AstraZeneca

Statistical Analysis Plan

Study Code	D933IC00003
Edition Number	3.0
Date	4th October 2019

A Phase II, Randomized, Multi-Center, Double-Blind, Comparative Global Study to Determine the Efficacy and Safety of Durvalumab in Combination With Olaparib for First-Line Treatment in Platinum-Ineligible Patients With Unresectable Stage IV Urothelial Cancer

A Phase II, Randomized, Multi-Center, Double-Blind, Comparative Global Study to Determine the Efficacy and Safety of Durvalumab in Combination With Olaparib for First-Line Treatment in Platinum-Ineligible Patients With Unresectable Stage IV Urothelial Cancer

Study Statistician

PPD

Date

A Phase II, Randomized, Multi-Center, Double-Blind, Comparative Global Study to Determine the Efficacy and Safety of Durvalumab in Combination With Olaparib for First-Line Treatment in Platinum-Ineligible Patients With Unresectable Stage IV Urothelial Cancer

Global Product Statistician

PPD

Date

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ADA	anti-drug antibody
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AML	acute myeloid leukemia
AST	aspartate aminotransferase
BICR	Blinded Independent Central Review
BID	twice daily
BoR	best overall response
BP	blood pressure
BRCA	breast cancer susceptibility gene
BRCAm	breast cancer susceptibility gene-mutated
CI	confidence interval
CR	complete response
CrCl	creatinine clearance
CRO	contract research organization
CSP	clinical study protocol
CSR	clinical study report
CT	computerized tomography
CTCAE	Common Terminology Criteria for Adverse Event
ctDNA	circulating tumor deoxyribonucleic acid
DCO	data cutoff
DoR	duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire

Abbreviation or special term	Explanation
ePRO	Electronic tablet
EQ-5D-5L	EuroQol 5-dimension, 5 level health state utility index
FAS	full analysis set
FFPE	formalin-fixed paraffin-embedded
HL	Hy's law
HR	hazard ratio
HRQoL	health-related quality of life
HRR	homologous recombination repair
HRRm	homologous recombination repair mutant
HRRwt	homologous recombination repair wild type
IDMC	Independent Data Monitoring Committee
ILD	interstitial lung disease
IC+	immune cells with staining at any intensity above background
ICP	immune cells present
IO	Immuno-oncology
IP	investigational product
ITT	Intent-to-Treat
IV	Intravenous
IVRS	interactive voice response system
IWRS	interactive web response system
KM	Kaplan-Meier
LIMS	Laboratory Information Management System
MDS	myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
NCI	National Cancer Institute
NE	not evaluable
NTL	non-target lesion
OAE	Other significant adverse events
ORR	objective response rate

Abbreviation or special term	Explanation
OS	overall survival
OS18	subjects alive at 18 months
PD	progressive disease
PD-L1	programmed cell death ligand 1
PFS	progression-free survival
PFS6	alive and progression free at 6 months
PGIC	Patient Global Impression of Change
PHL	potential Hy's law
PK	pharmacokinetic(s)
PR	partial response
PRO	patient-reported outcome
q12w	every 12 weeks
q28days	every 28 days
q4w	every 4 weeks
q8w	every 8 weeks
QoL	Quality of Life
QTc	QT interval corrected for heart rate
RECIST 1.1	Response Evaluation Criteria in Solid Tumors, version 1.1
SAE	serious adverse event
SAF	safety analysis set
SAP	statistical analysis plan
SD	stable disease
SoA	Schedule of Activities
TBL	total bilirubin
TCC	transitional cell carcinoma
TCGA	The Cancer Genome Atlas
TKI	tyrosine kinase inhibitor
TL	target lesion
TNM	Classification of Malignant Tumors (Tumor, Lymph Nodes, Metastasis)
UC	urothelial cancer
ULN	upper limit of normal

Abbreviation or special term	Explanation
WHO	World Health Organization

AMENDMENT HISTORY

Date	Brief description of change
17th October 2018	SAP updated to include changes for v2 of CSP. Analyses changed to take into account the revised study objectives and endpoints in terms of using the Full Analysis Set as the primary analysis population rather than the HRR mutant subgroup population.

Category change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Other	04-Oct-19	SAP author handover from PPD to PPD	NA	NA
Other	04-Oct-19	Change to list of important protocol deviations (Section 2.2.1)	NA	For consistency with NCHP
Other	04-Oct-19	Updated 'patient' to 'subject'	NA	For consistency with ICH guidelines
Derivation of primary or secondary endpoints	04-Oct-19	Definition of PK analysis set updated to match protocol (Section 2.1.3)	Y (v2)	For consistency with protocol
Primary or secondary endpoints	04-Oct-19	Removed references to confirmation of response / confirmed response as confirmation of response not required for this study	N	RECIST 1.1 criteria state that in randomized trials confirmation of response is not necessary

Derivation of primary or secondary endpoints	04-Oct-19	Clarification added that randomization date will be used to derive study day for RECIST efficacy (Section 3.2.2)	NA	Clarification on derivation
Derivation of primary or secondary endpoints	04-Oct-19	More detail added to best objective response section (Section 3.2.8)	NA	Clarification on derivation
Derivation of primary or secondary endpoints	04-Oct-19	Detail added for best percentage change in TL, including imputation rules (Section 3.2.9)	NA	Clarification on derivation
Other	04-Oct-19	Removed Time to Subsequent Therapy from discontinuation of study treatment section (SAP v2 Section 4.2.9.9)	NA	Not required
Primary or secondary endpoints	04-Oct-19	PRO sections revised: Removed EORTC-QLQ-C30 symptom improvement and deterioration. Added definition for overall compliance rate, Line plots for EORTC QLQ-C30, EQ-5D VAS.	N	Removed time to deterioration and improvement rate analyses
Derivation of primary or secondary endpoints	04-Oct-19	Visit window definitions updated (Section 3.3.1)	NA	Correction to derivation
Derivation of primary or secondary endpoints	04-Oct-19	For best TL change, include visits up to **and including** progression (Section 3.2.9)	NA	Correction to derivation

Derivation of primary or secondary endpoints	04-Oct-19	Derivation of Metastatic disease for Bajorin risk index updated	NA	Correction to derivation
Data presentations	04-Oct-19	3 subgroups added: 1. smoking vs non smoking 2. prior treatment vs no prior treatment 3. prior IO therapy vs no prior IO therapy (Section 4.2.2)	NA	Additional subgroups of interest
Data presentations	04-Oct-19	OS attrition bias sensitivity analysis removed	N	Analysis no longer required
Data presentations	04-Oct-19	Immune mediated AEs removed (imAE)	N	Analysis no longer required
Data presentations	04-Oct-19	Reduced list of separate AESI tables and included extra categories in the high level AE summary table	N	Separate summaries of these AESI categories not required by SOC/preferred term

1. STUDY DETAILS

1.1 Study objectives

Table 1 Study objectives

Primary Objective:	**Endpoint/Variable:**
To assess the efficacy of durvalumab + olaparib combination therapy compared with durvalumab + placebo in terms of PFS	PFS as determined by Investigator assessment according to RECIST 1.1
Secondary Objectives:	**Endpoint/Variable:**
Key Secondary Objective:	
To assess the efficacy of durvalumab + olaparib combination therapy compared with durvalumab + placebo	OS
Additional Secondary Objectives:	
To assess the efficacy of durvalumab + olaparib combination therapy compared with durvalumab + placebo	DoR, ORR, and PFS6 according to RECIST 1.1 using Investigator assessment OS18
To assess the efficacy of durvalumab + olaparib combination therapy compared with durvalumab + placebo in the subset of subjects with HRRm	PFS, DoR, ORR, and PFS6 according to RECIST 1.1 using Investigator assessment
To assess the PK of durvalumab and olaparib in both treatment arms	Concentration of durvalumab and olaparib
To investigate the immunogenicity of durvalumab in both treatment arms	Presence of ADAs for durvalumab
To assess disease-related symptoms and HRQoL in UC subjects treated with durvalumab + olaparib combination therapy compared with durvalumab + placebo	EORTC QLQ-C30: all scales analyzed. Main pre-specified endpoints: Global health status/QoL, functioning (physical), and multi-term symptoms (fatigue and pain)
Safety Objective:	**Endpoint/Variable:**
To assess the safety and tolerability profile of durvalumab + olaparib combination therapy compared with durvalumab + placebo	AEs/SAEs, physical examinations, laboratory findings (including clinical chemistry, hematology and urinalysis), WHO/ECOG performance status and vital signs
Exploratory objectives	**Endpoint/Variable:**

Table 1 Study objectives

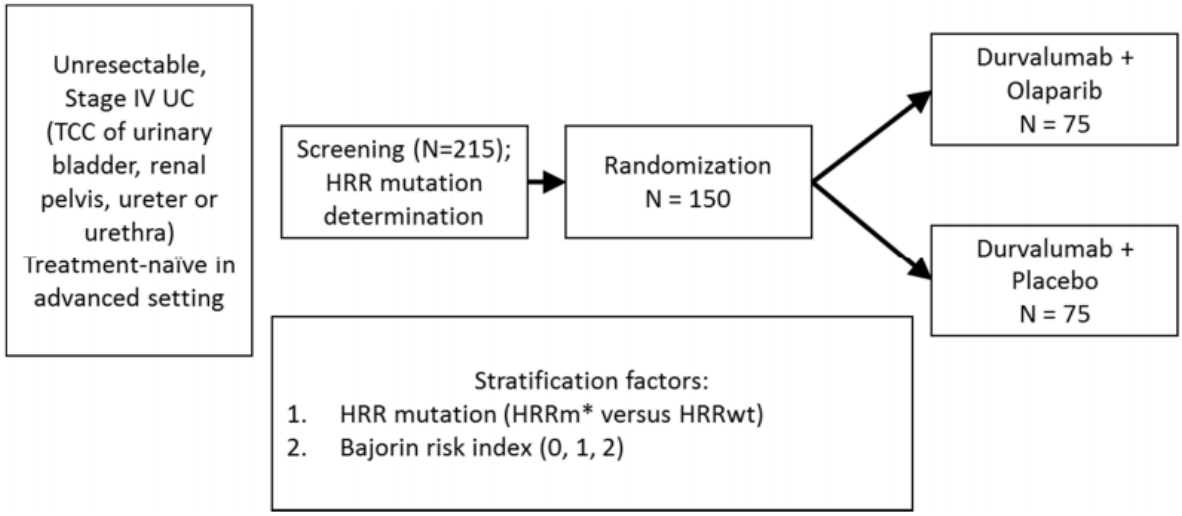
To assess overall change in health status since the start of study treatment in UC subjects treated with durvalumab + olaparib combination therapy compared with durvalumab + placebo	PGIC
To explore the impact of treatment and disease state on health state utility using the EQ-5D-5L during assigned treatment	EQ-5D-5L health state utility index will be used to derive health state utility based on subject-reported data
To collect blood, urine, and tissue samples for defining biological responses to durvalumab + olaparib and for identifying candidate markers that may correlate with likelihood of clinical benefit	Biomarkers (including but not limited to DNA or ctDNA alterations, Protein expression detected by IHC, change in ctDNA levels, mRNA expression) correlating with clinical response

ADA Anti-drug antibody; AE Adverse event; CR Complete response; ctDNA Circulating tumor deoxyribonucleic acid; DNA Deoxyribonucleic acid; DoR Duration of response;; ECOG Eastern Cooperative Oncology Group; EORTC QLQ-C30 European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire; EQ-5D-5L EuroQoL 5 dimension, 5-level health state utility index; HRQoL Health-related quality of life; HRRm Homologous recombination repair mutant; IHC Immunohistochemistry; mRNA Messenger ribonucleic acid; ORR Objective response rate; OS Overall survival; OS18 Subjects alive at 18 months; PFS Progression-free survival; PFS6 Alive and progression free at 6 months; PGIC Patient Global Impression of Change; PK Pharmacokinetic(s); PRO Patient-reported outcome; RECIST 1.1 Response Evaluation Criteria in Solid Tumors, version 1.1; SAE Serious adverse event; UC Urothelial cancer; WHO World Health Organization.

1.2 Study design

This is a Phase II, randomized, double-blind, placebo controlled, multi-center, comparative global study to determine the efficacy and safety of durvalumab + olaparib combination therapy versus durvalumab + placebo (durvalumab monotherapy) as first-line treatment in subjects ineligible for platinum-based therapy with unresectable Stage IV UC. Approximately 150 subjects globally will be randomized in a 1:1 ratio to either the durvalumab + olaparib treatment group or the durvalumab + placebo treatment group, 75 subjects per arm. The randomization will be stratified based on HRR status (mutant versus wild type) and Bajorin risk index (a composite stratum for lymph node only metastasis versus metastasis to any other organ system and Eastern Cooperative Oncology Group [ECOG] performance status [0, 1 versus 2]) (refer to Table 2).

Figure 1 Study design



*The expected prevalence of HRRm is approximately 14%.
HRR homologous recombination repair; HRRm homologous recombination repair mutation;
HRRwt homologous recombination repair wild-type; TCC transitional cell carcinoma; UC urothelial carcinoma

Table 2: Bajorin Risk Index

Bajorin Risk index	ECOG performance status	Disease status
0 (no risk factors)	0 or 1	No metastasis or Lymph node only metastasis
1 (1 risk factor)	0 or 1	Metastatic disease to any other organ system
	2	No metastasis or Lymph node only metastasis
2 (2 risk factors)	2	Metastatic disease to any other organ system

All subjects must provide an FFPE tumor sample for tissue-based HRR gene panel mutation testing. HRR mutant is defined as a subject with loss of function alterations in 15 pre-specified HRR genes: ATM, BRCA1, BRCA2, BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2, PPP2R2A, RAD51B, RAD51C, RAD51D, and RAD54L. If the test results indicate that the subject has at least 1 qualifying mutation in any of these genes, the subject will be considered HRRm for the purposes of the study. Subjects with no detected mutations will be considered HRRwt.

A small number of subjects may be enrolled and randomized based upon a historical HRR result, in this case an additional sample will be taken and submitted for HRR testing in

parallel. If there is a discrepancy between the historical result and subsequent test result, there will be no change to enrolment or randomization, following the ITT principle. All results of historical testing will be listed alongside the subsequent test result as part of the analysis. It is expected that only a small number of subjects will have a historical HRR result and thus most subjects will be enrolled based on the screening result directly.

Based on the HRR status and Bajorin risk index stratification factors, subjects will be stratified into 1 of the following 6 strata:

- Stratum 1: HRRm and Bajorin risk index 0
- Stratum 2: HRRm and Bajorin risk index 1
- Stratum 3: HRRm and Bajorin risk index 2
- Stratum 4: HRRwt and Bajorin risk index 0
- Stratum 5: HRRwt and Bajorin risk index 1
- Stratum 6: HRRwt and Bajorin risk index 2

This study will utilize a fixed dose for durvalumab treatment (1500 mg every 4 weeks [q4w] intravenously [IV]) + olaparib/placebo. Based on an average body weight of 75 kg, a fixed dose of 1500 mg of durvalumab q4w is equivalent to 20 mg/kg q4w.

Dose modifications for durvalumab are permitted in the management of certain IP-related toxicities as described in Section 8.4.5.1 of the Clinical Study Protocol (CSP).

Olaparib/placebo will be dosed orally at an initial dose of 300 mg BID to subjects with creatinine clearance (CrCl) ≥51 mL/min and an initial dose of 200 mg BID to subjects with CrCl ≥31 mL/min but <51 mL/min. At the start of each subsequent cycle the subjects CrCl level is checked and the olaparib dose is modified based on the observed level according to Section 6.7.2 in the CSP.

Durvalumab and olaparib/placebo will be administered beginning on Day 1 until confirmed progressive disease (PD) as per Investigator assessment of RECIST 1.1 and Investigator determination that the subject is no longer benefiting from treatment with the IP, unless there is unacceptable toxicity, withdrawal of consent, evidence of clinical progression, or another discontinuation criterion is met.

1.3 Number of subjects

The study is sized to characterize the PFS benefit of durvalumab in combination with olaparib versus durvalumab monotherapy in first-line subjects with unresectable Stage IV UC.

Approximately 150 subjects globally will be randomized in a 1:1 ratio to either the durvalumab + olaparib treatment group or the durvalumab + placebo treatment group, 75 subjects per arm. The randomization will be stratified based on HRR status (mutant versus wild type) and Bajorin risk index (0 versus 1 versus 2).

The DCO for the primary analysis of PFS will occur when approximately 118 PFS events have occurred across both treatment groups (79% maturity). Assuming a median survival of 3 months for durvalumab monotherapy and recruitment of 150 subjects in 6 months, it is estimated that this DCO will occur 7 months following recruitment of the last subject.

If the true PFS hazard ratio (HR) is 0.55 (likely to correspond to an 82% prolongation of PFS), the study will have 90% power to demonstrate a statistically significant difference for PFS at the 2-sided 5% significance level.

The smallest treatment difference that could be statistically significant at the primary analysis of PFS is an HR of 0.69 (assuming a 2-sided p-value of 0.05).

There will be up to 2 analysis timepoints for OS. The first OS analysis will occur at the same time as the primary PFS analysis and will be based on an estimated 44 OS events across therapies (29% maturity), and the second OS analysis will occur when approximately 100 OS events have occurred (67% maturity). With an approximate 6-month recruitment period and an assumed median OS of 16 months in the durvalumab + placebo arm, it is anticipated that the final analysis will be performed at approximately 27 months after the last subject has been recruited. Details of alpha spending rules to control the overall type 1 error rate for the analysis of OS are provided in Section 4.2.1.

Assuming the true OS HR is 0.75, the study will have 30% power to demonstrate a statistically significant OS effect with a 5% 2-sided significance level, allowing for 2 analyses of the data. The smallest treatment difference at the second analysis of OS that could be statistically significant is an HR of 0.68

2. ANALYSIS SETS

2.1 Definition of analysis sets

2.1.1 Full analysis set (FAS) (Intention to treat)

The full analysis set (FAS) will include all randomized subjects with treatment groups assigned in accordance with the randomization, regardless of the treatment actually received. Subjects who were randomized but did not subsequently receive study treatment will be included in the analysis in the treatment arm to which they were randomized. The FAS therefore follows the principles of ITT.

Subjects will be included in the analysis based on their HRR status and Bajorin risk index reported in the IxRS system that they were subsequently randomized on. Sensitivity analyses may be performed based on the CRF data (for Bajorin risk index) and HRR source data (for HRR status).

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

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

2.1.2 Safety analysis set (SAF)

The safety analysis set (SAF) will consist of all subjects who received at least 1 dose of any study treatment. Subjects will be classified based on the treatment actually received, e.g., subjects randomized to durvalumab + placebo who receive 1 or more doses of olaparib in error, will be reported in the durvalumab + olaparib group. Safety and tolerability summaries will be produced using the safety analysis set.

2.1.3 PK analysis set

All subjects who receive at least 1 dose of study treatment 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. The PK analysis set will be summarized according to the treatment actually received.

Table 3 Summary of outcome variables and analysis populations

Outcome variable	Populations
Efficacy data	
PFS	FAS (all subjects for the primary analysis and the subset of subjects with HRRm for the secondary analysis)
ORR, DoR, PFS6, OS, OS18, PROs	FAS (all subjects, additionally ORR, DoR, and PFS6 will be repeated in the subset of subjects with HRRm) DoR will be based on the subset of subjects in the analysis population who achieved objective tumor response
Demography	FAS
PK data	PK analysis set
Safety data	
Exposure	Safety analysis set
AEs	Safety analysis set
Laboratory measurements	Safety analysis set

Table 3 Summary of outcome variables and analysis populations

Outcome variable	Populations
WHO/ECOG performance status	Safety analysis set
Vital signs	Safety analysis set

AE Adverse event; DoR Duration of response; ECOG Eastern Cooperative Oncology Group; ORR Overall response rate; OS Overall survival; OS18 subjects alive at 18 months; PFS Progression-free survival; PFS6 Alive and progression-free at 6 months; PK Pharmacokinetics; PRO Patient-reported outcome; WHO World Health Organization.

2.2 Violations and deviations

2.2.1 Important protocol deviations

According to ICH E3 (ICH 1995) guidelines,

> "Protocol deviations consist of any change, divergence or departure from the study design or procedures defined in the protocol. Important protocol deviations (IPDs) are a subset of protocol deviations that may significantly affect the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety or well-being."

The following general categories will be considered IPDs and will be listed and discussed in the CSR:

- Subjects who deviate from inclusion criteria 5,6 and 8 and exclusion criteria 1, 9, 10, 11 and 18 per the CSP (Deviation 1).
- Subjects randomized who received treatment other than that to which treatment arm they were randomized to (Deviation 2).
- Received prohibited concomitant medications (any anticancer therapy other than investigational products including mAbs against CTLA-4, PD-1, or PD-L, concurrent chemotherapy, radiotherapy (radiation with palliative intent is allowed), immunotherapy, or biologic or hormonal therapy for cancer treatment and EGFR TKI) or use of known strong or moderate CYP3A inhibitors administered concomitantly with olaparib without appropriate dose reduction for olaparib during the period of concomitant administration. (Deviation 3)
- No baseline RECIST 1.1 assessment (Deviation 4)
- Baseline RECIST scan > 42 days before date of randomization (Deviation 5). (Based upon a 28 day screening period plus 2 weeks allowance, so that only serious violators are identified)

- Overdose of any study medication i.e. Use of durvalumab or olaparib in doses greater than 110% of the monthly prescribed dose. (Deviation 6)

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

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

- Did not have the intended disease of indication, or

- Did not receive any randomized therapy

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

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

2.2.2 Monitoring of important protocol deviations

The IPDs will be programmatically identified within the clinical database by programmed edit checks or via manual validation checks (see Appendix A). A programmatically derived IPD report will be created listing all potential IPDs and the data used to identify them. This report will be reviewed at regular IPD review meetings held on a bi-monthly basis. At this meeting, programmatically-derived IPDs will be checked to ensure that they have been correctly classified and reviewed to determine whether the potential IPD is important, non important or due to missing data in the database at the time of the review. On an ongoing basis throughout the study, monitoring notes or summaries will be reviewed to determine any important post-entry deviations that are not identifiable via programming.

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

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST visit responses

For all subjects, the RECIST tumor response data will be used to determine each subject's visit response according to RECIST version 1.1. It will also be used to determine if and when

a subject has progressed in accordance with RECIST and also their best objective response to study treatment.

Baseline radiological tumor assessments are to be performed within 28 days before randomization and ideally as close as possible to the start of study treatment. Tumor assessments are then performed every 8 weeks ±1 week following randomization for the first 48 weeks and every 12 weeks ±1 week thereafter until progressive disease (PD) as per Investigator assessment of RECIST 1.1 and Investigator determination that the subject is no longer benefiting from treatment with the IP.

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

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

Definitions of CR, PR, SD and PD are provided in Section 3.1.1.

RECIST derived efficacy outcomes (i.e. PFS, ORR, DoR and PFS6) will be calculated programmatically from the site investigator data (see Section 3.2).

3.1.1 Target lesions (TLs)

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (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 subject can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (representative of all lesions involved suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to randomization will be used to define the baseline sum of TLs. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

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

Measurable disease (i.e.at least one TL) is one of the entry criteria for the study. However, if a subject who does not have measurable disease and is enrolled into the study (i.e. no TLs), evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions (see 3.1.3 for further details). If a subject does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 4 TL visit responses

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

Rounding of TL data

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

Missing TL data

If all TL measurements are missing, then the TL visit response is NE.

If the sum of available 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, the TL visit response is PD. Note: the nadir can only be taken from assessments where all the TLs had a LD recorded.

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

Lymph nodes

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

TL visit responses subsequent to CR

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

- Step 1: If all lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node LD increases by 20% but remains < 10mm.

- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response will be set to NE irrespective of whether, when referencing the sum of TL diameters, the criteria for PD are also met.

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

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

TL too big to measure

If a TL becomes too big to measure this should be indicated in the database and a size ('x') above which it cannot be accurately measured should be recorded. If using a value of x in the calculation of TL response would not give an overall visit response of PD, then this will be

flagged and reviewed by the study team blinded to treatment assignment. It is expected that a visit response of PD will remain in the vast majority of cases.

TL too small to measure

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

Irradiated lesions/lesion intervention

Previously irradiated lesions (i.e. 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 and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumors:

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

- Step 2: If there was no evidence of progression after step 1, treat the lesion diameter (for those lesions with intervention) as missing and if ≤ 1/3 of the TLs have missing measurements then scale up as described in the 'Scaling' section below. If the scaling results in a visit response of PD then the subject would be assigned a TL response of PD.

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

At subsequent visits the above steps will be repeated to determine the TL and overall visit response. When calculating the previous minimum, lesions with intervention 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 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	Longest diameter at nadir visit	Longest diameter at follow-up visit
1	7.2	7.1
2	6.7	6.4
3	4.3	4.0
4	8.6	8.5
5	2.5	missing
Sum	**29.3**	**26**

Lesion 5 is missing at the follow-up visit.

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

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

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

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

Lesions that split in two

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

Lesions that merge

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

Change in method of assessment of TLs

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

If a change in method involves clinical examination (e.g. CT changes to clinical examination or vice versa), any affected lesions should be treated as missing. The TL visit response may still be evaluable if the number of missing TL measurements at a visit is ≤1/3 of the total number of TLs.

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

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

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

Table 5 NTL Visit Responses

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

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

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

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

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

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

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

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

3.1.3 Overall visit response

Table 6 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 6 Overall visit responses

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
CR	CR or NA	None recorded	**CR**
CR	Non- CR/Non-PD or NE	None recorded	**PR**
PR	Non-PD or NE or NA	None recorded	**PR**
SD	Non-PD or NE or NA	None recorded	**SD**
PD	Any	Any	**PD**
Any	PD	Any	**PD**

TARGET	NON-TARGET	NEW LESIONS	OVERALL VISIT RESPONSE
Any	Any	Yes	**PD**
NE	Non-PD or NE or NA	None recorded	**NE**
NA	CR	None recorded	**CR**
NA	Non-CR/Non-PD	None recorded	**SD**
NA	NE	None recorded	**NE**

3.1.4 Independent Review

BICR may be performed at AstraZeneca's discretion.

3.2 Outcome variables

3.2.1 Survival calls

Survival calls are made for both the primary PFS analysis and subsequent OS analysis.

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

3.2.2 Progression Free Survival

PFS (per RECIST 1.1 via investigator assessments) is defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the subject withdraws from randomized therapy or receives another anticancer therapy prior to progression:

PFS = date of event or censoring – date of randomization + 1

Subjects who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment. However, if the subject progresses or dies after 2 or more consecutive missed visits, the subject 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 a missed visit).

Given the scheduled visit assessment scheme (i.e. eight-weekly for the first 48 weeks then twelve-weekly thereafter) the definition of 2 missed visits will change. Study day will be calculated in relation to date of randomization.

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

- If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from eight-weekly to twelve-weekly this will equate to 22 weeks (i.e., take the average of 8 and 12 weeks which gives 10 weeks and then apply same rationale hence 2 x 10 weeks + 1 week for an early assessment + 1 week for a late assessment = 22 weeks). The time period for the previous RECIST assessment will be from study days 274 to 329 (i.e. week 39 to week 47). That is, the last visit is at week 40, and therefore two missing visits will mean that a subject has a missing visit on both the 8 week and 12 week schedule.

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

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

The PFS time will always be derived based on scan/assessment dates 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 RECIST 1.1 assessment/scan dates of the component that indicates progression.

- When censoring a subject for PFS, the subject will be censored at the latest of the dates contributing to a particular overall visit assessment.

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

3.2.3 Overall survival

Overall survival is defined as the time from the date of randomization until death due to any cause regardless of whether the subject withdraws from randomized therapy or receives another anti-cancer therapy;

OS = date of death or censoring – date of randomization + 1

Any subject not known to have died at the time of analysis will be censored based on the last recorded date on which the subject was known to be alive and DCO date i.e. min(SUR_DAT, recorded within the SURVIVE module of the eCRF, DCO date) (see Section 3.2.1 for details of date of death).

If a patient is known to have died where only a partial death date is available, then the date of death will be imputed according to Section 3.3.1.

3.2.4 Objective response rate

ORR (per RECIST 1.1 using Investigator assessment) is defined as;

ORR = the number (%) of subjects with at least one overall visit response of CR or PR

and will be based on a subset of all randomized subjects with measurable disease at baseline per the site investigator. All subjects in the FAS should have measurable disease at baseline in accordance with the eligibility criteria, however this will be confirmed programmatically.

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

3.2.5 Duration of response

Duration of response (per RECIST 1.1 using Investigator assessment) will be defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression;

DoR = date of PFS event or censoring – date of first response + 1

The end of response should coincide with the date of progression or death from any cause used for the RECIST 1.1 PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR.

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

3.2.6 Proportion of subjects alive and progression free at 6 months after randomization (PFS6)

The proportion of subjects alive and progression free at 6 months (i.e., PFS6) will be defined as the Kaplan-Meier estimate of PFS (per RECIST 1.1 as assessed by the site investigator) at 6 months after randomization.

3.2.7 Proportion of subjects alive at 18 months after randomization (OS18)

The proportion of subjects alive at 18 months (OS18) will be defined as the Kaplan-Meier estimate of OS at 18 months after randomization.

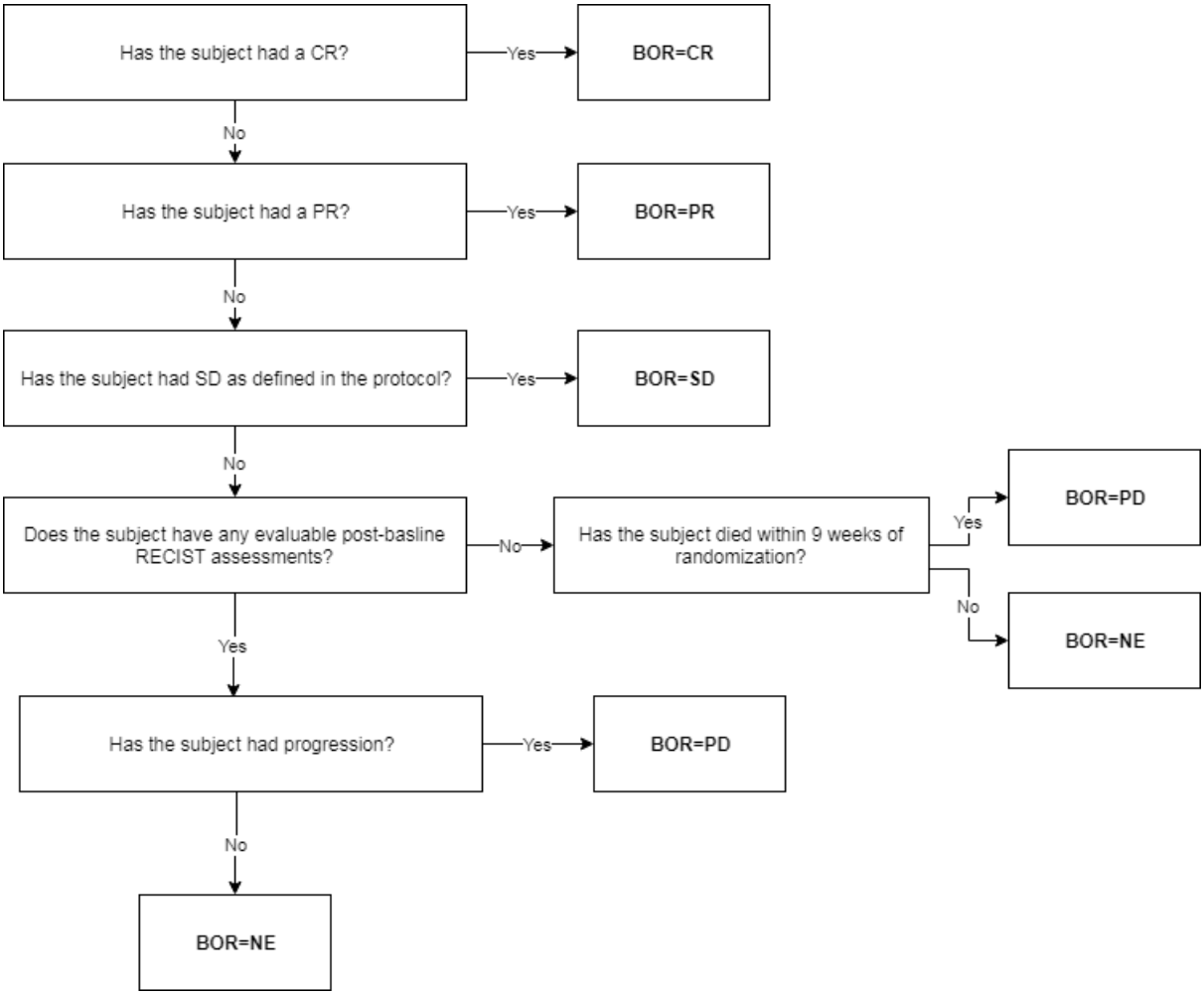
3.2.8 Best objective response (BoR)

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in Section 3.1. It is the best response a subject has had following randomization but prior to starting any subsequent anti-cancer therapy and up to and including RECIST progression or the last evaluable assessment in the absence of RECIST progression.

Categorization of BoR will be based on RECIST using the following response categories: CR, PR, SD, PD and NE (see Figure 2). For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. In order to have SD as BoR, SD should be recorded at least 7 weeks (8 weeks minus 1 week to allow for an early assessment within the assessment window), after randomization. For CR/PR, the initial overall visit assessment which showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

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

Figure 2 Flowchart for determining BoR



BoR will be determined programmatically based on RECIST from the overall visit response using all site investigator data up until the first progression event or start of subsequent anti-cancer therapy. The denominator will be consistent with those used in the ORR analysis.

3.2.9 Change in TL tumor size

The percentage change from baseline in the sum of tumor size at each assessment will be calculated. The best change in tumor size (i.e. depth of response) is the largest decrease from baseline or the smallest increase from baseline in the absence of a reduction and will include all assessments up to and including any evidence of progression (or prior to death in the absence of progression) prior to the start of subsequent anti-cancer therapy. Otherwise the last

evaluable RECIST assessment if the subject has not died, progressed or started subsequent anti-cancer therapy.

If best percentage change cannot be calculated due to missing data (including if the subject has no TLs at baseline), a value of +20% will be imputed as the best percentage change from baseline in the following situations (otherwise best percentage change will be left as missing):

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

3.2.10 Patient-reported outcome variables

The following PROs will be administered in this study: The European Organization for Research and Treatment of Cancer 30-item core quality of life questionnaire (EORTC QLQ-C30), 5-level health state utility index (EQ-5D-5L), Patient Global Impression of Change (PGIC), and the Vulnerability Elders Survey-13 (VES-13) (see CSP for questionnaire specimens). The VES-13 will only be given at baseline to assess the subject's frailty and not to assess the efficacy of treatment (see Section 3.3.9). All questionnaires will be scored according to published scoring guidelines. All PRO analyses will be based on the FAS, unless otherwise stated.

3.2.10.1 EORTC QLQ-C30

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

- 5 functional scales: physical, role, cognitive, emotional, and social
- 3 multi-item symptom scales: fatigue, pain, and nausea/vomiting
- 6 single item symptom scales: dyspnea, insomnia, appetite loss, constipation, diarrhea, financial difficulties
- Global health status/QoL scale

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

For each subscale, if <50% of the subscale items are missing, then the equation provided in the scoring manual can be applied excluding the missing items. If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded.

The global health status/HRQoL will be assessed using the EORTC-QLQ-C30 global QoL scale which includes 2 items from the QLQ-C30: "How would you rate your overall health during the past week? (Item 29) and "How would you rate your overall QoL during the past week?" (Item 30).

Additionally at each post-baseline assessment, the change from baseline in symptom scales, functional scales and HRQoL scores will be calculated.

3.2.10.2 EQ-5D-5L

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (1: no problems, 2: slight problems, 3: moderate problems, 4: severe problems, and 5: extreme problems). A unique EQ-5D health state, termed EQ-5D profile, is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index, termed the EQ-5D index, by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses).

In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analog scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

3.2.10.3 PGIC

The PGIC item assesses how a subject perceives their overall change in health status since the start of study treatment and helps to determine overall impact of treatment. The PGIC question is "Overall, how would you rate the change in your bladder cancer symptoms since you started this study?" with response options for the PGIC as follows: Very Much Improved (+3), Much Improved (+2), Minimally Improved (+1), No Change (0), Minimally Worse (-1), Much Worse (-2), and Very Much Worse (-3). No scoring will be done using the assigned numerical values.

3.2.10.4 PRO Compliance rates

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30, PGIC and EQ-5D-5L respectively. These will be based upon:

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

- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time e.g. a questionnaire from a subject who has not withdrawn from the study at the scheduled assessment time but excluding subjects in countries with no available translation. For subjects that have progressed, the latest of progression and safety follow-up will be used to assess whether the subject is still under HRQoL follow-up at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.

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

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

- Overall subject compliance rate is defined for each randomized treatment group as:

  = [(Total number of subjects with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above)) / the total number of subjects expected to have completed at least a baseline questionnaire)] x 100.

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

3.3 Safety variables

3.3.1 General considerations for safety and PRO assessments

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

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

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

- The window for the visits following baseline will be constructed in such a way that

the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two consecutive visits then the upper limit will be taken as the midpoint value minus 1 day.

For example, the visit windows for vital signs data (with 4 weeks between scheduled assessments) are:

- Cycle 1, Day 1, visit window N/A
- Cycle 2, Day 1 (i.e. Day 29), visit window 2 – 43
- Cycle 3, Day 1 (i.e. Day 57), visit window 44 - 71
- Cycle 4, Day 1 (i.e. Day 85), visit window 72 - 99
- Cycle 5, Day 1 (i.e. Day 113), visit window 100 - 127
- Cycle 6, Day 1 (i.e. Day 141), visit window 128 - 155
- Cycle 7, Day 1 (i.e. Day 169), visit window 156 - 183
- Cycle 8, Day 1 (i.e. Day 197), visit window 184 - 211
- Cycle 9, Day 1 (i.e. Day 225), visit window 212 to 239
- Cycle 10, Day 1 (i.e. Day 253), visit window 240 - 267
- Cycle 11, Day 1 (i.e. Day 281), visit window 268 - 295
- Cycle 12, Day 1 (i.e. Day 309), visit window 296 - 323
- Cycle 13, Day 1 (i.e. Day 337), visit window 324 - 351

Note, due to the differing assessment schedules, the visit windows may be different for the different endpoints.

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).

- Listings should display all values contributing to a time point for a subject.

- For visit based summaries
    - If there is more than one value per subject within a time window then the closest value to the scheduled visit date should be summarized, or the earlier, in the event the values are equidistant from the nominal visit date. The listings should highlight the value for the subject 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 should only be summarized if the number of observations is greater than the minimum of 20 and > 1/3 of subjects dosed.

- For summaries at a subject level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a subject level statistic such as a maximum.

- Baseline will generally be the last value obtained prior to the first dose of study medication. Alternatively, if two visits are equally eligible to assess subject 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. In the scenario where there are two assessments on day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. Where safety data are summarized over time, study day will be calculated in relation to date of first study treatment.

The following considerations are made for missing safety data, diagnostic dates and AE dates;

- Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e., below the lower limit of quantification) or > x (i.e., above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings.

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

- Adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

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

    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.

Note: When imputing a start date ensure that the new imputed date is sensible i.e. is prior to the end date of the AE or treatment.

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

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

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

- If an AE has a completely missing end date then it will be treated as ongoing.
- 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

- For partial subsequent anti-cancer therapy dates, the following will be applied:

  a. Missing day – if the month is the same as treatment end date then impute to the day after treatment, otherwise first day of the month.

  b. Missing day and month – if year is the same as treatment end date then impute to the day after treatment, otherwise 1st January of the same year as anti-cancer therapy date.

3.3.2 Exposure and dose interruptions

Exposure will be defined as follows:

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

Actual exposure of study treatment of durvalumab

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

Actual exposure of study treatment of olaparib

- Actual exposure = intended exposure – (total duration of dose delays+ total duration of dose interruptions) where intended exposure will be calculated as above.

The duration of dose delays and dose interruption are defined as follows:

- Total duration of dose delays of durvalumab = Sum of (Date of the dose – Date of previous dose – 28 days)
    - If there are no delays, the duration sums to 0, as infusions are performed every 4 weeks
- Total duration of dose delays + interruptions of olaparib/placebo = Sum of (Date of the dose – Date of previous dose – 1 day)
    - If there are no delays/interruptions, the duration sums to 0, as olaparib/placebo is administered daily

Dose modification of durvalumab are permitted in the management of certain IP-related toxicities as per Section 6.7 of the CSP.

Dose modification are also permitted for olaparib and placebo per Section 6.7.2 of the CSP based on the subjects CrCl level. Subjects will be administered olaparib/placebo orally at 200 or 300 mg BID continually based on the subject's CrCl level.

Subjects initiating olaparib at 200 mg BID due to reduced CrCl may modify the dose at the beginning of the next treatment cycle to 300 mg BID if the CrCl increases to ≥51 mL/min or to 150 mg or 100 mg BID if the CrCl ≥31 mL/min but <51 mL/min.

Subjects initiating olaparib at 300 mg may modify the dose at the beginning of the next treatment cycle if CrCl increases to ≥51 mL/min to 250 mg BID or 200 mg BID.

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

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

If a subject 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.

3.3.3 Dose Intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation.

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

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

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

3.3.4 Adverse events

AEs and SAEs for both treatment arms will be collected throughout the study, from date of informed consent until 90 days after the last dose of study treatment.

A treatment emergent adverse event (TEAE) is an AE with an onset date or a pre-existing AE worsening (by investigator report of a change in intensity) following the first dose of study treatment up to and including min(date of last dose of study treatment + 90 days, day before the first dose of subsequent anti-cancer therapy).Any AE occurring before study treatment (i.e. before the administration of the first dose on Study Day 1) and without worsening after initial of study treatment will be referred to as 'pre-treatment'.

The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03).

For the durvalumab + olaparib arm and the durvalumab + placebo arm, in the event of the components being administered separately then date of first dose/last dose will be considered as the earliest/latest dosing date of either component.

Other significant adverse events (OAEs)

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

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

AEs of special interest

Some clinical concepts (including some selected individual preferred terms and higher level terms) have been considered "AEs of special interest" (AESI) to the durvalumab and olaparib programs. These AESI's have been identified by the patient safety team. Other categories may be added as necessary. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which higher level terms and which preferred terms contribute to each AESI. Further reviews will take place prior to Database lock (DBL) to ensure any further terms not already included are captured within the categories.

AEs of special interest for durvalumab

AESIs for durvalumab include, but are not limited to, events with a potential inflammatory or immune-mediated mechanism and which may require more frequent monitoring and/or interventions such as steroids, immunosuppressants, and/or hormone replacement therapy.

AESIs observed with durvalumab include the following:

- Diarrhea/Colitis and intestinal perforation
- Pneumonitis/ILD
- Hepatitis/transaminase increases
- Endocrinopathies (i.e., events of hypophysitis/hypopituitarism, adrenal insufficiency, hypothyroidism and hyperthyroidism, and type I diabetes mellitus)
- Rash/Dermatitis
- Nephritis/blood creatinine increases
- Myocarditis
- Myositis or polymyositis
- Pancreatitis/serum lipase and amylase increases
- Neuropathy/neuromuscular toxicity (e.g., Guillain-Barré and myasthenia gravis)
- Other inflammatory responses that are rare/less frequent with a potential immune-mediated etiology include, but are not limited to, pericarditis; sarcoidosis; uveitis and other events involving the eye; and skin, hematological, and rheumatological events

- infusion-related reactions and hypersensitivity/anaphylactic reactions with a different underlying pharmacological etiology

Adverse events of special interest for olaparib

AESIs for olaparib are the important potential risks of:

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

3.3.5 Laboratory data

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

Change from baseline in hematology and clinical chemistry variables will be calculated for each post-dose visit on treatment. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 3.3 will be used.

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

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

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

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

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

Project reference ranges will be used throughout for reporting purposes. If the project range is unavailable for a particular test, local ranges will be used. The denominator used in laboratory summaries of CTCAE grades will only include evaluable subjects, in other words those who had sufficient data to have the possibility of an abnormality.

For example:

- If a CTCAE criterion involves a change from baseline, evaluable subjects 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 subject need only have 1 post dose-value recorded.

3.3.6 ECGs

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

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

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

3.3.7 Vital signs

The following vital signs will be collected in the eCRF every 4 weeks throughout the study, from screening to the 1st follow up visit, 30 days after the last dose of study treatment: systolic and diastolic blood pressure [BP], pulse, respiratory rate, temperature, height (screening only) and weight.

Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 3.3 will be used.

3.3.8 WHO/ECOG performance status

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

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

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

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

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

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

Any significant change from baseline or screening are reported as an AE.

3.3.9 Concomitant medication

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

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

3.4 Pharmacokinetic variables

PK concentration data will be collected within 10 minutes of the end of infusion for durvalumab during cycle 1 and pre-dose during cycle 2, cycle 4 and 90 days post treatment discontinuation of durvalumab. For olaparib, PK concentration data will be collected pre-dose for cycle 1, 2, 4 and 30 days post treatment discontinuation of olaparib. The actual sampling times will be used in the PK calculations.

3.5 Immunogenicity variables

Serum samples for antidrug antibodies ADA assessments will be conducted utilizing a tiered approach (screen, confirm, titer), and ADA data will be collected at scheduled visits shown in the CSP (Section 8.5.2). ADA result from each sample will be reported as either positive or negative (note that not detected is equivalent to negative). If the sample is positive, the ADA titer will be reported as well. In addition, the presence of neutralizing antibody (nAb) may be tested for all ADA-positive samples using a ligand-binding assay. The nAb results will be reported as positive or negative.

The number of ADA-evaluable subjects who fulfil the following the criteria will be determined. The percentage of ADA-positive subjects in each of the category will be calculated, using the number of ADA evaluable subjects in each treatment group as the

denominator. A subject is defined as being ADA positive if a positive ADA result is available at any time, including baseline and all post-baseline measurements; otherwise ADA negative.

- ADA positive at any visit; the percentage of ADA-positive subjects in the ADA evaluable subjects is known as ADA prevalence
- The sum of both treatment-induced and treatment-boosted ADA; the percentage of subjects fulfilling this criterion in the ADA analysis set is known as ADA incidence
- ADA positive post-baseline and positive at baseline
- ADA positive post-baseline and negative at baseline (treatment-induced ADA)
- ADA negative 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 with at least 16 weeks (112 days) between the first and last positive measurement. or an ADA positive result at the last available assessment. The category includes subjects meeting these criteria who are ADA positive at baseline
- Transiently positive ADA, defined as having at least one post-baseline ADA positive measurement and not fulfilling the conditions for persistently positive. The category includes subjects meeting these criteria who are ADA positive at baseline
- nAb positive at any visit

Subjects in the safety analysis set with non-missing baseline ADA sample and at least 1 post-baseline ADA sample will be considered ADA evaluable.

3.6 Biomarker variables

3.6.1 Tumor based biomarker variables

HRR Status

All subjects must provide an FFPE tumor sample at baseline for tissue-based HRR gene panel mutation testing.

A subject may be enrolled and randomized based upon a historical HRR result, in this case an additional sample will be taken and submitted for HRR testing in parallel. If there is a

discrepancy between the historical result and subsequent test result, there will be no change to enrollment or randomization, following the ITT principle.

The HRR status of the subject must be identified prior to the subject being randomized into the study thus missing data should not be feasible.

PD-L1 Status

The tumor specimen will also be used to establish the subjects PD-L1 status. PD-L1 status is determined by the percentage of tumor cells with any membrane staining above background or by the percentage of tumor associated immune cells with staining (IC+) at any intensity above background. The percent of tumor area occupied by any tumor-associated immune cells (Immune Cells Present, ICP) is used to determine IC+, which is the percent area of ICP exhibiting PD-L1 positive immune cell staining.

PD-L1 is believed to be a predictive factor for the benefit of durvalumab, with subjects with low PD-L1 levels having less benefit and is thus used as a covariate in efficacy analyses. A subject's PD-L1 status will be defined as the following:

- High Expression, if any of the following criteria are met:
  - $\geq$ 25% of tumor cells exhibit membrane staining; or,
  - ICP > 1% and IC+ $\geq$ 25%; or,
  - ICP = 1% and IC+ = 100%

- Low Expression: if the High criteria are not met.

based on previously validated cut-offs for durvalumab in bladder cancer (VENTANA PD-L1 (SP263) Assay [package insert] 2017).

As a subjects' PD-L1 status will be determined post randomization, it is possible that some subjects PD-L1 status will be unevaluable, and therefore missing. The expected % of subjects that will have missing PD-L1 status is expected to be low.

The method of handling missing PD-L1 data will be based on the % of subjects with missing PD-L1 result within the FAS. If $\leq$ 10% of subjects have missing value, the impact is expected to be low, and so for subjects with missing values, the covariate of PD-L1 status will be set to "unknown" in analysis models. If greater than 10% of data are missing, the impact of the missing data will in addition be assessed using a multiple imputation method, which is outlined below;

1. A logistic regression model will be fitted with PD-L1 missing status as the response variable. Covariates will be included in order to assess their impact on missingness of PD-L1 status, covariates will include relevant subject demography, specimen characteristics and any other clinical covariates of interest. Terms will be classified as significant if their p-value is < 0.20.

2. An imputation model (using PROC MI) will be fitted for PD-L1 status, including any covariates identified from 1), and any other clinically relevant covariates that may impact PD-L1 missingness. Multiple imputation will be performed using a logistic regression model, assuming a monotone missing pattern. 50 imputations will be performed. The pattern of missingness will be assessed to check the validity of the monotonic assumption. If the missing pattern is not monotone (i.e. arbitrary missingness), FCS will be used instead.

3. Imputed datasets will be analyzed according to the primary analysis model, with the log hazard ratio, standard error and corresponding two-sided 95% confidence limits being calculated for each dataset

4. PROC MIANALYZE will be used to combine the analysis results from the 50 imputations runs (using Rubin's rules), in order to generate an overall estimate of the log hazard ratio and corresponding confidence interval. The log hazard ratio, hazard ratio, and corresponding two-sided 95% confidence limits will be reported.

An FFPE tumor sample may also be taken at progression, subject to the subject's consent, and analyzed for HRR and PD-L1 status as part of the exploratory biomarker analyses to be reported outside of the CSR.

3.7 Genetic variables

In the case of genetic data, only the date that the subject gave consent to participation in the genetic research and the date the blood sample was taken from the subject will be recorded in the eCRF and database. The genetic data generated from the study will be stored in the AstraZeneca Laboratory Information Management System (LIMS) database or other appropriate system. Data will be reported outside the CSR and thus not included in this SAP.

4. ANALYSIS METHODS

The primary objective is to assess the efficacy of durvalumab + olaparib combination therapy compared with durvalumab + placebo in terms of PFS for subjects in the full analysis set (FAS). The secondary objective is to assess OS in the FAS. Thus, formal statistical analyses will be performed to test the main hypotheses:

- Null Hypothesis 1 (H_{10}): No difference between durvalumab + olaparib and durvalumab + placebo in terms of PFS

- Alternative Hypothesis 1 (H_{11}): Difference between durvalumab + olaparib and durvalumab + placebo in terms of PFS

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

- Null Hypothesis (H_{20}): No difference between durvalumab + olaparib compared with durvalumab + placebo in terms of OS
- Alternative Hypothesis (H_{21}): There is a difference between durvalumab + olaparib compared with durvalumab + placebo in terms of OS

To control for type 1 error, a significance level of 5% will be used for the analysis of PFS. If the null hypothesis (H_{10}) of no effect is rejected, 100% of the alpha can be passed to the OS secondary endpoint to test H_{20}.

The study will be considered positive (i.e., a success) if durvalumab + olaparib combination therapy is statistically different from durvalumab + placebo in terms of PFS in the FAS.

A secondary analysis of PFS in the HRRm subgroup of the FAS will be performed at the time of the primary PFS analysis and at the time of the OS analysis. However the study has not been sized sufficiently to characterize the PFS benefit of durvalumab + olaparib combination therapy vs durvalumab monotherapy in the HRRm subgroup.

Note that the primary analysis of PFS in the FAS will be only be performed at 1 timepoint, when there are approximately 118 events and the study has been unblinded. At this time all secondary analyses, including an interim analysis of OS, will also be performed excluding OS18 due to insufficient data expected at this time. RECIST data for the HRRm subgroup only along with survival and safety data for the FAS will continue to be collected to support a final analysis of OS, OS18, safety for the FAS/SAF and an updated analysis of PFS for the HRRm subjects only.

4.1 General principles

Recruitment into the study was paused to new subjects on the 3rd July 2018 after 9 subjects had been randomized due to the need for a CSP amendment (updated to CSP v2.0). The CSP amendment changed the inclusion criteria from cisplatin-ineligible to platinum-ineligible, and in addition removed the requirement to randomize an equal number of HRRm and HRRwt subjects. The intention is that the final analyses will be conducted in all subjects (pooled prior to amendment and post amendment), however the appropriateness of this approach will be investigated further prior to unblinding. There is no anticipated bias from having excluded some HRRwt subjects due to the enrichment process applied during CSP v1.0.

The primary source for the stratification factors, HRR status and Bajorin Risk index, will be those captured in IVRS during randomization. However, source data will be used to check concordance against the data entered into IVRS. For HRR status, source data provided by the relevant vendor will be used and for Bajorin risk index, it will be programmatically calculated

based on whether a subject has visceral metastasis and their ECOG PS (per the definition in Section 3.3.8) from the following CRF data;

- Metastatic disease to any organ system other than lymph nodes - yes
    - LOCADMET: metastatic/locally advanced = “metastatic” or “both” and DISSITES: site of local/metastatic disease = anything other than:
        - “lymph node”, “bladder”, “genitourinary”, “other locally advanced sites”
        - “other metastatic sites” and the specified site (S_MSOTHJ) is deemed a local site, determined by medic review.
- Metastatic disease to any organ system other than lymph nodes - no
    - LOCADMET: metastatic/locally advanced =– “locally advanced” only or
    - LOCADMET: metastatic/locally advanced = “metastatic” or “both” and DISSITES: site of local/metastatic disease =
        - “lymph node”, “bladder”, “genitourinary”, “other locally advanced sites”
        - “other metastatic sites” and the specified site (S_MSOTHJ) is deemed a local site, determined by medic review.

ECOG PS is defined directly in the CRF, with the relevant responses as follows;

- 0 - Normal Activity
- 1 - Restricted Activity
- 2 - In Bed Less Than or Equal to 50% of the Time

Sensitivity analyses using the source data may be performed, if a high level of discordance is observed.

Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum, and maximum. For log transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarized by frequency counts and percentages for each category.

Unless otherwise stated, percentages will be calculated out of the population total and for each treatment group. For continuous data the mean and median will be rounded to 1 additional

decimal place compared to the original data. The standard deviation and confidence intervals 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 laboratory data this is specified in the AstraZeneca durvalumab project reference ranges documentation. For summaries and analyses based on calculated percentages or ratios (including hazard ratios) the mean and median will be rounded to 2 decimal places. The standard deviation and confidence intervals will be rounded to 3 decimal places. Minimum and maximum will be rounded to 1 decimal place.

For categorical data, percentages will be rounded to 1 decimal place.

For PK data the mean, median, standard deviation, geometric mean and CV will be presented to 4 significant figures (sf), minimum and maximum will be presented to 3 sf and n will be presented as an integer.

The primary and secondary efficacy analyses (including PROs) will be performed on all subjects in the FAS, and additional secondary analyses will be performed on the HRRm subgroup of the FAS for selected efficacy endpoints. PK data will be summarized and analyzed based on the PK analysis set. Safety and treatment exposure data will be summarized based upon the safety analysis set. Study population and demography data will be summarized for the FAS.

Outputs will be summarized by treatment arm.

For all efficacy analyses involving summaries at a given month (e.g. OS18), the number of days is calculated as:

Number of days = 365.25/12*number of months, rounded up to the integer number

SAS® version 9.2 or above will be used for all analyses.

In general, for efficacy endpoints the last observed measurement prior to randomization will be considered as the baseline measurement. However, if an evaluable assessment is only available after randomization but before the first dose of randomized treatment then the assessment closest to randomization will be used as baseline.

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

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

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

For all analyses radiotherapy is not considered a subsequent anti-cancer therapy.

4.2 Analysis methods

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

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

Table 7 details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint. Note: all endpoints compare durvalumab + olaparib versus durvalumab + placebo in all randomized subjects (FAS), unless otherwise indicated.

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

Endpoints analyzed	Notes
Progression free survival (PFS)	Primary analysis for all-comers in the FAS and a secondary analysis in the HRRm subgroup of the FAS. Analysis is conducted using a stratified log-rank test using Investigator assessment per RECIST 1.1 with the following covariates: 1) HRR status (mutant versus wildtype) 2) PD-L1 tumor status (high vs low) [following the missing data methods in Section 3.6] 3) Bajorin risk index (0 versus 1 versus 2) Sensitivity analyses: 1) Interval censored analysis – evaluation time bias 2) Analysis using alternative censoring rules – attrition bias
Overall survival (OS)	Secondary analysis for all-comers in the FAS. Analyzed using a stratified log-rank test with the following covariates: 1) HRR status (mutant versus wildtype) 2) PD-L1 tumor status (high vs low) [following the missing data methods in Section 3.6] 3) Bajorin risk index (0 versus 1 versus 2)
Objective response rate (ORR)	Secondary analysis for all-comers, and for the HRRm subgroup of the FAS. Logistic regression using Investigator assessment per RECIST 1.1

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

Endpoints analyzed	Notes
Duration of response (DoR)	Secondary analysis for all-comers, and for the HRRm subgroup of the FAS. KM estimates using Investigator assessments per RECIST 1.1
Proportion of subjects alive and progression-free at 6 months (PFS6)	Secondary analysis for all-comers, and for the HRRm subgroup of the FAS. KM estimates of subjects alive and progression free at 6 months
Proportion of subjects alive at 18 months (OS18)	Secondary analysis for all-comers, and for the HRRm subgroup of the FAS. KM estimates of survival at 18 months
Best objective response (BoR)	N (%) using Investigator assessment
Change from baseline in symptoms (EORTC QLQ-C30 endpoints)	Secondary analysis for all-comers in the FAS. Average change from baseline using a Mixed Model Repeated Measurements (MMRM) analysis
EQ-5D-5L (health state utility values and Visual Analog Scale)	Summary statistics for health state utilities and visual analogue scale, including change from baseline
PGIC	Descriptive summaries of response categories at each visit

DoR Duration of response; EORTC QLQ-C30 European Organisation for Research and Treatment of Cancer Quality of life questionnaire; EQ-5D-5L EuroQoL 5 dimension, 5-level health state utility index; FAS full analysis set; HR hazard ratio; HRRm homologous recombination repair mutant; KM Kaplan-Meier; MMRM mixed effect model repeat measurement; ORR Objective response rate; OS Overall survival; PD-L1 Programmed death ligand 1; PFS Progression-free survival; PFS6 alive and progression free at 6 months; RECIST Response Evaluation Criteria In Solid Tumors.

4.2.1 Multiple testing strategy

Formal adjustment of alpha for multiple comparisons is not planned, but if required at the time of reporting, 100% of the alpha used for the primary PFS analysis will be used for the OS analysis if the PFS analysis is statistically significant. That is, formal statistical testing of the OS endpoint is planned only if the primary analysis of PFS is significant. If PFS is not significant, any analysis of OS will be considered exploratory.

For the OS endpoint, there are two analyses planned (1st OS analysis concurrent with the primary PFS analysis), and the alpha level will be controlled at the 1st OS analysis and final OS timepoints by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach. The O'Brien Fleming boundaries for the OS interim and final analyses will be adjusted depending on the timing of each OS analysis. Adjusted confidence limits will be reported along with nominal 95% confidence limits. Details on the timing of the OS analyses and alpha-spending are contained in Section 5.

4.2.2 Primary endpoint: Progression free survival

The primary analysis of the primary endpoint PFS will occur when it is expected that approximately 118 PFS events have occurred (79% maturity). PFS will be based on the programmatically derived RECIST 1.1 using investigator data.

The analysis will be performed for subjects in the FAS using a stratified log-rank test adjusting for HRR status (mutant versus wildtype), PD-L1 tumor status (high versus low) and Bajorin risk index (0 versus 1 versus 2) for generation of the p-value, and using the Breslow approach for handling ties (Breslow 1974).

The model will include these effects regardless of whether the inclusion of effects significantly improves the fit of the model. Handling of missing PD-L1 data is described in Section 3.6, and is applicable to the primary analysis.

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

KM plots of PFS will be presented by treatment arm, by treatment arm and HRR status, by treatment arm and PD-L1 tumor status, and by treatment arm and Bajorin risk index regardless of whether any of the planned covariates have been unable to be fitted due to the potential limited number of subjects per strata.

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

The assumption of proportionality will be assessed. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, by fitting a time-dependent covariate to assess the extent to which this represents random variation.

If a lack of proportionality is evident, the variation in treatment effect will be described by presenting piecewise HR calculated over distinct time-periods. In such circumstances, the HR can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. If lack of proportionality is found, this may be a result of treatment-by-covariate interactions, which will be investigated. In addition, the KM curve along with landmark analyses (e.g., 1 year PFS rate) will also help in understanding the treatment benefit.

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

Sensitivity Analyses

The HRR status will be compared between the data entered into the IVRS and the vendor source data. A sensitivity analysis for PFS in the HRR mutant group may be performed using the source data for the HRR status if there is a large number of subjects that have had the mutation result incorrectly entered into the IVRS and thus assigned the incorrect strata. The Bajorin risk index will be separately derived from the data in the CRF, and the concordance between the values in the IVRS and CRF reported. A sensitivity analysis may be performed if there is a substantial number of discrepant values observed.

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled timepoints. The midpoint between the time of progression and the previous evaluable RECIST assessment will be analyzed using a stratified log-rank test as described for the primary analysis of PFS. For subjects whose death was treated as a PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust even in highly asymmetric assessment schedules (Sun and Chen 2010). To support this analysis, the mean of subject-level average inter-assessment times will be tabulated for each treatment.

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of subjects who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, and within the same sensitivity analysis, subjects who take 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 anti-cancer therapy. This analysis will be supported by a KM plot of the time to censoring using the PFS data from the primary analysis where the censoring indicator of the PFS analysis is reversed.

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

Subgroup Analyses

The following subgroup analyses will be conducted comparing PFS between durvalumab + olaparib and durvalumab + placebo in the following subgroups of the FAS (but not limited to):

- Sex (male versus female)
- Age at randomization (<65 versus ≥65 years of age)
    - This will be determined from the date of birth (BIRTHDAT in the DM module) and date of randomization (IERNDDAT in the IE module) on the eCRF at screening, or AGE in DM module if AGE is available but BIRTHDAT is completely or partial missing

- Subjects with a partial date of birth (i.e. for those countries where year of birth only is given) will have the 1st of the month imputed if the date is missing, and 1st Jan imputed if date and month is missing.
- Subjects with a missing age value will be included using the mean age (overall) and categorized accordingly.

The subgroup analysis of HRR mutant subjects will be the primary interpretation of the secondary endpoint of PFS in HRRm subjects although this endpoint will also be assessed using a stratified log rank test.

- HRR status (mutant versus wildtype)
- BRCA status
  - HRRm with a BRCAm versus HRRm with a BRCAwt
  - HRRm with a BRCAm versus HRRwt
  - HRRm with a BRCAwt versus HRRwt
- PD-L1 status (high versus low)
- WHO/ECOG performance status (0 and 1 versus 2)
- VES-13 score (<3 versus ≥3)
- Extent of disease (lymph-node-only metastasis versus metastatic disease to any other organ system)
  - Visceral metastasis is defined as metastatic disease to any other organ system apart from lymph-nodes.
- Bajorin risk index (0 versus 1 versus 2)
  - see Table 2 for definition
- Smoking status (smoking versus non-smoking)
- Prior treatment (prior treatment versus no prior treatment)
- Prior immuno-oncology (IO) (prior IO therapy versus no prior IO therapy)

For these subgroup analyses any subject with missing values will be excluded from that particular subgroup.

The subgroup analyses for the HRR status and Bajorin Risk Index will be based on the values reported in the IVRS, all other factors will be based on values recorded on the eCRF as indicated above. Other baseline variables may also be assessed if there is clinical justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors. If however there are too few subjects in the certain categories of the subgroup, a combination of some categories may be applied. If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered as appropriate to present analyses where there are less than 20 events across both treatment groups in a subgroup), the relationship between that subgroup and PFS will not be formally analyzed. In this case, only descriptive summaries will be provided.

For each subgroup level of a factor, the HR (durvalumab + olaparib:durvalumab + placebo) and 95% profile likelihood CIs will be calculated from a Cox proportional hazards model with treatment as the only covariate. 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 will be presented on a forest plot, along with the HR and 95% CI from the overall primary analysis.

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

Consistency of treatment effect between subgroups

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

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 (Gail and Simon 1985).

Additional supportive summaries/graphs

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

Summary statistics for the number of days between RECIST assessments and the number of weeks between the time of progression and the last evaluable tumor assessment prior to progression will be presented for each treatment group. Summary statistics will also be given for the number of days from censoring to data cut-off for all censored subjects.

In addition, duration of follow-up will be summarized using descriptive summary statistics:

- In censored subjects: Time from randomization to date of censoring (date last known to be alive) by treatment arm.
- In all subjects: Time from randomization to the date of death (i.e. overall survival) or to the date of censoring for censored subjects regardless of treatment arm.

The number of subjects prematurely censored will be summarized by treatment arm. A subject would be defined as prematurely censored if they had not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than one scheduled tumor assessment interval plus 1 weeks (9 weeks) prior to the DCO date.

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

All of the derived RECIST 1.1 responses used in analysis will be listed for all randomized subjects. In addition, a summary of new lesions (i.e., sites of new lesions) will be produced.

4.2.3 Overall survival

The OS analysis will be conducted at the following two timepoints;

- 1st OS analysis (OS1 - conducted at the same time as the PFS) - OS and OS18 analyses conducted when approximately 44 subjects have died (29% maturity).
- 2nd OS analysis (OS2) - OS and OS18 analyses conducted when approximately 100 subjects have died (67% maturity).

The alpha for each analysis will be adjusted as described in section 5.

OS will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. The effect of treatment (durvalumab + olaparib vs durvalumab + placebo) will be estimated by the HR together with its 95% CI from a stratified Cox model. KM plots of OS will be presented by treatment arm, by treatment and HRR status, by Treatment and bajorin risk index and by treatment and PD-L1.

Summaries of the number and percentage of subjects who have died, those still in survival follow-up, those lost to follow-up, and those who have withdrawn consent will be provided along with the median OS and 95% CI for each treatment group. Subgroup analyses may be performed if there are a sufficient number of OS events.

4.2.4 PFS6

PFS6 will be summarized with a landmark estimate from the KM curve and corresponding 95% CI by treatment arm. Note: 6 months equates to study day 183 (365.25/12*6).

The PFS6 analysis described above will be repeated for a subset of subjects in the FAS with HRRm.

4.2.5 OS18

The OS18 will be conducted at the same time as the final OS analysis, as detailed in Section 4.2.3.

Overall survival at 18 months will be summarized (using the KM curve) with a landmark estimate from the KM curve and corresponding 95% CI by treatment arm. Note: 18 months equates to study day 548.

4.2.6 Objective response rate

The ORR will be based on the programmatically derived RECIST 1.1 assessment using the Investigator tumor data. The ORR will be compared between durvalumab + olaparib and

durvalumab + placebo using logistic regression models adjusting for the same factors as the primary endpoint PFS (HRR status, PD-L1 and Bajorin risk index). The results of the analysis will be presented in terms of an odds ratio (an odds ratio of greater than 1 will favor the durvalumab and olaparib combination therapy over the durvalumab and placebo) together with its associated profile likelihood CI (using the option 'LRCI' in SAS PROC GENMOD) and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model).

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

The ORR analysis described above will be repeated for a subset of subjects in the FAS with HRRm adjusting for the factors PD-L1 and Bajorin risk index.

4.2.7 Duration of response

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

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

4.2.8 Change in TL tumor size

A waterfall plot will be included of Best percentage change from baseline tumor size (sum of target lesion size) presenting each subject as a separate bar, with the bars ordered from the largest increase to the largest decrease. Reference lines at the –30% and +20% change in TL tumor size level will be added to the plots, which correspond with the definition of 'partial response' and 'progressive disease' respectively. The scale in these plots will be fixed to be from -100 to 100 to avoid presenting extreme values. All progressions will be marked with a '●' and imputed values are clearly marked with '*'.

4.2.9 Patient-reported outcomes

PRO analyses will be conducted for the FAS.

Compliance rates summarizing questionnaire completion at each visit will be tabulated. By visits summaries will use visits windows defined in Section 3.3.

4.2.9.1 EORTC QLQ-C30

Summary tables (mean, SD, median, inter-quartile range, minimum, and maximum) of absolute and unadjusted change from baseline of multi-item symptoms scales , single item scales, functional scales and HRQoL will be presented for each visit by treatment group. Line

plots of the mean score, with corresponding 95% CI, will also be produced by visit and treatment group for the functional scales, multi-item symptom scales, Global health status/HRQoL scale and for the single item symptom scales appetite loss only.

The mean change from baseline in the multi-item symptom scales, the functional scales and HRQoL scale will be analyzed using a mixed model for repeated measures (MMRM) and estimation will be based on Restricted maximum likelihood method (REML). The model will include treatment, the same covariates as the primary endpoint (HRR status, PD-L1 expression and Bajorin risk index), visit and treatment by visit interaction as explanatory variables, and the appropriate baseline EORTC QLQ C30 score as a covariate.

The analysis will make use of all data from baseline up to a selected timepoint cutoff. The cut off point for the MMRM analysis will be determined using the following rules. If at least one of the following conditions is met at a visit, then that visit and any subsequent visits afterward will not be included in the MMRM model.

1. Compliance rate at a visit is <50% in any treatment arm
2. There are less than 20 subjects in a visit in any treatment arm

The study discontinuation visit and the safety follow-up visit will be included in the analysis subject to the rules for assigning visit windows in Section 3.3.1.

The MMRM model will include the fixed categorical effects of treatment, the covariates, visit and treatment by visit interaction. A random intercept term for subject will also be included. The treatment by visit interaction will remain in the model regardless of significance. The adjusted mean change from baseline estimates (obtained from LSMEANS statement as the calculated least square means, adjusted for the random component of the model) and corresponding 95% CIs will be presented by visit (and an overall average across all visits) for each treatment group. A plot of the adjusted mean change from baseline over time and corresponding 95%CI will also be produced by treatment group.

An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom.

If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be tried in order until convergence is reached: Toeplitz with heterogeneity, autoregressive with heterogeneity, Toeplitz, and autoregressive.

PRO assessments are planned to be taken every 4 weeks (±7 days) from the 1st dose of study treatment until 3 months post treatment discontinuation. However, if a subject discontinues study treatment due to toxicity or symptomatic deterioration, PRO assessments are to be performed very 4 weeks (± 7 days) from the last dose of study treatment for the first 3 months and then every 8 weeks until 3 months post PD. This change in the PRO frequency in addition

to the timings of discontinuation and follow up visits, may result in a miss-alignment of visits between subjects. Thus if required a sensitivity analysis may be performed using a piecewise linear model to further support data interpretation.

4.2.9.2 EQ-5D-5L

The evaluable population will comprise a subset of the FAS who have a baseline EQ-5D-5L assessment.

Descriptive statistics will be calculated for each scheduled visit/time point in the study, for each treatment arm. These will report 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. Additionally summary statistics (e.g. n, mean, median, SD, min, max) may 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.

Graphical plots of the mean EQ-5D index score and EQ-VAS score, including change from baseline and associated 95% CI 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, which will be reported outside of the CSR.

4.2.9.3 PGIC

The proportion of subjects for each response category on the PGIC will be summarized descriptively as number of subjects and corresponding percentages for each category in the questionnaire at baseline, each visit, and overall i.e. best response by treatment arm.

4.2.10 Safety analyses

Safety and tolerability data will be presented by treatment arm using the safety analysis set. Safety data will be summarized only. No formal statistical analyses will be performed on the safety data.

Data from all cycles of treatment will be combined in the presentation of safety data. Safety data will be assessed in terms of AEs, physical examination, clinical chemistry, hematology, urinalysis, TSH, vital signs and ECGs. Exposure to durvalumab monotherapy and durvalumab + olaparib, will be summarized, and broken down into individual treatments per arm (that is, durvalumab exposure will be reported separately for the monotherapy vs combination arm). Time on study, dose delays/interruptions, and dose increases and reductions will also be summarized.

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

4.2.10.1 Adverse Events

All AEs, both in terms of current MedDRA preferred term and CTCAE grade, will be summarized descriptively by count (n) and percentage (%) for each treatment group. 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 treatment (whichever occurs first). This will more accurately depict AEs attributable to study treatment only as a number of AEs up to 90 days following discontinuation of treatment are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, some of the AE summaries may also be produced containing AEs observed up until 90 days following discontinuation of study treatment (i.e. without taking subsequent anti-cancer therapy into account).

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

- All AEs
- All AEs possibly related to either study medication (as determined by the reporting investigator)
- AEs with CTCAE grade 3 or 4
- AEs with CTCAE grade 3 or 4, possibly related to either study medication (as determined by the reporting investigator)
- AEs with outcome of death
- AEs with outcome of death possibly related to either study medication (as determined by the reporting investigator)
- All SAEs
- All SAEs possibly related to either study medication (as determined by the reporting investigator)
- AEs leading to discontinuation of study medications, separately for both study medications, durvalumab, and olaparib/placebo
- AEs leading to dose interruption of study medications, separately for both study medications, durvalumab, and olaparib/placebo
- SAEs leading to discontinuation of study medications, separately for both study medications, durvalumab, and olaparib/placebo

Multiple events per subject will not be accounted for.

A high level summary of the number and percentage of subjects in each of the above categories will be presented, with additional categories as required, including but not limited to:

- AEs possibly related to study treatment for each of durvalumab and olaparib/placebo
- AESI categories (See 4.2.10.3)
- Any durvalumab AESI leading to discontinuation of study medications, separately for both study medications, durvalumab, and olaparib/placebo
- Any olaparib AESI leading to discontinuation of study medications, separately for both study medications, durvalumab, and olaparib/placebo
- Grade 3 or 4 durvalumab AESI
- Grade 3 or 4 olaparib AESI
- Durvalumab AESI possibly related to study medication (separately for related to durvalumab and olaparib/placebo)
- Olaparib AESI possibly related to study medication (separately for related to durvalumab and olaparib/placebo)
- Other significant AEs
- Infusion reaction AEs (as determined by the reporting investigator)

A truncated AE table of most common AEs and another table showing most common AEs with CTCAE grade 3 or 4, showing all events that occur in at least 5% of subjects overall will be summarized by preferred term, by decreasing frequency. This cut-off may be modified after review of the data. When applying a cut-off (i.e., x %), the raw percentage should be compared to the cut-off, no rounding should be applied first (i.e., an AE with frequency of 4.9% will not appear if a cut-off is 5%).

Each AE event rate (per 100 subject years) and SAE event rate will also be summarized by preferred term within each system organ class. For each preferred term, the event rate is defined as the number of subjects with at least 1 event during the treatment period plus 90 days follow-up (the initiation of the first subsequent anti-cancer therapy following discontinuation of treatment, in days) divided by the total treatment duration only (excluding the follow up period, in days), summed over subjects and then multiplied by 365.25 x 100 to present in terms of per 100 subject years.

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

In addition, all AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug. Pre-treatment AE's and AE's that occur after a subject has received further therapy for cancer (following discontinuation of IP) will be included in the AE listings. A separate data listing of AEs occurring more than 90 days post-last dose of the latest IP will also be produced.

4.2.10.2 Deaths

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

- Total number of deaths (regardless of date of death)
- Death related to disease under investigation only as determined by investigator
- Death related to disease under investigation and an AE with an outcome of death
    a. AE onset prior to subsequent anti-cancer therapy. Which includes AEs with onset date (or pre-treatment AEs that increase in severity) on or after the date of first dose and up to and including 90 days following the last dose of study treatment, or AE start date <= the date of initiation of the first subsequent therapy (whichever occurs first).
    b. AE onset after start of subsequent anti-cancer therapy. Which includes AEs with start date > 90 days following the last dose of study medication and/or AE start date > the date of initiation of the first subsequent anti-cancer therapy
- AE with outcome of death only
    a. AE onset prior to subsequent anti-cancer therapy. Which includes AEs with an onset date (or pre-treatment AEs that increase in severity) on or after the date of first dose and up to and including 90 days following the last dose of study medication, or AE start date ≤ the date of initiation of the first subsequent anti-cancer therapy (whichever occurs first)
    b. AE onset after the start of subsequent anti-cancer therapy. Which includes AEs with a start date >90 days following the last dose of study medication and/or AE start date > date of initiation of the first subsequent anti-cancer therapy (whichever occurs first)

- Death after end of safety follow-up period (last dose of study medication + 90 days) and not due to disease under investigation
- Unknown reason for death
- Other deaths

This summary will then be repeated for all deaths on treatment or within 90 days of the last dose of study treatment.

A listing of all deaths will also be produced.

4.2.10.3 Adverse events of special interest (AESI)

Preferred terms of durvalumab and olaparib AESI's will be listed separately before DBL and documented in the Study Master File. Grouped summary tables of certain MedDRA preferred terms will be produced. For each 'grouped' term, the number (%) of subjects experiencing any of the specified terms will be presented by maximum CTCAE grade.

Separate summaries for durvalumab AESI and olaparib AESI by grouped term and preferred term for the safety analysis set to be provided are listed below.

- All AESI
- Serious AESI
- AESI with outcome death

4.2.10.4 Infection Adverse events

Infection AEs will be summarized by pooled terms and PTs in two ways: (1) using MedDRA HLGT/HLT pooled terms (2) Custom pooled terms. Preferred terms of infection AE's will be listed before DBL and documented in the Study Master File. For each 'grouped' term, the number (%) of patients experiencing any of the specified terms will be presented by maximum CTCAE grade.

4.2.10.5 Laboratory assessments

Specific outputs should be produced for Hy's Law, ALT, AST and total bilirubin.

Data collected up to and including min(date of last dose of study treatment + 90 days, day before the first dose of subsequent anti-cancer therapy). will be used for reporting. This will more accurately depict laboratory toxicities attributable to study treatment only as a number of toxicities up to 90 days following discontinuation of study treatment are likely to be attributable to subsequent anti-cancer 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 discontinuation of the study treatment (i.e. without taking subsequent anti-cancer therapy into account). Any data post 90 days last dose of study treatment will not be summarized.

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

Shift tables for laboratory values by worst CTCAE grade will be produced, and for specific parameters separate shift tables indicating hyper- and hypo- directionality of change from baseline will be produced. The laboratory parameters for which CTCAE grade shift outputs will be included but not limited are:

- Hematology: Hemoglobin, Leukocytes, Lymphocytes, absolute count-hypo and hyper, Neutrophils, absolute count, Platelets

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

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

Additionally, a summary table of the creatine clearance level and change in creatinine clearance from baseline will be presented by treatment arm and visit. Shift plots of the baseline creatine clearance versus the maximum observation on treatment and by visit may also be presented by subject.

4.2.10.6 Liver Enzyme Elevations and Potential Hy's law

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

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

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

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

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

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

- ALT or AST ≥3x ULN and Total bilirubin ≥2x ULN during the study (Potential Hy's law): The onset date of ALT or AST elevation should be prior to or on the date of Total Bilirubin elevation
  Narratives will be provided in the CSR for subjects who have potential Hy's Law as defined above.

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

Plots of ALT and AST vs. total bilirubin by treatment group will also be produced with reference lines at 3xULN for ALT, AST, and 2xULN for Total bilirubin. In each plot, Total bilirubin will be in the vertical axis.

4.2.10.7 Abnormal Thyroid function

Elevated thyroid stimulating hormone (TSH) will be summarized per treatment group in terms of number (%) of patients with elevated TSH (higher than the upper normal range), low TSH (lower than lower normal range), elevated TSH post-dose and within normal range at baseline, low TSH post-dose and within normal range at baseline. Shift tables showing baseline to maximum and baseline to minimum will be produced.

4.2.10.8 ECGs

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

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

4.2.10.9 Vital signs

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

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

4.2.10.10 ECOG performance status

All Eastern Cooperative Oncology Group (ECOG) performance status will be summarized over time for the FAS.

4.2.11 Pharmacokinetic analyses

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

4.2.12 Immunogenicity analyses

Immunogenicity data will be analyzed descriptively by summarizing the number and percentage of subjects who develop detectable ADAs in the safety analysis set by treatment group.

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

4.2.13 Pharmacokinetic/pharmacodynamic relationships

PK concentration data and summary statistics will be tabulated. Further exploratory analysis of PK data, if conducted, will be reported separately from the main CSR

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modeling approach.

4.2.14 Demographic and baseline characteristics data

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

- Subject disposition (including screening failures and reason for screening failure)
- Important protocol deviations
- Stratification factors from IVRS versus source data (HRR, visceral metastases, Bajorin risk index)
- Inclusion in analysis sets
- Demographics (age, age group, sex, race and ethnicity)
- Subject characteristics at baseline (height, weight, weight group)
- Subject recruitment by country and centre
- Previous disease-related treatment modalities
- Previous chemotherapy prior to this study
- Disease characteristics at initial diagnosis, screening and baseline (Platinum eligibility, HRR status, PD-L1, CrCl (recorded in CISELOM module), VES-13, ECOG

performance status, primary tumor location, histology type, tumor grade and overall disease classification)

- Extent of disease at baseline
- TNM classification at baseline
- Disease related medical history (past and current)
- Relevant surgical history
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use

The medications will be coded following WHO Drug dictionary.

4.2.15 Treatment exposure

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

- Total exposure of each treatment group.
- Actual exposure of each treatment group.
- Total number of cycles received for each treatment group.
- Number of, reasons for, and duration of dose delays/interruptions of durvalumab (per treatment arm) and olaparib/placebo. Dose interruptions will be based on investigator initiated dosing decisions. In addition, interruptions due to AEs and due to reasons other than AEs will be summarized separately.
- RDI (relative dose intensity) of durvalumab (per treatment arm) and olaparib/placebo

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

Exposure over time will be graphically represented by treatment group, showing time since first dose versus the percentage of subjects still on treatment.

4.2.16 Biomarker data

PD-L1 expression status (high versus low) will be summarized with disease characteristics (See 4.2.14).

Summaries and analyses for exploratory biomarkers will be documented in a separate analysis plan and will be reported outside the CSR separately.

4.2.17 Genetic data

DNA variants associated with response, including AEs and/or cancer or its evolution, are exploratory. These exploratory analyses are not expected to be reported in the CSR, but any contribution to the interpretation of the results may be included and will be clearly labeled as exploratory.

4.2.18 Listings

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

- Discontinued subjects (Appendix 12.2.1 in the AZ CSR)
- Protocol deviations (Appendix 12.2.2 in the AZ CSR)
- Subjects excluded from the efficacy analysis (Appendix 12.2.3 in the AZ CSR)
- Demographic data (Appendix 12.2.4 in the AZ CSR)
- Compliance and/or drug concentration data, if available (Appendix 12.2.5 in the AZ CSR)
- Individual efficacy response data (Appendix 12.2.6 in the AZ CSR)
- Adverse events (AEs): all AEs and serious adverse events (SAEs) (Appendix 12.2.7 in the AZ CSR)
- Listing of individual laboratory measurements by subject

5. INTERIM ANALYSES

5.1 Analysis Methods

The study will be unblinded for the primary analysis of PFS, which will be performed at 1 timepoint only, when approximately 118 events have occurred (79% maturity). At the time of the primary analysis, the secondary analyses of ORR, DoR, PFS6, and PRO will also be performed, conditional on the significance of the primary PFS analysis. An interim analysis of OS will also be analyzed at the time of the primary PFS analysis. Data will continue to be

collected after the primary analysis to support a further final analysis of OS and an updated analysis of PFS in the HRRm subgroup. The two OS analysis timepoints are

- 1st OS analysis (OS1) - OS analysis conducted at the same time as the primary PFS analysis when approximately 44 subjects have died (29% maturity).
- 2nd OS analysis (OS2) - OS and OS18 analyses conducted when approximately 100 subjects have died (67% maturity), and an updated analysis of PFS in the HRRm subgroup.

The primary PFS analysis, as well as any OS analyses, will be conducted by AstraZeneca or its delegate.

If the primary PFS analysis is significant, all of the alpha (i.e. 5%) will be recycled for use in the OS analyses. The available alpha will be controlled among the first and second OS timepoints by using the Lan-DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, where the significance level applied at the first analysis (i.e., at the time of 44 events) is dependent upon the proportion of information available. This proportion of information will be calculated at the first analysis using the data available at that time. For example, if the overall alpha level is 5% and if 44% of OS events required at the time of the second OS analysis are available at the time of the interim analysis (i.e., 44/100 events have occurred), the 2-sided significance level to be applied for the first OS analysis would be 0.10% and the 2-sided significance level to be applied for the final OS analysis would be 4.85%. The smallest treatment difference that could be statistically significant at the first analysis of OS is an HR of 0.38 and 0.68 at the second analysis of OS.

5.2 Independent data monitoring committee

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

An IDMC composed of independent experts will be established to perform an interim assessment of the safety of durvalumab + olaparib combination therapy in this population. The first safety review will take place approximately 6 months after the study has started. Safety reviews will be carried out by the IDMC in an unblinded manner.

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

The IDMC will meet approximately every 6 months, unless OTHERWISE requested by the IDMC. IDMC members will be consulted to ensure appropriate frequency.

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

6. CHANGES OF ANALYSIS FROM PROTOCOL

The SAP was updated to version 2.0 following the CSP that was updated to version 2.0 (dated 20 July 18). A full list of changes to study protocol can be found in the version history of the CSP.

The following changes were made from version 1.0 of the CSP (dated 16 November 2017);

In the protocol inconsistent analysis techniques for PFS are defined in 9.4.2 Table 17 and 9.4.2.1. In 9.4.2 the primary analysis is defined as using a stratified Cox proportional hazards model and in 9.4.2.1 the analysis is defined as a Cox proportional hazards model with covariates.

The planned analysis of PFS and OS, as detailed in sections 4.2.2 and 4.2.3 of the SAP, are a Cox proportional hazards model with covariates for PD-L1 status and Bajorin risk index as per 9.4.2.1 of the CSP.

Table 17 (Section 9.4.2) in the CSP incorrectly stated that the HR would be calculated for the PFS6 and OS18 analyses. The correct analyses were specified in the protocol in the analysis text (section 9.4.2.3 and 9.4.2.6 in the CSP). Table 7 has been updated to include the corrected analysis text.

The PRO analysis for time to deterioration has been amended to use the same analysis methods as the primary PFS analysis (Cox proportional hazards model), rather than the log-rank test.

Table 17 (Section 9.4.2) in the CSP incorrectly included EQ-5D-5L endpoints in the change from baseline analysis, indicating that an MMRM analysis would be performed for this PRO. Only summary statistics will be produced for the EQ-5D-5L endpoints.

The text the biomarker exploratory analysis in the study objectives listed in Table 1 has been amended to include the collection of urine. This matches with the planned collection of urine according to the SoA (Table 1 in CSP).

For efficacy endpoints in Table 1, text has been added to clarify that tumor assessments are performed by investigator as per RECIST 1.1.

For PRO endpoints in Table 1, text has been added to clarify that all scales will be analyzed for EORTC QLQ-C30, the indicated endpoints in the table are of primary interest.

The SAP was updated to version 3.0 and the current CSP version is version 2.0 (dated 20 July 2018). The following updates to the SAP are inconsistent with the current CSP:

Sensitivity analysis for attrition bias for OS (KM plot of time to censoring, where the censoring indicator of the primary analysis is reversed) has been removed.

7. REFERENCES

Bajorin et al 1999
Bajorin DF, Dodd PM, Mazumdar M, Fazzari M, McCaffrey JA, Scher HI, et al. Long-term survival in metastatic transitional-cell carcinoma and prognostic factors predicting outcome of therapy. J Clin Oncol 1999;17(10):3173-81.

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

EuroQoL Group 1990
EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208.

EuroQoL Group 2013
EuroQol Group. EQ-5D-5L user guide: basic information on how to use the EQ-5D-5L instrument, version 2.0 [2013 Oct]. Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L_v2.0_October_2013.pdf. Accessed 21 November 2013.

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

VENTANA PD-L1 (SP263) Assay [package insert] 2017
Available from: URL:
https://www.accessdata.fda.gov/cdrh_docs/pdf16/P160046C.pdf
Accessed 23 March 2018

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

Oemar and Janssen 2013
Oemar M, Janssen B. EQ-5D-3L user guide: basic information on how to use the EQ-5D-3L instrument V5.0 [2013 Oct]. Available from: URL:
http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

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

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's method with the conventional approach for interval-censored data analysis. Stat Biopharm Res 2010;2;97-108.

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

Saliba et al 2000
Saliba D, Orlando M, Wenger NS, Hays RD, Rubenstein LZ. Identifying a short functional disability screen for older persons. J Gerontol A Biol Sci Med Sci 2000;55:M750-M756.

Saliba et al 2001
Saliba D, Elliott M, Rubenstein LZ, Solomon DH, Young RT, Kamberg CJ, et al. The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. J Am Geriatr Soc 2001;49(12):1691-99.

8. APPENDIX

Appendix A Important protocol deviation handling

The master list of IPDs is shown in Error! Reference source not found..

Table 8: IPD identification

Deviation Code	Deviation	Rationale for CSR reporting
1	Patients who deviate from inclusion criteria 5, 6 and 8 and exclusion criteria 1, 9, 10, 11 and 18 per the CSP	Potential impact on the specific subject population intended for the study and potential safety concerns.
1.1	Enrolled but not met the inclusion criteria 5: Histologically or cytologically documented TCC/UC of the urothelium at screening with Unresectable, Stage IV disease and no prior systemic therapy for unresectable, Stage IV disease	

Deviation Code	Deviation	Rationale for CSR reporting
1.2	Enrolled but not met the inclusion criteria 6: Ineligible for platinum based chemotherapy	
1.3	Enrolled but not met the inclusion criteria 8: World Health Organization (WHO)/ECOG performance status of 0,1 or 2 at enrolment and randomization	
1.4	Enrolled but not met the exclusion criteria 1: Active or prior documented autoimmune or inflammatory disorders	
1.5	Enrolled but not met the exclusion criteria 9: Prior exposure to a PARP inhibitor or immune- mediated therapy.	
1.6	Enrolled but not met the exclusion criteria 10: Any concurrent chemotherapy, IP, biologic, or hormonal therapy for cancer treatment.	
1.7	Enrolled but not met the exclusion criteria 11: Current or prior use of immunosuppressive medication within 14 days before the first dose of the IP	
1.8	Enrolled but not met the exclusion criteria 18: Concurrent enrollment in another clinical study, unless it is an observational(non-interventional) clinical study or the follow-up period of an interventional study.	
2	Patients randomized who received treatment other than that to which treatment arm they were randomized to	Potential impact on the Full Analysis Set and consequently the efficacy results

Deviation Code	Deviation	Rationale for CSR reporting
3	Received prohibited concomitant medications (including other anticancer agents)	Potential confounding impact on the treatment effect and safety of the patients
4	No baseline RECIST 1.1 assessment on or before date of randomization	Potential impact on the primary efficacy results
5	Baseline RECIST scan > 42 days before date of randomization	Potential impact on the primary efficacy results

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: d933ic00003-sap-ed-3		
Document Title:	Statistical Analysis Plan Edition 3	
Document ID:	CCI [redacted]	
Version Label:	1.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
07-Oct-2019 10:49 UTC	PPD	Qualified Person Approval
08-Oct-2019 14:36 UTC	PPD	Qualified Person Approval
04-Oct-2019 10:22 UTC	PPD	Author Approval

Notes: CCI [redacted]