AstraZeneca

A Phase III, Randomised, Double Blind, Placebo Controlled, Multicentre Study of Olaparib Maintenance Monotherapy in Patients with BRCA Mutated Advanced (FIGO Stage III-IV) Ovarian Cancer following First Line Platinum Based Chemotherapy

A Phase III, Randomised, Double Blind, Placebo Controlled, Multicentre Study of Olaparib Maintenance Monotherapy in Patients with BRCA Mutated Advanced (FIGO Stage III-IV) Ovarian Cancer following First Line Platinum Based Chemotherapy

Study Statistician

A Phase III Randomised, Double Blind, Placebo Controlled, Multicentre Study of Olaparib Maintenance Monotherapy in Platinum Sensitive Relapsed BRCA Mutated Ovarian Cancer Patients who are in Complete or Partial Response Following Platinum based Chemotherapy

Global Product Statistician

TABLE OF CONTENTS   PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

The following abbreviations and special terms are used in this study Statistical Analysis Plan.

Abbreviation or special term	Explanation
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
Baseline	Refers to the most recent assessment of any variable prior to dosing with study treatment
BD	Twice daily
BICR	Blinded independent central review
BoR	Best overall RECIST response
BP	Blood pressure
BRCA	Breast cancer susceptibility gene
BRC*Analysis*®	Gene sequencing and large rearrangement analysis for Hereditary Breast and Ovarian Cancer, registered trademark of Myriad Genetics, Inc
BRCA mutation or BRCAm	Breast cancer susceptibility gene mutation (see gBRCA mutation or gBRCAm)
CA-125	Cancer Antigen – 125
cfDNA	Circulating free DNA
CLIA	Clinical Laboratory Improvement Amendments
CCR	Clinical complete response. 'Response' is used throughout the protocol and refers to patients being, in the opinion of the investigator, in clinical complete response or partial response on the post-treatment scan. Clinical complete response is defined as no evidence of either RECIST measurable or non-measurable disease on the post-treatment scan and a normal CA-125. Partial response is defined as ≥30% reduction in tumour volume demonstrated from the start to finish of chemotherapy OR no evidence of RECIST measurable disease on the post-treatment scan with a CA-125 which has not decreased to within the normal range.
CI	Confidence interval
CR	Complete response
CRF / eCRF	Case Report Form (electronic)
CRO	Clinical Research Organisation
CSP	Clinical Study Protocol

Abbreviation or special term	Explanation
CSR	Clinical Study Report
CT	Computed tomography
CTC / CTCAE	Common Terminology Criteria for Adverse Event
DAE	Discontinuation of investigational product due to adverse event
DBP	Diastolic blood pressure
DCO	Data cut-off
DCR	Disease control rate
DNA	Deoxyribonucleic acid
DOR	Duration of response
d.p.	Decimal places
ECG	Electrocardiogram
E-code	Enrolment code (allocated by IVRS/IWRS)
ECOG	Eastern Cooperative Oncology Group: A performance status using scales and criteria to assess how a patient's disease is progressing
EQ-5D-5L / EQ-5D	EuroQoL five dimensions, five level (EQ-5D-5L) health state utility index
EWB	Emotional well being
FACIT	Functional Assessment of Chronic Illness Therapy
FACT-O	Functional Assessment of Cancer Therapy – Ovarian: A multidimensional questionnaire for patients with ovarian cancer
FAS	Full analysis set
FDA	Food and Drug Administration
FIGO	International Federation of Gynecology and Obstetrics
FSI	First subject in
FWB	Functional well being
gBRCA	Germline BRCA
gBRCA mutation or gBRCAm	The term "gBRCA mutation" is used to refer to a germline BRCA1 or BRCA2 mutation classified as "deleterious" or "suspected deleterious" in accordance with the American College of Medical Genetics and Genomics recommendations for standards for interpretation and reporting of sequence variants.
gBRCA wt	gBRCA wildtype
GCIG	Gynecologic Cancer Intergroup
GOG	Gynecologic Oncology Group

Abbreviation or special term	Explanation
HR	Hazard ratio
HRQoL	Health-related quality of life
ICR	Independent Central Review
ICU	Intensive care unit
IDMC	Independent Data Monitoring Committee
IP	Intraperitoneal
IPCW	Inverse probability of censoring weighting
ITT	Intention to treat
IV	Intraveneous
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LD	Longest Diameter
LPLV	Last patient last visit
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milli-gram
MMRM	Mixed model for repeated measures
MRI	Magnetic resonance imaging
MTP	Multiple testing procedure
NA	Not applicable
NC	Not calculable
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
NTL	Non-target lesions
OAE	Other significant adverse event (see definition in Section 3.4.1)
ORR	Objective response rate
OS	Overall survival
PARP	Polyadenosine 5'diphosphoribose [poly (ADP ribose)] polymerisation
PD	Progressive disease
PFS	Progression free survival

Abbreviation or special term	Explanation
PFS2	Time from randomisation to second progression
PID	Percentage intended dose
p.o.	Per os (by mouth, orally)
PR	Partial response
PRO	Patient reported outcome
PWB	Physical well being
QoL	Quality of life
QSR	Quality System Regualtion
QTcB	QT interval (corrected for heart rate using Bazett's correction)
QTcF	QT interval (corrected for heart rate using Fridericia's correction)
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumours. This study will use modified RECIST version 1.1
REML	Restricted maximum likelihood
RPSFT	Rank preserving structural failure time
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SWB	Social well being
Study treatment	Olaparib or matching placebo
SBP	Systolic blood pressure
tBRCA mutation or tBRCAm	The term "tBRCA mutation" is used to refer to a somatic tumour BRCA1 or BRCA2 mutation classified as "deleterious" or "suspected deleterious" in accordance with the American College of Medical Genetics and Genomics recommendations for standards for interpretation and reporting of sequence variants.
TDT	Time from randomisation to study treatment discontinuation or death
TFST	Time from randomisation to first subsequent therapy or death
TL	Target lesions
TOI	Trial Outcome Index
TSST	Time from randomisation to second subsequent therapy or death
UCL	Upper confidence limit

Abbreviation or special term	Explanation
ULN	Upper limit of normal
VUS	Variant of uncertain significance
wt	Wildtype (patients without evidence of BRCA1 or BRCA2 deleterious or suspected deleterious mutations)

1 STUDY DETAILS

1.1 Study Objectives

Primary:

To determine the efficacy by progression free survival (PFS) using investigator assessment according to modified Response Evaluation Criteria In Solid Tumours (RECIST) 1.1) of olaparib maintenance monotherapy compared to placebo in Breast Cancer susceptibility gene (BRCA) mutated high risk advanced ovarian cancer patients who are in clinical complete response or partial response following first line platinum based chemotherapy.

Secondary:

1.   To determine the efficacy of olaparib maintenance monotherapy compared to placebo in BRCA mutated high risk advanced ovarian cancer patients who are in clinical complete response or partial response following first line platinum based chemotherapy by assessment of overall survival (OS), time to earliest progression by RECIST or Cancer Antigen-125 (CA-125), or death, and time from randomisation to second progression (PFS2).

2.   To compare the effects of olaparib maintenance monotherapy compared to 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 BRCA mutated high risk advanced ovarian cancer patients who are in clinical complete response or partial response following first line platinum based chemotherapy.

3.   To assess efficacy of olaparib in patients identified as having a deleterious or suspected deleterious variant in either of the BRCA genes using variants identified with current and future BRCA mutation assays (gene sequencing and large rearrangement analysis).

4.   To determine the efficacy of olaparib maintenance monotherapy compared to placebo in BRCA mutated high risk advanced ovarian cancer patients who are in clinical complete response or partial response following first line platinum based chemotherapy by assessment of 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 (TDT).

Safety:

1.   To assess the safety and tolerability of olaparib maintenance monotherapy in BRCA mutated high risk advanced ovarian cancer patients who are in clinical complete response or partial response following first line platinum based chemotherapy.

Exploratory:

1.

2.   To explore the impact of treatment and disease state on health state utility by EuroQoL five dimensions, five level (EQ-5D-5L).

3.   To explore the impact of treatment and disease on resource use.

4.   To explore the effects of olaparib maintenance monotherapy as assessed by the individual domains of the trial outcome index (TOI) of the Functional Assessment of Cancer Therapy – Ovarian (FACT-O).

5.   To explore the efficacy of olaparib by assessment of overall survival (OS) adjusting for the impact of spontaneous switching [outside of study design] to Polyadenosine 5'diphosphoribose [poly (ADP ribose)] polymerisation (PARP) inhibitors or other potentially active investigational agents.

6.

7.   Future exploratory research into factors that may influence development of cancer and/or response to study treatment (where response is defined broadly to include efficacy, tolerability or safety) may be performed on the collected and stored archival tumour samples that were mandatory for entry into the study or on optional tumour biopsy samples collected during the course of the study.

8.   To collect and store Deoxyribonucleic acid (DNA) (according to each country's local and ethical procedures) for future exploratory research into genes/genetic variation that may influence response (i.e., distribution, safety, tolerability and efficacy) to study treatments and or susceptibility to disease (optional).

The exploratory analyses may not be reported in the Clinical Study Report (CSR). If not, they will be reported separately.

1.2   Study Design

This is a Phase III, randomised, double-blind, placebo-controlled, multi-centre study to assess the efficacy of olaparib maintenance monotherapy in high risk advanced ovarian cancer patients (including patients with primary peritoneal and / or fallopian tube cancer) with BRCA mutations [documented mutation in BRCA1 or BRCA2] that is predicted to be deleterious or suspected deleterious (known or predicted to be detrimental/lead to loss of function) who have responded following first line platinum based chemotherapy.

Patients will be randomised (using an Interactive Voice Response System (IVRS) / Interactive Web Response System (IWRS)) in a 2:1 ratio to the treatments as specified below:

- olaparib tablets p.o. 300 mg twice daily

- placebo tablets p.o. twice daily

In addition a separate China cohort will be randomised in the same way (see Section 3.6 for further details).

Eligible patients will be those patients with newly diagnosed, histologically confirmed, high risk advanced (FIGO stage III-IV) BRCA mutated high grade serous or high grade endometrioid (based on local histopathological findings) ovarian cancer, primary peritoneal cancer and / or fallopian-tube cancer who are in clinical complete response or partial response following completion of first line platinum-based chemotherapy. Patients who re-present following prior diagnosis at an earlier stage of disease are not eligible. Stage III patients should have had one attempt at optimal debulking surgery (upfront or interval debulking). Stage IV patients must have had either a biopsy and/or upfront or interval debulking surgery.

Patients must have completed a minimum of six and maximum of nine treatment cycles of first line platinum-based therapy (e.g., carboplatin or cisplatin) before randomisation to the study and should be in the opinion of the investigator in clinical complete response or partial response. However, if platinum based therapy must be discontinued early as a result of toxicities specifically related to the platinum regimen, patients must have received a minimum of four cycles of the platinum regimen.

Patients must not have received bevacizumab (either in combination or as maintenance therapy following combination therapy) or any investigational agent during their first line course of treatment.

Patients known to have deleterious or suspected deleterious germline BRCA mutation/s (gBRCAm i.e., blood) prior to randomisation can enter the study based on this result. The result must be made available to AstraZeneca. In addition the patients must consent to provide 2 blood samples. One sample will be used for a confirmatory Myriad gBRCA test post randomisation using the current commercial Myriad BRC*Analysis*® (gene sequencing and large rearrangement analysis),

Patients with unknown BRCA status must consent to provide 2 blood samples for germline BRCA testing and follow all local ethical procedures for genetic testing. One sample will be used to test for BRCA mutations using the current commercial Myriad BRC*Analysis*® test prior to study entry. These samples will be required for the study even if the patients are found not to have a BRCA mutation. When the result from the Myriad test indicates the

patient does have a deleterious or suspected deleterious BRCA mutation, the patient can be randomised into the study.

Patients will be randomised within 8 weeks after their last dose of chemotherapy (last dose is the day of the last infusion).

Randomisation will be stratified by:

- Response to first line platinum chemotherapy (in the opinion of the investigator, clinical complete response (CCR) or partial response (PR)).

Following randomisation patients in both treatment arms will attend clinic visits weekly for the first 4 weeks of treatment (Days 8, 15, 22 and 29). Patients will then attend clinic visits every 4 weeks whilst on study. Patients should continue to receive study treatment for up to two years or until objective radiological disease progression as per RECIST as assessed by the investigator and as long as in the investigator's opinion they are benefiting from treatment and they do not meet any other discontinuation criteria (see Section 5.8 of the Clinical Study Protocol (CSP)). Patients who continue to have evidence of disease that remains stable (i.e. no evidence of disease progression) at two years or those who have progressed may continue to receive study treatment if, in the opinion of the investigator, it is in the patient's best interest. However, if at two years the patient has no evidence of disease, study treatment should be discontinued.

Once a patient has progressed and discontinued study treatment, clinic visits will be reduced to every 12 weeks. Following discontinuation of study treatment, further treatment will be at the discretion of the investigator. Any further systemic anti-cancer treatment data will be collected until death, loss to follow-up or withdrawal of consent. In addition to their regular 12 weekly contact, patients will be contacted in the 7 days following a specified date (data cut off date) for each survival analysis. Assessments will be performed as described in Tables 1, 2, 3 and 4 of the CSP. Patients in both treatment arms will have tumour assessments according to RECIST at baseline and every 12 weeks (±1week) up to 156 weeks and then every 24 weeks (±1 week) relative to date of randomisation, until objective radiological disease progression according to RECIST. All Computed tomography (CT)/ Magnetic resonance imaging (MRI) scans will be sent to an AstraZeneca appointed Clinical Research Organisation (CRO) for blinded independent central review. All treatment decisions will be based on site assessment of scans. After the primary PFS analysis, central review of scans will no longer be required and investigators will be advised when to stop sending copies of the scans to the CRO conducting the central review. Ongoing collection of site review tumour assessment is required and must be recorded in the electronic case report form (eCRF).

RECIST will be modified to assess patients with CR at entry who will be assessed as having no evidence of disease (NED) unless they have progressed based on the appearance of new lesions.

Any patient who discontinues study treatment for reasons other than objective radiological progression should continue to undergo scheduled objective tumour assessments according to the study plan (see Tables 3 and 4 of the CSP) in order to assess objective radiological progression of disease. Failure to do so may result in bias to the study results.

Once a patient has progressed, the patient will be followed for second progression (PFS2) every 12 weeks and then survival until the final analysis.   Patients will be contacted in the week following last patient last visit (LPLV) for each analysis of survival.

The primary analysis will be based on investigator assessment of disease progression by RECIST; however, a sensitivity analysis will be performed using blinded independent central review (BICR). All efficacy variables including overall survival will be analysed at the time of the primary analysis (providing sufficient events are available to make the analyses meaningful).

The overall study design is shown in Figure 1 of the CSP, the screening plan is shown in Figure 2 of the CSP and the study flow chart in Figure 1 and Figure 2 below. Screening schedules are detailed in Tables 1 and 2 of the CSP. The study schedule is detailed in Tables 3 and 4 of the CSP.

Figure 1   **Study Flow Chart Up to 108 Weeks on Treatment**
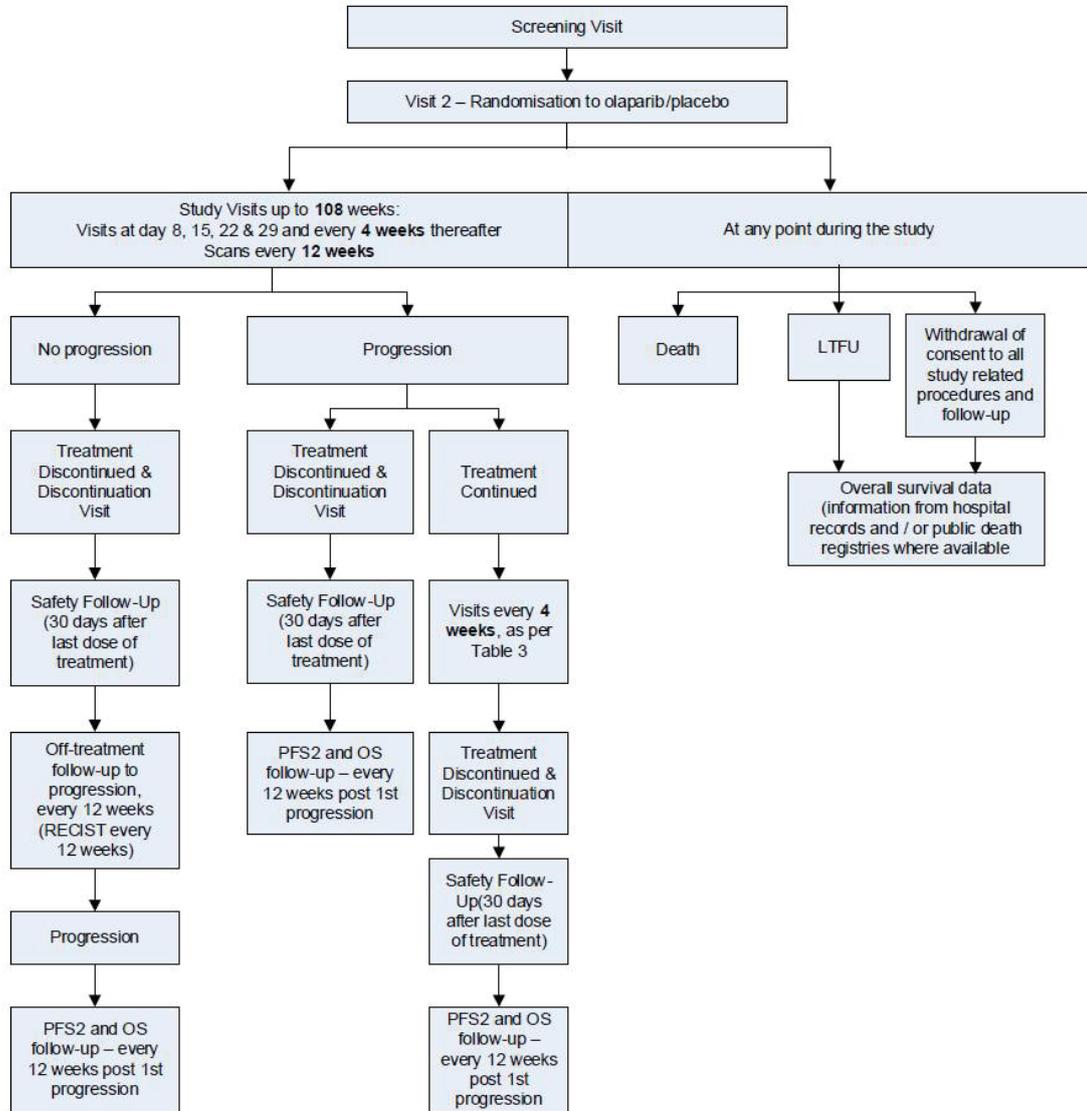
Figure 2   Study Flow Chart At 108 Weeks on Treatment
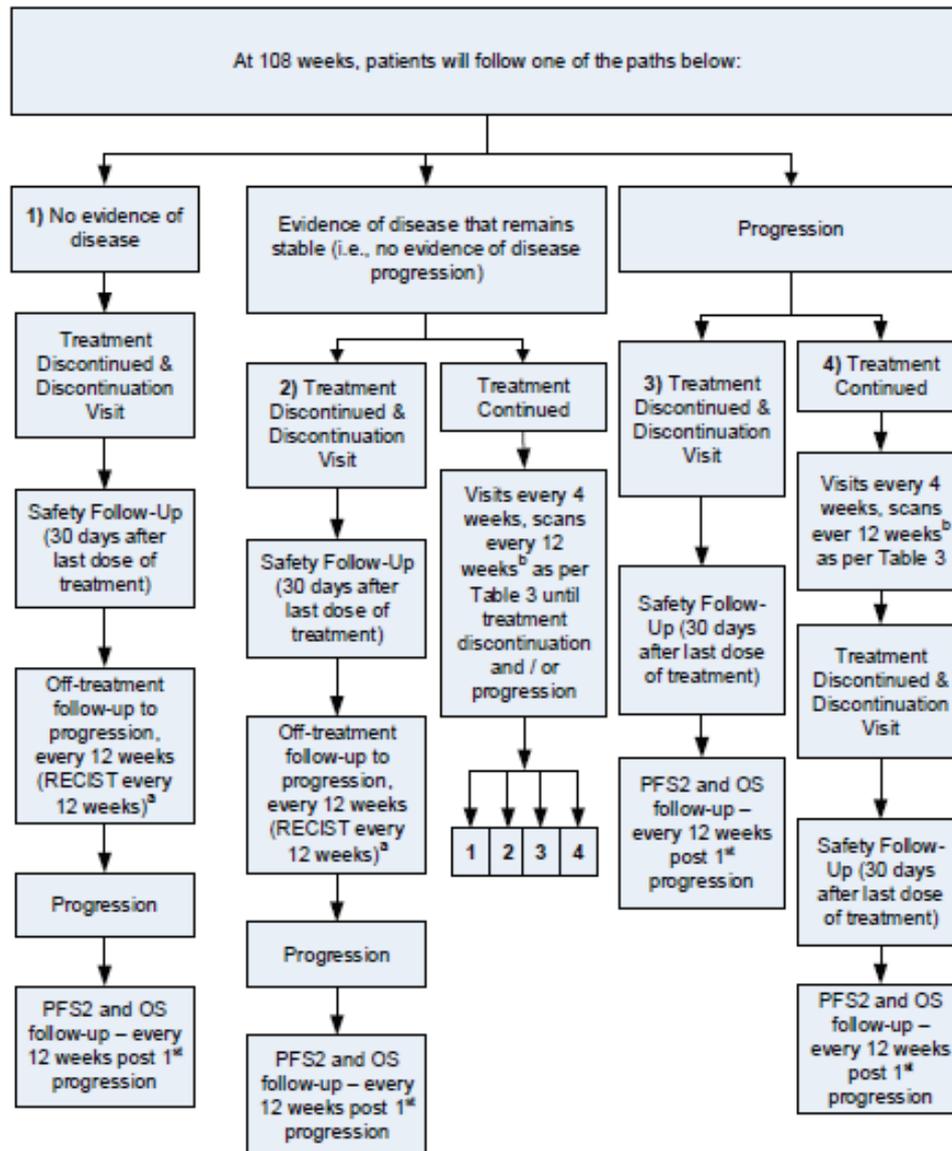
[a]   Patients off treatment: Off-treatment follow-up visits and scans will continue to be conducted every 12 weeks (±1 week) up to 156 weeks (3 years), then every 24 weeks (±1 week) relative to date of randomization.

[b]   Patients continuing treatment: Treatment visits will continue every 4 weeks (±3 days) up to 156 weeks (3 years), then every 12 weeks (±1 week) relative to date of randomization. Scans will continue every 12 weeks (±1 week) up to 156 weeks (3 years), then every 24 weeks (±1 week) relative to date of randomization.

18

1.3 Number of Subjects

In total 206 PFS events in the study would have 90% power to show statistically significant PFS at the 2-sided 5% level if the assumed true treatment effect were hazard ratio (HR) 0.62; this translates to a 8 month benefit in median PFS over 13 months on placebo (estimated from data reported by Alsop et al 2012) if PFS is exponentially distributed.

Approximately 344 patients will be recruited (2:1 ratio) so that data maturity for the PFS analysis is approximately 60%.

Assuming 18 months non-linear recruitment, 206 PFS events are expected to occur approximately 36 months after first subject in is enrolled in the study (FSI).   At this time, an analysis of OS will also be performed.

PFS will be analysed when approximately 196 events have occurred or after the last patient randomised has had the opportunity to have been on the study for at least 36 months, whichever comes first. No further analyses of PFS are planned beyond this point unless requested by Health Authorities.

A further analysis of OS may be performed at approximately 60% maturity (~206 events); this is anticipated to occur approximately 80 months after FSI. Assuming that the true OS treatment effect is 0.85 and this point estimate for the HR of 0.85 was observed, the 95% UCL for the HR would be 1.13. Note that these estimates are based on the assumption that no confounding will occur. AstraZeneca considers there will be potential confounding of OS data due to availability of PARP inhibitors for BRCA mutated ovarian cancer patients during follow up in this study, which are likely to disproportionately affect OS in one arm of the study (placebo-treated patients).

2   ANALYSIS SETS

2.1   Definition of Analysis Sets

Two main analysis sets are defined for this study for the analysis of patients who have been included in the study as part of the global enrolment.

Full analysis sets

Intention to treat (ITT): The primary statistical analysis of the efficacy of olaparib will include all patients who are randomised as part of the global enrolment. The primary analysis will compare the treatment groups on the basis of randomised treatment, regardless of the treatment actually received. Patients who were randomised as part of the global enrolment but did not subsequently go on to receive study treatment are included in the Full Analysis Set (FAS). Therefore, all efficacy and HRQoL data will be summarised and analysed using the

FAS on an intention to treat (ITT) basis. No centres will be provided with all of the treatment codes within that centre until completion of the study following the final analysis of overall survival.

Safety analysis set

All patients who received at least one dose of randomised investigational product, olaparib or placebo and are part of the global enrolment will be included in the safety analysis set (regardless of whether that was the randomised therapy intended or indeed whether, in rare cases, they received therapy without being randomised). Throughout the safety results sections, erroneously treated olaparib patients (those randomised to olaparib but actually received placebo) will be accounted for in the placebo treatment group. Erroneously treated placebo patients (those randomised to placebo but actually received olaparib) will be accounted for in the olaparib treatment group. Patients receiving treatment from more than one treatment arm will be accounted for based upon their initial treatment started

Table 1   Summary of Outcome Variables and Analysis Populations

Outcome Variable	Populations
Efficacy Data	
- PFS	FAS (ITT)
- OS, PFS2, TFST, TSST, TDT, symptom/HRQoL endpoints	FAS (ITT)

Table 1   **Summary of Outcome Variables and Analysis Populations**

Outcome Variable	Populations
Study Population/ Demography Data	
- Demography characteristics (e.g. age, sex etc)	FAS (ITT)
- Baseline and disease characteristics	FAS (ITT)
- Important deviations	FAS (ITT)
- Medical/Surgical history	FAS (ITT)
- Previous anti-cancer therapy	FAS (ITT)
- Concomitant medications/procedures	FAS (ITT)
- Subsequent anti-cancer therapy	FAS (ITT)
Safety Data	
- Exposure	Safety
- Adverse Events	Safety
- Laboratory measurements	Safety
- ECGs	Safety
- Vital Signs	Safety

2.2 Violations and Deviations

The important protocol deviations will be listed and summarised by randomised treatment group. None of the deviations will lead to any patients being excluded from any of the analysis sets described in Section 2.1. If the deviations are serious enough to have the potential to impact the primary analysis, sensitivity analyses may be performed.

The following general categories will be considered important deviations and be listed and discussed in the CSR as appropriate.

- Patients randomised but who did not receive olaparib/matching placebo.

- Patients who deviate from key entry criteria, which will be documented ahead of database lock.

- Baseline RECIST scan > 28 days before study treatment is started

- Baseline RECIST scan after randomised treatment is started.

- Patients who have a RECIST scan outside of a scheduled visit window on > 2 occasions.

- Disallowed concomitant medication use.

Misrandomisations in terms of errors in treatment dispensing patients receiving treatment other than that to which they were randomised as well as incorrect stratifications, will also be summarised and listed. A misrandomisation is when a patient is not randomised or treated according to the randomisation schedule. It is envisaged that there will be 2 sub categories of this:

(a)   Patients who receive no treatment whatsoever for a period of time due to errors in dispensing of medication. Note, this is not due to tolerability issues where patients may stop taking drug.

(b)   The patient receives a treatment pack with a different code to their randomisation code. However, the actual treatment may still match the randomised treatment. For example, a patient is given randomisation code 0001, which according to the randomisation schedule is olaparib. However, at the randomisation visit they are given treatment pack 0003, but this still contains olaparib.

Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1. During the study, decisions on how to handle misrandomisations will be made on an individual basis with written instruction from the study team leader and the study statistician.

In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification will be made prior to database lock and all decisions will be made whilst blinded to study treatment allocation.

3   PRIMARY AND SECONDARY VARIABLES

3.1   Derivation of RECIST Visit Responses

Patients with measurable or non measurable disease or no evidence of disease assessed at baseline by CT/MRI will be entered in this study. RECIST 1.1 has been modified to allow the assessment of progression due to new lesions in patients with no evidence of disease at baseline.

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

Baseline radiological tumour assessments are to be performed no more than 28 days before randomisation and ideally should be performed as close as possible to the start of study treatment. Tumour assessments are then performed every 12 weeks (±1 week) up to 156 weeks and then every 24 weeks (±1 week) following randomisation until disease progression.

If an unscheduled assessment was performed and the patient had 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.

At each visit, an overall visit response will be determined - using the information from target lesions (TL), non-target lesions (NTL) and new lesions. For the investigator assessments this will be done programmatically and the RECIST outcomes will be calculated using a computer program.

3.1.1   Target lesions (TLs)

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD) (except lymph nodes which must have short axis ≥ 15 mm) with CT or MRI and which is suitable for accurate repeated measurements.

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

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

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

For patients with no evidence of disease (NED) 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 NED.

Table 2   **TL Visit Responses**

Visit Responses	Description
Complete Response (CR)	Disappearance of all target lesions since baseline. Any pathological lymph nodes selected as target lesions must have a reduction in short axis to <10 mm
Partial Response (PR)	At least a 30% decrease in the sum of diameters of TL, taking as reference the baseline sum of diameters
Progressive Disease (PD)	At least a 20% increase in the sum of diameters of target lesions taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also indicate an absolute increase of at least 5 mm
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not Evaluable (NE)	Only relevant if any of the target lesions 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 target lesion response
Not Applicable (NA)	No target lesions are recorded at baseline

The percentage increase/decrease taking baseline as a reference for post-baseline visit will be calculated as:

[Post-baseline sum of diameters of TL – Baseline sum of diameters of TL] *100/ (Baseline sum of diameters of TL)

And similarly for the percentage change at a post-baseline visit using previous study minimum as a reference. If more than one baseline scan is recorded then measurements from the one that is closest to and prior to randomisation will be used to define the baseline sum of diameters of TLs.

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 target lesion measurements should be recorded. However, a visit response of PD should still be assigned if any of the following occurred:

- A new lesion is recorded.

- A NTL visit response of PD is recorded.

- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared. Note: the nadir can only be taken from assessments where all the TLs had a LD recorded.

Lymph nodes

For lymph nodes, if the size reduces to < 10 mm then these are considered non-pathological. However a size will still be given and this size should still be used to determine the TL visit response as normal. In the special case where all lymph nodes are < 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 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 TLs meet the CR criteria (i.e. 0 mm or < 10 mm for lymph nodes) then response will be set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node LD increases by 20% but remains < 10 mm.

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

- Step 3: If not all TLs meet the CR criteria and the sum of lesions meets the criteria for PD then response will be set to PD

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

TL too big to measure

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

25

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

TL too small to measure

If a TL becomes too small to measure a value of 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 blinded to treatment assignment.

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolisation), should be handled in the following way and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of 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 scale up as described below as long as there remain $\leq 1/3$ of the TLs with missing measurements. If the scaling results in a visit response of PD then the patient would be assigned a TL response of PD.

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

If $\leq 1/3$ of the TL measurements have interventions then the results will be scaled up based on the measurements at the nadir visit to give an estimated sum of diameters and this will be used in calculations; this is equivalent to comparing the visit sum of diameters of the non-intervention lesions to the nadir sum of diameters excluding the lesions with interventions.

Example of scaling

Lesion	Longest diameter at nadir visit	Longest diameter at follow-up visit
1	7.2	7.1
2	6.7	6.4
3	4.3	4.0
4	8.6	8.5
5	2.5	Intervention
Sum	**29.3**	**26**

Lesion 5 has had an intervention at the follow-up visit.

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

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

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

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 where appropriate (as per step 2 above).

Lesions that split in two

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

Lesions that merge

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

Change in method of assessment of TLs

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

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

3.1.2 Non-target lesions (NTLs) and new lesions.

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

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

Table 3   NTL Visit Responses

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

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

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

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

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

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

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

If the question 'Any new lesions since baseline' has not been answered with Yes or No and the new lesion details are blank this is not evidence that no new lesions are present and should be treated as NE in the derivation of overall visit response.

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

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

3.1.3 Overall visit response

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

Table 4   Overall Visit Responses

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

Table 4   **Overall Visit Responses**

Target Lesions	Non-target lesions	New Lesions	Overall Response
NE	CR, Non CR/Non PD or NE, NA	No (or NE)	NE
NA	NE	No (or NE)	NE
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD
NA	NA	No	NED

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, NE = not evaluable, NED = no evidence of disease, NA = not applicable (only relevant if there were no TL/NTL at baseline).

3.1.4   Independent review

The independent review charter contains the details of the independent central review conducted by the AstraZeneca-appointed central Core Imaging Laboratory and will be developed in advance of the start of the study. The independent data review will provide RECIST measurements for each visit for each patient at the time of primary data cut-off (DCO). After the primary PFS analysis, central review of scans will no longer be required.

For each patient, the independent reviewer will provide time point response data and the relevant scan dates for each time point (i.e. for visits where progression is/is not identified) with supporting measurements and assessments.

3.2 Outcome Variables

At each visit patients will be assigned a RECIST visit response of CR, PR, SD, PD, NED (applies only to those patients entering the study with no evidence of disease at baseline), NE depending on the status of their disease compared to baseline and previous assessments, using programmatically derived overall visit response from Investigator RECIST assessments. This will be repeated using the time point responses and relevant dates from the BICR.

Where applicable, outcome variables will be programmatically derived using data from Investigator RECIST assessments unless otherwise stated.

3.2.1   Progression free survival (PFS)

PFS is defined as the time from randomisation until the date of objective radiological disease progression according to RECIST or death (by any cause in the absence of progression) regardless of whether the patient discontinues randomised therapy or receives another anti-cancer therapy prior to progression (i.e. date of RECIST progression/death or censoring – date of randomisation + 1). Patients who have not progressed or died at the time of analysis, or who progress or die after two or more missed visits, are censored at the time of the latest evaluable visits or does not have a baseline assessment they will be censored at day 1 unless they die within two visits of baseline (25 weeks allowing for visit window).

Given the scheduled visit assessment scheme and the change in scanning frequency after 156 weeks then the following rules will be used to define two missed visits:

- If the latest evaluable assessment was on or prior to Week 133/Day 931 (Week 132 + one week) then two missed visits will equate to more than 26 weeks (12 x 2 + 2)

- If the latest evaluable assessment was post Week 133/ Day 931 and prior to Week 155/Day 1085 (Week 156 - one week) then two missed visits will equate to more than 38 weeks (12 + 24 + 2)

- If the latest evaluable assessment was on or post Week 155/Day 1085 (Week 156 - one week) then two missed visits will equate to more than 50 weeks (24 x 2 + 2)

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

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

(a)   For BICR assessment, date of progression will be determined based on the earliest of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or for either reviewer where both select PD as time point response and there is no adjudication.

(b)   For investigational site assessments, date of progression will be determined based on the earliest of the RECIST assessment/scan dates of the component that triggered the progression.

(c)   When censoring a patient for PFS the patient will be censored at the latest of the RECIST assessment/scan dates contributing to the last evaluable overall visit assessment.

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

Objective progression is defined as at least a 20% increase in the sum of the diameters of the target lesions (compared to previous minimum sum) and an absolute increase of > 5 mm, or an overall non-target lesion assessment of progression or a new lesion. For patients with no evidence of disease at baseline, following a complete radiological response to chemotherapy, progression is defined by the detection of new lesions on follow-up radiological assessments.

The primary analysis will be based on investigator-recorded assessment of the radiological scans.

A sensitivity analysis based on the BICR review of the radiological scans will be carried out.

The baseline RECIST assessment should be performed prior to randomisation but if an evaluable RECIST assessment occurs after randomisation but before treatment (and there is no available RECIST assessment before randomisation), then this assessment will be used as the baseline assessment. If a patient does not have a baseline RECIST scan performed prior to the date of first dose of study treatment (olaparib/placebo) then the patients will be censored at Day 1 in the analysis.

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

Time from randomisation to second progression or death is defined as the time from the date of randomisation to the earliest of the progression event subsequent to that used for the primary variable PFS or death. The date of second progression will be recorded by the Investigator and defined according to local standard clinical practice and may involve any of; objective radiological, CA-125 or symptomatic progression or death. Second progression status will be reviewed every 12 weeks following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression, i.e. censored at the latest of the PFS or PFS2 assessment date if the patient has not had a second progression or death. If, however the patient progresses for the second time or dies after two or more missed visits, the patient will be censored at the time of the latest evaluable investigator - recorded assessment (see Section 3.2.1 for details).

3.2.3   Overall survival (OS)

Overall survival (OS) is defined as the time from the date of randomisation until death due to any cause. Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made in the week following the date of DCO for the analysis, and if patients are confirmed to be alive or if the death date is post - DCO date these patients will be censored at the date of DCO. Death dates may be found by checking publicly available death registries.

3.2.4 Time to first subsequent therapy or death (TFST)

Time to start of first subsequent therapy or death (TFST) will be assessed. TFST is defined as the time from randomisation to the earlier of first subsequent therapy start date following study treatment discontinuation, or death. Subsequent therapies will be reviewed to assess which represent clinically important treatments intended to control ovarian cancer. Any patient not known to have died at the time of the analysis and not known to have had a further intervention of this type will be censored at the last known time to have not received subsequent therapy, i.e. the last follow-up visit where this was confirmed.

3.2.5 Time to second subsequent therapy or death (TSST)

Time to start of second subsequent therapy or death (TSST) will be assessed. TSST is defined as the time from randomisation to the earlier of the second subsequent therapy start date following study treatment discontinuation, or death. Any patient not known to have died at the time of the analysis and not known to have had a further intervention of this type 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.

3.2.6 Time to study treatment discontinuation or death (TDT)

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

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

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

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 Gynecologic Cancer Intergroup (GCIG) criteria (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, 2 times the nadir value in the 28 day period before day 1 on 2 occasions at least 1 week apart

(b) Where CA-125 does fall to within the normal range post-randomisation then there must be evidence of CA-125 greater than, or equal to, 2 times the ULN on 2 occasions at least 1 week apart

- Patients with CA-125 in the normal range on or before the date of randomisation and no results greater than ULN on or before the date of randomisation must show evidence of CA- 125 greater than, or equal to, 2 times the ULN on 2 occasions post-randomisation at least 1 week apart.

- CA-125 progression will be assigned the date of the first measurement that meets the above criteria.

Time to progression by RECIST or CA-125 progression or death is defined as the time from randomisation to the earlier date of RECIST 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 or their last available CA-125 measurement, whichever is the earliest at the time of the analysis. Since CA-125 is assessed more frequently than RECIST the two missed visit rule is based upon the RECIST schedule. Therefore if a patient dies, has RECIST progression or has CA-125 progression after two or more missed RECIST assessments, then the patient will be censored using at the last evaluable RECIST assessment where CA-125 was also collected. This will be defined as a RECIST assessment where the date of CA-125 sample is +/- 11 days (note the earliest date of the RECIST/CA-125 assessment will be used).

If only one assessment is missing during this period, no censoring is required. Patients that do not have any evaluable RECIST assessments or any CA-125 results post-randomisation will be censored at the date of randomisation.

Similarly to the primary analysis of PFS, RECIST progression will be determined from the investigator assessment of RECIST data.

3.2.8 Best overall RECIST response (BoR)

Best overall response (BoR) is calculated based on the overall 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 BoR 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 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 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 to the progression (PD) category. 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.

A patient will be classified as a responder if the RECIST criteria for a CR or PR are satisfied at any time prior to RECIST progression or the last evaluable assessment in the absence of RECIST progression, up to the earliest of the defined analysis cut-off point or the start of subsequent therapy. For each treatment group, the objective response rate (ORR) is the number of patients with a CR or PR post-baseline divided by the number of patients in the group in the FAS with measurable disease at baseline. Only patients with measurable disease at enrolment can achieve an objective response of CR or PR.

Duration of response will be defined as the time from the date of first documented response until date of documented progression or death in the absence of disease progression. The end of response should coincide with the date of progression or death from any cause used for the PFS endpoint. The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR. If a subject does not progress following a response, then their duration of response will use the PFS censoring time.

Time to onset of objective response is defined as the time from the date of randomisation until the date of first documented response. The date of first documented response should coincide with that used for the RECIST 1.1 BoR endpoint. Time to response will not be defined for those patients who do not have documented response.

The disease control rate (DCR) is defined as the percentage of patients who have at least one confirmed visit response of CR or PR or have demonstrated SD or NED for at least 23 weeks (i.e. 24 weeks ±1 week) prior to any evidence of progression. In the case of SD and NED, follow up assessments must have met the SD or NED criteria for a minimum interval of 23 weeks following randomisation.

3.3 Patient Reported Outcome (PRO) Variables

3.3.1 FACT-O

Patient reported health-related quality of life (HRQoL) will be assessed using the Functional Assessment of Cancer Therapy – Ovarian (FACT-O) questionnaire (Basen-Enquist K et al 2001). The FACT-O is composed of the following sub-scales: physical, social/family, emotional, and functional well-being as well as the additional concerns scales consisting of specific ovarian cancer symptoms.

The endpoint for HRQoL analysis will be the Trial Outcome Index (TOI), (Cella D et al 1993) an established single targeted index derived from the FACT-O questionnaire and it is considered to target the most relevant symptoms together with function and physical well-being and can be directly related to signs and symptoms and AEs. The TOI is composed of the following scales of the FACT-O: physical and functional well-being and additional concerns (ovarian cancer subscale).

Data relating to the FACT-O will be self-reported through patient questionnaires according to the study plan. Patients will be asked to report their HRQoL over the course of the previous 7 days. All patients will be asked to complete the FACT-O. The FACT-O questionnaire will be administered at baseline, at Day 29, then in line with the RECIST assessments every 12 weeks (+/- 7 days) for 156 weeks, then every 24 weeks (+/- 7 days) up to the data cut off for the primary analysis. In addition, HRQoL questionnaires will be collected at the discontinuation of study treatment visit and 30 days post last dose. Patients who had RECIST 1.1 disease progression will complete the questionnaires during the 12 weekly survival follow-up visits either in person or over the phone.

The Trial Outcome Index (TOI) score will be derived from the sum of the scores of three of the derived subscale scores; physical well-being (7 items), functional well-being (7 items), and ovarian cancer subscale (11 items) of the FACT-O questionnaire version 4. The total FACT-O score will also be calculated which is made up of the sum of the individual subscale scores: physical well being (PWB), social well being (SWB), emotional well being (EWB), functional well being (FWB) and ovarian cancer subscale (Additional Concerns) (11 items).

The scores will be derived in accordance with the FACT-O Scoring Manual. A number of items are negatively stated and need to be reversed by subtracting the response from "4". The scoring manual identifies that the following items need to be reversed prior to summarising: GP1-7, GE1, GE3-6, O1-3, C2, and B5. After reversing proper items, scores are summarised and multiplied by the number of items in the domain. For each subscale (domain), if less than 50% of the subscale items are missing, the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscale. If at least 50% of the items are missing, that subscale also will be treated as missing. The reason for any missing assessment will be identified. If data are missing at random, the above techniques will be used. If there is evidence that the missing data are systematic, missing values will be

handled to ensure that any possible bias is minimised. The TOI score ranges from 0-100 and the FACT-O from 0-152. For all Functional Assessment of Chronic Illness Therapy (FACIT) scales and symptom indices, a higher score indicates a higher HRQoL.

The actual change from baseline in TOI score will be derived for each visit where there is available data. For example; at visit X, the calculation will be (TOI score at visit X – baseline TOI score). Actual change from baseline for the individual FACT-O subscale scores will be calculated in a similar way. The baseline score is defined as last non-missing score prior to dosing with study treatment (olaparib or placebo). Any questionnaires completed on day 1 of dosing will be considered pre-dose.

A change of at least 10 points in TOI will be considered as a clinically relevant or a minimally important difference (Osobo et al 2005).

The threshold for a clinically important change is outlined below (Table 5):

Table 5   **Health-Related Quality of Life**

Score	Change from baseline	Visit response
TOI	$\geq +10$	Improved
	≤ -10 or "Subject too heavily affected by symptoms of disease under investigation" is answered as the reason for not completing HRQoL at visit.	Worsened
	Otherwise	No change

Best overall TOI improvement (in absence of starting any subsequent cancer therapy) will be defined as a change from baseline in the TOI of + 10 points or more (Osoba et al 2005) sustained for at least 28 days, the denominator consisting of a subset of the FAS population who have baseline TOI. It will be derived as the best symptom improvement response the patient achieved, based on evaluable HRQoL data collected from randomisation up to the earliest of starting any subsequent cancer therapy or death. Therefore, at the conclusion of the trial, the following criteria, will be used to assign a best overall score response based on the individual visit responses (Table 6).

Table 6   **Health-Related Quality of Life : Change rates - overall score**

Best overall TOI score response	Criteria

Table 6 **Health-Related Quality of Life : Change rates - overall score**

Best overall TOI score 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 subsequent response of "improved" or "no change" within 28 days.
Other	Does not qualify for one of the above.

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 evaluable patients with a TOI score at baseline.

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 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; or where REVPRDI form is ticked 'Subject too heavily affected by symptoms of disease under investigation'.

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 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 (per definition above), divided by the number of received forms.

3.3.2   EQ-5D-5L (exploratory analysis)

The EQ-5D-5L, developed by the EuroQol Group, is a generic questionnaire that provides a simple descriptive profile and a single index value for health status for economic appraisal. The questionnaire comprises six questions that cover five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and how the patient feels. For each dimension, patients select which statement best describes their health on that day from a possible five options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems and unable to/ extreme problems). A visual analogue scale (VAS) which ranges from 0 (worst imaginable health) to 100 (best imaginable health) is used to assess how the patient feels.

The EQ-5D-5L index comprises six questions that cover five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a VAS. A unique health state, the EQ-5D-5L profile, based on the five dimensions is reported as a five-digit code with a possible total of 3,125 health state. For example, state 11111 indicates no problems on any of the five dimensions. This EQ-5D-5L profile will be converted into a weighted health state utility value, termed the EQ-5D index, by applying a country-specific equation to the profile that represents the comparative value of health states. This equation is based on national valuation sets elicited from the general population and the base case will be the UK perspective. Where a valuation set has not been published, the EQ-5D-5L profile will be converted to the EQ-5D index using a crosswalk algorithm (van Hout et al 2012). The visual analogue scale, the EQ-VAS, is reported separately.

The evaluable population will comprise the FAS population.

to be 54 days between Day 29 and Day 85). If an odd number of days exist between two

3.4 Safety

Safety and tolerability will be assessed in terms of AEs (including SAEs), deaths, laboratory data, vital signs, ECG and exposure. These will be collected for all patients.

3.4.1 Adverse events

AEs and SAEswill be collected throughout the study, from informed consent until 30 days after the last dose of olaparib/placebo.

Other significant adverse events (OAE)

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

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

3.4.2   Treatment exposure

Exposure will be defined as follows:

Total (or intended) exposure of olaparib/placebo

- Total (or intended) exposure = last dose date – first dose date + 1

Actual exposure of olaparib/placebo

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

3.4.3   Dose intensity

Relative dose intensity (RDI) is the percentage of the actual dose intensity 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. Both will be derived using study treatment data up to three years or until the date of objective disease progression (whichever is earliest) as defined by RECIST using the investigator site assessments. If the investigator considered that it was in the patient's best interest to continue study treatment past this time, this was not included in the derivation of dose intensity.

Relative dose intensity (RDI) and percent intended dose (PID) will be defined as follows:

- RDI = 100% * 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 plus the protocol-defined post-dose rest period.

- PID = 100% * 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 3   **Example of Dose Intensity Calculations for Olaparib**
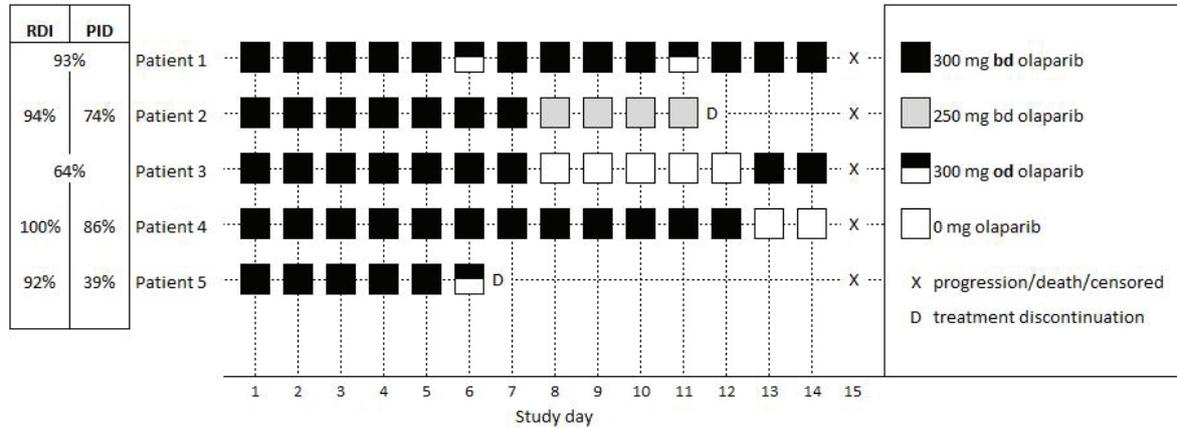
In this example, patients 1-4 progressed or were censored on Day 15. All four patients received less treatment than intended due to:

• 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)

Patient 1:   RDI = PID = ((12 * 300 mg * 2) + (2 * 300 mg)) / (14 * 300 mg * 2) = 93%

Patient 2:   RDI = ((7 * 300 mg * 2) + (4 * 250 mg * 2)) / (11 * 300 mg * 2) = 94%

PID = ((7 * 300 mg * 2) + (4 * 250 mg * 2)) / (14 * 300 mg * 2) = 74%

Patient 3:   RDI = PID = (9 * 300 mg * 2) / (14 * 300 mg * 2) = 64%

Patient 4:   RDI = (12 * 300 mg * 2) / (12 * 300 mg * 2) = 100%

PID = (12 * 300 mg * 2) / (14 * 300 mg * 2) = 86%

Patient 5:   RDI = ((5 * 300 mg * 2) + (1 * 300 mg)) / (6 * 300 mg * 2) = 92%

PID = ((5 * 300 mg * 2) + (1 * 300 mg)) / (14 * 300 mg * 2) = 39%

3.4.4　Laboratory data

Laboratory data will be collected throughout the study, from screening to follow-up visit as described in Table 1, 2 and 3 of the CSP. Blood and urine samples for determination of haematology, clinical chemistry, and urinalysis will be collected as described in Section 3.1.9 of the CSP. For derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 3.4.7 below will be used.

3.4.5　ECGs

ECG data obtained up until the 30 days from date of last dose of olaparib/placebo treatment will be used for reporting. At screening, overall evaluation of ECG, QTcF (QT interval corrected for heart rate using Fridericia's correction) and QTcB (QT interval corrected for heart rate using Bazett's correction) will be collected. For all post-baseline visits overall evaluation of ECG will be collected.

3.4.6　Vital signs

Vital signs data obtained up until the 30 days from date of last dose of olaparib/placebo treatment will be used for reporting. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 3.4.7 below will be used.

3.4.7　General consideration for safety assessments

Time windows will need defining for any presentations that summarise values of laboratory and vital signs data 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 up to and including 30 days after the last dose of study drug 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. The halfway point is assumed to be the midpoint of the number of days between the visits, excluding both visit days (for example there are assumed to be 27 days between day 29 and day 57). 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 days.

  For example, the visit windows for vital signs data are:

- Day 29, visit window 2 – 43

- Day 57, visit window 44 – 71

- Day 85, visit window 72 – 99

- Day 113, visit window 100 – 127

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 summaries to be presented for these visits separately as well as the inclusion of this data in mapped on-treatment visit summaries.

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

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

- For visit based summaries:

  - If there is more than one value per patient within a time window then the closest value should be summarised, or the earlier in the event the values are equidistant from the nominal visit date. The listings should highlight the value for that patient that went into the summary table, wherever feasible.

  - To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group visit data should only be summarised if the number of observations is greater than the minimum of 20 or > 1/3 of patients dosed.

- 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 (olaparib or placebo). For laboratory data and vital signs data, any assessments made on day 1 will be considered pre-dose. Where safety data are summarised over time, study day will be calculated in relation to date of first treatment (olaparib or placebo)

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.

3.5 Health Care Resource Use

The assessments to be carried out at each visit are detailed in the study schedule and include the HOSPAD module.

The assessment of health care resource use will increase the understanding regarding the relationship between treatment and tumour related cancer symptoms on resource use, such as the need for palliative procedures to address obstruction and bleeding. This will be captured and analysed to inform submissions to payers.

Calculation or derivation of health care resource use

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

- Planned and unplanned hospital attendances beyond trial protocol mandated visits (including physician visits, emergency room visits, day cases and admissions)

- Primary sign or symptom the patient presents with

- Length of hospital stay

- Length of any time spent in an intensive care unit (ICU)

Where admitted overnight, the length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation (length of hospital stay = end date of hospitalisation – start date of hospitalisation + 1). Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation. The length of ICU stay will be calculated using the same method.

Approximately 53 patients from sites in China will be recruited and randomised in a 2:1 ratio to receive olaparib or placebo and will follow the same study plan and procedures as patients recruited to the global study. The safety and efficacy data collected will be summarised and analysed separately to the global study safety and ITT analysis sets (as defined in section 2.1). A standalone SAP will be written for the China cohort.

Analyses of the China cohort will be performed after a minimum of 29 PFS events have been observed. The primary statistical analysis of the efficacy of olaparib for China cohort ITT patients will be an assessment of progression free survival based on investigator assessment. A KM plot of PFS will be presented by treatment group. Summaries of the number and percentage of subjects experiencing a PFS event, and the type of event (RECIST or death) will be provided along with median PFS for each treatment. In addition, an analysis of PFS will be performed. The HR (olaparib:placebo) and associated CI will be calculated from a Cox proportional hazards model (ties = Efron) that contains the treatment term and a term for response to first line platinum chemotherapy. The HRs and 95% CIs will be presented on a forest plot including the HR and 95% CI from the overall ITT population (using the primary analysis). Exploratory p-values from a log rank test stratified by response to first line platinum chemotherapy will also be presented. If there are issues with this approach a Cox proportional hazards model with a treatment term only and an unadjusted log rank test will be implemented.

Summaries and analysis of secondary supportive efficacy endpoints (including OS, PFS2, ORR, time to first subsequent therapy (TFST), time to second subsequent therapy (TSST) time to earliest progression by RECIST or CA-125 or death and TDT) will be performed for the China cohort as described for the ITT analysis set.

In addition, an exploratory analysis may be performed using a weighted average or "borrowing evidence" approach (Huang 2012) to combine data from non-China cohort and China cohort patients to estimate the PFS treatment effect if required to support a future China HA request. This will be done outside of the China CSR

When assessing safety and tolerability, summaries will be produced separately for the China cohort based on the China Safety Analysis Set as defined in section 2.1. The China safety data will be summarised descriptively and will not be formally analysed.

Updated safety summaries may also be produced which include safety data from all subjects who received at least one dose of randomised treatment (olaparib or placebo) if requested by health authorities.

4   ANALYSIS METHODS

PFS will be analysed when approximately 196 events have occurred or after the last patient randomised has had the opportunity to have been on the study for at least 36 months,

whichever comes first. No further analyses of PFS are planned beyond this point unless requested by Health Authorities.

An initial analysis of OS, PFS2, TFST and TSST will be performed at the same time as the primary analysis of PFS and will use the same methodology and model.

Supportive analyses of time to earliest progression by RECIST or CA-125 or death and TDT will be provided, using the same methodology as specified for the primary analyses of PFS; however no multiple adjustment will be applied as these are viewed as supportive endpoints.

4.1 General Principles

The primary statistical analysis of the efficacy of olaparib will include all randomised patients and will compare the treatment groups 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 Full Analysis Set (FAS). Therefore, all efficacy and HRQoL data will be summarised and analysed using the FAS on an ITT basis.

When assessing safety and tolerability, summaries will be produced based on the Safety Analysis Set. This will include all patients who receive at least one dose of randomised treatment (olaparib or placebo). Patients in the safety analysis set will be analysed by their initial treatment received. The safety data will be summarised descriptively and will not be formally analysed.

Relevant parametric analyses of time to event endpoints required for payer submissions will be outlined in a separate payer analysis plan.

4.2 Analysis Methods

The treatment comparison is olaparib 300 mg bd versus placebo.

All efficacy analyses will be performed on the ITT population. In addition, as a sensitivity to the main analyses of PFS, PFS2, OS, TDT, TFST and TSST, analyses of these endpoints will be performed in those patients whose gBRCAm status is confirmed by Myriad centrally. Each randomised patient will have one of either the Myriad BRACAnalysis CDx performed under Quality Systems Regulations (QSR) or the Myriad Integrated BRACAnalysis assay that is performed under Clinical Laboratory Improvement Amendments (CLIA). Patients randomised in the China Cohort had gBRCAm testing performed by BGI.

Key efficacy (PFS, PFS2, OS, TDT, TFST and TSST) and safety estimates associated with patients whose tBRCAm status is confirmed by Foundation Medicine will also be reported as a separate subset.

Results of all statistical analysis will be presented using 95% confidence intervals and two-sided p-values.

The following table details which endpoints are to be subject to formal statistical analysis, together with pre-planned sensitivity analyses making clear which analysis is regarded as primary for that endpoint.

Table 7   Formal Statistical Analyses to be Conducted and Pre-Planned Sensitivity Analyses

Endpoints Analysed	Notes
PFS (Time from randomisation to first progression or death)	Primary analysis: stratified log-rank test using investigator data
	Key sensitivity analyses: • stratified log rank test using investigator data in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test using investigator data in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
	Additional sensitivity analyses:
	1) Evaluation time bias analysis; stratified log-rank test using investigator data
	2) Attrition bias analysis (using alternative censoring rules); stratified log-rank test using investigator data
	3) Ascertainment bias analysis; stratified log-rank test using BICR data
	4) Deviation bias analysis (if meaningful to do); stratified log-rank test using investigator data
	5) Stratified log rank test using U and V statistics to calculate HR and CI based on investigator data
OS (Time from randomisation to death due to any	Stratified log-rank test

Table 7   **Formal Statistical Analyses to be Conducted and Pre-Planned Sensitivity Analyses**

Endpoints Analysed	Notes
cause)	Key sensitivity analyses: • stratified log rank test in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
PFS2 (Time from randomisation to second progression or death)	Stratified log rank test based on investigator assessment of second progression
	Key sensitivity analyses: • stratified log rank test using investigator data in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
	Additional sensitivity analysis: marginal model approach of Wei et al 1989
TFST (Time to first subsequent therapy or death)	Stratified log rank test
	Key sensitivity analyses: • stratified log rank test in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
TSST (Time to second subsequent therapy or death)	Stratified log rank test
	Key sensitivity analyses: • stratified log rank test in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
TDT (Time to study treatment discontinuation or	Stratified log rank test

Table 7   Formal Statistical Analyses to be Conducted and Pre-Planned Sensitivity Analyses

Endpoints Analysed	Notes
death)	Key sensitivity analyses: • stratified log rank test in randomised patients confirmed as gBRCA mutation positive by Myriad centrally • stratified log rank test in randomised patients confirmed as tBRCA mutation positive by Foundation Medicine
Change from baseline in TOI score	MMRM analysis of the change from baseline in TOI score
Time to earliest progression by RECIST 1.1, CA-125 or death	Stratified log rank test using investigator data

4.2.1   Multiplicity

In order to describe the nature of the benefits of olaparib maintenance treatment, PFS, PFS2, TFST, TSST, TDT change from baseline in TOI score and OS will be tested at a 2-sided significance level of 5%.

In addition to these planned analyses, which will be performed and reported in the CSR, in order to strongly control the type I error at 2.5% one-sided for key label claims, a multiple testing procedure (MTP) will also be employed across the primary endpoint (PFS) and key secondary endpoints (PFS2 and OS). There is no requirement to adjust for multiplicity due to PFS interim analyses, since there are no planned interim PFS analyses with the opportunity to make an early claim of efficacy.

A hierarchical testing strategy will be employed where PFS is tested first using the full test mass (full test mass = alpha) and key secondary endpoints of PFS2 and OS will then be tested using a MTP with a recycling strategy (i.e., the MTP will recycle the test mass to the endpoint not yet rejected in the hierarchy outlined in Figure). The MTP is detailed below.

Figure 4   **Multiple Testing Procedure**

PFS2 will only be tested (with the test mass split between interim and final PFS2 analyses) if statistical significance is shown for PFS. OS will only be tested if statistical significance is shown for PFS2.

Both PFS2 and OS will be tested at the time of the primary analyses of PFS and again when there are approximately 60% deaths. A proportion of alpha will be spent at this first analysis time point for both endpoints to control for multiple testing however different spending functions will be applied for each.

An interim analysis for PFS2 will be performed at the time of the PFS analysis   Statistical significance will be declared at the interim analysis for PFS2 if the 1-sided p<0.0125.

If the null hypothesis for PFS2 is not rejected at this first analysis time point then PFS2 will be tested again when the final analysis of OS occurs. The type I error will be controlled at 2.5% 1-sided by assigning approximately 1.8% significance level (1-sided) to the final analysis of PFS2 (final significance level to be determined accounting for correlation between the interim and final PFS2 analyses) (Stone 2010). If PFS2 is significant at either the interim or final analyses, the full test mass (alpha) will be carried forward to OS.

An interim analysis for OS will be performed at the time of the PFS analysis. Statistical significance will be declared at the interim analysis for OS if the null hypothesis for PFS2 is rejected at the PFS analysis and the observed p-value for OS is p<0.0001. This allows the significance level at the final analysis for OS to be controlled at the 2.5% level (1-sided) (Haybittle J L 1971).

4.2.2 Primary variable - progression free survival (PFS)

PFS will be analysed using a log-rank test stratified by response to first line platinum chemotherapy (in the opinion of the investigator, clinical complete response (CR) or partial response (PR)) for generation of the p-value and using the Breslow approach for handling ties. The hazard ratio (HR) and confidence interval (CI) will be estimated from a Cox Proportional

Hazards model (with ties = Efron and the stratification variable as a covariate) and the CI will be calculated using a profile likelihood approach.

Stratification variables will be defined according to data from the randomisation. If there are any patients who were mis-stratified, a sensitivity analysis will be carried out using the (correct) baseline data collected in the eCRF. Although not anticipated, if patients are randomised in error when they have not previously had a response to first line platinum chemotherapy, they will be categorised in the "PR" category for the sensitivity analysis using eCRF stratification data.

The HR (olaparib versus placebo) together with its corresponding 95% CI and p-value will be presented (a HR less than 1 represents the reduction in risk for those patients allocated olaparib).

A KM plot, with tick marks to identify censored observations, of PFS will be presented by treatment group. Summaries of the number and percentage of subjects experiencing a PFS event, and the type of event (RECIST or death) will be provided along with median PFS and corresponding 95% CI for each treatment.

The assumption of proportionality will be assessed. 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. The primary analysis will be based on the PFS based on investigator assessments, and using all scans regardless of whether they were scheduled or not.

The proportion and corresponding 95% CI of patients alive and progression free at 6 monthly intervals will be summarised (using the KM plot) and presented by treatment group.

The number of patients prematurely censored will be summarised by treatment arm together with baseline prognostic factors of the prematurely censored patients. A patient is defined as prematurely censored if they did not progress and the latest scan prior to DCO was more than one scheduled tumour assessment interval (+ 2 weeks) prior to the DCO date.

In addition, duration of follow-up will be summarised using median time from randomisation to date of censoring (date last known to be non-progressor) in censored (not progressed) patients only, presented by treatment group. The interquartile range will also be presented by treatment group.

As patients will be randomised, imbalances in demographic factors between the treatment groups are not anticipated. However, if any imbalances should occur, the HR and associated CI calculated from a Cox Proportional Hazards model containing treatment, the stratification variable and these additional demographic variables, may be reported.

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

Subgroup analyses will be conducted comparing PFS between treatments. Median PFS will be reported. The purpose of the subgroup analyses is to assess the consistency of treatment effect across potential or expected prognostic factors. The results observed in Phase II (D0810C00019) do not suggest that these factors will be predictive factors for a qualitatively different treatment effect. If there are too few events available for a meaningful analysis of a particular subgroup (where there are less than 20 events per subgroup level no formal statistical tests will be performed), the relationship between that subgroup level and PFS will not be formally analysed. In this case, only descriptive summaries will be provided. When there are less than 20 events per a subgroup, consideration will be given to combing relevant subgroups if appropriate to do so.

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

- Response to previous platinum chemotherapy (obtained from the randomisation schedule)

- gBRCAm status-confirmed by Myriad test or gBRCAwildtype (wt) or gBRCA variant of uncertain significance (VUS) or missing by Myriad test]*.This will be determined from the Myriad central laboratory tests.

- ECOG performance status at baseline (normal activity [PSTAT=0] or restricted activity [PSTAT=1]).

    This will be determined from the response to "Performance status" (PSTAT module) on the eCRF at screening.

- Baseline CA-125 value (≤ ULN or > ULN).

    The baseline CA-125 value will be defined as the measurement nearest to but prior to date of randomisation.

- BRCA mutation type (BRCA1 or BRCA2 or BRCA1/2 (both)).

This will be determined from the Myriad central laboratory tests. If there are less than 20 events in the "BRCA1/2 both" category, these patients will be excluded from this analysis.

- Age (<65 or $\geq$ 65).

  This will be determined from the date of birth (BIRTHDAT in the DEM module) and date of randomization. See section 4.2.13 for rules on deriving age for partial dates of birth.

- Stage of disease at initial diagnosis (III [FIGO_STG=30 or 31 or 32 or 33] or IV [FIGO_STG=40]).

  This will be determined from the response to "FIGO stage" (PATHGEN module) on the eCRF at screening.

- Residual macroscopic disease following debulking surgery prior to entry into the study [HISPOUT=1] or no residual macroscopic disease [HISPOUT=2].

  This will be determined from the response to the "Last debulking surgery outcome" on the history of debulking surgery page (HISHC module) of the eCRF at screening.

- Region 1 (North America or Rest of World).

  This will be determined from the centre number (CENTRE). A 'North America' patient is defined as any patient randomised at a site in USA or Canada. Otherwise a patient is defined as 'Rest of World'. This selection is based on potential differences in clinical practice between the regions.

- Region 2 (Brazil, Poland, Russia, Japan, Korea or Rest of World)

  This will be determined from the centre number (CENTRE). A 'Brazil, Poland, Russia, Japan, Korea' patient is defined as any patient randomised at a site in either Brazil, Poland, Russia, Japan or Korea. Otherwise a patient is defined as 'Rest of World'. This selection is based on a comparison of Asia/non Western Europe versus Rest of World.

- Race (White or Black/African-American or Asian or Native Hawaiian/Pacific Islander or American Indian/Alaska Native or Others).

  This will be determined from the response to "Race" (DEM module) on the eCRF at screening.

* Some patients may only have a tumour mutation or are from China where no sample will be supplied for Myriad testing.

Other baseline variables may also be assessed if there is clinical justification, for example, patients who received intravenous (IV) chemotherapy versus patients who received intraperitoneal (IP) chemotherapy in their first line regimen.

A minimal number of patients that are tBRCAm (by local testing) and gBRCA wt (by Myriad testing) are expected to be randomised into this study. Assuming the number of progression events in this population is less than 20, the number and proportion of events will be summarised by treatment arm. If the number of events is ≥20, this factor will be added to the forest plot (i.e. gBRCA by Myriad testing and tBRCA mutated versus tBRCA mutated only).

For each subgroup, the HRs (olaparib: placebo) and associated CIs will be calculated from a Cox Proportional Hazards model (ties = Efron) that contains the treatment term, factor (subgroup) and treatment-by-factor interaction term. The treatment effect HRs for each treatment comparison along with their CIs will be obtained for each level of the subgroup from this single model. The HRs and 95% CIs will be presented on a forest plot including the HR and 95% CI from the overall population (using the primary analysis). In addition, summaries of the number and percentage of patients experiencing a PFS event for each subgroup will be provided along with the median PFS for each treatment.

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

The presence of quantitative interactions will be assessed by means of an overall global interaction test. This will be performed in the overall population by comparing the fit of a Cox Proportional Hazards model including treatment, all covariates, and all covariate-by-treatment interaction terms, with one that excludes the interaction terms and will be assessed at the 2-sided 10% significance level. If a covariate does not have ≥20 events per level (of the covariate) it will be included as a covariate in the model but the covariate-by-treatment interaction term will be omitted. If the fit of the model is not significantly improved then it will be concluded that overall the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. Throughout this process all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

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

A further analysis of PFS (using investigator assessed RECIST) may be performed at the time of the OS analyses, if requested by health authorities.

PFS sensitivity analyses

As a key sensitivity analysis to the primary endpoint of PFS, the primary analysis will be repeated excluding any patients who did not have a gBRCA mutation status confirmed by Myriad centrally. The same methodology and model will be used and the HR and associated 95% CI from a Cox Proportional Hazards model will be reported. A KM plot of PFS in this subset of patients will be presented by treatment group. This sensitivity analysis will also be repeated excluding any patients who did not have a tBRCA mutation status confirmed by Foundation Medicine.

Sensitivity analyses will be performed to assess the possible presence of time-assessment bias (i.e. differential assessment times between treatment groups)

Summary statistics for the number of weeks between the time of progression and the last evaluable RECIST assessment prior to progression will be presented for each treatment group.

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

(a)   Evaluation-time bias

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

(b)   Attrition bias

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

was originally a censored event in the primary PFS analysis becomes an actual event and what originally was a PFS event becomes a censored event) will be presented.

(c)   Ascertainment bias

A stratified log-rank test will be repeated using BICR RECIST data to programmatically derive PFS. The HR and 95% CI will be presented.

The type of event (RECIST or death) will also be provided.

If there is an important discrepancy between the primary analysis using investigator assessments and this sensitivity analysis using BICR assessments, then the proportion of subjects with site but no central confirmation of progression will be summarised. This scenario is possible as at the time of the primary analysis data cut off all scans will be sent for central review assessment regardless of whether or not a patient has been assessed as having progressed by the site. The approach of imputing an event at the next visit in the BICR assessed analysis may help inform the most likely HR value, but only if an important discrepancy exists.

Disagreements between investigator and BICR assessment of RECIST progression will be presented for each treatment group. The summary will include the early discrepancy rate which is the frequency of BICR progressions declared before the investigator review progressions (≥2 weeks earlier and including progressions declared by BICR but not investigator) as a proportion of all BICR progressions, and the late discrepancy rate which is the frequency of BICR progressions declared after the investigator review (≥2 weeks later and including progressions declared by investigator but not BICR as a proportion of all discrepancies (including early and late discrepancies) (Amit et al 2011).

(d)   Deviation bias (if meaningful to do)

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

- did not have the intended disease or indication or

- did not receive any randomised therapy

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

An additional sensitivity analysis of PFS will be performed based on a log-rank test stratified by the stratification variable, i.e. response to first line platinum chemotherapy, and using the Breslow approach for handling ties. The HR and CI will be estimated from the U and V

statistics obtained directly from the LIFETEST model with inclusion of STRATA terms for the stratification variable.

The HR and its CI will be estimated from the log-rank as follows (Berry et al 1999 and Sellke and Siegmund 1983):

HR = exp(U/V)

95% CI for HR = (exp{U/V – 1.96/ $\sqrt{V}$}, exp{U/V + 1.96/ $\sqrt{V}$})

Where $U = \sum_k U_k = \sum_k \sum_i (d_{1ki}, -e_{1ki})$ is the stratified log-rank test statistic obtained from the SAS LIFETEST procedure, $\sqrt{V} = \sqrt{\sum_k V_k}$, is its standard deviation, k denotes the stratum and d_{1ki} and e_{1ki} are the observed and expected events in Group 1, stratum k, ith event time.

(e)   Event rate at 24 months

The proportion of patients progression free at 24 months (PFS24) will be defined as the Kaplan-Meier estimate of PFS at 24 months.

The stratified KM estimates (95% CI) of PFS24 stratified by response to first line platinum chemotherapy will be estimated by treatment group. The difference (95% CI) in the stratified estimates of PFS24 between treatments will also be calculated. For estimating the confidence intervals, Greenwood's estimate of the variance of the KM proportion will be used (Collett 2003). The hypothesis of no difference in PFS24 between treatment groups across the levels of response to first line platinum chemotherapy will also be tested.

4.2.3   Time from randomisation to second progression (PFS2)

PFS2 analysis will be performed using the same methodology and model as PFS. If PFS2 is not statistically significant at the primary PFS analysis, then a further analysis of PFS2 will be performed when the OS data are approximately 60% mature. If PFS2 is statistically significant at the primary PFS analysis, then no further analysis of PFS2 will be performed unless requested by a regulatory authority. The type of progression (objective progression by RECIST, progression by CA-125, symptomatic progression or other) will also be summarised by treatment arm.

As a key sensitivity, the analysis of PFS2 will be repeated in those patients whose gBRCAm status is confirmed by Myriad centrally. A KM plot of PFS2 in this subset of patients will be presented by treatment group. This sensitivity analysis will also be repeated excluding any patients who did not have a tBRCA mutation status confirmed by Foundation Medicine.The sensitivity analysis outlined for PFS in Section 4.2.2 will not be repeated for PFS2 with the exception of a KM plot of the time to censoring where the censoring indicator of the primary

PFS2 is reversed (what was originally a censored event in the primary PFS2 analysis becomes an actual event and what originally was a PFS2 event becomes a censored event).

In addition, a sensitivity analysis of PFS2 using the marginal model approach of Wei, Lin and Weissfeld (Wei LJ et al 1989) will be undertaken where patient risks will be partitioned from randomisation to PFS and from PFS to PFS2.

Time from second progression to previous assessment will be summarised by treatment arm.

4.2.4   Overall survival (OS)

OS data will be analysed at the time of the primary analysis of PFS and will use the same methodology and model (provided there are sufficient events available for a meaningful analysis [$\geq$20 deaths], if not descriptive summaries will be provided). A further analysis of OS will be performed when the OS data are approximately 60% mature.

As a key sensitivity, the analysis of OS will be repeated in those patients whose gBRCAm status is confirmed by Myriad centrally. A KM plot of OS in this subset of patients will be presented by treatment group. This sensitivity analysis will also be repeated excluding any patients who did not have a tBRCA mutation status confirmed by Foundation Medicine.

The remaining sensitivity analyses outlined for PFS in Section 4.2.2 will not be repeated for OS with the exception of a KM plot of the time to censoring where the censoring indicator of the primary OS is reversed (what was originally a censored event in the primary PFS2 analysis becomes an actual event and what originally was a PFS2 event becomes a censored event).

A summary of survival status at the time of analysis will be produced. This will summarise the number of patients who have died, who are still in survival follow-up, who are lost to follow-up or who have withdrawn consent.

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

- In censored (not died) patients only: Time from randomisation to date of censoring (date last known to be alive)

- In all patients: Time from randomisation to the date of death or to the date of censoring for censored patients.

At the time of the further analysis of OS, the subgroup analyses and global interaction test detailed for PFS in Section 4.2.2 will be repeated for OS.

Exploratory analyses of OS

Exploratory analyses of OS adjusting for impact of subsequent PARP inhibitor trial or treatment may be performed if a sufficient proportion of patients switch. The decision to adjust and final choice of methods will be based on a blinded review of the data and the plausibility of the underlying assumptions. Baseline and time-dependent characteristics will be explored, and summaries of baseline characteristics will be summarised for placebo patients, splitting between those that have and haven't switched at the time of the analyses. This will be performed outside of the CSR.

4.2.5   Time to first subsequent therapy or death (TFST) and time to second subsequent therapy or death (TSST)

Time to first subsequent therapy or death (TFST) and time to second subsequent therapy or death (TSST) will be analysed at the same time as the primary analysis of PFS and using the same methodology and model. The HRs for the treatment effect together with 95% CIs will be presented. KM plots will be presented by treatment arm. In addition, the number of patients who received further therapy relative to progression (before, after, no progression) will also be presented by treatment arm.

Summary tables of first and second subsequent therapies by treatment arm will be provided, as well as response to first and second subsequent therapy by treatment arm.

Further analyses of these endpoints will be performed when the OS data are approximately 60% mature.

As a key sensitivity, the analyses of TFST and TSST will be repeated in those patients whose gBRCAm status is confirmed by Myriad centrally. KM plots of TFST and TSST in this subset of patients will be presented by treatment group. These sensitivity analyses will also be repeated excluding any patients who did not have a tBRCA mutation status confirmed by Foundation Medicine.

4.2.6   Time to study treatment discontinuation or death (TDT)

Time to study treatment discontinuation or death (TDT) will be analysed at the same time as the primary analysis of PFS and using the same methodology and model. The HR for the treatment effect together with 95% CIs will be presented. A KM plot will be presented by treatment arm. No multiplicity adjustment will be applied as this is viewed as a supportive endpoint.

Further analysis of this endpoint will be performed when the OS data are approximately 60% mature.

As a key sensitivity, the analyses of TDT will be repeated in those patients whose gBRCAm status is confirmed by Myriad centrally. A KM plot of TDT in this subset of patients will be

presented by treatment group. This sensitivity analysis will also be repeated excluding any patients who did not have a tBRCA mutation status confirmed by Foundation Medicine.

4.2.7   Time to earliest progression by RECIST 1.1, 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 (%) of patients reporting a CA-125 progression, and a combined objective progression and/or CA- 125 progression will be tabulated.

No multiplicity adjustment will be applied as this is viewed as a supportive endpoint (to PFS).

4.2.8   Best overall RECIST response (BoR)

For each treatment arm, Best Overall 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.

The ORR and DCR will be summarised (i.e., number of patients (%)) by treatment group, in patients in the FAS (ITT population) with measurable disease at baseline. In addition, for patients who have an objective response, the duration and onset of the response will be summarised.

4.2.9   Patient reported outcomes (PROs)

The analysis population for HRQoL data will be the subset of the FAS (ITT set).

It should be noted that some centres erroneously collected patient data using the FACT-O at non-protocol visits. These data will not be used in any analyses or summaries, but will be listed in the appendices and flagged accordingly.

The HRQoL benefit of the long PFS delay with olaparib is expected to be observed when placebo patients require chemotherapy earlier as this is when HRQoL may be expected to deteriorate. It is expected that the stability in "HRQoL" over the 24 months following start of randomised treatment will be longer for patients randomised to olaparib than placebo. This will be investigated using the TOI derived from the FACT-O questionnaire.

Change from baseline TOI scores

Change from baseline in TOI score will be regarded as the primary analysis of the FACT-O questionnaire and will be analysed using a mixed model for repeated measures (MMRM) analysis of the change from baseline (defined as prior to first dose) in TOI scores for each visit.

The primary analysis will be to compare the average treatment effect from the point of randomisation for the first 24 months (which will include analysis visits obtained within the

first 24 months, i.e. baseline, day 29 (week 4), weeks 12, 24, 36, 48, 60, 72, 84, 96, see Section 3.3.2) unless there is excessive missing data (defined as >75% missing data). If the time to first subsequent chemotherapy when approximately 50% of placebo patients receive chemotherapy does not occur by 24 months post-randomisation then additional time periods will be analysed and will be included on supportive summaries and graphical displays as appropriate.

The MMRM model will include patient, treatment, visit (analysis) and treatment by visit interaction as explanatory variables, the baseline TOI score as a covariate along with the baseline TOI score by visit interaction. Treatment, visit and treatment by visit interaction will be fixed effects in the model; patient will be included as a random effect. Restricted maximum likelihood (REML) estimation will be used. An overall adjusted mean estimate will be derived that will estimate the average treatment effect over visits giving each visit equal weight.   For this overall treatment comparison, adjusted mean estimates per treatment group and corresponding 95% CIs will be presented along with an estimate of the treatment difference, 95% CI and p-value. The treatment by visit interaction will remain in the model regardless of significance.

An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. The following provides sample code for implementing the MMRM analysis:

proc mixed data=TOI method = reml;

```
  class TRT VISIT SUBJECT;

  model TOISC = TRT VISIT TRT*VISIT TOIBL TOIBL*VISIT  / s ddfm=kr;

  repeated VISIT / type=UN subject=SUBJECT;

  random intercept / subject= SUBJECT;

 lsmeans TRT / at means pdiff diff alpha=0.05 cl;
```

where TRT is the randomised treatment, VISIT is the visit, TOISC is the change from baseline in the TOI score, and TOIBL is the baseline TOI score.

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

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

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

fit of the model or estimation of the treatment effects, SUBJECT will be treated as a fixed effect.

For each treatment and visit, the adjusted (least squares) mean estimates, corresponding 95% CIs, estimates of the treatment difference, corresponding 95% CIs and p-values will be presented.

During the randomised treatment phase, it is expected that olaparib will not cause harm to HRQoL, this will be demonstrated by comparing the Trial Outcome Index (TOI) best response rates (improved/no change/worsened) whilst on randomised treatment. A contingency table summarising the TOI best change rates will be provided and a CMH analysis including the randomisation stratification factor will be performed.

An AUC (Area Under the Curve) analysis will also be performed for TOI. For each patient, the AUC will be derived using the trapezoidal rule and including all available scheduled visits. Time-adjusted AUC will be calculated as AUC divided by time from baseline to last HRQoL visit where the patient provides data. The mean time-adjusted AUC will be analysed using analysis of variance (ANOVA) including using the same stratification factor as described for the primary analysis of PFS. Time-adjusted AUC summary statistics (mean, median, minimum, maximum, StD) per treatment arm will be presented. In addition, the analysis table will provide the estimated mean difference between treatment groups, 95% CI and p-value. This analysis will done using the first 24 months of data provided by a patient. For this analysis, for any patient who dies during the first 24 months, data for each question in the FACT-O from the time of death until 24 months will be set to 0. In addition, an analysis using all available data for each patient will be performed.

If deemed appropriate alternative methods may be used (e.g. piecewise linear modelling).

Descriptive statistics and graphs will be reported for the TOI by visits as well as change in these scores from baseline.

Total FACT-O and subscales scores

Descriptive statistics, by visit and change from baseline, will be reported for physical well-being (PWB), functional well-being (FWB) and the ovarian cancer subscale (Additional Concerns) domains. Descriptive statistics will also be provided for the individual questions that make up the additional concerns subscale.

Descriptive statistics, by visit and change from baseline, will also be reported for the emotional well being (EWB) and social well being (SWB) subscales, these two subscales do not form part of the TOI itself.

FACT-O questionnaire compliance (overall compliance and by visit compliance) will be summarised for each 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 3.3.3 for description of visit windowing). These visits will not be presented in corresponding figures.

In order to assess the amount of missing data and the reasons for missing data, plots including the number of questionnaires completed/partially completed/missing at each visit will be produced.

4.2.10 Exploratory analyses

EQ-5D-5L

The evaluable population will comprise all patients in the FAS (ITT population).

Descriptive statistics will be calculated for each scheduled visit/time point in the study, for each trial arm and as a total. These will report the number of patients, the number of EQ-5D questionnaires completed at each visit, the number and proportion responding to each dimension of the EQ-5D-5L. Additionally summary statistics (e.g. n, mean, median, StD, min, max) will be reported for the EQ-5D index score and the EQ-VAS score, and the change from baseline for the EQ-5D index score and the EQ-VAS score.

Graphical plots of the mean EQ-5D index score and EQ-VAS score, including change from baseline, and associated 95% CI by scheduled visits/time points in the study will be produced. To support submissions to payers, additional analyses may be undertaken and these will be outlined in a separate Payer Analysis Plan.

Health care resource use

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

Tumour BRCA

A summary table will be produced comparing BRCA mutation/s in tumour to germline BRCA mutation status.

Subsequent therapy

Subsequent therapies received after discontinuation of olaparib will be summarised and listed by treatment group, together with number of regimens received. Patients who subsequently received a PARP inhibitor or entered a PARP inhibitor trial will be summarised and listed by treatment arm according to line of subsequent therapy.

4.2.11 Safety

Safety data will be summarised and listed only. No formal statistical analyses will be performed on the safety data. All safety data will be summarised by actual treatment group (olaparib or placebo) including patients who have dose reduction for blinded period of study. Any patient who received an initial dose of olaparib will be included in the olaparib group, even if the patient was planned to receive placebo. Similarly, any patient who received an initial dose of placebo will be included in the placebo group, even if the patient was planned to receive olaparib. However, some listings such as AEs listings will display the actual dose the patient received at onset of an AE.

Adverse events

All AEs, both in terms of Medical Dictionary for Regulatory Activities (MedDRA) preferred term and Common Toxicity Criteria for Adverse Events (CTCAE) grade, will be listed and summarised descriptively by count (n) and percentage (%) for each treatment arm. MedDRA dictionary will be used for coding. Any AE occurring before olaparib/placebo treatment (i.e. before Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as 'pre-treatment'.

The summary tables will include all AEs that occurred after the start of treatment up until the end of the 30 day follow-up period. The 30 day follow-up period will be defined as 30 days following discontinuation of olaparib/placebo treatment.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved), investigator's assessment of severity and relationship to study drug. Frequencies and percentages of patients reporting each preferred term will be presented (i.e. multiple events per patient will not be accounted for apart from on the episode level summaries).

 Partially missing dates will be imputed as per the phUSE guidelines (http://www.phusewiki.org/wiki/index.php?title=Imputing_Partial_Dates)

For date of onset:

- Missing day - Impute the 1st of the month unless month is same as month of first dose of study drug then impute first dose date

- Missing day and month – impute 1st January unless year is the same as first dose date then impute first dose date

- Completely missing – impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date.

When imputing a start date ensures that the new imputed date is sensible i.e. is prior to the end date of the AE.

For date of resolution:

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

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

- Completely Missing – need to look at whether the AE is still ongoing before imputing a date and also when it started in relation to study drug. If the ongoing flag is missing then assume that AE is still present (i.e. do not impute a date). If the AE has stopped and start date is prior to first dose date then impute the 1st dose date, if it started on or after first dose date then impute a date that is after the last dose date.

Summary information (the number and percent of patients by treatment) will be tabulated for:

- All AEs

- All AEs causally related to study medication

- AEs with CTCAE grade 3 or higher

- AEs with CTCAE grade 3 or higher, causally related to study medication

- AEs with outcome of death

- AEs with outcome of death causally related to study medication

- All serious adverse events (SAEs)

- All SAEs causally related to study medication

- AEs leading to discontinuation of olaparib/placebo

- AEs leading to discontinuation of olaparib/placebo, causally related to olaparib/placebo

- Other significant AEs

- Other significant AEs causally related to olaparib/placebo

An overall summary of the number and percentage of patients in each category will be presented, as will an overall summary of the number of episodes in each category. In addition, a truncated AE table of most common AEs, showing all events that occur in at least

5% of patients overall will be summarised by preferred term, by decreasing frequency. This cut-off may be modified after review of the data.

Each AE event rate (per 1000 patient years) will also be summarised by preferred term within each system organ class. For each preferred term, the event rate will be presented and will be defined as the number of patients with that AE divided by the sum of the duration from the start of treatment to 30 days after the last treatment dose (for patients without the event) and the time to the AE (for patients with the event) in each group multiplied by 1000.

AEs will be assigned CTCAE grades (National Cancer Institute (NCI) CTCAE version 4.0) and summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, system organ class, preferred term and actual treatment group. Fluctuations observed in CTCAE grades during study will be listed.

Summaries of the number and percentage of patients with AEs leading to dose modification of olaparib/placebo by preferred term and treatment group will be presented for the following:

- AEs leading to a dose reduction of olaparib/placebo

- AEs leading to a dose interruption of olaparib/placebo

- AEs leading to a dose modification, defined as a dose interruption and/or dose reduction of olaparib/placebo.

In addition, AEs with outcome of death, SAEs, AEs leading to discontinuation of treatment, AEs causally related to olaparib/placebo and OAEs will be listed.

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

- Death related to disease under investigation only

- Death related to disease under investigation only (death > 30 days after last treatment dose)

- AE with outcome of death only

- AE related to disease under investigation and with AE outcome of death

- AE with outcome of death only (AE start date falling >30 days after last treatment dose)

- Deaths > 30 days after last treatment dose, unrelated to AE or disease under investigation

- Deaths >30 days after last treatment dose, AE related to disease under investigation and with AE outcome of death,

- Patients with unknown reason for death.

A corresponding listing will also be produced.

A separate summary will be produced that presents any events that occur prior to dosing or starting more than 30 days after discontinuing therapy.

Summary tables will also be produced for the following common adverse events, based on grouped preferred terms:

- Anaemia

- Neutropenia

- Thrombocytopenia

- Nausea

- Vomiting

- Fatigue/Asthenia

The specific preferred terms to be included in the summaries will be provided prior to database lock.

Summary of long term tolerability

For each AE, median time to first onset of the AE will be presented in patients in the safety analysis set by actual treatment group. Patients who did not experience the AE will not be included in the summaries. Summary tables of time to first onset for each AE will also be produced (e.g. 1-28 days, 29-56 days, 57-84 days, 85-112 days, >112 days). Median duration of the AE will be presented in patients who experienced each AE.

Laboratory assessments

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

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

Shift tables for laboratory values by worst common toxicity criteria (CTC) grade will be produced, and for specific parameters separate shift tables indicating hyper- and hypo-directionality of change will be produced. For parameters with no CTCAE grading, shift tables from baseline to worst value on-treatment will be provided (i.e. on-treatment is defined as data collected up until 30 days after the last dose of olaparib/placebo).

A scatter plot of ALT versus total bilirubin, both expressed as multiples of ULN, will be produced. The scatter plot will be repeated for AST versus total bilirubin.

Liver biochemistry test results over time for patients with elevated ALT or AST, and elevated total bilirubin (at any time) will be plotted. Individual patient data where ALT or AST plus total bilirubin are elevated at any time will be listed also.

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

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

ECGs

ECG data will be listed by actual treatment group.

Vital signs

Vital signs (SBP, DBP, pulse rate, body temperature and weight) will be summarised over time in terms of absolute values and changes from baseline at each scheduled measurement by actual treatment group. Vital signs data will be listed by actual treatment group.

Selected safety data will also be summarised and listed at the time of updated OS analysis (approximately 60% maturity for OS).

4.2.12 Demographic and baseline characteristics data

The following will be listed and summarised by randomised treatment group:

- Patient disposition (including screening failures and reason for screening failure) This will also include the number (%) of patients who at two years had no evidence of disease and so discontinued study treatment

- Important deviations

- Inclusion in analysis populations

- Demographics (age, age group, sex, race and ethnicity)

- Patient characteristics at baseline (height, weight, weight group (<65, 65-90, > 90), body mass index (BMI) and body mass index group ('Normal (<25)', 'Overweight (25-30)' and 'Obesity (>30)')

- Stratification factors recorded on the eCRF

- Stratification factors according to the randomisation

- Patient recruitment by country and centre

- Previous ovarian cancer therapy

- Previous therapy for other cancer

- Disease characteristics at baseline (ECOG performance status, *BRCA* gene name at screening [*BRCA1, BRCA2* or Both or Missing] confirmed locally and by Myriad, *gBRCA* status confirmed locally and by Myriad testing [*gBRCAm, gBRCA wt, gBRCA VUS*, Missing], *tBRCA* status confirmed by Foundation Medicine testing [*BRCAm, BRCA wt, BRCA VUS*, Missing], primary tumour location, histology type, tumour grade, International Federation of Gynecology and Obstetrics (FIGO) stage, time from previous platinum chemotherapy to reandomisation, baseline CA-125 value, On-study baseline tumour biopsy and overall disease classification)

- Extent of disease

- Disease related medical history (including number of regimens of previous platinum chemotherapy)

- Relevant surgical history

- History of debulking surgery

- Number of blood transfusions

- Physical examination at baseline

- Time from completion of first line platinum chemotherapy to randomisation

- Disallowed concomitant medications

- Allowed concomitant medications

- Post-discontinuation cancer therapy

- Radiotherapy (previous, on randomised treatment and post-treatment discontinuation)

- BRCA mutation status (by Local germline and Myriad germline, by Myriad germline and tumour BRCA and by Local germline and tumour BRCA)

- Deleterious and suspected deleterious mutation types in germline and tumour BRCA patients

AZ drug dictionary (AZDD) will be used for concomitant medication coding.

Age will be derived as age at last birthday in whole years using the date of randomisation and date of birth. Where a partial date of birth has been collected, the following imputation rules will be applied in order to calculate the patient's age for use in listings and summaries tables presenting age and/or age group and subgroup analyses based on age:

- If only the month and year of birth has been collected, the day of birth will be imputed as 15th

- If only the year of birth has been collected the day and month of birth will be imputed as 1st July

- If the date of birth is completely missing, the derived age of the patient will also be missing.

Date of birth will be listed as it has been collected on the eCRF.

Overall disease classification will be categorised as follows:

- If a patient has any site of metastic disease, they are classified as 'Metastatic',

- If a patient has no sites of metastatic disease and does have sites of local disease they are classified as 'local'

- Else if they have no sites of disease identified and no target lesions present at baseline and no non-target lesions present at baseline (as assessed by the investigator) and they do have a valid investigator assessed baseline scan, they are classified as 'NED',

- Else classified as 'unknown'

Partial dates for start and end of concomitant medications will be derived as per date of onset and date of resolution for adverse events, respectively.

Patients who were unblinded (a) prior to disease progression and (b) prior to or on the day of treatment discontinuation will be listed.

A listing containing both the eCRF BRCA results and the Myriad gBRCA and Foundation Medicine tBRCA results will be produced.

Patients disposition data will also be summarised and listed at the time of updated OS analysis (approximately 60% maturity for OS).

4.2.13 Treatment exposure

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

- Total exposure of olaparib/placebo.

- Actual exposure of olaparib/placebo.

- Reasons for dose reductions, dose interruptions, and dose modifications of olaparib/placebo. Dose reductions and dose interruptions will be based on investigator initiated dosing decisions. Dose interruptions due to "Subject Forgot to Take Dose" will be omitted from these summaries.

- Number of dose reductions, dose interruptions, and dose modifications of olaparib/placebo that last for a period of three days or more.

- PID and RDI of olaparib/placebo (entire intended treatment period).

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

All treatment information data will be listed for the safety analysis set by actual treatment group.

Selected exposure data will also be summarised and listed at the time of updated OS analysis (approximately 60% maturity for OS).

4.2.14 Data cut-offs

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

5   INTERIM ANALYSES

No interim analyses of PFS prior to the primary analysis will be performed; however additional analyses of PFS and/or OS may be performed to meet Regulatory Agency requests.

This study will use an external independent data monitoring committee (IDMC) to perform interim reviews of accumulating study safety data. This committee will be composed of therapeutic area experts and statisticians, who are not employed by AZ, and do not have any major conflict of interest. Following the review the IDMC will recommend whether the study should continue unchanged, be terminated, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to AstraZeneca. The report will only include the recommendation and any potential protocol amendments and will not contain any unblinding information.

A separate IDMC charter will be developed which will contain details of the IDMC members and clearly define the responsibilities of the IDMC.

A set of outputs using the data from subjects recruited in Japan will be produced at the time of the primary analysis for Japanese regulatory purposes. The outputs will be repeats of a selection of the main set of outputs and the results will not be described in the clinical study report, but rather will be described in a separate standalone report. In addition some repeats of output will be produced for subjects recruited in Japan at OS timepoint.

6   CHANGES OF ANALYSIS FROM PROTOCOL

All efficacy and HRQoL will be analysed using FAS.

Section 3.2.1 Progression free survival
Added clarification that according to the RECIST 1.1 guidelines an overall visit response of PD also requires an absolute increase of > 5 mm in the sum of the diameters of the target lesions.

Section 3.2.7 Time to earliest progression by RECIST or CA-125 or death
Derivations will be relative to randomization (CSP states start of study treatment). Also, censoring will use the earliest of the last evaluable RECIST assessment or the last available CA-125 measurement (CSP states most recent).

7   REFERENCES

Alsop et al 2012
Alsop K, Fereday S, Meldrum C, DeFazio A, Emmanuel C, George J, et al. BRCA Mutation Frequency and Patterns of Treatment Response in BRCA Mutation–Positive Women With

Ovarian Cancer: A Report From the Australian Ovarian Cancer Study Group. J Clin Oncol 2012;30:2654-63.

Amit et al 2011
Amit O, Mannino F, Stone Am et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis and recommendations from a PhRMA working group. European J Cancer 2011; 47:1772-1778.

Basen-Enquist K et al 2001
Basen-Engquist K, Bodurka-Bevers D, Fitzgerald MA, Webster K, Cella D, Hu S, et al. Reliability and validity of the functional assessment of cancer therapy-ovarian. J Clin Oncol. 2001 Mar 15;19(6):1809-17.

Berry et al 1999
Berry G, Kitchin RM and Mock PA. A comparison of two simple hazard ratio estimators based on the logrank test. Statistics in Medicine 1991; 10(5):749-755

Cella D et al 1993
Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993 Mar;11(3):570-9.

Collett 2003
Modelling Survival Data in Medical Research, Second Edition. CRC Press, 2003

FACT-O Scoring Manual
FACT-O Scoring Manual. See URL: www.facit.org

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

Haybittle J L 1971
Haybittle, J L 1971. Repeat assessment of results in clinical trials of cancer treatment. British Journal of Radiology. 1971 Oct; 44:793-7

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

Huang 2012
Qin Huang, Gang Chen, Zhilong Yuan, K K Gordan Lan: Design and sample size considerations for simultaneous global drug development program. J Biopharm Stat 2012 Sep;22(5):1060-73

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

Osoba et al 2005
Osoba D, Bezjak A, Brundage M, et al: Analysis and interpretation of health-related quality-of-life data from clinical trials: Basic approach of The National Cancer Institute of Canada Clinical Trials Group. Eur J Cancer 41:280-287, 2005

Sellke and Siegmund 1983
Sellke, T. and Siegmund, D. Sequential analysis of the proportional hazards model. Biometrika 1983 70: 315-326

Stone 2010
Stone, A. The application of bespoke spending functions in group sequential designs and the effect of delayed treatment crossover in survival trials. Pharm. Statistics 5 (2010) 151-161

Sun and Chen 2010
Sun X and Chen C. Comparison of Finkelstein's Method With the Conventional Approach for Interval-Censored Data Analysis. Stat Biopharm Research 2010;2:1:97-108.

van Hout et al 2012
Ben van Hou, M.F. Janssen, You-Shan Feng, Thomas Kohlmann,, Jan Busschbach et al. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. Value in Health 2012; 15: 708-715.

Wei LJ et al 1989
Wei LJ, Lin DY, Weissfeld L. Regression-analysis of multivariate incomplete failure time data by modelling marginal distributions. Journal of the American Statistical Association 1989; 84:1065–73.