NOVARTIS

Clinical Development

BYL719/Alpelisib

CBYL719K12301/ NCT04729387

EPIK-O: A Phase III, multi-center, randomized (1:1), open-label, active-controlled study to assess the efficacy and safety of alpelisib (BYL719) in combination with olaparib as compared to single agent cytotoxic chemotherapy, in participants with no germline BRCA mutation detected, platinum-resistant or refractory, high-grade serous ovarian cancer

Statistical Analysis Plan (SAP)

Author:	PPD
Document type:	SAP Documentation
Document status:	Final Amendment 1
Release date:	20-Jun-2023
Number of pages:	77

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
CCI				

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
CCI				

Table of contents

List of tables

List of abbreviations

aBC	advanced breast cancer
AE	Adverse event
AI	aromatase inhibitor
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
BC	breast cancer
BOR	Best overall response
BRCA	Breast Cancer gene
BRCAnm	BRCA non-mutated
CBR	Clinical benefit rate
CR	Complete response
CRS	Case retrieval strategy
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
ΔQTcF	Change from baseline in QTcF
DMC	Data Monitoring Committee
DOR	Duration of Response
FAS	Full Analysis Set
eCRF	Electronic Case Report Form
ECOG	Eastern Cooperative Oncology Group
EOT	End of treatment
FAS	Full analysis set
HER2	Human epidermal growth factor receptor 2
HGSOC	High-grade serous ovarian cancer
HR	Hazard Ratio
IRT	Interactive Response Technology
MedDRA	Medical Dictionary for Drug Regulatory Affairs
NCI	National Cancer Institute
NI	Non-inferiority
NMQ	Novartis MedDRA queries
NSAI	Non-steroidal aromatase inhibitor
ORR	Overall response rate
PAS	Pharmacokinetic analysis set
PD	Progressive disease
PDI	Planned dose intensity

PDS	Programming Datasets Specifications
PFS	Progression-Free Survival
PK	Pharmacokinetics
PPS	Per-Protocol Set
PR	Partial response
PT	Preferred term
qd	Qua'que di'e / once a day
QTcF	QT interval corrected by Fridericia method
RAP	Report and Analysis Process
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SEC	Safety Event Categories
SMQ	Standardized MedDRA queries
SOC	System Organ Class
TA	Tumor assessment
TBIL	Total Bilirubin
TEAE	Treatment-emergent adverse event
TFLs	Tables, Figures, Listings
WHO	World Health Organization
NE	Not Evaluable

1 Introduction

This document describes the detailed statistical methodology to be used for the clinical study report (CSR) for the primary analysis of study CBYL719K12301, a Phase III, multicenter, randomized (1:1), open-label, active-controlled study to assess the efficacy and safety of alpelisib (BYL719) in combination with olaparib compared to single agent cytotoxic chemotherapy in participants with platinum resistant or refractory HGSOC with no germline BRCA mutation detected.

The content of this SAP is based on the CBYL719K12301 protocol amendment 1 released on April 28th, 2021. All decisions regarding primary PFS analysis, as defined in this document, have been made prior to the database lock of study data. The primary CSR will be produced after the primary PFS analysis.

CSR deliverables (shells for tables, figures, listings) and further programming specifications are described in the Tables, Figures & Listings (TFL) shells and Programming Datasets Specification (PDS), respectively.

1.1 Study design

This is an open-label, randomized, active-controlled, international, multicenter Phase III trial designed to evaluate the efficacy and safety of alpelisib in combination with olaparib compared to single agent cytotoxic chemotherapy in participants with platinum resistant or refractory HGSOC with no germline BRCA mutation detected.

During the Molecular Pre-Screening period, all participants will be required to have their BRCAnm status determined. Once BRCAnm status is confirmed, participants will enter into Screening. Screening is up to 28 days in duration and will be used to assess eligibility of the participants to enter the Treatment period and to collect baseline values for some variables. Participants will be eligible regardless of their tumor PIK3CA mutation status.

Approximately 358 participants will be randomized in a 1:1 ratio to receive:

- Experimental Arm (Arm 1): alpelisib 200 mg orally once daily with olaparib 200 mg orally twice daily starting on Cycle 1 Day 1 in a 28-day cycle.

OR

- Control Arm (Arm 2): Investigator's choice of paclitaxel 80 mg/m^2 intravenously weekly in a 28 day treatment cycle, starting on Cycle 1 Day 1; or pegylated liposomal doxorubicin (PLD) 40-50 mg/m^2 (physician discretion) intravenously every 28 days in a 28 day treatment cycle, starting on Cycle 1 Day 1.

Randomization will be stratified by:

- Time to relapse from last platinum dose (< 3 months vs. ≥ 3 and ≤ 6 months)
- Prior PARP inhibitor use (yes vs. no)
- Prior bevacizumab use (yes vs. no)

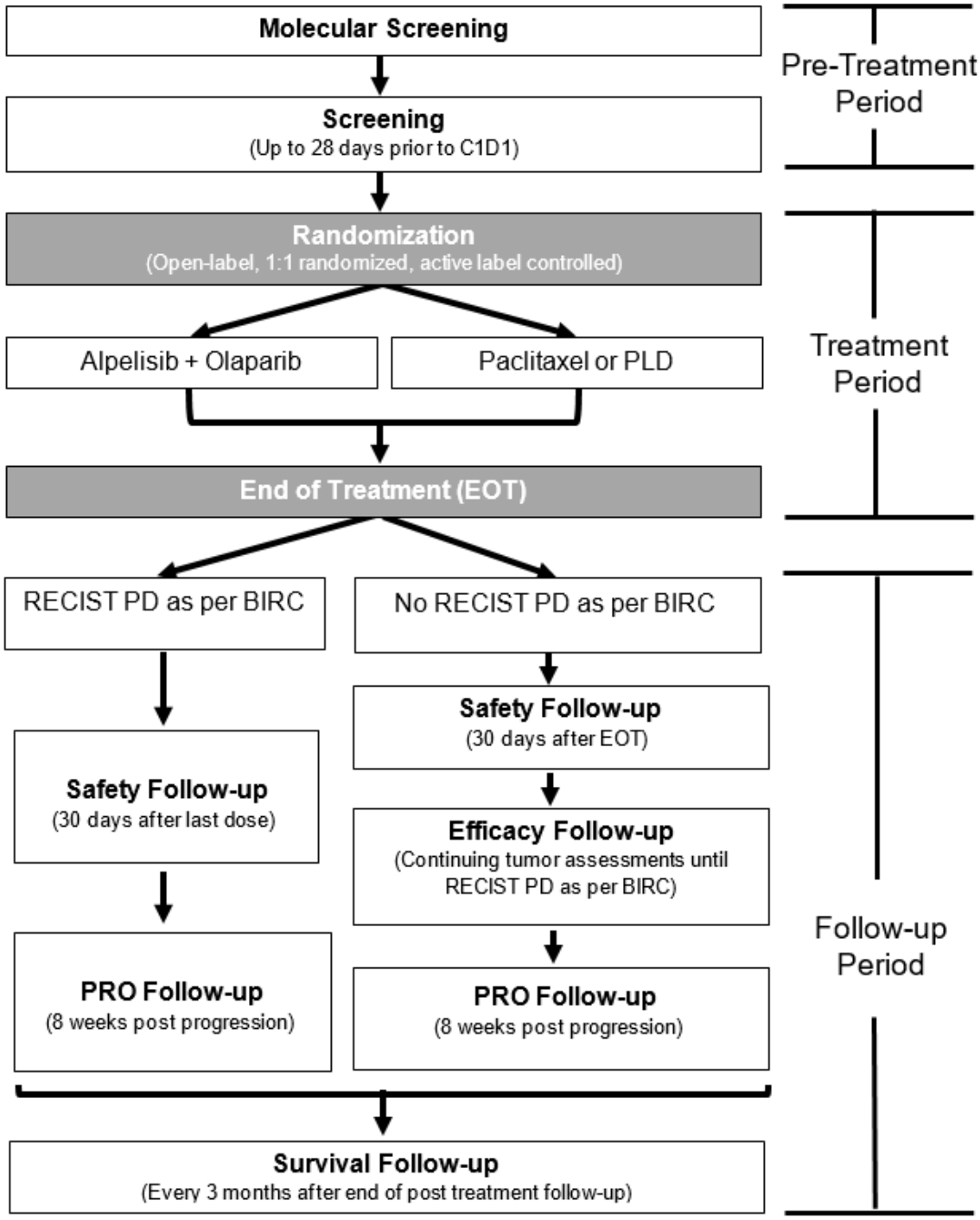
The total number of participants with non-measurable disease will be limited to up to CCI of the overall study population.

Treatment cross-over from one arm to another arm will not be permitted in this study; no switching between chemotherapy agents is allowed within Arm 2 once the first dose of chemotherapy is administered.

Progression-free survival (PFS) as assessed by Blinded Independent Review Committee (BIRC) assessments and using RECIST 1.1 criteria is the primary endpoint. Overall survival (OS) is the key secondary endpoint. For details of the interim analyses for PFS and OS, please refer to Section 2.10. An independent Data Monitoring Committee (DMC) will monitor unblinded safety and efficacy data accruing during the trial. The interim PFS futility analysis (see Section 2.10 for details) will be performed by an independent statistician external to Novartis and the results will be provided to the DMC by the independent statistician. A separate DMC SAP will specify the analyses to be performed for the DMC reviews.

After treatment discontinuation, all participants will enter the post-treatment follow-up period which consists of a safety follow-up visit and a 8-week post-progression visit. Participants who discontinued for reasons other than disease progression, death, lost to follow-up, or withdrawal of consent will enter a post-treatment efficacy follow-up. Participant will then enter survival follow-up as described in the study flow Figure 1-1.

Figure 1-1 Study Flow


Molecular Screening
Screening
(Up to 28 days prior to C1D1)
Pre-Treatment Period
Randomization
(Open-label, 1:1 randomized, active label controlled)
Alpelisib + Olaparib
Paclitaxel or PLD
Treatment Period
End of Treatment (EOT)
RECIST PD as per BIRC
No RECIST PD as per BIRC
Safety Follow-up
(30 days after last dose)
Safety Follow-up
(30 days after EOT)
Efficacy Follow-up
(Continuing tumor assessments until RECIST PD as per BIRC)
PRO Follow-up
(8 weeks post progression)
PRO Follow-up
(8 weeks post progression)
Follow-up Period
Survival Follow-up
(Every 3 months after end of post treatment follow-up)

1.2 Study objectives, endpoints and estimands

Objectives and related endpoints are described in Table 1-1 below.

Table 1-1 Objectives and related endpoints

	Objective	Endpoint(s)
Primary	To determine whether treatment with alpelisib in combination with olaparib prolongs PFS compared to single agent cytotoxic chemotherapy in participants with platinum-resistant or refractory, gBRCAnm HGSOC	PFS based on BIRC assessment using RECIST 1.1 criteria
Key Secondary	To determine whether treatment with alpelisib in combination with olaparib prolongs OS compared to single agent cytotoxic chemotherapy in participants with platinum-resistant or refractory, gBRCAnm HGSOC	OS
Other Secondary	To assess safety and tolerability of alpelisib in combination with olaparib when administered to participants with platinum-resistant or refractory, gBRCAnm HGSOC	**Safety**: Incidence, type and severity of adverse events per CTCAE v4.03 criteria including changes in laboratory values, vital signs, hepatic, renal and cardiac assessments. **Tolerability:** dose interruptions, reductions, dose intensity and duration of exposure for all drug components.
	To evaluate alpelisib in combination with olaparib compared to single agent cytotoxic chemotherapy in participants with platinum-resistant or refractory, gBRCAnm HGSOC with respect to time to deterioration of ECOG (Eastern Cooperative Oncology Group) performance status	Time to definitive deterioration of the ECOG performance status from baseline

	Objective	**Endpoint(s)**
	To assess additional efficacy parameters	Overall Response Rate (ORR) with confirmed response, Clinical Benefit Rate (CBR) with confirmed response, Duration Of Response (DOR) with confirmed response and Time To Response (TTR) (based on BIRC assessment using RECIST 1.1 criteria)
	To characterize the PK of alpelisib and olaparib when administered in combination in participants with platinum-resistant or refractory, gBRCAnm HGSOC.	Summary of statistics of PK parameters (including but not limited to AUCtau, AUClast, Cmax, Tmax) of alpelisib and olaparib (full PK subset only), Summary of statistics of plasma alpelisib and olaparib concentrations by time point.
	To evaluate patient reported-outcomes of alpelisib in combination with olaparib compared to single agent cytotoxic chemotherapy in adult participants with platinum-resistant or refractory, gBRCAnm HGSOC	Change from baseline in Function Assessment of Cancer Therapy-Ovarian Trial Outcome Index (FACT-O)
Exploratory		

1.2.1 Primary estimand

The estimand is the precise description of the treatment effect and reflects strategies to address events occurring during trial conduct which could impact the interpretation of the trial results (e.g premature discontinuation of treatment). The primary scientific question of interest of this study is: what is the treatment effect based on PFS for alpelisib in combination with olaparib versus single agent cytotoxic chemotherapy (paclitaxel or PLD) in participants with gBRCAnm, platinum-resistant or refractory high grade serous ovarian cancer, regardless of study treatment discontinuation or start of new anti-neoplastic therapy?

The justification for targeting this treatment effect is to assess the treatment effect not only during the on-treatment period but also after the discontinuation of study treatment, i.e., during the entire course of the study; and compare not just alpelisib + olaparib vs. SoC, but alpelisib + olaparib followed by any new anti-neoplastic therapy vs. SoC followed by any new anti-neoplastic therapy.

The primary estimand is characterized by the following attributes:

1. Population: all participants randomized with gBRCAnm, platinum-resistant or refractory high-grade serous or high-grade endometrioid ovarian cancer. Further details on the population are provided in Section 5 of the study protocol.
2. Treatment: the investigational treatment is alpelisib in combination with olaparib plus any subsequent anti-neoplastic therapy as needed. The control treatment is single agent cytotoxic chemotherapy (paclitaxel or PLD) plus any subsequent anti-neoplastic therapy as needed. Further details about the investigational treatment and control treatment are provided in Section 6 of the study protocol.
3. Variable: PFS based on BIRC assessment and using RECIST 1.1 criteria. Further details on PFS are provided in Section 2.5.1.
4. Intercurrent events:
   - discontinuation of study treatment for any reason

Details on how to handle intercurrent events are provided in Section 2.5.3.

5. Summary measure: PFS hazard ratio (alpelisib + olaparib versus single agent cytotoxic chemotherapy (paclitaxel or PLD)) along with 95% confidence interval, estimated using a Cox proportional hazard model stratified by the randomization stratification factors. Further details on how the summary measure will be tested are provided in Section 2.5.2.

1.2.2 Key secondary estimand

The key secondary scientific question of interest is: what is the treatment effect based on OS for alpelisib in combination with olaparib versus single agent cytotoxic chemotherapy (paclitaxel or PLD) in participants with gBRCAnm, platinum-resistant or refractory high-grade serous or high-grade endometrioid ovarian cancer regardless of study treatment discontinuation or start of new anti-neoplastic therapy?

The justification for targeting this treatment effect is to assess the treatment effect based on OS and compare not just alpelisib + olaparib vs. SoC, but alpelisib + olaparib followed by any new anti-neoplastic therapy vs. SoC followed by any new anti-neoplastic therapy, i.e., any subsequent anti-neoplastic therapy is part of treatment attribute.

The key secondary estimand is characterized by the following attributes:

- Population: all participants randomized with gBRCAnm, platinum-resistant or refractory high grade serous ovarian cancer. Further details on the population are provided in Section 5 of the study protocol.
- Treatment: the investigational treatment is alpelisib in combination with olaparib plus any subsequent anti-neoplastic therapy as needed. The control treatment is single agent cytotoxic chemotherapy (paclitaxel or PLD) plus any subsequent anti-neoplastic therapy as needed. Further details about the investigational treatment and control treatment are provided in Section 6 of the study protocol.
- Variable: OS. Further details on OS are provided in Section 2.6.1.
- Intercurrent events:
  - discontinuation of study treatment for any reason

Details on how to handle intercurrent events are provided in Section 2.6.3.

- Summary measure: OS hazard ratio (alpelisib + olaparib versus single agent cytotoxic chemotherapy (paclitaxel or PLD)) along with 95% confidence interval, estimated using a Cox proportional hazard model stratified by the randomization stratification factors. Further details on how the summary measure will be tested are provided in Section 2.6.2.

2 Statistical methods

2.1 Data analysis general information

The final PFS analysis will be performed by Novartis. The interim PFS analysis will be performed by an independent statistician external to Novartis for review by a DMC. SAS version 9.4 or later will be used to perform all data analyses and to generate tables, figures, and listings.

Data included in the analysis

The primary efficacy and safety analyses for the study will be performed after observing approximately 224 PFS events per BIRC assessment. The primary CSR will be produced after the primary PFS analysis.

There is one futility interim (at approximately 40% information fraction) and one final PFS analysis planned for the primary efficacy endpoint. Up to two interim and one final OS analysis may be performed for the key secondary endpoint. A unique cut-off date will be established after the targeted number of events for each of the planned interim and final analyses has been documented. For each of the analyses, all statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as ‘ongoing’. The same rule will be applied to events starting before or on the cut-off

date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

Following the cut-off date for the primary analysis reported in the primary CSR, the study will remain open. Any additional data for participants continuing to receive study treatment past this time and for participants continuing for efficacy follow-up (PFS, OS), as allowed by the protocol, will be further summarized in a study report at the time of the final OS analysis after observing approximately 252 OS events, or when statistical significance is reached at any interim OS analysis.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of participants enrolled at centers, no center effect will be assessed.
Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by treatment group; a missing category will be included as applicable. Percentages will be calculated using the number of participants in the relevant population or subgroup as the denominator.
Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, minimum, and maximum) by treatment group.

2.1.1 General definitions

2.1.1.1 Investigational drug and study treatment

Investigational drug will refer to alpelisib and olaparib. **Control drug** will refer to investigator's choice of paclitaxel or PLD. Whereas, **study treatment** will refer to alpelisib + olaparib and investigator's choice of paclitaxel or PLD.

The term investigational treatment may also be referred to as **study treatment** which is used throughout this document.

2.1.1.2 Date of first administration of study treatment component

The date of first administration of certain component of study treatment is defined as the first date when a non-zero dose of that component of study treatment is administered and recorded on the respective component-specific Study Treatment eCRF. The date of first administration of component of study treatment will also be referred to as the start of component of study treatment.

2.1.1.3 Date of last administration of study treatment component

The date of last administration of certain component of study treatment is defined as the last date when a non-zero dose of that component of study treatment is administered and recorded

on the respective component-specific Study Treatment eCRF. The date of last administration of component of study treatment will also be referred to as the end of component of study treatment.

2.1.1.4 Date of first administration of study treatment

The date of first administration of study treatment is defined as the first date when a non-zero dose of any component of study treatment is administered and recorded on the respective component-specific Study Treatment eCRF. The date of first administration of study treatment will also be referred to as the start of study treatment. (Example: if 1st dose of alpelisib is administered on 05-April-2021, and 1st dose of olaparib is administered on 03-April-2021, then the date of first administration of study treatment is on 03-April-2021.)

2.1.1.5 Date of last administration of study treatment

The date of last administration of study treatment is defined as the last date when a non-zero dose of any component of study treatment was administered and recorded on the respective component-specific Study Treatment eCRF. (Example: if the last dose of alpelisib is administered on 22-Apr-2021, and the last dose of olaparib is administered on 24-Apr-2021, then the date of last administration of study treatment is on 24-Apr-2021.)

2.1.1.6 Study day

The study day, describes the day of the event or assessment date, relative to the reference start date.

The reference start date is designated as Study Day 1. Study Day –1 is the day that precedes Day 1. Study Day 0 is not defined. Study day is not to be used in numerical computations, for example in calculating exposure.

The study day will be calculated as:

- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference start date for **safety assessments** (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, PK etc.) is the start of study treatment.

The reference start date for **all other, non-safety assessments** (i.e., tumor assessment, death, disease progression, tumor response, ECOG performance status, and patient reported outcomes (PRO)) is the date of randomization. (Example: if randomization date is 15-DEC-2021, start of study treatment is on 18-DEC-2021, and the date of death is 28-DEC-2021, then the study day when the death occurred is 14).

The study day will be displayed in data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

2.1.1.7 Time unit

A year length is defined as 365.25 days. A month length is 30.4375 days (365.25/12). If duration is reported in months, duration in days will be divided by 30.4375. If duration is reported in years, duration in days will be divided by 365.25.

2.1.1.8 Baseline

For efficacy evaluations, the last available assessment, including unscheduled assessments, on or before the date of randomization will be used as the "baseline" value or "baseline" assessment. In the context of baseline definition, the efficacy evaluations also include PRO and ECOG performance status. For RECIST-based endpoints only, including PFS, ORR, CBR, time to response and duration of response, a window of 7 days after the start of study treatment will be allowed, i.e. the investigator/BIRC-reported responses will be considered as candidate for the baseline assessment if the assessment is within 7 days after the treatment start date.

For safety evaluations (e.g. laboratory assessments, ECGs and vital signs), the last available assessment, including unscheduled assessments, on or before the start of study treatment as described in Section 2.1.1.4, will be used as the "baseline" assessment. Assessments specified to be collected post-dose on the first date of treatment are not considered as baseline values.

If participants have no value as defined above, the baseline results will be considered missing.

2.1.1.9 On-treatment assessment/event

Safety summaries and selected summaries of deaths will summarize only on-treatment assessments/events.

An on-treatment assessment/event is defined as any assessment/event in the following time interval:

[date of first administration of study treatment, date of last administration of study treatment + 30 days], i.e. including the lower and upper limits.

Note: the calculation of study treatment duration will use different rules as specified in Section 2.4.1.1.

An AE starting in the screening phase and ongoing in the on-treatment phase will not be considered an on-treatment AE unless it has worsened in severity.

If the last date of study treatment is missing, any assessment/event occurring after the start of study treatment will be considered as on-treatment.

Data listings will include all assessments/events, flagging those which are not on-treatment assessment/events.

Note: The date of first administration of study treatment and the date of last administration of study treatment are defined in Sections 2.1.1.4 and 2.1.1.5, respectively.

2.1.1.10 Windows for multiple assessments

In order to summarize ECOG, laboratory, PRO, and other data collected over time (including unscheduled visits), the assessments will be time slotted. The following general rule will be applied in creating the assessment windows (except for ECOG, see Section 2.7.2 for ECOG-specific rules): If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the earlier of the 2 assessments will be used. If multiple assessments are taken on the same date then the worst case will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

Please refer to Table 2-9, Table 2-10 and Table 2-14 for the time windows for ECOG performance status, lab and PRO, respectively.

2.1.1.11 Last contact date

The last contact date will be derived for participants not known to have died at the analysis cut-off using the last complete date among the following:

Table 2-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No Condition
Last date participant was known to be alive from Survival Follow-up eCRF	Participant status is reported to be alive or unknown.
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term.
Start/End dates from each component-specific Study Treatment eCRF	Non-missing dose. Doses of 0 are allowed.
Date of treatment disposition or post-treatment follow-up disposition from disposition eCRF	No condition.
Date of PRO assessment	At least one non-missing answer to questionnaire
Tumor (RECIST) assessment date	For non-target lesion: non-missing lesion status For target lesion: non-missing lesion diameter For new lesion: “Is there a new lesion?” = yes
Laboratory collection dates	At least one non-missing parameter value
PK collection dates	Was sample taken? = Yes
Vital signs or ECG assessment date	At least one non-missing parameter value
ECOG Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

Source data	Conditions
Concomitant medication date	Non-missing medication
Biomarker collection date	At least one non-missing biomarker measurement
Hospitalization admission/discharge date	Non-missing verbatim term
Cardiac imaging date	Non-missing LVEF or overall interpretation

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date. The cut-off date will not be used for last contact date, unless the participant was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. For participants who died, the last contact date is defined as the date of death. Partial date imputation is allowed for event (death)/censoring if it is recorded on the 'Survival' eCRF.

The last contact date will be used for censoring of participants in the analysis of overall survival.

2.2 Analysis sets

The Full Analysis Set (FAS) is comprised of all participants to whom study treatment has been assigned by randomization. According to the intent to treat (ITT) principle, participants will be analyzed according to the treatment and strata they have been assigned to during the randomization procedure. The FAS will be the primary population for all efficacy analyses.

The Safety Set includes all participants who received at least one dose of study treatment (i.e. at least one dose of any component of alpelisib, olaparib, paclitaxel or PLD). Participants will be analyzed according to the study treatment received, where treatment received is defined as the randomized treatment if the participant took at least one dose of that treatment or the first treatment received if the randomized treatment was never received. The safety analysis set will be the population for all safety analyses.

The Pharmacokinetic Analysis Set (PAS) consists of all participants who receive at least one dose of alpelisib or olaparib and provide at least one evaluable PK concentration (see Section 2.8.2.1 for definition of evaluable PK concentration).

The Per-Protocol Set (PPS) includes the subset of the participants in the FAS without any major protocol deviation who took at least one dose of study treatment. Participants with any of the following protocol deviations will be excluded from the PPS.

- Written informed consent not obtained;
- BRCA mutation not determined (or assessed) using the FDA-approved BRCA Analysis CDx test
- Not having histologically confirmed high-grade serous or high-grade endometrioid cancer, fallopian tube cancer or primary peritoneal cancer

- Not having any measurable or non-measurable disease at baseline. If only non-measurable disease is present at baseline, not evaluable by GCIG
- No prior systemic treatment regimen at all or received more than three prior systemic treatment regimens
- Not having platinum-resistant or platinum refractory disease or participant has Primary Platinum Refractory Disease
- Received prior treatments with PI3K-, mTOR-, or AKT-Inhibitor
- Concurrently using other anti-cancer therapy
- With a known somatic BRCA mutation (sBRCAm)
- Having non-stable CNS involvement, or completed prior treatment for CNS mets <=28 days prior to randomization, or is receiving steroids and/or enzyme inducing anti-epileptic meds for brain mets
- Not having an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 at baseline
- Currently receiving strong inducers of CYP3A4 or Inhibitors of BCRP within 7d of treatment start
- Germline BRCA1/2 mutation or BRCA VUS, Genetic variant, favor polymorphism present and patient randomized

A sensitivity analysis of the primary endpoint (i.e., PFS per BIRC) will be performed using PPS if the primary efficacy analysis is significant and the FAS and PPS differ.

2.2.1 Participant classification

Participants may be excluded from the analysis sets defined above based on the protocol deviations entered in the database and/or on specific participant classification rules as defined in Table 2-2.

Table 2-2 Participant classification based on protocol deviations and non-protocol déviations criteria

Analysis set	Protocol deviations leading to exclusion	Non-protocol deviation criteria leading to exclusion
FAS	No written informed consent	NA
Safety set	No written informed consent	No dose of study treatment
PAS	No written informed consent	No dose of study treatment or no evaluable PK concentration
PPS	Any major protocol deviation as listed in definition of per rotocol set	NA

2.2.2 Withdrawal of informed consent

Any data collected in the clinical database after a participant withdraws informed consent from further participation in the trial, will not be included in the analysis. The date on which a participant withdraws full consent is recorded in the eCRF. When an end of treatment visit occurs after withdrawal of informed consent, the end of treatment disposition status and reasons are retained.

Death events may be used in the analysis of PFS/OS if captured from public records (registers), local law and participant informed consent permitting.

Additional data for which there is a separate informed consent, e.g. biomarker, collected in the clinical database without having obtained that consent will not be included in the analysis. These data will be excluded by the presence of the appropriate protocol deviation criterion.

2.2.3 Subgroup of interest

Subgroup analyses will be performed for efficacy and safety as outlined below. Subgroups will be formed using eCRF data (with the exception of biomarker subgroups, for which the applicable third party biomarker data will be used). This includes variables related to stratification factors (i.e. time to relapse from last platinum dose, prior PARP inhibitor use, and prior bevacizumab use), i.e., eCRF data will be used to define these subgroups. Analyses by stratum/stratification factor based on IRT data are covered by the analyses described in Section 2.5.

Geographic Region and race subgroups specified below will be defined as follows:
Geographic Region:

- Asia: China, Japan, Korea, Malaysia, Singapore, Taiwan
- Europe: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Italy, Netherlands, Portugal, Russia, Slovakia, Spain, UK
- Latin America: Argentina, Brazil, Mexico
- North America: Canada, United States
- Other: Australia, Turkey.

Race (as reported on demography CRF):

- White: race = White
- Asian: race = Asian
- Black or African American: race = Black or African American
- Other: Any other race except participants with missing race information, who will be excluded from the subgroup analyses.

Efficacy

The primary efficacy endpoint will be summarized by the following subgroups to examine the homogeneity of treatment effect provided that the primary efficacy analysis based on the FAS is statistically significant:

- Time to relapse from last platinum dose (stratification factor: < 3 months vs. ≥ 3 and ≤ 6 months)

- Prior PARP inhibitor use (stratification factor: Yes vs. No)
- Prior bevacizumab use (stratification factor: Yes vs. No)
- Presence of ascites at baseline [data from Diagnosis and Extent of Cancer eCRF, no information prior to the screening period should be included] (Yes vs. No)
- Baseline ECOG performance status (0 vs. ≥ 1)
- CA-125 at baseline (< 100 U/mL vs. ≥ 100 U/mL) (Lee et al. 2019)
- Age (< 65 years vs. ≥ 65 years)
- Race (White vs. Asian vs. Black or African American vs. Other)
- Number of prior regimens at baseline (1 or 2 vs. ≥ 3)
- Geographic Region (Europe, North America, Asia, Latin America and Other)
- Stage at baseline (I-III vs. IV)
- Longest diameter of the largest lesion at baseline (< 50mm vs. ≥ 50mm) (Lee et al. 2019)
- HRD status (positive vs. negative) [will be analyzed only if these data are available]

For each of the subgroups, the following analyses will be performed:

- Median Kaplan-Meier estimates of the survival distribution of PFS
- Hazard ratio with 95% CI using stratified Cox proportional hazards model. For subgroups based on time to relapse from last platinum dose, prior PARP inhibitor use, and prior bevacizumab use, unstratified analyses will be performed to avoid estimation issues, since these variables are related to the stratification factors.

Efficacy analyses in subgroups will be purely exploratory and are intended to explore the consistency of treatment effect. Forest plot (HR, 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

If OS is found to be statistically significant, additional descriptive subgroup analyses of OS will also be performed for the subgroups listed above.

Safety

The main safety analyses will be repeated on the safety set in the following subgroups:

- Stage at baseline (I-III vs. IV)
- Age (< 65 years vs. ≥ 65 years)
- Race (White vs Asian vs Black or African American vs Other)
- Geographic Region (Europe, North America, Asia, Latin America and Other)
- Number of prior regimens at baseline (1 or 2 vs. ≥ 3)
- Presence of ascites at baseline [data from Diagnosis and Extent of Cancer eCRF, no information prior to the screening period should be included] (Yes vs. No)

These main safety analyses include:

- AEs by system organ class, preferred term and maximum grade
- Treatment-related AEs by system organ class, preferred term and maximum grade

- Serious AEs by system organ class, preferred term and maximum grade
- For the AESI of Hyperglycemia only, a subgroup analysis by hyperglycemia diagnosis status at baseline per American Diabetes Association (ADA) 2017 will be presented:
    - Diabetic: FPG ≥ 7.0 mmol/l or 126 mg/dl or HbA1c ≥ 6.5 % vs.
    - Pre-diabetic: FPG 5.6- <7.0 mmol/l or 100-125 mg/dl or HbA1c 5.7- <6.5% vs.
    - Normal: [FPG <5.6 mmol/l or <100 mg/dl] and HbA1c <5.7%

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to a subgroup of participants, or safety issues that are more commonly observed in a subgroup of participants.

2.3 Participant disposition, demographics and other baseline characteristics

Demographic and other baseline data including disease characteristics/prognostic data will be summarized descriptively by treatment group for the FAS and Safety Set (for DSUR). Relevant medical history and current medical conditions at baseline will be summarized separately by system organ class, preferred term, and treatment group in the FAS.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented. For selected parameters, 25th and 75th percentiles will also be presented.

2.3.1 Participant disposition

Enrollment by country and center will be summarized for all screened participants and also by treatment group using the FAS. The number (%) of randomized and treated participants included in the FAS will be presented overall and by treatment group. The number (%) of screened but not-randomized participants and the reasons for screening failure will also be displayed. The number (%) of participants in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented overall and by treatment group.

The following summaries will be provided (with % based on the total number of FAS participants):

- Number (%) of participants who were randomized (based on data from IRT system)
- Number (%) of participants who were randomized but not treated (based on ‘Study Treatment’ eCRF not completed for any study treatment component)
- Primary reason for not being treated (based on ‘Treatment disposition’ eCRF)
- Number (%) of participants who were treated (based on ‘Study Treatment’ eCRF of each study treatment component entered with non-zero dose administered)
- Number (%) of participants who are still on-treatment (based on the ‘Treatment disposition’ eCRF not completed);
- Number (%) of participants who discontinued the study treatment phase (based on the ‘Treatment disposition’ eCRF)

- Primary reason for study treatment phase discontinuation (based on the 'Treatment disposition' eCRF)
- Number (%) of participants who have entered the post-treatment follow-up (based on the 'Treatment disposition' page);
- Number (%) of participants who have discontinued from the post-treatment follow-up (based on the 'Post-treatment follow-up disposition' eCRF);
- Reasons for discontinuation from the post-treatment follow-up (based on 'Post-treatment follow-up disposition' eCRF);
- Number (%) of participants who have entered the survival follow-up (based on the 'Treatment disposition' or 'Post-treatment follow-up disposition' eCRF).

Protocol deviations

The number and percentage of participants in the FAS with any protocol deviations will be tabulated by deviation category (as specified in the Edit Check Specification Document) and treatment group. Major protocol deviations leading to exclusion from analysis sets will be tabulated separately overall and by treatment group. All protocol deviations will be listed.

Analysis sets

The number and percentage of participants in each analysis set will be summarized by treatment group and randomization stratum.

2.3.2 Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment group. Categorical data (e.g. race, ECOG performance status, etc.) will be summarized by frequency counts and percentages; the number and percentage of participants with missing data will be provided. Continuous data (e.g. age, weight, etc.) will be summarized using descriptive statistics (N, mean, standard deviation, median, minimum, and maximum) by treatment group. BMI (kg/m^2) will be calculated as weight[kg] / (height[m])^2 using weight at baseline.

2.3.3 Baseline stratification factors

The number (%) of participants in each stratum based on data obtained from the IRT system will be summarized overall and by treatment group for the FAS. Discordances between the stratum recorded in IRT at the time of randomization and the actual stratum recorded in the clinical database through the data collected on eCRF will be cross-tabulated and listed.

Unless otherwise specified, stratified analyses and analyses "by stratum" will be based on IRT stratification data, while data for other subgroup analyses will be based on eCRF data.

In particular, for the stratification factor "Time to relapse from last platinum dose" based on eCRF data, the derivation instructions below should be followed.

- Complete dates of the last platinum dose and the associated progression will be used in the derivation if they are available.
- If progression date or last platinum dose date is partially missing or completely missing, no imputation for them and consider it as Unknown for the stratification factor "Time to relapse from last platinum dose" based on eCRF data.

2.3.4 Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer by treatment group. This analysis will include the following: primary site of cancer, stage at initial diagnosis, time since initial diagnosis, stage at time of study entry, presence/absence of target and non-target lesions, number and type of metastatic sites involved, baseline CA-125, presence of ascites at baseline (data from Diagnosis and Extent of Cancer eCRF, no information prior to the screening period should be included).

The presence/absence of target and non-target lesions will be based on the data collected on RECIST target/non-target lesion assessment eCRFs. Metastatic sites will be based on Diagnosis and Extent of Cancer eCRF. Time since initial diagnosis will be summarized in months. A month is defined as 365.25/12=30.4375 days.

2.3.5 Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on the eCRF, will be summarized and listed by treatment group. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by system organ class (SOC) and preferred term (PT). Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

2.3.6 Other

All data collected at baseline, including participants' referrals, child bearing potential and biomarker informed consent, will be listed.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Duration of study treatment exposure, cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by treatment group. The number of participants with dose reductions/interruptions and the reasons, will be summarized and listed. Details of the derivations and summaries are provided in the following sections. Participants with no exposure to the study treatment component will be excluded from the corresponding tabular summaries.

Categorical data will be summarized as frequencies and percentages. For continuous data, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum will be presented.

The safety set will be used for all summaries and listings of study treatment.

2.4.1.1 Duration of study treatment exposure

The duration of exposure to study treatment will be calculated as:

Duration of exposure to study treatment (days) = (**last date of exposure** to any study treatment component) – (date of first administration of study treatment) + 1.

Duration of exposure to any single component of study treatment will be calculated as:

Duration of exposure (days) = (last date of exposure to study treatment component) – (date of first administration of study treatment component) + 1.

The **last date of exposure** is defined as follows for the study treatment components:

- For alpelisib or olaparib, the last date of exposure is defined as the date of last administration of alpelisib or olaparib;
- For paclitaxel, the last date of exposure is defined as the date of last administration of paclitaxel + 6 days (given that paclitaxel is administered on a weekly basis).
  - If a participant died or was lost to follow-up within date of last administration + 6 days, then the last date of exposure is the date of death or the date of last contact, respectively.
  - If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.
- For PLD, the last date of exposure is defined as the date of last administration of PLD + 27 days (given that PLD is administered every 4 weeks).
  - If a participant died or was lost to follow-up within date of last administration + 27 days, then the last date of exposure is the date of death or the date of last contact, respectively.
  - If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.

This duration of exposure includes the periods of temporary interruption (of any component of the study treatment for any reason). The duration of study treatment exposure and exposure to each treatment component will be summarized by treatment group. In addition, the duration of exposure will be categorized into time intervals (e.g. <1 month; <3 months; 3-<6 months; 6-<9 months, etc.); frequency counts and percentages will be presented for the number of participants in each interval.

2.4.1.2 Cumulative dose

Cumulative dose for any component of study treatment is defined as the total dose of the medication given during the study treatment exposure.

Cumulative dose will be summarized using descriptive statistics for each component of study treatment by treatment group.

For alpelisib or olaparib, the cumulative dose is the sum of the non-zero doses recorded over the dosing period.

For paclitaxel or PLD, the cumulative dose should be defined based on the days when the participant is assumed to have taken a non-zero dose during dosing periods.

2.4.1.3 Dose intensity and relative dose intensity

Dose intensity (DI) for participants with non-zero duration of exposure to each study treatment component is defined as follows:

For alpelisib or olaparib, DI is defined as

DI (mg/day) = Cumulative dose (mg) / duration of exposure to alpelisib or olaparib (days).

For paclitaxel, DI is defined as

DI (mg/m^2/7 days) = Cumulative dose (mg/m^2) / {[(date of last administration of study treatment component + 6) – (date of first administration of study treatment component) + 1] / 7},

to account for the paclitaxel dosing schedule in the DI computation.

For PLD, DI is defined as

DI (mg/m^2/28 days) = Cumulative dose (mg/m^2) / {[(date of last administration of study treatment component + 27) – (date of first administration of study treatment component) + 1] / 28},

to account for the PLD dosing schedule in the DI computation.

For participants who did not take any drug, the DI is equal to zero. Planned dose intensity (PDI) is defined as the assigned dose by unit of time planned to be given to participants as per protocol in the same dose unit and unit of time as that of the Dose Intensity. The PDI for each study treatment component is displayed in Table 2-3. Note that DI will also be calculated and reported in the units displayed in Table 2-3, whereas duration of exposure itself will be summarized in months.

Table 2-3 Planned dose intensity

Medication	PDI (dose unit/unit of time)
Alpelisib	200 mg/day
Olaparib	400 mg/day
Paclitaxel	80 mg/m^2 / 7 days
PLD	40 or 50 mg/m^2 / 28 days

Relative dose intensity (RDI) is defined as:

RDI (%) = [DI (dosing unit / unit of time) / PDI (dosing unit / unit of time)] x 100.

DI and RDI will be summarized separately for each of the study treatment components.

2.4.1.4 Dose reductions, interruptions or permanent discontinuations

The number and percentage of participants with dose reductions, interruptions or permanent discontinuations, and associated reasons, will be summarized by treatment group as outlined below.

Dose administered (mg) and dosing frequency from the component-specific Study Treatment eCRF will be used to determine the dose reductions and interruptions.

'Dose permanently discontinued' ticked box from the component-specific Study Treatment eCRF will be used to determine permanent discontinuation.

Interruption: An interruption is defined as a 0 mg dose given on one or more days between two non-zero dosing records. Any two or more consecutive zero doses of alpelisib (e.g. in the sequence 200 mg daily to 0 mg to 0 mg to 200 mg daily) or olaparib or paclitaxel / PLD will be counted as one interruption if the reasons for these two consecutive dose interruptions are the same. It will be counted as two different interruptions only if the reasons are different. For participants who have dose interruption checked but never resume non-zero dose, the dose interruption will not be counted. For example, in the sequence of 200 mg daily to 0 mg (dose interruption) to 0 mg (dose permanently discontinued), the 0 mg (dose interruption) will not be counted as a dose interruption. Interruptions will be summarized for each component of study treatment.

Reduction: A dose reduction for alpelisib or olaparib is defined as a decrease from the previous non-zero dose to another non-zero dose less than the protocol-planned dose, even if this decrease has been directly preceded by an interruption. For example, in the sequence 200 mg daily to 0 mg to 150 mg daily, the 150 mg dose will be counted as a reduction. On the other hand, if the dose decrease is followed by an interruption, with the dose resuming at the same level prior to the interruption (e.g. in the sequence 150 mg daily to 0 mg to 150 mg daily), the second dose decrease will not be counted as dose reduction.

If, due to a dosing error, a participant receives a higher than planned starting dose and moves down to the planned starting dose then this is not considered a dose reduction. However if the dose change is from a higher than planned starting dose down to a lower than protocol planned starting dose, then this is considered a dose reduction (e.g. in the sequence: 250 mg daily to 200 mg daily to 150 mg daily, 150 mg is considered as a dose reduction assuming that 200 mg daily is the protocol planned starting dose).

If, due to a dosing error, a participant receives a lower than previous non-zero dose and resumes later at the protocol specified dose reduction, then the lower dose received due to dosing error and protocol specified dose reduction are dose reductions (e.g. in the sequence: 200 mg daily to 100 mg daily to 150 mg daily, then 100 mg and 150 mg are considered as dose reductions assuming that 200 mg daily is the protocol planned starting dose).

If a participant receives a protocol specified reduced dose and receives later at a lower than previous non-zero dose due to a dosing error, then 2 dose reductions will be counted (e.g. in the sequence: 200 mg daily to 150 mg daily to 100 mg daily, 150 mg and 100 mg are dose reductions assuming that 200 mg daily is the protocol planned starting dose).

For olaparib specifically, if a participant misses one single dose in a BID frequency due to a dosing error, then that missed dosing will be counted as a dose reduction (e.g. in the sequence: 200 mg BID to 150 mg BID to 150 mg daily [due to dosing error] to 150 mg BID, both 150 mg BID and 150 mg daily are dose reductions assuming that 200 mg BID is the protocol planned starting dose).

Table 2-4 Examples of Dose Reduction for alpelisib (assuming protocol-planned starting dose of 200 mg daily)

Sequence	Reduction
With dose change	
200 mg daily to 150 mg daily to 0 mg to 150 mg daily	1 reduction (the 1st 150 mg)
200 mg daily to 200 mg daily to 0 mg to 150 mg daily	1 reduction (150 mg)
200 mg daily to 0 mg to 150 mg daily	1 reduction (150 mg)
With interruption	
200 mg daily to 0 mg to 200 mg daily	0 reductions
With dosing error	
200 mg daily to 150 mg daily to 100 mg daily*	2 reductions (150 mg, 100 mg)
200 mg daily to 100 mg daily* to 150 mg daily	2 reductions (100 mg, 150 mg)
200 mg daily to 400 mg daily* to 200 mg daily	0 reductions since 400 mg is dose escalation not reduction
200 mg daily to 150 mg daily* to 200 mg daily	1 reduction (150 mg)
With dosing error at the 1st administration	
150 mg daily* to 200 mg daily	1 reduction (150 mg)
150 mg daily* to 0 mg to 150 mg daily* to 200 mg daily	1 reduction (150 mg)
50 mg daily* to 200 mg daily to 0 mg to 150 mg daily	2 reductions (50 mg and 150 mg)

*dosing error

Table 2-5 Examples of Dose Reduction for olaparib (assuming protocol-planned starting dose of 200 mg BID)

Sequence	Reduction
With dose change	
200 mg BID to 150 mg BID to 0 mg to 150 mg BID	1 reduction (the 1st 150 mg BID)

Sequence	Reduction
200 mg BID to 200 mg BID to 0 mg to 150 mg BID	1 reduction (150 mg BID)
200 mg BID to 0 mg to 150 mg BID	1 reduction (150 mg BID)
With interruption	
200 mg BID to 0 mg to 200 mg BID	0 reductions
With dosing error	
200 mg BID to 150 mg BID to 150 mg daily*	2 reductions (150 mg BID, 150 mg daily)
200 mg BID to 200 mg daily* to 150 mg BID	2 reductions (200 mg daily, 150 mg BID)
200 mg BID to (300 mg + 200 mg) daily* to 200 mg BID	0 reductions since (300 mg + 200 mg) daily is dose escalation not reduction
200 mg BID to 150 mg BID^ to 200 mg BID	1 reduction (150 mg BID)
With dosing error at the 1st administration	
200 mg daily* to 200 mg BID	1 reduction (200 mg daily)
200 mg daily* to 0 mg to 200 mg daily* to 200 mg BID	1 reduction (200 mg daily)
200 mg daily* to 200 mg BID to 0 mg to 150 mg BID	2 reductions (200 mg daily and 150 mg BID)

* dosing error

^ dispensing error

Missing data: If dose is recorded but frequency is missing or entered as ‘none’, it is assumed that the study drug was taken as per-protocol.

In this study, dose reductions for paclitaxel or PLD will be at investigator's discretion and in accordance with institutional practice and local labeling. Therefore, it will be difficult to summarize them. However, all dosing data will be listed.

In addition, the reasons for permanent discontinuation of each study treatment component will be summarized by treatment group based on the information on the respective component-specific Study Treatment eCRF.

2.4.2 Prior, concomitant and post-treatment therapies

Prior anti-cancer therapy

The number and percentage of participants who received any prior anti-neoplastic medications, radiotherapy or surgery (biopsy and non-biopsy separately) will be summarized by treatment group both separately and in a combined fashion.

- Prior anti-neoplastic medications will be summarized by therapy type (e.g. chemotherapy, hormonal therapy etc.), and also by lowest ATC class, preferred term and treatment group. The total number of regimens along with the type (e.g. hormonal therapy), best response and time from treatment end date of the last therapy to progression will be summarized by treatment group. The medication therapy type of any combination therapy will be classified based on the following order: chemotherapy, biologic therapy, targeted therapy, hormonal therapy. For example, a combination therapy of chemotherapy and targeted therapy will be classified as 'chemotherapy'. This classification will be based on medical review of observed trial data and the excel spreadsheet created will be stored in GPS.
- For radiotherapy, the location for the last therapy will be summarized.
- For surgery (excluding biopsies), procedure at last surgery and the time since last surgery will be summarized.

Separate listings will be produced for prior anti-neoplastic medications, radiotherapy, and surgery.

The above analyses will be performed using the FAS.

Concomitant therapy

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment that was administered to a participant, coinciding with the study treatment period. Concomitant therapy includes medications (other than study treatment) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system.

Concomitant medications will be summarized by lowest ATC class, preferred term and treatment group. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term.

These summaries will include:

1. medications starting on or after the start of study treatment but no later than 30 days after the last dose of study treatment, and
2. medications starting prior to the start of study treatment and continuing after the start of study treatment.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 30 days after the last date of study treatment will be flagged in the listing.

The safety set will be used for all concomitant medication tables and listings.

Post-treatment anti-cancer therapy

Anti-neoplastic therapies after discontinuation of study treatment will be listed and tabulated by ATC class, preferred term and treatment group; by means of frequency counts and percentages using the FAS.

2.5 Analysis of the primary objective

The primary objective of the study is to determine whether treatment with alpelisib in combination with olaparib prolongs PFS compared to treatment with single agent cytotoxic chemotherapy (paclitaxel or PLD) in participants with platinum resistant or refractory HGSOC with no germline BRCA mutation detected.

2.5.1 Primary endpoint / estimand

The primary endpoint (variable attribute of the primary estimand; refer to Section 1.2.1) of the study is PFS, defined as the time from the date of randomization to the date of the first documented progression or death due to any cause. For the primary efficacy analysis, PFS will be based on BIRC tumor assessments and using RECIST 1.1 criteria (see Section 16.3 of the study protocol). The primary analysis will be based on the FAS and will include all data observed upto the cut-off date. If a participant has not progressed or died at the analysis cut-off date, PFS will be censored at the date of the last adequate tumor assessment before the cut-off date. PFS events (i.e. RECIST 1.1. documented disease progression as per BIRC or death) documented after the initiation of new anti-neoplastic therapy will be considered for the primary analysis provided tumor assessments continue after initiation of new cancer therapy (see Section 2.5.4 for additional details regarding censoring rules and determination of date of last adequate tumor assessment). Discontinuation due to disease progression (collected on the 'End of treatment' and 'End of post treatment follow up' disposition pages) without supporting objective evidence satisfying progression criteria per RECIST 1.1 will not be considered as disease progression for PFS derivation. Clinical deterioration without objective radiological evidence will not be considered as documented disease progression.

The tumor endpoint derivation is based on the sequence of overall lesion responses at each assessment/time point. However, the overall lesion response at a given assessment/time point may be provided from different sources as illustrated in Table 2-6.

Table 2-6 Sources for overall lesion response

Source 1	Investigator (local radiology) reported overall lesion response
Source 2	Novartis-calculated overall lesion response based on raw (i.e. individual lesion) measurements from investigator (local radiology)
Source 3	Final BIRC-reported overall lesion response
Source 4	Novartis-calculated overall lesion response based on raw (i.e. individual lesion) measurements from BIRC

The primary efficacy analysis will be based on the BIRC assessment. The BIRC comprises of two independent radiologists and an adjudicator, if applicable; as well as an oncologist. The BIRC-reported overall lesion response at each assessment/time point (Source 3 in Table 2-6) will be used to derive the efficacy endpoints. Data from independent readers will be listed together with the data from adjudicator and oncologist.

The overall response at each assessment will also be calculated using the raw lesion measurements (Source 2 and Source 4 in Table 2-6). The calculated responses will be listed along with the responses given by BIRC and investigator. PFS based on calculated overall lesion response (Source 2 in Table 2-6) will also be summarized.

Tumor assessment data based on investigator/local radiology review (Source 1) will be used for sensitivity efficacy analyses. The investigator-reported overall lesion response data will be used to derive the investigator-based endpoints. Differences in overall responses between local radiology (Source 1) and central radiology (Source 3) will be listed.

2.5.2 Statistical hypothesis, model, and method of analysis

The primary efficacy analysis will be the comparison of PFS between the two treatment groups using a stratified log-rank test at an overall one-sided 2.5% level of significance.

Assuming proportional hazards model for PFS, the following statistical hypotheses will be tested to address the primary efficacy objective:

$H_{01}: \theta_1 \geq 0$ vs. $H_{a1}: \theta_1 < 0$

where θ_1 is the log-hazard ratio (alpelisib + olaparib arm vs. paclitaxel or PLD arm) of PFS.

The primary efficacy endpoint of PFS (variable attribute of the primary estimand; refer to Section 1.2.1) will be analyzed at the interim futility look and final analysis of a group sequential design based on the data observed in the FAS up to the cut-off date, according to the treatment group participants were randomized to and the strata they were assigned to at randomization (strata formed using the randomization factor as obtained via IRT). Refer to Section 2.10 for more details on interim analysis. PFS will be estimated using the Kaplan-Meier method. The results will be plotted graphically by treatment group. The median PFS along with 95% confidence intervals will be presented by treatment group. A stratified Cox regression model will be used to estimate the hazard ratio of PFS, along with 95% confidence interval using the same strata information as the primary efficacy comparison.

2.5.3 Handling of intercurrent events of primary estimand

The primary estimand will account for different intercurrent events as explained in the following:

- **Discontinuation of study treatment**: tumor assessment data collected after discontinuation of study treatment will be used for the primary analysis irrespective of the study treatment discontinuation reason (treatment policy strategy).

2.5.4 Handling of missing values/censoring/discontinuations

This is an event-driven trial and the final analysis for PFS will be performed after approximately 224 PFS events have been documented based on BIRC review of tumor assessments. An analysis cut-off date will be established after approximately 224 PFS events have been documented.

In the primary analysis, PFS will be censored at the date of the last adequate tumor assessment if no PFS event is observed prior to the analysis cut-off date.

PFS events (i.e. RECIST 1.1. documented disease progression as per BIRC or death) documented after the initiation of new anti-neoplastic therapy will be considered for the primary analysis provided tumor assessments continue after initiation of new cancer therapy.

If a PFS event is observed after one or more missing or non-adequate tumor assessments, the actual date of event will be used (see RECIST 1.1 in Section 16.3 of the study protocol).

The term "missing adequate tumor assessment" is defined as a tumor assessment (TA) not performed or tumor assessment with overall lesion response of "NE".

The date of last adequate tumor assessment is the date of the last tumor assessment with overall lesion response of CR, PR or SD or non-CR/non-PD before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment will be used. If no post-baseline assessments are available (before an event or a censoring reason occurred) then the date of randomization will be used.

Refer to Table 2-7 for censoring and event date options and outcomes for PFS.

Table 2-7 Outcome and event/censor dates for PFS analysis

Situation	Date	Outcome
No baseline assessment	Date of randomization	Censored
Progression or death at or before next scheduled Assessment	Date of progression (or death)	Progressed
Progression or death after one or more missing assessments	Date of progression (or death)	Progressed
No progression (or death)	Date of last adequate assessment	Censored
Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	Ignore the treatment discontinuation and follow situations above	As per above situations
New anticancer therapy given prior to protocol defined progression	Ignore the new anticancer therapy and follow situations above	As per above situations
Death before first PD assessment	Date of death	Progressed

2.5.5 Sensitivity and supplementary analyses

Sensitivity analyses

If there is a high rate of discrepancy (at least 5%) between the strata classifications constructed using the eCRF data and those obtained from the IRT, a sensitivity analysis will be performed

in which a stratified Cox regression model will be used to estimate the treatment hazard ratio and the associated 95% confidence interval based on the eCRF-derived strata. No other inferential statistics will be provided.

In addition, to assess the impact of stratification, the hazard ratio and 95% confidence interval for PFS per BIRC assessment will also be obtained from an unstratified and covariate unadjusted Cox model.

Furthermore, PFS per local review will be analyzed using a stratified Cox model, with the same analysis conventions as the primary efficacy analysis, and the treatment effect will be summarized by the hazard ratio with its 95% confidence interval. Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group.

Supplementary analyses

As supplementary analyses, the primary efficacy endpoint will be summarized based on the data obtained below to assess different censoring mechanisms. Kaplan-Meier estimates of median PFS along with 95% confidence intervals, and hazard ratio and corresponding 95% confidence interval obtained using the Cox proportional hazards model will be provided. No other inferential statistics will be provided.

- using the primary analysis source (i.e., BIRC assessment) on the FAS and censoring PFS at the date of the last adequate tumor assessment if PFS event is observed after two or more missing tumor assessments (TAs). The rule to determine number of missing TAs is based on the time interval between the date of last adequate tumor assessment and the date of an event (see Section 5.7.5). If the interval is greater than twice the protocol-specified interval between the TAs and 2 times the protocol-allowed time window around assessments, then the number of missing assessments will be 2 or more. In the summary tables, this approach is referred as ‘missing TA PFS supplementary analysis’.
- using the primary analysis source (i.e., BIRC assessment) on the FAS where events occurring after one or more missing TA are set equal to the date of the next scheduled assessment after the last adequate TA. In the summary tables, this approach is referred as ‘backdating PFS supplementary analysis’.
- using the primary analysis source (i.e., BIRC assessment) on the FAS and censoring PFS at the date of the last adequate tumor assessment before the start of new anticancer therapy if no PFS event is observed prior to the start of new antineoplastic therapy. In the summary tables, this approach is referred as ‘new anticancer therapy leading to PFS censoring supplementary analysis’.
- using the primary analysis source (i.e., BIRC assessment) on the FAS and considering initiation of new antineoplastic therapy or RECIST PD or death as a PFS event. In the summary tables, this approach is referred as ‘new anticancer therapy leading to PFS event supplementary analysis’.
- using the primary analysis source (i.e., BIRC assessment) on the FAS and considering either RECIST PD or clinical PD (as defined in the protocol) or death as a PFS event. In the summary tables, this approach is referred as ‘BIRC RECIST or clinical PD supplementary analysis’.

- using data source from local investigator assessment on the FAS and considering either RECIST PD or clinical PD (as defined in the protocol) or death as a PFS event. In the summary tables, this approach is referred as 'local RECIST or clinical PD supplementary analysis'.

Further supplementary analyses will include:

- A multivariate Cox regression model stratified by randomization stratification factors (per IRT) will be fitted to evaluate the effect of other baseline demographic and disease characteristics on the estimated hazard ratio. The fitted model adjusting the treatment difference for key baseline and prognostic factors will include as covariates the following: presence of ascites at baseline only (data from Diagnosis and Extent of Cancer eCRF, no information prior to the screening period should be included) (yes vs. no), CA-125 at baseline (< 100 U/mL vs. ≥ 100 U/mL), baseline ECOG performance status (0 vs. ≥ 1), longest diameter of the largest lesion at baseline (< 50mm vs. ≥ 50mm), stage at baseline (I-III vs. IV). All covariates will be included in the model regardless of their observed significance (p-value for given covariate).
- Repeating the primary efficacy analysis using Novartis-derived overall lesion response per local radiology assessment.

Additionally, the following analyses will be performed:

- Timing of all tumor assessments will be depicted graphically separately for central radiology and investigator/local radiology and displayed by treatment group.
- Comparison of PFS event type/censor between local radiology review and BIRC review
- Number of participants with a PFS event and number of participants censored for the PFS analysis will be summarized. In addition, a summary of reasons for PFS censoring will be provided by treatment group based on the reasons defined below. These summaries on censoring reasons will be produced for PFS by investigator radiology and BIRC review.

Censoring pattern of PFS

Number of participants with a PFS event and number of participants censored for the PFS analysis will be summarized. In addition, a summary of reasons for PFS censoring will be provided by treatment group based on the following reasons:

1: Ongoing without event
2: Lost to follow-up
3: Withdrew consent
4: Adequate assessment no longer available

The PFS censoring reasons are defined in the following way.

If the time interval between the last adequate TA date and the earliest of the following dates is smaller or equal to interval of 2 missing tumor assessments:

1. Analysis cut-off date,
2. Date of consent withdrawal,

3. Visit date of study treatment discontinuation or end of post-treatment follow-up discontinuation due to lost to follow-up.

Then the PFS censoring reason will be:

1. 'Ongoing',
2. 'Withdrew consent',
3. 'Lost to follow-up',

If the time interval is larger than the interval of 2 missing tumor assessments with no event observed, then the PFS censoring reason will always default to 'Adequate assessment no longer available'.

These summaries on censoring reasons will be produced for PFS by investigator and BIRC assessment. The censoring patterns will be compared between treatment groups within each of the two comparisons and also between investigator and BIRC assessment. Summary of the difference in days to progression as per BIRC and investigator assessment will also be generated.

Subgroup analyses for the primary endpoint

The primary endpoint of PFS will be summarized for the subgroups of stratification factors (per eCRF) using the unstratified analysis. If the primary efficacy analysis is statistically significant, additional subgroup analyses as specified in Section 2.2.3 will be performed using the same conventions as for the primary analysis.

For each of the additional subgroups, the following analyses will be performed:

- Kaplan-Meier estimates of median and its corresponding 95% CI of the survival distribution of PFS
- Hazard ratio with 95% CI using stratified (per IRT) Cox proportional hazards model.

Efficacy analyses in subgroups are intended to explore the consistency (homogeneity) of treatment effect. Forest plot (including sample size/number of PFS events and HR with 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

2.6 Analysis of the key secondary objective

The key secondary objective is to determine whether treatment with alpelisib in combination with olaparib prolongs OS compared with single agent cytotoxic chemotherapy (paclitaxel or PLD).

2.6.1 Key secondary endpoint / estimand

OS is identified as the key secondary endpoint (variable attribute of the key secondary estimand). It is defined as the time from date of randomization to date of death due to any cause. All deaths occurring on or before the cut-off date in the FAS will be used in the OS analysis. If a participant is not known to have died at the time of analysis cut-off, OS will be censored at the date of last contact.

A hierarchical testing strategy will be used to control the overall type I error rate, where OS will only be formally tested and interpreted if the primary analysis of PFS is statistically significant.

2.6.2 Statistical hypothesis, model, and method of analysis

Assuming proportional hazards for OS, the following statistical hypotheses will be tested only if PFS is statistically significant:

$$H_{02}: \theta_2 \geq 0 \quad \text{vs.} \quad H_{A2}: \theta_2 < 0$$

where θ_2 is the log-hazard ratio (alpelisib + olaparib arm vs. paclitaxel or PLD arm) of OS. The analysis to test this hypothesis will consist of a stratified log-rank test at an overall one-sided 2.5% level of significance.

OS will be analyzed using a group sequential design using a Lan-Demets (O'Brien-Fleming) alpha spending function (Lan and DeMets 1983), which is independent of the one used for PFS. Refer to Section 2.10 for more details on interim analysis. Analyses will be based on the FAS according to the randomized treatment group and strata assigned at randomization. All deaths recorded up to the cut-off date will be included in the analysis. The survival distribution of OS will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves, medians and 95% confidence intervals of the medians along with the proportion of participants alive at selected time points and the associated 95% confidence intervals will be presented for each treatment group. The hazard ratio for OS will be calculated, along with its 95% confidence interval, using a stratified Cox model with the same stratification factors as the log-rank test.

OS will be tested hierarchically as follows:

1. If the final PFS analysis is statistically significant:

   The time point for the first OS interim analysis will be at the time of final PFS analysis after approximately 50% of deaths (126 deaths) are expected to have been recorded in the clinical database. If PFS is statistically significant at this stage, OS will also be tested.

   If OS is not statistically significant at the first interim analysis, the second OS analysis will be planned after approximately 75% of deaths (189 deaths) have been recorded in the clinical database. If OS is not statistically significant at this stage, a final analysis is planned at the time approximately 252 deaths have been recorded.

2. If the final PFS analysis is not statistically significant, then OS will not be tested.

2.6.3 Handling of intercurrent events of key secondary estimand

The analysis of key secondary estimand will account for different intercurrent events as explained in the following:

- **Discontinuation of study treatment**: OS will take into account all deaths irrespective of the study treatment discontinuation reasons (treatment policy strategy).

2.6.4 Handling of missing values/censoring/discontinuations

If a participant is not known to have died at the time of analysis cut-off, then OS will be censored at the last known date participant was alive, i.e., last contact date (see Section 2.1.1.11).

For rare cases when either day is missing or both month and day are missing for the date of death, the follow imputation rules will be implemented:

- If only day is missing, then impute max [(1 mmm-yyyy), min (last contact date +1, cutoff date)].
- If both day and month are missing, then impute max [(1 Jan-yyyy, min (last contact date +1, cutoff date)].

2.6.5 Sensitivity and supplementary analyses

Sensitivity analyses

If there is a high rate of discrepancy (at least 5%) between the strata classifications constructed using the eCRF data and those obtained from the IRT, a sensitivity analysis will be performed in which a stratified Cox regression model will be used to estimate the treatment hazard ratio and the associated 95% confidence interval based on the eCRF-derived strata. No other inferential statistics will be provided.

In addition, to assess the impact of stratification, the hazard ratio and 95% confidence interval for OS will also be obtained from an unstratified and covariate unadjusted Cox model.

Supplementary analyses

A multivariate Cox regression model stratified by randomization stratification factors (per IRT) will be fitted to evaluate the effect of other baseline demographic and disease characteristics on the estimated hazard ratio. The fitted model adjusting the treatment difference for key baseline and prognostic factors will include as covariates the following: presence of ascites at baseline only (data from Diagnosis and Extent of Cancer eCRF, no information prior to the screening period should be included) (yes vs. no), CA-125 at baseline (< 100 U/mL vs. ≥ 100 U/mL), baseline ECOG performance status (0 vs. ≥ 1), longest diameter of the largest lesion at baseline (< 50mm vs. ≥ 50mm), stage at baseline (I-III vs. IV). All covariates will be included in the model regardless of their observed significance (p-value for given covariate).

The key secondary endpoint of OS will be summarized for the subgroups of stratification factors (per eCRF) using the unstratified analysis. If the key secondary efficacy analysis of OS is statistically significant, additional subgroup analyses as specified in Section 2.2.3 will be performed using the same conventions as for the primary analysis.

For each of the additional subgroups, the following analyses will be performed:

- Kaplan-Meier estimates of median and its corresponding 95% CI of the survival distribution of OS
- Hazard ratio with 95% CI using stratified (per IRT) Cox proportional hazards model.

A forest plot (sample size/number of deaths, HR, 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

The pattern of censored data will be examined between the treatment groups: reasons for censoring (‘Alive’ or ‘Lost to follow-up’) and death cause will be summarized by treatment

group. In addition, survival status, reason for censoring and death cause will be listed. Participants not known to have died will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than the protocol specified interval between the survival follow-up assessments plus 2 weeks, i.e., 12 + 2 weeks = 14 weeks for this study (i.e., the planned interval between two OS follow-up visits plus the 1-week window on either side).

Additional sensitivity/supplementary analysis for OS may be considered in the FAS population if deemed necessary, e.g., excludes deaths due to COVID-19, and OS will be censored on the date of death due to COVID-19.

2.7 Analysis of secondary efficacy objective(s)

The other secondary efficacy objectives are to:

- Evaluate the two treatment groups with respect to overall response rate and clinical benefit rate
- Describe time to response and duration of response in each treatment group
- Evaluate the two treatment groups with respect to time to definitive deterioration of ECOG performance status.

2.7.1 Secondary endpoints

Overall response rate (ORR)

ORR with confirmed response is defined as the proportion of participants with best overall response (BOR) of confirmed complete response (CR) or confirmed partial response (PR), as per BIRC assessment and according to RECIST 1.1 (see Section 16.3 of the study protocol). ORR with confirmed response will be calculated based on the FAS. Participants with only non-measurable disease at baseline will be included in the numerator if they achieve a complete response. Only tumor assessments performed before the start of any further anti-neoplastic therapies (i.e. any additional anti-neoplastic therapy or surgery) and within 30 days after the last administration of study treatment will be considered in the assessment of best overall response. Further anti-neoplastic therapies will be identified from the data collected on 'Anti-neoplastic therapies since discontinuation of study treatment' CRFs. Palliative radiotherapy is the only setting of radiotherapy allowed during the study. Therefore, palliative radiotherapy will not be considered as an anti-neoplastic therapy for assessment of BOR. Continuation of combination partner therapy alone after end of study treatment will also not be considered as a new anti-neoplastic therapy.

Clinical benefit rate (CBR)

CBR with confirmed response is defined as the proportion of participants with a best overall response of confirmed complete response (CR) or confirmed partial response (PR), or an overall response of stable disease (SD) lasting for at least 24 weeks. CR, PR, and SD are defined as per BIRC assessment according to RECIST 1.1 (see Section 16.3 of the study protocol). A participant will be considered to have SD for 24 weeks or longer if a SD response is recorded at 24-1=23 weeks or later from randomization, allowing for the ±1 week visit window for tumor assessments. Participants with only non-measurable disease at baseline will be included in the

numerator only if they achieve a complete response or have a 'Non-CR/Non-PD' response at 24-1=23 weeks or later from randomization. CBR will be calculated using the FAS based on BIRC.

Time to response (TTR)

TTR (CR or PR) is defined as the time from the date of randomization to the first documented response of CR or PR, which must be subsequently confirmed (although date of initial response is used, not date of confirmation), using BIRC data and according to RECIST 1.1 (see Section 16.3 of the study protocol).

All participants in the FAS will be included in TTR calculations. Participants without a confirmed CR or PR will be censored at the study-maximum follow-up time (i.e. LPLV-FPFV) for participants with a PFS event (i.e. disease progression or death due to any cause), or at the date of the last adequate tumor assessment for participants without a PFS event.

Duration of response (DOR)

DOR with confirmed response only applies to participants whose best overall response is confirmed CR or confirmed PR according to RECIST 1.1 based on tumor response data per BIRC assessment. The start date is the date of first documented response of CR or PR (i.e. the start date of response, not the date when response was confirmed), and the end date is defined as the date of the first documented progression or death due to underlying cancer. Participants continuing without progression or death due to underlying cancer will be censored at the date of their last adequate tumor assessment.

ECOG performance status

The ECOG PS scale (Table 2-8) will be used to assess physical health of participants, ranging from 0 (most active) to 5 (least active):

Table 2-8 ECOG Performance Scale

Score	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

The following intervals will be used to group the ECOG PS data over time. Day in columns 2 and 3 is defined as date of ECOG PS assessment date – randomization date + 1. The correspondence with Day in column 1 assumes that a participant is treated on the day of randomization; however the definition of Day in columns 2 and 3 still applies if this is not the case, i.e. randomization date is taken as the reference for the windows.

Table 2-9 Time windows for ECOG PS assessments

Assessment	Target day of assessment	Time Interval
Baseline		Day 1 (if not available use screening)
Cycle 2 Day 1	29	Day 2 to day 42
Cycle 3 Day 1	57	Day 43 to day 70
Cycle 4 Day 1	85	Day 71 to day 98
Cycle k Day 1 (k≥5)	d=(k-1)*28+1	Day d-14 to day d+13
End of Treatment		Assessment taken at the end of treatment visit

If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the worst of the 2 assessments will be used.

Time to definitive deterioration in ECOG PS is defined as the time from the date of randomization to the date when ECOG PS has definitively deteriorated by at least one category compared with baseline. Deterioration is considered definitive if there is no subsequent improvement in ECOG PS back to the baseline category or above. Participants will be censored if no definitive deterioration in ECOG PS is observed before the analysis cut-off date. The censoring date will be the date of the last ECOG PS assessment prior to cut-off. For the situation of baseline assessment only, if death or deterioration occurred early (e.g., occurred within the period of first 2 assessments), then it would be considered as an event. Otherwise, it would be censored at randomization date. For the situation of post-baseline assessments available, if death or deterioration occurred after 2 or more missing assessments, censor it at the last valid assessment. For new antineoplastic therapy, ignore it and consider death or deterioration occurred after new antineoplastic therapy as an event.

2.7.2 Statistical hypothesis, model, and method of analysis

Secondary objectives fall outside of the statistical testing framework specified for the primary and key secondary objectives thus no statistical testing will be undertaken.

ORR/CBR

ORR/CBR with confirmed response per BIRC assessment based on the FAS will be summarized using descriptive statistics (N, %) by treatment group along with two-sided exact binomial 95% CIs (Clopper and Pearson 1934). As a sensitivity analysis, ORR/CBR with confirmed response as assessed by local investigator review will be calculated by treatment group and presented along with 95% confidence intervals. In addition, the following ORR/CBR will also be calculated and presented by treatment group together with exact binomial 95% confidence intervals:

- ORR/CBR with confirmed response per BIRC for participants with measurable disease at baseline,
- ORR/CBR with unconfirmed response per BIRC based on the FAS,
- ORR/CBR with unconfirmed response per BIRC for participants with measurable disease at baseline.

Time to response

Time to response per BIRC assessment will be listed and summarized by treatment group. The distribution of time to response will be estimated using the Kaplan-Meier method and the median time to response will be presented along with 95% confidence interval. A descriptive summary of time to response for the responders will also be presented. No inferential analysis that compares time to response between the two treatment groups will be performed.

Duration of Response

DOR with confirmed response per BIRC assessment will be listed and summarized by treatment group for all participants in the FAS with confirmed BOR of CR or PR. The distribution of duration of response will be estimated using the Kaplan-Meier method and the median duration of response will be presented along with 95% confidence interval. As a supplemental analysis, DOR with unconfirmed response based on the FAS will also be listed and summarized by treatment group. No inferential analysis that compares duration of response between the two treatment groups will be performed.

ECOG performance status

Time windows are applicable for descriptive summary of ECOG data by visit only. For time to deterioration analysis described hereafter, all post-baseline assessments will be considered. Only assessments collected while the participant is on treatment and at the end of treatment visit will be included in the time to definitive deterioration of ECOG.

Frequency counts and percentages of participants in each score category will be provided by treatment group and time point.

Time to definitive deterioration in ECOG PS will be analyzed in the FAS according to the randomized treatment group and strata assigned at randomization. Kaplan-Meier curves, medians and 95% confidence intervals of the medians will be presented for each treatment group. The hazard ratio for the time to definitive deterioration in ECOG PS will be calculated, along with its 95% confidence interval, using a stratified Cox model.

2.8 Analysis of other secondary objective(s)

2.8.1 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will be presented by treatment group.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for deaths, including on treatment and post treatment deaths will be provided.

The on-treatment period lasts from the date of first administration of study treatment to 30 days after the date of the last actual administration of any study treatment.

The overall observation period will be divided into three mutually exclusive segments:

1. Pre-treatment period: from day of participant's informed consent to the day before first dose of study treatment.
2. On-treatment period: from day of first dose of study treatment to 30 days after last dose of study treatment.
3. Post-treatment period: starting at day 31 after last dose of study treatment.

2.8.1.1 Adverse events (AEs)

2.8.1.1.1 General rules for AE Reporting

Summary tables for adverse events (AEs) will include only AEs that started or worsened during the on-treatment period, the treatment-emergent AEs. However, all safety data (including those from the pre and post-treatment periods) will be listed and those collected during the pre-treatment and post-treatment period are to be flagged. All information obtained on adverse events will be displayed by treatment group and participant.

The number (and percentage) of participants with treatment emergent adverse events may be summarized in the following ways:

- by treatment, system organ class and preferred term.
- by treatment, system organ class, preferred term and maximum severity.
- by treatment, Standardized MedDRA Query (SMQ) and preferred term.

A participant with multiple adverse events within a system organ class is only counted once towards the total of the system organ class.

Separate summaries will be provided for study treatment related adverse events, death, serious adverse events, other significant adverse events leading to discontinuation and adverse events leading to dose adjustment.

The incidence of treatment-emergent adverse events (new or worsening from baseline) will be summarized by system organ class and or preferred term, severity (based on CTCAE grades), type of adverse event, relation to study treatment. A participant with multiple grades for an AE will be summarized under the maximum grade recorded for the event. AEs with a missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries the system organ class will be presented alphabetically and the preferred terms will be sorted within SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the alpelisib treatment group.

The frequency of grade 3 and above AEs will be summarized separately.

Any information collected (e.g. grades, relationship to study treatment, serious, action taken etc.) will be summarized as appropriate.

All AEs, deaths, and serious adverse events (including those from the pre and post-treatment periods) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

2.8.1.1.2 Adverse events of special interest / grouping of AEs

An adverse event of special interest (AESI) is a grouping of adverse events that are of scientific and medical concern specific to alpelisib. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HLGTs (high level group terms), HLTs (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. For each specified AESI, the number and percentage of participants with at least one event of the AESI occurring during the on-treatment period will be summarized.

Summaries of these AESIs will be provided by treatment group (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, etc.). All AEs of special interest will be listed. A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will also be generated.

An electronic Case Retrieval Sheet (eCRS) with the exact composition of the AE groupings is to be used to map reported AEs to the AESI groupings. This file may be updated (i.e. it is a living document) based on review of accumulating trial data, and therefore the groupings are also subject to potential change. The most up-to-date version of the eCRS will be used at the time of the cut-off date for each analysis.

2.8.1.1.3 Time to first occurrence of any CTC grade ≥ 2 [≥ 3] AESI

Time to first occurrence of CTC grade ≥ 2 AESI will be summarized using the Kaplan-Meier method. Median time to first occurrence and 95% CI will be provided. Simple descriptive statistics, median, minimum and maximum as well as the 25th and 75th percentiles, will also be presented. Ascending Kaplan-Meier plots will be generated. For the AESI of 'Hyperglycemia', 'Rash' and 'GI Toxicity Nausea, vomiting and Diarrhea', additional analyses for time to first occurrence of any CTC grade ≥ 3 AESI will be presented. In addition, the median time to first occurrence for the subset of participants who experienced the event of interest will also be provided.

Time to first occurrence of CTC grade ≥ 2 [≥ 3] AESI is defined as the time from the start of study treatment to the start date of the first incidence of an AESI of CTC grade ≥ 2 [≥ 3] i.e. time in days is calculated as (start date of first occurrence of the AESI) – (date of first dose of study treatment) +1.

In the absence of an AESI during the on-treatment period, the censoring date applied will be the earliest of the following dates:

- end date of on-treatment period (up to end of study treatment + 30 days)
- death date
- start date of new antineoplastic therapy (with the exception of palliative radiotherapy) before experiencing any CTC grade ≥ 2 [≥ 3] AESI
- data cut-off date
- withdrawal of informed consent date.

2.8.1.2 Deaths

Separate summaries for on-treatment and all deaths (including post-treatment death) will be produced by treatment group, system organ class and preferred term.

All deaths will be listed, post-treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened participants.

2.8.1.3 Laboratory data

Regarding laboratory assessments, data from all sources (central and local laboratories) will be combined. For the purposes of handling multiple assessments at each visit, all scheduled/unscheduled assessments should be assigned to time windows (Table 2-10).

The summaries will include all laboratory assessments collected no later than 30 days after the last administration of study treatment. All laboratory assessments will be listed and those collected later than 30 days after the last treatment date will be flagged in the listings.

Table 2-10 Time windows for laboratory assessments

Assessment	Target day of assessment	Time Interval
Baseline		≤ Day 1
Cycle 1 Day 8	8	Day 2 to Day 11
Cycle 1 Day 15	15	Day 12 to Day 21
Cycle 2 Day 1	29	Day 22 to Day 32
Cycle 2 Day 8	36	Day 33 to Day 39
Cycle 2 Day 15	43	Day 40 to Day 49
Cycle 3 Day 1	57	Day 50 to day 70
Cycle 4 Day 1	85	Day 71 to day 98
Cycle k Day 1 (k≥5)	=(k-1)*28+1	Day d-14 to Day d+13
End of Treatment		Assessment taken at the EOT visit

All laboratory data will be listed by treatment group, participant, and visit/time. If normal ranges are available, abnormalities will be flagged. Summary statistics will be provided by treatment group and visit/time.

Grading of laboratory values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. The criteria to assign CTC grades are given in Section 5.3.

The following summaries will be produced for the laboratory data (by laboratory parameter):

- Number and percentage of participants with each CTC grade as their worst post-baseline value (regardless of the baseline status). Each participant will be counted only for the worst grade observed post baseline.

- Shift tables using CTC grades to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters with CTC grades.
- For laboratory parameters where CTC grades are not defined, shift tables to the worst post-baseline value will be produced using the low/normal/high classifications based on laboratory reference ranges.
- HbA1c and fasting plasma glucose data will be summarized in tables by treatment group and time point. Summary statistics include number of participants with available data, mean, standard deviation, median, minimum and maximum. Figures of mean fasting glucose with two-sided 95% confidence intervals over time by treatment group may also be produced to view the trends over time.
- A plot of baseline HbA1c vs. worst post-baseline HbA1c (%) will be presented.

The following listings will be produced for the laboratory data:

- Listing of participants with CTC grade 3 or 4 laboratory abnormalities;
- Listing of all laboratory data with values flagged to show the corresponding CTC grades and the classifications relative to the laboratory reference ranges.

The following time to laboratory event analyses will be conducted.

- Time to first occurrence of CTC Grade ≥2 glucose
- Time to first occurrence of CTC Grade ≥3 glucose
- Time to resolution of CTC Grade ≥2 glucose
- Time to resolution of CTC Grade ≥3 glucose

Note that CTC grade 3 and 4 glucose lab events use both fasting and non-fasting glucose laboratory values.

Time to first occurrence

Median time to first occurrence and 95% CI will be provided based on the Kaplan-Meier method. In addition, Kaplan-Meier plots will be generated.

Time to first occurrence of grade X or worse laboratory toxicity is defined as the time from the start of treatment to the start date of the first incidence of grade X or worse laboratory toxicity, i.e. time in days is calculated as (start date of first occurrence) – (date of first dose of study treatment) +1. A participant will be censored if:

- The participant did not report any post-baseline grade X or worse event on or before the analysis cut-off date.
- The participant discontinued treatment without reporting any grade X or worse event up to 30 days after study treatment discontinuation.
- The participant died without reporting any grade X or worse event.
- The participant received a new anticancer therapy before reporting any grade X or worse event up to 30 days after study treatment discontinuation.

The censoring date will be the earliest of the following dates: end of treatment + 30 days, analysis cut-off, new anti-cancer therapy start date (with the exception of palliative radiotherapy), death date, withdrawal of informed consent date and last non-missing

assessment for the lab parameter. Note that participants who have grade X or worse toxicity at baseline or missing baseline evaluation will be excluded from this analysis.

In addition, the median time to first occurrence for the subset of participants who experienced the event of interest will be calculated. Simple descriptive statistics, median, min and max as well as 25th percentile and 75th percentile, will be presented.

Time to resolution (i.e. duration)

Time to resolution of first occurrence of laboratory toxicity may also be summarized. Time to resolution of the first event is defined as time from onset of the first event to the date of resolution of the first event: (date of resolution of the first event) – (date of onset of the first event) +1. Resolution of the first event means that there is a lab value showing disappearance of the first event.

Time to resolution of an event will be presented for the subset of participants of the Safety Set who experience the event and will be summarized using the Kaplan-Meier method. Median time to resolution and 95% CI will be presented. In addition, Kaplan-Meier plots will be generated.

A participant will be censored for time to resolution if there is no resolution during the on-treatment period. The same censoring rules as described above for time to first occurrence will apply to time to resolution.

Liver function parameters

The liver function parameters of interest in this study are total bilirubin (TBIL), ALT, AST and alkaline phosphatase (ALP).

The number and percentage of participants meeting the following categorical liver function test criteria will be summarized:

- ALT > 3xULN, 5xULN, 10xULN, 20xULN
- AST > 3xULN, 5xULN, 10xULN, 20xULN
- ALT or AST > 3xULN, 5xULN, 8xULN, 10xULN, 20xULN
- TBIL > 2xULN, 3xULN

For the following combined categories, the assessments need not be concurrent, i.e. participants are counted based on their most extreme value for each parameter (highest in the case of ALT, AST and TBIL; lowest in the case of ALP). Further medical review will be conducted to assess potential confounding factors such as liver metastases, liver function at baseline, etc.

- If AST and ALT ≤ ULN at baseline
  - ALT or AST >3x ULN & TBIL>2x ULN
  - ALT or AST >3x ULN & TBIL>2x ULN & ALP ≥2x ULN
  - ALT or AST >3x ULN & TBIL>2x ULN & ALP <2x ULN
- If AST and ALT > ULN at baseline
  - Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN)

- Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN) & ALP ≥2x ULN
- Elevated ALT or AST (>3x Baseline value or 8x ULN) & TBIL(>2x Baseline value and 2x ULN) & ALP <2x ULN

Additional categories may be added to the above list based on any updates to the internal guidelines on collection, analysis, and presentation of liver safety data.

2.8.1.4 ECG

ECG data will be summarized by presenting summary statistics of the raw data and change from baseline by treatment group and time point.

Notable elevations of ECG summarize the number of participants meeting or exceeding predefined limits in terms of absolute QT/QTc interval data/PR/RR/QRS or changes from baseline as defined in Table 2-11, and notable elevations of ECG summary includes only newly occurring ECG abnormality. A newly occurring ECG abnormality is defined as an abnormal post-baseline ECG finding that is not present at baseline. The percentage of participants having notable ECG interval values is based on the number of participants at risk for the change with a value at baseline and post-baseline.

Table 2-11 Clinically notable ECG values

ECG parameter (unit)	Clinically notable criteria
QT, QTcF (ms)	New > 450 and ≤ 480
	New > 480 and ≤ 500
	New > 500
	Increase from Baseline > 30 and ≤ 60
	Increase from Baseline > 60
PR duration (ms)	Increase > 25% from baseline and to PR duration > 200 New > 200
QRS duration (ms)	Increase > 25% from baseline and to QRS duration > 120 New > 120
Heart Rate (bpm)	< 50 and decrease from Baseline of > 25%
	> 100 and increase from Baseline of > 25%

All ECG data will also be listed by treatment group, participant and visit/time. Abnormalities will be flagged.

2.8.1.5 Cardiac imaging

Shift tables comparing baseline to worst post-baseline cardiac imaging (MRI/MRA or ECHO or MUGA) overall interpretation will be provided. Percentages will be based on all participants in the Safety set.

Note: If there is any change in the methodology used throughout the study compared to baseline, the post-baseline values for which the methodology differs from baseline will be discarded in the tables presenting comparisons to baseline.

Descriptive statistics of the left ventricular ejection fraction (LVEF) at baseline, worst post-baseline value and change from baseline to worst post-baseline value will be provided.

A listing of participants with newly occurring clinically significant abnormality will be produced by treatment group.

2.8.1.6 Vital signs

Vital signs assessments are performed in order to characterize basic body function. The parameters expected to be collected include: weight, body temperature, pulse rate, and systolic and diastolic blood pressure.

The criteria for clinically notable abnormalities are defined in Table 2-12 below.

Table 2-12 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Systolic blood pressure (mmHg)	≥180 with increase from baseline of ≥20	≤90 with decrease from baseline of ≥20
Diastolic blood pressure (mmHg)	≥105 with increase from baseline of ≥15	≤50 with decrease from baseline of ≥15
Pulse rate (bpm)	≥100 with increase from baseline of >25%	≤50 with decrease from baseline of > 25%
Body temperature	≥ 39.1	-
Weight (kg)	increase > 10% from Baseline	decrease > 10% from Baseline

The following summaries will be produced for each vital sign parameter by treatment group:

- Summary statistic for change from baseline to the worst post-baseline value (in both directions, i.e. from baseline to highest post baseline and from baseline to lowest post baseline value).
- Number and percentage of participants with at least one post-baseline vital sign abnormality (in both directions, i.e. both elevated and below normal values).

In addition, the following two listings will be produced by treatment group:

- Participants with clinically notable vital sign abnormalities.
- All vital sign assessments will be listed by participant and vital sign parameter.

In both listings, the clinically notable values will be flagged and also assessments collected later than 30 days after the last treatment date will be flagged.

2.8.1.7 Other safety data

Data from other tests will be listed, notable values will be flagged, and any other information collected will be listed as appropriate.

All assessments collected later than 30 days after the last treatment date will be flagged in the listings.

Subgroup analyses will be explored as described in Section 2.2.3.

2.8.2 Pharmacokinetic endpoints

2.8.2.1 General principle

The PAS will be used in all pharmacokinetic data analysis and PK summary statistics for alpelisib and olaparib, unless otherwise specified.

Evaluable concentrations are defined as those for which all of the following apply:

- samples taken within the following time windows around the scheduled time points:
  - Pre-dose: prior to dosing on the assessment day and collected at 24 ± 2 hours after the last dose
  - 1 h post-dose: within ± 10 minutes of the scheduled time point
  - 2, 3, 4, 6 or 8 h post-dose: within ± 30 minutes of the scheduled time point (**only applicable to the full PK cohort**)
  - 3 h post-dose: within + 30 minutes of the scheduled time point (i.e., sample should be collected within 30 min after recommended scheduled timepoint of 3 hours, not before; **only applicable to the sparse PK cohort**)
  - 24 h post-dose: within ± 120 minutes of the scheduled time point (**only applicable to alpelisib in the full PK cohort**)
- no vomiting occurs within the first 4 hours of the last dose (pre-dose samples)
- no vomiting occurs within the first 4 hours of the current dose (post-dose samples)
- the concentration has not been flagged for exclusion by the pharmacokineticist
- assessments with at least 3 consecutive days of daily dosing at the planned dose (dose assigned at study entry) immediately prior to the PK collection

A PK parameter will be considered as NOT evaluable if any of the following conditions are satisfied:

- vomiting occurs within 4 hours of the current or last dose
- participant do not receive at least 3 consecutive days of the protocol-planned dose of the respective drug prior to and on the PK collection day
- the parameter is flagged for exclusion by the pharmacokineticist

Only evaluable PK concentrations/parameters which are not flagged for exclusion will be used for figures and summaries. However, concentration and parameter listings will include all values, with flags indicating those excluded from analyses.

PK data collected from this study may be combined with data from other studies to support a population PK analysis using non-linear mixed effect modeling; details will be provided in a separate analysis plan.

PK parameters

The PK parameters that will be determined are shown in Table 2-13. The PK parameters are derived based on the non-compartmental methods using Phoenix WinNonlin® software version 8.0 or higher.

Table 2-13 Non-compartmental PK parameters for alpelisib and olaparib

AUC%Extrap[1]	Area under the plasma concentration-time curve extrapolated from the time t to infinity as a percentage of total AUC (%)
AUClast	The AUC from time zero to the last measurable concentration sampling time (tlast) (mass x time x volume^{-1})
AUCtau[2]	The AUC calculated to the end of a dosing interval (tau) at steady-state (amount x time x volume^{-1})
Clast	Last measurable concentration (mass x volume^{-1})
Cmax	The maximum (peak) observed plasma, blood, serum, or other body fluid drug concentration after single dose administration (mass x volume^{-1})
Tmax	The time to reach maximum (peak) plasma, blood, serum, or other body fluid drug concentration after single dose administration (time)
T1/2	The elimination half-life associated with the terminal slope (λz) of a semi logarithmic concentration-time curve (time).
Tlast	Last measurable concentration sampling (time)
Rsqadj[1]	Square of the correlation coefficient associated with lambda_z

[1] AUC%Extrap and Rsqadj will be used in the interpretation of the primary PK parameters and therefore will be included in the listings only.
[2] For alpelisib the dosing interval is 24 hrs, while for olaparib it is 12 hrs

Descriptive statistics (n, arithmetic mean, CV% mean, standard deviation (SD), median, geometric mean, CV% geo-mean, minimum and maximum) will be presented for Pharmacokinetic analysis set for all PK parameters defined in Table 2-13 except Tmax, where only n, median, minimum and maximum will be presented.

2.8.2.2 PK concentrations

Alpelisib and olaparib plasma concentration data will be listed by participant and visit/sampling time point. Descriptive summary statistics will be provided by visit/sampling time point.

Descriptive statistics of concentrations will include n, m (number of non-zero concentrations), mean, CV%, SD, median, geometric mean, geometric CV%, minimum and maximum. Coefficient of variation CV (%) is calculated as follows: 100*(SD/arithmetic mean).

Geometric CV (%) is calculated as follows from non-zero values:

$$CV(\%) = 100 \cdot \sqrt{\exp\left(\hat{\sigma}^2\right) - 1}$$

where $\hat{\sigma}^2$ denotes the variance of the log-transformed values.

Geometric mean and arithmetic mean (SD) concentration-time profiles on C1D8, as well as individual concentration-time profiles with median on C1D8 of alpelisib and olaparib will be graphically presented separately. These plots will also be repeated for pre-dose samples.

2.8.2.3 Handling missing and invalid values

Plasma samples will be assayed for alpelisib and olaparib concentrations by Novartis or Novartis designated laboratory using validated LC-MS/MS methods with a LLOQ of approximately CCI ng/mL for alpelisib and approximately CCI ng/mL for olaparib.

All concentrations below the LLOQ will be displayed in listings as zero with a flag and handled as zero in any calculations of summary statistics, but handled as missing for the calculation of the geometric means and their CV.

Any missing PK data will not be imputed.

2.8.3 Patient reported outcomes

The FAS will be used for analyzing PRO data unless specified differently. The TOI score of the FACT-O, a composite endpoint of physical well-being (PWB), functional well-being (FWB), and Ovarian Cancer Subscale (OCS) (see protocol Section 8.5.1.1.1), will be analyzed as the primary PRO measure to evaluate the quality of life of participants between treatment groups.The PRO instruments are planned to be administered during screening and every 8 weeks after randomization during the first 18 months, and every 12 weeks thereafter until the end of treatment. PRO assessments will continue to be collected during the efficacy follow-up after the end of treatment and at 8 weeks after disease progression.

The baseline is defined as the last PRO assessment on or prior to randomization.

The TOI score from the FACT-O will be displayed as mean profiles for each treatment group, presented over time using time windows as described below in Table 2-14. Data obtained after the end of treatment will be summarized separately. Change from baseline in the TOI score of the FACT-O at the time of each assessment will also be summarized.

Time windows will be defined for descriptive summary of PRO data by visit and longitudinal data analysis. If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to visit will be considered. Data obtained at the end of treatment will be classified as other assessment in the corresponding time window.

Table 2-14 Time windows for PRO

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline	On or before Study Day 1*	≤ Study Day 1*
Cycle 3 Day 1	Study Day 57	Study Days 2 – 85
Cycle 5 Day 1	Study Day 113	Study Days 86 – 141
Cycle k Day 1 (k=7, 9, ..., 17)	d=(k-1)*28+1	Study Days d-27 to d+28

Cycle 19 Day 1 Every 12 weeks thereafter	Study Day 505	Study Days 478 - 547
Cycle 22 Day 1	Study Day 589	Study Days 548 - 631
Cycle 25 Day 1	Study Day 673	Study Days 632 - 715
Cycle k Day 1 (k=28, 31, …)	d=(k-1)*28+1	Study Days d-41 to d+42 Note: EOT data visit are included if obtained within 7^ days of last non-zero dose intake.
End of treatment		
End of treatment	N.A.	Data collected under EOT visit, if no data were collected at the EOT visit last available data obtained before EOT
Post treatment		
30-day safety follow-up	Post treatment study day 30	Post treatment study day 27 – 33
Post treatment follow-up k	Day of post treatment follow-up k	NA

Follow-up after disease progression	8 weeks after disease progression	Data collected under this visit, if no data were collected at this visit last available data obtained before this visit

* Study Day 1 = randomization date
Post treatment study day 1=end of treatment date + 1 day
^ 7 days is considered to be the time until total drug elimination

Scoring of raw data and methods for handling of missing items or missing assessments will be handled according to FACIT Administration and Scoring Guidelines.

As the main analysis and to best utilize the repeated PRO assessments, a repeated measures model for longitudinal data will be used to estimate differences in the TOI score of FACT-O between treatment groups. The modeling will mainly be done on the actual score. Note that the modeling of the change in score or the actual score is equivalent since adjustment for baseline score is considered (CHMP Guideline on adjustment for baseline covariates). This repeated measures model will include terms for treatment, the stratification factors, time, baseline value as main effects, and an interaction term for treatment by time. This analysis will be restricted to participants with an evaluable baseline score and at least one evaluable post-baseline score. All data collected until end of treatment (including the end of treatment assessment) will be included in the analysis. Note that only data collected under treatment (i.e. while the patient is treated) will be included. The end of treatment assessment will be included if collected within 7 days of the last dose intake. Additionally, the differences in least square means between

treatment and control group, and the corresponding 2-sided 95% confidence interval (CI) at selected time points will be presented.

Time will be considered as a continuous variable expressed in weeks, i.e. considering that PRO data follow a linear trend.

As a first approach, an unstructured correlation matrix will be used to model the correlation within participants. The structure of the correlation matrix will be investigated and simplified using likelihood ratio tested if appropriate.

If PRO data is found not to follow a linear trend, the time variable might be considered as a categorical variable instead of continuous in the model. While increasing flexibility and aiming for a better data fit, this approach may however complicate the interpretation.

Analysis of the time to definitive 10% deterioration in the primary PRO variable of interest will be performed. Definitive 10% deterioration is defined as a worsening in score by at least 10% compared to baseline, with no later improvement above this threshold observed during the treatment period, or death due to any cause. A higher score of FACT-O represents better QoL.

The time to definitive deterioration is calculated from the date of randomization to the date of definitive deterioration event. Only data while on treatment are included in the analysis. A single measure reporting a decrease of at least 10% is considered definitive only if it is the last one available for the participant. Participants with no events at cut-off date are censored at date of last assessment before cutoff. Participants who discontinued the study treatment prior to the analysis data cutoff will be censored at the date of their last assessment before study treatment discontinuation. Participants receiving any further anti-neoplastic therapy before definitive worsening will be censored at the date of their last assessment before starting this therapy. Participants with no baseline data will be censored at Study Day 1. Additional analyses by using different approaches may be conducted after the completion of the Clinical Study Report (CSR) and will be documented in another separate report.

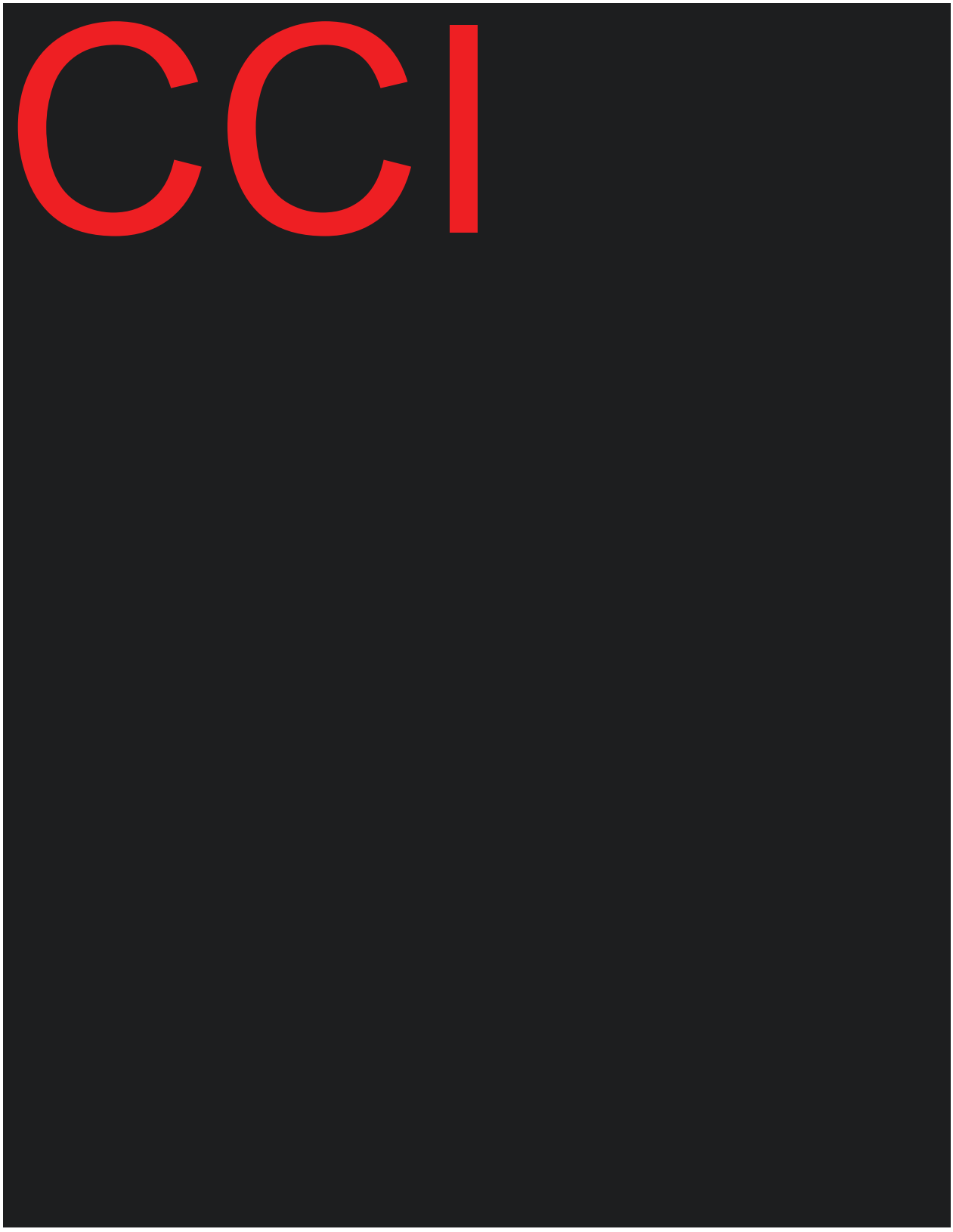
Death is considered as an event when it occurs within a period of time defined by 2 times the period between two assessments as planned in the study protocol. This avoids overestimating the time to definitive worsening in participants dying after an irregular assessment scheme. Participants who die after more than twice the planned period between two assessments since the last assessment are censored at the date of their last available questionnaire. For participants with no baseline data, death is considered as an event when it occurs within a period of time defined by 2 times the period after baseline, otherwise the participants will be censored at Study Day 1.

The survival distributions will be presented descriptively using Kaplan-Meier curves. Summary statistics from the Kaplan-Meier distributions will be determined, including the median time to 10% deterioration and the proportions of participants without 10% deterioration at 2-monthly intervals. Both point estimates and 95% CIs will be presented. A stratified Cox regression model will be used to estimate the hazard ratio (HR) of time to deterioration, along with 95% confidence interval.

2.9 Analysis of exploratory endpoints

2.9.1 Efficacy endpoint(s)

2.9.2 Patient reported outcomes



2.9.3 Pharmacokinetics and PK/PD

Details of the analysis method will be developed in a separate analysis plan and will be documented in a separate report or submission document.

2.9.4 Biomarkers

2.9.4.1 Outline of the data analysis

Additional analyses that may be performed after the completion of the end-of-study Clinical Study Report (CSR) will be documented in separate reports. These analyses may include but are not limited to the meta-analysis of data from this study combined with data from other studies or the analysis of biomarkers generated from samples collected during the study but analyzed after the database lock and completion of the CSR. The data analysis may be described in an addendum of the SAP or in a stand-alone analysis plan document, as appropriate.

The following exploratory endpoints may be assessed and the details of the analysis may be outlined in a biomarker SAP:

2.9.4.2 Data handling principles

The screening assessment will be used as the baseline assessment for tumor biopsy while the Cycle 1 Day 1 assessment (pre-dose) will be used as the baseline value for other biomarkers.

For assessments performed in tumor biopsies, fresh biopsy results will be used for baseline if both archival and fresh tumor samples are available.

When more than one biomarker data values are available for a participant at any time point, the mean of the replicate values will be used for all statistical analyses.

Data preprocessing and transformations will be described in detail in the Programming Datasets Specifications (PDS) document.

2.9.4.2.1 Analysis sets

The FAS will be used for all biomarker analysis. Unless otherwise specified, all statistical analyses of biomarker data will be performed on participants with biomarker data.

2.9.4.2.2 Basic tables, figures and listings

Unless otherwise stated, as project standard, all biomarker data collected will be listed and summarized. In the event of the collection of large biomarker data such as next generation sequencing, gene expression or protein expression panels, some pragmatic considerations will be applied to limit output, as these may easily top thousands, if not tens of thousands of pages offering little to no scientific value.

Depending on the endpoint of interest, graphical displays such as box plots or strip plots may be used to assess the relationship of different biomarkers with clinical benefit. These may be separated by treatment group and include either baseline or change from baseline values, where applicable.

For categorical markers such as mutation status, 2x2 contingency tables may be used to assess the relationship with clinical benefit. Kaplan-Meier curves may be generated given the number of PFS events warrant such an assessment.

2.9.5 Duration of follow-up

Study follow-up will be summarized using the following methods:

- Summary of duration between randomization and cut-off date, and follow-up times for PFS/OS, which are defined as follows:
    - Randomization (recruitment) period = (Date of last participant randomized - Date of first participant randomized + 1) / 30.4375 (months)
    - Duration between randomization and data cut-off date = (Cut-off date – Date of randomization + 1) / 30.4375 (months). This item will be summarized overall.
    - Follow-up time = (Date of event or censoring – Date of randomization + 1) / 30.4375 (months) regardless of censoring. Date of censoring is defined as the last adequate tumor assessment date for PFS or last contact date for OS. This item will be summarized by treatment group.

All summaries will be reported in months. The calculations for PFS will be based on BIRC assessment. Date of censoring is the same as defined for the PFS and OS analysis.

The time from PFS/OS censoring date to data cut-off date will be summarized by time intervals in months: <3, 3 to < 6, 6 to < 12, 12 to < 18, 18 to <24 and by 12 month intervals thereafter if necessary. The gap time is calculated as ([analysis cut-off date] - [censoring date] + 1)/30.4375.

2.10 Interim analysis

Primary endpoint: Progression free survival (PFS)

One futility interim analysis is planned when approximately 90 of the 224 targeted PFS events (approximately 40% information fraction) have been documented (expected around 13.9 months from the date of first participant randomized in the study). The primary intent of the interim analysis is to stop early for lack of efficacy (futility); [redacted]. Approximately 209 participants (approximately 58.5%) are expected to be randomized at the time of the interim futility analysis, i.e., when approximately 90 PFS events have occurred.

An user-defined gamma spending function ($\gamma = 2.5$) (Hwang et al. 1990) will be used as a beta-spending function to determine the non-binding futility boundary (as implemented in East 6.4).

Based on the choice of β-spending functions described above, and if the interim analysis is performed exactly at 90 PFS events, the futility boundary in terms of p-value scale (or Z-statistic scale) at the first interim is calculated as p=0.236 (or Z=-0.72). The observed (i.e. nominal) p-value has to be greater than the p-value scale futility boundary = 0.236 to conclude futility. In addition, the posterior predictive probability of success based on the final planned number of PFS events will be calculated, given the interim data, and provided to the DMC at the time of the futility interim analysis as supportive information.

Since the observed number of events at the interim analysis may not be exactly equal to the planned 90 PFS events, the futility boundary will need to be re-calculated using the pre-specified β-spending function and based on the actual number of observed events at interim and the total number of targeted events to calculate the exact information fraction. The observed p-value (or Z-test statistic) at the interim analysis will then be compared against the re-calculated futility boundary.

If the study continues to the final PFS analysis, the final PFS analysis will be performed when approximately 224 PFS events have been documented. If exactly 90 events are observed at the interim analysis, the study continued and exactly 224 events are obtained at the final analysis, the observed p-value will have to be less than 0.025 to declare statistical significance. In practice, the final analysis will be based on the actual number of PFS events documented at the cut-off date for the final PFS analysis and the alpha already spent at the interim analysis. The boundary for the final analysis will be derived accordingly such that the overall significance level across all analyses is maintained at 0.025.

Statistical properties of the group sequential design are summarized in Table 2-15.

Table 2-15 Simulated probabilities to stop for efficacy or futility at the interim or final PFS analysis

Scenario	Look	Number of PFS events	Simulated cumulative probabilities (%)		Simulated incremental probabilities (%)	
			Stop for efficacy	Stop for futility	Stop for efficacy	Stop for futility
Under H_{01} (HR=[redacted])	Interim	90	CCI	CCI	CCI	CCI
	Final	224				
Under H_{a1} (HR=CCI)	Interim	90				
	Final	224				
Under other H_{a1} (HR=CCI)	Interim	90				
	Final	224				

Simulation is performed in East 6.4 with number of simulations = 10,000 and seed=2020.

[redacted] CCI [redacted].

The interim analyses will be performed by an independent statistician (not involved with the conduct of the study). Further details will be described in the DMC charter. The results of the interim analyses will be provided to the DMC by the independent statistician.

Key secondary endpoint: overall survival (OS)

A hierarchical testing procedure will be adopted and the OS analyses will be performed only if the primary efficacy endpoint PFS is statistically significant. A maximum of three analyses is planned for OS; at the time of the final analysis for PFS (provided PFS is significant), at which point approximately 126 deaths are expected; and at the time when approximately 189 deaths are expected (expected approximately 31 months from date of first participant to be randomized) as well as a final analysis for OS when approximately 252 deaths are expected (expected approximately 44 months from date of first participant to be randomized). An α-spending function according to Lan-DeMets (O'Brien-Fleming) as implemented in East 6.4, independent of the one used for the primary efficacy analysis, along with the testing strategy outlined below will be used to maintain the overall type I error probability (Lan and DeMets 1983). This guarantees the protection of the overall level $\alpha = 2.5\%$ across the two hypotheses and the repeated testing of the OS hypotheses in the interim and the final analysis (Glimm et al. 2010). The trial allows for the stopping of the study for a superior OS result, provided the primary endpoint PFS has already been shown to be statistically significant favoring the treatment group. The exact nominal p-values that will need to be observed to declare statistical significance at the time of these analyses for OS will depend on the number of OS events that have been observed at the time of these analyses and the α for OS already spent at the time of earlier analyses.

The projected timing of interim analyses is summarized in Table 2-16.

Table 2-16 Estimated timelines for interim and final analyses

Months after randomization of the first participant (approximation)	# PFS Events	Cumulative Power against a hazard ratio of CCI	# OS events	Cumulative Power [b] against a hazard ratio of CCI
14	90 (40 %)	CCI	-	CCI
22	224 (100 %)		126 (50 %)[a]	
31	-		189 (75 %)[a]	
44	-		252 (100 %)[a]	

CCI

Note: Simulation is performed in East 6.4 with number of simulations = 10,000 and randomization seed = 2020

Confidentiality of interim results

At the time of final PFS analysis, both PFS and interim OS analysis will be performed by the Sponsor's clinical team. All participants will continue to be followed for OS until the final OS analysis (or earlier if OS reaches statistical significance at any of the interim analyses). The 2nd OS interim and the final OS analysis will also be performed by the Sponsor's clinical team.

3 Sample size calculation

3.1 Primary analysis

Based on the AURELIA study (Pujade-Lauraine et al 2014), the median PFS in the control arm of current study is assumed to be around 3.4 months for the purpose of sample size calculation. It is expected that treatment with alpelisib + olaparib will result in a CCI% reduction in the hazard rate for PFS, i.e. an expected hazard ratio of CCI (which corresponds to an increase in median PFS to CCI months under the exponential model assumption).

Then in order to ensure CCI % power to test the null hypothesis: PFS hazard ratio = CCI, versus the specific alternative hypothesis: PFS hazard ratio = CCI, it is calculated that a total of 224 PFS events need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment arms in a 1:1 ratio, and a 2-look group sequential design CCI CCI as well as a gamma spending function (Hwang et al. 1990) to define a non-binding futility rule at the interim analysis, using information fractions of (0.40, 1).

Assuming that enrolment will continue for approximately 19.2 months at a stepwise rate: (1) 8 participants per month for the first 6 months; (2) 18 participants per month from 6 months to 12 months; and (3) 28 participants per month after 12 months, as well as a 20% annual dropout rate by the time of the final PFS analysis, a total of 358 participants will need to be randomized

to observe the targeted 224 PFS events at about 22.4 months after the randomization date of the first participant. The sample size of 358 participants will be randomly assigned to each treatment group in a 1:1 ratio (179 participants in the treatment group and 179 participants in the control group). These calculations were made using the software package East 6.4.

3.2 Power for analysis of key secondary variables

OS, as the key secondary variable, will be formally statistically tested, provided that the primary variable PFS is statistically significant. Based on available data from AURELIA study (Pujade-Lauraine et al 2014), the median OS in the control arm is expected to be around 13.3 months. It is hypothesized that treatment with alpelisib will result in a CCI% reduction in the hazard rate for OS, i.e., an expected hazard ratio of CCI (which corresponds to an increase in median OS to CCI months under the exponential model assumption). Then in order to ensure CCI% power to test the null hypothesis: OS hazard ratio = CCI, versus the specific alternative hypothesis: OS hazard ratio = CCI it is calculated that a total of 252 deaths need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment arms in a 1:1 ratio, and a 3-look group sequential design with a Lan-Demets (O'Brien-Fleming) alpha spending function (Lan and DeMets 1983) using information fractions of (0.50, 0.75, 1). Based on the same number of participants that are planned to be enrolled in this study to provide sufficient power for the primary endpoint (i.e., 358 participants), and assuming a 5% annual dropout rate by the time of the final OS analysis, it is estimated that these 252 deaths will be observed approximately 43.5 months after the randomization date of the first participant. Therefore the cut-off date for the final analysis of OS will be approximately 21 months after the cut-off date for the final analysis of PFS. These calculations were made using the software package East 6.4.

4 Change to protocol specified analyses

No change from protocol specified analysis was made.

5 Appendix

5.1 Imputation rules

The missing or partial date imputation rules will be described in the programming datasets specification document.

5.2 AEs coding/grading

5.2.1 Coding of AEs

Adverse events are coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

5.2.2 Grading of AEs

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.3.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE v4.0.3 grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death.

If CTCAE grading does not exist for an adverse event, grades 1 – 5 corresponding to the severity of mild, moderate, severe, life-threatening and death will be used. Information on deaths will also be collected on the 'Death' eCRF.

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version v4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters shown below.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 is not applicable in lab data. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values will be set equal to X-0.0001 or X+0.0001, respectively.

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010) **Page 1**

Lab test (toxicity)	SI unit	Lab test (NCDS)	Normal ranges (Merck manual, July 2015) and *conversion factors*	CTC Grades[1] 0	1	2	3	4
Hematology								
WBC ↓	10^9/L	WBC	**3.9 – 10.7 x 10^9/L**	≥ LLN	< LLN - 3.0 x 10^9/L	< 3.0 – 2.0 x 10^9/L	< 2.0 – 1.0 x 10^9/L	< 1.0 x 10^9/L
WBC [2] (Leukocytosis)	10^9/L	WBC			-	-	> 100 x 10^9/L	-
Hemoglobin [2] (Anemia)	g/L	HGB	**120 - 160 g/L** or 7.4 - 9.9 mmol/L (F) **140 - 170 g/L** or 8.7 – 10.6 mmol/L (M)	≥ LLN	< LLN - 100 g/L < LLN - 6.2 mmol/L	< 100 - 80 g/L < 6.2 - 4.9 mmol/L	< 80 g/L < 4.9 mmol/L	-
Hemoglobin ↑	g/L	HGB	*(16.113 x mmol/L = g/L)*		Increase >0-20 g/L above ULN	Increase >20-40 g/L above ULN	Increase >40 g/L above ULN	-
Platelets ↓	10^9/L	PLAT	**150 - 350 x 10^9/L**	≥ LLN	< LLN - 75.0 x 10^9/L	< 75.0 - 50.0 x 10^9 L	< 50.0 - 25.0 x 10^9/L	< 25.0 x 10^9/L
Neutrophils [3] ↓	10^9/L	NEUT		≥2x10^9/L	< 2.0 - 1.5 x 10^9/L	< 1.5 - 1.0 x 10^9/L	< 1.0 - 0.5 x 10^9/L	< 0.5 x 10^9/L
Lymphocytes [3] ↓	10^9/L	LYM		≥1.5x10^9/L	< 1.5 - 0.8 x 10^9/L	< 0.8 - 0.5 x 10^9/L	< 0.5 - 0.2 x 10^9/L	< 0.2 x 10^9/L
Lymphocytes ↑	10^9/L	LYM			-	> 4 - 20 x 10^9/L	> 20 x 10^9/L	-
Biochemistry								
AST ↑	U/L	AST	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
ALT ↑	U/L	ALT	**0 - 35 U/L** or 0 – 0.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN – 3.0 x ULN	> 3.0 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Total bilirubin ↑	umol/L	BILI	**5.1 – 20.5 umol/L** or 0.3 – 1.2 mg/dL *(17.1 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 10.0 x ULN	> 10.0 x ULN
Alk. Phosphatase ↑	U/L	ALP	**36 - 92 U/L** or 0.5 - 1.5 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 20.0 x ULN	> 20.0 x ULN
Creatinine [4] ↑	umol/L	CREAT	**61.9 - 115 umol/L** or 0.7 – 1.3 mg/dL *(88.4 x mg/dL = umol/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 3.0 x ULN	> 3.0 - 6.0 x ULN	> 6.0 x ULN
Creatinine kinase [4] ↑	U/L	CK	**30 - 170 U/L** or 0.5 – 2.83 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 2.5 x ULN	> 2.5 - 5.0 x ULN	> 5.0 - 10.0 x ULN	> 10.0 x ULN
Albumin [2] (Hypoalbuminemia)	g/L	ALB	**35 - 55 g/L** or 3.5 to 5.5 g/dL	≥ LLN	< LLN - 30 g/L	< 30 - 20 g/L	< 20 g/L	-
Total Cholesterol ↑	mmol/L	CHOL	**3.88 – 5.15 mmol/L** or 150 - 199 mg/dL *(38.67 x mg/dL = mmol/L)*	≤ ULN	> ULN - 7.75 mmol/L > ULN - 300 mg/dL	> 7.75 -10.34 mmol/L > 300 – 400 mg/dL	>10.34-12.92 mmol/L > 400 – 500 mg/dL	>12.92 mmol/L > 500 mg/dL
Lipase ↑	U/L	LIPASE	**<95 U/L** or <1.58 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Amylase ↑	U/L	AMYLASE	**0 - 130 U/L** or 0 – 2.17 ukat/L *(60 x ukat/L = U/L)*	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.0 x ULN	> 2.0 - 5.0 x ULN	> 5.0 x ULN
Uric acid [2] (Hyperuricemia)	umol/L	URATE	**150 - 470 umol/L** or 2.5 – 8 mg/dL *(59.48 x mg/dL = umol/L)*	≤ ULN	> ULN – 10 mg/dL > ULN – 595 umol/L	-	-	> 10 mg/dL > 595 umol/L

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

LAB - CTC grades in Novartis Oncology (26Oct15)

CTC grades for laboratory values in Novartis Oncology (based on CTCAE v4.03 – June 2010) **Page 2**

				CTC Grades[1]				
Lab test (toxicity)	**SI unit**	**Lab test (NCDS)**	**Normal ranges (Merck manual, July 2015) and *conversion factors***	**0**	**1**	**2**	**3**	**4**
Phosphorus[2] (Hypophosphatemia)	mmol/L	PHOS	**0.97 – 1.45 mmol/L** or 3.0 - 4.5 mg/dL *(0.32 x mg/dL = mmol/L)*	≥ LLN	< LLN - 2.5 mg/dL < LLN - 0.8 mmol/L	< 2.5 - 2.0 mg/dL < 0.8 - 0.6 mmol/L	< 2.0 - 1.0 mg/dL < 0.6 - 0.3 mmol/L	< 1.0 mg/dL < 0.3 mmol/L
Calcium (corrected) [2] (Hypercalcemia)	mmol/L	CACALC	**2.2 - 2.6 mmol/L** or 9 - 10.5 mg/dL *(0.2495 x mg/dL = mmol/L)*	≤ ULN	> ULN - 11.5 mg/dL > ULN - 2.9 mmol/L	> 11.5 - 12.5 mg/dL > 2.9 - 3.1 mmol/L	> 12.5 - 13.5 mg/dL > 3.1 - 3.4 mmol/L	> 13.5 mg/dL > 3.4 mmol/L
Calcium (corrected) [2] (Hypocalcemia)	mmol/L	CACALC		≥ LLN	< LLN - 8.0 mg/dL < LLN - 2.0 mmol/L	< 8.0 - 7.0 mg/dL < 2.0 - 1.75 mmol/L	< 7.0 - 6.0 mg/dL < 1.75 - 1.5 mmol/L	< 6.0 mg/dL < 1.5 mmol/L
Magnesium [2] (Hypermagnesemia)	mmol/L	MG	**0.62 – 0.99 mmol/L** or 1.5 – 2.4 mg/dL *(0.4114 x mg/dL = mmol/L)*	≤ ULN	> ULN - 3.0 mg/dL > ULN - 1.23 mmol/L	-	> 3.0 – 8.0 mg/dL > 1.23 – 3.3 mmol/L	> 8.0 mg/dL > 3.3 mmol/L
Magnesium [2] (Hypomagnesemia)	mmol/L	MG		≥ LLN	< LLN - 1.2 mg/dL < LLN - 0.5 mmol/L	< 1.2 - 0.9 mg/dL < 0.5 - 0.4 mmol/L	< 0.9 - 0.7 mg/dL < 0.4 - 0.3 mmol/L	< 0.7 mg/dL < 0.3 mmol/L
Glucose (non-fasting) [2] (Hyperglycemia)	mmol/L	GLUCSN	**<7.8 mmol/L** or <140 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	-	> ULN - 250 mg/dL > ULN - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose (fasting) [2] (Hyperglycemia)	mmol/L	GLUCSF	**3.9 – 5.8 mmol/L** or 70 - 105 mg/dL *(0.05551 x mg/dL = mmol/L)*	≤ ULN	> ULN - 160 mg/dL > ULN - 8.9 mmol/L	> 160 - 250 mg/dL > 8.9 - 13.9 mmol/L	> 250 - 500 mg/dL > 13.9 - 27.8 mmol/L	> 500 mg/dL > 27.8 mmol/L
Glucose [2] (Hypoglycemia)	mmol/L	GLUCSN/ GLUCSF		≥ LLN	< LLN - 55 mg/dL < LLN - 3.0 mmol/L	< 55 - 40 mg/dL < 3.0 - 2.2 mmol/L	< 40 - 30 mg/dL < 2.2 - 1.7 mmol/L	< 30 mg/dL < 1.7 mmol/L
Potassium [2] (Hyperkalemia)	mmol/L	K	**3.5 - 5.0 mmol/L** *(0.2558 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 5.5 mmol/L	> 5.5 - 6.0 mmol/L	> 6.0 - 7.0 mmol/L	> 7.0 mmol/L
Potassium [2] (Hypokalemia)	mmol/L	K		≥ LLN	< LLN - 3.0 mmol/L	-	< 3.0 - 2.5 mmol/L	< 2.5 mmol/L
Sodium [2] (Hypernatremia)	mmol/L	SODIUM	**136 - 145 mmol/L** *(0.435 x mg/dL = mEq/L = mmol/L)*	≤ ULN	> ULN - 150 mmol/L	> 150 - 155 mmol/L	> 155 - 160 mmol/L	> 160 mmol/L
Sodium [2] (Hyponatremia)	mmol/L	SODIUM		≥ LLN	< LLN - 130 mmol/L	-	< 130 - 120 mmol/L	< 120 mmol/L
Triglyceride [2] ↑	mmol/L	TRIG	**< 2.82 mmol/L** or < 250 mg/dL *(0.01129 x mg/dL = umol/L)*	< 150 < 1.71	≥ 150 - 300 mg/dL ≥ 1.71 – 3.42 mmol/L	> 300 - 500 mg/dL > 3.42 – 5.7 mmol/L	> 500 - 1000 mg/dL > 5.7 – 11.4 mmol/L	> 1000 mg/dL > 11.4 mmol/L
Coagulation								
INR [2]↑	1	INR	**0.8 – 1.2**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Activated partial thromboplastin time[2]↑	sec	APTT	**25 - 35 sec**	≤ ULN	> ULN - 1.5 x ULN	> 1.5 - 2.5 x ULN	> 2.5 x ULN	-
Fibrinogen [4]↓	g/L	FIBRINO	**1.5 – 3.5 g/L** or 150 – 350 mg/dL *(0.01 x mg/dL = g/L)*	≥ LLN	< LLN - 0.75 x LLN	< 0.75 - 0.5 x LLN	< 0.5 - 0.25 x LLN	< 0.25 x LLN

ULN = Upper Limit of Normal range; LLN = Lower Limit of Normal range

(1) = **LAB CTC grades 1, 2, 3, 4 overrule the study specific (central or local) normal range criteria**, e.g. if ULN of Sodium is 151 mmol/L and the value is 151 mmol/L, CTC grade 2 is assigned although the value is ≤ ULN.
(2) = **Life-threatening consequences and/or hospitalization are not considered** for determination of LAB CTC grades 3 and 4. Concomitant usage of anticoagulation therapy (for INR and Fibrinogen) is not considered either.
(3) = Values and LNRs for blood differentials can be given as %, absolute values should then be calculated using WBC. Generally, **≥ 1.5 x 10^9/L (lymphocytes) and ≥ 2 x 10^9/L (neutrophils) are considered as LAB CTC grade 0**
(4) = For Creatinine and Fibrinogen, the **comparison with baseline is not considered** for derivation of LAB CTC grades

LAB - CTC grades in Novartis Oncology (26Oct15)

5.4 Statistical models

5.4.1 Baseline comparability

Appropriate descriptive summary statistics of baseline variables (see Section 2.3) will be provided as in-text tables in the core CSR and also in Section 14 in the post-text tables. The summaries will be grouped by dose cohort for the safety run-in part or by treatment group for the randomized part, but no p-values will be provided.

5.4.2 Analysis of time to event data

The following sections present a general methodology to be used to analyze time-to-event variables. Inferential testing however will only be conducted for PFS and OS, i.e., the primary and key secondary endpoints as detailed in Section 2.5 and Section 2.6. The following parameters are considered:

- Progression-free survival
- Overall survival
- Time to definitive deterioration of the ECOG score by at least one category of the score from baseline
- Time to response: defined as the time between date of randomization until first documented response (CR or PR) according to RECIST 1.1
- Duration of response

- Time to first occurrence of grade 2/grade3 or worse laboratory toxicity/adverse events
- Time to resolution of first occurrence of grade 2/grade 3 or worse laboratory toxicity/adverse events
- Time to definitive deterioration of PRO scores

5.4.2.1 Analysis of time-to-event data with ties

The STRATA statement in LIFETEST procedure will be used to analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

5.4.2.2 Checking proportionality of hazard assumption

Plots (SURVIVAL LOGSURV LOGLOGS) generated by LIFETEST procedure in SAS will be used to provide visual checks of the proportional hazard assumption.

- SURVIVAL plots estimated survivor functions. The shape of the curves should be basically the same if hazards are proportional.
- LOGSURV plots the cumulative hazard functions. The larger cumulative hazard should be a multiple of smaller if hazards are proportional.
- LOGLOGS plots log (cumulative hazard). The LOGLOG plot will show parallel curves if hazards are proportional.

5.4.2.3 Hazard ratio

The hazard ratio as a measure of treatment effect will be derived from the Cox proportional hazards model using SAS procedure PHREG with TIES=EXACT option in the MODEL statement. The stratified unadjusted Cox model will be used (where the baseline hazard function is allowed to vary across strata) for the primary analysis, i.e. the MODEL statement will include the treatment group variable as the only covariate and the STRATA statement will include stratification variable(s).

Hazard ratio with two-sided 95% confidence interval will be based on Wald test.

5.4.2.4 Hypothesis and test statistic

The primary efficacy analysis will be the comparison of the distribution of PFS/OS between the two treatment groups using a stratified log-rank test at one-sided 2.5% level of significance.

Assuming proportional hazards model for PFS/OS, the following statistical hypothesis will be tested to address the primary efficacy objective:

$H_{01}: \theta_1 \geq 0$ vs. $H_{a1}: \theta_1 < 0$

where θ_1 is the log-hazard ratio (alpelisib + olaparib arm vs. paclitaxel or PLD arm) of PFS/OS.

The ***stratified log-rank*** test (strata based on IRT data) will be implemented as follows: For each of the K=8 strata, the LIFETEST procedure will be run with the STRATA statement including only the treatment variable. The TIME statement will include the survival time and a (right) censoring variable. The rank statistic S_k and the corresponding variance $var(S_k)$ (k=1, 2, …, K) will be estimated from this analysis.

The final test statistics will then be reconstructed using the formula:

$Z=[S_1+ ... + S_K]/\sqrt{[var(S_1) + ...+ var (S_K)]}$. One-sided p-value will be computed using this Z statistic. Note: Under the null hypothesis, the asymptotic distribution of the test statistic Z is approximately normal (and correspondingly, Z^2 is approximately distributed as chi-square with one degree of freedom).

5.4.2.5 Kaplan-Meier estimates

The survival function in each treatment group will be estimated using the Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

Median survival for each treatment group will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of (Brookmeyer and Crowley 1982). Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula (Collett 1994).

5.4.2.6 Calculation of posterior predictive probability of success for hazard ratio

The posterior predictive probability of observing a significant treatment effect at the end of study, i.e. Probability ($HR_{final} \leq$ CCI $\mid HR_{interim}$) will be calculated and will be provided to DMC as supportive information at the time of futility interim analysis.

The following sections provide the calculation of posterior predictive probability in a general framework.

Let θ denote the natural logarithm of the hazard ratio (HR) for PFS, Y_I and Y_F be the observed ln(HR) at the interim and final analysis respectively. Further assume an r:1 randomization ratio (r for experimental group and 1 for the control).

Assuming a non-informative prior, the posterior distribution of θ at interim can be derived as

$$\theta \mid Y_I=y_I \sim N\left(y_I, \frac{a}{D_1}\right),$$

where

$a = \frac{(r+1)^2}{r}$ and D_1 = Number of events at interim analysis.

The posterior predictive distribution of $Y_F \mid Y_I=y_I \sim N\left(m_{y_I}, v_{y_I}\right)$,

where

D_2 = Number of events between interim and final analysis,

$$m_{y_I} = \frac{D_1\, y_I + D_2\, y_I}{D_1 + D_2} = y_I,$$

$$v_{y_I} = a\,\frac{D_2^2\,(\frac{1}{D_1}+\frac{1}{D_2})}{(D_1+\,D_2)^2} = a\,\frac{D_2}{D_1(D_1+\,D_2)}\,.$$

The predictive probability can be calculated from the normal distribution using the above mentioned mean and variance.

5.5 Group sequential design used in Phase III studies

The statistical methodology for the interim analyses will be based on group sequential methodology with efficacy stopping boundaries defined by type I error spending functions.

Since the exact number of events available for interim and final analyses cannot be predicted exactly in the clinical trial setting, the group sequential design will be implemented using the α spending function approach. This approach is flexible in dealing with any deviations from the targeted event totals, or unexpected changes to the plan.

If the exact number of events observed at the interim and final analyses deviates from the target numbers described in the protocol, the actual critical boundaries will be derived using the pre-specified error spending functions and the actual numbers of events observed.

- At interim analyses, information fractions will be computed as the ratio of the number of events observed at the considered interim analysis relative to the number targeted for the final analysis, as described in the sample size section of the protocol.
- At the final analysis, the critical value will be calculated using the exact number of observed events at the final cut-off date, considering the α-levels spent at interim analyses and considering the actual correlation among the test statistics, in order to achieve a cumulative type I error smaller than the desired significance level (i.e. smaller than 2.5% for a one-sided test).

5.5.1 Alpha-spending function

The stopping boundaries to be used for the efficacy test will be calculated using the α-spending function approach described in Lan and DeMets (Lan and DeMets, 1983). The spending function for one-sided test has the following functional form:

$\alpha(t) = 2 - 2\Phi(Z_{\alpha/2}/\sqrt{t})$

This function generates stopping boundaries that closely resemble the O'Brien-Fleming boundaries (O'Brien and Fleming, 1979).

5.5.2 Methodology

A maximum of three analyses are planned for OS: at the time of the primary analysis for PFS (provided PFS is significant), at which point a total of approximately 126 deaths are expected, after approximately 189 events have been documented, and a final analysis for OS when approximately 252 deaths are expected (expected approximately 44 months from date of first participant to be randomized).

If the exact number of events observed at the interim and final analyses deviates from numbers defined in the protocol, the actual critical boundaries to be used at each of the interim analyses will be recalculated and will be derived from the pre-specified error spending functions using the actual number of events observed and assuming the final events number is the number derived in the sample size section of the protocol. The critical value for the final analysis will be calculated using the exact number of observed events at the cut-off date, and considering the α-levels spent at interim analyses, in order to achieve a cumulative type I error smaller than 2.5%.

It is recognized that circumstances (that are either internal or external to the trial) may require changes in the scheduling of the interim analyses. In case an additional unscheduled interim analysis is requested (e.g. the DMC might request this analysis if the study duration is much longer than expected) the procedure to calculate stopping boundaries needs to be adapted accordingly. An adaptation is also required if the interim analysis is skipped, e.g. is not considered necessary anymore. Both scenarios can be implemented without inflating the type-I error, thanks to the error spending approach used for the group sequential design. The calculation of stopping boundaries needs to be adapted accordingly.

5.5.3 Calculation of stopping boundaries

At the interim and the final analysis the critical boundaries for the group sequential test (stratified log-rank test) will be derived from the predefined spending functions. The calculations will be performed with East 6.4 software.

5.6 Confidence intervals for response rate and clinical benefit rate

Responses will be summarized in terms of percentage rates with $100(1-\alpha)\%$ confidence interval using exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way table (Clopper and Pearson 1934).

5.7 Implementation of RECIST

Response and progression evaluation will be performed according to Novartis RECIST guideline (as described in detail in Section 16.3 of the study protocol), which is based on RECIST version 1.1 (Eisenhauer et al 2009). The text below gives instructions and rules to provide details needed for programming.

5.7.1 Overall lesions response for participants with only non-measurable lesions at baseline

Participants without measurable disease per RECIST 1.1 are eligible to enter the study if they have at least one predominantly lytic bone lesion. For participants with non-measurable lesions only at baseline, the overall lesion response will be based solely on non-target lesion response or an occurrence of a new lesion. Non-measurable lesions will be entered as non-target lesions. Therefore, the best overall response is determined from non-target lesion response and presence of new lesions (refer to Table 16-8 in Section 16.3.3.2.9 of the study protocol).

5.7.2 Best overall response (BOR)

The best overall tumor response will be assessed as per RECIST 1.1 criteria. The definitions and the details on the derivation are given in Section 16.3 of the study protocol.

Only tumor assessments performed before the start of any anti-neoplastic therapies (i.e. any additional anti-neoplastic therapy or surgery) and within 30 days after the last administration of study treatment will be included in the assessment of best overall response.

- New anti-neoplastic therapies will be identified from the data collected on 'Anti-neoplastic therapies since discontinuation of study treatment' eCRF.
- Palliative radiotherapy is the only setting of radiotherapy allowed during the study. Therefore, palliative radiotherapy will not be considered as an anti-neoplastic therapy for assessment of BOR unless reported on the "Antineoplastic Radiotherapies Since Discontinuation of Study Treatment" eCRF. As per RECIST 1.1, it should not be delivered to a target lesion.
- Continuation of combination partner therapy alone after end of study treatment without confirmed progression will also not be considered as a new anti-neoplastic therapy.

The standard definition of a best overall response evaluation of 'stable disease', 'disease progression' or 'not evaluable' given in the Section 16.3 of the study protocol will be used for this study. Best overall response with confirmation of response for each participant is determined from the sequence of overall (lesion) responses (as reported by the investigator for local BOR, and as reported by BIRC for central BOR) according to the following rules:

- CR = at least two determinations of CR at least 4 weeks apart before progression.
- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR).
- SD = at least one SD assessment (or better) > 7 weeks after randomization (and not qualifying for CR or PR).
- Non-CR/non-PD = at least one non-CR/non-PD assessment (or better) > 7 weeks after randomization date (and not qualifying for CR). This applies only for participants with non-measurable disease alone at baseline.
- PD = progression ≤ 17 weeks after randomization (and not qualifying for CR, PR, SD and non-CR/non-PD)
- NE = all other cases (i.e. not qualifying for confirmed CR or PR and without SD or non-CR/non-PD after more than 7 weeks or without progression within the first 17 weeks).

Participants with best overall response "not evaluable" will be summarized by reason for having not-evaluable status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall response NE
- New anti-neoplastic therapy started before first post-baseline assessment
- SD too early (≤7 weeks after randomization)
- PD too late (>17 weeks after randomization and not qualifying for CR, PR, SD or non-CR/non-PD)

Special (and rare) cases where BOR is not evaluable due to both early SD and late PD will be classified as "SD too early".

5.7.3 Disease progression

Progressive disease should only be assigned if it is confirmed by an assessment method as per RECIST 1.1 guidelines (e.g. radiologic assessment, photos for skin lesions, etc.). If a new lesion is detected using an objective assessment method other than radiologic assessment, then it should also be entered as a new lesion in the eCRF with the appropriate method. Discontinuation due to disease progression or death due to study indication, without corresponding supportive data in the RECIST CRF (as defined above), will not be considered as progressive disease in the calculation of best overall response and in the analysis of PFS.

5.7.4 Change in imaging modality

Per RECIST 1.1, a change in methodology can be defined as either a change in contrast use (e.g. keeping the same technique, like CT, but switching from 'with' to 'without' contrast use or vice-versa, regardless of the justification for the change), a major change in technique (e.g. from CT to MRI, or vice-versa), or a change in any other imaging modality. A change from conventional to spiral CT or vice versa will not constitute a major "change in method" for the purposes of response assessment. A change in methodology will result by default in a NE (not evaluable) overall lesion response based on the Novartis calculation. However, a response from the investigator or the central blinded reviewer that differs from the Novartis calculated NE is acceptable, if a definitive response assessment can be justified based on the available information.

Potential discrepancies between the modality used and overall lesion response (e.g. change in modality but response is different from 'not evaluable') will be queried during the data validation process.

5.7.5 Determination of missing adequate assessments

The term 'missing adequate assessment' refers to assessments that are not done or for which the overall lesion response is 'not evaluable'. 'Missing adequate assessment' will also be referred to as 'missing assessment'.

As detailed in Section 16.3.3.2.10 of the study protocol, the PFS censoring and event date options depend on the presence and the number of missing tumor assessments.

An exact rule to determine whether there are no, one or two missing TAs is therefore needed. This rule is based on the interval between the last adequate tumor assessment (LATA) date and the event date. The scheduled date of tumor assessments (in weeks from randomization), protocol specified window for tumor assessments, and the thresholds for LATA that belong to a visit can be found in the following table.

Table 5-1 Schedule for tumor assessment and time windows

Assessment schedule		Scheduled date – 1 week	Scheduled date (weeks from randomization)	Scheduled date +1 week	Threshold*
	Baseline	0	0	0	0

Every 8 weeks for the first 18 months	C3D1	7	8	9	12
	C5D1	15	16	17	20
	C7D1	23	24	25	28
	C9D1	31	32	33	36
	C11D1	39	40	41	44
	C13D1	47	48	49	52
	C15D1	55	56	57	**60**
	C17D1	63	64	65	**68**
	C19D1	71	72	73	78
Every 12 weeks after 18 months	C22D1	83	84	85	90
	C25D1	95	96	97	102
	C28D1	107	108	109	114
	C31D1	119	120	121	126
* The mid-point between current and next visit (except for baseline) and the upper limit for LATA to be matched to a certain scheduled assessment, e.g. if LATA is at week 13, this is after threshold for C3D1 and before that for C5D1, so the matching scheduled assessment is C5D1.					

To calculate the number of missing tumor assessments, the LATA before an event is matched with a scheduled tumor assessment using the time window in Table 5-1 (essentially whichever scheduled assessment it is closest to). Two thresholds, D1 and D2 are calculated for that scheduled assessment based on the protocol-specified schedule and windows

- An event after LATA+D1 will be considered as having >=1 missing assessment
- An event after LATA+D2 will be considered as having >=2 missing assessments

Since there is a change of schedule for tumor assessments at 18 months, D1 and D2 are defined differently depending on when LATA happens.

Rule 1: if LATA happens within 60 weeks from randomization (the matched scheduled tumor assessment is C15D1 or before)

- D1=8+2=10 weeks
- D2=2*8+2=18 weeks

Rule 2: if LATA happens after 60 weeks but within 68 weeks from randomization (the matched scheduled tumor assessment is C17D1)

- D1=8+2=10 weeks
- D2=8+12+2=22 weeks

Rule 3: if LATA happens after 68 weeks from randomization (the matched scheduled tumor assessment is C19D1 or later)

- D1=12+2=14 weeks
- D2=2*12+2=26 weeks.

Therefore, using the D2 definition above, the censoring of an event occurring after ≥2 missing TAs (in the PFS supplementary analysis) can be refined as follows: if the distance between the last adequate TA date and the PFS event date is larger than D2, then the participant will be

censored and the censoring reason will be 'Event documented after two or more missing tumor assessments'.

The same definition of D2 will be used to determine the PFS censoring reason. If the distance between the last adequate tumor assessment date and the earliest of the following dates (analysis cut off, consent withdrawal etc.) is less than or equal to D2:

1. Analysis cut-off date
2. Date of consent withdrawal
3. Date of loss to follow-up

then the censoring reason will be 1. 'Ongoing without event', 2. 'Withdrew consent' or 3. 'Lost to follow-up', respectively. However, if this distance is larger than D2 with no event observed, then the censoring reason will be 'Adequate assessment no longer available'.

5.7.6 No baseline tumor assessments

For the PFS analysis, as specified in Table 16-9 in Section 16.3.3.2.10 of the study protocol, since the timing of disease progression cannot be determined for participants with missing baseline tumor assessment, these participants are censored in the PFS analysis at the date of randomization. This rule however only applies to the disease progression component of the PFS assessment, and not to the survival component. Participants without baseline tumor assessments who die within D2 distance (see Section 5.7.5 for definition) of randomization will be counted as having an event in the derivation of PFS at the date of death (Note: all deaths will be counted in the overall survival analysis regardless of presence or absence of the baseline tumor assessment).

5.7.7 Construction of waterfall graphs

Waterfall graphs will be used to depict the anti-tumor activity. These plots will display the best percentage change from baseline in the sum of diameters of all target lesions for each participant. Only participants with measurable disease at baseline will be included in the waterfall graphs.

Special consideration is needed for assessments where the target lesion response is CR, PR or SD, but the appearance of a new lesion or a worsening of non-target lesions results in an overall lesion response of PD. As a conservative approach, such assessments will not be considered for display as bars in the graph, since the percentage change in the sum of diameters of target lesions reflects the non-PD target lesion response, but the overall lesion response is PD. A participant with only such assessments will be represented by a special symbol (e.g. ✶) in the waterfall graph.

Assessments with "not evaluable" target lesion response and assessments with not-evaluable overall response will be excluded from the waterfall plots. Participants without any valid assessments will be completely excluded from the graphs.

The total number of participants displayed in the graph will be shown and this number will be used as the denominator for calculating the percentages of participants with tumor shrinkage and tumor growth. A footnote will explain the reason for excluding some participants (due to absence of any valid assessment).

All possible assessment scenarios are described in Table 5-2.

Table 5-2 Assessments considered for calculation of best percentage change for waterfall graphs

	Criteria for inclusion/exclusion			Possible sources of contradictions	
case	Target response	Overall lesion response	Include in waterfall?	Non-target response	New lesion?
1	CR/PR/SD	PD	Yes but as ✶ only	PD	any
2	CR/PR/SD	PD	Yes but as ✶ only	any	Yes
3	NE	NE or PD	No	any	any
4	CR/PR/SD	NE	No	NE	No
5	CR/PR/SD	CR/PR/SD	Yes as a bar	SD/IR	No
6	PD	PD	Yes as a bar	any	any

6 References

Amit O, Mannino F, Stone AM, et al (2011) Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. Eur. J. Cancer p. 1772-8.

Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41

Cleeland CS (2009) The Brief Pain Inventory. User Guide. Houston, Texas.

Clopper, C.; Pearson, E.S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". Biometrika 26: 404–413

Collett D (1994). Modelling survival data in medical research. London, Chapman & Hall

Committee for Medicinal Products for Human Use (CHMP). Guideline on adjustment for baseline covariates in clinical trials, EMA/CHMP/295050/2013.

Dent S, Zee (2001) application of a new multinomial phase II stopping rule using response and early progression, J Clin Oncol; 19: 785-791

Dodd LE, Korn EL, Freidlin B, et al (2011) An audit strategy for progression-free survival. Biometrics p. 1092-9.

Eisenhauer E, et al (2009) New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). European Journal of Cancer, Vol.45: 228-47

Eton D, et al (2004) A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. Journal of Clinical Epidemiology 2004; 57: 898-910

FDA (2012) FDA briefing Document Oncologic Drugs Advisory Committee Meeting, Evaluation of radiologic review of progression-free survival in Non-hematologic malignancies.

Food and Drug Administration (2005). Guidance for Industry. E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs.

Glimm E, Maurer W, Bretz F (2010) Hierarchical testing of multiple endpoints in group-sequential trials. Stat Med p. 219-28.

Hwang IK, Shih WJ and De Cani JS (1990). Group sequential designs using a family of type 1 error probability spending functions. Statistics in medicine, 9(12), 1439-1445.

Lan KK, DeMets DL (1983) Discrete sequential boundaries for clinical trials. Biometrika; 70: 659-63.

Lee CK, et al (2019) Development and validation of a prognostic nomogram for overall survival in patients with platinum-resistant ovarian cancer treated with chemotherapy. European Journal of Cancer; 117: 99-106.

O'Brien PC and Fleming TR (1979). A multiple testing procedure for clinical trials. Biometrics, 35, 549-53.

Oemar M, Jannsen B. (2013) EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. Version 2.0, October 2013. © EuroQol Group 2013

Pujade-Lauraine E, et al (2014) Bevacizumab Combined With Chemotherapy for Platinum-Resistant Recurrent Ovarian Cancer: The AURELIA Open- Label Randomized Phase III Trial. J Clin Oncol. p. 1302-08.

Zhang JJ, Chen H, He K, Tang S, Justice R, Keegan P, Pazdur R and Sridhara R. Evaluation of blinded independent central review of tumor progression in oncology clinical trials: a meta-analysis. Therapeutic innovation & regulatory science: 2012, 47(2): 167-174.