AstraZeneca

Statistical Analysis Plan

Study Code	D081DC00008
Edition Number	Final v2.0
Date	23OCT2017

A Randomised, Double-Blind, Placebo-Controlled, Multicentre Phase II Study to Compare the Efficacy, Safety and Tolerability of Olaparib Versus Placebo When Given in Addition to Abiraterone Treatment in Patients With Metastatic Castrate-Resistant Prostate Cancer Who Have Received Prior Chemotherapy Containing Docetaxel

A Randomised, Double-Blind, Placebo-Controlled, Multicentre Phase II Study to Compare the Efficacy, Safety and Tolerability of Olaparib Versus Placebo When Given in Addition to Abiraterone Treatment in Patients With Metastatic Castrate-Resistant Prostate Cancer Who Have Received Prior Chemotherapy Containing Docetaxel

QuintilesIMS Study Statistician PPD PPD 23 OCT 17

Date

A Randomised, Double-Blind, Placebo-Controlled, Multicentre Phase II Study to Compare the Efficacy, Safety and Tolerability of Olaparib Versus Placebo When Given in Addition to Abiraterone Treatment in Patients With Metastatic Castrate-Resistant Prostate Cancer Who Have Received Prior Chemotherapy Containing Docetaxel

AZ Study Statistician

PPD

Robert Kozarski

23OCT2017

Date

A Randomised, Double-Blind, Placebo-Controlled, Multicentre Phase II Study to Compare the Efficacy, Safety and Tolerability of Olaparib Versus Placebo When Given in Addition to Abiraterone Treatment in Patients With Metastatic Castrate-Resistant Prostate Cancer Who Have Received Prior Chemotherapy Containing Docetaxel

Global Product Statistician PPD

Nigel Baker

23 OCT 2017

Date

TABLE OF CONTENTS PAGE

TITLE PAGE 1
SIGNATURE OF STUDY STATISTICIAN 2
SIGNATURE OF GLOBAL PRODUCT STATISTICIAN 3
SIGNATURE OF GLOBAL PRODUCT STATISTICIAN 4
TABLE OF CONTENTS 5
LIST OF ABBREVIATIONS 9
AMENDMENT HISTORY 13
1. STUDY DETAILS 14
1.1 Study objectives 14
1.1.1 Primary Objective 14
1.1.2 Secondary Objectives 14
1.1.3 Exploratory Objectives 16
1.2 Study design 17
1.2.1 Part A: Safety run-in/dose escalation 18
1.2.2 Part B: Randomised part 19
1.3 Number of subjects 19
1.3.1 Part A: Safety run-in/dose escalation 19
1.3.2 Part B: Randomised part 20
2. ANALYSIS SETS 21
2.1 Definition of analysis sets 21
2.1.1 Full analysis set 21
2.1.2 PK analysis set 21
2.1.3 Safety analysis sets 21
2.2 Violations and deviations 21
2.2.1 Part A 21
2.2.2 Part B 22
3. PRIMARY AND SECONDARY VARIABLES 23
3.1 Derivation of RECIST visit responses – malignant soft tissue 23
3.1.1 RECIST Target lesions (TLs) 23
3.1.2 RECIST Non-Target Lesions (NTLs) and new lesions. 27
3.1.3 RECIST overall visit response 28
3.1.4 RECIST Best objective response 29
3.2 Evaluation of PCWG2 progression status – bone lesions 30
3.3 Primary variable – radiologic progression free survival (rPFS) 31

3.4 Secondary outcome variables 33
3.4.1 Time to second progression (PFS2) 33
3.4.2 Overall survival (OS) 33
3.4.3 Objective response rate (ORR) and associated variables 34
3.4.4 Time to subsequent therapies 35

3.5 Safety 35
3.5.1 Dose limiting toxicities (DLTs) 35
3.5.2 Adverse Events 36
3.5.3 Laboratory data 36
3.5.4 Vital signs 37
3.5.5 Electrocardiograms (ECG) 37
3.5.6 Physical examination 38
3.5.7 Duration of exposure 38

3.6 Pharmacokinetic variables and PK analysis methods 39

3.7 Calculation or derivation of biomarker variable(s) 40
3.7.1 Prostate specific antigen 40
3.7.1.1 PSA response 40
3.7.1.2 PSA changes on continuous scale 40
3.7.2 Circulating tumour cell counts (CTCs) 40
3.7.3 ATM mutation 41
3.7.4 BRCA mutation 41
3.7.5 Composite BRCA/ATM mutation 42
3.7.6 Composite HRR mutation (excluding BRCA1, BRCA2 and ATM) 43

3.8 Exploratory variables 43
3.8.1 Patient reported outcomes (PRO) and HRQL 43
3.8.1.1 Functional Assessment of Cancer therapy – Prostate Cancer (FACT-P) 43
3.8.1.2 Brief Pain Inventory – short form (BPI-SF) 49
3.8.1.3 Bone pain item 50
3.8.2 Health economics 51
3.8.2.1 Utilities: EQ-5D-5L 51
3.8.2.2 Resource use 52

4. ANALYSIS METHODS 52

4.1 General principles 52

4.2 Analysis methods 53
4.2.1 Multiplicity 53
4.2.2 Analysis of the primary efficacy variable (rPFS) 54
4.2.2.1 Sensitivity analyses 56
4.2.3 Analysis of the secondary outcome variables 57
4.2.3.1 Time to second progression (PFS2) 57
4.2.3.2 Overall survival (OS) 57
4.2.3.3 Objective response rate (ORR) and associated variables 57

4.2.3.4 Time to first subsequent therapy or death (TFST) and time to second subsequent therapy or death (TSST) 58
4.2.4 Tolerability and safety 58
4.2.4.1 Adverse events (AE) 58
4.2.4.2 Dose-limiting toxicities 62
4.2.4.3 Exposure and compliance 62
4.2.4.4 Laboratory data 62
4.2.4.5 Concomitant medications 63
4.2.4.6 Vital signs 63
4.2.5 Biomarkers 63
4.2.6 Patient reported outcomes (PRO) and HRQL 64
4.2.6.1 FACT-P 64
4.2.6.2 BPI-SF 65
4.2.6.3 Bone Pain 66
4.2.7 Health economics data 66
4.2.7.1 Utilities: EQ-5D-5L 66
4.2.7.2 Resource use 66
4.2.8 Demographic data and baseline data 66
4.2.9 Statistical analysis of PK data 67
5. INTERIM ANALYSES (NOT APPLICABLE) 71
6. CHANGES OF ANALYSIS FROM PROTOCOL 71
7. REFERENCES 71
8. APPENDIX (NOT APPLICABLE) 73

LIST OF TABLES

Table 1 Primary objectives 14
Table 2 Secondary objectives 14
Table 3 Exploratory objectives 16
Table 4 TL visit responses 24
Table 5 Example of scaling-up 27
Table 6 Overall visit response for soft tissue lesions 29
Table 7 Overall PCWG-2 progression status (bone) 31
Table 8 Overall radiological visit response 32
Table 9 Description of ATM mutation categories 41
Table 10 Description of BRCA mutation categories 41
Table 11 Description of composite BRCA/ATM mutation categories 42
Table 12 Description of composite HRR categories 43

Table 13 Definition of visit response for FACT-P, FAPSI-8, FAPSI-6, TOI, PCS, FWB, PWB and SWB 45
Table 14 Overall score response for FACT-P, FAPSI-8, FAPSI-6, TOI, PCS, FWB, PWB, SWB 47

LIST OF FIGURES

Figure 1 Multiple Testing Procedure 54

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
ALT	Alanine aminotransferase
Amean	Arithmetic mean
ATM	Ataxia telangiectasia mutated
AST	Aspartate aminotransferase
AUC_{ss}	Area under the plasma concentration-time curve across the dosing interval at steady state
AUC_{0-t}	Area under the plasma concentration-time curve from time zero to last quantifiable concentration
Bid	Twice daily (Latin: *bis die*)
BOR	Best objective response
BP	Blood pressure
BPI-SF	Brief Pain Inventory – Short Form
BRCA	Breast cancer gene, i.e., BRCA1 and BRCA2
CI	Confidence interval
CL_{ss}/F	Apparent clearance of drug from the body as steady state
$C_{ss,max}$	Maximum plasma concentration at steady state
$C_{ss,min}$	Minimum plasma concentration at steady state
CR	Complete response
CRPC	Castrate-resistant prostate cancer
CSR	Clinical Study Report
CT	Computed tomography
CTC	Circulating tumour cell
CTCAE	Common Terminology Criteria for Adverse Event
%CV	% Coefficient of variation
DAE	Discontinuation of Investigational Product due to Adverse Event
DCO	Data cut-off
DDI	Drug-drug interaction
DLT	Dose-limiting toxicity
DNA	Deoxyribonucleic acid

Abbreviation or special term	Explanation
DOR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EQ-5D-5L	EuroQuol-5 Dimensions, five level
EWB	Emotional well-being
FACIT	Functional assessment of chronic illness
FACT-P	Functional Assessment of Cancer Therapy – Prostate Cancer
FAPSI-6	Functional Assessment of Prostate Cancer Symptoms Index 6
FAPSI-8	Functional Assessment of Prostate Cancer Symptoms Index 8
FPI	First patient in
FWB	Functional well-being
Gmean	Geometric mean
GSD	Geometric standard deviation factor
%GCV	% Geometric coefficient of variation
HR	Hazard ratio
HRR	Homologous recombination repair
HRQL	Health-Related Quality of Life
ICU	Intensive care unit
IP	Investigational Product
ITT	Intention to treat
LD	Longest diameter
LOQ	Limit of quantification
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed model for repeated measures
MRI	Magnetic resonance imaging
MTP	Multiple testing procedure
NA	Not applicable
NC	Not calculable
NCA	Non-compartmental analysis
NE	Not-evaluable
NED	No evidence of disease

Abbreviation or special term	Explanation
NGS	Next-generation sequencing
NQ	Non-quantifiable
NRS	Numerical rating scale
NTL	Non-target lesion
OAE	Other Significant Adverse Event
ORR	Objective response rate
OS	Overall survival
PARP	Polyadenosine 5'-diphosphoribose polymerase
PCS	Prostate cancer symptoms
PCWG-2	Prostate Cancer Working Group 2
PD	Progressive disease
PFS2	Second progression
PID	Percent intended dose
PK	Pharmacokinetics
PR	Partial response
PRO	Patient Reported Outcome
PSA	Prostate specific antigen
PWB	Physical well-being
QTc	Corrected QT interval
RDI	Relative dose intensity
RECIST 1.1	Response Evaluation Criteria in Solid Tumours version 1.1
REML	Restricted maximum likelihood
rPFS	Radiologic progression-free survival
RPFST	Rank preserving structural failure time
RR	Respiratory rate
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Stable disease
SD	Standard deviation based on non log-transformed data
SI	International system of units (*systeme international*)
SOC	System organ class
SRC	Safety Review Committee

Abbreviation or special term	Explanation
SWB	Social well-being
TFST	Time to first subsequent therapy
TL	Target lesion
$t_{ss,max}$	Time to reach maximum plasma concentration at steady state
$t_{ss,last}$	Time to last quantifiable plasma concentration at steady state
TOI	Trial outcome index
TSST	Time to second subsequent therapy

AMENDMENT HISTORY

Date	Brief description of change
10OCT2017	SAP v2.0 vs SAP v1.0
	• Replaced efficacy analysis use of Cox's proportional hazards modelling with the log rank test. Added 95% CI for HR to be reported. • Two-sided 5% significance level for the endpoint testing replaced with 1-sided 2.5%. • No interim analysis for PFS2 and OS is planned. Both endpoints will be tested at the time of rPFS analysis. • Updated primary efficacy subgroup analysis plans by inclusion of the other (beside ATM, BRCA1/2) genes mutation statuses from the HRR group, and exclusion of ERG expression/fusion status. The minimum 20 endpoint events per subgroup/arm condition replaced with minimum 5 events. • The PTEN and AR marker factors were excluded from the current exploratory analysis. Any futures exploratory analysis and outcome variables were left yet to be defined. No exploratory analysis is planned to assess an impact from possible treatment switching on the OS adjustment. • In the PK analysis, the DDI ratio was changed from mono:combination to combination:mono for olaparib and abiraterone. • Updated modelling approach in PRO supportive analysis • List of important deviation categories for Part B was extended. Table 7 (Overall PCWG-2 response), Table 8 (Overall radiological visit response classification), and Table 14 (Overall score response) content updated. The FACT-P missing data subscale scoring details updated, and HRQL visit response extended with PWB and SWB scores. Weight (part of vital signs) levels were recorded at the baseline only. In the summary of deaths the number of categories increased. • Completed list of the SAP v2.0 changes from the CSP • Global product statistician and study statistician changed • Added 80% CI to the sensitivity analysis.

1. STUDY DETAILS

1.1 Study objectives

1.1.1 Primary Objective

Table 1 Primary objectives

Primary objective	Primary outcome variable
Part A, Safety run-in:	
To assess the safety and tolerability of olaparib when given in addition to abiraterone and to recommend, by assessment of dose-limiting toxicities (DLTs) and other safety and tolerability data, a dose of olaparib for further study when given in addition to abiraterone.	Assessment of adverse events (AEs) graded by common terminology criteria for adverse events (CTCAE v4.0), vital signs (including blood pressure (BP), pulse), and evaluation of laboratory parameters (clinical chemistry and haematology). Incidence of DLTs during the initial evaluation period.
Part B:	
To compare the efficacy of olaparib when given in addition to abiraterone, with placebo given in addition to abiraterone, by assessment of radiologic progression-free survival (rPFS) using Response Evaluation Criteria in Solid Tumours version 1.1 (RECIST 1.1) and Prostate Cancer Working Group 2 (PCWG-2) criteria.	rPFS, defined as the time from randomisation to disease progression according to RECIST 1.1 (for soft tissue disease) and/or PCWG-2 criteria (for bone disease), or death.

1.1.2 Secondary Objectives

Table 2 Secondary objectives

Secondary objective	Secondary outcome variables
Part A, Safety run-in:	
To evaluate the presence of any drug interaction between olaparib and abiraterone by determination of steady state exposure to olaparib in the presence and absence of abiraterone, and determination of steady state exposure to abiraterone in the presence and absence of olaparib.	Olaparib and abiraterone pharmacokinetic steady state (PK) parameters (where the data allow): maximum plasma concentration at steady state ($C_{ss,max}$), time to reach maximum plasma concentration at steady state ($t_{ss,max}$), minimum plasma concentration at steady state ($C_{ss,min}$), area under the plasma concentration-time curve at steady state (AUC_{ss}), area under the plasma concentration-time curve from time zero to time of last quantifiable concentration (AUC_{0-t}), time to last quantifiable plasma concentration at steady state ($t_{ss,last}$), along with ratios for $C_{ss,max,}$ $C_{ss,min}$ and AUC_{ss} of drug in combination versus monotherapy.

Table 2 Secondary objectives

Secondary objective	Secondary outcome variables
Part B:	
To compare the safety and tolerability of olaparib when given in addition to abiraterone, with placebo given in addition to abiraterone.	Assessment of AEs graded by CTCAE v4.0, vital signs (including BP, pulse), and evaluation of laboratory parameters (clinical chemistry and haematology).
To assess the anti-tumour activity of olaparib when given in addition to abiraterone, compared with placebo given in addition to abiraterone, by measurement of changes in circulating prostate specific antigen (PSA) and circulating tumour cells (CTCs), calculation of overall radiological objective response rate (ORR) (by RECIST 1.1 and PCWG-2 bone scan criteria) and malignant soft tissue ORR (by RECIST 1.1), duration of response (DOR), time to first subsequent therapy (TFST) for prostate cancer, and time to second subsequent therapy (TSST) for prostate cancer.	Absolute and percentage change from baseline in PSA levels and PSA response. Change and best change from baseline in CTC numbers, and CTC conversion rates. Tumour response in terms ORR (malignant soft tissue response and overall radiological response [malignant soft tissue response by RECIST 1.1, overall radiological response by RECIST 1.1 and PCWG-2]) and DOR. TFST, defined as the time from randomisation until the first subsequent therapy for prostate cancer, and TSST, defined as the time from randomisation until the second subsequent therapy for prostate cancer.
To assess the efficacy of olaparib when given in addition to abiraterone, compared with placebo given in addition to abiraterone, by assessment of overall survival (OS).	OS, defined as time from randomisation to the date of death from any cause.
To assess the efficacy of olaparib when given in addition to abiraterone, compared with placebo given in addition to abiraterone, by assessment of time from randomisation to second progression (PFS2).	PFS2. Progression defined by local standard clinical practice. May involve any of: objective radiological progression, symptomatic progression, rises in PSA or death.
To investigate BRCA and ATM mutation as candidate predictors of response to olaparib. In addition to BRCA and ATM mutations, mutations in 12 other homologous recombination repair (HRR) genes will be explored. Note: This objective is dependent upon the number of evaluable samples obtained from the study.	BRCA, ATM, and HRR mutation status; if the number of events in these groups is sufficient then the primary analysis of rPFS will be repeated in all BRCA and/or ATM, and HRR mutation patients.

1.1.3 Exploratory Objectives

Table 3 Exploratory objectives

Exploratory objectives	Exploratory outcome variables
To explore the effects of olaparib on pain and other prostate cancer-related symptoms compared to placebo.	Change from baseline in worst pain, general pain and pain interference in daily activities scales of the Brief Pain Inventory – Short Form (BPI-SF) and the worst bone pain item. Change from baseline in the Functional Assessment of Prostate Cancer Symptoms Index 8 (FAPSI-8), as derived from 8 items within the Functional Assessment of Cancer Therapy – Prostate Cancer (FACT-P), Functional Assessment of Prostate Cancer Symptoms Index 6 (FAPSI-6) and the Prostate Cancer Symptoms (PCS), as derived from the 12 items in the prostate specific module of the FACT-P.
To explore the effects of olaparib on health-related quality of life (HRQL) compared to placebo.	Change from baseline, as measured by the FACT-P scales: functional well-being (FWB), physical well-being (PWB), emotional well-being (EWB), social well-being (SWB), the FACT-P total score, and the trial outcome index (TOI) score (the sum of the PWB, FWB and PCS scores).
To assess the time to deterioration in pain.	Time to deterioration in the BPI-SF worst pain item and time to deterioration in the worst bone pain item.
To assess time to deterioration in HRQL.	Time to deterioration in HRQL, as measured by FACT-P TOI score.
To explore the impact of treatment and disease state on health state utility.	EuroQuol-5 Dimensions, five level (EQ-5D-5L) health state utility index.
To investigate the impact of treatment and disease progression on metastatic castrate-resistant prostate cancer (CRPC) management resource use.	Resource use will be captured, focussing on in-patient and ICU admissions, length of stay, palliative interventions and reason for admission into hospital and interventions.
Future exploratory research into factors that may influence development of cancer and/or response to treatment (where response is defined broadly to include efficacy, tolerability or safety) may be performed on the collected and stored archival tumour samples (where available), blood samples (mandatory), urine and CTC samples (mandatory).*	Analysis and outcome variables yet to be defined.

Table 3 Exploratory objectives

Exploratory objectives	Exploratory outcome variables
To explore whether resistance mechanisms to olaparib can be identified through analysis of tumour and blood samples – archival tumour sample (where available), blood sample at baseline and on disease recurrence (mandatory), and disease recurrence (mandatory), and CTC samples (mandatory).*	Analysis and outcome variables yet to be defined.
CCI [redacted]	Analysis and outcome variables yet to be defined.

*: These exploratory objectives will be reported separately from the clinical study report (CSR) and the details of these analyses will not be specified in this statistical analysis plan (SAP).

1.2 Study design

This is a 2-part study in patients with metastatic castrate-resistant prostate cancer (CRPC). Part A is an open-label safety run-in study to assess the safety, tolerability and pharmacokinetics (PK) of olaparib when given in addition to abiraterone 1000 mg once daily. Part B is a randomised, double-blind, placebo-controlled comparison of the efficacy, safety and tolerability of the dose of olaparib selected from Part A when given in addition to abiraterone, versus placebo in addition to abiraterone.

Abiraterone is indicated in combination with prednisone or prednisolone for the treatment of patients with metastatic CRPC. Prednisone or prednisolone 5 mg twice daily (bid) will be administered with the abiraterone in this study, but throughout this SAP the treatment will be referred to simply as abiraterone.

For Part A of the study, 15 to 18 evaluable patients (Cohorts 1 and 2) are planned to be enrolled from approximately 4 sites in approximately 1 or 2 countries, and a further 12 patients may be recruited into a 3rd cohort if necessary.

For Part B of the study, approximately 140 patients who have received prior chemotherapy containing docetaxel will be randomised from approximately 40 sites in North America and Europe. Patients who have been dosed in Part A of the study may not participate in Part B.

1.2.1 Part A: Safety run-in/dose escalation

Patients will attend the clinic for assessments on the first day of study treatment, at 1 and 2 weeks, then every 4 weeks up to Week 52, and every 12 weeks thereafter. They will also attend for study treatment supplies at 16 weeks and every 4 weeks thereafter.

Cohort 1 (up to 6 patients)

At least 3 and up to 6 evaluable patients will be enrolled in Cohort 1. Patients will receive olaparib 200 mg bid and abiraterone 1000 mg once daily. Dose-limiting toxicities (DLTs) from all patients will be assessed by a Safety Review Committee (SRC) after a minimum of 3 patients have received a minimum of 14 days' treatment.

Cohort 2 (12 patients)

If the combination of olaparib 200 mg bid and abiraterone 1000 mg once daily is tolerated, a cohort of 12 patients (split into 2 groups of 6 patients) will be treated with olaparib 300 mg bid given in addition to abiraterone 1000 mg once daily, and will be evaluated for safety, tolerability and PK. Patients will be recruited to Group 1 and Group 2 concurrently, recruitment will alternate between Group 1 and Group 2.

Group 1:

Patients will receive olaparib alone (300 mg bid) for between 3 and 7 days. Blood samples will be collected to determine the steady state PK profile for olaparib; this may be done on any day between Days 3 and 7, starting after the morning dose.

Patients will then receive abiraterone (1000 mg once daily) starting from the day after the olaparib PK profile has been collected. Olaparib and abiraterone will continue to be dosed in combination for at least 5 days, then blood samples will be collected again to determine both olaparib and abiraterone PK profiles.

Group 2:

Patients will receive abiraterone alone (1000 mg once daily) for between 5 and 7 days. Blood samples will be collected to determine the steady state PK profile for abiraterone (any day between Days 5 and 7).

Patients will receive olaparib (300 mg bid) starting immediately after the 24-hour abiraterone PK sample has been collected. Olaparib and abiraterone will continue to be dosed in combination for at least 3 days, then blood samples will be collected again to determine both olaparib and abiraterone PK profiles.

Dose-limiting toxicities will be assessed by the SRC after a minimum of 14 days' treatment with both olaparib and abiraterone. If <2 DLTs occur in the first 9 patients that have completed at least 14 days' dosing with both olaparib and abiraterone, then the randomised part of the study, Part B, may commence with this dose combination and the remaining 3 patients to be dosed in Part A may be recruited in parallel with Part B. If 2 or more DLTs

occur then all 12 patients must be dosed before a decision is made to progress to Part B. If <4 DLTs occur in this cohort of 12 patients (Groups 1 and 2), then the randomised part of the study, Part B, will commence with this dose combination.

Cohort 3 (12 patients)

If 4 or more DLTs occur in Cohort 2, a further 12-patient cohort may be recruited and treated with olaparib 200 mg bid in combination with abiraterone 1000 mg once daily and evaluated for safety, tolerability and PK as described for Cohort 2.

If Cohort 3 is dosed and there are <4 DLTs, then Part B may proceed using a 200 mg bid dose of olaparib. If 4 or more DLTs are observed, then the study will be terminated.

1.2.2 Part B: Randomised part

Progression to Part B will be based on a review of all available safety, tolerability and PK data from Part A of the study. This will be performed when sufficient evaluable patients in both of the patient groups in Part A Cohort 2 (or Cohort 3 if necessary) have completed a minimum of 14 days' treatment. The selection of the dose for Part B will be taken by the SRC and AstraZeneca. This review may be conducted on pre-database lock data that have not been formally analysed.

Patients will receive olaparib at the dose determined by Part A of the study, and abiraterone 1000 mg once daily.

Patients will be evaluated until disease progression regardless of whether study treatment is discontinued. The collection of Response Evaluation Criteria in Solid Tumours version 1.1 (RECIST 1.1)/prostate cancer working group 2 (PCWG-2) data will then stop and patients will be followed for second progression (PFS2) and survival.

The primary analysis will be performed once approximately 100 progression events have occurred.

1.3 Number of subjects

1.3.1 Part A: Safety run-in/dose escalation

The primary objective of Part A of this study is to assess the safety and tolerability of olaparib when combined with abiraterone and to recommend a dose of olaparib for study in Part B. Hence the number of patients has been based on the desire to obtain adequate tolerability, safety and PK data while exposing as few patients as possible to the investigational product (IP) and procedures.

Cohorts of 3 to 6 evaluable patients will be required. The total number of patients will depend upon the number of dose escalations/reductions necessary.

The cohort sizes are based upon accepted methodology for Phase I oncology studies.

1.3.2 Part B: Randomised part

The sample size for Part B of this study was selected to be consistent with the hypothesis that administration of olaparib when given in addition to abiraterone will result in greater efficacy, as determined by radiological progression free survival (rPFS), in patients with metastatic CRPC than treatment with abiraterone alone.

In total, 100 rPFS events in the study would provide approximately 80% power to show a statistically significant rPFS at the 1-sided 10% level if the assumed true treatment effect was a hazard ratio (HR) of 0.65; this translates to a 3.75 month benefit in median rPFS on olaparib/abiraterone combination over 7 months (based on the mean of 2 trials: de Bono et al 2011 and Scher et al 2012) on abiraterone monotherapy if rPFS is exponentially distributed. Given the number of rPFS events stays unchanged, the power to find the proposed effect size to be significant at the 1-sided 2.5% level drops down to approximately 58% and the smallest detectable effect becomes HR=0.67 (which translates to a 3.4 months improvement in the olaparib/abiraterone combination over the abiraterone/placebo arm).

Approximately 140 patients will be recruited (1:1 ratio) so that data maturity for the rPFS analysis is approximately 70% at the time of the primary analysis. Assuming 12 months non-linear recruitment, 100 rPFS events are expected to occur approximately 24 months after the first patient is enrolled in the study (FPI) (12 months accrual + 12 months follow-up). This will be the primary analysis of rPFS. With 100 events, the smallest treatment difference that would be statistically significant at the 1-sided 10% level is rPFS HR= 0.77 (which translates to approximately a 2.1 months median difference). `

Overall survival will be analysed at the time of the rPFS primary analysis. At this time we expect approximately 60% of the patients to have died (84 deaths). Eighty four death events gives 80% power, to show a statistically significant OS effect at the 10% 1-sided level if the true hypothesised improvement in median time to death of 10 months on olaparib/abiraterone combination over 17 months on abiraterone alone (HR=0.63). Keeping the event number fixed, the power to detect a similar effect at the 1-sided 2.5% significance goes down to approximately 56%. Note, if the true OS improvement is similar to the anticipated rPFS improvement (approximately 4 months median improvement represented by a HR of 0.81) then this study is not powered to detect an OS difference at the 1-sided 10% significance level.

The primary and all secondary efficacy endpoints will be analysed in the full analysis set unless otherwise stated. The safety endpoints will be analysed in the safety analysis set.

For sample size information relating to the pharmacogenetic component of the study, please see the clinical study protocol Appendix D (‘Pharmacogenetics research’).

2. ANALYSIS SETS

2.1 Definition of analysis sets

The study physician, pharmacokineticist, and statistician will agree on the strategy for dealing with data affected by protocol deviations before any formal statistical analysis is performed and prior to the unblinding of part B.

2.1.1 Full analysis set

Intention to treat (ITT): The primary statistical analysis of the efficacy of olaparib when given in addition to abiraterone will include all randomised patients in Part B, 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.

2.1.2 PK analysis set

All patients who receive at least one dose of olaparib per the protocol, for whom there are at least one reportable PK concentration and who have no important protocol deviations or adverse events that impact PK on all PK Days, will be included in the PK analysis set.

2.1.3 Safety analysis sets

Part A safety analysis set: All patients in Part A of the study who receive at least 1 dose of olaparib. Treatment group comparisons will be based on the initial dose of olaparib actually received.

Part B safety analysis set: All patients randomised into Part B of the study who receive at least 1 dose of olaparib/placebo. Treatment group comparisons will be based on the initial dose of olaparib/placebo actually received.

When assessing safety and tolerability, summaries will be produced based on the safety analysis set.

2.2 Violations and deviations

Deviations from the inclusion/exclusion criteria detailed in section 4 of the protocol are not allowed in either part of the study and will be considered important deviations.

2.2.1 Part A

Protocol deviations will be assessed on an ongoing basis to ensure there are sufficient evaluable patients for the decision points detailed in Section 1.2.1.

Important protocol deviations include changes to the procedures that may impact the quality of the data or any circumstances that can alter the evaluation of the PK. Examples include, but are not be limited to, sample processing errors that lead to inaccurate bioanalytical results, incomplete dose administered, incomplete PK profile collected, and/or use of disallowed concomitant medication. In the case of an important protocol deviation which could impact

the robustness of the data, affected PK data collected will be excluded from the descriptive summary statistics but will still be reported in the study result listings. Important deviations will be listed and discussed in the CSR.

2.2.2 Part B

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 for study.

- Patients randomised but who did not receive olaparib/matching placebo.
- Patients who deviate from entry criteria, which will be documented ahead of database lock.
- Baseline RECIST scan > 28 days before study treatment is started.
- Baseline RECIST scans after randomisation.
- Patients who have RECIST scans outside of a scheduled visit window on > 2 occasions.
- Disallowed concomitant medication use.

The categorisation of these as important deviations is not automatic and will depend on duration and the perceived effect on efficacy and safety. 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. For example, details of disallowed concomitant medication use will be reviewed by a physician using blinded data and may be determined as key.

Misrandomisations in terms of errors in treatment dispensing, in addition to incorrect stratifications, will also be summarised and listed separately to the important protocol deviations. 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.

The summary will include all patients with a dispensing error but will also include information on how many of those patients received at least one dose of the wrong treatment (olaparib/placebo) at any time. 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/or statistician.

3. PRIMARY AND SECONDARY VARIABLES

For Part A, the primary assessment is the safety and tolerability of olaparib when given in addition to abiraterone and to recommend, by assessment of DLTs and other safety and tolerability data, a dose of olaparib for further study when given in addition to abiraterone. See section 3.3 of this SAP for further details.

For Part B, the primary assessment of efficacy is rPFS, defined as disease progression according to RECIST 1.1 (for soft tissue disease) and/or PCWG-2 criteria (for bone disease), or death. To ensure comparability, identical imaging techniques should be used for the assessment of response at baseline and throughout the study. Further details of the methods used to determine the RECIST response are detailed below and also in Appendix F of the clinical study protocol.

3.1 Derivation of RECIST visit responses – malignant soft tissue

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

Baseline radiological tumour assessments are to be performed no more than 28 days before the start of randomised treatment. Tumour assessments are then performed every 12 weeks following randomisation up to and including Week 72. Thereafter, assessments will be conducted every 24 weeks until objective progression.

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

3.1.1 RECIST Target lesions (TLs)

Measurable disease is defined as having at least one measurable lesion which is ≥ 10 mm in the longest diameter (except lymph nodes which must have short axis ≥ 15 mm) with CT or

MRI. A patient can have a maximum of 5 measurable lesions recorded at baseline and these are referred to as TLs. If more than 1 baseline scan is recorded then measurements from the one that is closest to randomisation will be used to define the baseline sum of TLs.

Table 4 provides details for TL visit responses.

Table 4 TL visit responses

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

Rounding of TL data

For calculation of progressive disease (PD) and partial response (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 then all TL measurements should be recorded. However, a visit response of PD should still be assigned if any of the following occurred:

- A new lesion is recorded

- A NTL visit response of PD is recorded

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

If any TL measurements are missing, the visit response will automatically be set to not-evaluable (NE) unless lesion measurements are missing due to an intervention (such as irradiation/palliative surgery/embolisation) during the course of the study, in which case they will be handled using the 3 steps outlined below.

TL Visit responses subsequent to CR

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

- Step 1: If all lesions meet the CR criteria (i.e., 0 mm or < 10 mm for lymph nodes) 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 longest diameter (LD) increases by 20% but remains < 10 mm.

- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0 mm or < 10 mm for lymph nodes) 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 lesions meet the CR criteria and the sum of lesions meets the criteria for PD then response will be set to PD

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

Lymph nodes

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

TL too big to measure

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

TL too small to measure

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

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation/palliative surgery/embolisation), should be handled in the following way and once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of 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. If the scaling-up results in a visit response of PD then the patient would be assigned a TL response of PD.

  Note: scaling up will only be performed if ≤ 1/3 of the TL measurements have been set to missing following an intervention. If > 1/3 of TL measurements are set to missing then TL response will be set to NE).

- Step 3: If after both steps PD has not been assigned, then if appropriate, a scaled sum of diameters will be calculated, treating the lesion with intervention as missing, and PR or stable disease (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 the lesions that have been patient to intervention also has a value of 0 recorded. If scaling-up is not appropriate due to too few non-missing sizes then the visit response will be set to NE.

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

Table 5 provides an example of scaling-up.

Table 5 Example of scaling-up

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

Lesion 5 has received intervention prior to the follow-up visit and has been set to missing. 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:

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

Lesions that split in two

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

Lesions that merge

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

Change in method of assessment of target lesions

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.

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

Non-target lesion response will be derived based on the Investigator's overall assessment of NTLs as follows:

Progressive disease: Unequivocal progression of existing NTLs, which 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.

Complete response:	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10mm short axis).
Non-CR/Non-PD:	Persistence of one or more NTLs with no evidence of progression.
Not evaluable:	Only relevant when one or some of the NTLs have not been assessed and in the Investigator's opinion they are not able to provide an evaluable overall NTL assessment.
Not applicable:	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 1 or more new soft-tissue lesions is assessed as progression.

A soft-tissue 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.

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

Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping study therapy.

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

3.1.3 RECIST overall visit response

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

Table 6 Overall visit response for soft tissue lesions

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

3.1.4 RECIST Best objective response

RECIST best objective response (BOR) for soft tissue represents the best response for malignant soft tissue a patient has had during their time in the study up until disease progression, or last evaluable assessment in the absence of progression.

Each patient will be assigned a best objective response of CR, PR, SD, PD, NE or no evidence of disease (NED), which will be calculated as follows:

CR: Any overall visit response of CR

PR: Any overall visit response of PR in the absence of CR

SD: Stable disease recorded at least 12 weeks - 1 week, i.e. at least 77 days after the date of randomisation in the absence of CR/PR. Stable disease recorded for less than 77 days is assigned to NE.

PD: Progression or Death in the absence of CR/PR or SD

NED: No evidence of disease at all visits

NE: No evidence of CR/PR or SD or PD or Death

3.2 Evaluation of PCWG2 progression status – bone lesions

Metastatic bone disease status will be reported from the bone lesion assessment on bone scan as non-progressive disease (non-PD), PD and NE, separately from RECIST 1.1 soft tissue and PSA response assessments, according to PCWG-2 criteria.

Bone lesions will be assessed by bone scintigraphy (bone scans). Guidelines for the assessment of bone lesions are presented in protocol Appendix F. A baseline bone scan should be performed no more than 28 days before randomisation, and ideally should be performed as close as possible to the start of study treatment. Subsequent assessments will be performed every 12 weeks (±1 week) from the start of treatment until disease progression, death, or withdrawal of consent (every 24 weeks after Week 72).

Bone lesions will be assessed by bone scan and will not be part of the RECIST v1.1 malignant soft tissue assessment. Positive hot spots (if considered to constitute bone metastases) are on the bone scan should be considered significant and unequivocal sites of malignant disease to be recorded as metastatic bone lesions.

If an unscheduled assessment, e.g., confirmation of progression scan, is performed and there is no evidence of disease progression, every attempt should be made to perform the subsequent assessments at their scheduled visits while the patient remains on study treatment.

Progression on a bone scan is assessed as:

- **At the 12 week scan:**

  If 2 or more new metastatic bone lesions are observed on the first 12-week scan, the confirmatory scan performed, preferably at the next scheduled visit for a bone scan (i.e., Week 24) and at least 6 weeks later (week 18 or later), must show 2 or more additional new metastatic bone lesions (for a total of 4 or more new metastatic bone lesions since the baseline assessment) for progression to be documented.

  As bone scans at Week 12 do not have a PCWG-2 response (because a confirmatory scan is required), they will be classified as non-PD, unless Bone Lesion Status cannot be determined from that scan.

- **After the 12 week scan:**

  If 2 or more new metastatic bone lesions are observed on scans obtained after the first 12-week assessment, confirmatory scan performed preferably at the next scheduled visit for a bone scan (e.g., Week 36) and at least 6 weeks later needs to show the persistence of or an increase in the number of metastatic bone lesions compared to the prior scan.

  Additionally, for any scan with 2 or more new bone lesions that does not have a subsequent scan, the response will be classified as non-PD, however is unusable and deemed not contributing for the primary endpoint.

The date of progression is the date of the first scan that shows the change.

If the Investigator is in doubt as to whether progression has occurred, it is advisable to continue study treatment and reassess the bone lesion status at the next scheduled assessment, or sooner if clinically indicated.

It is important to follow the assessment schedule as closely as possible.

Table 7 provides the definitions of the criteria to determine tumour visit response for bone lesions.

Table 7 Overall PCWG-2 progression status (bone)

Non–progressive disease (non-PD)	Persistence of one or more bone lesions, disappearance of all bone lesions or a bone scan with 2 or more new lesions that is not confirmed.
Progressive disease (PD)	Bone lesions fulfilling the requirements for new lesions and confirmation of progression
Non evaluable (NE)	Bone scan not performed at the visit or image quality not sufficient for evaluation

Disease progression will only occur when a response of progressive disease is recorded at two visits at least 6 weeks apart. The date of the progression is designated as the first scan showing 2 or more new bone lesions.

Disease progression and the date of disease progression will be recorded on the eCRF by the investigator. Progression will be determined by the investigator using the definition of progression on bone scan as described above and in Appendix F of the clinical study protocol.

3.3 Primary variable – radiologic progression free survival (rPFS)

The primary endpoint in Part B is rPFS (defined by RECIST 1.1 and/or PCWG-2 as assessed by the Investigator, and expressed in months). The visit responses for the soft tissue assessment (based on RECIST 1.1) will be used in conjunction with the PCWG-2 bone scan evaluations assessed by the investigator, in that a patient is progression-free only if they are not PD by RECIST 1.1 or PCWG-2.

Table 8 defines how the visit responses for soft tissue (according to RECIST1.1 criteria) and bone (according to PCWG-2 criteria) will be combined with bone lesion information to give an overall radiological visit response.

Table 8 Overall radiological visit response

RECIST1.1 visit response (soft tissue/visceral)	PCWG-2 progression status (bone)	Bone lesions Present/Absent	Overall radiological visit response
CR	Non-PD	Absent	CR
CR	Non-PD	Present	PR
CR	NE [a]	Any [b]	PR
PR	Non-PD or NE [a]	Any	PR
SD	Non-PD or NE [a]	Any	SD
NED [c]	Non-PD	Any	non-PD
NED [c]	NE [a]	Any	NE
NE [d]	Non-PD or NE [a]	Any	NE
PD	Any	Any	PD
Any	PD	Any	PD[e]

a - From an imaging point of view, a PCWG2 visit response of 'NE' will mainly be because a patient did not have a scan at all at that time-point or a bone scan was done however there were issues with the image quality, e.g. poor contrast, the wrong acquisition parameter was used.

b - 'Any' refers to any result whether measured or missing.

c - NED – No evidence of soft tissue disease at all visits.

d - NE – Not evaluable.

e - Following the confirmation of PD at the subsequent bone scan, the overall radiological visit response of progression for the visit will be programmatically derived as PD using the Progression Date (Previous Bone scan date where new lesions first seen) captured on the eCRF.

Radiologic progression-free survival (months) is defined as the time from randomisation until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to progression.

Patients will undergo tumour assessments at baseline, and thereafter every 12 weeks following randomisation up to and including Week 72 (every 24 weeks thereafter), until objective disease progression or death. The rPFS time will always be derived based on scan/assessment dates, not visit dates.

Patients who are not known to have progressed (defined as CR, PR, SD or NED by RECIST 1.1 and/or non-PD by PCWG-2 bone scan) or died at the time of analysis will be censored at the time of the latest date of assessment from either their last evaluable RECIST 1.1 assessment or their last evaluable PCWG-2 assessment. Also, if the patient progresses or dies after 2 or more missed visits (i.e. 2 or more visits where neither a RECIST 1.1 visit scan nor a PCWG-2 visit scan were done; not evaluable scans that were completed are not considered "missed") the patient will be censored at the time of the latest evaluable RECIST 1.1 or

PCWG-2 assessment. If the patient has no evaluable visits or does not have baseline data they will be censored at Day 1 unless they die within 2 visits of baseline (in which case their date of death will be used).

When the Investigator is in doubt as to whether PD has occurred and therefore reassesses the patient at a later date, the date of the initial scan should be declared as the date of progression if the repeat scans confirm progression.

If a patient discontinues treatment prior to progression then the patient will continue to be followed until evidence of objective disease progression as defined by RECIST 1.1 or PCWG-2 occurs and their rPFS time will be derived as defined above.

RECIST 1.1/PCWG-2 assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

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

3.4 Secondary outcome variables

3.4.1 Time to second progression (PFS2)

PFS2 (months) 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 rPFS, or death. The date of PFS2 will be recorded by the Investigator and defined according to local standard clinical practice and may involve any of: objective radiological progression, symptomatic progression, rises in PSA or death. This is calculated for Part B only.

PFS2 status will be reviewed every 12 weeks following the progression event used for the primary variable rPFS (the first progression) and investigator opinion of progression status (progressed or non-progressed) at each assessment will be recorded in the electronic case report form (eCRF). RECIST 1.1 and PCWG-2 assessments will not be required for assessment of PFS2. 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 rPFS or PFS2 assessment date if the patient has not had a second progression or death).

PFS2 status will be assessed during the 2 weeks following the DCO.

3.4.2 Overall survival (OS)

Overall survival (months) 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. This is calculated for Part B only. This should be based on information in the date subject last known to be alive

variable recorded within the survival status module of the eCRF (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note, survival calls will be made in the 2 weeks following the date of data cut-off (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 in the analyses as of the date of DCO. Death dates may be found by checking publicly available death registries.

Patients will enter the OS phase of the study following confirmed progression by the Investigator. All patients will be followed for survival data until approximately 60% of patients have died.

3.4.3 Objective response rate (ORR) and associated variables

Objective response rate (ORR) rate is defined as the number (%) of patients with at least one visit response of CR or PR. Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of response rate. This will be irrespective of whether or not patients discontinued treatment. Any responses observed after the start of the first cancer therapy subsequent to treatment discontinuation will not be included.

For soft tissue disease ORR, only patients with measurable disease (target lesions) at entry will be included in the denominator. A responder will be any patient with a RECIST best overall response of PR or CR in their soft tissue disease assessed by RECIST 1.1, irrespective of the bone scan status assessed by PCWG-2.

A soft tissue visit response of CR is defined when all TL and NTL lesions present at baseline have disappeared (with the exception of lymph nodes which must be <10 mm to be considered non-pathological) and no new lesions have developed since baseline. A soft tissue visit response of PR is defined when the sum of diameters of the TLs has decreased by 30% or more compared with baseline (with no evidence of progression) and the NTLs are at least stable with no evidence of new lesions.

For overall disease ORR, again only patients with visceral measurable disease (TLs) at entry will be included in the denominator whether or not they have documented metastatic bone scan at entry. A responder demonstrating overall radiological response will be any patient with a best overall response of PR or CR in soft tissue disease assessed by RECIST 1.1 and also bone scan status of non-PD or NE for their bone scans assessed by PCWG-2. Soft tissue response rate will thus be equal to or higher than the overall radiological ORR.

Duration of overall radiological response (RECIST 1.1 soft tissue disease response of PR or CR, with a PCWG-2 response of non-PD/NE) will be defined as the time from overall radiological response until date of documented radiological progression by RECIST 1.1 or PCWG-2, 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 rPFS endpoint. The time of the initial overall radiological response will be defined as the latest of the dates contributing

towards the first visit where an overall radiological response is seen. The time to the initial overall radiological response will be defined as the time from randomisation until the earliest of the dates where an overall radiological response is seen.

If a patient does not progress following an overall radiological response, then their duration of response (DOR) will be censored at the rPFS censoring time.

Duration of response will not be defined for those patients who do not have documented overall radiological response or visceral measurable disease at entry.

3.4.4 Time to subsequent therapies

Subsequent anti-cancer therapies, which does not include radiotherapy, will be reviewed prior to unblinding to assess which represent clinically important treatments intended to control prostate cancer. As a supportive summary to rPFS, time to start of first subsequent anti-cancer therapy or death (TFST) will be assessed. TFST is defined as the time from randomisation to the earlier of first subsequent anti-cancer 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 subsequent therapy, i.e. the last follow-up visit where this was confirmed.

As a supportive summary to rPFS, time to start of second subsequent anti-cancer therapy or death (TSST) will be assessed. TSST is defined as the time from randomisation to the earlier of the second subsequent anti-cancer 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.

3.5 Safety

3.5.1 Dose limiting toxicities (DLTs)

Dose limiting toxicities will be reviewed in Part A, the incidence of DLTs will be summarised by cohort and listed.

A DLT is defined as any toxicity which is not a recognised adverse effect of abiraterone or prednisolone, and not attributable to the disease or disease-related processes under investigation, which occurs during a minimum of 14 days' treatment and which includes:

Haematological toxicity ≥ common terminology criteria for adverse events (CTCAE) Grade 4 present for more than 4 days

- Except anaemia

Non-haematological toxicity ≥ CTCAE Grade 3 including:

- Infection including febrile neutropenia
- Corrected QT interval (QTc) prolongation (> 500 msec)

Any other toxicity that is greater than that at baseline, is clinically significant and/or unacceptable, does not respond to supportive care, results in a disruption of dosing schedule of 7 days or more, or is judged to be a DLT by the SRC.

3.5.2 Adverse Events

The definitions of adverse events (AEs) and serious AEs (SAEs) are given in Sections 6.4.1 and 6.4.2 of the clinical study protocol. AEs will be listed for each patient and summarised by treatment received according to the system organ class (SOC) and preferred term assigned to the event using the Medical Dictionary for Regulatory Activities (MedDRA). AEs will be graded according to the National Cancer Institute CTCAE (version 4.0). The CTCAE grade will be assigned by the investigator as will an assessment of causality. All summaries will include treatment emergent events, defined as adverse events with an onset date on or after the first dose and up to and including 30 days following the date of last dose of study medication, unless otherwise stated

Grouped AEs

Specific preferred terms describing similar clinical concepts will be considered collectively as Grouped AEs. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician will confirm the list of Grouped AEs and the specific preferred terms contributing to each grouping at blind data review.

AEs of special interest

Adverse events of special interest (AESIs) for olaparib are events of scientific and medical interest specific to the further understanding of the olaparib safety profile. An AESI may be serious or non-serious. For olaparib these AESI's have been identified as the Important Potential Risks of myelodysplastic syndrome (MDS)/acute myeloid leukemia (AML), new primary malignancy (other than MDS/AML), pneumonitis and diffuse alveolar damage. The preferred terms associated with these AESIs are listed in Appendix B.

Some clinical concepts (including Grouped AEs and some selected individual preferred terms) are considered "AEs of interest". An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician will confirm the list of AEs of special interest (to include Grouped AEs and the specific preferred terms) at prior to unblinding.

3.5.3 Laboratory data

Blood samples will be used for determination of clinical chemistry and haematology. Clinical chemistry and haematology will be taken at all scheduled visits. The laboratory parameters to be collected are given in the protocol Section 6.4.5 of the clinical study protocol.

Change from baseline in haematology and clinical chemistry variables will be calculated for each post-dose visit on treatment. Baseline will be the last non-missing assessments prior to initiation of olaparib/placebo, in most cases this will be the assessments performed on Day 1. CTCAE grades will be defined at each visit according to the CTCAE grade criteria using AZ standard ranges as required, after conversion of laboratory results to corresponding SI units. The following parameters have CTCAE grades defined for both high and low values: potassium, sodium, magnesium and calcium so high and low CTCAE grades will be calculated. CTCAE grades are not defined for total protein, urea nitrogen, absolute eosinophil count, absolute basophil count, absolute monocyte count.

Absolute values will be compared to the AZ standard reference range and classified as low (below lower limit), normal (within range or on limits of range) and high (above upper limit). The maximum or minimum on-treatment value (depending on the direction of an adverse effect) will be defined for each laboratory parameter as the maximum (or minimum) post-dose value at any time.

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

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

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

Any data collected at unscheduled visits (including retests) will be listed, included in the derivation of baseline and also in the derivation of maximum overall CTCAE grade during treatment and maximum value during treatment, but will not be included in any summaries by visit.

3.5.4 Vital signs

Vital sign assessments (resting blood pressure [BP], pulse rate, body temperature), including weight, will be performed as per the study. Height will be assessed at screening only.

Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. The denominator in vital signs data should include only those patients with recorded data. Any data collected at unscheduled visits will be listed and included in the derivation of baseline, but will not be included in any summaries by visit.

3.5.5 Electrocardiograms (ECG)

Twelve-lead ECG will be performed as per the study plan. Patients should be supine and at rest 5 minutes prior to recording the ECG. At each time point the Investigator's assessment of the ECG (normal, borderline or abnormal) and heart rate, duration of QRS complex, respiratory rate (RR), pulse rate and QT intervals will be collected.

For triplicate ECGs, the mean of the 3 ECG assessments will be used to determine the value at that time point.

QTcF will be derived during creation of the reporting database using the reported ECG values (RR and QT).

QTcF = QT / RR^ (1/3) where RR is in seconds

3.5.6 Physical examination

A physical examination will be performed as per the study plan.

3.5.7 Duration of exposure

Study drug exposure (days) will be defined as time from first dose of olaparib, up to and including the, last day that the dose of olaparib was greater than 0mg. Exposure to abiraterone will be calculated in the same way.

Exposure will be defined as: Last dose date – first dose date + 1.

Actual exposure (days) to study drug, defined as time from first dose of olaparib, up to and including the, last day of dosing minus total number of days where the dose was 0 (including missed or forgotten doses), will also be calculated. Actual exposure to abiraterone will be calculated in the same way.

The number of days each patient took the assigned dose of study drug in part B, (this will depend on the result of part A but is assumed to be 300mg olaparib/placebo b.i.d), will also be calculated and summary statistics of this data will be presented.

Percentage compliance will be defined as:

{(No. tablets dispensed in period – no. tablets returned from period)/
(no. days of study drug exposure in period * expected tablets per day)} * 100%

where expected tablets per day will take into account once daily or bid dosing.

Overall compliance may be calculated over various periods if the dose has been modified, to take into account the differing expected tablets per day or the protocol-specified dose interruptions. Missed doses will not be adjusted for; the overall compliance will be reduced.

For summaries, compliance will be capped at 100%, however uncapped compliance rates will be listed by subject.

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 two years or until the date

of objective disease progression (whichever is earliest) as defined by RECIST and/or PCWG-2 criteria 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.

RDI and 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.

3.6 Pharmacokinetic variables and PK analysis methods

The PK analysis of the plasma concentration data for olaparib and abiraterone will be performed by Covance Clinical Pharmacokinetic Alliance (CPKA) on behalf of AstraZeneca R&D. The actual sampling times will be used in the PK calculations, except for the pre-dose sample for which the time will be set to zero. All PK computations will be performed using Phoenix™ WinNonLin® v6.4 for non-compartmental analysis (NCA).

Patients who withdraw from the study following dosing, but prior to study completion, will be included in the PK analysis provided they have evaluable profiles over the planned collection period. Patients with partial data will be evaluated on a case-by-case basis to determine if sufficient data are available for meaningful analysis.

Where plasma concentration-time data allow the following steady state PK parameters will be determined in Part A for olaparib (Cohort 2 Group 1 Visits 3 and 4 and Cohort 2 Group 2 Visit 4) and abiraterone (Cohort 2 Group 1 Visit 4 and Cohort 2 Group 2 Visits 3 and 4): $C_{ss,max}$, $C_{ss,min}$, $t_{ss,max}$, AUC_{ss}, AUC_{0-t}, $t_{ss,last}$, CL_{ss}/F and ratios of $C_{ss,max}$, $C_{ss,min}$ and AUC_{ss} for each analyte in combination compared to monotherapy.

Plasma concentrations that are non-quantifiable (NQ) (i.e. are below the LOQ limit of quantification) from the time of pre-dose sampling (t=0) up to the time of the first quantifiable concentration will be set to a value of zero. After this time point, NQ plasma concentrations will be set to missing for all concentration profiles. Where two or more consecutive concentrations are NQ at the end of a profile, the profile will be deemed to have terminated and therefore any further quantifiable concentrations will be set to missing for the calculation of the PK parameters unless it is considered to be a true characteristic of the profile of the drug. If an entire concentration-time profile is NQ, the profile will be excluded from the PK analysis.

$C_{ss,max}$, $C_{ss,min}$, $t_{ss,max}$ and $t_{ss,last}$ will be determined by visual inspection of the concentration-time profiles. $C_{ss,min}$ will be taken as the concentration in the pre-dose sample taken immediately prior to dosing on the PK day. AUC_{ss} will be calculated using the linear trapezoidal method when concentrations are increasing and the logarithmic trapezoidal method when concentrations are decreasing (linear up, log down). The minimum requirement for the calculation of AUC_{ss} will be the inclusion of at least three consecutive plasma concentrations above the lower limit of quantification (LLOQ), with at least one of these concentrations following $C_{ss,max}$. The apparent steady state clearance (CL_{ss}/F) will be determined from the ratio of dose/AUC_{ss}. Ratios of $C_{ss,max}$, $C_{ss,min}$ and AUC_{ss} for each analyte in combination compared to monotherapy will be determined in Phoenix using non log-transformed parameter data .

3.7 Calculation or derivation of biomarker variable(s)

3.7.1 Prostate specific antigen

The distribution of the PSA data will be assessed prior to unblinding and if necessary, a log-transformation may be applied.

3.7.1.1 PSA response

- A patient will be regarded as having a single PSA visit response if their PSA level at any post-dose visit is reduced by 50% or more compared with baseline
- A patient will be regarded as having a confirmed PSA response if they have a reduction in PSA level of 50% or more compared with baseline that is confirmed at the next assessment at least 4 weeks later (i.e., decrease relative to baseline of at least 50% documented on 2 consecutive occasions at least 4 weeks apart).

3.7.1.2 PSA changes on continuous scale

- PSA levels will be evaluated in terms of percentage change from baseline which will be derived for each post baseline visit where PSA data are available
- % change = [(post-dose PSA level – baseline PSA level) / baseline PSA level] *100
- Best percentage change from baseline in PSA will be derived as the biggest reduction in PSA level compared with baseline (or the smallest increase in the absence of a reduction) taking account of all PSA values collected for each patient.

3.7.2 Circulating tumour cell counts (CTCs)

- Conversion: change of circulating tumour cells (CTC) from ≥5 at baseline to <5 post baseline
- Absolute change in CTC counts as a continuous variable following treatment with olaparib/placebo and abiraterone (calculated as post-dose CTC count - baseline CTC count).

- Absolute change in CTC count at 12 weeks.

3.7.3 ATM mutation

Patient tumour samples will be assessed for the presence of Ataxia telangiectasia (ATM) mutation utilising a targeted NGS assay (Foundation Medicine). Results from the tumour analysis will be used to classify the data into the categories as listed in Table 9 to define the subgroups for analysis. The tumour-derived ATM mutation call may be supplemented by germline data if that becomes available during study (e.g. if noted on eCRF). The presence of ATM mutation from any source is dominant (whether considered of germline origin or tumour-specific).

Table 9 Description of ATM mutation categories

Category	Description
Detected	Patient with a deleterious or suspected deleterious ATM mutation identified in their tumour.
Variant of unknown significance (*VUS*)	Patient who was not ATM mutated as defined above, but tumour analysis indicated the presence of a variant of unknown significance.
Not Detected	Patient who was not ATM mutated/ATM VUS as defined above, but tumour analysis was performed successfully and no mutation was detected
Failed	Patient for whom a tumour sample was analysed but the analysis failed to provide sufficient data to enable classification of their ATM mutation status.
Status unknown	Patient did not have an ATM result recorded from tumour analysis.

3.7.4 BRCA mutation

Patient tumour samples will be assessed for the presence of breast cancer gene (BRCA1 or BRCA2) mutation utilising a targeted NGS assay (Foundation Medicine). Results from the tumour analysis will be used to classify the data into the categories as listed in Table 10 to define the subgroups for analysis. The tumour-derived BRCA mutation call may be supplemented by germline data if that becomes available during study (e.g. if noted on eCRF). The presence of BRCA mutation from any source is dominant (whether considered of germline origin or tumour-specific).

Table 10 Description of BRCA mutation categories

Category	Description
Detected	Patient with a deleterious or suspected deleterious BRCA (BRCA1 or BRCA2) mutation identified in their tumour.

Category	Description
Variant of unknown significance (*VUS*)	Patient who was not BRCA mutated as defined above, but tumour analysis indicated the presence of a variant of unknown significance.
Not Detected	Patient who was not BRCA mutated/BRCA VUS as defined above, but tumour analysis was performed successfully and no mutation was detected
Failed	Patient for whom a tumour sample was analysed but the analysis failed to provide sufficient data to enable classification of their BRCA mutation status.
Status Unknown	Patient did not have a BRCA result recorded from tumour analysis or a BRCA result recorded in the eCRF.

3.7.5 Composite BRCA/ATM mutation

Results from the tumour analysis for BRCA1, BRCA2 and ATM mutation as described above in sections 3.7.3 and 3.7.4 will be used to classify the data into composite categories as shown in Table 11 to define the subgroups for the analysis.

Table 11 Description of composite BRCA/ATM mutation categories

Category	Description
Detected	Patient with a deleterious or suspected deleterious mutation identified in their tumour in at least one of the BRCA1, BRCA2 and ATM genes.
Variant of unknown significance (*VUS*)	Patient was not mutated as defined above, but tumour analysis indicated the presence of a variant of unknown significance in at least one of the BRCA1, BRCA2 and ATM genes.
Not Detected	Patient was not mutated/VUS as defined above, but tumour analysis was performed successfully and no mutation was detected in the BRCA1, BRCA2 and ATM genes.
Failed	Patient for whom a tumour sample was analysed but the analysis failed to provide sufficient data to enable classification of their mutation status.
Status Unknown	Patient did not have a result recorded from tumour analysis or a result recorded in the eCRF.

3.7.6 Composite HRR mutation (excluding BRCA1, BRCA2 and ATM)

Patient tumour samples will also be assessed for the presence of an additional panel of 12 HRR gene mutations (denoted as follows: BARD1, BRIP1, CDK12, CHEK1, CHEK2, FANCL, PALB2, PPP2R2A, RAD51B, RAD51C, RAD51D, and RAD54L) utilising a targeted NGS assay (Foundation Medicine). Results from the tumour analysis will be used to classify the data into composite HRR categories as shown in Table 12 to define the subgroups for the analysis.

Table 12 Description of composite HRR mutation categories

Category	Description
Detected	Patient with a deleterious or suspected deleterious mutation identified in their tumour in at least one of the 12 HRR genes (as denoted above).
Variant of unknown significance (*VUS*)	Patient was not mutated as defined above, but tumour analysis indicated the presence of a variant of unknown significance in at least one of the 12 HRR genes.
Not Detected	Patient was not mutated/VUS as defined above, but tumour analysis was performed successfully and no mutation was detected in all 12 HRR genes.
Failed	Patient for whom a tumour sample was analysed but the analysis failed to provide sufficient data to enable classification of their mutation status.
Status Unknown	Patient did not have a result recorded from tumour analysis.

3.8 Exploratory variables

3.8.1 Patient reported outcomes (PRO) and HRQL

Patient reported outcomes (PRO) will be measured using the Brief Pain Inventory – short form (BPI-SF), and the individual item on bone pain and the Functional Assessment of Cancer Therapy - Prostate Cancer (FACT-P;- Esper et al 1997, Cella et al 1993) questionnaires.

3.8.1.1 Functional Assessment of Cancer therapy – Prostate Cancer (FACT-P)

The patient-reported FACT-P will be used to assess health-related quality of life. The questionnaire will be administered in Part B of the study, at baseline, Week 4, 8 and 12 and then continue to be administered to all patients (who have not withdrawn consent) every 12 weeks.

The following outcome measures will be calculated from the FACT-P questionnaire, the resulting value is the total score for the associated questions or scaled scores:

- Physical well-being subscale (PWB) (Questions GP1 to GP7)
- Social/family well-being subscale (SWB) (Questions GS1 to GS7)
- Emotional well-being subscale (EWB) (Questions GE1 to GE6)
- Functional well-being subscale (FWB) (Questions GF1 to GF7)
- Prostate cancer subscale (PCS) (Questions C2, C6, P1 to P8, BL2 and BL5)
- Trial Outcome Index (TOI), sum of PWB, FWB and PCS
- Functional Assessment of Prostate Cancer Symptoms Index 8 (FAPSI-8) (Questions P1 to P3, GP1, C2, P7, P8 and GE6)
- Functional Assessment of Prostate Cancer Symptoms Index 6 (FAPSI-6) (Questions P1 to P3, GP1, C2 and GE6).
- FACT-P total score (sum of scores of all the sub-scales)

Items to be reversed:

- Each question in the FACT-P questionnaires has a choice of 5 responses, “Not at all”, “A little bit”, “Somewhat”, “Quite a bit” and “Very much”. The scores range from 0 (“Not at all”) to 4 (“Very much”) for positively phrased questions. Negatively phrased questions have a reverse scoring, from 0 (“Very much”) to 4 (“Not at all”). This results in a consistent approach, where higher scores indicate a better quality of life.
- Note, questions that are reversed (via subtraction of the response from 4) are: GP1-7, GE1, GE3-6, C2, P1-3, P6-P8 and BL2.

Missing data

As per the FACIT scoring guidelines (Cella et al 1993, Cella 1994, Esper et al 1997),

- More than 80% of questions in a questionnaire must be completed for the questionnaire to have the FACT-P total score evaluable. If 80% or less of questions are completed, the FACT-P total scores will not be calculated.
- For each domain (PWB, SWB, EWB, FWB and PCS) if more than 50% of the items were answered (e.g., a minimum of 4 of 7 items, 4 of 6 items, etc), the subscale score will be calculated by multiplying the sum of subscale by the number of items in the subscale, then dividing by the number of items actually answered:

Subscale score= (sum of item scores x N of items in subscale)/ N of items answered

- If at least 50% of the domain items are missing, that domain will be treated as missing and thus NE. The total score for each variable (FACT-P TOI and FACT-P total) is then calculated as the sum of the un-weighted prorated scores. If a domain score is NE, any health related quality of life (HRQL) variable which these domains contribute to is also termed NE. For example, for the FACT-P TOI variable, if PWB is NE at a visit, the FACT-P TOI variable is also NE at this visit. Also, the FACT-P total score cannot be computed if any of the domain scores is NE.

Visit responses

The last non-missing assessment before randomisation will be assigned to be the baseline assessment. At each post-baseline visit, the following criteria as listed below in Table 13 will be used to assign a visit response for the FACT-P total score, FACT-P TOI, FAPSI-8, FAPSI-6, PCS and FWB scores (Cella et al 2009). This response should be maintained for 2 consecutive visits.

Table 13 Definition of visit response for FACT-P, FAPSI-8, FAPSI-6, TOI, PCS, FWB, PWB and SWB

FACT-P scale	Possible scores	Change from baseline	Visit response
FACT-P-Total	0-156	≥ +6	Improved
		≤ -6	Worsened
		Otherwise (i.e., >-6 and <+6)	No change
		Missing/non-calculable score	Not evaluable
FAPSI-8	0-32	≥ +3	Improved
		≤ -3	Worsened
		Otherwise (i.e., >-3 and <+3)	No change
		Missing/non-calculable score	Not evaluable
FAPSI-6	0-24	≥ +3	Improved
		≤ -3	Worsened
		Otherwise (i.e., >-3 and <+3)	No change
		Missing/non-calculable score	Not evaluable
TOI	0-104	≥ +5	Improved
		≤ -5	Worsened
		Otherwise (i.e., >-5 and <+5)	No change
		Missing/non-calculable score	Not evaluable
PCS	0-48	≥ +3	Improved
		≤ -3	Worsened

FACT-P scale	Possible scores	Change from baseline	Visit response
		Otherwise (i.e., >-3 and <+3)	No change
		Missing/non-calculable score	Not evaluable
FWB	0-28	≥ +2	Improved
		≤ -2	Worsened
		Otherwise (i.e., >-2 and <+2)	No change
		Missing/non-calculable score	Not evaluable
PWB	0-28	≥ +2	Improved
		≤ -2	Worsened
		Otherwise (i.e., >-2 and <+2)	No change
		Missing/non-calculable score	Not evaluable
SWB	0-28	≥ +2	Improved
		≤ -2	Worsened
		Otherwise (i.e., >-2 and <+2)	No change
		Missing/non-calculable score	Not evaluable

Note for some patients it will not be immediately possible to obtain a visit response for a particular subscale, for example:

- Patients with no baseline score for a particular subscale, or no baseline data at all
- Patients whose baseline subscale score is too close to the maximum or minimum possible score to allow an increase or decrease of the specific size to be observed.
  - For patients whose baseline score is greater than the maximum possible score for that subscale minus the score needed to satisfy improvement, the best visit response possible will be “No Change”.
  - For patients whose baseline score is less than the threshold needed for worsening (e.g., a baseline FACT-P TOI < 5) all post-baseline visit responses will be considered not-calculable.

For those patients who meet the criteria above (where it is not possible to improve or worsen), descriptive data will be provided.

At the conclusion of the study, the criteria listed in Table 14 will be used to assign a best overall response score based on the individual visit responses.

Table 14 Overall score response for FACT-P, FAPSI-8, FAPSI-6, TOI, PCS, FWB, PWB, SWB

Overall score response	Criteria
Improved	Two consecutive visit responses of 'improved'
No change	Does not qualify for overall score response of 'improved'. Two consecutive visit responses of either 'no change', or 'improved' and 'no change'
Worsened	Does not qualify for overall score response of 'improved' or 'no change'. A visit response of 'worsened'
Other	Does not qualify for one of the above
Not evaluable	Missing or non-calculable scores

Time to deterioration for FACT-P

Time to deterioration of HRQL as measured by FACT-P TOI will be defined as the interval from the date of randomisation to the first assessment of worsened without an improvement in the next 12 weeks in FACT-P TOI, or the date of death (by any cause in the absence of symptom deterioration). Time to deterioration as measured by the FACT-P total score and FAPSI-6, FAPSI8, PCS, FWB, PWB and SWB will be derived similarly.

A worsening is as described in Table 13 for example, for FACT-P TOI a decrease in score from baseline of ≥ 5, or "Subject too affected by symptoms of disease under investigation" answered as the reason for not completing HRQL at a post-baseline visit will constitute a deterioration. Improvement is also as defined within Table 13.

A deterioration followed by an improvement within 12 weeks means that this first deterioration cannot be used in determining the time to deterioration of symptoms. However, if a subsequent worsening is seen, with no improvement in the following 12 weeks, this second deterioration will be used to derive the time to deterioration.

Radiologic progression or a reason for not completing the questionnaire of "Subject unwilling", "Subject too sick, other than disease under investigation" or "Administrative failure to distribute questionnaire to Subject" will not be considered as deterioration in symptoms.

Note, under the same principles applied to the primary outcome variable (rPFS), time to deterioration will be derived regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to symptom deterioration. In addition, if a patient either meets the criteria for deterioration after 2 or more missed HRQL, then the patient will be censored at the time of the latest evaluable HRQL assessment. These patients will be presented as "Censored FACT-P TOI Score" in summaries.

A number of situations will lead to a patient's time to deterioration of HRQL endpoints being censored. These are:

- Patients who have not met the criteria for symptom deterioration or died at the time of analysis will be censored at the time of the latest evaluable HRQL assessment:
  - The censoring date will be the date of the last assessment that led to evaluable being assigned for FACT-P TOI. These patients will be presented as alive and deterioration-free in summaries.
  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1 unless they die within 2 visits of baseline. These patients will be presented as alive and deterioration-free in summaries.
- Patients whose baseline subscale score is close to the minimum possible
  - For patients whose baseline score is less than the threshold needed for worsening (e.g., a baseline FACT-P TOI of < 6), time to deterioration will be censored at Day 1 unless they die within 2 visits of baseline. Patients who haven't died will be presented as "Censored FACT-P TOI Score" in summaries.

The time to deterioration of HRQL will be derived based on assessment dates, not visit dates, unless the subject is too affected by symptoms of disease under investigation, in which case there are no assessment dates and visit dates will be used.

PRO Compliance

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

- Received forms = number of FACT-P 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 HRQL 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 HRQL 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-P forms with at least one subscale that can be determined; or where REVPRDI form is ticked 'Subject too affected by symptoms of disease under investigation'.

Thus the overall compliance rate is defined as 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-P form.

Compliance over time will be calculated separately for each visit, including baseline, as the number of patients with an evaluable 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.

3.8.1.2 Brief Pain Inventory – short form (BPI-SF)

The BPI-SF will be used to assess the impact of pain on daily life. Worst pain, general pain and pain's interference with daily life will be assessed during Part B of the study at baseline, Week 4, 8 and 12 and then continue to be administered to all patients (who have not withdrawn consent) every 12 weeks.

The BPI-SF comprises a total of 15 items measuring 2 domains: pain severity and pain interference. Items measuring pain severity (including 'worst pain') are rated on an 11-point numerical rating scale (NRS) ranging from 0=No pain to 10=Pain as bad as you can imagine. All BPI-SF items are measured using a 24-hour recall period.

The following outcome measure will be calculated for the BPI-SF:

- Pain severity (Questions 3, 4, 5 and 6)
- Pain Interference (Question 9 A-G)

Note that the database contains an item number in addition to the Question number. For each outcome measure the mean score of non-missing items will be calculated. For pain interference at least 50% of the items must have a response for a mean score to be calculated. The individual responses to the questions comprising the pain severity domain will also be summarised.

Absolute change in the BPI-SF pain and pain interference scores will be calculated as change from baseline at each clinic visit.

Time to deterioration in BPI-SF

Time to deterioration (or worsening of pain) is defined as the time from the date of randomisation to the date of first assessment of worsening of pain, or death from any cause (as long as the death occurs within 2 BPI-SF assessment visits of the last evaluable assessment of BPI-SF and regardless of whether the patient withdraws from study treatment or receives another anti-cancer therapy). "Subject too heavily affected by symptoms of disease under investigation" answered as the reason for not completing HRQL at a post-baseline visit will also signify a clinically important deterioration in BPI-SF. Worsening of pain is defined as an increase in worst pain BPI-SF score by at least 2 points on the 0 to 10 scale.

Note, under the same principles applied to time to deterioration for FACT-P, time to deterioration in BPI-SF will be derived regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to symptom deterioration. In addition, if a patient either meets the criteria for deterioration, or dies after 2 or more missed HRQL assessments, then the patient will be censored at the time of the latest evaluable HRQL assessment. These patients will be presented as "Censored BPI-SF Worst Pain Score" or "Censored Death" in summaries.

A number of situations will lead to a patient's time to deterioration of HRQL endpoints being censored. These are:

- Patients who have not met the criteria for symptom deterioration or have not died at the time of analysis will be censored at the time of the latest evaluable HRQL assessment:
  - The censoring date will be the date of the last assessment that led to evaluable being assigned for BPI-SF. These patients will be presented as alive and deterioration-free in summaries.
  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1 unless they die within 2 visits of baseline. These patients will be presented as alive and deterioration-free in summaries.

The time to deterioration of HRQL will be derived based on assessment dates, not visit dates, unless the subject is too affected by symptoms of disease under investigation, in which case there are no assessment dates and visit dates will be used.

PRO Compliance

Summary measures of overall compliance and compliance over time will be derived for the BPI-SF. These will be based upon the compliance derivations described for FACT-P.

3.8.1.3 Bone pain item

The bone pain item is an individual question asking patients to rate their worst bone pain severity on a 0 to 10 point scale. This will be summarised as captured on the eCRF. Absolute change in the worst bone pain item will be calculated as change from baseline at each clinic visit.

Time to deterioration in bone pain item

Time to deterioration (or worsening of bone pain) is defined as the time from the date of randomisation to the date of first assessment of worsening of bone pain, or death from any cause (as long as the death occurs within 2 PRO assessment visits of the last evaluable assessment of worst bone pain and regardless of whether the patient withdraws from study treatment or receives another anti-cancer therapy). "Subject too heavily affected by symptoms of disease under investigation" answered as the reason for not completing HRQL at a post-

baseline visit will also signify a clinically important deterioration in bone pain. Worsening of bone pain is defined as an increase in worst bone pain score by at least 2 points.

Note, under the same principles applied to time to deterioration for FACT-P, time to deterioration in bone pain will be derived regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy prior to symptom deterioration. In addition, if a patient either meets the criteria for deterioration, or dies after 2 or more missed HRQL, then the patient will be censored at the time of the latest evaluable HRQL assessment. These patients will be presented as "Censored Worst Bone Pain Score" or "Censored Death" in summaries.

A number of situations will lead to a patient's time to deterioration of HRQL endpoints being censored. These are:

- Patients who have not met the criteria for symptom deterioration or died at the time of analysis will be censored at the time of the latest evaluable HRQL assessment:
  - The censoring date will be the date of the last assessment that led to evaluable being assigned for bone pain. These patients will be presented as alive and deterioration-free in summaries.
  - Patients with no evaluable baseline or post-baseline data will be censored at Day 1 unless they die within 2 visits of baseline. These patients will be presented as alive and deterioration-free in summaries.

The time to deterioration of HRQL will be derived based on assessment dates, not visit dates, unless the subject is too affected by symptoms of disease under investigation, in which case there are no assessment dates and visit dates will be used.

PRO Compliance

Summary measures of overall compliance and compliance over time will be derived for the bone pain item. These will be based upon the compliance derivations described for FACT-P.

3.8.2 Health economics

3.8.2.1 Utilities: EQ-5D-5L

The euroqol-5 dimensions, five level (EQ-5D-5L) index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems and unable to/ extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the UK valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-

3L crosswalk will be applied. In addition to the descriptive system, respondents also assess their health today on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately. The evaluable population will comprise the full analysis set.

The proportion of patients with each score for each item may be summarised for each time point and by treatment allocation.

Further Payer required analyses such as mapping quality of life instruments unto utilities will be described in detail in the Payer analysis plan.

3.8.2.2 Resource use

Resource use will be analysed from procedures captured on the 'Oncology Hospital Admission' form and the standard concomitant medication eCRF, and, where required, supplemented with data on hospitalisation length of stay logged as part of SAE recording. Frequency of palliative interventions and the reasons for the intervention will be estimated from the 'Oncology Hospital Admission' form. The evaluable population will comprise the full analysis set. The analysis will be conducted on blinded data.

Further Payer required analyses involving resource use will be described in detail in the Payer analysis plan.

4. ANALYSIS METHODS

4.1 General principles

The DCO date for the statistical analysis for the primary objective of the study will be established when ~100 confirmed progression events are expected to have occurred.

The study will be centrally unblinded after database lock, but patients and Investigators will remain blinded. As per the protocol (Section 6.3.5), after the primary rPFS analysis, patients who are still receiving olaparib can either choose to discontinue from the study or, where the Investigator believes patients are gaining clinical benefit, patients may continue to receive olaparib. All patients will receive follow-up care in accordance with standard local clinical practice. Patients that are on placebo will not be offered olaparib as a study treatment.

Serious adverse events will continue to be reported to AstraZeneca Patient Safety Department, for any patients who continue on olaparib until 30 days after study treatment is discontinued, in accordance with protocol Section 6.4.4. Additionally, as stated, any SAE or non-serious AE that is ongoing at the end of the study must be followed up to resolution unless the event is considered by the Investigator to be unlikely to resolve, or the patient is lost to follow-up. If an Investigator learns of any SAEs, including death, at any time after a patient has completed the study, and he/she considers there is a reasonable possibility that the event is causally related to the investigational product, the Investigator should notify AstraZeneca, Patient Safety.

Study day will be calculated as follows:

Days prior to first dose: Study day = date – first dose date.

Days on or after first dose: Study day = date – first dose date + 1.

Where month is derived, days will be divided by 30.4375.

In general, missing data will not be imputed. For the date variables of historical data (i.e., any data referring to the period prior to the informed consent date), if the year is missing then the value will not be imputed. If the month or day is missing, the value will be imputed: month will be imputed with June; day will be imputed as 15th.

Efficacy data will be summarised and analysed on the full analysis set. Safety data will be summarised and analysed on the respective safety population.

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

All descriptive statistics will be presented by treatment group. Continuous variables will be summarised by the number of observations, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarised by frequency counts and percentages for each category. Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment group.

Data for exploratory PRO endpoints will be analysed descriptively on an ITT basis. For PRO endpoints, data will be presented in terms of minimum, maximum, mean, standard deviation and median scores together with 95% confidence interval (CIs) at each visit as well as change from baseline to end of treatment.

4.2 Analysis methods

4.2.1 Multiplicity

In order to describe the nature of the benefits of olaparib treatment, rPFS, PFS2, TFST, TSST, radiological ORR, soft tissue ORR, time to deterioration in FACT-P TOI and OS will be tested at a 1-sided significance level of 2.5%.

However if strong results are observed for the primary analysis (rPFS 1-sided p-value < 0.025) then a multiple testing strategy will be implemented in order to describe the nature of the benefits of olaparib treatment.

In order to strongly control the type I error at 2.5% 1-sided, the multiple testing procedure will be employed across the primary endpoint and secondary endpoints intended for key label claims (i.e. PFS2 and OS). There is no requirement to adjust for multiplicity due to rPFS

interim analyses, since there are no planned interim rPFS analyses with the opportunity to make an early claim of efficacy.

A hierarchical testing strategy will be employed where rPFS 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 multiple testing procedure (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 1).

Figure 1 Multiple Testing Procedure

PFS2 will only be tested if statistical significance is shown for rPFS. OS will only be tested if the null hypothesis (of no difference) is rejected for PFS2.

All planned analyses will be performed, regardless of the outcome of the MTP. If the full test mass has been used on an earlier test in the MTP, any claims regarding the statistical significance of a subsequent analysis in the MTP will be affected. Both PFS2 and OS will be tested at the time of the primary analyses of rPFS.

4.2.2 Analysis of the primary efficacy variable (rPFS)

The primary endpoint of Part B is rPFS (months). Once approximately 100 progression events have occurred, the primary analysis will be performed. Radiological progression-free survival will be analysed using the log rank test. The HR 80% CI and 95% CI will be estimated from the U and V statistics obtained directly from PROC LIFETEST in SAS (as below) using the Breslow approach for handling ties. A HR less than 1 will favour olaparib.

The HR and its confidence intervals will be estimated from the log-rank test as follows (Berry et al 1991, Sellke et al 1983):

$HR = \exp(U/V)$

80% CI for HR = $(\exp\{U/V - 1.28/\sqrt{V}\}, \exp\{U/V + 1.28/\sqrt{V}\})$

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

Where $U=\Sigma_i(d_{1i}-e_{1i})$ is the log-rank test statistic (with d_{1i} and e_{1i} the observed and expected events in group 1) and $\sqrt{V}$ the standard deviation of the log-rank test statistic as produced in the LIFETEST output. The syntax for running LIFETEST is as follows.

```
PROC LIFETEST DATA=XXX METHOD=KM;
  ODS OUTPUT LOGUNICHISQ=XX;
  TIME TIMEVAR*CENSORED(1);
  TEST TREAT;
RUN;
```

In addition, a one-sided p-value will be calculated to test the hypothesis of HR<1 (olaparib improves survival) versus the hull hypothesis of HR=1(no treatment effect). The one-sided p-value will be found using a Z-score test statistic calculated as Z =log(HR)/SE(log(HR)) which equates to $Z=(U/V)/(1/\sqrt{V})=U/\sqrt{V}$.

Kaplan-Meier survival curves (product-limit estimates) of rPFS will be presented by treatment group, together with a summary of associated statistics (median rPFS time, and 6-, 12-, 18-, 24-month survival rate estimates). Summaries of the number and percentage of patients experiencing a rPFS event, and the type of event (RECIST progression, PCWG-2 progression, both or death) will also be presented.

Listings of rPFS per patient will be produced.

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 examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, a time-dependent covariate would be fitted into the Cox proportional hazards model with treatment factor to assess the extent to which this represents random variation.

Separate log rank tests comparing rPFS between treatments in following subsets of the full analysis set based on key baseline demographics and disease characteristics (e.g. Age, Race, ECOG) as well as ATM, BRCA, composite BRCA/ATM, and composite HRR mutation categories (Sections 3.7.3, 3.7.4, 3.7.5, and 3.7.6). The listed gene mutation statuses will be aggregated to the *Detected* and the *Other* subgroups, where the latter pools together: *Not Detected, VUS, Failed,* and *Status Unknown* subtypes. The HRR mutation status will be aggregated to the *Detected*, *Not Detected*, and *Unknown* subgroups, where the last category pools together *VUS, Failed, and Status Unknown* subtypes. The results will be presented on a forest plot. If there are too few events available for a meaningful subgroup analysis (it is not considered appropriate to present analyses with less than 5 events within each treatment arm within the subgroup), only descriptive statistics will be provided along with the number (%) of patients and events.

The primary rPFS analysis will also be repeated for patients who had not responded to treatment with abiraterone by Week 12 (Visit 5), a non-responder is defined as in Section

3.7.1.1. Patients who had responded to treatment at this time will be excluded from the analysis.

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

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.

For the subset of patients who have measureable disease at baseline, summary statistics for the change from baseline in the sum of the target lesion LDs will be presented.

Additionally, summaries will be provided showing the reason for progression, i.e. new lesion, progression of target lesions, progression of non-target lesions or a combination of these. The reasons will be summarised separately for patients with RECIST only progressions and those with bone and RECIST progressions.

Summary statistics for the number of months between rPFS time and the last evaluable assessment prior to progression will be presented for each treatment group for subjects who have progressed.

4.2.2.1 Sensitivity analyses

To assess the sensitivity of the primary rPFS analysis, the following supportive analyses will be performed:

- To assess time evaluation bias (i.e., differential assessment times between treatment groups), the primary rPFS analysis, will be repeated, but using the midpoint between the time that progression was detected (as calculated for the primary endpoint) and the previous evaluable assessment (RECIST or PCWG-2).

- To assess attrition bias, the primary rPFS analysis will be repeated, but the actual rPFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following 2, or more, NE tumour assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy.

- To assess censoring bias, a Kaplan-Meier plot of the time to censoring will be produced, where the censoring indicator of the primary rPFS analysis is reversed.

- 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

  Using the same methodology as described for the primary analysis of rPFS. The HR and its corresponding 80% and 95% CI will be presented.

- The primary rPFS analysis will be repeated using only PCWG-2 progression events to address any possible bias introduced as a result of including any soft tissue progressions as assessed by RECIST 1.1 in the primary definition of rPFS.

4.2.3 Analysis of the secondary outcome variables

4.2.3.1 Time to second progression (PFS2)

The secondary endpoint of PFS2 (months) will be analysed using the same time-to analysis as rPFS: log rank test, Kaplan-Meier and summaries of associated statistics (median second progression time, and 6-, 12-, 18-, 24-month survival rate estimates).

The type of progression event will also be summarised by treatment group.

4.2.3.2 Overall survival (OS)

The secondary efficacy endpoint for Part B, OS (months), will be performed at the time of the analysis of rPFS when approximately 60% of patients have died; patients who have not died at that point will be censored. Overall survival will be analysed using the same approach as for rPFS: log rank test, Kaplan-Meier and summaries of associated statistics (median second progression time, and 6-, 12-, 18-, 24-month survival rate estimates). A listing of OS time per patient will also be produced.

4.2.3.3 Objective response rate (ORR) and associated variables

The secondary efficacy endpoint for Part B, ORR for soft tissue disease response and overall radiological response, will be performed at the time of the rPFS analysis. A summary of ORR will be presented by treatment group. Objective response rate will be formally analysed for both soft tissue ORR and overall radiological ORR using logistic regression, with treatment group as a factor. The effect of treatment will be estimated using the adjusted odds ratio and its corresponding 80% CIs and 95% CIs. Only patients with measurable disease (target lesions) at baseline will be included in the analysis

The other associated endpoints (best overall soft tissue disease response, best overall radiological response, DOR, time to response) will be summarised only.

Best objective response (soft tissue disease and overall radiological response) at each scheduled visit will be summarised.

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

TFST and TSST will be analysed at the same time as the primary analysis of rPFS using the same methodology and model. Patients who did not receive randomised treatment will be censored on the date of randomisation. The HRs for the treatment effect together with 80% and 95% CIs will be presented. Kaplan-Meier plots will be presented by treatment arm.

Summary tables of first and second subsequent therapies event type by treatment arm will be provided.

A listing of TFST and TSST, per patient, will also be produced.

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

4.2.4 Tolerability and safety

Summaries will be presented for scheduled visits only. Any unscheduled assessments will be listed. The baseline value is defined as the latest result obtained prior to the start of IP.

4.2.4.1 Adverse events (AE)

Adverse events will be summarised separately for Parts A and B of the study, and also by dose for AEs occurring within Part A.

All AEs, both in terms of MedDRA preferred term and CTCAE grade, will be listed and summarised descriptively by count (n) and percentage (%) and dose (Part A) or treatment group (Part B). 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’.

An overall summary of the number and percentage of patients in each category will be presented for Parts A and B and an overall summary of the number of episodes in each category will be presented for Part B only.

Frequencies and percentages of patients reporting each preferred term will be presented for Part B only. Total number of events will also be reported separately.

All reported AEs will be included in listings along with the date of onset, date of resolution (if AE is resolved), investigator’s assessment of severity and relationship to study drug.

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

- All AEs
- All AEs causally related to olaparib/placebo

- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to olaparib/placebo
- AEs with outcome of death
- AEs with outcome of death causally related to olaparib/placebo
- All SAEs
- All SAEs causally related to olaparib/placebo
- AEs leading to discontinuation of olaparib/placebo
- AEs leading to discontinuation of olaparib, causally related to olaparib/placebo
- Other significant AEs
- Other significant AEs causally related to olaparib/placebo

Key patient information tables will be produced for:

- AEs causally related to olaparib
- AEs with outcome of death
- All SAEs
- AEs leading to discontinuation of olaparib/placebo
- Other significant AEs

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 for Part B only. This cut-off may be modified after review of the data.

For Part B, 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 of therapy (for patients without the event) and the time to the AE (for patients with the event) in each group multiplied by 1000.

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

Adverse events with CTCAE grade 3 or higher will be summarised as the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

Causally related adverse events will be summarised as the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

Causally related adverse events of CTCAE grade 3 or higher will be summarised for the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

Summaries of the number and percentage of patients with AEs leading to dose change of olaparib/placebo and also dose interruptions of olaparib/placebo will be presented by preferred term and treatment group for Part B only.

Summaries of the number and percentage of patients with AEs which started prior to first dose or > 30 days following date of last dose will be presented by preferred terms and treatment group for Part B only.

A summary of deaths will be provided with number and percentage of patients by study part and actual treatment group (Part B), categorised as:

- Related to disease under investigation
- AE outcome=death
- Both related to disease under investigation and with AE outcome=death
- Unrelated to AE or disease under investigation
- Deaths ≥ 30 days after last treatment dose, related to disease under investigation
- AE with outcome=death ≥30 days after last treatment dose
- Deaths ≥ 30 days after last treatment dose, related to AE or disease under investigation
- Deaths ≥ 30 days after last treatment dose, unrelated to AE or disease under investigation
- Patients with unknown reason for death, and
- Other deaths (not captured above).

Causally related adverse events with an outcome of death will be summarised for the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

Causally related serious adverse events will be summarised for the number and percentage of patients by SOC, preferred term and actual treatment group for Parts A and B.

Adverse events leading to discontinuation of olaparib/placebo will be summarised for the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

Causally related adverse events leading to discontinuation of olaparib/placebo will be summarised for the number and percentage of patients by SOC, preferred term and actual treatment group for Part B only.

In addition, AEs with outcome of death, SAEs, OAEs, AEs leading to discontinuation of treatment and AEs causally related to olaparib/placebo will be listed in key patient information tables for Parts A and B.

Listings of AE data will also be produced for Parts A and B.

Summary of long term tolerability

For Part B only, to assess long term tolerability, prevalence plots, life table plots and cumulative incidence plots will be presented for:

- Nausea
- Vomiting
- Any other events considered important after review of the safety data, provided there are ≥10 events.

A prevalence plot provides information on the extent to which the events may be an ongoing burden to patients. The prevalence at time t after first dose of study treatment is calculated as the number of patients experiencing the event divided by the number of patients receiving study treatment or in safety follow-up at time t; generally, t is categorised by each day after dosing. The prevalence is plotted over time split by treatment group. Multiple occurrences of the same event are considered for each patient but a patient is only counted in the numerator whilst they are experiencing one of the occurrences of the event.

For each AE, median time to first onset of the AE will be presented for patients in the safety analysis set by actual treatment group. Patients who did not experience the AE will be censored at the end of their safety follow-up. 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 for patients who experienced each AE.

A life table plot can be used to describe the time to onset of the event and specifically when patients are at most risk of first experiencing the event. The hazard, or in other words, the probability of having an AE in a specified time period (e.g. 0-1 months, 1-3 months, 3-6

months, etc.) given that the patient reaches that time period without having an event is plotted for each time period split by treatment.

A cumulative incidence plot is a plot of the raw cumulative incidence and cumulative incidence function over time with the treatment groups presented on separate plots. The raw cumulative incidence is the actual probability that a patient will have experienced their first occurrence of the event by a given time point. The cumulative incidence function estimates the cumulative incidence if the DCO had not been imposed and all patients had completed safety follow-up (O'Neill 1987).

4.2.4.2 Dose-limiting toxicities

The incidence of DLTs will be summarised by treatment cohort for Part A.

4.2.4.3 Exposure and compliance

Listings and summaries of exposure, actual exposure, and compliance will be produced for olaparib and abiraterone, by treatment cohorts, for Part A and treatment groups in Part B. RDI, PID will also be summarised, but not listed.

Compliance for olaparib will be derived as per section 3.5.7 details. Compliance for abiraterone will be summarised using the eCRF measurement 'Study drug administered according to protocol'.

The number of patients with study drug interruptions or reductions for olaparib/placebo, and the reasons, will be summarised by treatment group for Part B only. These data will also be listed.

Treatment duration will be summarised. Treatment duration is based on the dates of first and last dose.

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

4.2.4.4 Laboratory data

Laboratory data (clinical chemistry and haematology) will be summarised. Laboratory data will be summarised separately for Parts A and B by dose group and by study day.

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 CTCAE grade will be produced, within each part of the study and overall, 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 the last dose of olaparib/placebo). A scatter plot of ALT versus total bilirubin, both expressed as multiples of upper limit of normal range, 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 for Part B only.

Box-plots of change from baseline will be presented.

All laboratory summaries will be presented by actual treatment group.

4.2.4.5 Concomitant medications

Concomitant medications will be summarised by the coded terms. Astra Zeneca drug dictionary (AZDD) will be used for concomitant medication coding. The number of patients receiving a medication will be summarised overall and for each part of the study. A medication taken during the course of the study is considered concomitant. A patient is only counted once if receiving the medication more than once. Allowed and disallowed medications will be summarised separately with allowed concomitant medications summarised for Part B only.

4.2.4.6 Vital signs

Vital signs, including BP (mmHg), heart rate (beats/min) and body temperature (°C) actual values and change from baseline values, will be summarised separately for Parts A and B by dose group and by study day. Vital signs will be summarised at baseline and each scheduled visit. The baseline value is the last pre-dose assessment. Descriptive statistics will include number of non-missing patients (n), mean, standard deviation, median, first and third quartiles (Q1 and Q3), minimum and maximum.

Box-plots of absolute values by study day by treatment group, and change from baseline by study day by treatment group will also be produced for pulse rate, systolic BP, diastolic BP and temperature for Part B only.

4.2.5 Biomarkers

All Biomarkers analysis will be performed for Part B only.

- Proportion of patients achieving a PSA response (single visit response and also a confirmed response) with 80% CIs and best percentage change from baseline values will be presented.
- PSA absolute change and percentage change from baseline values will be presented.
- CTC change from baseline values will be presented.

- Proportion of patients achieving a CTC conversion with 80% CIs and best change from baseline will be presented.

4.2.6 Patient reported outcomes (PRO) and HRQL

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

4.2.6.1 FACT-P

Summary statistics for FACT-P score will be presented by treatment group (including means, standard deviation, median and range) for all visits until there are less than one third of patients with evaluable data.

Time to FACT-P deterioration will be analysed using the same methodology and model as described for the primary analysis of rPFS. However sensitivity analyses will not be performed.

FACT-P improvement rates will be analysed using a logistic regression model with treatment group as a factor. If the overall response rate is < 5%, no analysis will be performed. (Note that if the response rate in only one of the treatment groups is < 5% but ≥ 5% in the other treatment group then the analysis will still be performed). If the expected response rate is low (< 20%) a Fisher's exact test will be considered and mid p-values used.

The results of the analysis will be presented in terms of an odds ratio together with its associated profile likelihood CIs and p-values (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model).

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

Descriptive statistics will be reported for the FACT-P by visits as well as change in these scores from baseline. These will also be reported for PWB, FWB, PCS and the individual items for PCS. Summary tables of FACT-P best change rates (improvement, worsening, and no change) will be provided.

Supportive analyses

Supportive analyses will be performed for the FACT-P total score, and scales (TOI, FAPSI-6) for both improvement rates, time to deterioration and best change rates (best overall response). Compliance will be analysed for the FACT-P total only.

Change from baseline in the FACT-P total score, and scales (TOI, FAPSI-6) will be analysed using a mixed model for repeated measures (MMRM) analysis of all the post-baseline FACT-P (or equivalent) scores for each visit. The study discontinuation visit and the safety follow-up visit will be excluded from this analysis. Restricted maximum likelihood (REML) estimation will be used. The model will include treatment, visit and treatment by visit interaction as explanatory variables and the baseline FACT-P total score as a covariate, along with the baseline FACT-P total score by visit interaction. Treatment, visit, treatment by visit

interaction, baseline FACT-P total score, and the baseline FACT-P score by visit interaction will be fixed effects in the model. 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= FACTP method = reml;
class TRT VISIT SUBJECT;
model CHBL = TRT VISIT TRT*VISIT BL BL*VISIT / s ddfm=kr;
repeated VISIT / type=UN subject=SUBJECT;
lsmeans TRT / at means pdiff diff alpha=0.05 cl;
```

where TRT is the randomised treatment, VISIT is the visit, CHBL is the change from baseline in the FACT-P total score, and BL is the baseline FACT-P total 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.

The adjusted mean estimates and corresponding 95% confidence intervals will be presented by visit for each treatment group.

The minimally important difference in score is ≥20%, ≥50% is an important difference and ≥80% a large difference.

4.2.6.2 BPI-SF

Summary statistics for BPI-SF pain severity scale and pain interference scale, means, standard deviations, medians and ranges will be presented by treatment group for visits until there are less than 20% of patients with evaluable data.

Analysis of time to deterioration, and change from baseline in the aforementioned BPI-SF scores will be carried out, as described for FACT-P in Section 4.2.7.1. Supportive analyses will also be repeated, including MMRM for the worst pain severity score (item nbr 3) and the total pain severity score (item numbers 3-6).

BPI-SF compliance (overall compliance and by visit compliance) will be summarised for each treatment group.

4.2.6.3 Bone Pain

Summary statistics for mean bone pain score, standard deviation, median and range will be presented by treatment group for visits until there are less than 20% of patients with evaluable data.

Analysis of time to deterioration, improvement rates and change from baseline in bone pain score will be carried out as described for FACT-P in Section 4.2.7.1. Supportive analyses will also be repeated, including MMRM.

Bone pain compliance (overall compliance and by visit compliance) will be summarised for each treatment group.

4.2.7 Health economics data

4.2.7.1 Utilities: EQ-5D-5L

Descriptive statistics will be reported for health state utility index values and visual analogue scale by visits as well as change in these scores from baseline. To support future economic evaluations of olaparib, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post-progression. Further details will be outlined in the Payer analysis plan.

4.2.7.2 Resource use

An exploratory health economic analysis of the frequency of metastatic prostate cancer related palliative interventions, time to interventions, and reason for the intervention will be undertaken. In addition, length of stay, intensive care unit (ICU) use, concomitant medications and analgesic use will be examined.

These analyses will examine the impact of disease and treatment on resource use to primarily support the economic evaluation of olaparib in castrate resistant metastatic prostate cancer.

4.2.8 Demographic data and baseline data

The following baseline characteristics will be listed for each patient and summarised for all patients in the safety analysis set (Part A) and the FAS (Part B) (unless otherwise specified):

The following collected data will be listed.

- Patient disposition (including confirmed eligible (Part A)/randomised (Part B), received treatment and completed the study)
- Important deviations
- Inclusion in analysis populations
- Demographics (age, age group, race and ethnicity)

- Patient characteristics at baseline (height, weight, and body mass index (BMI)
- Patient recruitment by country and centre (presented for Part B only)
- Previous disease related treatment modalities
- Previous radiotherapy (tabulated for Part B only and not listed)
- Disease characteristics at baseline (ECOG performance status at baseline, PSA at baseline, ATM mutation status at screening BRCA status, time since original diagnosis, histology type at diagnosis, tumour grade at diagnosis, AJCC stage at diagnosis and Gleason score at diagnosis)
- Extent of disease
- Time from prior disease progression to randomisation (Part B only)
- Post-discontinuation disease-related anticancer therapy (Part B only)
- Current signs and symptoms at baseline
- Current and past medical history
- Relevant surgical history (presented for Part B only and not listed)
- Physical examination at baseline (presented for Part B only)

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

4.2.9 Statistical analysis of PK data

The PK analysis will be performed for Part A only.

PK descriptive statistics and precision of data for reporting

All plasma concentration data received for olaparib and abiraterone and all derived reportable PK parameters will be presented in data listings for the safety analysis set. All plasma concentrations, $C_{ss,max}$ and $C_{ss,min}$ will be reported to the same precision as the source data (3 significant figures). All other PK parameters will be listed to 3 significant figures except for $t_{ss,max}$ and $t_{ss,last}$ which will be presented to 2 decimal places.

Reportable pharmacokinetic concentration and parameter summary data will be presented for patients in the PK analysis set. Extra measurements (such as unscheduled or repeated assessments) will be included in the patient listings but will only be included in summary tables and figures if the extra measurement is used in the PK analysis e.g. if a repeat value is substituted for the original at the discretion of the PK scientist.

The exclusion of any patients from the PK analysis set or from the analysis in a specific treatment period will be agreed by the study physician, statistician and PK Scientist during the data review meeting (DRM) held once all data is clean and final. Any data to be excluded from the PK analysis and/or summary tables and corresponding figures will be flagged with the reasons by the PK Scientist in the PK handover document and flagging file for programming and these data will be flagged in the listings and the appropriate footnotes added.

Quantitative PK variables for olaparib and abiraterone will be summarised using descriptive statistics by Cohort, Group and treatment period (Visit/Day)) (Cohort 2 Group 1 Visits 3 and 4 and Cohort 2 Group 2 for olaparib and Cohort 2 Group 1 Visit 4 and Cohort 2 Group 2 Visits 3 and 4 for abiraterone).

The PK parameters AUC, AUC_{ss}, AUC_{0-t}, $C_{ss,max,}$ and $C_{ss,min}$ will present:

- Number of observations (n)
- Geometric mean (Gmean, calculated as exp[μ], where μ is the mean of the data on a log scale)
- Geometric coefficient of variation (%GCV, calculated as 100 √ [exp(s^2)-1], where s is the standard deviation of the data on a log scale)
- Arithmetic mean (Amean, calculated using non log-transformed data)
- Standard Deviation (SD, using non log-transformed data)
- %CV (using non log-transformed data)
- Median
- Minimum (min)
- Maximum (max)

Plasma concentrations of olaparib and abiraterone at each time point will use the same summary statistics with the addition of n> LLOQ and Gmean error bars calculated as exp[μ ± s], where μ and s are the mean and standard deviation of the log-transformed data, respectively. These Gmean error bars will be labelled 'Gmean +/- GSD'.

CL_{ss}/F will present only n, Amean, SD, %CV, median, min and max.

$t_{ss,max}$ and $t_{ss,last}$ will present only n, median, min and max

The descriptive statistics for all PK data will be presented to 4 significant figures except for min and max which will be presented to 3 significant figures and n and n>LLOQ which will be presented as a whole integer.

Reporting of plasma concentrations that are Below Limit of Quantification (BLQ)

Individual olaparib and abiraterone concentrations below their Lower Limit of Quantification (LLOQ) of the bioanalytical assay will be reported as NQ (Not Quantifiable) in the listings with the LLOQ defined in the footnotes of the relevant TFLs. Individual plasma concentrations that are Not Reportable will be reported as NR and those that are missing will

be reported as NS (No Sample) in the listings. Plasma concentrations that are NQ, NR or NS will be handled as follows for the provision of descriptive statistics.

Descriptive Statistics

- Any values reported as NR or NS will be excluded from the summary tables and figures.
- At a time point where less than or equal to 50% of the concentration values are NQ, all NQ values will be substituted with the LLOQ concentration, and all descriptive statistics will be calculated accordingly.
- At a time point where more than half (but not all) of the values are NQ, the Gmean, Gmean with range, %GCV, Amean, SD, and %CV will be set to NC. The maximum value will be reported from the individual data, and the minimum and median will be set to NQ.
- If all concentrations are NQ at a time point, the Gmean, Amean, minimum, median and maximum will be reported as NQ, and the Gmean with range, %GCV, SD, and %CV will be reported as NC.
- The number of values above LLOQ (n >LLOQ) will be reported for each time point together with the total number of collected values.

Three observations > LLOQ are required as a minimum for a plasma concentration or PK parameter to be summarised. Two values are presented as a minimum and maximum with the other summary statistics as NC. For consistency, the same plasma concentration values are used in the mean data graphs as those given in the descriptive statistics summary table for each time point.

Evaluation of DDI

A secondary objective of Part A of the study is to evaluate the presence of any drug interaction between olaparib and abiraterone by looking for any marked difference in steady state exposure when each is dosed in the presence and absence of the other. $C_{ss,max}$, AUC_{ss} and $C_{ss,min}$ ratios will be calculated using non log-transformed parameter data for olaparib from Cohort 2 Group 1 patients (olaparib + abiraterone treatment : olaparib alone treatment) and for abiraterone from Cohort 2 Group 2 patients (abiraterone + olaparib treatment : abiraterone alone treatment). These PK parameter ratios will be summarised for each Group and displayed graphically. No formal statistical analysis of the PK data will be performed. The study was not sized to show statistically significant differences.

PK figures

The following PK figures will be provided:

- Plot of individual steady state plasma concentrations (µg/ml) of olaparib versus time following 300mg twice daily dosing of olaparib – Cohort 2 – 1 semilogarithmic plot per subject for Group 1 Visits 3 and 4 and 1 semilogarithmic plot per subject for Group 2 Visit 4 (Safety Analysis Set)

- Plot of individual steady state plasma concentrations (µg/ml) of abiraterone versus time following 1000mg once daily dosing of abiraterone – Cohort 2 – 1 semilogarithmic plot per subject for Group 1 Visit 4 and 1 semilogarithmic plot per subject for Group 2 Visits 3 and 4 (Safety Analysis Set)

- Combined plot of individual steady state olaparib plasma concentrations (µg/mL) versus time following 300mg twice daily dosing of olaparib – Cohort 2 - 1 linear plot per treatment for Group 1 Visit 3, Group 1 Visit 4 and Group 2 Visit 4 (Safety Analysis Set)

- Combined plot of individual steady state abiraterone plasma concentrations (ng/mL) versus time following 1000mg once daily dosing of abiraterone – Cohort 2 - 1 linear plot per treatment for Group 1 Visit 4 , Group 2 Visit 3 and Group 2 Visit 4 (Safety Analysis Set)

- Plot of Gmean (+/- GSD) steady state olaparib plasma concentrations (µg/mL) vs time following 300mg twice daily dosing of olaparib – Cohort 2 – 1 semilogarithmic plot and 1 linear plot (PK Analysis Set)

- Plot of Gmean steady state olaparib plasma concentrations (µg/mL) vs time following 300mg twice daily dosing of olaparib – Cohort 2 – 1 semilogarithmic plot (PK Analysis Set)

- Plot of Gmean (+/- GSD) steady state abiraterone plasma concentrations (ng/mL) versus time following 1000mg once daily dosing of abiraterone – Cohort 2 – 1 semilogarithmic plot and 1 linear plot (PK Analysis Set)

- Plot of Gmean steady state abiraterone plasma concentrations (ng/mL) versus time following 1000mg once daily dosing of abiraterone – Cohort 2 – 1 semilogarithmic plot (PK Analysis Set)

- Line plots of individual and Gmean olaparib steady state pharmacokinetic parameters alone and in combination with abiraterone following 300mg twice daily dosing of olaparib – Cohort 2 Group 1 (PK Analysis Set)

- Line plots of individual and Gmean abiraterone steady state pharmacokinetic parameters alone and in combination with olaparib following 1000mg once daily dosing of abiraterone – Cohort 2 Group 2 (PK Analysis Set)

- Scatter plot for the individual and Gmean ratio of steady state pharmacokinetic parameters of olaparib in combination with abiraterone : alone (Visit 4:Visit 3) following 300mg twice daily dosing of olaparib – Cohort 2 Group 1 (PK Analysis Set)

- Scatter plot for the individual and Gmean ratio of steady state pharmacokinetic parameters of abiraterone in combination with olaparib : alone (Visit 4:Visit 3)

following 1000mg once daily dosing of abiraterone – Cohort 2 Group 2 (PK Analysis Set)

5. INTERIM ANALYSES (NOT APPLICABLE)

6. CHANGES OF ANALYSIS FROM PROTOCOL

- Changed the description of the log-rank test such that it does not suggest using a stratified log-rank test. The randomisation was not stratified and the primary analysis will not be stratified. Added 95% CI for HR to be reported. The 30-month survival rates will not be provided in the associated statistics.

- Updated primary efficacy subgroup analysis plans by inclusion of the other (beside ATM, BRCA1/2) genes mutation statuses from the HRR group, and exclusion of ERG expression/fusion status. The minimum 20 endpoint events per subgroup/arm condition replaced with minimum 5 events.

- The PTEN and AR marker factors were excluded from the current exploratory analysis. Any futures exploratory analysis and outcome variables were left yet to be defined. No exploratory analysis is planned to assess an impact from possible treatment switching on the OS adjustment.

- For the PRO endpoints, the time to worsening will be analysed similar as in primary analysis, and not as planned with a Cox PH model.

- In the PK analysis, the DDI ratio was changed from mono:combination to combination:mono for olaparib and abiraterone.

- New category ('Not applicable (NA)') was added to the TLs and NTLs response outcomes. Weight (part of vital signs) levels were recorded at the baseline only.

7. REFERENCES

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

Cella 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.

Cella 1994
Cella, D.F. Manual: Functional Assessment of Cancer Therapy (FACT) Scales and Functional Assessment of HIV Infection (FAHI) Scale, Chicago: Rush-Presbyterian-St Luke's Medical Center. 1994

Cella et al 2009
Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating clinically meaningful changes for the Functional Assessment of Cancer Therapy-Prostate: results from a clinical trial of patients with metastatic hormone-refractory prostate cancer.
Value Health 2009; Jan-Feb;12(1):124-9.

de Bono et al 2011
de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, Chi KN, Jones RJ, Goodman OB Jr, Saad F, Staffurth JN, Mainwaring P, Harland S, Flaig TW, Hutson TE, Cheng T, Patterson H, Hainsworth JD, Ryan CJ, Sternberg CN, Ellard SL, Fléchon A, Saleh M, Scholz M, Efstathiou E, Zivi A, Bianchini D, Loriot Y, Chieffo N, Kheoh T, Haqq CM, Scher HI, COU-AA-301 Investigators, Abiraterone and increased survival in metastatic prostate cancer.
N Engl J Med. 2011 May 26; 364(21):1995-2005.

Esper et al 1997
Esper P, Mo F, Chodak G, Sinner M, Cella DF, Pienta KJ. Measuring quality of life in men with prostate cancer using the Functional Assessment of Cancer Therapy-Prostate instrument.
Urology 1997;50:920-8.

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.

Robins et al 1991
Robins JM, Tsiatis AA. Correcting for noncompliance in randomized trials using rank preserving structural failure time models. Communications in Statistics - Theory and Methods 1991; 20:2609-2631.

Robins 1993
Robins J. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. Proceedings of the Biopharmaceutical Section, American Statistical Association 1993; 24-33.

Scher et al 2012
Scher HI, Fizazi K, Saad F, Taplin ME, Sternberg CN, Miller K et al. Increased survival with enzalutamide in prostate cancer after chemotherapy.
N Engl J Med. 2012 Sep 27;367(13):1187-97.

Sellke et al 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.

O'Neill 1987
O'Neill R.T Statistical analyses of adverse event data from clinical trials. Special emphasis on serious events, Drug Information Journal, 21, 9 - 20, 1987.

8. APPENDIX (NOT APPLICABLE)