Statistical Analysis Plan	
Study Code	D8488C00001
Edition Number	2.0
Date	20th September 2019

A single arm, open label, phase IIb study to assess the efficacy and safety of the combination of cediranib (30 mg orally daily) and olaparib (200 mg orally twice daily) tablets in women with recurrent platinum resistant epithelial ovarian cancer, including fallopian tube and/or primary peritoneal cancer who do not carry a deleterious or suspected deleterious germline *BRCA* mutation

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse Event
AESI	Adverse Event of Special Interest
ALP	Alkaline Phosphatase
ALT	Alanine Transaminase
AML	Acute myeloid leukaemia
ANC	Absolute Neutrophil Count
AST	Aspartate Transaminase
ATC	Anatomic-Therapeutic-Chemical
BDRM	Blinded Data Review Meeting
BoR	Best Objective Response
CI	Confidence Interval
CLIA	Clinical Laboratory Improvement Amendments
CR	Complete Response
CRF	Case Report Form
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database Lock
DCO	Data Cut-off
DCR	Disease Control Rate
DoR	Duration of Response
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EFR	Evaluable For Response
EORTC-QLQ	European Organisation for Research and Treatment of Cancer – Quality of Life Questionnaire
FAS	Full Analysis Set

Abbreviation or special term	Explanation
FFPE	Formalin Fixed, Paraffin Embedded
GCIG	Gynaecologic Cancer InterGroup
HLT	High Level Term
HRD	Homologous Recombination Deficiency
HRQoL	Health-Related Quality of Life
HRR	Homologous Recombination Repair
ICR	Independent Central Review
IP	Investigation Product
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
LD	Longest Diameters
LFT	Liver Function Test
LVEF	Left Ventricular Ejection Fraction
MCV	Mean Cell Volume
MDS	Myelodysplastic Syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic Resonance Imaging
NA	Not Applicable
NE	Not Evaluable
NTL	Non-Target Lesion
OAE	Other significant Adverse Events
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PFS	Progression Free Survival
PK	Pharmacokinetic
PR	Partial Response
PRES	Posterior Reversible Encephalopathy Syndrome
PRO-CTCAE	Patient Reported Outcomes – Common Terminology Criteria for Adverse Events
PT	Preferred Term

Abbreviation or special term	Explanation
QoL	Quality Of Life
RDI	Relative Dose Intensity
REACT	Real-time Analytics for Clinical Trials
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SOC	System Organ Class
TBL	Total Bilirubin
TDT	Time to Discontinuation or Death
TL	Target Lesion
ULN	Upper Limit of Normal
VEGF	Vascular Endothelial Growth Factor
WBC	White Blood Cells

AMENDMENT HISTORY

Date	Brief description of change
19th September 2019	Amended signature pages with latest details of study statistician and global product statistician. Updated subject to patient in SAP for consistency. Removed reference to PK analysis set population as no PK will be analysed for study CSR. Updated sample size and OS analysis per latest Protocol Amendment version 5. Added clarity to last date calculation for OS analysis in section 3.2.2. Clarified that OAEs are different to AESIs in section 3.3.3. Amended disposition section 4.2.1 for clarification of enrolment. Added platinum sensitivity and HRR/BRCA information to section 4.2.3. Added PRO-CTCAE analysis to section 4.2.8. Removed post treatment AE analysis from section 4.2.9. Added AEs leading to dose interruption/modification/reduction to section 4.2.9. Changes from protocol section updated as new protocol amendment means that the changes mentioned are now present in version 5 which this SAP is based off. Replaced references to randomisation with start of treatment
9th September 2016	New document

1. STUDY DETAILS

1.1 Study objectives

1.1.1 Primary objective

To determine the efficacy of the combination of cediranib and olaparib in patients with recurrent platinum resistant epithelial ovarian, fallopian tube and/or primary peritoneal cancer who do not carry deleterious or suspected deleterious germline BRCA mutations, as assessed by Objective Response Rate (ORR) by independent central review (ICR), using RECIST version 1.1.

1.1.2 Secondary objectives

To assess the efficacy of the combination of cediranib and olaparib in this patient population in terms of:

- Objective Response Rate by Investigator assessment using RECIST 1.1
- Duration of Response (DoR) by ICR and investigator assessment using RECIST 1.1
- Disease Control Rate (DCR) by ICR and investigator assessment using RECIST 1.1
- Progression Free Survival (PFS) by ICR and investigator assessment using RECIST 1.1
- Time to discontinuation or death (TDT)
- Overall Survival (OS)

To assess the efficacy of the combination of cediranib and olaparib in subgroups of patients

  - carrying and not carrying a somatic deleterious or suspected deleterious variant in either of the *BRCA* genes (*BRCA1* or *BRCA2)*
  - carrying or not carrying a deleterious or suspected deleterious variant in homologous recombination repair (HRR)-associated genes identified with current and potential future tumor based *BRCA* or HRR gene mutation assays

To evaluate disease related symptoms and health-related quality of life in patients treated with combination of cediranib and olaparib tablets when compared with baseline data in terms of:

- European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ) C30
- EORTC QLQ OV-28

1.1.3 Safety objective

To evaluate the safety and tolerability of the combination of cediranib and olaparib using Adverse Events (AEs), and treatment emergent changes in vital signs and laboratory parameters.

1.1.4 Exploratory objectives

To assess adverse events by patient self-reporting of specific Patient Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™) symptoms.

To assess exploratory biomarkers of potential homologous recombination deficiency (HRD), including genomic scarring, *BRCA1* methylation, and other genes that might modify HRD.

To assess exploratory angiogenic biomarkers such as vascular endothelial growth factor (VEGF)-A, C, D; sVEGFR2, Tie2, Ang1, 2; leptin; interleukin 6 (IL6) and exploratory immunologic biomarkers.

To evaluate exploratory biomarkers of potential resistance/sensitivity to cediranib or olaparib.

To assess exploratory biomarkers of gene expression signature for angioimmune/angio/immune/metabolic categorization.

Data relating to exploratory biomarkers may be reported separately to the CSR.

1.2 Study design

This is an open label, Phase IIb, single arm, multi-center study to assess the efficacy and safety of the combination of cediranib and olaparib in platinum resistant relapsed high grade serous, high grade endometroid or clear cell ovarian, fallopian tube or primary peritoneal carcinoma who have received at least 3 prior lines of chemotherapy and who have no evidence of deleterious or suspected deleterious *gBRCA* mutation(s) as defined by a local test conducted in an appropriately accredited laboratory (e.g., Clinical Laboratory Improvement Amendments [CLIA]-certified) that includes comprehensive DNA sequencing and large rearrangement analysis of *BRCA1* and *BRCA2*. All patients should have received prior antiangiogenic treatment e.g. bevacizumab, either in a first line or a recurrent setting. To be eligible to enter the study, all patients should have measurable disease (as assessed by the Investigator), defined as at least 1 lesion that can be accurately assessed at baseline by computed tomography (CT)/ magnetic resonance imaging (MRI) and is suitable for repeated assessment per RECIST 1.1. All patients will have RECIST 1.1 tumor assessments at screening (within 28 days prior to the start of study treatment) and every 8 weeks (±1 week) after start of treatment until objective radiological disease progression or withdrawal of consent. Objective radiological response on treatment (complete response [CR] or PR) will need to be confirmed during next RECIST 1.1 visit assessment to ensure that identified responses are not result of measurement error.

Patients will receive the combination of:

- Cediranib tablets 30 mg oral once daily
- Olaparib tablets 200 mg oral twice daily

Patients should continue to receive IPs until objective radiological disease progression, unacceptable toxicity or withdrawal of consent.

The study originally planned to recruit 100 patients over approximately 12 months. Due to slower than projected recruitment, the study now plans to recruit approximately 60 patients. An interim futility analysis will be undertaken once the first 20 evaluable (by Investigator assessment) patients have received at least 1 dose of IP and have had opportunity to be followed for at least 4 months. If fewer than 4 radiological tumor responses are observed at this point, the trial may be considered futile and stop enrollment. Enrollment will continue whilst the futility analysis is being conducted.

The primary analysis of the study will be based on measurement of ORR in the evaluable for response (EFR) analysis set (CR+PR; confirmed responses only) by Independent Central review (ICR) using RECIST 1.1. The primary analysis will also report analyses of DoR and disease control rate (DCR) by ICR in the EFR analysis set. Other variables analyzed at the primary analysis (ORR, DoR, DCR based on investigator assessment; PFS [ICR and investigator assessment], time to discontinuation or death [TDT], OS, patient-reported outcome [PRO] variables and safety variables) will be summarized using the FAS.

The data cut-off for the primary analysis will take place 8 months after the last patient has received their first dose of IP and is the final analysis for the study. This primary data cut-off will be followed by clinical database lock.

An analysis of OS will also occur using this data cut-off point.

Patients continuing to receive clinical benefit without meeting any discontinuation criteria during the study, will complete a Final Protocol Visit (FPV) which will occur at their last scheduled visit prior to the primary data cut-off. Beyond the FPV, patients may continue to receive cediranib and/or olaparib if they are deriving clinical benefit in the opinion of the investigator, and not fulfilling any discontinuation criteria.

All patients will be required to provide an archival formalin fixed, paraffin embedded (FFPE) tumor sample from the primary or preferably recurrent cancer. If an archival tumor sample is not available, a tumor sample from a fresh biopsy is acceptable. Please note that a lesion that has been biopsied during screening cannot be selected as a target lesion for RECIST assessment. Tumor samples will be tested for tumor (i.e., somatic [non-germline] or non-germline) deleterious or suspected deleterious mutation in either of the *BRCA* genes (*BRCA1* or *BRCA2*) and the following homologous recombination repair (HRR)-associated genes: *ATM*, *BRIP1*,

PALB2, *RAD51C*, *BARD1*, *CDK12*, *CHEK1*, *CHEK2*, *FANCL*, *PPP2R2A*, *RAD51B*, *RAD51D*, and *RAD54L*.

1.3 Number of patients

The study originally planned to recruit 100 patients over approximately 12 months. Due to slower than projected recruitment, the study now plans to recruit approximately 60 patients. The primary analysis scheduled at 8 months after the last patient received her first dose of IP (approximately 20 months after study start), allows all patients to have at least 4 RECIST follow-up assessments.

An interim futility analysis of ORR (confirmed and/or unconfirmed CRs and PRs based on investigator assessment) will take place after 20 patients have received at least 1 dose of IP and have had opportunity to be observed for at least 4 months. If less than 4 of the 20 have responded the study may be stopped for futility. If 4 or more patients have responded the study will continue to the final analysis.

The study design was formulated within a Bayesian framework (Fisch et al 2015) based on dual criteria to show improvement over the expected response rate (RR) under current standard treatments (20% RR targeted as an improvement over <15% RR based on chemotherapy, Davis et al 2014) and to have a RR of a relevant size (approximately 30%, similar to the RR seen for bevacizumab plus chemotherapy in the AURELIA trial, Pujade-Lauraine et al 2014) defined as follows:

1. to have high confidence (>90%) that the true ORR is at least 20% or probability (ORR≥ 20%) >0.9 and
2. an observed rate of 30/100 is seen at the final analysis.

The futility rule above (stop if number of responses is <4), was chosen so that the trial may stop if there is a low chance (<10% predictive probability) of 30 or more responses being observed at the final analysis. Computation of posterior and predictive probabilities used a neutral prior i.e. beta (1/3, 1/3) distribution (Kerman 2011), for the ORR. The predictive probability of observing a given number of responses at the end of the trial given the observed responses at the interim analysis is computed using a beta binomial distribution.

With 30 responders out of 100 at the final analysis we would be >99% confident that the ORR is greater or equal to 20% and 94.6% confident that the ORR is ≥23%. It is noted that the interim analysis is based on unconfirmed, investigator assessed RECIST data, whereas the final analysis will be based on ICR assessed data with responses confirmed at least 4 weeks after the initial response. Data from the OCEANS study indicated an approximate 1.5% reduction in response by ICR compared with investigator response (Aghajanian et al 2014). If only 26 patients out of 100 were to meet the latter definition of response at the final analysis there would still be >90% confidence that the true rate was greater than or equal to 20%.

Assuming median PFS of 6 months, 73% PFS events are expected at the primary analysis at 8 months after the last patient receives her first dose of IP.
The original analysis for OS at 12 months after last patient receives her first dose of IP would expect 59% deaths.

The study now plans to recruit approximately 60 patients and analysis of OS, originally planned at 12 months after the last patient received her first dose of IP, will now occur at 8 months after the last patient received her first dose of IP. Based on the revised duration of recruitment and assuming a median PFS of 6 months as originally planned and a median of 12 months for OS, 80% PFS events and 57% OS events are expected to occur at the time of the analysis.

Table 1 Response Rates and associated 95% CIs for a sample size of 60

Response (n)	Response (%)	95% CIs
5	8.3%	(1.3%,15.3%)
10	16.7%	(7.2%, 26.1%)
15	25.0%	(14.0%, 36.0%)

2. ANALYSIS SETS

2.1 Definition of analysis sets

Table below details the analysis set for each type of endpoint.

2.1.1 Full analysis set

All patients who received at least one dose of either IP will be included in the full analysis set.

2.1.2 Efficacy analysis set (evaluable for response population [EFR])

The evaluable for response (EFR) analysis set will be all patients who have received at least one dose of treatment and have measurable disease at baseline according to the independent review of baseline imaging data.

The primary analysis of ORR, and analyses of DoR and DCR by ICR will be produced based on the EFR analysis set. All other efficacy variables (ORR, DoR, DCR based on investigator assessment; PFS [ICR and investigator assessment], OS, TDT) and PRO variables will be summarized using the FAS.

Demographic data will be summarized for the FAS and to support the primary analysis, also for the EFR set (if different).

2.1.3 Safety analysis set

This is the same set as the FAS; i.e. all patients who received at least one dose of either IP.

Table 2 Analysis sets used for each endpoint

Endpoint	Analysis set
Primary endpoint:	
ORR by ICR (primary analysis)	Evaluable for response
- ORR by ICR (sensitivity analysis)	Full analysis set
- ORR by Investigator assessment (sensitivity analysis)	Full analysis set
Secondary endpoints:	
DoR, DCR by ICR	Evaluable for response
PFS by ICR	Full analysis set
DoR, DCR, PFS by investigator assessment (sensitivity analyses)	Full analysis set
Time to discontinuation or death	Full analysis set
Overall survival	Full analysis set
Demography and disposition	Full analysis set
	Evaluable for response (if different to FAS)

2.2 Violations and deviations

None of the deviations will lead to patients being excluded from the analysis sets described in Section 2. A per-protocol analysis excluding patients with significant protocol deviations is not planned; however, a 'sensitivity analysis on the primary endpoint will be performed excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients:

- did not have the intended disease or indication or
- fulfil deviation 3 (as labelled in brackets below) .

The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock.

Eligibility criteria deviations are deviations from the protocol inclusion and exclusion criteria. Post-entry deviations are deviations from the protocol that occurred after the patient was assigned to the study.

The following general categories will be considered important deviations and be listed and discussed in the CSR as appropriate for the study. If a sensitivity analysis is conducted then patients with these deviations will be excluded from the sensitivity analysis:

- Patients who deviate from key entry criteria (inclusion criteria 2, 3, 4, 8, 9, and exclusion criteria 4 from the CSP section 3.1 and 3.2) (Deviation 1).
- Baseline RECIST scan > 28 days before date of first dose of study IP (Deviation 2).
- No baseline RECIST 1.1 assessment on or before date of first dose of study IP (Deviation 3).
- Received prohibited concomitant medications (including other anti-cancer agents) (Deviation 4). Please refer to the Clinical Study Protocol (CSP) section 7.7 for those medications that are detailed as being 'excluded' from permitted use during the study. This will be used as a guiding principle for the physician review of all medications prior to database lock.

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

In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification of deviations will be made at the blinded data review meeting (BDRM) prior to database lock or data cut-off. Decisions made at the BDRM will be documented and approved by AstraZeneca prior to analysis.

All important deviations related to the study inclusion or exclusion criteria and study conduct will be identified.

Patients who were enrolled into the study but did not receive study treatment will be listed.

3. PRIMARY AND SECONDARY VARIABLES

The primary analysis of ORR (primary endpoint), DoR, PFS, and DCR (secondary endpoints) will be reported based on the Independent Central Review of radiological data. ORR will also be reported based on the Investigator assessment (secondary endpoint).

In addition, sensitivity analyses will be performed of DoR, PFS, and DCR using investigators assessments of RECIST.

3.1 Derivation of RECIST visit responses

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

Baseline radiological tumor assessments are to be performed no more than 28 days before the start of IPs and ideally as close as possible to the start of IPs. Tumor assessments are then performed every 8 weeks following the first dose of IPs until disease progression.
If an unscheduled assessment is performed, and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits. This schedule is to be followed in order to minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

An ICR of all radiological imaging data will be carried out using RECIST version 1.1 and the overall visit response will be provided at each visit (i.e. the reviewers will provide the overall visit response according to RECIST 1.1 and no programmatic derivation of visit response is necessary).

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

Please refer to sections 3.1.1 and 3.1.2 for the definitions of CR, PR, SD and PD.

RECIST outcomes (ie PFS, ORR etc.) will be calculated programmatically for both the ICR and 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 (LD) (except lymph nodes which must have short axis ≥ 15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. A patient can have a maximum of 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 start of treatment 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.

Table 3 TL visit responses

Visit Reponses	Description
Complete Response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD. Stable disease can only be declared at or after the first scheduled RECIST scan.
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

Rounding of TL data

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

Missing TL data

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

- A new lesion is recorded
- A NTL visit response of PD is recorded
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared

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

If there is at least one TL measurement missing and a visit response of PD cannot be assigned, the 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 and the sum of lesions meets the criteria for PD then response will be set to PD
- Step 4: If after steps 1 – 3 a response can still not be determined the response will be set to remain as CR

TL too big to measure

If a 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 / embolisation), 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, ignore the lesion diameter for those lesions with intervention and scale up as described below as long as there remain ≤ 1/3 of the TLs with intervention. If the scaling results in a visit response of PD then the patient would be assigned a TL response of PD.

- Step 3: If after both steps PD has not been assigned, then, if appropriate, a scaled sum of diameters will be calculated (as long as ≤ 1/3 of the TLs have interventions), and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10mm for lymph nodes) and each lesion that has experienced intervention also has a value of 0 recorded. If scaling up is not appropriate due to too few lesions without intervention 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 sum of TL diameters, lesions with intervention should not be used and the sum of diameters scaled up where appropriate (as per step 2 above).

If ≤ 1/3 of the TL measurements have interventions 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-intervention lesions to the nadir sum of diameters excluding the lesions with interventions..

If > 1/3 of TL measurements have interventions then TL response will be NE, unless the sum of diameters of TL without intervention would result in PD (i.e. if using a value of 0 for non-intervention 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).

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	Intervention
Sum	**29.3**	**26**

Lesion 5 has had an intervention 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 are lesions with intervention. 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 with intervention, 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.

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 4 NTL Visit Responses

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

To achieve 'unequivocal progression' on the basis of NTLs, there must be an overall level of substantial worsening in non-target disease such that, even in the presence of SD or PR in TLs, the overall 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.

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

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

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

3.1.3 Overall visit response

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

Table 5 Overall visit responses

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

3.2 Efficacy Variables

The endpoints (ORR, DoR, DCR, PFS) will be derived from the overall visit response date and the scan dates. RECIST assessments/scans contributing towards a particular visit may be performed on different dates and the date of progression will be determined based on the earliest of the scan dates of the component that triggered the progression.

Independent Central Review of RECIST based assessments

The ICR of all radiological imaging data will be carried out using RECIST version 1.1. All radiological scans for all patients (including those at unscheduled visits, or outside visit windows) will be provided to the ICR. Prior radiotherapy reports for patients (at baseline) and information on any lesions that were biopsied to provide a tumor sample for study entry will also be provided to allow the selection of appropriate target lesions. The imaging scans will be reviewed by two independent radiologists using RECIST 1.1 criteria and will be adjudicated if required. For each patient, the ICR will define the overall visit response data (CR, PR, SD, PD or NE) and the relevant scan dates for each time point (i.e. for visits where response or progression is/is not identified). If a patient has had a tumor assessment which cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response will be assigned as PD).

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

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

Investigator RECIST based assessments

From the investigators review of the imagining scans, the RECIST tumor response data will be used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD or PD depending on the status of their disease compared with baseline and previous assessments. If a patient has had a tumor assessment which cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response will be assigned as PD).

3.2.1 Primary Endpoint

The primary efficacy endpoint variable is Objective Response Rate (ORR), defined as the percentage of patients with objective response according to RECIST 1.1 by ICR using the evaluable for response analysis set (all patients who have received at least 1 dose of IPs and have measurable disease at baseline according to ICR).

The secondary endpoint variable of ORR using the investigators assessment of RECIST and based on the full analysis set will also be analysed.

An additional sensitivity analysis of ORR will be performed using the ICR data summarized using the full analysis set.

ORR will only include patients whose response has been confirmed by a second scan at least 4 weeks after the initial response. A confirmed response of CR/PR means that a response of CR/PR is recorded at 1 visit and confirmed by repeat imaging not less than 4 weeks after the visit when the response was first observed with no evidence of progression between the initial and CR/PR confirmation visit. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Patients who receive a subsequent anti-cancer therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy) and then respond will not be included as responders in the ORR (i.e. both visits contributing to a response must be prior to subsequent therapy for the patient to be considered as a responder).

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

Table 6 below outlines scenarios for the first and subsequent visit responses and the rules that will be followed to determine whether such patients would be counted in the numerator of ORR for the primary analysis and their classification for Best Objective Response (BOR) for the primary analysis. BOR is included as supportive data for the ORR endpoint.

Table 6 Confirmed ORR and BoR criteria

First time point	Second time point	Confirmed Responder for ORR	Best objective response
CR	CR	Yes	CR
CR	PR	Scenario not possible, CR can only be followed by CR PD, NE	SD, PD or PR[a]
CR	SD	Scenario not possible, CR can only be followed by CR PD, NE	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	No	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	No	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	Yes	PR
PR	PR	Yes	PR
PR	SD	No	SD
PR	PD	No	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	No	SD provided minimum criteria for SD duration met, otherwise NE
CR/PR	Second timepoint not reached at time of analysis	No*	SD provided minimum criteria for SD duration met, otherwise NE
CR/PR	Patient withdrew consent before second timepoint reached	No*	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE		NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable. (Eisenhauer et al 2009).

* Note for the interim analysis, ORR will include both confirmed and unconfirmed responses, so patients in these scenarios would be counted as responders.

[a] If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met (scan at least 7 weeks post-baseline). However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

3.2.2 Secondary endpoints

Duration of response (DoR)

DoR will be defined as the time from the date of first documented response (which is subsequently confirmed) until date of documented progression or death in the absence of disease progression (ie, date of PFS event or censoring – date of first response + 1). The end of response should coincide with the date of progression or death from any cause used for the PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first visit that was CR or PR that was subsequently confirmed.

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

Disease control rate

Disease control rate is defined as the percentage of patients who have a best overall response of CR or PR or SD (at 6 months). For SD to be counted, patients should have demonstrated SD for a minimum interval of 23 weeks (24 weeks minus 1 week to allow for an early assessment within the assessment window, ie, 161 days) following the start of treatment.

Progression free survival (PFS)

PFS is defined as the time from date of first dose until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from study therapy or receives another anti-cancer therapy prior to progression (i.e. date of PFS event or censoring – date of first dose + 1). Patients who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST assessment. However, if the patient progresses or dies after two or more missed RECIST assessments visits, the patient will be censored at the time of the latest evaluable RECIST assessment prior to the two missed visits.

If the patient has no evaluable 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).

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

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

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

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

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

Overall visit assessments will be determined for each assessment (scheduled or unscheduled) and will contribute to the derivation of PFS. Objective disease progression is defined as: at least a 20% increase in the sum of the diameters of the target lesions (compared with a previous minimum sum) and an absolute increase of >5 mm, an overall assessment of progression or a new lesion.

Overall survival (OS)

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

Survival calls will be made in the 2 weeks following the date of the data cut-off for the primary analysis, and if patients are confirmed to be alive (or if the death date is post the relevant data cut-off date) these patients will be censored in the associated analyses at the date of the data cut off. Data will be recorded on the SURVIVE module of the eCRF and this dataset will be used as priority for analysis.

For the OS analysis for the primary analysis at the 8 month DCO the Date last known alive on survival status CRF (SURVIVE: SUR_DAT) will be used.

If this date is missing the last date for each individual patient is defined as the latest among the following dates recorded on the case report forms (CRFs):

• AE start and stop dates (AE: AESTDAT, AEENDAT)

• Admission and discharge dates of hospitalization (SERAE: AESHODAT, AESDIDAT)

• Study treatment date (EX & EX1: EXSTDAT)

• End of treatment date (EX & EX1: EXENDAT)

• Laboratory collection dates (LB & LB1-6: LBDAT)

• Date of vital signs (VS & VS1: VSDAT)

• Disease scan dates on RECIST CRF (1,3RECIST1: R1ASMDAT)

• Start and stop dates of alternative anticancer treatment (CAPRX1: CXSDAT, CXEDAT)

• Start and stop dates of post-IP radiotherapy (CAPRXR1: CXRDSDAT, CXRDEDAT)

• End of study date (DS: DSSTDAT)

• Visit date (VISIT: VIS_DAT)

The status of ongoing, withdrawn (from the study) and "lost to follow-up" patients at the time of the final OS analysis should be obtained by the site personnel by checking the patient's notes, hospital records, contacting the patient's general practitioner and checking publicly-available death registries. In the event that the patient has actively withdrawn consent to the processing of their personal data, the vital status of the patient can be obtained by site personnel from publicly-available resources where it is possible to do so under applicable local laws.

Best objective response (BoR)

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in section 3.1.3. It is the best response a patient has had following start of treatment, but prior to starting any subsequent cancer therapy and up to and including RECIST progression or the last evaluable assessment in the absence of RECIST progression. Categorisation of BoR will be based on RECIST using the following response categories: CR, PR, SD, PD and NE.

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

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

For patients 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 start of

treatment, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST assessments, if the death occurs > 9 weeks after start of treatment then BoR will be assigned to the NE category.

Time to discontinuation of IPs or death (TDT)

Time to discontinuation of IPs or death is defined as the time from the date of first dose of IP to the earlier of the date of discontinuation of both IPs, or death date. Both IPs should be discontinued at the same time but in rare circumstances discontinuation of one of the IPs before the other may be allowed. In such a case, TDT will reflect the time from the date of first dose to the earliest IP discontinuation date.

Any patient not known to have died at the time of analysis and not known to have discontinued study treatment will be censored based on the last recorded date on which the patient was known to be alive.

3.3 Safety and tolerability variables

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

Safety and tolerability will be assessed in terms of AEs, deaths, laboratory data, vital signs including blood pressure and ECGs. Refer to section 4.1 for general considerations for safety assessments. The safety analysis set will be used for all summaries.

3.3.1 Exposure and dose interruptions

Exposure will be defined as follows:

Total (or intended) exposure of study treatment

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

Actual exposure of study treatment

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

The actual exposure calculation makes no adjustment for any dose reductions that may have occurred. An interruption is defined to be any time a patient has not received study treatment for at least 1 day.

Missed or forgotten doses

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

Patients who permanently discontinue during a dose interruption

If a patient permanently discontinues study treatment during a dose interruption, then the date of last administration of study medication recorded on DOSDISC will be used in the programming

Safety Follow-up

- Total Safety Follow-up = min((last dose date + 30 days), date of withdrawal of consent, date of death, date of DCO) – first dose date +1

3.3.2 Dose Intensity

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

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

Intensity of cediranib and olaparib will summarised separately. The intended cumulative dose is defined as 200mg olaparib twice daily and 30mg cediranib once daily.

3.3.3 AEs

AEs and SAEs will be collected throughout the study, from date of first dose of either IP until 30 days after the last dose of study medication. The Medical Dictionary for Regulatory Activities (MedDRA)(using the latest MedDRA version available at database lock) 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).

If partial dates of onset occur, a conservative approach will be followed and events will be assumed to be treatment emergent unless there is convincing evidence to the contrary.

The following variables will be collected for each AE;

- AE verbatim
- The date when the AE started and stopped
- Changes in CTCAE grade will be reported for AEs of diarrhoea, fatigue, hypertension and nausea. For all other AEs, only the highest attained CTCAE grade will be reported.
- Whether the AE is serious or not
- Investigator causality assessment against each of the IP(s) (yes or no)
- Action taken with regard to each of the IP(s)
- Outcome (if resolved with sequelae then sequelae should be recorded).

In addition, the following variables will also be collected for SAEs:

- Date AE met criteria for SAE
- Date Investigator became aware of SAE
- AE seriousness criteria
- If the patient was hospitalized:
- Date of hospitalization
- Date of discharge

The following variables will also be collected for deaths:

- Probable cause of death
- Date of death
- Autopsy performed
- Causality assessment in relation to IPs and/or Study procedure(s).

Groups of preferred terms will be used to identify adverse events of special interests (AESIs); these preferred terms will be listed before database lock and documented in the Study Master File. The AESIs for cediranib will include PRES, arterial thromboembolic events, symptomatic left ventricular dysfunction and cardiac failure, nephrotic syndrome, renal thrombotic microangiopathy, liver failure, diarrhea, fatigue, hypertension and nausea. The AESIs for olaparib will include MDS/AML, new primary malignancy (other than MDS/AML) and pneumonitis.

For the purposes of this study all AESIs identified above will be reported for both IPs.

Any AE occurring before the start of IP or more than 30 days after discontinuation of IPs will be included in the data listings but will not be included in the summary tables of AEs.

Any AE occurring within 30 days of discontinuation of IPs (i.e., the last dose of cediranib+olaparib) or the last dose of the last IP to be discontinued, in the rare case where discontinuation dates are different will be included in the AE summaries. Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of cediranib+olaparib) will be flagged in the data listings.

3.3.4 Clinical chemistry, haematology and urinalysis

Blood and urine samples for determination of clinical chemistry, hematology, and urinalysis will be taken at the following times:

- Hematology and Biochemistry data will be taken during screening, at day 1, weekly till week 4, every 2 weeks till week 8, and every 4 weeks from week 8 onwards. If the study treatment is discontinued then a blood sample will be taken at this time and during the day 30 follow-up visit

- A urine dipstick will be conducted during screening, at day 1, and every 4 weeks from week 8 onwards. If the study treatment is discontinued then a blood sample will be taken at this time and during the day 30 follow-up visit

Additional safety samples may be collected if clinically indicated at the discretion of the Investigator. The date, time of collection and results (values, units and reference ranges) will be recorded on the appropriate eCRF.

The clinical chemistry, hematology and urinalysis will be performed at a local laboratory at or near to the Investigator site. Sample tubes and sample sizes may vary depending on laboratory method used and routine practice at the site.

The following laboratory variables will be measured:

Table 7 Laboratory safety variables

Clinical chemistry (Serum/Plasma)	Haematology (Blood)	Urinalysis (Urine)
Albumin	Hemoglobin	Protein
Alanine Transaminase (ALT)	Platelet counts	
Aspartate Transaminase (AST)	Mean Cell Volume (MCV)	
Alkaline Phosphatase (ALP)	White Blood Cell count (WBC)	
Bilirubin, total	Absolute Neutrophil Count (ANC)	
Calcium	Absolute Lymphocyte Count	
Creatinine		
Magnesium		
Potassium		
Sodium		

All values will be classified as low (below range), normal (within range) and high (above range) based on project-specific reference ranges (Appendix A). Patient with urine protein ≥3.5 g/24 hours (CTCAE Grade 3 according to CTCAE v4) or urinary protein/creatinine ratio of >3.5 should be considered as nephrotic-range proteinuria.

Changes from baseline to each visit for all patients who have a baseline laboratory test and the corresponding post-baseline laboratory test will be calculated as the post-baseline test value minus the baseline test value, where baseline is the latest non-missing value collected prior to the first dose of study drug

3.3.4.1 Disease specific tumor marker samples (CA-125)

As part of the routine blood samples, all patients will have CA-125 assessment at screening, every 8 weeks on treatment, and at study discontinuation.

A rise in CA-125 alone is not sufficient to declare progression, and discontinue IPs; progression events should be determined by radiographic evidence of progression.

Retrospectively based on the data, CA-125 progression will be defined based on the Gynaecological Cancer Intergroup (GCIG) criteria, which are well-established and followed in clinical practice (Rustin et al 2011).

A patient is considered to have CA125 progression if any of the following are met:

1. Patients with elevated CA125 pre-treatment and normalisation of CA125 must show evidence of CA125 ≥2 × upper limit of normal (ULN) on 2 occasions at least 1 week apart or

2. Patients with elevated CA125 before treatment, which never normalises, must show evidence of CA125 ≥2 × the nadir value on 2 occasions at least 1 week apart or

3. Patients with CA125 in the reference range before treatment must show evidence of CA125 ≥2 × ULN on 2 occasions at least 1 week apart.

3.3.5 Physical examination, height and weight

A complete physical examination as per standard clinical practice should be performed by the Investigator at baseline and every 4 weeks. If there are any clinically significant abnormal finding not related to disease progression, they should be recorded as an AE on the eCRF. Height will be assessed at screening only. Weight will be assessed at baseline, at every 4 weeks until treatment discontinuation and the 30 day safety follow-up visit.

Change from baseline weight will be calculated.

3.3.6 ECG

3.3.6.1 Resting 12-lead ECG

ECGs are required within 7 days prior to starting IPs, and when clinically indicated thereafter.

All ECGs should be assessed by the investigator as to whether they are clinically significantly abnormal / not clinically significantly abnormal. If there is a clinically significant abnormal finding at baseline, the Investigator should record this in the baseline eCRF. If there is a treatment emergent significant abnormal finding, the Investigator will record it as an AE on the eCRF.

3.3.7 Vital signs

Any clinically significant changes in vital signs (pulse rate, systolic and diastolic blood pressure, respiration rate, pulse oxygen saturation, and temperature) should be recorded as an AE.

3.3.7.1 Pulse and blood pressure

Patient blood pressure will be measured during routine study visits: weekly during the first 4 weeks, every 2 weeks during the next 4 weeks and every 4 weeks thereafter to ensure that blood pressure guidelines are being correctly followed.

3.3.7.2 Body temperature

Body temperature will be measured in degrees Celsius weekly during the first 4 weeks, every 2 weeks during the next 4 weeks and every 4 weeks thereafter.

3.3.7.3 Serum or urine pregnancy test

Pregnancy tests on blood or urine samples will be performed for pre-menopausal women of childbearing potential, within 7 days prior to the start of IPs. Tests will be performed by the hospital's local laboratory. If results are positive the patient is ineligible and must not be recruited into the study. In the event of a suspected pregnancy during the study, the test should be repeated. A confirmed pregnancy must be recorded in eCRF and immediately reported; the patient will be discontinued from the IPs, but should continue in the study.

3.4 Patient reported outcomes

Patient reported outcomes will be assessed by EORTC-QLQ-C30, and EORTC-QLQ-OV28, according to the schedule listed in Table 8. They will also be assessed using PRO-CTCAE.

The PROs will be collected (using an electronic diary) according to the following schedule and assessments will be mapped into scheduled visits using a windowing rule similar to other endpoints collected by visit:

Table 8 PRO Schedule

PRO Tool	Week 1-12	Week 12 - Progression	Post-progression until end of study
EORTC QLQ-C30	Day1, then every 8 weeks	Every 8 weeks	Day 30 follow-up
EORTC QLQ-OV28	Day 1, then every 8 weeks	Every 8 weeks	Day 30 follow-up
PRO-CTCAE	Day 1, then weekly	Every 4 weeks	Day 30 follow-up

An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/ items, the functional scales and the global health status scale in the QLQ-C30 and for each of the symptom scales/items in the QLQ OV28 according to the EORTC QLQ-C30 Scoring Manual and EORTC QLQ OV28 instructions. Higher scores on the global health status and functional scales indicate better health status/function. Higher scores on the symptom scales indicate greater symptom burden.

The following endpoints will be used to assess the HRQoL:

- Absolute scores and change from baseline of all subscales
- Proportions of patients with a best observed change of improved, stayed the same or deteriorated for all subscales (symptom and functioning scales). A minimum clinically meaningful change is defined as a 10 point change in score from baseline.
- Best observed change of improvement will be defined as the number (%) of patients with two assessments at least 18 days apart (i.e. 21 days allowing a visit window of 3 days) which showed a clinically meaningful improvement (a decrease from baseline score ≥10) in that symptom from baseline.
- Symptom improvement rate at 8 weeks for dyspnea, pain and constipation from QLQ-C30 and abdominal/GI symptoms from QLQ-OV28.

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30 and OV28 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 patient who has not withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. For patients that have progressed, the latest of progression and safety follow-up will be used to assess whether the patient 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 patient compliance rate is defined as: Total number of patients with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above), divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

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

3.4.1 EORTC QLQ C30

The EORTC QLQ-C30 questionnaire was developed to assess HRQoL and is a commonly used cancer-specific HRQoL tool. The EORTC QLQ-C30 comprises 30 questions designed for all cancer types. Questions can be grouped into the following scales:

- 5 multi-item functional scales (physical, role, emotional, cognitive and social)
- 3 multi-item symptom scales (fatigue, pain, nausea vomiting)
- A 2-item global HRQoL scale
- 5 single items assessing the following common cancer symptoms: dyspnea, loss of appetite, insomnia, constipation, and diarrhea
- 1 item on the financial impact of the disease

3.4.2 OV-28

The EORTC QLQ-OV28 is a module which is specific for ovarian cancer. It consists of 28 items assessing the following:

- 6 items on abdominal/GI symptoms
- 2 items on peripheral neuropathy
- 5 items on other chemotherapy side effects
- 2 items on hormonal symptoms
- 2 items on body image
- 3 items on attitudes to disease/treatment
- 4 items on sexuality
- And four other single items

3.4.3 PRO CTCAE

The PRO-CTCAE questionnaire will be used to derive patient reporting of selected CTCAE symptoms.

PRO-CTCAE is an item library of symptoms of the CTCAE experienced by patients while undergoing treatment of their cancer that has been converted to patient terms (i.e., CTCAE term

"myalgia" converted to "aching muscles"). The following 6 PRO-CTCAE AEs will be collected in the study: nausea, vomiting, diarrhoea, decreased appetite, fatigue, and dizziness.

3.5 PK variables

Blood samples for determination of cediranib and olaparib concentrations in plasma will be taken at week 4 and week 8. Samples will be collected, labelled stored and shipped as detailed in the Laboratory Manual.

Samples for determination of cediranib and olaparib concentrations in plasma will be analyzed by Covance on behalf of AstraZeneca, using appropriate bioanalytical methods.

4. ANALYSIS METHODS

An interim futility analysis of ORR (confirmed and/or unconfirmed complete responses [CR] and partial responses [PR]) will take place after 20 evaluable patients (i.e., with measurable disease [by Investigator assessment] at baseline and at least 1 dose of study drugs) have received first dose of IPs and had opportunity to be observed for at least 4 months. Assessment of measurable disease and of ORR will be based on investigator assessment. The criterion for declaring futility is based on predictive probability within a Bayesian design framework as described in Section 1.3. If less than 4 of the 20 patients have responded, the study may be stopped for futility. If 4 or more patients have responded the study will continue to the final analysis.

Evaluation of baseline data to determine measurable disease will also be performed by ICR to assess concordance with the investigator determined measurable disease status. The concordance evaluation will be done for information and is not required in order to make the decision to stop or continue the study.

Assuming the study passes the futility assessment, the data cut-off for the primary (and final) analysis of ORR (confirmed CRs + PRs only, based on ICR assessed using the EFR analysis set) will take place 8 months after the last patient has received their first dose of IPs, to allow opportunity for all patients to complete at least 4 RECIST follow-up assessments and estimate duration of response (DoR). The primary analysis for the clinical study report (CSR) will be the confirmed ORR by ICR using RECIST 1.1 criteria and supported by Investigator assessment of ORR and by other efficacy endpoints (DoR, DCR, PFS, OS and TDT), PROs and safety and tolerability data.

4.1 General principles

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. Categorical variables will be summarized by frequency counts and percentages for

each category. Unless otherwise stated, percentages will be calculated out of the full analysis set. A row or column denoted 'Missing' will be included in count tabulations where necessary to account for missing values

For continuous data the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data. For categorical data, percentages will be rounded to 1 decimal place.

4.1.1 Baseline

Baseline will be the last assessment of the variable under consideration prior to the intake of the first dose of investigational product.

4.1.2 Multiplicity adjustments

No adjustments for multiplicity will be made.

4.1.3 Visit windows

The following conventions will apply:

- Listings should display all values contributing to a time point for a patient.
- For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit based summary, when deriving a patient level statistic such as a maximum.

- The time windows should be exhaustive so that data recorded at any timepoint has the potential to be summarised. 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 CA-125 data (with 8 weeks between scheduled assessments) are:

    - Day 57, visit window 2 – 84
    - Day 113, visit window 85 – 140
    - Day 169, visit window 141 – 196, etc

Data up to and including 30 days after last dose of either IP will be considered on treatment and presented by scheduled visit except for PRO data which will be reported as follow-up visit; thereafter data will be assigned to a post-treatment visit category.

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

- For visit based summaries

  – If there is more than one value per patient within a time window then the closest value to the scheduled visit date should be summarised, or the earlier, in the event the values are equidistant from the nominal visit date. . 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 summarised if the number of observations is greater than the minimum of 20 and > 1/3 of patients dosed,

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

4.2 Analysis methods

4.2.1 Patient disposition and data sets analysed

Patient disposition will be listed and summarised and will include the number and percentage of patients:

- Screened i.e. those patients who received an E-code.
- Enrolled. i.e. those who were confirmed as eligible for the study and signed informed consent.
- Treated.
- Patients ongoing study treatment at the data cut-off.
- Included in each analysis set (full, evaluable for response, PK).

In addition, the number and percentage of patients discontinuing treatment and reasons for discontinuing will be presented.

4.2.2 Protocol deviations

All important protocol deviations will be listed and summarised for the full analysis set. All important protocol deviations will be defined by the study team and identified before database lock.

4.2.3 Demographic and other baseline characteristics

Demographic and baseline patient characteristics will be summarised for the full analysis set. If the sizes of other analysis sets are different from the FAS then the demographic and baseline characteristics tables will be repeated for the relevant analysis sets. Standard descriptive statistics will be presented for the continuous variables of:

- Age (years).
- Weight (kg).
- Height (cm).
- Body mass index (kg/m^2) [calculated as ($weight/height^2$) where weight is in kg and height is in m].

The total counts and percentages of patients will be presented for the categorical variables of:

- Age group (years) (grouped as <50, >=50-<65,≥= 65-<75, >=75).

- Race.
- Ethnic group.
- Stage of disease.
- Germline BRCA status.
- 1st sensitivity to first platinum line calculated as <6, 6-<12, >=12 months from date of completing first-line platinum-based chemotherapy to date of progression
- 2nd sensitivity to the last platinum line calculated as <6, 6-<12, >=12 months from date of last receipt of platinum-based chemotherapy to date of progression
- Ascites at baseline.
- Histology
- ECOG performance status at baseline
- gBRCA classification at study entry
- suspected sBRCAm status
- Patients carrying a tumor deleterious or suspected deleterious variant in homologous recombination repair (HRR) associated genes

4.2.4 Medical history

Disease related medical history and relevant surgical history will be coded using MedDRA (version 17.0 or later). The number and percentage of patients with any disease related medical history or relevant surgical history will be summarised for the full analysis set by system organ class (SOC) and preferred term (PT) for each cohort and overall.

4.2.5 Concomitant medications (prior, during, post treatment with IP)

Medications received prior to, concomitantly, or post study IPs will be coded using the AstraZeneca Drug Dictionary Anatomical Therapeutic Chemical (ATC) Classification codes.

The period of study treatment is defined as the period in which a patient is receiving either of the IPs. Olaparib and cediranib should be initiated and stopped at the same time, however if this doesn't occur then the date of the first and last dose of either IP will be used to define the period of study treatment.

Prior medications, concomitant and post-treatment medications are defined as follows:

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

All medications other than anti-cancer medications will be collected until 30 days after the last dose of study treatment. Anti-cancer medications will be collected as part of survival follow-up until the end of the study.

A summary of disallowed medications taken by the patients during the study treatment will be produced.

4.2.5.1 Previous anti-cancer treatment modalities

All prior anti-cancer therapies will be summarised for the full analysis set.

Treatments will be grouped by ATC code and summarised by line of therapy individually and grouped (3, >=4). In addition a similar summary will be presented where treatments are grouped as platinum and non-platinum based.

The time from most recent disease progression until first dose of first IP will be summarised. Where partial dates are given and only day is missing, the last day of the month will be imputed, otherwise if month or year is missing then the entire date will be regarded as missing and not included in the summary.

In addition all prior anti-cancer therapies will also be summarised by the medically reviewed categories:

a. Platinum containing regimen (excluding bevacizumab or PARPi containing)

b. Platinum containing regimen with PARPi maintenance

c. Platinum in combination with bevacizumab

d. Other bevacizumab containing regimen

e. PARPi monotherapy regimen

f. Any other PARPi containing regimen

g. Any other chemotherapy regimen (including immunotherapy; excluding platinum, bevacizumab, or PARPi containing)

h. Hormonal agent

i. Investigational agent

4.2.5.2 Concomitant anti-cancer medications and radiotherapy

Concomitant anti-cancer medications during treatment are not allowed by the protocol.

4.2.5.3 Post treatment anti-cancer medications and radiotherapy

All post treatment anti-cancer medications and surgical procedures will be summarised for the full analysis set.

In addition all post treatment anti-cancer medications will be summarised by the medically reviewed categories in section 4.2.5.1.

4.2.6 Treatment exposure

Exposure to Cediranib, Olaparib, and the combination of Cediranib and Olaparib will each be summarised separately. Two criteria will be used to summarise and plot the combination of Cediranib and Olaparib:

1. The period of treatment will include the first date either IP was initiated till the last date either IP was discontinued. An interruption to both IPs will be considered an interruption for this summary.
2. The period of treatment will include the first date the combination was taken till the last date the combination was taken. An interruption to either IP will be considered an interruption for this summary.

An overall summary of total treatment duration and actual treatment duration will be presented. The Relative dose intensity (RDI) and cumulative exposure over time will also be included.

A summary of duration of therapy at starting dose and each level of dose reduction will be created for each of cediranib and olaparib.

The number of patients who ended on each possible dose combination will be summarised along with the number of patients who took each combination at least once.

Dose reductions will be summarised using a bar chart where each individual patient is represented by two bars, one for cediranib and one for olaparib. The bars represent the dose level over time, including any interruptions, and indicate whether a reduction has taken place. The plot

will also show whether a discontinuation or discontinuation due to death has occurred, and whether treatment was ongoing at data cut-off.

Treatment interruptions, dose reductions, and the duration of interruptions (any) will be summarised for both cediranib and olaparib.

4.2.7 Efficacy

Tables, figures and listings will be summarised according to the analysis sets for the relevant variable as detailed in Table .

4.2.7.1 Analysis of the primary efficacy variable ORR

The primary analysis of ORR will be based on ICR assessments of tumor response using the EFR analysis set.

The number of responders (confirmed CR or PR) will be modelled using a binomial distribution with a neutral Beta(1/3,1/3) prior distribution on the rate.

The mean and median of the posterior distribution of ORR will be presented, along with the standard deviation and a 95% credible interval around the mean (based on the highest posterior density [Lee, 1997]). The observed ORR by ICR will also be presented with 95% exact (Clopper Pearson) confidence intervals.

The posterior probabilities of the ORR being larger than 15, 20 and 24% will be also computed. To further characterize the response rate, the posterior probability of showing unacceptable (<15%), low (>15-20%), moderate (>20-24%) or substantial (>24%) efficacy will be presented.

Summaries of the number of patients with best response in each of the follow categories will be presented: CR, PR, SD, PD and Non-Evaluable (NE). The number of patients with an unconfirmed CR or PR will be presented as a sub-category of SD. For the interim analysis the CR (PR) category will be split into confirmed CR (PR) and unconfirmed CR (PR).

4.2.7.2 Analysis of the secondary efficacy variable(s)

DoR

Duration of response in responding patients will be summarized and the number (%) of responding patients with a duration of response >3; >6; >9; >12 months will be presented. A Kaplan Meier plot and median duration of response with 95% CI (calculated from the Kaplan-Meier plot) will be presented.

Further plots showing the duration of response in responding patients will be presented as appropriate. For example, a plot showing a horizontal bar for each patient for the duration of response which are ordered from longest to shortest response. Any patients who are censored will have an arrow at the end of the bar to indicate response is on-going.

The time to onset of response from first dose will also be summarised using the median calculated using the Kaplan-Meier method. The number and percentage of responding patients with an onset by 6, 12, 18 and 24 weeks will be summarised.

DCR

DCR will be presented together with 95% exact (Clopper Pearson) confidence intervals.

PFS

The progression status at the time of the PFS analysis will be summarised, including the number and percentage of patients who progressed due to RECIST progression or due to death, those who did not progress due to being progression free or due to being lost to follow up, and those who have withdrawn consent. The number of patients censored on Day 1 will also be reported.

PFS will be displayed using a Kaplan Meier plot. The median time to event (calculated from the Kaplan-Meier plot), and the proportion of patients without an event at 3, 6, 12, and 18 months will be summarized.

OS

A summary of the number and percentage of patients who have died, are still in survival follow-up, are lost to follow-up and have withdrawn consent will be presented.

When there is sufficient data (>20 events), OS will be displayed using a Kaplan-Meier plot. The number of events, median time to event (calculated from the Kaplan-Meier plot), and the proportion of patients without an event at 3, 6, 12 and 18 months will be summarized.

Time to discontinuation of study treatment or death (TDT)

TDT will be displayed using a Kaplan-Meier plot. The number of events, median time to event (calculated from the Kaplan-Meier plot), and the proportion of patients without an event at 3, 6, 12, and 18 months will be summarized.

4.2.7.3 Sensitivity analysis

ORR, DCR, DoR and PFS will be re-analyzed at the time of the primary analysis using the investigator assessment of tumor response instead of ICR assessment, as sensitivity analyses.

ORR and exact 95% CI (Clopper Pearson) will be presented for the following:

- ORR by investigator assessment. The numerator (number of confirmed responses) will be based on RECIST as per investigator assessment and the denominator will be based on the full analysis set

- ORR for both confirmed and unconfirmed responses based on ICR assessment. The numerator will include unconfirmed responses in addition to confirmed responses and the denominator will be based on the EFR set.

Sensitivity analyses of PFS, DoR and DCR using the investigators assessment of RECIST will be performed in an analogous manner to those using the ICR described in Section 4.2.7.2 above.

In addition, ORR and BOR using the ICR will be summarised for the FAS.

For patients where measurable disease at baseline is not confirmed by the ICR and who therefore appear in the sensitivity analysis using FAS, but not the primary analysis set (evaluable for response), a detailed description of the reasons for being considered not evaluable will be described in the CSR.

A further comparison of each type of overall objective response (CR, PR, SD, PD, NE or without measureable disease at baseline) between investigator and ICR will be presented using the full analysis set. Details for the patients identified as responders by ICR and not by the investigator, as well as by the investigator and not by ICR will be given.

The concordance between response as assessed by ICR and by investigator with the full analysis set will be summarised by cross-tabulation of the number and percentage of patients with:

- Response declared by investigator and ICR
- Response declared by investigator but not ICR
- Response declared by ICR but not investigator
- No response declared by both investigator and ICR

The overall concordance rate will be calculated as the total number of agreements between investigator and ICR divided by the total sample size.

4.2.7.4 Subgroup analysis

The analyses of ORR, DCR, DoR, PFS at the time of the primary analyses and OS at the time of the final analysis will be repeated in the following subgroups of patients:

1. patients carrying a somatic deleterious or suspected deleterious variants in either of the *BRCA* genes (*BRCA1* or *BRCA2*) identified with current and potential future tumor based *BRCA* mutation assays

2. patients who don't carry a somatic deleterious or suspected deleterious variants in either of the *BRCA* genes (*BRCA1* or *BRCA2*) identified with current and potential future tumor based *BRCA* mutation assays

3. patients identified as having a deleterious or suspected deleterious variants in HRR-associated genes identified with current and potential future HRR gene mutation assays

4. patients identified as not having a deleterious or suspected deleterious variants in HRR-associated genes identified with current and potential future HRR gene mutation assays

5. Patients in either group 1 or 3, (if the number in group 1 or 3 separately is small [less than 10])

6. Patients not in group 1 or 3 (if the number in group 1 or 3 separately is small [less than 10])

If a subgroup has <10 patients then the subgroup will not be analyzed and a listing including response status will be produced instead.

4.2.7.5 Exploratory endpoints

Biomarkers

The following biomarkers may be analyzed using plots and descriptive summary statistics.

- Biomarker of potential HRDs, including genomic scarring, *BRCA1* methylation, and other genes that might modify HRD.
- Biomarkers of gene expression signature for angioimmune/angio/immune/metabolic categorization
- Angiogenic biomarkers such as VEGF-A, C, D; sVEGFR2, Tie2, Ang1, 2; leptin; IL6

The exploratory biomarker analyses will be reported separately to the CSR.

4.2.8 PROs

Patients without a baseline PRO assessment will not be included in summaries of PRO endpoints.

4.2.8.1 PRO-CTCAE

Data from the PRO-CTCAE will be summarised for those patients with evaluable data. The number (%) of patients with each level of response for each PRO-CTCAE item at baseline and over time will be summarised.

4.2.8.2 EORTC QLQ- C30 and EORTC QLQ- OV28

The EORTC-QLQ-C30 and EORTC-QLQ-OV28 questionnaires will be summarized. Descriptive statistics of absolute scores and change from baseline of all subscales will be presented and plots will be generated as appropriate to interpret changes over time.

Proportions of patients who had a best observed change of improved, stayed the same or deteriorated whilst on treatment (excluding follow-up) for all subscales (symptom and functioning scales) will be presented. A 10 point change in score will be used as cut-off point. Improved will be defined as the number (%) of patients with two assessments at least 18 days apart (i.e. 21 days allowing a visit window of 3 days) which showed a clinically meaningful improvement (a decrease from baseline score ≥10) in that symptom from baseline. The quantity and impact of missing data will be evaluated. If warranted, some investigation of HRQoL data by whether patients are still taking study drug will be performed.

Symptom Improvement Rate at 8 weeks

For dyspnea, pain and constipation from QLQ-C30 and abdominal/GI symptoms from QLQ-OV28, the symptom improvement rate at 8 weeks (follow-up data will be excluded) will be defined as the number (%) of patients showing a clinically meaningful improvement (a decrease from baseline score >=10; Osoba et al 2005) in that symptom from baseline. The denominator will consist of a subset of the FAS population who have a baseline symptom score >=10. A 95% confidence interval for the symptom improvement rate will be constructed using the Clopper-Pearson method.

4.2.9 Safety and tolerability

Safety and tolerability data will be presented using summaries and descriptive statistics.

The safety analysis set will be used for all safety and tolerability tables, figures and listings except where expressly noted.

For change from baseline summaries for vital signs, laboratory data, LVEF, the baseline value will be the latest result obtained prior to the start of investigational product.

The denominator used in laboratory summaries will only include evaluable patients, in other words those who had sufficient data to have the possibility of an abnormality. For example:

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

The denominator in vital signs data should include only those patients with recorded data.

REACT
To enable continuing assessment of the safety profile of the cediranib and olaparib combination, and toxicity management for study patients, a data visualisation tool (e.g., REACT) will be used. No efficacy data (RECIST response; progression status) will be included. The "CONCERTO principles for use of REACT" document will describe appropriate access to and use of the tool.

AEs

All AE data will be listed individually by patient. Any AE occurring before treatment with either IPs will be included in the data listings and flagged but will not be included in the summary tables of AEs.

An overview table will summarise the number and percentage of patients with at least one of the following AEs where patients with more than one AE in a particular category are counted only once in that category:

- Any AE
- Any AE causally related to either IP, listed separately
- Any AE of CTCAE grade 3 or higher
- Any AE of CTCAE grade 3 or higher, causally related to either IP, listed separately
- Any AE with outcome of death
- Any AE with outcome = death, causally related to either IP, listed separately
- Any SAE (including events with outcome=death)
- Any SAE (including events with outcome=death), causally related to either IP, listed separately
- Any SAE leading to discontinuation of either IP, listed separately Any AE leading to discontinuation of either IP, listed separately
- Any AE leading to discontinuation of either IP, listed separately, causally related to either IP, listed separately
- Any AE leading to dose interruption
- Any AE leading to dose reduction
- Any AE leading to dose modification (interruption and/or reduction)
- Any other significant AE (if applicable)

The number and percentage patients reporting each AE will be summarised, by SOC (sorted alphabetically) and PT (sorted by descending overall total) The following summaries will be produced:

- AEs by SOC and PT.
- AEs by maximum reported CTCAE grade, by SOC and PT.
- AEs of CTCAE grade 3 or higher, by SOC and PT.
- AEs causally related by IP, by SOC and PT.
- AEs leading to hospitalisation, by SOC and PT.
- AEs leading to dose interruption, by SOC and PT.
- AEs leading to dose reduction, by SOC and PT.
- AEs with outcome of death, by SOC and PT.
- AEs with outcome of death, causally related by IP, by SOC and PT.
- SAEs by SOC and PT.
- SAEs causally related by IP, by SOC and PT.

- AEs leading to discontinuation of either IP (listed separately), by SOC and PT.
- OAEs by SOC and PT (if applicable).

In the above summaries, patients with more than one AE within a particular SOC are counted only once for that SOC. Similarly, patients with more than one AE within a particular PT are counted only once that PT. For summaries by maximum reported CTCAE grade with multiple AEs within a particular SOC or PT will be counted under the category of their most severe AE within that SOC or PT.

A summary of deaths will be provided, detailing deaths due to disease only, deaths where AE outcome is death only or both.

AEs of special interest

The AESIs for Cediranib are as follows: PRES, arterial thromboembolic events, and the Important Potential Risks: symptomatic left ventricular dysfunction and cardiac failure; nephrotic syndrome; renal thrombotic microangiopathy, liver failure, and (per eCRF) diarrhea, fatigue, hypertension and nausea. The AESIs for Olaparib are as follows: potential risks of MDS/AML, new primary malignancy (other than MDS/AML) and pneumonitis.

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

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

- At least one AESI presented by outcome
- At least one AESI causally related to either study medication separately
- At least one AESI leading to discontinuation of either study medication separately

A summary of total duration (days) of AESI will be provided for events which have an end date and this will be supported by summaries of ongoing AESIs at death and, separately, at data cut-off.

To reflect the particular interest in diarrhoea, fatigue, hypertension, and nausea, life tables and prevalence plots will be produced for these AEs.

In addition as all CTCAE grade changes for diarrhoea, fatigue, hypertension, and nausea will be recorded, a plot will be generated for each AE which shows the AE grade over time for each patient.

Laboratory evaluations

All laboratory data recorded in the eCRF will be listed. If any additional analyses to those in Table 7 are also recorded then these will be listed only. Out-of-reference range values will be flagged as high (H) or low (L) in the listings.

For clinical chemistry and haematology, numerical summaries will be generated which include the mean, standard deviation, median, minimum value, maximum value, and lower and upper quartiles at each visit in which laboratory values are taken. Plots of both the maximum post-baseline ALT and AST versus the maximum post-baseline total bilirubin, expressed as multiples of their upper limit of reference range will be produced.

As a primary interpretation of potential drug induced liver injury, outputs related to patients with liver function tests (LFTs) that meet biochemical criteria for Hy's Law and with elevations of LFTs will be produced. Biochemical criteria for Hy's Law are defined as any situation where a study patient has an increase in both AST or ALT $\geq 3\times$ULN and TBL $\geq 2\times$ULN, irrespective of ALP, at any point during the study. The elevations do not have to occur at the same time or within a specified time frame.

For all patients who meet the criteria for potential Hy's Law, a plot of AST, ALT, ALP and total bilirubin over time will be produced.

The proportion of patients who met GCIG criteria for CA-125 progression (any of the criteria, and each criterion individually) will be summarised by visit. The time to CA-125 progression will be presented using a Kaplan Meier plot and the median time to CA-125 progression (from the Kaplan Meier) calculated. Patients without CA-125 progression will be censored at their latest lab evaluation visit.

Vital signs

Absolute values and change from baseline will summarised, overall and by visit. The minimum, maximum, and most recent measurement will also be listed. Plots of observed and changes in vital sign variables data may be generated to illustrate the data.

Physical examination

Abnormalities identified from physical examination will be captured as AEs.

4.2.10 PK

PK data will be analysed and reported separately from the CSR, along with other studies as part of a pooled population PK analysis.

5. INTERIM ANALYSES

An interim futility analysis will take place after 20 evaluable (by Investigator assessment) patients are recruited and have had opportunity to be observed for at least 4 months, and datasets required are clean. The primary endpoint of the interim futility analysis will be the number of responders using the investigators assessment of RECIST 1.1. The study may be stopped for futility if there are less than 4 responders out of the first 20 patients in which case the predictive probability of observing 30 responders or more at the end of trial is ≤4.6%.

A summary of the ORR will be presented for the interim population, along with a two-way table illustrating the concordance of the baseline scans between local assessment and ICR with respect to whether there was measurable disease at baseline.

For the purposes of the interim analyses a responder is defined as a patient with measurable disease at baseline, with at least one confirmed or unconfirmed visit response of CR or PR.

No formal outputs beyond assessment of ORR, and concordance of the population with ICR, are planned to be generated at the interim analysis.

6. CHANGES OF ANALYSIS FROM PROTOCOL

CSP version 5 mentions “For dyspnea, pain and constipation from QLQ-C30 and abdominal/GI symptoms from QLQ-OV28, the symptom improvement rate at 8 weeks will be defined as the number (%) of patients showing a clinically meaningful improvement (a decrease from baseline score >10)”. This is an error and the decrease from baseline score should be >=10.

7. REFERENCES

Fisch et al 2015. Bayesian Design of Proof-of-Concept Trials. Therapeutic Innovation & Regulatory Science, 2015, Volume 49(1) 155-162

Kerman, J. Neutral noninformative and informative conjugate beta and gamma prior distributions. Electronic Journal of Statistics, 2011, Volume 5; 1450-1470.

Aghajanian C, Goff B, Nycum LR, Wang Y, Husain A, Blank S. Independent radiologic review: bevacizumab in combination with gemcitabine and carboplatin in recurrent ovarian cancer. Gynecol Oncol. 2014; 133(1):105-10.

Rustin, G, et al. Definitions for Response and Progression in Ovarian Cancer Clinical Trials Incorporating RECIST 1.1 and CA 125 Agreed by the Gynecological Cancer Intergroup (GCIG). International Journal of Gynecological Cancer, 2011; 21:419-423.

Osoba D, Bezjak A, Brundage M, Zee B, Tu D, Pater J; Quality of Life Committee of the NCIC CTG: Analysis and interpretation of health-related quality-of life data from clinical trials: Basic approach of The National Cancer Institute of Canada Clinical Trials Group. Eur J Cancer 2005; 41:280-287

Eisenhauer, E. et al. New response evaluation criteria in solid tumours: revised RECIST guideline. European Journal of Cancer, 2009; 45(2):228-47

8. APPENDIX

8.1 Appendix A

AZ project-specific normal range for laboratory tests

Test	Gender	SI Unit	Lower Limit of Normal Range	Upper Limit of Normal Range
Albumin	M and F	g/L	32	56
Alkaline phosphatase	M and F	U/L	20	130
ALT	M and F	U/L	0	50
AST	M and F	U/L	0	45
CA125	M and F	U/mL	0	55
Calcium	M and F	mmol/L	2.2	2.6
Creatinine	M and F	umol/L	53	120
Magnesium	M and F	mmol/L	0.70	1.00
Phosphate	M and F	mmol/L	0.8	1.4
Potassium	M and F	mmol/L	3.5	5.0
Sodium	M and F	mmol/L	136	145
T4, free	M and F	pmol/L	10	30
Total Bilirubin	M and F	umol/L	2	21
Total Protein	M and F	g/L	60	80
TSH	M and F	mU/L	0.5	5.0
Urea	M and F	mmol/L	2.5	6.7
APTT	M and F	sec	25	35
Haemoglobin	M	g/L	135	180
Haemoglobin	F	g/L	120	160
Leucocytes	M and F	10**9/L	4.4	11.3
Neutrophils	M and F	10**9/L	1.8	7.8
Platelets	M and F	10**9/L	150	450
Prothrombin time	M and F	sec	10.5	14.5
Protein (24hr)	M and F	mg/24 h	0	199

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: d8488c00001-sap-ed-2		
Document Title:	Statistical Analysis Plan Edition 2	
Document ID:	CCI [redacted]	
Version Label:	2.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
20-Sep-2019 15:03 UTC	PPD [redacted]	Author Approval
20-Sep-2019 14:52 UTC	PPD [redacted]	Content Approval

Notes: CCI [redacted]