A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

ClinicalTrials.gov Identifier: NCT03106987

Statistical Analysis Plan Amendment No. 3: Final 4.0, 16 Apr 2021

Statistical Analysis Plan

Drug Substance	Olaparib, AZD2281, KU-0059436
Study Code	D0816C00014
ENGOT Number	ENGOT-ov38/OReO
Edition Number	4.0
Date	16-April-2021

A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

Study Statistician

PPD

Date

A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

Global Product Statistician

PPD Date

A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

ENGOT Statistician

PPD Date

A Phase IIIb, Randomised, Double-blind, Placebo-controlled, Multicentre Study of Olaparib Maintenance Retreatment in Patients with Epithelial Ovarian Cancer Previously Treated With a PARPi and Responding to Repeat Platinum Chemotherapy (OReO)

ENGOT Statistician

PPD

Date

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AML	Acute myeloid leukaemia
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
bd	Twice a day
BMI	Body mass index
BRCA1 and BRCA2	Breast cancer susceptibility genes
CA-125	Cancer antigen 125
COVID-19	Coronavirus Disease 2019
CR	Complete response
CSP	Clinical study protocol
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Event
CCI	[redacted]
DCO	Data cut-off
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EOC	Epithelial ovarian cancer
EWB	Emotional well-being
FACIT	Functional Assessment of Chronic Illness Therapy
FACT-G	Functional Assessment of Cancer Therapy – General
FACT-O	Functional Assessment of Cancer Therapy – Ovarian
FAS	Full analysis set
FPI	First patient in
FWB	Functional well-being
gBRCA1/2	Germline *BRCA1/2*
GCIG	Gynaecologic Cancer InterGroup
GGT	Gamma-glutamyl transpeptidase
HR	Hazard ratio
HRD	Homologous recombination deficiency
HRQoL	Health-related quality of life
HRR	Homologous recombination repair
IDMC	Independent data monitoring committee
ITT	Intent-to-treat
IVRS	Interactive voice response system
K-M	Kaplan-Meier
LD	Longest diameter
MAR	(Data being) missing at random
MCAR	(Data being) missing completely at random

Abbreviation or special term	Explanation
MCMC	Markov Chain Monte Carlo
MDS	Myelodysplastic syndrome
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimally important difference
MMRM	Mixed model for repeated measures
mPFS	Median progression-free survival
MRI	Magnetic resonance imaging
NED	No evidence of disease
NTL	Non-target lesion
OCS	Ovarian cancer subscale
OS	Overall survival
PARPi	Polyadenosine 5diphosphoribose [poly (ADP ribose)] polymerisation inhibitor
PD	Progressive disease
PFS	Progression-free survival
PID	Percentage intended dose
p.o.	Oral
PR	Partial response
PRO	Patient reported outcome
PT	Preferred term
PWB	Physical well-being
QoL	Quality of life
RECIST	Response Evaluation Criteria In Solid Tumours
RDI	Relative dose intensity
SAE	Serious adverse event
sBRCA1/2	Somatic *BRCA1/2*
SD	Stable disease
SOC	System organ class
SOLO2	A previous AstraZeneca sponsored confirmatory Phase III trial of Olaparib in patients with ovarian cancer
SWB	Social/family well-being
TDT	Time to study treatment discontinuation or death
TEAE	Treatment-emergent adverse event
TFST	Time to first subsequent treatment commencement or death
TL	Target lesion
TOI	Trial outcome index
TSST	Time to second subsequent treatment commencement or death
ULN	Upper limit of normal
WHO	World Health Organisation

AMENDMENT HISTORY

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
	20 March 2017 (v1)	First approved version.	Y (V1)	-
Other	22 May 2017 (v2)	Second approved version.	Y (V1)	-
Other	05 March 2019 (v3)	1) Modification of inclusion criterion #5 to align with SOLO2 (ENGOT Ov-21; NCT01874353), to allow patients to have at least a partial radiological response or not have a rising CA-125 following optimal debulking surgery with no measurable disease. 2) Updated statistical design assumptions. 3) Revised sample size and power justification.	Y (V3)	Updated following CSP update to Version 3.
Derivation of primary or secondary endpoints	16 April 2021 (v4)	Updates to text regarding time to earliest progression by RECIST or CA-125 or death	Y(V3)	Clarified text regarding derivation of this endpoint and application of the 'two missing visits' rule to align with the AZ Oncology TA SAP.

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Derivation of primary or secondary endpoints	16 April 2021 (v4)	Removed subscales from PRO analyses	Y(V3)	Text regarding PRO analyses was inconsistent across the SAP. Updated to clarify that only TOI will be formally analysed.
Derivation of primary or secondary endpoints	16 April 2021 (v4)	Additions/amendments to text regarding RECIST derivations	Y (V3)	Updated to be consistent with the AZ Oncology TA SAP.
Data presentations	16 April 2021 (v4)	Added summary for AEs leading to dose modification	NA	Was previously missed, added for completeness.
Data presentations	16 April 2021 (v4)	New AE summary for AE groups added	NA	AE groups of interest added for consistency with other Olaparib studies
Data presentations	16 April 2021 (v4)	Added summary of stratification factors based on actual stratification values	NA	For completeness, so any differences between randomized stratification and actual stratification values are clear.
Statistical analysis method for primary or secondary endpoints	16 April 2021 (v4)	PRO analysis set amended.	N	For clarity and to make the PRO analysis set a subset of the FAS.

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Statistical analysis method for primary or secondary endpoints	16 April 2021 (v4)	Added a sensitivity analysis for cohort adjustment.	NA	To allow the assessment of the impact of patients who have been found to be BRCA+ve after randomisation to the study as BRCA-ve.
Statistical analysis method for primary or secondary endpoints	16 April 2021 (v4)	Analysis method for PRO endpoint clarified.	Y(V3)	Missing details regarding specifics of the mixed modelling approach were added to the SAP text.
Other	16 April 2021 (v4)	Update study information	Y (V3)	To make SAP consistent with updated study documents.
Other	16 April 2021 (v4)	Correction of typos and amendments to make terms consistent.	Y (V3)	For clarity.
Other	16 April 2021 (v4)	Rewording of secondary endpoints	Y(V3)	For completeness, so that information regarding both the objective and measure are included.
Other	16 April 2021 (v4)	Added text to state exploratory objectives will not be assessed as part of this SAP	Y (V3)	For clarity.
Other	16 April 2021 (v4)	Important protocol deviations list amended	Y (V3)	Updated to be consistent with other Olaparib studies.

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Other	16 April 2021 (v4)	Added text to clarify that important protocol deviations occurring due to the COVID-19 pandemic will be identified and presented.	N/A	For completeness as protocol deviations are classified separately if they occur due to the COVID-19 pandemic.
Other	16 April 2021 (v4)	Text regarding identification of protocol deviations amended.	Y(V3)	Text updated to be consistent with how protocol deviations are identified and assessed during the study.
Other	16 April 2021 (v4)	Text regarding the timing of the primary analysis for PFS and the OS analysis amended.	Y(V4)	Text updated to be consistent with the updated study protocol.
Statistical analysis method for primary or secondary endpoints	16 April 2021 (v4)	Text added to clarify what a definitive deterioration is for TOI score and the related PRO deterioration-free survival analysis.	N	For clarity.
Other	16 April 2021 (v4)	CCI [redacted]	Y(V4)	New exploratory analysis of interest identified.
Other	16 April 2021 (v4)	Text added regarding the addition of the COVID-19 listings.	N/A	Inclusion of mandatory COVID-19 listings.

Category: Change refers to	Date	Description of change	In line with CSP? Y (version) / N / NA	Rationale
Other	16 April 2021 (v4)	Text added regarding the time points when subsequent treatment or death free survival and treatment discontinuation or death free survival will be evaluated.	N/A	This additional information is of interest.

1. STUDY DETAILS

1.1 Study objectives

Primary:

To determine the efficacy by progression free survival (PFS) (using investigator assessed scans according to Response Evaluation Criteria In Solid Tumours (RECIST v1.1) of Olaparib maintenance retreatment compared to matching placebo in patients with epithelial ovarian cancer previously treated with a Polyadenosine 5'diphosphoribose [poly (ADP ribose)] polymerisation inhibitor (PARPi) and responding to repeat platinum chemotherapy.

Secondary:

1. To determine the efficacy of Olaparib maintenance retreatment compared to matching placebo in patients with epithelial ovarian cancer previously treated with PARPi and responding to repeat platinum chemotherapy by assessment of overall survival (OS).
2. To determine the efficacy of Olaparib maintenance retreatment compared to matching placebo in patients with epithelial ovarian cancer previously treated with PARPi and responding to repeat platinum chemotherapy by assessment of time to progression by Gynaecologic Cancer Intergroup (GCIG) criteria which is defined as the time from randomisation to the earliest of progression by RECIST 1.1 or Cancer Antigen-125 (CA-125), or death (by any cause in the absence of progression).
3. To determine the efficacy of Olaparib maintenance retreatment compared to matching placebo in patients with epithelial ovarian cancer previously treated with PARPi and responding to repeat platinum chemotherapy by assessment of the use of subsequent therapies and study treatment discontinuation. This will be measured by examining time from randomisation to first subsequent therapy or death (TFST), time from randomisation to second subsequent therapy or death (TSST) and time from randomisation to study treatment discontinuation or death if this occurs before discontinuation of study treatment (TDT).
4. To compare the effects of Olaparib maintenance retreatment compared to matching placebo on Health-related Quality of Life (HRQoL) as assessed by the trial outcome index (TOI) of the Functional Assessment of Cancer Therapy – Ovarian (FACT-O) in patients with epithelial ovarian cancer previously treated with PARPi and responding to repeat platinum chemotherapy. This will be measured by examining change from baseline, time to deterioration and proportion improved in TOI.

Safety:

1. To assess the safety and tolerability of Olaparib maintenance retreatment in patients with epithelial ovarian cancer previously treated with PARPi and responding to repeat platinum chemotherapy.

Exploratory:

CCI

[redacted]

1.2 Study design

This is a Phase IIIb, randomised, double-blind, placebo-controlled, multicentre study to assess the efficacy and tolerability of Olaparib retreatment, versus matching placebo, in non-mucinous epithelial ovarian cancer (EOC) patients (including patients with primary peritoneal and/or fallopian tube cancer). To be eligible, patients must have received maintenance therapy with a PARPi, and must have had at least a partial response to their most recent course of platinum-based chemotherapy, or may have no evidence of disease (if optimal cytoreductive surgery was conducted prior to chemotherapy), and no evidence of a rising CA-125. All patients must have a confirmed genetic status for *BRCA1/2*.

Patients will be randomised into one of two cohorts depending on their known *BRCA* status (see Figure 1):

- The first cohort will enrol patients with confirmed *BRCA1/2* (+ve) status (somatic, *sBRCA1/2*, or germline, *gBRCA1/2*)
- The second cohort will enrol patients who are known *gBRCA1/2* (-ve) and may include some patients who have an undetected *sBRCA1/2* mutation.

Within each cohort, patients will be randomised by prospective allocation in a 2:1 ratio (Olaparib: matching placebo) to the treatments as specified below:

- Olaparib tablets (oral [p.o.]), 300 mg twice a day (bd) (except where this dose and formulation was previously not tolerated; see Section 6.7 of the protocol)

- placebo tablets to match, p.o. bd

Randomisation, allocated by the interactive voice response system (IVRS) from a block randomisation schedule, will be stratified by:

- Use of prior bevacizumab (yes versus no)
- Number of prior regimens of platinum-containing chemotherapy (≤3 regimens versus ≥4 regimens)

The minimum periods for which patients must have taken maintenance PARPi without progression to be eligible for this retreatment study are:

- For the *BRCA1/2* (+ve) cohort, the duration of exposure must have been ≥18 months following a first line of chemotherapy or ≥12 months following a second line or subsequent line of chemotherapy
- For the *BRCA1/2* (-ve) cohort, the duration of exposure must have been ≥12 months following a first line of chemotherapy or ≥6 months following a second line or subsequent line of chemotherapy

A month is defined as being from the start date to the same date in the next month (e.g. 01 January to 01 July is 6 months). These periods are of continuous PARPi administration and are measured from the date of the first dose to the date of the last dose of PARPi. During this period, no new anticancer treatment for disease progression should have been administered, but patients may have had short breaks (maximum 14 days on any occasion), for example, for holidays, control of toxicity, non-cancer treatments, etc.

Patients may have, but do not need to have, a specific confirmatory homologous recombination repair (HRR) or homologous recombination deficiency (HRD) test. Patients with *BRCA* variants of unknown significance will be regarded as *BRCA1/2* (-ve) for the purposes of this study.

Patients should also have received subsequent platinum-based chemotherapy, excluding bevacizumab, following progression during or after prior PARPi therapy and remain platinum sensitive. For the purposes of this study, platinum sensitivity means that the patient had a RECIST version 1.1 partial or complete response (as determined by the Investigator) to the most recent line of platinum-based chemotherapy or may have no evidence of disease (if optimal cytoreductive surgery was conducted prior to chemotherapy), and no evidence of a rising CA-125.

In addition, patients must be randomised into the study within 8 weeks of their last dose of platinum-based chemotherapy (last dose is the day of the last infusion). If the last platinum-based chemotherapy is taken in conjunction with another chemotherapy treatment (as part of a combined regimen) then the patient should be randomised within 8 weeks of their last dose of the combined regimen.

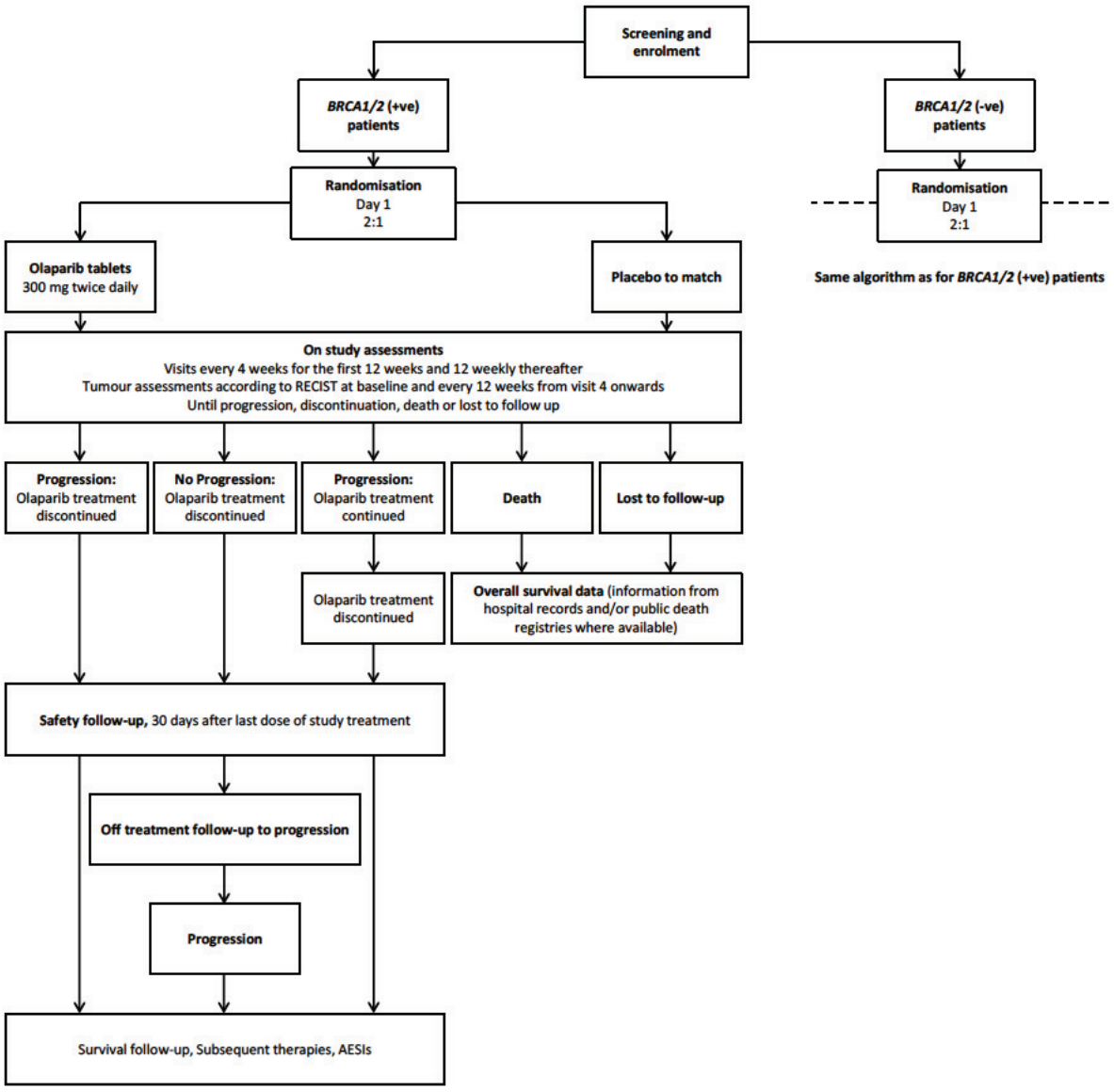
Investigators will be required to provide tumour assessment information using RECIST v1.1 and HRQoL questionnaires at baseline (a maximum of 28 days prior to randomisation). Following randomisation, patients in all study arms must have tumour assessments every 12 weeks (±7 days) until objective disease progression. More information is provided in Section 5.1 of the protocol.

Patients should continue to receive study treatment until objective radiological disease progression as per RECIST v1.1 or as long as in the Investigator's opinion they are benefiting from treatment and they do not meet any other discontinuation criteria. Such patients should additionally:

- Not have symptoms and signs (including worsening of laboratory values) indicating unequivocal progression of disease
- Not have a decline in Eastern Cooperative Oncology Group (ECOG) performance status that can be attributed to disease progression
- Not have tumour progression at critical anatomical sites that cannot be readily managed and stabilized by protocol allowed medical interventions
- Be provided information deferring any standard treatment options that may exist in favour of continuing investigational product treatment at the time of initial progression.

Once a patient has stopped treatment, a 30 days post last study drug visit will be performed during which AEs, concomitant medications and the completed HRQoL questionnaire will be collected. Thereafter, data on subsequent therapies according to routine clinical practice, adverse events of special interest (AESI) and survival will be collected until the data cut-off date for the final analysis.

Figure 1. OReO study flow chart



1.3 Number of patients

In the *BRCA1/2* (+ve) cohort it is assumed that the median PFS from randomisation for patients in the placebo arm will be approximately 4.5 months. In total, 85 progression or death events among 120 patients will have 85% power to demonstrate significant PFS benefit at the 2-sided 5% level if the assumed true treatment effect resulted in a HR of 0.5; this translates to a 4.5 month (100%) increase in median PFS beyond the 4.5 months expected for patients on placebo, if PFS is exponentially distributed and allowing a 10% drop-out rate. An observed HR of 0.63 or less will be required to achieve this level of significance. Assuming 34 months of non-linear recruitment, 85 events are expected to occur approximately 41 months after the first patient in (FSI) is enrolled into this cohort of the study.

In the *BRCA1/2* (-ve) cohort it is assumed that the median PFS from randomisation for patients in the placebo arm will be approximately 4.5 months. In total, 74 progression or death events from 108 patients will have 80% power to demonstrate a significant PFS benefit at the 2-sided 5% level if the assumed true treatment effect resulted in a hazard ratio (HR) of 0.5; this translates to a 4.5 month (100%) increase in median PFS beyond the 4.5 months expected for patients on placebo, if PFS is exponentially distributed and allowing a 10% drop-out rate. An observed HR of 0.61 or less will be required to achieve this level of significance. Assuming 36 months of non-linear recruitment, 74 events are expected to occur approximately 42 months after the FSI is enrolled into this cohort of the study.
Considering both cohorts it is expected that approximately 228 patients in total will be enrolled into the study. The assumptions made in calculating the sample size will be evaluated at the time of the planned interim analysis.

The Primary Analysis will be performed separately for the two cohorts and will occur at the same time, after the later of the two cohorts reaches the defined number of progression or death events.

2. ANALYSIS SETS

2.1 Definition of analysis sets

2.1.1 Full analysis set (Intent-to-Treat Principle)

The full analysis set (FAS) will include all randomised patients and will compare the treatment groups using the intent-to-treat (ITT) principle, i.e., on the basis of randomised treatment, regardless of the treatment actually received. Patients who were randomised but did not subsequently go on to receive study treatment are included in the FAS. Therefore, all efficacy data will be summarised and analysed using the FAS on an ITT basis as the primary analysis set.

2.1.2 Safety analysis set

All patients who received at least one dose of randomised study treatment, Olaparib or placebo, will be included in the safety analysis set. If a patient receives at least one dose of Olaparib study treatment they will be summarised in the Olaparib arm for safety summaries (e.g., Olaparib arm will include patients randomised to Olaparib who receive at least one dose of Olaparib or placebo patients who receive at least one dose of Olaparib study treatment in error at any time). If a patient randomised to Olaparib receives only placebo treatment then the patient will be summarised as part of the placebo arm.

Safety data will be summarised and analysed using the safety analysis set.

2.1.3 Patient reported outcome (PRO) analysis set

The PRO analysis set will consist of the FAS patients with a baseline PRO assessment. HRQoL data will be summarized and analysed using the PRO analysis set.

Table 1 summarizes the population analysis sets for each outcome variable.

Table 1. Summary of outcome variables and analysis populations

Outcome variable	Population
Efficacy data	
PFS, OS, TFST, TSST, TDT, Time to earliest progression by RECIST or CA-125, or death	FAS (ITT)
HRQoL	PRO
Demography and baseline characteristics	FAS (ITT)
Important protocol deviations	FAS (ITT)
Safety data	
Exposure	Safety
AEs including AESI	Safety
Laboratory measurements	Safety
Vital signs	Safety

2.2 Protocol deviations

The following general categories will be considered important deviations which will be tabulated, listed and discussed in the clinical study report (CSR):

- Patients randomised but did not receive Olaparib or matching placebo (Deviation 1)
- Patients who deviate from key entry criteria per the protocol (Deviation 2) including:
    - Female patients with histologically diagnosed relapsed non-mucinous EOC (including primary peritoneal and/or fallopian tube cancer). (Non-mucinous EOC includes patients with serous, endometrioid, and transitional cell tumours, and those with mixed histology where one of these subtypes is predominant (>50%). Inclusion of other subtypes should first be discussed with the Medical Monitor).
    - Documented *BRCA1/2* status
        - To be regarded as *BRCA1/2* (+ve), the patient must have a mutation that is predicted to be deleterious or suspected deleterious (known or predicted to be detrimental/lead to loss of function)
    - Patients must have received one prior PARPi therapy if they were enrolled from study protocol v2.0 onwards
        - PARPi therapy includes any agent (including Olaparib) used in a maintenance setting
        - For the *BRCA1/2* (+ve) cohort, the duration of first PARPi exposure must have been ≥ 18 months following a first line of chemotherapy or ≥ 12 months following a second or subsequent line of chemotherapy

        - For the *BRCA1/2* (-ve) cohort, the duration of first PARPi exposure must have been ≥ 12 months following a first line of chemotherapy or ≥ 6 months following a second or subsequent line of chemotherapy
    - For the last chemotherapy course immediately prior to randomisation on the study
        - Patients must have received a platinum-based chemotherapy regimen (carboplatin, cisplatin or oxaliplatin)) and have received at least 4 cycles of treatment
        - Patients must be, in the opinion of the investigator, in response (partial or complete radiological response), or may have no evidence of disease (if optimal cytoreductive surgery was conducted prior to chemotherapy), and no evidence of a rising CA-125, as defined below, following completion of this chemotherapy course
        - Pre-treatment CA-125 measurements must meet criterion specified below:
            - If the first value is within upper limit of normal (ULN) the patient is eligible to be randomised and a second sample is not required
            - If the first value is greater than ULN a second assessment must be performed at least 7 days after the 1st. If the second assessment is ≥ 15% more than the first the patient is not eligible
        - Patients must not have received bevacizumab during this course of treatment. Bevacizumab use as part of an earlier line of chemotherapy is permitted
        - Patients must not have received any investigational agent during this course of treatment
        - Patients must be randomised within 8 weeks of their last dose of chemotherapy (last dose is the day of the last infusion)
    - At least one lesion (measurable and/or non-measurable) that can be accurately assessed at baseline with computed tomography (CT) or magnetic resonance imaging (MRI) and is suitable for repeated assessment OR no measurable disease following a complete response to most recent chemotherapy (+/- surgery).
    - Patients receiving any systemic chemotherapy or radiotherapy (except for palliative radiotherapy) within 3 weeks prior to study treatment.
    - Concomitant use of known strong cytochrome P450 (CYP) subfamily 3A (CYP3A) inhibitors or moderate CYP3A inhibitors. The required washout period prior to starting study treatment is 2 weeks.
    - Concomitant use of known strong or moderate CYP3A inducers. The required washout period prior to starting study treatment is 5 weeks for enzalutamide or phenobarbital and 3 weeks for other agents. Persistent toxicities (Common Terminology Criteria for Adverse Event [CTCAE] Grade 2 or higher) caused by previous cancer therapy, excluding alopecia and stable Grade 2 peripheral neuropathy.

  - Patients with current or previous myelodysplastic syndrome (MDS)/acute myeloid leukaemia (AML) or with features suggestive of MDS/AML.
  - Patients with symptomatic uncontrolled brain metastases. A scan to confirm the absence of brain metastases is not required. The patient can receive a stable dose of corticosteroids before and during the study as long as these were started at least 4 weeks prior to treatment. Patients with spinal cord compression unless considered to have received definitive treatment for this and evidence of clinically stable disease for 28 days.
  - Major surgery within 2 weeks of starting study treatment and patients must have recovered from any effects of any major surgery.
- Baseline RECIST 1.1 scan more than 42 days prior to randomisation (Deviation 3)
- No baseline RECIST 1.1 assessment on or before randomisation (Deviation 4)
- Received prohibited concomitant medications (including other anti-cancer agents) (Deviation 5). Please refer to protocol section 7.7 for those medications[1] that are detailed as 'not to be administered' or being 'restricted concomitant medications'. These will be used as a guiding principle for the physician review of all medications prior to database lock.
- Patients randomised who received their randomised study treatment at an incorrect dose or received an alternative study treatment to that they were randomised to (Deviation 6).

Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1.

Deviation 1 will lead to exclusion from the safety analysis set. No other deviations will lead to patients being excluded from the analysis sets described in Section 2.1.

Important protocol deviations will be listed and summarised separately by cohort and treatment group. Any important protocol deviations that occur specifically due to the Coronavirus Disease of 2019 (COVID-19) pandemic will be clearly distinguished within the outputs. A per-protocol analysis excluding patients with specific important protocol deviations is not planned.

In addition, other major and minor study deviations will be captured during the study. The final classification of protocol deviations will be made prior to database lock and all decisions will be made whilst blinded to study treatment allocation.

All COVID-19 protocol deviations will be included in listings related to reported issues in the Clinical Trials Management System.

[1] Administered medications as documented in the eCRF which fall into the 'no other anti-cancer therapy' category.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Efficacy

3.1.1 Derivation of RECIST visit responses

RECIST 1.1 criteria will be used to assess each patient's visit response to treatment and to assess when a patient has progressed for determining PFS times. The RECIST 1.1 guidelines for measurable, non-measurable, target and non-target lesions and the objective tumour response criteria (complete response [CR], partial response [PR], stable disease [SD], or progression of disease [PD]) are presented in the Appendix F of the protocol.

The methods of assessment of tumour burden used at baseline - CT or MRI scans of chest, abdomen, pelvis, must be used at each subsequent follow-up assessment. Any other areas of disease involvement should be additionally imaged based on the signs and symptoms of individual patients.

Baseline radiological tumour assessments are to be performed no more than 28 days before randomisation and ideally as close as possible to the start of study treatment. Following the baseline assessment, efficacy for all patients will be assessed by objective tumour assessments every 12 weeks ± 7 days after randomisation, until objective disease progression as defined by RECIST 1.1. Any other sites at which new disease is suspected should also be appropriately imaged. Patients will be evaluated until objective radiological disease progression by RECIST 1.1 as per the study schedule regardless of whether study treatment is discontinued or delayed and/or the occurrence of protocol violations, unless they withdraw consent.

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.

If the Investigator is in doubt as to whether progression has occurred, particularly with response to non-target lesion or the appearance of a new lesion, it is advisable to continue treatment until the next scheduled assessment or sooner if clinically indicated and reassess the patient's status. If repeat scans confirm progression, then the date of the initial scan should be declared as the date of progression.

If a patient discontinues treatment (and/or receives a subsequent cancer therapy) prior to progression, then the patient should continue to be followed-up until objective disease progression as defined by RECIST 1.1.

From the investigators review of the imaging scans, the RECIST 1.1 tumour response data will be used to determine each patient's visit response according to RECIST 1.1. At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD), using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a

patient has had a tumour assessment that cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE), (unless there is evidence of progression in which case the response will be assigned as PD).

Please refer to Section 3.1.1.1 for the definitions of CR, PR, SD and PD.

All RECIST 1.1 assessments, scheduled or unscheduled, will be included in the calculations. This is also regardless of whether a patient discontinues study treatment or receives another anti-cancer therapy. A sensitivity analysis may be performed without the data after treatment discontinuation and after starting another anti-cancer therapy to assess whether patients who have discontinued for whatever reason other than PD appear to have benefited.

3.1.1.1 Target lesions

Measurable disease is defined as having at least one measurable (by RECIST 1.1) lesion, that has not been previously irradiated, which is ≥10 mm in the longest diameter (except lymph nodes, which must have short axis ≥15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. If only one measurable lesion exists, it is acceptable to be used (as a target lesion) as long as it has not been previously irradiated and baseline tumour assessment scans are done at least 14 days after the screening biopsy is performed.

A patient can have a maximum of five measurable lesions, with a maximum of two lesions per organ (representative of all lesions involved suitable for accurate repeated measurements) and these are referred to as TLs at baseline. Target lesions should be selected on the basis of their size (longest diameter for non-nodal lesions or short axis for nodal lesions), but in addition should be those that also lend themselves to reproducible repeated measurements. 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.

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

Table 2 provides the definition of the criteria used to determine objective tumour visit response for TL.

Table 2. TL visit responses

Visit Responses	Description
CR	Disappearance of all TLs since baseline. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10 mm.
PR	At least a 30% decrease in the sum of the diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
SD	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.
PD	At least a 20% increase in the sum of the diameters of TLs, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on the study). In addition to the relative increase of at least 20%, the sum must also demonstrate an absolute increase of at least 5 mm.
NE	Only relevant if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit. Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response.
NA	No TLs are recorded at baseline.

CR Complete response, NA Not applicable, NE Not evaluable, PD Progressive disease, PR Partial response, SD Stable disease, TL Target lesion.

Rounding of TL data

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

Missing TL data

For a visit to be evaluable, 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 ≥5 mm, from nadir even assuming the non-recorded TLs have disappeared.

The nadir (i.e., the smallest measurement based on the same set of lesions at baseline and on a given visit) can only be taken from assessments where all the TLs had a longest diameter (LD) recorded.

For patients with at least one TL at baseline, if there is at least one TL measurement missing and a visit response of PD cannot be assigned, the visit response is NE.

Lymph nodes

For lymph nodes, if the size reduces to <10 mm, these are considered non-pathological.

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

TL visit responses subsequent to CR

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

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

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

- Step 3: If not all lesions meet the CR criteria, or a new lesion appears, the response will be set to PD.

- Step 4: If, after steps 1 through 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 an estimated 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. 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 5 mm 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.

Irradiated lesions/lesion intervention

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

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

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

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

- Step 3: If after both steps PD has not been assigned, then, if appropriate (i.e., if ≤1/3 of the TLs have missing measurements), the scaled sum of diameters as calculated in Step 2 should be used, and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or <10 mm for lymph nodes) and the lesions that have been subject to intervention also has a value of 0 recorded. If scaling up is not appropriate due to too few non-missing measurements, then the visit response will be set to NE.

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

Scaling (applicable only for irradiated lesions/lesion intervention)

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

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

Example of scaling

Lesion 5 is missing at the follow-up visit; lesions 1 to 5 had a nadir measure of 29.3 cm.

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

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

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

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

Lesions that split

If a TL splits into two or more parts, 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 or more TLs merge, then the LD of the merged lesion should be recorded for one of the TL sizes and the other TL size(s) should be recorded as 0 cm.

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 involves clinical examination (e.g., CT changes to clinical examination or vice versa), any affected lesions should be treated as missing.

3.1.1.2 Non-target lesions and new lesions

At each visit, an overall assessment of the NTL response should be recorded by the Investigator. Table 3 provides the definitions of the criteria used to determine and record overall response for NTLs at the investigational site at each visit.

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

Table 3. NTL visit responses

Visit Responses	Description
CR	Disappearance of all NTLs since baseline. All lymph nodes must be non-pathological in size (<10 mm short axis).
Non CR/Non PD	Persistence of one or more NTLs (with no evidence of progression).
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.
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.
NA	Only relevant if there are no NTLs at baseline.

CR Complete response, NA Not applicable, NE Not evaluable, PD Progressive disease, NTL Non-target lesion

To achieve 'unequivocal progression' on the basis of NTLs, there must be an overall level of substantial worsening in non-target disease such that, even in presence of SD or PR in TLs, the overall tumour burden has increased sufficiently to merit a determination of 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 will be considered as progression.

A lesion identified at a follow-up assessment in an anatomical location that was not scanned at baseline is considered as 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 a tumour.

If a new lesion is equivocal, for example because of its small size, the treatment and tumour assessments should be continued until the new lesion has been confirmed. If repeat scans confirm there is a new lesion, then the progression date should be declared using the date of the initial scan.

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

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 and should be treated as NE in the derivation of overall visit response.

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

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

3.1.1.3 Overall visit response

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

Table 4. Overall visit response

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

CR Complete response, NA Not applicable (only relevant if there were no TLs or NTLs at baseline), NE Not evaluable, NED No evidence of disease, PD Progressive disease, PR Partial response, SD Stable disease.

3.2 Efficacy variables

3.2.1 Primary endpoint

3.2.1.1 Progression-free survival

PFS (per RECIST 1.1 according to the Investigator's assessment) is defined as the time from the date of randomisation until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from

allocated therapy or receives another anti-cancer therapy prior to progression, in months and is calculated as follows:

$$PFS = \frac{date\ of\ objective\ disease\ progression/death\ or\ censoring - date\ of\ randomisation + 1}{30.4375}$$

Patients who have not progressed or died at the date of data cut-off (DCO) will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment.

However, if the patient progresses or dies after two or more missed visits, the patient will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the two missed visits.

Given the scheduled visit tumour assessment scheme, two missing visits will equate to more than 26 weeks since the previous RECIST assessment, allowing for early and late visits (NE is not considered a missed visit).

If the patient has no evaluable visits[2] or does not have baseline data, they will be censored at Study Day 1 unless they die within two visits of baseline (25 weeks allowing for visit window).

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

RECIST 1.1 assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- Date of progression will be determined based on the earliest of the dates of the component that triggered the progression.
- When censoring a patient for PFS, the patient will be censored at the latest of the RECIST assessment/scan dates contributing to a particular overall visit assessment.

Overall visit assessments will be determined for each visit (scheduled or unscheduled) and will contribute to the derivation of PFS.

[2] Evaluable visits are visits where tumour response or NED can be determined according to RECIST.

3.2.2 Secondary endpoints

3.2.2.1 Overall survival

OS is defined as the time from the date of randomisation until the date of death due to any cause, in months and is calculated as follows:

$$OS = \frac{date\ of\ death\ or\ censoring - date\ of\ randomisation + 1}{30.4375}$$

Any patient not known to have died at the date of the DCO for the analysis will be censored based on the last recorded date on which the patient was known to be alive (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made in the week following the date of data cut-off for the analysis, and if patients are confirmed to be alive or if the death date is post the data cut-off date, these patients will be censored at the date of data cut-off.

For any OS analysis performed prior to the final OS analysis, in the absence of survival calls being made (or where a last known to be alive data cannot be determined), it may be necessary to use all relevant eCRF fields to determine the last recorded date on which the patient was known to be alive. The last date for each individual patient is defined as the latest among the following dates recorded on the eCRFs:

• AE start and stop dates

• Admission and discharge dates of hospitalization

• Study treatment date

• End of treatment date

• Laboratory test dates

• Date of vital signs

• Disease assessment dates on RECIST CRF

• Start and stop dates of alternative anticancer therapy

• Patient end of study date .

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

a. For Missing day only – using the 1st of the month

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

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

3.2.2.2 Time to first and second subsequent therapy or death

As a supportive summary to PFS, TFST will be assessed. TFST is defined as the time from the date of randomisation to the earlier of first subsequent therapy start date (excluding radiotherapy), or death date, in months and is calculated as follows:

$$TFST = \frac{date\ of\ first\ subsequent\ cancer\ therapy/death\ or\ censoring - date\ of\ randomisation + 1}{30.4375}$$

Any patient not known to have died and not known to have had a further subsequent therapy will be censored at the last date that the patient was known not to have received subsequent therapy (i.e. the last follow-up visit where this was confirmed). If a patient terminated the study for reasons other than death before first subsequent therapy, these patients will be censored at the earliest of the DCO, their last known to be alive and termination dates.

Additionally, TSST will be assessed. TSST is defined as the time from the date of randomisation to the earlier of the date of second subsequent therapy start date, or death date, in months and is calculated as follows:

$$TSST = \frac{date\ of\ second\ subsequent\ cancer\ therapy/death\ or\ censoring - date\ of\ randomisation + 1}{30.4375}$$

Any patient not known to have died and not known to have had a further second subsequent therapy will be censored at the last known time to have not received second subsequent therapy (i.e., the last follow-up visit where this was confirmed). If a patient terminated the study for reasons other than death before a second subsequent therapy, these patients will be censored at the earliest of the DCO, their last known to be alive date or termination date.

3.2.2.3 Time to study treatment discontinuation or death

TDT is defined as the time from randomisation to the earlier of the date of permanent study treatment discontinuation or death, in months and is calculated as follows:

$$TDT = \frac{date\ of\ study\ treatment\ discontinuation/death\ or\ censoring - date\ of\ randomisation + 1}{30.4375}$$

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

3.2.2.4 Time to earliest progression by RECIST or CA-125 or death

Time to progression by RECIST or CA-125 or death is defined as the time from randomisation to the earlier date of RECIST progression or CA-125 progression or death by any cause. Patients without a CA-125 progression or a RECIST progression who are still alive at the time of analysis will be censored at the time of their last evaluable RECIST assessment and/or their last available CA-125 measurement, whichever is the earliest at the time of analysis. Patients that do not have any evaluable RECIST assessments or any CA-125 results post randomisation will be censored at the date of randomisation. Patients who have RECIST progression or CA-125 progression denoted after 2 missed assessment visits will be censored at the date of the visit prior to progression.

Progression or recurrence based on serum CA-125 levels will be defined on the basis of a progressive serial elevation of serum CA-125, according to the following modified GCIG criteria (Rustin 2003) (note GCIG criteria is not validated for this trial population):

- For patients with elevated CA-125 on or before the date of randomisation (i.e., greater than the upper limit of normal (ULN)):
  a) If CA-125 does not fall to within the normal range post randomisation then there must be evidence of CA-125 greater than, or equal to, two times the nadir value in the 28-day period before day 1 on two occasions at least 1 week apart.
  b) Where CA-125 does fall to within the normal range post randomisation (and the patient has not already progressed by way of a) above), then there must be evidence of CA-125 greater than, or equal to, two times the ULN on two occasions at least 1 week apart.
- Patients with CA-125 in the normal range on or before the date of randomisation must show evidence of CA- 125 greater than, or equal to, two times the ULN on two occasions at least 1 week apart.
- CA-125 progression will be assigned the date of the first measurement that meets the above criteria.

Although it is planned that CA-125 will be assessed at the same time as RECIST the two missed visit rule will be based upon the RECIST schedule to account for any instances of differences in assessment dates. The two missed visit rule will be applied as follows:

1. Consider RECIST assessments and CA-125 assessments separately.
2. Evaluate the progression/death or censoring date of RECIST, applying the two missed visit rule without considering CA-125.
3. Evaluate the progression/death or censoring date of CA-125 (last evaluable CA-125), applying the two missed visit rule for censoring to match the same duration of the 2 missed visit RECIST schedule.
4. Take the earliest progression date from 2 and 3 (still take the progression date even if progression is on RECIST but not CA-125, or vice versa).
5. If both are censored take the earliest censoring date.
6. If only one is censored, PFS will be an event at the progression date for either endpoint.

3.2.2.5 Best Overall RECIST Response (BoR)

Best overall response will be reported in order to support the interpretation of PFS and is calculated based on the visit responses from each RECIST assessment (Table 4). It is the best response a patient has had following randomisation but prior to starting any subsequent cancer therapy and prior to RECIST progression or the last evaluable assessment in the absence of RECIST progression.

Categorisation of best overall response will be based on the RECIST criteria using the following response categories: complete response (CR), partial response (PR), stable disease (SD), No Evidence of Disease (NED; applies only to those patients entering the study with no disease at baseline), progressive disease (PD) and not evaluable (NE).

Best overall response will be programmatically derived from the investigator data.

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 12 weeks +/- 1 week, i.e. at least 77 days (to allow for the assessment window), after randomisation. For CR/PR, the initial post-baseline overall visit assessment which showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

For patients whose progression event is death, BoR will be calculated based on data up until and including the last evaluable RECIST assessment prior to death.

For patients who die with no evaluable RECIST assessments, if the death occurred ≤25 weeks (i.e., 24 weeks ±1 week) after randomisation then BoR will be assigned as PD. For patients who die with no evaluable RECIST assessments, if the death occurred >25 weeks (i.e., 24 weeks ±1 week) after randomisation then BoR will be assigned to the non-evaluable (NE) category.

Progression events that have been censored due to them being more than two missed visits after the last evaluable assessment will not contribute to the BOR derivation.

3.2.2.6 PRO endpoints

The FACT-O questionnaire consists of 39 questions: 27 Functional Assessment of Cancer Therapy – General (FACT-G) items and 12 additional concerns items consisting of specific ovarian cancer symptoms). The questionnaire will be scored into subscales and composite

scores according to the Functional Assessment of Chronic Illness Therapy (FACIT) scoring guidelines as follows:

- Physical Well-Being (PWB): Score range 0-28
- Social/Family Well-Being (SWB): Score range 0-28
- Emotional Well-Being (EWB): Score range 0-24
- Functional Well-Being (FWB): Score range 0-28
- Ovarian Cancer Subscale (OCS): Score range 0-44
- FACT-O TOI, Score range 0-100. Derived using PWB+FWB+OCS
- FACT-O total score, Score range 0-152, Derived using PWB+FWB+SWB+EWB+OCS
- FACT-G total score: Score range 0-108. Derived using PWB+FWB+EWB+SWB.

The higher the score, the better the HRQoL for all subscales. Missing items will be dealt with as described in the FACIT Administration and Scoring Guidelines. If at least 50% of items in a subscale have been answered the subscale score will be prorated. Subscale scores will be missing if less than 50% of items within that subscale have been answered. Total scores will only be calculated if all component subscales have valid scores. The reason for any missing assessment will be collected in the CRF.

The TOI score will be of primary interest. Other scores and subscales will be considered as exploratory.

Four outcome measures will be calculated:

- Change from baseline score

  The actual change from baseline score will be derived for each visit where there is available data. For example; at visit X, the calculation will be (subscale score at visit X – baseline subscale score). Actual change from baseline for the individual subscale scores will be calculated in a similar way. Subscale-specific minimally important differences (MID) will be used for interpretation where available (Yost and Eton 2005).

- Proportion of patients with a PRO response (improved PRO Score) – TOI score only

  The proportion of patients with an improved score (≥MID of 10 points increase from baseline), worsened (≥MID of 10 points decrease from baseline) or no change (changes of less than MID of 10 points in either direction) will be calculated at each visit. The denominator will consist of all patients in the PRO analysis set.

- Best overall response – TOI score only

  Best overall improvement (improvement in the absence of subsequent cancer therapy) will be defined as an increase from baseline of ≥MID of 10 sustained for at least 28 days; the denominator consisting of all patients in the PRO analysis set. It will be derived as the best symptom improvement response the patient achieved, based on evaluable QoL data

collected from randomisation up to the earliest of starting any subsequent cancer therapy or death. Therefore, the following criteria will be used to assign a best overall score response for each patient based on the individual visit responses (Table 5).

Table 5. HRQoL: Best Overall Response

Best overall response	Criteria
Improved	Two visit responses of "improved" a minimum of 28 days apart without an intervening visit response of "worsened"
No change	Does not qualify for overall score response of "improved". Two visit responses of either "no change" or "improved and "no change" a minimum of 28 days apart without an intervening visit response of "worsened"
Worsened	Does not qualify for overall score response of "improved" A visit response of "worsened" without a response of "improved" or "no change" within 28 days.
Other	Does not qualify for one of the above.

An improvement rate (in the absence of subsequent cancer therapy) will be calculated as the percent of all analysed patients with a best overall response of improved. In the calculation of the proportion of patients that have a response of Improved, No Change or Worsened, the denominator used in the calculation will use the number of evaluable patients for the subscale score at baseline.

- PRO deterioration-free survival – TOI score only

  Time from randomisation to definitive deterioration of TOI score (≥MID of 10 points decrease from baseline with another ≥MID of 10 points decrease from baseline a minimum of 28 days apart without an intervening improvement on the ≥MID of 10 points decrease from baseline or subsequent missing data).

  Patients that did not experience PRO deterioration or who died or progressed without a prior definitive deterioration will be censored at the latest PRO assessment date.

Summary measures of overall compliance and compliance over time will be derived for the FACT-O questionnaire. These will be based upon:

- Received forms = number of FACT-O forms received back plus the number not received back where the reason was 'Subject too heavily affected by symptoms of disease under investigation'

- Expected forms = number of patients still under HRQoL follow-up at the specified assessment time excluding patients in countries with no available translation. For patients that have progressed, the latest of progression and safety follow-up dates 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 forms = subset of expected FACT-O forms with at least one subscale that can be determined.

Thus the overall compliance rate is defined as the number of patients with an evaluable baseline and at least one evaluable follow-up form (as defined above), divided by the number of patients expected to have completed at least a baseline FACT-O form.

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

3.2.2.7 Time to gastrointestinal symptom deterioration

Abdominal symptoms are common symptoms associated with recurrence or progression of ovarian cancer. The time to first recorded gastrointestinal (GI) symptom deterioration will be compared between the two treatment arms as it is expected that abdominal symptoms will be delayed for patients randomised to Olaparib. Delay in abdominal symptoms between treatment groups will be assessed using the following four items from the FACT-O physical wellbeing subscale and ovarian cancer subscale (Additional Concerns):

1. I have swelling in my stomach area
2. I have cramps in my stomach area
3. I have control of my bowels
4. I have pain

Patients will be included in this analysis if they have a zero score (with appropriate reverse scoring applied) on their baseline questionnaire and the four questions above will be assessed separately. Time to event will be calculated as time from randomisation to first time the patient records a response of ≥1 (‘a little bit’) for the statement ‘I have swelling in my stomach area’ on the FACT-O questionnaire. If the patient does not record a response of ≥1 (‘a little bit’) at any time point then they will be censored at time of starting subsequent cancer therapy or death (whichever occurs first). Any patient not known to have died and not known to have had a subsequent therapy will be censored at the last date that the patient was known not to have received subsequent therapy (i.e. the last follow-up visit where this was confirmed). If a patient terminated the study for reasons other than death before first subsequent therapy, these patients will be censored at the earliest of the DCO, their last known to be alive and termination dates.

The same method will be used to assess the other three statements specified above.

This analysis will only be performed if there are sufficient numbers of patients available. The decision as to whether to perform this analysis will be determined prior to study unblinding.

3.2.2.8 Time to progressive disease for patients with significant symptoms at baseline

It is expected that the time to protocol defined progressive disease will be shorter in patients with significant symptoms at baseline. A subgroup of patients with significant symptoms at baseline will be defined according to the following four items from the FACT-O physical wellbeing subscale and ovarian cancer subscale (Additional Concerns):

1. I have swelling in my stomach area
2. I have cramps in my stomach area
3. I have control of my bowels
4. I have pain

Patients that answered ‘quite a bit’ or ‘very much’ to any of the four questions on their baseline questionnaire will be defined as having significant symptoms at baseline.

3.2.2.9 Assessing the impact of post-progression chemotherapy on HRQoL

It is expected that commencing chemotherapy post-progression will be associated with a decline in health-related quality of life but that the initiation of chemotherapy for progression will be delayed for longer in patients receiving Olaparib.

3.2.3 General consideration for patient reported outcome variables

Due to the practicality of scheduling patient visits, not all patients will have their visits on the same study day. In order to allow for any presentations that summarise values by visit, visit windows will be defined.

The window for the visits following baseline up to and including 30 days after the date of last dose of study drug will be constructed as per the safety visit windows, 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). The half-way point will be assumed to be the midpoint of the number of days between the visits, excluding both visit days (for example there are

assumed to be 54 days between Day 29 and Day 85). If an odd number of days exist between two consecutive visits, then the upper limit will be taken as the midpoint value plus 0.5 day.

For example:

Day 29, visit window 2 – 57

Day 85, visit window 58– 127

Day 169, visit window 128 – 211

For post follow-up visits the windows will be taken relative to the end of treatment, for example:

12 weeks post follow up, visit window 32 - 127 days post treatment

24 weeks post follow up, visit window 128 - 211 days post treatment

In addition, an End of Treatment visit will be identified as the visit occurring between 1 and 8 days (inclusive) after the end of treatment. And similarly, a 30 day follow up visit will be identified as the visit between 9 and 31 days (inclusive) following end of treatment. These additional points will allow separate summaries to be presented for these visits as well as allow the inclusion of this data in mapped on-treatment visit summaries.

3.3 Safety variables

Safety and tolerability will be assessed in terms of AEs, SAEs, AEs leading to discontinuation of study drug from randomisation to 30 days after last study treatment, AESIs, laboratory data (including chemistry and haematology) and vital signs. The latest version of the Medical Dictionary for Regulatory Activities (MedDRA) will be used to code the AEs. AEs will be graded according to version 4 of the National Cancer Institute Common Terminology Criteria for AEs (CTCAE).

3.3.1 Adverse events

AEs and SAEs will be collected from date of informed consent, throughout the treatment period and including the 30-day follow-up period after the last dose of treatment. Events will be defined as treatment emergent adverse events (TEAEs), if they onset or worsen in severity (from baseline severity) on or after the first day of Olaparib or placebo administration up to and including 30 days after the last dose of Olaparib or placebo (defined as the treatment period).

3.3.2 Adverse events of special interest

AESIs are events of scientific and medical interest specific to the further understanding of Olaparib's safety profile and require close monitoring and rapid communication by the Investigators to AstraZeneca. An AESI may be serious or non-serious. AESIs for Olaparib are

the important potential risks of myelodysplastic syndrome (MDS)/acute myeloid leukaemia (AML), new primary malignancy (other than MDS/AML) and pneumonitis.

These AESIs have been identified as a list of categories provided by the patient safety team. Other categories may be added as necessary or existing terms may be merged. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which higher-level terms and which preferred terms contribute to each AESI. Further reviews may take place prior to database lock to ensure any further terms not already included are captured within the categories.

3.3.3 Duration of exposure and dose interruptions

Exposure will be reported separately for each cohort and treatment group.

Olaparib and placebo are taken daily. Hence, total (or intended) exposure time (months) of study treatment (Olaparib or placebo) will be calculated as follows:

- [the total treatment period from the date of first dose of study treatment to the earliest of (last dose, death date) +1] / 30.4375

Actual exposure of Olaparib (or placebo) will be calculated as follows:

- 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 dose (taking into account the scheduled off treatment period). The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

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

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

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.

3.3.3.1 Safety follow-up

Total safety follow-up will be calculated as follows:

- 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

3.3.4 Dose intensity

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

RDI and PID will be defined as follows:

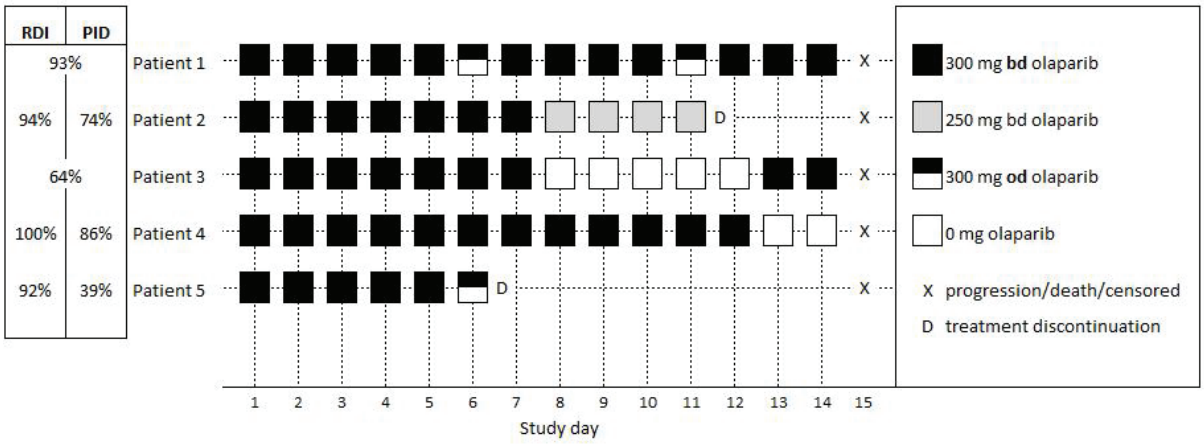
$$RDI\ (\%) = 100\ \times \frac{d}{D}$$

where d is the actual cumulative dose delivered up to the earlier of progression (or a censoring event) or the actual last day of dosing and D is the intended cumulative dose up to the earlier of progression (or a censoring event) or the actual last day of dosing or the actual last day of dosing plus the protocol-defined post-dose rest period.

$$PID\ (\%) = 100\ \times \frac{d}{D}$$

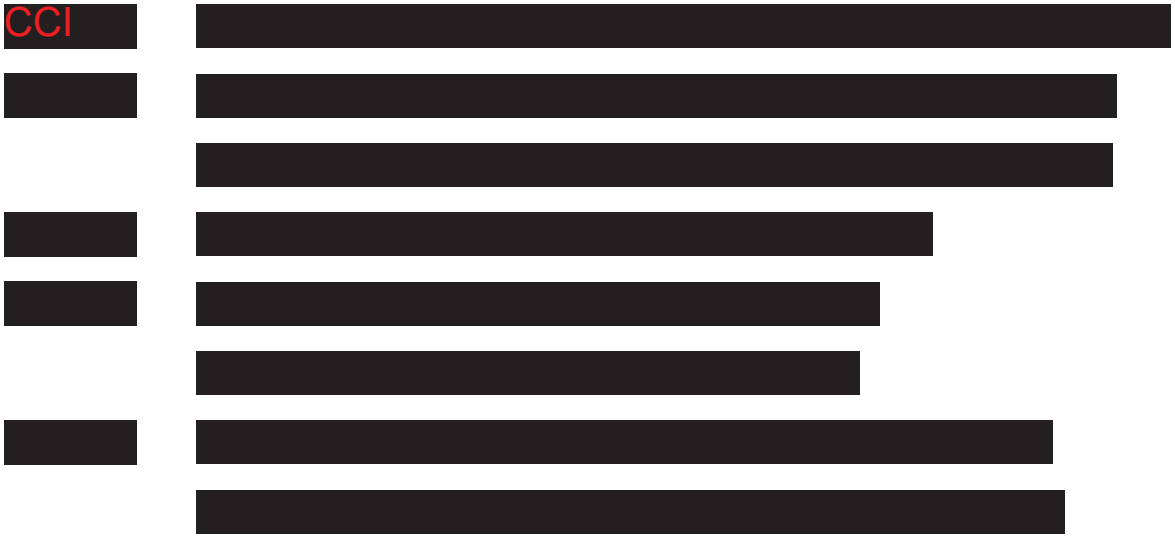
where d is the actual cumulative dose delivered up to progression (or a censoring event) and D is the intended cumulative dose up to progression (or a censoring event). D is the total dose that would be delivered, if there were no modification to dose or schedule.

Figure 2. Example of Dose Intensity Calculations for Olaparib



In this example, patients 1-4 progressed or were censored on Day 15. All four patients received less treatment than intended due to:

- Missed/forgotten doses (Patient 1)
- Dose reduction and early stopping (Patient 2)
- Dose interruption (Patient 3)
- Progression whilst on dose interruption (Patient 4)
- Early stopping (Patient 5)



3.3.5 Laboratory variables

Laboratory data will be collected throughout the study, from screening to the follow-up visits as per study plan for each cohort and treatment arm described in the protocol. These include blood samples for determination of clinical chemistry and haematology. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 3.3.7 will be used.

Change from baseline in clinical chemistry and haematology variables will be calculated for each post-dose visit on treatment and up to 30 days follow-up. Common toxicity criteria (CTC) grades will be defined at each visit according to the CTC grade criteria using local or project ranges as required, after conversion of lab result to corresponding project-wide preferred units. The following parameters have CTC grades defined for both high and low values, so high and low CTC grades will be calculated: potassium, sodium, magnesium, glucose and corrected calcium.

Corrected calcium (mmol/L) will be derived using the following formula (where total calcium is measured in mmol/L and albumin is measured in G/L):

$$Corrected\ calcium = Total\ calcium + \ (40 - albumin) \times 0.02$$

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

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

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

For example:

- If a CTC criterion involves a change from baseline, evaluable patients would have both a baseline and at least 1 post-dose value recorded.

- If a CTC criterion does not consider changes from baseline, to be evaluable the patient needs only to have 1 post dose-value recorded.

3.3.6 Vital signs

Vital signs data (blood pressure, pulse, body temperature and respiration rate) obtained up to 30 days from the date of last dose of study treatments will be used for reporting. Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post-baseline visit values, considering visit window, and to handle multiple records, derivation rules as described in Section 3.3.7 will be used.

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

3.3.7 General considerations for safety assessments

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

- The time windows should be exhaustive so that data recorded at any time point has the potential to be 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 one day.

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

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

- For visit-based summaries, if there is more than one value per patient within a time window, then the value closest to the planned study day should be summarised, or the earlier in the event the values are equidistant from the planned study day. The listings should highlight the value for each patient that was included in the summary table, wherever feasible.

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

- Baseline will be defined as the last non-missing measurement prior to dosing with study treatment. For laboratory data, any assessments made on Study Day 1, when time is not captured, will be considered pre-dose. Alternatively, if two visits are equally eligible to assess patient status at baseline (egg, 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. Where safety data are summarized over time, study day will be calculated in relation to date of first study treatment.

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.

3.4 Biomarker variables

Biomarker data from biopsies and plasma at diagnosis, baseline, and progression on PARPi will be used for additional CCI work, which may be conducted to elucidate the mechanism of response, understand the mode of action of study treatment, or improve the understanding of disease progression. These analyses however will be described and reported separately from the current SAP.

4. ANALYSIS METHODS

4.1 General principles

The following general principles will be followed throughout the study:

- All analyses and summaries will be performed and presented separately for each cohort.
- All summaries will be presented by treatment arm.
- Continuous variables will be summarised by the number of observations, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarised by frequency counts and percentages for each category.
- Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment arm.
- The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean and median will be reported to one more decimal place, and the standard deviation will be reported to two more decimal places, than the raw data recorded in the database.
- Percentages will be presented to one decimal place.
- SAS® version 9.3 will be used for all analyses.
- For time interval analyses in months, duration in months will be calculated as total duration in days/30.4375.

Study Day 1 is defined as the date of randomisation. For visits (or events) that occur on or after randomisation, study day is defined as (date of visit [event] - date of randomisation + 1). For visits (or events) that occur prior to randomisation, study day is defined as (date of visit [event] - date of randomisation). There is no Study Day 0.

For efficacy, safety and PRO endpoints, the last observed measurement prior to first dose of study treatment will be considered the baseline measurement. However, if the date of the first dose of study treatment is not available, then the last observed measurement prior to randomisation will be used as baseline.

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

Efficacy and safety data will be summarised and analysed based on the FAS and safety analysis sets respectively. HRQoL data will be summarised and analysed based on the PRO Analysis Set. Study population and demography data will be summarised based upon the FAS.

4.2 Analysis methods

4.2.1 Patient disposition

The total number of patients screened (including screening failures), enrolled (i.e. patients with informed consent), consisting of patients assigned to treatment and patients not assigned to treatment, patients who received any study treatment and patients who did not receive any study treatment, patients ongoing treatment at DCO, patients who completed treatment, patients who discontinued treatment and the reason for discontinuation, patients who discontinued study and the reason for discontinuation will be summarised.

The number and percentage of patients included in the analysis populations and the number of patients recruited by country and centre will be also presented.

The COVID-19 pandemic was declared by the World Health Organisation on 11th March 2020. Due to this, some COVID-19 specific listings will be presented to show the impact of the pandemic.

By-patient listings of disposition details for patients affected by the COVID-19 pandemic and patients with reported issues in the Clinical Trials Management System due to COVID-19 pandemic will be provided.

4.2.2 Demographics and other baseline characteristics

Demographic and baseline patient characteristics (age, race, weight [kg], weight group [<70, ≥70 – ≤90, >90], Eastern Cooperative Oncology Group (ECOG) performance status and previous cancer therapy, smoking history and stratification factors recorded at randomisation: use of prior bevacizumab [yes, no] and number of prior regimens of platinum-containing chemotherapy [≤3, ≥4]) (and any deviations from the stratification) will be summarised for the FAS.

Descriptive statistics will be presented for the continuous variables and total counts and percentages of patients will be presented for the categorical variables.

Additionally, the following variables will be summarized:

- Age as a continuous variable
- Age by
    - class (<50; ≥50 - <65; ≥65 - <75; ≥75) and
    - class (<65; ≥65)

4.2.3 Medical history

Relevant medical history (past and current) and relevant surgical history will be coded using the latest version of MedDRA.

All medical history will be summarised (number and percentage of patients) for the FAS by system organ class (SOC) and preferred term (PT).

All relevant surgical history (ovarian and other) will be summarised similarly.

4.2.4 Disease history

The following disease characteristics will be summarized for all patients in the FAS (unless specified otherwise):

- Disease characteristics at baseline (time from completion of most recent platinum-containing chemotherapy to randomisation, time from diagnosis to randomisation [months], FIGO stage, primary tumour location, histology type, best response to previous therapy [CR, PR, SD, PD, NE, NA], CA-125 values, *BRCA* status [positive, negative])
- Extent of disease upon entry to study [site of metastatic disease]
- Previous disease-related treatment modalities
- Type of most recent platinum-containing chemotherapy [carboplatin, cisplatin, other]
- Number of regimens of previous platinum-containing chemotherapy

4.2.5 Concomitant medication

Prior and concomitant medications are defined as follows:

- Prior medications are medications taken prior to or during screening with a stop date prior to the first dose of study treatment
- Concomitant medications are medications with or without a stop date on or after the date of first dose of study treatment (and could have started prior to or during treatment).

Concomitant medication will be summarised using frequency tables by Anatomical therapeutic chemical (ATC) classification code (based on World Health Organisation (WHO) classification).

Moreover, the following concomitant medications or therapies will be summarized:

- Disallowed concomitant medications (other anti-cancer therapies, investigational agents, and radiotherapy should not be given while the patient is on study drug)
- Post-treatment anti-cancer therapies (post-treatment medications are those with a start date after the last dose date of study treatment)

4.2.6 Exposure

The following summaries will be produced for the safety analysis set:

- Total (or intended) exposure to study treatments,
- Actual exposure to study treatments
- Number of cycles of study treatments initiated
- Number of and reasons for dose interruptions

- Number of and reasons for dose reductions
- RDI and PID
- Time on study – defined as the time in months from the date of randomisation to the date of last study assessment or the date of withdrawal, whichever comes first.

4.2.7 Efficacy

Table 6 presents the formal statistical analyses and the pre-planned sensitivity analyses to be conducted for efficacy endpoints.

Table 6. Formal efficacy analyses and sensitivity analyses

Endpoints analysed	Notes
PFS	Primary analysis stratified log-rank test Sensitivity analyses 1. Time assessment bias 2. Attrition bias 3. Deviation bias 4. Adjustment for additional prognostic factors 5. Cohort adjustment
Overall survival	Secondary endpoints
Time to subsequent therapies or death	
Time to study treatment discontinuation or death	
Time to earliest progression by RECIST or CA-125 or death	

4.2.7.1 Progression-free survival

The primary analysis for PFS for each cohort will be performed when both 85 progression or death events have occurred in the *BRCA1/2* (+ve) cohort and 74 progression or death events in the *BRCA1/2* (-ve) cohort (whichever occurs later). The primary analysis will be performed separately for the two cohorts. The primary analysis will be based on Investigator assessment of disease progression according to RECIST 1.1.

PFS will be analysed using a stratified log-rank test at the 5% significance level (two-sided), adjusting for treatment, use of prior bevacizumab (yes versus no) and number of prior regimens of platinum containing chemotherapy (≤3 versus ≥4 regimens), for generation of the p-value and using the Breslow approach for handling ties (Breslow, 1974).

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

PFS will be analysed using Kaplan-Meier (K-M) methodology and the median and its 95% CI will be provided for each treatment group, together with PFS rates at clinically relevant time points (6 months and 12 months). K-M curves will also be provided, with tick marks to identify censored observations.

The HR and its confidence interval (CI) will be estimated (HR less than 1.0 favours Olaparib) from a Cox Proportional Hazards model (with ties = Efron and the stratification factors as covariates) and the CI calculated using a profile likelihood approach.

The assumption of proportionality will be assessed. The results of these model checks will not be presented in the CSR as part of the formal outputs; however, any deviation from this assumption will be considered in the interpretation. Note that in the presence of non-proportionality, the HR will be interpreted as an average HR over the observed extent of follow-up. Proportionality will be tested firstly by producing plots of complementary log-log (event times) versus log (time) and, if these raise concerns, a time dependent covariate would be fitted to assess the extent to which this represents random variation.

Kaplan-Meier plots of PFS will be presented by cohort and treatment arm. The treatment status at progression of patients at the time of analysis will be summarised. This will include the number (%) of patients who were on treatment at the time of progression, the number (%) of patients who discontinued study treatment prior to progression, the number (%) of patients who have not progressed and were on treatment or discontinued treatment. This will also provide the distribution of the number of days that treatment was discontinued prior to progression for the patients who have discontinued treatment.

Additionally, the duration of follow-up will be summarised:

- In censored patients who have not progressed at the last assessment before DCO only: Time from date of first dose of randomised treatment in this study to date of censoring (date last known to have not progressed before the DCO).
- In all patients: Time from date of first dose of randomised treatment in this study to the date of progression/death or to the date of censoring for censored patients.

4.2.7.2 Overall survival

Similar to PFS, OS will be analysed using a stratified log-rank test at the 5% significance level (two-sided), adjusting for treatment, use of prior bevacizumab (yes versus no) and number of prior regimens of platinum containing chemotherapy (≤3 versus ≥4 regimens), for generation of the p-value and using the Breslow approach for handling ties.

Moreover, OS will be analysed using K-M methodology and the median and its 95% CI will be provided for each treatment group, together with OS rates at clinically relevant time points (e.g., at 6 months and 12 months). K-M curves will also be provided, with tick marks to identify censored observations.

The HR and its CI will be estimated (HR less than 1.0 favours Olaparib) from a Cox Proportional Hazards model (with ties = Efron and the stratification factors as covariates) and the CI calculated using a profile likelihood approach.

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

OS will be analysed at the time of the primary analysis for PFS and after 50% death events in either cohort, or 60 months after First Patient In (FPI), whichever is earlier.

In addition, the duration of follow-up will be summarised:

- In censored patients who are alive at the last assessment before DCO only: Time from date of first dose of randomised treatment in this study to date of censoring (date last known to be alive before the DCO).
- In all patients: Time from date of first dose of randomised treatment in this study to the date of death (i.e. overall survival) or to the date of censoring for censored patients.

The median follow-up time will be presented by cohort and treatment group. A summary of the number of patients prematurely censored (defined as patients with a status not defined at the DCO and with last available eCRF information more than 6 weeks prior to the DCO) will also be produced.

4.2.7.3 Time to subsequent therapies or death

Time to first subsequent treatment or death and time to second subsequent treatment or death for the two treatment groups will be analysed using a two-sided stratified log-rank test at the 5% significance level, adjusting for treatment, use of prior bevacizumab (yes versus no) and number of prior regimens of platinum-containing chemotherapy (≤3 versus ≥4 regimens), for generation of the p-value and using the Breslow approach for handling ties.

Moreover, TFST and TSST will be analysed using K-M methodology and the median and their 95% CI will be provided for each treatment group, together with first subsequent treatment or death free survival rates and second subsequent treatment or death free survival rates at clinically relevant time points (e.g., at 6 months and 12 months). K-M curves will also be provided, with tick marks to identify censored observations.

The HR and its 95% CI will be estimated (HR less than 1.0 favours Olaparib) from a Cox Proportional Hazards model (with ties = Efron and the stratification factors as covariates) and the CI calculated using a profile likelihood approach.

Kaplan-Meier plots of TFST and TSST will be presented by cohort and treatment arm. Summaries of the number and percentage of patients who received a subsequent anti-cancer therapy, a summary table of first (and second) subsequent anti-cancer therapies by treatment

arm will be provided, as well as response to first subsequent anti-cancer therapy by treatment arm (where available).

These endpoints will be analysed at the time of the primary analysis of PFS and at the time of any further OS analysis.

4.2.7.4 Time to study treatment discontinuation or death

Time to study treatment discontinuation or death will be compared between the two treatment groups using a two-sided stratified log-rank test as defined for the previous endpoints. The HR and its 95% CI will be estimated (HR less than 1.0 favours Olaparib) using the same methods as described for PFS and OS. A K-M analysis will also be performed (median time to study treatment discontinuation or death and 95% CI, and treatment discontinuation free survival rates at clinically relevant time points e.g., at 6 months and 12 months) and K-M curves presented.

This endpoint will be analysed at the time of the primary analysis of PFS and at the time of any further OS analysis.

4.2.7.5 Time to earliest progression by RECIST or CA-125 or death

Time to progression by RECIST 1.1, CA-125 or death will be performed at the same time as the primary analysis of PFS and will use the same methodology and model.

The number and percentage of patients reporting a CA-125 progression, an objective RECIST 1.1 progression and both a CA-125 and/or objective RECIST progression will be tabulated.

4.2.7.6 Best overall RECIST Response (BoR)

For each treatment arm, Best Overall RECIST Response (BoR) derived programmatically from investigator data will be summarised by n (%) for each category (CR, PR, SD, NED, PD, NE). No formal statistical analyses are planned.

4.2.8 Patient Reported Outcome data

Table 7 summarises the formal analyses and pre-planned sensitivity analysis for the HRQoL endpoints.

Table 7. Formal and sensitivity analysis for PRO endpoints

Endpoints analysed	Notes
FACT-O TOI score	Primary subscale Sensitivity analyses 1. Deviation from Missing At Random (MAR) assumption, if required
PWB score SWB score Functional Well-Being (FWB) score EWB score OCS score FACT-O total score FACT-G total score	Exploratory subscales

FACT-O questionnaire compliance (overall compliance and by visit compliance) will be summarised for each cohort and treatment group.

For the FACT-O questionnaire compliance table and subsequent tables of descriptive statistics on a visit basis, in addition to individual visits being presented, additional entries will be given for End of treatment and 30 day follow up visits (see section 3.3.7 for description of visit windowing).

Reasons for missing data will be summarised. Missing data will be explored in order to assess the assumption of data being MAR. Raw scores and change from baseline scores will be summarised descriptively for all visits and all scores/subscales.

The following analyses will be performed for the primary PRO endpoint of interest (TOI score):

- Change from baseline score - TOI

Change from baseline in TOI score will be analysed using a mixed model for repeated measures (MMRM) analysis of the change from baseline in TOI scores for each visit. Only visits with at least 25% of non-missing values in both treatment arms (calculated separately by treatment arm) are included in the model. The End of Treatment visit and follow-up visits will be excluded from this analysis. The analysis will compare the average treatment effect from the point of randomisation for all remaining on-treatment visits.

The MMRM model will include patient, treatment, visit, treatment by visit interaction and randomisation stratification factors (i.e., use of prior bevacizumab [yes versus no] and number of prior regimens of platinum-containing chemotherapy [≤3 versus ≥4 regimens]) as

explanatory variables and the baseline TOI and baseline visit interaction as covariates. An overall adjusted mean estimate will be derived that will estimate the average treatment effect over visits giving each visit equal weight. For the overall treatment comparison, adjusted mean estimates per treatment group and corresponding 95% CIs will be presented along with an estimate of the treatment difference, 95% CI and p-value. The treatment by visit interaction will remain in the model regardless of significance. An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom.

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

For each treatment and visit, the adjusted (least squares) mean estimates, corresponding 95% CIs, estimates of the treatment difference, corresponding 95% CIs and p-values will be presented. The clinical relevance of the estimated differences will be assessed using the MID (Yost et al, 2005).

Although the MMRM is generally robust to deviations from the assumption of MAR (Bell et all, 2014), this will be assessed after the study is unblinded and if a sensitivity analysis is required this may be carried out after unblinding and included in a CSR addendum or as data on file. To assess this, plots of the mean TOI stratified by treatment and dropout time will be reviewed to confirm the MAR assumption that the trajectory for patients who dropped out in each treatment group is similar to those patients observed in their own treatment group. Plots will be produced for the overall data and may before each stratification category.

- Proportion of patients with a PRO response - TOI

The proportion of patients with a PRO response will be reported and compared across treatment groups at key visits using the Cochran-Mantel-Haenszel test to account for the randomisation stratification factors unless there is excessive missing data (data for fewer than 20 patients available in a cohort at that timepoint).

- Best overall response - TOI

Best overall response will be summarised as number and proportion of patients with each response level by treatment group. The proportion with a best overall response of improved (against any other non-missing response) will be compared using the Cochran-Mantel-Haenszel test to account for the randomisation stratification factors unless there is excessive missing data (data for fewer than 20 patients available in a cohort at that timepoint).

- PRO deterioration-free survival - TOI

PRO deterioration-free survival will be analysed using Cox regression accounting for the randomisation stratification factors to compare treatment groups. Median time until PRO

deterioration-free survival will be reported for each group and for the difference between groups with 95% CIs. K-M curves will be used to show unadjusted time to deterioration.

4.2.9 Time to gastrointestinal symptom deterioration

Time to GI symptom deterioration will be compared between treatment groups using a proportional hazards cox regression (PH) model and presented as a HR with corresponding 95% CI and p-value and displayed using K-M curves. The PH model will be adjusted for baseline TOI score.

Descriptive summary statistics will be produced including graphical displays of change over time in GI symptoms by treatment group.

4.2.10 Time to progressive disease for patients with significant symptoms at baseline

Given that there are a sufficient number of patients, time to protocol defined progression disease for patients with significant symptoms at baseline (as defined in Section 3.2.2.8) will be compared between treatment groups using a PH model and presented as a HR with corresponding 95% CI and p-value and displayed using K-M curves. Graphical summaries will also be displayed.

4.2.11 Assessing impact of post-progression chemotherapy on HRQoL

Descriptive summary statistics (e.g. n, mean, median, sd, min, max) will be produced from the TOI score post-progression and after starting subsequent chemotherapy questionnaires overall and by treatment arm. Graphical displays of change over time in TOI score post-progression (figures will only show post-progression values, starting at the point of treatment discontinuation) may be produced by treatment arm.

4.2.12 Multiplicity

No multiplicity adjustment will be performed for the defined analyses.

4.2.13 Sensitivity analyses

The following sensitivity analyses may be performed for the primary efficacy endpoint (PFS):

- Time assessment bias: This refers to the possibility that the PFS effect is partly an artefact of one treatment arm being assessed more frequently. It applies in situations where scans are not performed at the protocol-scheduled time points.

  In this sensitivity analysis, the midpoint between the time of progression and the previous evaluable RECIST assessment (using the final date of the assessment) will be analysed using a stratified log-rank test, as described for the primary analysis of PFS. Note that midpoint values resulting in non-integer values should be rounded down. For patients whose death was treated as a PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust to even highly asymmetric assessment schedules (Sun et al, 2010). To support this

analysis, the mean of patient-level average inter-assessment times will be tabulated for each treatment.

- Attrition bias: An assessment of whether the rate and nature of censoring has resulted in bias.

  Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following two, or more, non-evaluable tumour assessments will be included. In addition, and within the same sensitivity analysis, patients who take subsequent therapy (note that for this analysis radiotherapy is not considered a subsequent anti-cancer therapy) prior to their last evaluable RECIST assessment or progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a K-M plot of the time to censoring using the PFS data from the primary analysis and where the censoring indicator of the PFS analysis is reversed (Schemper, 1996).

- Deviation bias: An assessment of excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients in either treatment group:

  - Did not have the intended disease or indication or
  - Did not receive any randomised therapy or
  - Did not receive a prior PARPi or unknown if received PARPi (i.e. blinded PARPi prior to protocol amendment 1)
  - Received more than one prior PARPi (for more than 15 days), including blinded PARPi/placebo.

  Deviation bias will be assessed by repeating the PFS analysis excluding patients with deviations that may affect the efficacy of trial therapy.

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

- Adjustment for additional prognostic factors:

  A sensitivity analysis will be performed using Cox proportional hazards regression to determine if adjustment for covariates – other than the stratification variables – will modify the conclusions of the primary PFS analysis. Variables to be considered are the stratification factors (i.e., use of prior bevacizumab [yes versus no] and number of prior regimens of platinum-containing chemotherapy [≤3 versus ≥4 regimens]), as well as number of prior lines of chemotherapy, whether a patient had a CR or PR with the most recent course of chemotherapy, the duration of platinum-free interval, defined as the last day of one course of platinum to the detection of relapse, following penultimate course of chemotherapy, duration of previous PARPi exposure, prior

surgery of ovarian cancer only if this occurred immediately before most recent line of chemotherapy.

- Cohort adjustment:

  If enough patients, based on the percentage of available data, are found to be allocated to the incorrect *BRCA1/2* status (correct based on data at the time of randomisation but later found to be incorrect), a per-protocol or sensitivity analysis may be carried by repeating the PFS analysis and either removing these patients entirely or re-allocating them to their alternative cohort.

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

For the PRO primary endpoint (FACT-O TOI), the following sensitivity analysis may be performed:

- Deviation from the MAR assumption

  The assessment of the MAR assumption will be made after database lock and unblinding, therefore any sensitivity analysis may be carried out separately to the main analysis and included as a CSR addendum or data on file. See Section 4.2.8 for further details.

For the gastrointestinal symptom deterioration endpoint, the following sensitivity analysis may be performed:

- A time to GI symptom deterioration analysis that classifies deterioration as a response of ≥2 (‘some-what’) instead of a response of ≥1 (‘a little bit’).

4.2.14 Safety

The following sections describe the planned safety summaries for AEs, SAEs, deaths, laboratory parameters and vital signs.

4.2.14.1 Adverse events

Only TEAEs, i.e., adverse events developing or worsening in severity (from baseline severity) on or after the first day of Olaparib or placebo administration up to and including 30 days after the last dose of Olaparib or placebo (defined as the treatment period) will be summarised.

Any AE occurring before the first dose of Olaparib or placebo and AEs occurring 30 days after last dose will be listed only and not included in the summaries.

Any AEs that occur after a patient has received further therapy for cancer (following discontinuation of Olaparib or placebo) will be flagged in the data listings.

The number of patients experiencing each AE will be summarised by treatment group by the MedDRA system organ class, MedDRA preferred term and worst CTCAE grade. The number and percentage of patients with AEs in different categories (i.e., causally related, maximum reported intensity, CTCAE grade ≥3) will be summarised by treatment group, and events in each category will be further summarised by MedDRA system organ class and preferred term. The number of patients experiencing AEs leading to dose modification will also be summarised by treatment group.

For the purposes of understanding the causal relationship of AEs, K-M curves for time to onset of first AE will be plotted. Time to onset of AE is defined as date of first AE – date of first dose+1 for those patients with an AE and min([date of last dose + safety follow-up period], OS date, DCO) – date of first dose + 1 for those without an AE (i.e. censored patients).

AE event rates (i.e., exposure adjusted incidence of patients with an AE) will also be summarized by preferred term within each system organ class. AE event rates will be calculated as the number of patients with at least one event divided by the total time at risk, i.e. the number of days of exposure to drug (including 30 days after the last treatment dose) summed over all patients. This rate is then multiplied by 1000 to present events per 1000 patient years. SAEs and deaths will be summarised.

Grouped AEs relevant to Olaparib will be also summarised, with the groups as follows:

- Anaemia
- Neutropenia
- Thrombocytopenia
- Nausea
- Vomiting
- Fatigue/Asthenia

Adverse events of special interest

Adverse events of special interest for Olaparib are:

- MDS/AML
- New primary malignancy (other than MDS/AML)
- Pneumonitis

Summaries of the above mentioned AESIs will include number (%) of patients who have:

- At least one AESI
- At least one AESI causally related to study treatment
- At least one AESI leading to discontinuation of study treatment

4.2.14.2 Laboratory evaluations

Laboratory data obtained until the 30-day safety follow-up visit will be included in the summary tables. Absolute values and change from baseline for all continuous haematology and clinical chemistry laboratory parameters will be summarised for each cohort and by treatment arm and visit.

Shift tables for change in grade from baseline by treatment group will be produced. The laboratory parameters for which CTC grade shift outputs will be produced are:

- Haematology: haemoglobin, lymphocytes (absolute count), neutrophils (absolute count), platelets, activated partial thromboplastin time, and international normalized ratio
- Clinical chemistry: Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), alkaline phosphatase, amylase, lipase, total bilirubin, albumin, magnesium (hypo- and hyper-), sodium (hypo- and hyper-), potassium (hypo- and hyper-), corrected calcium (hypo- and hyper-), glucose (hypo- and hyper-), gamma-glutamyl transferase or gamma-glutamyl transpeptidase (GGT), creatinine

Hy's law

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

- Elevated ALT, AST, and total bilirubin during the study
    - ALT ≥3x – ≤5x, >5x – ≤8x, >8x – ≤10x, >10x – ≤20x, and >20x upper limit of normal (ULN) during the study
    - AST ≥3x – ≤5x, >5x – ≤8x, >8x – ≤10x, >10x – ≤20x, and >20x ULN during the study
    - Total bilirubin ≥2x – ≤3x, >3x – ≤5x, >5x ULN during the study
    - ALT or AST ≥3x – ≤5x, >5x – ≤8x, >8x – ≤10x, >10x – ≤20x, >20x ULN during the study
    - ALT or AST ≥3x ULN and total bilirubin ≥2x ULN during the study (potential Hy's law): the onset date of ALT or AST elevation should be prior to or on the date of total bilirubin elevation

Liver biochemistry test results over time for patients with elevated ALT or AST (i.e., ≥ 3 × ULN) and elevated total bilirubin (i.e., ≥ 2 × ULN) at any time will be plotted.

Individual patient data where ALT or AST plus total bilirubin are elevated at any time will be listed (i.e., ALT and/or AST ≥3 × ULN and total bilirubin ≥2 × ULN, at any time).

4.2.14.3 Vital signs

Vital sign data obtained up until the 30-day safety follow-up visit will be included in the summary tables.

Vital signs will be summarised by cohort and treatment group using descriptive statistics (mean, median, standard deviation, minimum, maximum and number of patients).

4.2.15 Exploratory analysis

CCI

[redacted]

4.3 Subgroup analysis

Subgroup analyses will be conducted, in each cohort, comparing PFS between Olaparib and placebo in the following subgroups of the FAS:

- The use of prior bevacizumab [yes versus no]
- The number of prior regimens of platinum-containing chemotherapy [≤3 versus ≥4 regimens])
- Whether a patient had a CR or PR with the most recent course of chemotherapy at baseline,
- Duration of platinum-free interval following penultimate course of chemotherapy at baseline [≥6 months to <12 months, ≥12 months to <18 months, ≥18 months]
- Duration of previous PARPi exposure [*BRCA1/2* (+ve): <18 months, ≥18 months; *BRCA1/2* (-ve): <12 months, ≥12 months]
- Prior surgery for ovarian cancer only if this occurred immediately before or during current chemotherapy [Yes, No].

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

No adjustment to the significance level for testing will be made [redacted] and may only be supportive of the primary analysis of PFS.

For each subgroup level, the HR and 95% CI will be calculated from an un-stratified Cox proportional hazards model with treatment as the only covariate. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties, using the by statement to obtain a HR and 95% CI for each subgroup level separately.

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

An analysis will not be performed if there are too few events available for a meaningful analysis of a particular subgroup (so if there are less than 20 events in a subgroup). In this case, only descriptive summaries will be provided.

5. INTERIM ANALYSIS

In the *BRCA1/2* (+ve) cohort an interim analysis for futility will be performed after 50% of the target PFS events (i.e., after 43 events). Based upon the assumed accrual and event rate it is estimated that this will occur after approximately 29 months. Using a conditional power non-binding futility analysis, the cohort may stop for futility if the hazard ratio is > 1.056. Under the null hypothesis the probability of stopping for futility is 0.433 and under the alternative hypothesis it is 0.01.

In the *BRCA1/2* (-ve) cohort an interim analysis for futility will be performed after 50% of the target PFS events (i.e. after 37 events). Based upon the assumed accrual and event rate it is estimated that this will occur after approximately 30 months. Using a conditional power non-binding futility analysis, the cohort may stop for futility if the hazard ratio > 1.02. Under the null hypothesis the probability of stopping for futility is 0.477 and under the alternative hypothesis it is 0.02.

Results from the planned interim analyses will be shared with the IDMC who will make a recommendation on continuing or stopping a cohort.

The interim analysis for futility will be performed separately and at different times for each cohort as each pre-defined number of target events is reached.

6. CHANGES OF ANALYSIS FROM PROTOCOL

The PRO analysis set has been amended to only include patients present in the FAS with a baseline PRO assessment present.

The definition of definitive deterioration for TOI score and the related PRO deterioration-free survival analysis has been amended to match the advice of a PRO subject matter expert.

An AE summary for grouped AEs relevant to Olaparib has been added.

The following QoL related analyses have been added:

- Time to gastrointestinal (GI) symptom deterioration
- Time to progressive disease for patients with significant symptoms at baseline
- Assessing the impact of post-progression chemotherapy on HRQoL

CCI

7. REFERENCES

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

Schemper, M & Smith T.L. A note on quantifying follow-up in studies of failure time. Controlled Clinical Trials 17:343-346.

Yost, K. J. & D. T. Eton (2005) Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof, 28, 172-91.

Rustin GJ. Use of CA-125 to assess response to new agents in ovarian cancer trials. J Clin Oncol. 2003; 21(10 Suppl):187s-193s.

Bell, M. L. & D. L. Fairclough (2014) Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res, 23, 440-59.

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