STATISTICAL ANALYSIS PLAN

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Phase 3 Study of Rucaparib as Switch Maintenance Following Platinum-Based Chemotherapy in Patients with Platinum-Sensitive, High-Grade Serous or Endometrioid Epithelial Ovarian, Primary Peritoneal or Fallopian Tube Ovarian Cancer

PROTOCOL: CO-338-014

VERSION: Version 1.0

DATE FINAL: 21 April 2017

SPONSOR: Clovis Oncology
5500 Flatiron Parkway, Suite 100
Boulder, CO 80301
(303) 625-5000

APPROVAL PAGE

This document was reviewed by the following Clovis team members:

- [redacted]

The document was authored by:

- [redacted]

The document is hereby signed by

Statistician

Signature [redacted]

Printed Name [redacted] Date [redacted]

Clinical Development

Signature [redacted]

Printed Name [redacted] Date [redacted]

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ANALYSIS POPULATION DEFINTION

AE(s)	adverse event(s)
ATC	Anatomical Therapeutical Chemical (coding)
BRCA	breast cancer gene
CA-125	cancer antigen 125
CDx	companion diagnostic
CI	confidence interval
Cm	Centimeter
CR	complete response
CRF	case report form
CSR	clinical study report
CTA	Clinical Trial Assay (also referred to as ICTA)
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of response
DRS-P	disease related symptoms-physical subscale
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
FACT	Functional Assessment of Cancer Therapy
FMI	Foundation Medicine, Inc.
FOSI-18	FACT-Ovarian Symptom Index-18
gBRCA	germline *BRCA1/2* mutation
GCIG	Gynecologic Cancer InterGroup
HGSOC	high grade serous ovarian cancer
HR	hazard ratio
HRD	homologous recombination deficiency
HRR	homologous recombination repair
ICH	International Conference on Harmonisation
IDMC	Independent Data Monitoring Committee
invPFS	progression-free survival, investigator assessed
IRR	independent radiology review
irrPFS	progression-free survival, independent radiology review
ITT	intent-to-treat
LOH	loss of heterozygosity
MedDRA	Medical Dictionary for Regulatory Activities
nbHRD	non-tBRCA HRD
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NGS	next generation sequencing
ORR	overall response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival

PK	pharmacokinetic(s)
PR	partial response
PRO	patient reported outcome
PT	Preferred Term
QoL	quality of life
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
sNDA	supplemental New Drug Application
SOC	System Organ Class
tBRCA	*BRCA1/2* mutation in tumor tissue
TEAE(s)	treatment-emergent adverse event(s)
ULN	upper limit of normal
VAS	Visual Analogue Scale
WHO	World Health Organization

Analysis Population Definitions

tBRCA	Patients with BRCA mutation in their tumor
HRD	Patients with HRD^+ tumors (a NGS HRD assay result), composed of tBRCA and non-tBRCA LOH+
non-tBRCA LOH+	Patients without a tBRCA mutation and with percent of tumor genome LOH $\geq$ 16%
non-tBRCA LOH-	Patients without a tBRCA mutation and with percent of tumor genome LOH < 16%
non-tBRCA LOH unknown	Patients who do not have a tBRCA mutation and where the LOH result is unknown.
Intent-to-Treat (ITT)	All randomized patients
Safety Population	The safety population will consist of all patients who received at least one dose of protocol-specified treatment

1 INTRODUCTION

This document describes the statistical analyses and data presentations to be performed for the clinical study report (CSR) of Clovis Oncology protocol CO-338-014 "A Multicenter, Randomized, Double-Blind, Placebo-Controlled Phase 3 Study of Rucaparib as Switch Maintenance Following Platinum-Based Chemotherapy in Patients with Platinum-Sensitive, High-Grade Serous or Endometrioid Epithelial Ovarian, Primary Peritoneal or Fallopian Tube Cancer."

This statistical analysis plan (SAP) provides a comprehensive and detailed description of the strategy, rationale, and statistical techniques to be used to assess the efficacy and safety of rucaparib (CO-338) compared with placebo in patients with platinum-sensitive high-grade serous ovarian or endometrioid epithelial ovarian, primary peritoneal, or fallopian tube cancer who achieved a response to platinum-based therapy.

The purpose of the SAP is to ensure the credibility of the study findings by specifying the statistical approaches to the analysis of study data prior to snapshot and unblinding of treatment codes. This SAP provides additional details around the statistical analyses that are outlined in the original protocol dated 9 Sep 2013, and three protocol amendments (Amendment 1 dated 4 Nov 2014, Amendment 2 dated 9 Mar 2015, and Amendment 3 dated 7 Jul 2016).

All statistical analyses detailed in this SAP will be conducted using SAS® Version 9.3 or higher.

2 OVERALL STUDY DESIGN, OBJECTIVES, AND ENDPOINTS

2.1 Study Objectives and Endpoints Outlined in Original Protocol

Table 1. Primary, Secondary, and Exploratory Objectives and Endpoints

Primary Objectives	**Primary Endpoints**
To evaluate progression-free survival (PFS) by Response Evaluation Criteria in Solid Tumors (RECIST) Version 1.1, as assessed by the investigator (invPFS), in molecularly-defined homologous recombination deficiency (HRD) subgroups	Disease progression according to RECIST v.1, as assessed by the investigator (invPFS), or death from any cause, in molecularly-defined HRD subgroups
Secondary Objectives	**Secondary Endpoints**
To evaluate PFS by RECIST, as assessed by independent radiology review (IRR), in molecularly-defined HRD subgroups	Disease progression according to RECIST v1.1, as assessed by independent radiology review (irrPFS), or death from any cause, in molecularly-defined HRD subgroups
To evaluate patient-reported outcome (PRO) of disease-related symptoms utilizing the disease-related symptoms – physical (DRS-P) subscale of the National Comprehensive Cancer Network-Functional Assessment of Cancer Therapy (NCCN-FACT) FACT-Ovarian Symptom Index 18 (FOSI-18)	Time to a 4-point decrease in the DRS–P subscale of the FOSI-18
To evaluate PRO utilizing the complete FOSI-18 instrument	Time to a 8-point decrease in the total score of the FOSI-18
To evaluate survival benefit	Overall survival (OS)
To evaluate safety	The incidence of adverse events (AEs), clinical laboratory abnormalities, and dose modifications
To determine the population pharmacokinetics (PK) of rucaparib	Individual model parameter estimates of rucaparib and covariates identification
Exploratory Objectives	**Exploratory Endpoints**
To evaluate the relationship between cancer antigen 125 (CA-125) levels and invPFS	Association between the change from baseline in CA-125 measurements and invPFS
To evaluate PFS2 (PFS on the subsequent line of treatment)	Time to the next event of disease progression or death, as assessed by the investigator
To evaluate overall response rate (ORR)	ORR per RECIST v1.1, as assessed by both the investigator and IRR, in patients who have measureable disease at study entry

Table 1 continued. Primary, Secondary, and Exploratory Objectives and Endpoints

To evaluate duration of response (DOR)	DOR per RECIST v1.1, as assessed by both the investigator and IRR
To evaluate PRO utilizing the Euro-QoL 5D (EQ-5D) instrument	PRO as measured by the total score on the EQ-5D
To explore the relationship between rucaparib exposure, efficacy, and safety	Rucaparib PK, invPFS, irrPFS, CA-125, AEs, clinical laboratory abnormalities, and dose modifications

2.2 Trial Design

This double-blind, randomized, multicenter, Phase 3 study will compare rucaparib with placebo in patients with platinum-sensitive, high-grade serous or endometrioid epithelial ovarian, primary peritoneal, or fallopian tube cancer who have achieved a response to platinum-based therapy (ARIEL3).

The initial clinical trial assay (CTA) will be utilized to stratify patients for randomization at study entry into 1 of 3 HRD subgroups (tBRCA, nbHRD, and biomarker negative). The CTA sequences DNA extracted from tumor tissue and identifies mutations in 30 genes involved in HRD: *BRCA1/2* (stratify into the tBRCA) and 28 other HRR genes (stratify into nbHRD) (see Appendix A for the specific list of genes). Patients with no mutations identified in the 30 HRD genes are stratified into the biomarker negative subgroup. Eligible patients will be randomized (2:1) using Interactive Response Technology to receive either rucaparib or placebo.

The primary efficacy endpoint for this study is to evaluate PFS, as assessed by the investigator, in molecularly-defined HRD subgroups. Investigational centers will interpret tumor scans locally for the purpose of making treatment decisions. IRR will also be performed to support analyses by investigator assessment. Secondary endpoints include patient-reported outcome as assessed by the FOSI-18 instrument using an electronic PRO tool and overall survival.

For the efficacy analyses, patients will be assigned to molecularly-defined HRD subgroups based on the CTA results. Results from study CO-338-017 (ARIEL2) Part 1 demonstrated that genomic loss of heterozygosity (LOH) was a good predictor of rucaparib sensitivity. Based on these ARIEL2 results, an optimized rucaparib sensitivity algorithm, which includes a pre-specified LOH cut-off of 16% percent genomic LOH along with tBRCA mutation status, has been established and documented, and will be prospectively applied to the primary analysis of ARIEL3 [redacted] That is, all analyses will be performed using the optimized definition of HRD. Section 3.2 details the full development for the LOH cut-off used for HRD classification of all patients without a tBRCA mutation for the purpose of ARIEL3 efficacy analyses.

A separate analysis plan will be created for the clinical validation of the companion diagnostic (CDx) test. This report will evaluate the primary endpoint of invPFS based on the CDx results and will also compare the CTA and CDx results. A multiple imputation analysis will also be performed in order to evaluate the effect of patients with an unknown CDx result.

Safety analyses are based on adverse event, vital signs, and physical exam collection and central laboratories measurements.

2.3 Sample Size

Approximately 540 patients will be randomized (2:1) to receive either rucaparib or placebo in this study. A minimum of 180 and a maximum of 200 patients with a deleterious tBRCA mutation will be enrolled. Enrollment of patients with a known deleterious germline *BRCA1/2* mutation (gBRCA) documented in their medical record will not exceed 150. There is no minimum number of patients required for each of the nbHRD and biomarker negative subgroups; however, no more than 360 total patients will be randomized for stratification into these subgroups combined.

Tumor HRD status by the CTA (see Section 3.2) will be determined after randomization, but before the final efficacy analysis, so that the primary endpoint (PFS in molecularly-defined HRD subgroups) can be assessed prospectively.

Table 2. Sample Size and Power Calculations

Group	Hazard Ratio	Cumulative N	Minimum Number of Events (70%)	Median PFS Placebo vs Rucaparib (months)	Power	One-sided Alpha
tBRCA	0.50	180	126	6 vs 12	90%	0.025
HRD	0.60	300	210	6 vs 10	90%	0.025
ITT	0.70	540	378	6 vs 8.5	90%	0.025

The study will end after 70% of the patients in the tBRCA subgroup have an observed event of investigator-determined disease progression or death. If the minimum number of tBRCA patients are enrolled, then the study will end following the 126th event of investigator-determined disease progression or death. Similarly, if the maximum number of tBRCA patients are enrolled then the study will end following the 140th event of investigator-determined disease progression or death. The Independent Data Monitoring Committee (IDMC) will inform the sponsor when the required number of PFS events have been observed in order to ensure the sponsor remains blinded to which patients are in the tBRCA subgroup.

2.4 Data Cut-off Used

The IDMC will inform the sponsor when the required number of PFS events have been observed and when treatment blind can be broken for the study. All data up to and including this date will be cleaned and collected before the treatment blind can be broken. In addition, key vendor data such as HRD classification from the Foundation Medicine CTA, central laboratory measurements, and Independent Radiology Review up to this date will also need to be collected. This process can take up to 4-8 weeks after the established visit cut-off for the primary analysis. The dataset will include all data that are collected and dated prior to or on the date of the visit cut-off date.

3 GENERAL ANALYSIS CONVENTIONS

The summary tables will be presented by randomized treatment group (rucaparib vs placebo) within the nested populations outlined in the step-down procedure in Section 3.2 and other subgroups of interest.

Quantitative variables will typically be summarized using frequencies and percentages for appropriate categorizations and may also be summarized using descriptive statistics. For variables summarized with descriptive statistics, the following will be presented: N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be presented using frequencies and percentages. The Kaplan-Meier methodology will be used to summarize time-to-event variables. If estimable, the 25th, 50th (median), and 75th percentiles will be presented along with the Kaplan-Meier estimates of event rates at 6-month intervals. The number of patients with events and the number of censored patients will also be presented. The stratified log rank test will be used to compare the PFS distributions between the randomized treatment groups. In addition, the Cox proportional hazards model will be used to estimate the hazard ratio between the randomized treatment groups. Months will be calculated as number of days divided by 30.4375.

All data will be used to their maximum possible extent, but without any imputations for missing data.

Unless otherwise indicated, baseline is defined as the last measurement on or prior to randomization.

3.1 Populations Definitions

Intent-to-Treat Population: The intent-to-treat (ITT) population will consist of all randomized patients.

Safety Population: The safety population will consist of all patients who received at least one dose of protocol-specified treatment.

In addition to the population definitions above, further analysis subgroups within the ITT population based on tumor HRD status by the CTA are outlined in Section 3.2 below.

3.2 Definition of HRD Subgroups

The CTA is a tumor tissue-based next generation sequencing (NGS) test that has been developed in collaboration with Foundation Medicine, Inc. (FMI). Tumors with potential deficiency in homologous recombination repair (HRR) may be identified by a gene mutation analysis approach or using a phenotypic genomic instability indicator such as loss of heterozygosity (LOH). The FMI CTA test is capable of both types of analyses, i.e., it can identify gene mutations for the purpose of classifying patients into 1 of 3 molecularly defined HRD subgroups for stratification at randomization (i.e., tBRCA, nbHRD, or biomarker negative as described in Appendix A), and can also assess the percentage of the tumor genome with LOH in order to

classify patients as tBRCA, non-tBRCA LOH+, non-tBRCA LOH-, or non-tBRCA LOH unknown.

Clovis is blinded to treatment assignment for the efficacy data and also blinded to the percent genomic LOH results from ARIEL3, and thus no results from ARIEL3 were used to inform the pre-specification of the LOH cut-off used for HRD determination.

The purpose of the ARIEL2 study has always been to generate rucaparib efficacy data to identify the optimal % LOH cut-off for prospective testing in study ARIEL3. ARIEL2 Part 1 is a Phase 2, open-label study of rucaparib in patients with platinum-sensitive, relapsed, high-grade epithelial ovarian, fallopian tube, or primary peritoneal cancer.

The 14% LOH cut-off prospectively tested in ARIEL2 Part 1 was based on analysis of the public TCGA dataset in high-grade serous ovarian tumors. In this dataset, clinical outcome data were only available for platinum-based therapy and not for PARP inhibitor therapy. Therefore, platinum sensitivity was used as a surrogate of rucaparib sensitivity. This analysis of the TCGA dataset demonstrated that 14% was the LOH cut-off that best differentiated overall survival outcome associated with platinum-based therapy. Therefore 14% was used as the pre-specified LOH cut-off to distinguish between the LOH-high and LOH-low populations in ARIEL2 Part 1 and to test the hypothesis that LOH-high tumors would be more sensitive to rucaparib than LOH-low tumors. This was documented in a Clovis internal memo signed on February 18, 2014.

While the 14% pre-specified LOH cut-off determined from TCGA platinum data significantly discriminated patient benefit in LOH-high and LOH-low tumors, initial analysis of ARIEL2 Part 1 suggested that a range of potential LOH cut-offs (16-18%) could potentially better identify tumors that were sensitive to rucaparib, based on efficacy endpoints of investigator assessed confirmed response rate by RECIST v1.1. or investigator assessed Progression-Free Survival (invPFS). The LOH cut-off of 18% was pre-specified for ARIEL2 Part 2 because it enrolled only fourth or fifth line patients, whose tumors may have accumulated more genomic scars and therefore have higher genomic LOH (Clovis internal memo signed on August 21, 2015).

The pre-specified LOH cut-off of 16% for prospective testing of ARIEL3 was selected following an analysis of mature clinical data from ARIEL2 Part 1, which enrolled platinum-sensitive patients similar to ARIEL3. Based on an analysis of invPFS for the subgroup of non-tBRCA patients in ARIEL2 Part 1, it was found that the greatest separation in PFS distribution between LOH-high vs LOH-low to be optimal using the 16% LOH cut-off (HR=0.455, $p<0.0001$). The pre-specification of 16% LOH cut-off for prospective testing in ARIEL3 [redacted]

Patients with a percent genomic LOH ≥ 16% will be classified as LOH high (LOH+). Therefore the HRD defined nested populations used in the step-down procedure is as follows:

1. **tBRCA:** Patients with a tBRCA mutation per the CTA.
2. **HRD (tBRCA or non-tBRCA LOH+):** Patients who are found to have a BRCA mutation per the CTA and/or to have LOH ≥ 16%
3. **ITT Population**

In addition, non-nested subgroups as noted below will also be explored for efficacy:

- **Non-tBRCA LOH+:** Patients without a tBRCA mutation and with percent of tumor genome LOH ≥ 16%
- **Non-tBRCA LOH-:** Patients without a tBRCA mutation and with percent of tumor genome LOH < 16%
- **Non-tBRCA LOH unknown:** Patients without a tBRCA mutation and with percent of tumor genome LOH unknown

3.3 Step-down Procedure

All efficacy evaluations will be conducted using the ITT population in the 3 nested HRD subgroups, unless otherwise specified:

1. **tBRCA:** Patients with a tBRCA mutation per the CTA.
2. **HRD (tBRCA or non-tBRCA LOH+):** Patients who are found to have a BRCA mutation per the CTA or do not have a BRCA mutation, but have tumor genomic LOH ≥ 16%
3. **ITT Population**

The primary and key secondary endpoints will be tested among the tBRCA, HRD, and ITT subgroups using an ordered step-down multiple comparisons procedure. invPFS in the tBRCA subgroup will be tested first at a one-sided 0.025 significance level. If invPFS in the tBRCA subgroup is statistically significant then invPFS will be tested in the all HRD subgroup followed by invPFS in the ITT population. Continuing in an ordered step-down manner, the remaining secondary endpoints will be tested at the one-sided 0.025 significance level in the tBRCA, HRD, and ITT subgroups. Once statistical significance is not achieved for one test, the statistical significance will not be declared for all subsequent analyses in the ordered step-down procedure.

The primary and secondary efficacy endpoints will be evaluated in the non-nested, non-overlapping subgroups of tBRCA, non-tBRCA LOH+, non-tBRCA LOH-, and non-tBRCA LOH unknown in order to ensure that the results in the HRD and ITT subgroups are not solely driven by the results in the tBRCA subgroup. In order to claim a significant result in the HRD subgroup, the size of the estimated effect in the non-tBRCA LOH+ subgroup should be clinically relevant and should be at least as large as what would be needed to achieve ‘‘statistical significance’’ in an analysis conducted in the entire HRD population. Similarly, for the ITT results to be considered significant and not solely driven by the results of the tBRCA or HRD subgroups, the size of the estimated effect in the non-tBRCA LOH- and non-tBRCA LOH unknown subgroups should be clinically relevant and should be at least as large as what would be needed to achieve ‘‘statistical significance’’ in an analysis conducted in the entire ITT population.

3.4 BRCA Subgroups

The tumor tissue-based CTA used in this study identifies sequence variants in the *BRCA1* and *BRCA2* genes; however, it does not distinguish between the mutation type, i.e., germline or somatic. In order to determine whether a *BRCA* mutation detected by the CTA is germline or

somatic, a central laboratory will sequence DNA extracted from blood. The germline and somatic designation for a tBRCA mutation will be based on the results of the central germline BRCA testing as described in Table 3.

Table 3. Algorithm for Determine Origin of BRCA Mutation - Germline versus Somatic – in the tBRCA Population

CTA Result	Central Germline Blood Test Result	Designation
BRCA Positive	BRCA Positive	Germline[a]
BRCA Positive	BRCA Negative	Somatic
BRCA Positive	Not tested	Unknown

[a] The same BRCA mutation should be detected in the CTA and central germline blood test result for a mutation to be considered germline.

4 PATIENT DISPOSITION

Patient disposition (analysis population allocation, randomized, discontinued, and primary reason for discontinuation) will be summarized using frequency counts, and the corresponding percentages.

5 PROTOCOL DEVIATIONS

Major protocol deviations will be identified via use of both central lab data and EDC data. The categories for major protocol deviations are in accordance with the ICH E3 guidelines[1] and will be identified prior to database lock. The deviations that will be deemed major fall under the following categories:

- Incorrect IP treatment compared to randomization
- Inclusion or exclusion criteria deviations
- Over-dose or under-dose of study drug
- Prohibited medications such as other anti-cancer medications
- Any other significant deviation that had the potential to affect the primary efficacy or safety of the patients

A listing with all the major protocol deviations will be presented.

6 DEMOGRAPHICS AND BASELINE CHARACTERISTICS

All demographic and baseline characteristics will be summarized for the ITT population.

6.1 Demographics

The demographic variables will be summarized with frequency tabulations that will focus on identifying differences between treatment groups in the extreme values of the distributions.

Descriptive statistics may also be used to summarize the quantitative variables. The demographic variables presented will include age, height, weight, gender, race, and ECOG performance status using the following categorizations:

- Age (years): ≤ 50, 51-60, 61-70, 71-80, 81-90, > 90;
- Height (cm): ≤ 75, > 75-100, > 100-125, > 125-150, > 150-175, > 175;
- Weight (kg): ≤ 50, > 50-75, > 75-100, > 100-125, > 125-150, > 150;
- Race: American Indian or Alaska Native, Asian, Black, Native Hawaiian or Other Pacific Islander, White, Other
- ECOG performance status: 0, 1
- Region: North America, Western Europe, Australia/New Zealand, and Israel

These categorizations may be adjusted if the majority of the data lies in only 2 or 3 of the categories.

6.2 Baseline Clinical Characteristics

The following variables will be summarized with frequency tabulations

- Time since diagnosis (months): > 12-24, > 24
- Type of Cancer (Epithelial Ovarian Cancer, Fallopian Tube Cancer, Primary Peritoneal Cancer, Other)
- Histological Classification (Serous, Endometrioid, Mixed, Other).
- Stratification subgroups used in the treatment randomization:
  - tBRCA, nbHRD, biomarker negative
  - Best response to most recent platinum-based regimen (complete response [CR], partial response [PR])
  - Interval between completion of the penultimate platinum-based regimen and disease progression (6 to 12 months or > 12 months)
- Subgroups based on data recorded on eCRFs:
  - Best response to most recent platinum-based regimen
  - Interval between completion of the penultimate platinum-based regimen and disease progression (6 to 12 months or > 12 months)
- tBRCA and LOH subgroups based on the CTA
  - tBRCA
  - non-tBRCA LOH+
  - non-tBRCA LOH-
  - non-tBRCA LOH unknown
- tBRCA mutation (*BRCA1*, *BRCA2*) based on the CTA
- BRCA mutation origin (germline, somatic, or unknown) based on central laboratory and CTA results

Descriptive statistics may also be used to summarize these variables.

6.3 Medical History

Medical history events will be classified using the Medical Dictionary for Regulatory Activities (MedDRA) classification system version 19.1. Medical history data will be summarized using frequency tabulations by System Organ Class (SOC) and Preferred Term (PT).

7 STUDY DRUG EXPOSURE AND COMPLIANCE

The following variables will be summarized:

- Duration of treatment
- Number of patients with at least one dose reduction or interruption

The duration of treatment will be calculated as the number of days from the first dose of study drug to the day of the last dose of study drug +1. The number of patients with at least one dose reduction or interruption based on the dispensation log will be summarized with frequencies and percentages. In addition, the number of patients at each dose level will be summarized (i.e., 600 mg BID, 480 mg BID, 360 mg BID) in order to assess patients with multiple levels of dose reductions.

8 PRIOR AND CONCOMITANT MEDICATIONS

All concomitant treatments documented during the study period will be summarized in frequency tabulations for each randomized treatment group and overall. Prior/concomitant medication coding will utilize the World Health Organization (WHO) Drug Dictionary version 2016SEP01DDE (Enhanced).

Prior medications will only be presented in listings. Prior medications will be defined as those medications with both a start and a stop date that is before the day of the first dose of study drug administration. If either the start date and/or the stop date of the medication is missing so that it is unclear whether the medication was stopped prior to first dose of study drug administration then the medication will be included in the summary of the concomitant medications.

9 EFFICACY VARIABLES

9.1 Primary Efficacy Variable

The primary efficacy variable is investigator determined PFS (invPFS).

9.2 Secondary Efficacy Variables Included in the Step Down Analysis

- DRS-P subscale of the FOSI-18 instrument
- Total score of the FOSI-18 instrument
- Final Overall Survival (OS)

9.3 Secondary Efficacy Variables Not Included in the Step Down Analysis

- Interim OS
- Independent Radiology Review Determination of PFS

10 EFFICACY ANALYSIS

10.1 Primary Efficacy Analysis

The primary efficacy endpoint is PFS as assessed by the investigator (invPFS). The time to invPFS will be calculated in months as the time from randomization to disease progression +1 day, as determined by RECIST v1.1 criteria[2] or death due to any cause, whichever occurs first.

Only scans and deaths prior to the start of any subsequent anti-cancer treatment or within 90 days of treatment end date will be included in the analysis. Patients without a documented event of progression will be censored on the date of their last tumor assessment (i.e., radiologic assessment) prior to the start of any subsequent anti-cancer treatment or within 90 days of treatment end date. Patients who withdrew without a disease progression event and do not have any post-baseline tumor assessment will be censored at date of randomization.

The stratified log rank test will be considered the primary analysis for invPFS comparing rucparib to placebo. In addition, a stratified Cox proportional hazard model will be used to calculate the hazard ratio (HR) between the treatment groups. The following randomization strata are used to estimate the treatment effect:

- HRD classification by the CTA (tBRCA, nbHRD, biomarker negative)
- Best response to most recent platinum-based regimen (CR, PR)
- Interval between completion of the penultimate platinum-based regimen and disease progression (6 to 12 months or > 12 months)

10.1.1 Sensitivity Analyses for PFS

10.1.1.1 Censoring Distribution

Sensitivity analyses for invPFS will be performed to evaluate the impact of censored patients. The following analyses will be performed:

- According to the study protocol, tumor scans were to continue to be performed during follow up for patients who discontinued without a documented disease progression event by RECIST v1.1. As such, a sensitivity analysis will be performed in which all tumor scans or death events will be included for assessment of PFS even if the patient discontinued study treatment or initiated a subsequent anticancer therapy.
- Patients who discontinued the study due to clinical progression will be considered to have a PFS event on the date of their last dose of treatment

Additional sensitivity analyses may also be performed to evaluate the robustness of the study results. These analyses will be considered exploratory and will likely be motivated by the observed results.

10.1.1.2 Interaction Between Treatment and HRD Status

In order to further evaluate the effectiveness of rucaparib in the HRD subgroups, the interaction between treatment and HRD status will be tested using the Cox proportional hazards model for the primary endpoint of invPFs. The model will include:

- Indicator variable for treatment with rucaparib;
- Categorical variable for HRD status; and
- Interaction between treatment and HRD status.

The HRD status is defined as the following mutually exclusive subgroups: tBRCA, non-tBRCA LOH+, non-tBRCA LOH-, and non-tBRCA LOH unknown. In addition, further exploratory analyses may be performed using the HRD definitions based on the mutation in Appendix A.

10.2 Secondary Efficacy Analyses Included in the Step Down Procedure

Secondary efficacy analyses will be based on nested populations (tBRCA, HRD, ITT). The multiple comparisons procedure presented in Section 10.1 will be used to control the type 1 error rate for the key secondary endpoints.

10.2.1 Disease Related Symptoms – Physical Subscale of the FOSI-18

The time to an event of worsening in the DRS-P subscale of the FOSI-18 will be defined as the time from randomization to a 4 point reduction in the DRS-P subscale. Patients without a documented event of a 4 point reduction will be censored on the date of their last adequate FOSI-18 assessment or date of randomization if no FOSI-18 assessments have been completed. For patients without a baseline FOSI-18 assessment their values will be censored at date of randomization.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.2.2 Total Score of the FOSI-18

The time to an event of worsening in the total score of the FOSI-18 will be defined as the time from randomization to an 8 point reduction in the total score. Patients without a documented event of an 8 point reduction will be censored on the date of their last adequate FOSI-18 assessment or date of randomization if no FOSI-18 assessments have been completed. For patients without a baseline FOSI-18 assessment their values will be censored at date of randomization.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.2.3 Independent Radiologic Review of PFS (irrPFS)

The time to irrPFS will be calculated in months as the time from randomization to disease progression + 1 day, as determined by the IRR or death due to any cause, whichever occurs first.

Only scans and deaths prior to the start of any subsequent anti-cancer treatment or within 90 days of treatment end date will be included in the analysis. Patients without a documented event of progression will be censored on the date of their last tumor assessment (i.e., radiologic assessment) prior to the start of any subsequent anti-cancer treatment or within 90 days of treatment end date. Patients who withdrew without a disease progression event and do not have any post-baseline tumor assessment will be censored at date of randomization.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.3 Secondary Efficacy Analyses Not Included in the Step Down Procedure

10.3.1 Overall Survival

The time to overall survival will be calculated in months as the time from randomization to date of death due to any cause. Patients who are still alive will be censored on the date of their last available visit or last date known to be alive.

It is anticipated that the data for overall survival will be heavily censored at the time of the primary endpoint analysis. In order to adjust for multiple analyses of overall survival at a later stage, a stopping rule will be applied. The Haybittle-Peto[3, 4] stopping rule will be applied where an overall survival result with a p-value <0.001 can be used to claim superiority of rucaparib compared to placebo. This means that a p-value <0.05 can be utilized at the final analysis which is projected to be once 70% of the death events has been collected.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.3.2 Population PK: Individual Model Parameter Estimates of Rucaparib and Covariates Identification.

A population PK analysis will be performed and documented in a separate report.

10.4 Exploratory Endpoints

Exploratory endpoints will be presented separately for the three nested analysis populations (i.e., tBRCA, HRD, and ITT population). The exploratory endpoints include change in CA-125 level, tumor response, and DOR in patients who had measurable disease per the investigator at baseline. The purpose of the exploratory endpoint is to further explore the efficacy. No multiple adjustment is performed for these analyses.

10.4.1 Change from Baseline in CA-125 Measurement

Analyses of changes and/or percent changes from baseline for CA-125 measured by the central laboratory will be analyzed for each scheduled post-baseline visit and for the final visit. Patients who do not have both a baseline measurement and at least one post-baseline measurement will not be included in these analyses.

At a given visit, the change and/or percent change from baseline for CA-125 will be compared between the randomized treatment groups using an ANCOVA using the treatment as a categorical factor and baseline measurement for the parameter as a continuous covariate.

10.4.2 Progression Free Survival 2 (PFS2)

The second event of PFS (referred to as PFS2) is defined as the time from randomization to the second event of disease progression as assessed by the investigator, or death due to any cause. The first event of disease progression will be captured as the primary endpoint in this study and thus the second event will be the next event of disease progression as assessed by the investigator. This second event of PFS may be a documented event per RECIST guidelines or may be an event of symptomatic/clinical or CA-125 progression.

10.4.3 ORR per RECIST v1.1 as Assessed by the Investigator

The Overall Response Rate (ORR) as assessed by the investigator will be analyzed in the subgroup of patients who have measurable disease (i.e., measurable target lesions) at baseline. Both unconfirmed and confirmed response rate by RECIST v1.1 will be summarized. The confirmed response rate is defined as the proportion of patients with a confirmed CR or PR on subsequent tumor assessment at least 28 days after first response documentation. The ORR will be summarized with frequencies and proportion together with 95% confidence interval (CI) and compared between treatments by using a stratified Cochran-Mantel-Haenszel (CMH) test. In addition, the frequency and proportion of patients will be summarized for each of the unconfirmed and confirmed response categories:

- CR
- PR
- Stable disease (SD)
- Progressive disease (PD)
- Not evaluable (for example, discontinuation or death before first tumor assessment)

10.4.4 Duration of Response (DOR) per RECIST v1.1 as Assessed by the Investigator

The Duration of Response (DOR) as assessed by investigator will be analyzed in the subgroup of patients who have measurable disease (i.e., measurable target lesions) at baseline and have a confirmed response by RECIST v1.1. DOR for any confirmed RECIST CR or PR will be measured from the date of the first response until the first date that PD is documented. DOR will

be summarized as a time to event variable. For patients who continue treatment post-progression, the first date of progression will be used for the analysis. Any patients with an ongoing response will be censored at the date of the last post-baseline scan.

The Kaplan-Meier methodology will be used to summarize DOR. If able to be estimated, the 50th (median) together with a 95% CI, will be presented. The stratified log rank test comparing the treatment groups will also be presented. The number of patients with PD events and the number of censored patients will also be presented.

10.4.5 Change from Baseline in Sum of the Diameters of Target Lesions as Assessed by the Investigator

The change from baseline in sum of diameters of target lesions as assessed by the investigator will be analyzed in the subgroup of patients who have measurable disease (i.e., measurable target lesions) at baseline. The largest percent decrease from baseline in the sum of the diameters of target lesions as identified by RECIST v1.1 will be displayed graphically using a waterfall plot. This analysis will be based on the investigator assessed tumor responses and only presented for patients with measurable disease at baseline and one valid post-baseline evaluation of the target lesions.

10.4.6 Patient Reported Outcome of EQ-5D

Analyses of changes and/or percent changes from baseline will be analyzed for each scheduled post-baseline visit and for the final visit for EQ-5D VAS. Patients that do not have both a baseline measurement and at least one post-baseline measurement will not be included.

At a given visit, the change and/or percent change from baseline will be compared between the randomized treatment groups using an ANCOVA using the treatment as a categorical factor and baseline measurement for the parameter as a continuous covariate.

10.4.7 Exploratory Endpoint: Association Between the Change from Baseline in CA-125 Measurement and invPFS

The association between the change from baseline in CA-125 and invPFS will be analyzed by adding the maximum decrease from baseline in CA-125 to the stratified Cox proportional model as a covariate in the model. The change or percent change will be analyzed both as a continuous variable and as a categorical variable depending on the data distribution.

10.4.8 Exploratory Endpoint: Chemo-Free Interval

The chemo-free interval will be calculated in months as the time since the last dose of the most recent chemotherapy regimen to the date of the first dose of a subsequent chemotherapy after study drug + 1 day. Patients without a documented start of a subsequent chemotherapy after study drug will be censored on the date of their last available assessment.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.4.9 Exploratory Endpoint: Time to Start of First Subsequent Anti-cancer Treatment

The time to start of the first subsequent anti-cancer treatment will be calculated in months as the time from randomization to the date of the first dose of the first subsequent anti-cancer treatment regimen after study drug + 1 day. Patients without a documented start of a subsequent anti-cancer treatment after study drug will be censored on the date of their last available assessment.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.4.10 Exploratory Endpoint: Time to Start of Second Subsequent Anti-cancer Treatment

The time to start of the second subsequent anti-cancer treatment will be calculated in months as the time from randomization to the date of the first dose of the second subsequent anti-cancer treatment regimen after study drug + 1 day. Patients without a documented start of a second subsequent anti-cancer treatment after study drug will be censored on the date of their last available assessment.

The same statistical test used for the primary endpoint (i.e., stratified log rank test and a stratified Cox proportional model) will be used to compare rucaparib to placebo for this endpoint.

10.4.11 Exposure- Response PK Report: Rucaparib PK, invPFS, irrPFS, CA-125, AEs, Clinical Laboratory Abnormalities, and Dose Modifications

An Exposure-Response PK analysis will be performed and documented in a separate report.

11 STATISTICAL / ANALYTICAL ISSUES

11.1 Statistical Hypothesis

The primary statistical hypothesis for this study is to test for differences in PFS within each of the nested populations (tBRCA, HRD and ITT) using a stratified log-rank test.

H_0: HR (rucaparib/placebo) ≥ 1.
H_a: HR (rucaparib /placebo) < 1.

Stratification factors include:

- HRD classification tBRCA or nbHRD or biomarker negative) as determined by CTA
- Interval between completion of the penultimate platinum-based regimen and disease progression (6 to 12 months or >12 months)
- Best response (RECIST CR or PR [RECIST and/or GCIG CA-125] response) to platinum regimen received immediately prior to initiation of rucaparib maintenance therapy

All efficacy analyses will be based on the randomization strata given at randomization, however, a sensitivity analysis of invPFS may be performed using the actual strata if patients have been allocated incorrectly.

11.2 Handling of Dropouts or Missing Data

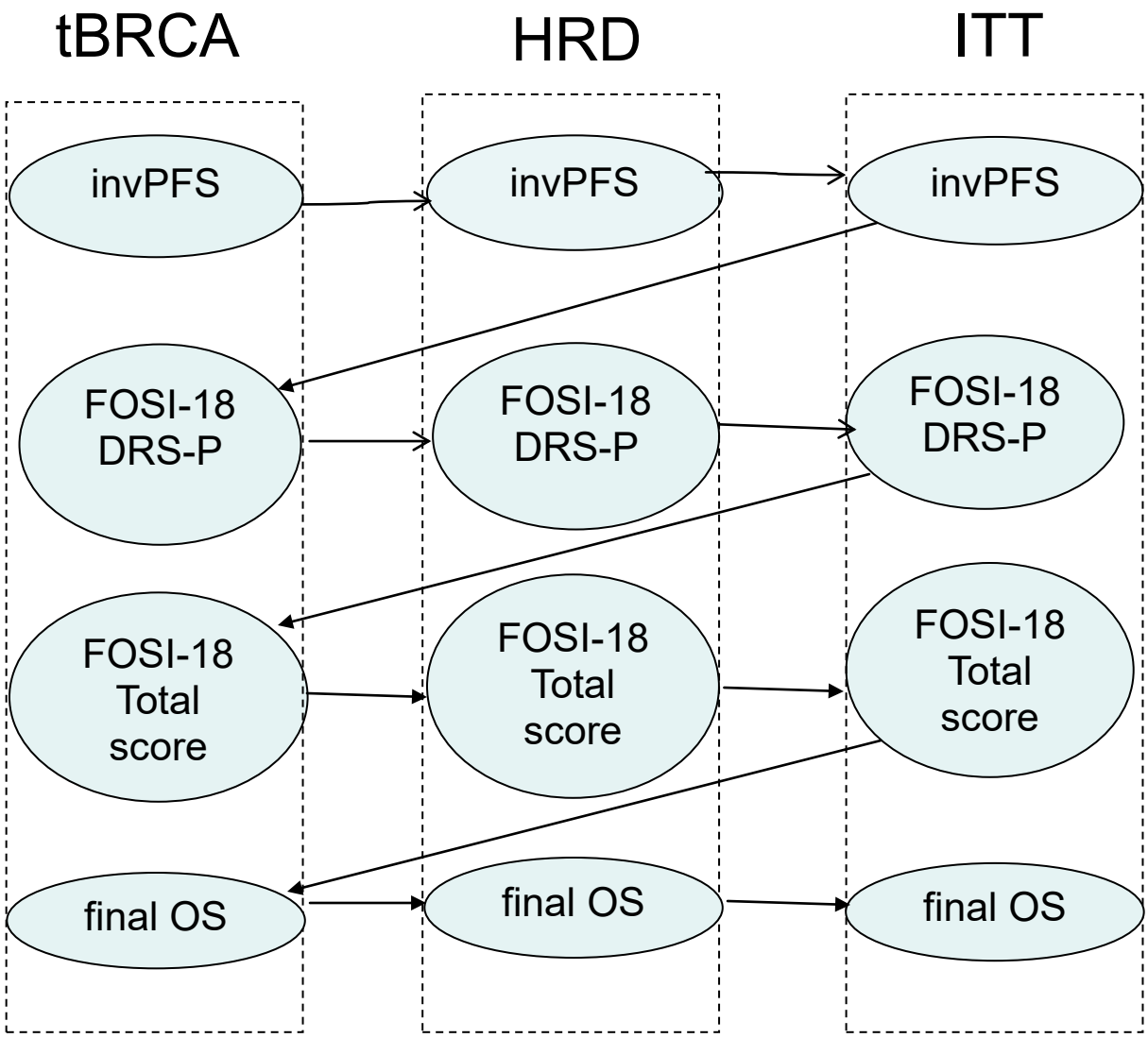
All patients with a deleterious tBRCA mutation based on the CTA will be included in the HRD positive subgroup even if the percent of genomic LOH is unknown. Patients without a deleterious tBRCA mutation and without a genomic LOH result will only be included in the analyses in the ITT and Safety Populations.

Imputation of missing baseline values for the FOSI-18 will be performed as supportive analyses. For patients without a baseline value, the baseline value will be imputed by using the FOSI-18 assessment performed closest to randomization and may include assessments performed up to and including Cycle 2 Day 1.

11.3 Multiple Comparison / Multiplicity

The primary and key secondary endpoints will be tested among the subgroups of tBRCA, HRD, and ITT using the ordered step-down procedure as illustrated in Figure 1 below.

Figure 1. Ordered Step-down Procedure



11.4 Examination of Efficacy in Subgroups

Subgroup analyses of the primary endpoint (invPFS) will be presented for the following:

- Randomization stratification factors
  - tBRCA, nbHRD, or biomarker negative as determined by CTA for randomization (per gene list, as defined in Appendix A)
  - Interval between completion of the penultimate platinum-based regimen and disease progression (6 to 12 months or > 12 months)
  - Best response (RECIST CR or PR [RECIST and/or GCIG CA-125 response]) to platinum regimen received immediately prior to initiation of rucaparib maintenance therapy

- HRD definition used for analysis (tBRCA, non-tBRCA LOH+, non-tBRCA LOH-, non-tBRCA LOH unknown), as further described in Section 3.3
- Age groups (< 65, 65-74, ≥ 75)
- Race (white, non-white)

Subgroups by Disease Burden at Baseline:

In addition, the primary endpoint (invPFS) will be further explored based on the disease burden at baseline in the following subgroups of patients:

- **Measurable Disease**: All patients who have measurable disease (i.e., target lesion of any size) at baseline as assessed by the investigator and/or independent radiology reviewer).
- **No Disease**: All patients who have no target lesions or non-target disease at baseline as assessed by the investigator and/or independent radiology reviewer).
- **No Bulky Disease**: All patients with any/all lesions < 2 cm in the shortest axis (lymph nodes) or longest axis (all other lesions) as assessed by the independent reviewer. A reassessment of baseline scans was conducted by independent radiology reviewer in order to define this subgroup. The shortest axis is used as the diameter for malignant lymph nodes, and the longest axis is used for all other measurable lesions, in accordance with RECIST 1.1.

Subgroups Based on Gene Mutation and Type:

In addition, the primary endpoint (invPFS) will be further explored based on mutation and mutation type.

- BRCA mutation (*BRCA1*, *BRCA2*)
- BRCA mutation origin (germline, somatic, unknown)
- Patients with tumors harboring a deleterious mutation within the pre-specified list of non-BRCA HRR genes [redacted] This pre-specified list of non-BRCA HRR genes was derived based on biomarker analysis of ARIEL2, and is more specific than the nbHRD gene list used for stratification.
- Combining the somatic tBRCA subgroup with the non-tBRCA LOH+ subgroup.
- Comparing germline tBRCA vs all others (i.e., all non-germline tBRCA and non-tBRCA patients)

In addition, a sensitivity analysis of invPFS will be performed for all BRCA patients, including those identified as BRCA by the CTA and those identified by a local or central BRCA test.

12 SAFETY ANALYSIS

The safety analyses will be performed using the safety population.

12.1 Adverse Events

AEs will be classified using the MedDRA version 19.1 classification system. The severity of the toxicities will be graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) whenever possible. Treatment-emergent adverse events (TEAEs) are defined as AEs with onset date on or after the date of first dose of study medication until the date of the last study medication dose plus 28 days. AEs will be considered treatment-emergent if all or part of the date of onset of the adverse event is missing and it cannot be determined if the AE meets the definition for treatment-emergent.

The number and percentage of patients who experienced TEAEs for each SOC and PT will be presented. Multiple instances of the TEAE in each SOC and multiple occurrences of the same PT are counted only once per patient. The number and percentage of patients with at least one TEAE will also be summarized.

Separate tables will be presented as follows:

- All TEAEs;
- TEAEs by CTCAE grade;
- Grade 3 or greater TEAEs;
- Treatment-related TEAEs;
- Serious TEAEs;
- TEAEs with an outcome of death;
- TEAEs leading to discontinuation of study medication;
- TEAEs resulting in reduction study medication;
- TEAEs resulting in interruption of study medication;
- TEAEs resulting in reduction or interruption of study medication;
- Time to the first TEAE that results in a reduction or interruption of study drug.

The incidence of TEAEs will be summarized by relationship to study drug according to the following categories: “treatment-related,” or “not treatment-related”. If a patient experiences multiple occurrences of the same AE with different relationship categories, the patient will be counted once, as a relationship category of treatment related.

If a patient experiences multiple occurrences of the same AE with different toxicity grades, the patient will be counted once for the maximum (most severe) toxicity grade. AEs with a missing toxicity grade will be presented in the summary table with a toxicity grade of “Missing.” For each toxicity grade, the number and percentage of patients with at least one TEAE of the given maximum grade will be summarized.

The time to the first TEAE and first treatment-related TEAE that results in a dose reduction, delay, interruption, or discontinuation of study drug is defined as 1+ the number of days from the first dose of study drug to the start of the first adverse event. The cumulative incidence is presented in a 1-KM graph for just the patients with an event and the median time to onset will be calculated together with the 95% CI.

Non-TEAEs (pre-treatment and post-treatment) will be presented in the by patient data listings for the safety population.

MedDRA PTs were combined for the following similar terms:

- Asthenia/Fatigue
- Alanine Aminotransferase (ALT)/ Aspartate Aminotransferase (AST) Increased
- Anaemia and/or Low/Decreased Haemoglobin
- Thromobocytopenia and/or Low/Decreased Platelets
- Neutropenia and/or Low/Decreased Absolute Neutrophil Count (ANC)

In addition, the analysis of combined terms for anemia is explored as a time to first event analysis as described above. Transfusions (blood or plasma) and concomitant medications / growth factor support are provided in patient listings. The number of transfusions and the time to fist transfusion is also summarized.

12.2 Clinical Laboratory Evaluations

Clinical laboratory evaluations include the continuous variables for hematology, serum chemistry, and urinalysis. The laboratory values will be presented in SI units. The on-treatment period will be defined as the time from randomization to 28 days after the last dose of study drug. Laboratory values collected during the on-treatment period will be included in the summary tables. The laboratory values collected after the on-treatment period will only be presented in the data listings.

The summary of laboratory data will include descriptive statistics (N, mean, SD, minimum, median, and maximum) of the maximum, minimum, and last value during the on-treatment period. Summaries using descriptive statistics of the change from baseline to the maximum, minimum, and last value during the on-treatment period will also be given.

Summary tables from baseline to the maximum on-treatment toxicity grade (CTCAE Version 4.03) for each lab parameter will be summarized. In addition, the number and percent of patients who has at least one worsening in toxicity grade on-treatment compared to baseline toxicity grade will be summarized by treatment group.

Laboratory data including normal ranges and flagged abnormal laboratory findings will be provided using by-patient listings.

12.3 Vital Signs

The on-treatment period will be defined as the time from randomization to 28 days after the last dose of study drug. Vital sign measurements collected during the on-treatment period will be included in the summary tables. The vital sign measurements collected after the on-treatment period will only be presented in the data listings.

The summary of vital sign data will include descriptive statistics (N, mean, SD, minimum, median, third quartile and maximum) of the maximum, minimum and last value during the on-treatment period. Summaries using descriptive statistics (N, mean, SD, minimum, median, and maximum) of the change from baseline to the maximum, minimum, and last value during the on-treatment period will also be given. The data will be presented separately for each randomized treatment group, and overall.

12.4 Examination of Safety in Subgroups

Safety will be further explored in the following subgroups:

- HRD subgroups per efficacy nested populations (tBRCA, HRD)
- tBRCA subgroups (germline, somatic)
- Age groups (< 65, 65-74, ≥ 75)
- Race (white, non-white)

13 REFERENCES

1. International Conference on Harmonisation E3 Implementation Working Group/ICH E3 Guideline: Structure and Content of Clinical Study Reports/Questions & Answers (version 7Jun2012). 2012.

2. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer. 2009;45(2):228-47.

3. Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol. 1971;44(526):793-7.

4. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer. 1976;34(6):585-612.

APPENDICES

Appendix A List of Homologous Recombination Repair Genes for HRD Stratification by the CTA

tBRCA	nbHRD	Biomarker-negative
BRCA1 *BRCA2*	*ATM* *ATR* *ATRX* *BARD1* *BLM* *BRIP1* *CHEK1* *CHEK2* *FANCA* *FANCC* *FANCD2* *FANCE* *FANCF* *FANCG* *FANCI* *FANCL* *FANCM* *MRE11A* *NBN* *PALB2* *RAD50* *RAD51* *RAD51B* *RAD51C* *RAD51D* *RAD52* *RAD54L* *RPA1*	Genes not included in the tBRCA or nbHRD groups

Appendix B

[redacted]

CLOVIS ONCOLOGY

Internal memo

Date: [redacted]
From: [redacted]
To:
Subject: [redacted]
Distribution: Note to file

[redacted]

CLOVIS ONCOLOGY

[redacted]

Prepared By:

[redacted]

Approved By:

[redacted]

[redacted]

[redacted]

Appendix C [redacted]