will be an assessment of programmatically derived PFS based on investigator assessments (RECIST 1.1) in the China cohort ITT population (China FAS). The China FAS comprises all patients from sites in China who are randomised regardless of whether they receive treatment or not. The data cut-off for the analysis of PFS in the China cohort will be undertaken at the same calendar time when approximately 55 PFS events have occurred (64% maturity) for the comparison of the Durvalumab+Placebo arm versus the control arm and approximately 52 PFS events have occurred (60% maturity) for the comparison of the Durvalumab+Olaparib arm versus the control arm. Where data permits, summaries and analysis of secondary supportive efficacy endpoints (including at least but not limited to OS) will be performed for the China cohort. The detailed analysis plan will be documented in the China supplementary SAP.

When assessing safety and tolerability, summaries will be produced separately for the China cohort based on the China safety analysis set. The China safety analysis set includes all subjects from sites in China who receive any amount of study treatment (ie, durvalumab/placebo or olaparib/placebo). The China safety data will be summarised descriptively and will not be formally analysed.

8. REFERENCES

Aaronson et al 1993
Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JCJM, Kaasa S, Klee MC, Osoba D, Razavi D, Rofe PB, Schraub S, Sneeuw KCA, Sullivan M, Takeda F. The European Organisation for Research and Treatment of Cancer QLQC30:A quality-of-life instrument for use in international clinical trials in oncology. Journal of the National Cancer Institute 1993; 85: 365-376.

Amit et al 2010
Amit O, Bushnell W, Dodd L, Roach N, Sargent D. Blinded independent central review of the progression-free survival endpoint. Oncologist, 2010, 15(5), 492-495.

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med. 2009; 28:739-761.

Collet 2003
Collet D. (2003) Modelling Survival Data in Medical Research 2nd edn. Chapman & Hall/CRC.

Eisenhauer et al 2009
Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009 Jan;45(2):228-47.

Fayers et al 2000
Fayers PM, Machin D. Quality of Life: Assessment, Analysis and Interpretation. J Wiley & Sons Ltd, Chichester, 2000. ISBN: 0-471-96861-7.

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001.

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

Gail and Simon 1985
Gail M, Simon R. Testing for qualitative interactions between treatment effects and subject subsets. Biometrics. 1985; 41:361-72.

Glimm et al 2009
Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group‐sequential trials. Statistics in Medicine 2009. 29; 2.

Greimel et al 2011
Greimel E, Daghofer F, Petru E. (2011). Prospective assessment of quality of life in long-term ovarian cancer survivors. International Journal of Cancer. 2011; 128(12):3005-3011.

Hsu and Taylor 2009
Hsu CH and Taylor J: Nonparametric comparison of two survival functions with dependent censoring via nonparametric multiple imputation, Statist. Med. 2009; 28:462–475.

Jackson et al 2014
Jackson D, White IR, Seaman S, Evans H, Baisley K, Carpenter J. Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation, Statist. Med. 2014; 33(27): 4681-94

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.

Patil et al 2012
Patil S, Figlin RA, Hutson TE, Michaelson MD, Negrier S, Kim ST, Huang X, Motzer RJ. Q-TWiST analysis to estimate overall benefit for patients with metastatic renal cell carcinoma treated in a phase III trial of sunitinib vs interferon-a Br J Cancer. 2012 May 8;106(10):1587-90.

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.

van Hout et al 2012
van Hout B., Janssen,M.F., Feng,Y.S., Kohlmann,T., Busschbach,J., Golicki,D., Lloyd,A., Scalone,L., Kind,P., Pickard,A.S. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012 Jul-Aug;15(5):708-15.

9. APPENDICES

9.1 Scoring algorithm for the EORTC-QLQC30

On the EORTC QLQ-C30 version 3 questionnaire, all items from 1 to 28 (Table 8) are four point scales with response categories 1 = “Not at all”, 2 = “A little”, 3 = “Quite a bit” and 4 = “Very much”. Items 29 and 30 that constitute the rating of a patient’s overall health status and quality of life are assessed on a 7-point Likert scale 1 – 7 with 1 = “Very poor” and 7 = “Excellent”.

All of the scales and single-item measures will be transformed to range in score from 0 to 100 (Fayers et al 2001). A high scale score represents a higher response level, that is, a higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms. Thus;

- A high score for a functional scale represents a high / healthy level of functioning,
- A high score for the global health status / QoL represents a high QoL,
- A high score for a symptom scale / item represents a high level of symptomatology problems.

Table 8 Scoring the QLQ-C30 version 3.0

Domain	Number of items	Item numbers
Global health status / QoL		
Global health status/QoL	2	29, 30
Functional scales		
Physical functioning	5	1, 2, 3, 4, 5
Role functioning	2	6, 7
Emotional functioning	4	21, 22, 23, 24
Cognitive functioning	2	20, 25
Social functioning	2	26, 27
Symptom scales / items		
Fatigue	3	10, 12, 18
Nausea and vomiting	2	14, 15
Pain	2	9, 19
Dyspnea	1	8
Insomnia	1	11
Appetite loss	1	13
Constipation	1	16
Diarrhea	1	17
Financial difficulties	1	28

The principle for scoring these scales is the same in all cases:

Step 1: Estimate the average of the items that contribute to the scale; this is the raw score. If items $Q_1, Q_2, \ldots, Q_n$ are included in a scale, then for that scale, the *RawScore, RS*, is derived as the mean of the component items;

$$RawScore = RS = \frac{(Q1 + Q2 + \ldots + Qn)}{n}$$

Step 2: Use a linear transformation to standardize the raw score, so that scores range from 0 to 100;

Then for Functional scales:

$$Score = \left\{1 - \frac{(RS-1)}{range}\right\} x\ 100$$

and for Symptom scales/items and Global health status / QoL:

$$Score = \left\{\frac{(RS-1)}{range}\right\} x\ 100$$

Note: *range* is the difference between the maximum possible value of the raw score *RS* and the minimum possible value. Therefore, the range of *RS* equals the range of the item values. For most items (1 to 28) scored 1 to 4, the *range* is 3 (=4-1). The exceptions are the items contributing to the global health status / QoL, which are 7-point questions with *range* = 6.

Example:

Emotional functioning which is derived from items $Q_{21}, Q_{22}, Q_{23}, Q_{24}$,

$$RawScore = RS = \frac{(Q21 + Q22 + Q23 + Q24)}{4}$$

and then the Emotional functioning transformed score

$$Score = \left\{1 - \frac{(RS-1)}{3}\right\} x\ 100$$

Fatigue which is derived from items Q_{10}, Q_{12}, Q_{18},

$$RawScore = RS = \frac{(Q10 + Q12 + Q18)}{4}$$

and then the Fatigue functioning transformed score

$$Score = \left\{\frac{(RS-1)}{range}\right\} x\ 100$$

Missing data

Missing data may be classified as either missing items (one or more missing answers to questions within a questionnaire), or missing forms (the whole questionnaire is missing for a patient). Fayers and Machin (2000) (Fayers et al 2000) describe methods of analysis to use when data on response to items are missing, including imputation techniques. This involves applying the equations already given under "Scoring procedures" for calculating the scale scores; the missing items in multi-item scales are simply ignored when making the calculations when less than half the items are completed.

The protocol for missing items then can be summarised as;

- Have at least half of the items from the multi-item scale been answered?

    o If Yes, use all the items that were completed, and apply the standard equations given on the previous pages for calculating the scale scores; ignore any items with missing values when making the calculations.

    o If No, set scale score to missing.

- For single-item measures, set score to missing

9.2 Score algorithm for the EORTC-EN24

The QLQ-EN24 incorporates 5 multi-item scales to assess lymphoedema, urological symptoms, gastrointestinal symptoms, body image and sexual/vaginal problems. In addition, 8 single items assess pain in back and pelvis, tingling/numbness, muscular pain, hair loss, taste change, sexual interest, sexual activity and sexual enjoyment. On the EORTC QLQ-EN24 questionnaire, all items (Table 9) are four point scales with response categories 1 = "Not at all", 2 = "A little", 3 = "Quite a bit" and 4 = "Very much".

Table 9 Scoring the EORTC QLQ-EN24

Domain	Number of Items	Item Number
Functional scales		
Sexual interest	1	49
Sexual activity	1	50
Sexual enjoyment	1	54
Symptom scales		
Lymphoedema	2	31-32
Urological symptoms	4	34-37
Gastrointestinal symptoms	5	38-42
Poor body image	2	47–48
Sexual/vaginal problems	3	51–53
Pain in back and pelvis	1	33
Tingling/numbness	1	43
Muscular pain	1	44
Hair loss	1	45

Domain	Number of Items	Item Number
Taste change	1	46

All of the scales and single-item measures will be transformed to range in score from 0 to 100. A high score for the functional items represents a high level of functioning, whereas a high score for the symptom scales and single-items represents a high level of symptomatology or problems. The principle for scoring these scales is the same in all cases:

Step 1: Estimate the average of the items that contribute to the scale; this is the raw score. If items $Q_1, Q_2, \ldots, Q_n$ are included in a scale, then for that scale, the *RawScore, RS*, is derived as the mean of the component items;

$$RawScore = RS = \frac{(Q1 + Q2 + \ldots + Qn)}{n}$$

Step 2: Use a linear transformation to standardize the raw score, so that scores range from 0 to 100;

$$Score = \left\{\frac{(RS-1)}{range}\right\} x\ 100$$

Note: *range* is the difference between the maximum possible value of the raw score *RS* and the minimum possible value. Therefore, the range of *RS* equals the range of the item values.

Missing data

Missing data will be handled the same as EORTC QLQ-C30 (see Appendix 9.1 for details).

9.3 Scoring algorithm for the PRO-CTCAE

The items selected for this study will be assessed relative to one or more distinct attributes, including frequency and severity. For each item, responses are provided on a 5-point Likert scale with corresponding response choices for frequency and severity.

For Frequency (Shivering or shaking chills):

- 0 = Never
- 1 = Rarely
- 2 = Occasionally
- 3 = Frequently
- 4 = Almost constantly

For Severity (Itchy skin or Shivering or shaking chills)

- 0 = None
- 1 = Mild
- 2 = Moderate
- 3 = Severe
- 4 = Very severe

There are no guidelines yet established for how to combine attributes into a single score so each item will be evaluated separately.

Missing data

No imputation will be made for missing items. The proportion of missing data should also be summarised to aid interpretation.

9.4 PGI-BR

The PGI-BR contain 5 items: one question is for the experience of the clinical trial without considering study medication and 4 questions are for the experience/opinion of the study medication. The responses for each item are as the following:

Experience participating in the clinical trial

- Very negative
- Somewhat negative
- Neither negative nor positive
- Somewhat positive
- Very positive

Has study medication helped your condition

- Not at all
- A little bit
- Somewhat
- Quite a bit
- Very much

Side effects of the study medication

- None
- Not at all bad
- Slightly bad
- Moderately bad
- Very bad
- Extremely bad

Convenience of the study medication

- Very inconvenient
- Somewhat inconvenient
- Neither convenient nor inconvenient
- Somewhat convenient
- very convenient

Positive and negative things about the study medication

- Negative far greater
- Negative somewhat greater
- Equal
- Positive somewhat greater
- Positive far greater

9.5 COVID-19

9.5.1 Definitions and Derivations

The following definitions will be used to identify patients with coronavirus disease 2019 (COVID-19) adverse events:

- **Confirmed or Suspected COVID-19 AEs**: All AEs within the AE search criteria developed by the latest MedDRA MSSO guidance for COVID-19 have been reviewed by AstraZeneca Global Patient Safety and a list of terms pre-defined. A further review will take place prior to DBL to ensure any further terms not already included are captured.
- **COVID-19 Associated Adverse Events:** All confirmed and suspected COVID-19 AEs defined above, plus all other AEs occurring within <7 days before and <30 days after the start date of all the confirmed COVID-19 events.

9.5.2 Presentation

Depending on the extent of any impact, summaries of data relating to patients diagnosed with coronavirus disease 2019 (COVID-19), and impact of COVID-19 on study conduct (in particular missed visits, delayed or discontinued study treatment, and other protocol deviations) may be generated, by treatment group, including:

- Disposition (discontinued treatment due to COVID-19 and withdrew study due to COVID-19)
- Deviations (indicating whether a deviation is due to COVID-19 or not)
- Summary of disruptions due to COVID-19 (visit impact, drug impacted)
- Listing for patients affected by the COVID-19 pandemic

- Listing for patients with reported issues in the Clinical Trial Management System due to the COVID-19 pandemic.

In addition, if there is sufficient number of patients (e.g. ≥5 and/or ≥ 2% of the patient population) with an event of interest then the following may be generated, else AE listings/narratives will only be generated:

Demographic and baseline characteristics

- Summaries of demographics and baseline characteristics repeated within the subset of patients with confirmed/suspected COVID-19 infection.
- Summaries of medical history repeated within the subset of patients with confirmed/suspected COVID-19 infection.

Efficacy

- **PFS:** A sensitivity analysis may be conducted to assess for the potential impact of COVID deaths on PFS. This will be assessed by repeating the primary PFS analysis except that any patient who had a PFS event due to death where primary/secondary cause of death was due to COVID-19 Infection, or a COVID-19 infection reported as a fatal AE, will be censored at their last evaluable assessment prior to their COVID infection death date
- **OS:** A sensitivity analysis to assess for the potential impact of COVID deaths on OS. This will be assessed by repeating the OS analysis except that any patient who had a death with primary/secondary cause as COVID-19 Infection, or a COVID-19 infection reported as a fatal AE will be censored at their COVID infection death date.

Safety

- The number of patients with Confirmed/Suspected COVID-19 infection and Confirmed/Suspected COVID-19 deaths
- Summary in subjects with / without Confirmed/Suspected COVID-19 infection
    - Overall TEAE summary table
    - TEAEs by SoC and PT by max grade
- Summary of TEAEs associated with COVID-19
- Summary of TEAEs (excluding AEs associated with COVID-19 infection)
- Summary of Confirmed/Suspected COVID-19 TEAEs
- Summary of TEAEs (excluding Confirmed/Suspected COVID-19 infection AEs)
- Summary of TEAEs associated with COVID-19 infection resulting in death
- Summary of TEAEs resulting in death (excluding AEs associated with COVID-19 infection)

- Summary of TEAEs associated with COVID-19 infection leading to study treatment discontinuation
- Summary of TEAEs leading to study treatment discontinuation (excluding AEs associated with COVID-19 infection)

Furthermore, patient narratives in all patients with Confirmed/Suspected COVID-19 SAEs will be generated. Also a separate AE listing of patients with Confirmed/Suspected COVID infection will also be generated, as required.

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: d9311c00001-sap-ed-4		
Document Title:	Statistical Analysis Plan Edition 4	
Document ID:	Doc ID-004242269	
Version Label:	4.0 CURRENT LATEST APPROVED	
Server Date **(dd-MMM-yyyy HH:mm 'UTC'Z)**	**Signed by**	**Meaning of Signature**
29-Nov-2022 15:10 UTC	PPD	Author Approval
29-Nov-2022 13:43 UTC		Content Approval
30-Nov-2022 02:56 UTC		Qualified Person Approval

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