ImmunoGen, Inc.

Protocol #: IMGN853-0416

MIRASOL: A Randomized, Open-label, Phase 3 Study of Mirvetuximab Soravtansine vs. Investigator's Choice of Chemotherapy in Platinum-Resistant, Advanced High-Grade Epithelial Ovarian, Primary Peritoneal, or Fallopian Tube Cancers with High Folate Receptor-Alpha Expression

Statistical Analysis Plan

Version 3.0
14 February 2023

Prepared by:

[redacted]
Director, Biostatistics
ImmunoGen, Inc.

SPONSOR SIGNATURE PAGE

[redacted]
Director Biostatistics
ImmunoGen, Inc.

Date

[redacted]
Executive Medical Director
ImmunoGen, Inc.

Date

[redacted]
VP Biometrics
ImmunoGen, Inc.

Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or Specialist Term	Explanation
ADA	Anti-drug antibodies
AE	Adverse event
AIBW	Adjusted ideal body weight
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BICR	Blinded Independent Central Review
BOR	Best overall response
BRCA	Breast cancer susceptibility gene
BSA	Body surface area
CA-125	Cancer antigen 125
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel test
CR	Complete response
CTCAE	Common Terminology Criteria for Adverse Events
DOR	Duration of response
ECG	Electrocardiogram
ECHO	Echocardiogram
ECOG PS	Eastern Cooperative Oncology Group Performance Status
eCRF	Electronic case report form
EOC	Epithelial ovarian cancer
EORTC	European Organisation for Research and Treatment of Cancer
FRα	Folate receptor alpha
GCIG	Gynecologic Cancer Intergroup
IC	Investigator's choice/selected standard-of-care chemotherapy
IDMC	Independent Data Monitoring Committee
IMGN	ImmunoGen
INV	Investigator
IRT	Interactive Response Technology

Abbreviation or Specialist Term	Explanation
ITT	Intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MUGA	Multigated acquisition scan
OR	Overall response
ORR	Objective response rate
OS	Overall survival
Pac	Paclitaxel
PD	Progressive disease
PFS	Progression-free survival
PFS2	Time to second disease progression
PK	Pharmacokinetics
PLD	Pegylated liposomal doxorubicin
PR	Partial response
PRO	Patient reported outcomes
PT	Preferred term
Q3W	Every 3 weeks
Q4W	Every 4 weeks
QoLP	Primary endpoint for quality of life
QT	The length of time it takes the electrical system in the heart to repolarize, adjusted for heart rate
QTc	Corrected QT interval
QTcF	Corrected QT using Fridericia's formula
RECIST	Response Evaluation Criteria in Solid Tumors
RMST	Restricted mean survival time
SAE	Serious adverse event
SAP	Statistical analysis plan
SOC	System organ class
TBL	Total bilirubin
TEAE	Treatment-emergent adverse event
Topo	Topotecan
TTR	Time to response

Abbreviation or Specialist Term	Explanation
ULN	Upper limit of normal
WHO Drug	World Health Organization Drug Dictionary

1. INTRODUCTION

1.1. Background

This Statistical Analysis Plan (SAP) is based on Version 2.0 of Protocol IMGN853-0416, dated 04DEC2020.

A brief history of protocol amendments is presented in Table 1.

Table 1: History of Protocol Amendments

Version	Approval Date	Salient Changes (if any)*
Original Protocol	10OCT2019	N/A
Amendment 1	09DEC2019	N/A
Amendment 2	04DEC2020	N/A

*Changes expected to require accommodation in analysis plan.

This SAP will govern the analysis of study data. The plan for the primary efficacy analysis and all other analyses may be modified until the time of interim futility analysis. Any deviations from the SAP will be documented as such in the study report.

2. STUDY DESIGN

2.1. Protocol Objectives

2.1.1. Primary Objective

- To compare the progression-free survival (PFS) of patients randomized to mirvetuximab soravtansine (MIRV) versus Investigator's Choice chemotherapy (IC Chemo), as assessed by the investigator, in patients with high Folate Receptor-α (FRα) level (≥ 75% of tumor staining at ≥ 2+ intensity).

2.1.2. Secondary Objectives

2.1.2.1. Key Secondary Objectives:

- To compare the objective response rate (ORR) of patients randomized to MIRV versus IC Chemo, as assessed by the investigator.
- To compare the overall survival (OS) of patients randomized to MIRV versus IC Chemo.

- To compare the primary patient-reported outcome (PRO) using the European Organization for Research and Treatment of Cancer (EORTC) QLQ-OV28 (Abdominal/GI Symptom Scale) assessment from patients randomized to MIRV versus IC Chemo.

2.1.2.2. Other Secondary Objectives:

- To compare the safety and tolerability of MIRV vs. IC Chemo.
- To compare the duration of response (DOR) in patients randomized to MIRV versus IC Chemo.
- To compare the cancer antigen (CA-125) response rate (RR) per Gynecologic Cancer Intergroup (GCIG) CA-125 criteria in patients randomized to MIRV versus IC Chemo.
- To compare the time to progression or death on the next line of treatment (PFS2) in patients randomized to MIRV versus IC Chemo.

2.1.2.3. Exploratory Objectives

- To assess patient-reported outcomes (PRO) using the EORTC QLQ-C30, EuroQol-5 Dimension 5-level (EQ-5D-5L) and Patient Global Impression of Severity (PGIS) questionnaires.
- To evaluate concentrations of MIRV, total antibody (TAb), DM4 and S-methyl DM4, using sparse sampling.
- To assess the immunogenicity of MIRV via anti-drug antibodies (ADAs).
- To evaluate potential biomarkers in blood and tumor tissue predictive of response to MIRV.

2.1.3. Study Endpoints

2.1.3.1. Primary Endpoint

- PFS: defined as the time from date of randomization until Investigator-assessed progressive disease (PD) or death, whichever occurs first (PFS_{INV}).

2.1.3.2. Secondary Endpoints

2.1.3.2.1. Key Secondary Endpoints:

- ORR: Objective response includes best response of complete response (CR) or partial response (PR) as assessed by the Investigator (ORR_{INV}).
- OS: Overall survival defined as the time from date of randomization until the date of death. Patients alive at the time of analysis will be censored at the last date known to be alive.

- QoLP: Primary PRO assessment, defined as the number of patients achieving at least 15 point absolute improvement at Week 8 or Week 9 in the abdominal/GI scale of EORTC QLQ-OV28.

2.1.3.2.2. Other Secondary Endpoints:

- Treatment-emergent adverse events (TEAEs) and laboratory test results, physical examination, or vital signs.
- DOR: Duration of response, defined as the time from initial response until Investigator-assessed PD for all patients who achieve a confirmed objective response (PR or CR).
- CA-125 response determined using the GCIG criteria defined in Appendix C of the Protocol. CA-125 response per GCIG criteria will be determined programmatically.
- PFS2 defined as the time from date of randomization until second disease progression or death whichever occurs first.

2.1.3.3. Exploratory Endpoints

- PRO from EORTC QLQ-C30/OV-28, EQ5D-5L and PGIS .
  - Analyses of PRO endpoints will be covered in a separate PRO SAP.
- Immunogenicity is defined as the presence of ADA to MIRV.
- Soluble FRα levels:

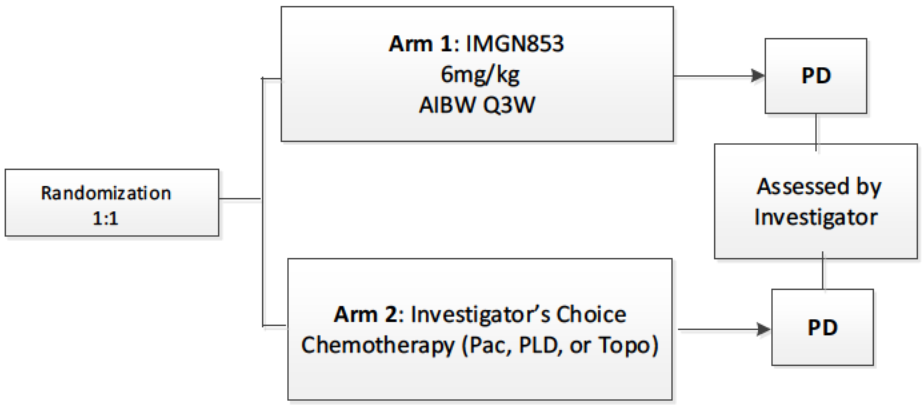
2.2. Study Overview

This study is designed to compare the efficacy of MIRV to that of IC Chemo in women with platinum resistant advanced epithelial ovarian cancer (EOC), primary peritoneal cancer, or fallopian tube cancer.

Patients will be stratified by the number of prior lines of treatment (1 vs 2 vs 3) and IC chemotherapy (paclitaxel [Pac] vs pegylated liposomal doxorubicin (PLD) vs topotecan [Topo]) which is chosen by the investigator prior to randomization. Patients will be randomized 1:1 into 1 of 2 arms as follows (Figure 1):

- ***Arm 1:*** IMGN853 6mg/kg AIBW Q3W.
- ***Arm 2:*** IC chemotherapy (weekly paclitaxel every 4 weeks [Q4W], PLD administered once Q4W, or topotecan administered either on Days 1, 8, and 15 Q4W or for 5 consecutive days Q3W).

Figure 1: Study Design Schema



AIBW = adjusted ideal body weight; Pac = paclitaxel; PD = disease progression; PLD = pegylated liposomal doxorubicin; Q3W = every 3 weeks; Q4W = every 4 weeks; Topo= topotecan.

Patients will continue to receive study drug until disease progression, unacceptable toxicity, withdrawal of consent, death, or until the Sponsor terminates the study (whichever comes first).

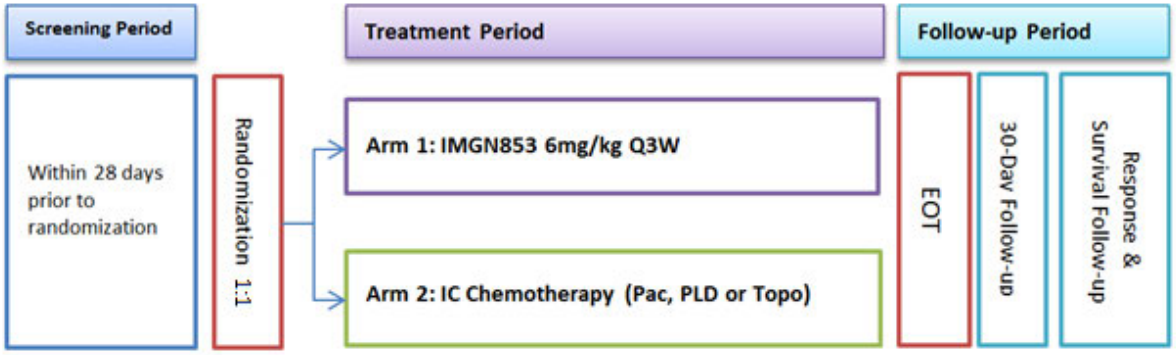
Tumor assessments, including radiological assessments by CT/MRI scans will be performed at Screening and subsequently every 6 weeks (± 1 week) from C1D1 (for all regimens) for the first 36 weeks then every 12 weeks (± 3 weeks) until disease progression, death, the initiation of subsequent anticancer therapy, or patient's withdrawal of consent, whichever occurs first.

Patients who discontinue study drug for reasons other than radiographic disease progression will continue with tumor assessments until documentation of disease progression per RECIST 1.1 or the start of new anticancer therapy, whichever comes first. Prior to Week 36 (from Cycle 1, Day 1), assessments should occur every 6 weeks (± 1 week) as allowed by local requirements but must occur at an interval of no more than 12 weeks. After Week 36, assessment will occur every 12 weeks (± 3 weeks) until documentation of PD, death, the start of new anticancer therapy or patient's withdrawal of consent (whichever comes first).

All patients who discontinue study drug will be followed every 3 months (±1 month) until death, lost to follow-up, withdrawal of consent for survival follow-up, or EOS, whichever comes first. All patients will be followed for PFS2 and survival until 300 deaths have occurred.

The study consists of a Screening period, a Treatment period, an End-of-Treatment visit, and a Follow-up period (see Figure 2).

Figure 2: Study Period Schema



EOT = end of treatment; Pac = paclitaxel; PLD = pegylated liposomal doxorubicin; Q3W = every 3 weeks; Topo = topotecan.

2.2.1. Study Population

All patients must have 1 of the following pathologically documented, definitively diagnosed tumor types: advanced platinum resistant EOC, primary peritoneal cancer, or fallopian tube cancer. Patients must also have the following:

- At least 1 lesion that meets the definition of measurable disease according to RECIST Version 1.1 (radiologically measured by the Investigator).
- Received at least 1 but no more than 3 prior systemic lines of anticancer therapy and for whom single-agent therapy is appropriate as the next line of treatment.
- Confirmation of FRα positivity as defined by the Ventana FOLR1 (FOLR-2.1) CDx assay in archival or fresh biopsy tumor sample.

See Section 3.1 of the protocol for a complete list of the inclusion/exclusion criteria.

Patients who have consented to the study and have been randomized are considered enrolled. Patients who are issued a patient number, but who do not successfully complete the screening process and who do not get randomized will be considered screen failures. Patient numbers assigned to patients who screen fail will not be re-issued.

2.2.2. Power and Sample Size

2.2.2.1. Primary Endpoint

The primary endpoint is PFS as per investigator assessment. The study is designed to test the null hypothesis that the survival function for PFS is the same between the MIRV arm and the IC chemotherapy arm versus the alternative hypothesis that the survival function for PFS is different between the MIRV and IC chemotherapy arms.

Approximately 430 patients will be randomized 1:1 (approximately 215 patients each in the MIRV and IC chemotherapy arms, respectively) over a period of approximately 18 months. The final analysis will be conducted when at least 330 PFS events have been observed.

The study will have 90% power to detect a hazard ratio of 0.7. An interim futility analysis will be conducted when at least 110 PFS events have been observed. The study may be terminated for

futility at interim analysis if the observed hazard ratio [MIRV to IC chemotherapy] is greater than 1. No alpha-spending is needed for this futility only analysis.

Sample size and power were determined using R-Project package gsDesign version 2 with the following assumptions:

- Median PFS for the IC arm is 3.5 months.
- Median PFS for the MIRV arm is 5.0 months.
- Exponential distribution for the event process.
- Overall attrition rate is approximately 13% in both arms.
- Uniform enrollment over an 18-month period.
- Duration of follow up is 6 months.
- Median OS for the IC Chemo arm is 12 months.
- Median OS for the MIRV arm is 17.5 months.

2.2.2.2. Power for Key Secondary Endpoints

Only if the primary endpoint of PFS is statistically significant, a hierarchical testing procedure will be used to control the study-wise error rate (SWER) for key secondary endpoints of ORR, OS, and the primary PRO, in that order.

ORR as Assessed by Investigator

With a sample size of 430 (215 in each arm), the study will have a power of 91.8% to detect a 14% difference in ORR (30% in MIRV vs 16% in IC chemotherapy).

OS

The final analysis of OS will be conducted when at least 300 events have occurred. There will be one interim analysis for OS at the time of final analysis of PFS at which time approximately 180 (60%) deaths will have been observed. A Lan-DeMets alpha-spending function using an O'Brien-Fleming stopping boundary is used to control overall Type I error for OS at 2-sided alpha level of 0.05. The exact boundary at the interim OS analysis will be determined based on the actual number of death events. The study will have 90% power to detect an OS hazard ratio of 0.6857 and 80% power to detect an OS hazard ratio of 0.7235. The OS alpha-spending is calculated using R-Project package gsDesign version 2.

Primary PRO Endpoint

Primary PRO endpoint is proportion of patients achieving at least a 15-point (or equivalently, 15%) improvement on the QLQ-OV28 abdominal/gastrointestinal symptom subscale [Items 31-36] at the Week 8/9 assessment. This endpoint will be analyzed using a responder approach. Assuming a 90% compliance rate in the PRO endpoint, the study will have a power of 86% to detect a 15% difference in response rate for the primary PRO endpoint (25% in MIRV versus 10% in IC chemotherapy).

2.2.3. Treatment Randomization and Stratification

The treatment randomization schedule was developed by 4G Clinical. Patients will be randomized 1:1 into 2 groups, as follows:

- ***Arm 1:*** MIRV six mg/kg AIBW Q3W.
- ***Arm 2:*** IC chemotherapy (weekly Pac Q4W, PLD Q4W, or Topo administered either on Days 1, 8, and 15 Q4W or for 5 consecutive days Q3W.

Randomization will be stratified as follows:

- Number of prior lines of therapy (1 vs 2 vs 3).
- IC chemotherapy (Pac, PLD, or Topo) choice prior to randomization.

The required FRα expression level is ≥ 75% of tumor staining at ≥ 2+ intensity.

2.2.4. Blinding and the BICR

This is an open-label study. Treatment assignment will not be blinded because each study drug has a unique dosing schedule. A blinded independent central review (BICR) will be used to provide an independent assessment of radiographic tumor assessments. The BICR will be blinded to the treatment assignment.

The Sponsor has also decided to blind certain members of the study team to the efficacy endpoints. Specifically, ImmunoGen statisticians and statistical programmers will not have access to the RECIST tumor assessments for either the Investigator's assessment captured in the EDC system or the BICR data.

Please see the Data Access Plan for this study for detailed information regarding the blinding of efficacy endpoints.

2.2.4.1. BICR

Copies of all imaging scans must be obtained and sent to a central imaging vendor designated by ImmunoGen as outlined in the Imaging Manual. The central imaging vendor will assess the quality of the images. The imaging vendor will be responsible for the formation and management of the BICR.

Tumor response will be assessed by the Investigator and by the BICR using RECIST Version 1.1. Response as determined by the Investigator will be recorded in the electronic case report forms (eCRFs). The BICR assessment will be used as sensitivity analysis.

2.2.5. Assessment Schedule

See protocol Table 2, Table 3, and Table 4 for study schedules of assessments.

2.3. Interventions

2.3.1. Clinical Trial Material

The dose levels and dosing schedule is provided in Table 2.

Table 2: Dosing and Dosing Schedule for Phase 3 Portion

Phase 3 (Randomization 2:1)			
Group	**Drug**	**Dose**	**Dosing Schedule**
Arm 1	IMGN853	6 mg/kg AIBW	Day 1 of a 3-week cycle
Arm 2	Paclitaxel	80 mg/m^2	Days 1, 8, 15, and 22 of a 4-week cycle
	PLD	40 mg/m^2	Day 1 of a 4-week cycle
	Topotecan (4-week cycle)	4 mg/m^2	Days 1, 8, and 15 of a 4-week cycle
	Topotecan (3-week cycle)	1.25 mg/m^2	Days 1 through 5 of a 3-week cycle

AIBW = adjusted ideal body weight; PLD = pegylated liposomal doxorubicin.

The protocol provides additional details in Section 4.1.2.

3. GENERAL ANALYTICAL CONSIDERATIONS

3.1. Data Sources

Data are recorded on eCRFs. Central laboratory data will be provided via electronic data transfers. Section 12 of the protocol provides additional details regarding data recording and handling. Randomization is completed through an Interactive Response Technology (IRT) system.

3.2. Definition of Baseline

Study Day 1 (ie, Cycle 1 Day 1) will be designated as the first day a patient receives study drug. The baseline value is defined as the last non-missing value on or before the date of first dose of study drug.

3.3. Missing Data

Partial dates are allowed on the eCRF for prior anti-cancer systemic therapy start and stop dates, subsequent anti-cancer therapy start date, adverse event (AE) onset and resolution dates, concomitant medication start and stop dates, and concomitant procedure dates. An entry for the year is required in the eCRF system for each of these dates. Only the month and day may be entered as unknown. Dates from these forms will be reported in listings as collected. Every effort will be made to query missing dates.

For records with missing AE onset date, the following procedure will be employed for use in determining whether the AE is treatment emergent:

- AE onset dates with missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for AEs occurring in the first month of dosing, in which case the date will be the first day of dosing.
- AE onset dates with missing month will be assumed to occur on the first day of the non-missing year (ie, January 1), except for AEs occurring in the first year of dosing, in which case the date will be the first day of dosing.

For records with a missing medication start and/or stop date, the following procedure will be employed for use in determining whether the medication is prior or concomitant:

- Medication start dates with a missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for medications occurring in the first month of dosing, in which case the date will be the first day of dosing.
- Medication start dates with missing month will be assumed to occur on the first day of the non-missing year (ie, January 1), except for medications occurring in the first year of dosing, in which case the date will be the first day of dosing.
- Medications that are not ongoing and have a medication stop date with a missing day and non-missing month will be assumed to occur on the last day of the non-missing month.
- Medications that are not ongoing and have a medication stop date with a missing month will be assumed to occur on the last day of the non-missing year (ie, December 31).

For records with a missing procedure date, the following procedure will be employed for use in determining whether the procedure is prior or concomitant:

- Procedure dates with a missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for procedures occurring in the first month of dosing, in which case the date will be the first day of dosing.
- Procedure dates with missing month will be assumed to occur on the first day of the non-missing year (ie, January 1), except for procedures occurring in the first year of dosing, in which case the date will be the first day of dosing.

For records with a missing prior anti-cancer systemic therapy start and/or stop date, and/or disease progression date from prior therapy, the following procedure will be employed for use in calculating the platinum-free interval (PFI):

- If the stop date of prior platinum therapy is completely missing OR only the year is available, the PFI will be set as missing.
- If both year and month of stop date of prior platinum therapy are available and only day is missing, the stop date will be imputed by the first day of the month.
- If the disease progression date from prior platinum therapy is completely missing OR only the year is available, the PFI will be set as missing.
- If both year and month of disease progression date from prior platinum therapy are available and only day is missing, the date of disease progression will be imputed by the last day of the month.

For records with a missing subsequent anti-cancer therapy start date, the following procedure will be employed for use in deriving efficacy variables such as PFS and best overall response (BOR).

- The start date of subsequent anti-cancer therapy will be imputed by the earliest possible date that is not contradicting with any other available data in the database.

For records with a partial death date, the following procedure will be employed for use in deriving time to event variables (PFS, OS, DOR, PFS2).

- The death date will be imputed by the earliest possible date that is not contradicting with any other available data in the database.

All other data will be reported as they are collected. No imputation methods will be used to replace missing data unless otherwise stated in this document.

3.4. Multiple Assessments for the Same Assessment Time Point

In the case of multiple observations at a specific visit, the first non-missing measurement will be used for analysis, unless multiple study assessments are expected (eg, pre-dose vs post-dose). When multiple study assessments are expected, the first non-missing measurement for the visit and assessment time point will be used for analysis.

3.5. Multiple Study Centers

No adjustment for study center is planned.

3.6. Covariate Adjustment in Primary Analysis

A stratified analysis using randomization stratification factors will used for primary inference. The details are described in Section 5.2.

3.7. Sample Size Reassessment

Not applicable.

3.8. Interim Analyses

For the primary endpoint of PFS by investigator, the final analysis will be conducted when at least 330 events have been observed. An interim futility analysis was designed to be conducted when at least 110 events have been observed. The study would be terminated for futility at interim analysis if the observed hazard ratio [MIRV to IC chemotherapy] was greater than 1. No alpha spending was planned for this futility analysis. This interim futility analysis (IA1) was conducted with 131 PFS events (data cut 31 January 2022). Independent Data Monitoring Committee (IDMC) reviewed PFS results in a closed session, confirmed the futility boundary was not hit and recommended the trial to continue without modifications.

Besides PFS, the interim futility analysis also included ORR and DOR analyses using the same method described in this document per FDA requests on 12 April 2022. No separate statistical analysis plan was prepared for the interim analysis.

For the secondary endpoint of OS, the final analysis will be conducted when at least 300 deaths have been observed. There will be one interim analysis for OS at the time of final analysis of PFS at which time approximately 180 (60%) deaths will have been observed. A Lan-DeMets alpha-spending function using an O'Brien-Fleming stopping boundary is used to control overall Type I error for OS at 2-sided alpha level of 0.05. The exact boundary at the interim OS analysis will be determined based on the actual number of death events.

Additionally, there will be ongoing analysis of safety data by the Independent Data Monitoring Committee (IDMC) as described in the IDMC charter.

3.9. Timing of Final Analysis

If the study continues to a minimum of 430 patients for full enrollment (ie, the study is not stopped for futility at the interim analysis), the final analysis for PFS will be conducted when at least 330 PFS events have been observed (as determined by the investigator).

3.10. Hypothesis Testing and Multiple Comparisons

3.10.1. Primary Endpoint

The null hypothesis for primary endpoint of PFS by investigator is:

- H_{01}: the survival function for PFS_{INV} is the same between the MIRV arm and the IC chemotherapy arm.

And the alternative hypothesis is:

- H_{a1}: the survival function for PFS_{INV} is different between the MIRV arm and the IC chemotherapy arm.

For a single primary endpoint, there is no need to adjust for multiple comparison.

3.10.2. Key Secondary Endpoints

For the 3 key secondary endpoints, a hierarchical testing procedure will be applied to control the study-wise Type I error only if the null hypothesis for the primary endpoint is rejected at 2-sided α-level of 0.05. The order of the hierarchical testing is as follows:

- ORR_{INV}.
- OS.
- Primary endpoint for quality of life (QoLP): number of patients achieving at least a 15% (or equivalently, 15-point) improvement on the QLQ-OV28 abdominal/gastrointestinal symptom subscale [Items 31-36] at the Week 8/9 assessment.

The null hypothesis for ORR_{INV} is as follows:

- H_{02}: the ORR_{INV} is the same between the MIRV arm and the IC Chemo arm.

And the alternative hypothesis is as follows:

- H_{a2}: the ORR_{INV} is different between MIRV arm and the IC Chemo arm.

The null hypothesis for OS is as follows:

- H_{03}: the survival function for OS is the same between the MIRV arm and the IC Chemo arm.

And the alternative hypothesis is as follows:

- H_{a3}: the survival function for OS is different between the MIRV arm and the IC Chemo arm.

The null hypothesis for QoLP is as follows:

- H_{04}: the QoLP is the same between the MIRV arm and the IC Chemo arm.

And the alternative hypothesis is as follows:

- H_{a4}: the QoLP is different between the MIRV arm and the IC Chemo arm.

Following this hierarchical testing procedure, if H_{03} is rejected favoring the MIRV arm at the final analysis for PFS, statistical significance will be claimed for OS and no further hypothesis testing is necessary for OS. In this case, however, an updated analysis of OS may still be conducted after 300 death events have occurred to provide updated estimate of treatment effect for OS.

3.11. Analysis Populations

Eight analysis populations will be defined for use with various analyses. Table 3 illustrates the relationship between each population and the analyses for which the data from the population will be used.

Table 3: Analysis Populations

Analysis Population	Analysis					
	Reason for Screen Failure	Baseline	Patient Disposition	Efficacy	Safety	PK
Screened	X					
ITT		X	X	X		
Per Protocol				X		
Response Evaluable				ORR only		
CA-125 Evaluable				CA-125 only		
Safety					X	
PK						X
LPP				QoLP		

CA-125=cancer antigen 125; ITT=intent-to-treat; LPP= Longitudinal Period Population; ORR=objective response rate; PK=pharmacokinetics; QoLP=Primary endpoint for quality of life.

3.11.1. Screened

The Screened population includes all patients who have signed an informed consent.

3.11.2. ITT Population

The intent-to-treat (ITT) population is defined as all patients randomized to the study, regardless of whether or not patients received study treatment (MIRV or IC Chemo).

Patients will be analyzed based on the randomized treatment.

3.11.3. Per Protocol Population

Per Protocol Population is defined as all patients randomized to the study who have received at least 1 dose of MIRV or IC Chemo, excluding patients with protocol deviations in the following categories: patients -

- whose tumors do not have FRα positivity as defined by the Ventana FOLR1 (FOLR-2.1) CDx assay high expression, and/or
- who have received four (4) or more prior lines of anticancer systemic therapy, and/or
- who do not meet the eligibility criterion of being platinum resistant, and/or
- who meet the protocol exclusion criterion of being primary platinum refractory.
- Or other critical/major protocol deviations that deemed affecting primary and/or key secondary objectives. The final list of critical/major protocol deviations leading to the removal of patients from the per protocol population will be finalized and signed off before the database lock for the final PFS analysis.

Patient will be analyzed based on the randomized treatment.

3.11.4. Response-Evaluable Population

The Response-Evaluable population is defined as all patients randomized to the study who have received at least 1 dose of MIRV or IC Chemo and have measurable disease at baseline per investigator or BICR.

Patient will be analyzed based on the randomized treatment.

3.11.5. CA-125-Evaluable Population

The CA-125-Evaluable population is defined as all patients randomized to the study who have received at least 1 dose of MIRV or IC Chemo, whose pretreatment CA-125 is ≥ 2.0 times the upper limit of normal (ULN), within 2 weeks prior to randomization, and who have at least 1 post-baseline CA-125 evaluation.

Patient will be analyzed based on the randomized treatment.

3.11.6. Safety Population

All randomized patients who received at least 1 dose of MIRV or IC Chemo will be included in the Safety population.

Patient will be analyzed based on the actual treatment received.

3.11.7. Pharmacokinetic (PK) Analysis Population

All randomized patients who received at least 1 dose of MIRV and who have at least one non-missing PK concentration data point will be included in the PK analysis population.

3.11.8. Longitudinal Period Population (LPP)

Longitudinal Period Population (LPP) defined as all randomized patients who have survived and maintained in the study from randomization through week 8/9 assessment and have available QoL data for both baseline and week 8/9.

3.12. Data Display Characteristics

Data displays produced for this study will include summary tables, data listings, and figures.

Data listings will report the data recorded on the eCRF or derived for each patient. Data will be ordered by treatment, patient number, and date/time of assessment. When expedient, additional levels of ordering hierarchy may reflect subsets of assessments within patient. Data listings will not display patient initials.

Summary tables will display summary statistics calculated for each of the treatment groups. For safety analyses, the columns of the summary tables will be MIRV, IC Chemo total, Pac, PLD, Topo. For efficacy analyses, the columns of the summary tables will be MIRV and IC chemo.

The summary statistics displayed will be a function of the type of data associated with the summarized assessment. Unless stated otherwise in relevant sections, continuous data will be summarized with the number of non-missing values, mean, standard deviation, minimum, median, and maximum. Categorical data will be summarized with the number of non-missing values and the numbers of values equal to each of the possible values. Percentages of patients with each of the possible values will be calculated from the number of patients in the corresponding analysis population, unless stated otherwise. Some continuous variables may also be grouped into categorical levels and evaluated in frequency tables.

4. PATIENT ACCOUNTABILITY

4.1. Patient Characteristics

Patient characteristics will be summarized and listed for the ITT population by the treatment to which they were randomized (planned treatment).

4.1.1. Demography

Data collected about the following patient characteristics at the screening visit will be summarized as follows:

- Age is collected in the IRT system at the time of patient registration. As this is a global study, there are certain regions where local regulations prohibit the collection of a complete date of birth. Therefore, age will not be recalculated for analysis purposes. The collected age will be used for summarization.
- Sex.
  - Childbearing potential (female only, yes/no).
- Ethnicity.
- Race.

- Region.
- Country.

All demographic data, including informed consent date, will be listed.

4.1.2. Height and Baseline Weight, AIBW, and Body Surface Area

Height and baseline weight, AIBW, and body surface area (BSA) will be summarized by treatment group and presented in a listing.

4.1.3. Medical History

Medical history will be coded using Medical Dictionary for Regulatory Affairs (MedDRA) (Version 24) and graded using NCI-CTCAE (version 5.0), associating lower-level terms with preferred terms (PTs) and system organ classes (SOCs) by the primary hierarchy. Medical histories will be summarized as the number and percentage of patients who reported at least 1 medical history event; and number and percentage of patients who reported at least 1 medical history event in each SOC. Within each SOC, tables will display the number and percentage of patients reporting at least 1 medical history event as designated by PT. All medical history information will also be listed.

4.1.4. Disease Characteristics, Prior Therapy, and Gene Mutations

Listings of all collected data related to disease characteristics and prior therapy will be provided. A summary of the following elements will also be provided:

- Eastern Cooperative Oncology Group Performance Status (ECOG) performance status
- Primary diagnosis (epithelial ovarian cancer, fallopian tube cancer, primary peritoneal cancer, other)
- FRα expression (percent staining at intensity 2+)
- Histology
- Disease stage at initial diagnosis
- Histologic grade
- Prior radiotherapy .
- Prior systemic therapy
  - The number of prior systemic therapies is defined as the number of distinct treatment regimens.
- Prior cancer related surgery.
- Prior exposure to bevacizumab
- Prior exposure to doxorubicin/PLD
- Prior exposure to topotecan

- Prior exposure to taxanes
- Prior exposure to PARP inhibitors
  - If patient participated in a double-blind randomized trial comparing PARP inhibitor to placebo and the actual treatment was unknown, patient will be summarized under 'uncertain' category.
- Any breast cancer susceptibility gene (*BRCA)* mutations
- Time since initial diagnosis (months)
- Primary platinum-free interval (months)
  - Defined as time from last dose of 1st line platinum therapy to the date of disease progression and/or relapse following 1st line therapy.
- Most recent Platinum-free interval
  - Defined as time from last dose of latest line platinum therapy to the date of disease progression and/or relapse following that line of therapy.

4.2. Patient Disposition

4.2.1. Screened and Enrolled Patients

The number and percentage of patients who were screened, enrolled (randomized), not enrolled (screen failure), and reason for screen failure will be summarized overall, and by geographic region (North America, Europe, middle east and Africa, APAC), country, and site.

4.2.2. ITT, Safety Disposition

A summary of patient disposition will summarize, for the ITT population, the number of patients randomized and the reason for treatment and study discontinuation. This summary will also be produced for the Safety population if more than 5% of patients in the ITT population are excluded from the Safety population.

Percentage of patients who withdrew for each reason on the End-of-Treatment and End-of-Study forms will be calculated using all members of the relevant population in the relevant treatment group for the denominator.

The number and percentage of ITT patients included in each analysis population defined in Section 3.12 will be presented for each treatment group. This table will be produced for all patients at all sites, pooled.

4.2.3. Protocol Deviations and Population Inclusions

Protocol deviations will be captured in a protocol deviation log. A listing of patients with any protocol deviation will be provided for the ITT population.

5. EFFICACY ANALYSES

The PFS and OS efficacy analyses will use data from the ITT population. Patients will be analyzed by the treatment to which they were randomized (planned treatment).

5.1. Efficacy Outcomes

5.1.1. PFS

PFS is defined as the time from the date of randomization until the date of PD or death from any cause, whichever occurs first. PFS is defined based on radiological assessments and determined by the Investigator or the BICR (depending on the analysis). Clinical progression is not considered a progression endpoint.

Table 4 summarizes the rules to be used for PFS. When an analysis cutoff date is implemented, only data (deaths and radiological assessments) occurring on or prior to the cutoff date will be used for analysis.

Table 4: PFS Definitions

Situation	Date of PFS Event or Censoring	Outcome
No baseline tumor assessments or post-baseline radiological assessments, and patient did not die within 105 days of randomization	Date of randomization	Censored
No baseline tumor assessments or post-baseline radiological assessments, and patient died within 105 days of randomization	Date of death	Death
Death	Date of death	Death
Radiological Progression	Date of first radiological assessment indicating progression (i.e., OR = PD).	Progression
New anti-cancer therapy prior to PD or death (including palliative radiotherapy during study treatment)	Date of last radiological assessment prior to the start of the new anticancer therapy	Censored
No death or PD	Date of last radiological assessment	Censored
PD or death after missing 2 or more consecutive radiological assessments (PD or death date - last radiological assessment date +1 ≥ 105 days or >= 231 days if the assessment schedule has changed from every 6 weeks [+-1 week] to every 12 weeks [+- 3 weeks] after Week 36 or after discontinuing from study treatment per local requirements)	Date of last adequate radiological assessment showing no PD	Censored

Includes radiographic progression only.
PFS_{INV} is based on the Investigator's assessment; PFS_{BICR} is based on BICR assessment.
OR = overall response; PD = disease progression; PFS = progression-free survival.

5.1.2. Best Overall Response

Best overall response (BOR) for a patient is the best response designation (as assessed by either the Investigator or the BICR, depending on the analysis) recorded between the date of randomization and the date of objectively documented PD per RECIST Version 1.1, the date of the start of new anti-cancer therapy, or the date of study discontinuation, whichever occurs first. When an analysis cutoff date is implemented, only radiological assessments occurring on or prior to the cutoff date will be used for analysis. Patients with an overall response of CR or PR must

have a repeat tumor assessment performed no less than 4 weeks (28 days) after the criteria for response are first met. When stable disease is the best overall response, it must meet the minimum duration of 35 days (6 weeks – 1 week window = 35 days from the date of first dose). The confirmatory scan is valid following treatment discontinuation as long as the patient has not started a new anti-cancer therapy.

5.1.2.1. ORR

The ORR will be calculated as proportion of patients with a BOR of CR or PR. Patients without post-baseline RECIST assessment will be treated as non-responders (ie, these patients will contribute to the denominator, but not the numerator).

5.1.2.2. DOR

DOR is defined as the time from the date of the first response (CR or PR), whichever occurs first, to the date of PD or death from any cause, whichever occurs first. DOR is only defined for patients who have a BOR of CR or PR.

Per the BOR definition, patients with an overall response of CR or PR must have a repeat tumor assessment performed no less than 4 weeks after the criteria for response are first met. The first date at which a CR or PR response was noted will be used to calculate DOR, not the date of the confirmatory tumor assessment. DOR end dates and censoring rules are the same as the PFS in Table 4 above.

5.1.2.3. Time to Response

Time to response (TTR) is defined as the time from the date of randomization until the date of the first observed CR or PR. TTR is only defined for patients who have a BOR of CR or PR.

5.1.3. OS

OS is defined as the time from the date of randomization until the date of death from any cause.

Patients who are alive or lost to follow-up at the analysis are censored at the last known date at which they were known to be alive. When an analysis cutoff date is implemented, only deaths occurring on or before the cutoff date are counted as OS events. Patients without death events will be censored on the latest date with a known alive status on or before the cutoff date.

5.1.4. CA-125 Response

A CA-125 response is defined as a ≥ 50% reduction in CA-125 levels from baseline. The response must be confirmed and maintained for at least 28 days. The CA-125 response will be conducted using the CA-125-Evaluable population. The date of response corresponds to the date when the CA-125 level is first reduced by 50%. The summary table for CA-125 will include the number (percentage) of patients in the CA-125-Evaluable population, and that sample size will then be used as the denominator for CA-125 response rate.

5.1.5. RECIST-Related Endpoints –Investigator's or BICR Assessment

The primary analyses will use endpoints based on the investigator's assessment of response. In the event that the primary endpoint of PFS by INV is positive, the BICR assessment of response

will be used for sensitivity analyses. A subscript of INV (Investigator) or BICR will indicate which assessment of response is used for the calculation of the endpoint. For example, BOR_{INV} and PFS_{INV} will be used to denote the BOR and PFS based on the Investigator's assessment of response; BOR_{BICR} and PFS_{BICR} will be used to denote BOR and PFS based on the BICR assessment of response.

5.1.6. PFS2

PFS2 is defined as time from randomization to second PD or death. Specifically, for the purposes of the PFS2 analysis, the date of event was defined as the following, whichever occurred first:

- Date the patient experienced an event of radiological or clinical progression reported on the response follow-up forms.
- Date of death.
- Date of end of next-line treatment.
  - For next-line treatment, any therapy other than radiotherapy or surgery will be considered.
- Any event of PD reported during the survival follow-up evaluation will be considered, including PD in patients who did not receive subsequent next-line treatment.
- Patients who did not experience PD by INV during the study period but who had PD documented on the response follow-up form will also be considered as having an event of PFS2.
- For patients who did not receive subsequent therapy, nor experience PD or death during any follow-up, the data will be censored at the time of the last follow-up contact.
- Data from patients lost to follow-up after at least 1 follow-up assessment will be included in the analysis as censored observations on the date the patient was last known to be alive.
- Patients who did not have follow-up contacts will be censored on the date of study discontinuation.

5.2. Primary Efficacy Outcome Analysis

The protocol specifies the following primary efficacy endpoint:

- PFS as assessed by the investigator (PFS_{INV}).

The distribution of PFS_{INV} will be summarized using the Kaplan-Meier method. PFS rates will be reported at 3-month intervals (eg, 3 months, 6 months, etc.). Median times will be estimated for each treatment from the 50^{th} percentile of the corresponding Kaplan-Meier estimates. 95% confidence intervals (CI) for the 3-month intervals and median times will also be provided. The primary comparison between treatments will use the log-rank test stratified by the randomization stratification factors. Stratified analysis using values collected on the eCRFs will be conducted as

sensitivity analysis. As a sensitivity analysis, the results from an unstratified analysis will also be provided. The chi-square p-values from the log-rank tests will be reported.

Additionally, the restricted mean survival time (RMST) for PFS_{INV} will also be reported at 3-month intervals (eg, 3 months, 6 months, etc.). The RMST at 12 months will be compared between treatments stratified by the randomization stratification factors using SAS PROC RMSTREG. As a sensitivity analysis, the RMST from an unstratified analysis will also be provided.

In stratified analysis, the strata will be IC Chemo (Pac vs PLD vs Topo) and the number of prior lines of therapy (1 vs 2 vs 3).

The hazard ratio for PFS_{INV} treatment comparisons will be estimated using a stratified Cox proportional hazards model. As a sensitivity analysis, the hazard ratio from an unstratified Cox proportional hazards model will also be provided. The Supremum test will be conducted to test for proportional hazards assumption.

The hazard ratio (MIRV arm vs IC Chemo arm) will be reported using the maximum likelihood estimate along with 95% CI. The IC Chemo arm will be used as the reference treatment (ie, denominator of the hazard ratio). A time-to-event hazard ratio less than 1.0 would indicate a MIRV benefit over IC Chemo.

As a sensitivity analysis, PFS_{INV} will also be analyzed in the Per-protocol population. Patients will be analyzed by the actual treatment they received in this sensitivity analysis.

5.3. Secondary Efficacy Analyses

5.3.1. Key Secondary Efficacy Analyses

The protocol describes the following key secondary efficacy endpoints:

- ORR per RECIST Version 1.1 criteria based on the investigator's assessment (ORR_{INV}).
- OS.
- Primary PRO endpoint (QoLP): Number of patients achieving at least a 15% (or equivalently, 15-point) improvement on the QLQ-OV28 abdominal/gastrointestinal symptom subscale [Items 31-36] at the Week 8/9 assessment).

Only if the primary endpoint of PFS by INV is positive, a hierarchical testing procedure will be applied to key secondary endpoints (in the order listed above) to control the study-wise Type I error.

Unless noted otherwise, all analyses of key secondary endpoints will be based on the ITT population. In addition, ORR will also be summarized using the Response-Evaluable population and QoLP will be summarized using the LPP population.

As a sensitivity analysis, ORR and OS will also be summarized on the Per-protocol population.

ORR will be analyzed using a Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors to test for differences between the MIRV arm and the IC

Chemo arm. P-value and 95% CIs for ORR will be provided. Odds ratio and its 95% CI will also be reported.

The distribution of OS will be summarized using the Kaplan-Meier method. OS rates will be reported at 3-month intervals (eg, 3 months, 6 months, etc.). Median OS will be estimated for each treatment from the 50^{th} percentile of the corresponding Kaplan-Meier estimates. 95% CIs for the 3-month intervals and median OS will also be provided. The primary comparison between treatments will use the log-rank test stratified by the randomization stratification factors. Stratified analysis using values collected on the eCRFs will be conducted as sensitivity analysis. As a sensitivity analysis, the results from an unstratified analysis will also be provided. The chi-square p-values from the log-rank tests will be reported.

The hazard ratio for the OS treatment comparisons will be estimated using a stratified Cox proportional hazards model. As a sensitivity analysis, the hazard ratio from an unstratified Cox proportional hazards model will also be provided. The Supremum test will be conducted to test for proportional hazards assumption.

The hazard ratio (MIRV arm vs IC Chemo arm) will be reported using the maximum likelihood estimate along with 95% CI. The IC Chemo arm will be used as the reference treatment (ie, denominator of the hazard ratio). A hazard ratio less than 1.0 would indicate a MIRV benefit over IC Chemo.

The median follow up time and its 95% CI will be estimated using reverse Kaplan-Meier method on OS.

The number of IC Chemo patients who received MIRV post study treatment will be closely monitored. Additional sensitivity analyses to adjust for bias due to crossover (e.g., two-stage estimation, inverse probability censoring weighting) will be evaluated and conducted, if deemed necessary.

For QoLP analysis, patients will be classified into the following categories based on their QLQ-OV28 abdominal/GI symptom subscale scores:

- If a patient's score on the QLQ-OV28 abdominal/GI symptom subscale increases by 15 points or greater from baseline to week 8/9 assessment, the patient will be categorized "Improved".
- If a patient's score on the QLQ-OV28 abdominal/GI symptom subscale increases by less than 15 points from baseline to week 8/9 assessment, the patient will be categorized "Unimproved".
- If a patient died or progressed prior to week 8/9 assessment, the patient will be categorized "Unimproved".

A Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors to test for differences between the MIRV arm and the IC Chemo arm will be performed. All two-sided P-values from these tests will be reported.

5.3.2. Other Secondary Efficacy Analyses

Additionally, the following efficacy endpoints will also be analyzed:

- DOR based on Investigator's assessment (DOR_{INV}).
- TTR based on Investigator's assessment (TTR_{INV}).
- GCIG CA-125 criteria clinical response rate.

Because there is no procedure in place to control the Type I error on other secondary efficacy endpoints, all p-values for treatment comparison on other secondary efficacy endpoints will be for information only and will be considered as nominal.

GCIG CA-125 response will be summarized using the CA-125 Response-Evaluable population only.

DOR and TTR will be summarized in patients with a BOR of CR or PR only.

GCIG CA-125 response will be analyzed using a CMH test stratified by the randomization stratification factors to test for differences between the MIRV arm and the IC Chemo arm. P-value and 95% CIs for the proportion of patients with GCIG CA-125 clinical response will be provided. Odds ratio and its 95% CI will also be reported.

The distribution of DOR will be summarized using the Kaplan-Meier method. Kaplan-Meier estimate of DOR will also be reported at 3-month intervals (eg, 3 months, 6 months, etc.). Median DOR will be estimated for each treatment from the 50^{th} percentile of the corresponding Kaplan-Meier estimates. 95% CIs for the 3-month intervals and median DOR will also be provided. Because DOR will be analyzed for the subset of patients who achieved a BOR of CR or PR, the primary comparison between treatments will use the unstratified log-rank test. The chi-square p-values from the log-rank tests will be reported.

The hazard ratio for DOR treatment comparisons will be estimated using an unstratified Cox proportional hazards model. The hazard ratio (MIRV arm vs IC Chemo arm) will be reported using the maximum likelihood estimate along with 95% CI. The IC Chemo arm will be used as the reference treatment (ie, denominator of the hazard ratio). A hazard ratio less than 1.0 would indicate a MIRV benefit over IC Chemo.

TTR will be summarized using descriptive statistics for patients who achieved a confirmed response of CR or PR.

If the primary endpoint of PFS by INV is statistically significant, PFS, ORR, DOR, and TTR by BICR will be analyzed using the same methods as in investigator assessed PFS, ORR, DOR and TTR as sensitivity analysis.

Analysis of other PRO endpoints will be described in a separate document to the SAP.

PFS2 analysis will be conducted using the same method as in PFS analysis.

5.4. Exploratory Efficacy Analyses

The protocol describes the following exploratory efficacy analyses:

- Evaluate association of *BRCA* mutation status and FRα expression level with anti-tumor activity of MIRV.

- Evaluate the association of anti-tumor activity and/or safety with the following:
  - Soluble FRα level in blood samples.

5.5. Efficacy Analysis on Subgroups of Patients

PFS_{INV}, ORR_{INV}, and OS will be analyzed with the following subgroups of patients:

- *BRCA* status (positive vs negative/unknown).
- Age (< 65 year vs ≥ 65 years).
- Baseline Eastern Cooperative Oncology Group Performance State (ECOG PS) (0 vs 1).
- Prior exposure to bevacizumab.
- Number of prior lines of therapy.
- Primary platinum-free interval (≤6 months vs >6 months).
- Most recent Platinum-free interval (≤3 months vs >3 months).
- Prior exposure to PARPi maintenance therapy (yes vs no vs uncertain).
- Country (USA vs rest of world).
- Stage at diagnosis (I-III vs IV).
- Weight at baseline (<=60kg, 60-80kg, >80kg).

The summaries for time to event variables include the number and percentage of events, median and its 95% CI, 2-sided P-value, and hazard ratio (MIRV to IC chemo) and its 95% CI.

The forest plot of time to event variables will display the number of subjects, events, hazard ratio (MIRV to IC chemo) and its 95% CI.

The summaries for ORR include the ORR and its 95% CI, and odds ratio and its 95% CI.

The forest plot of ORR will display the number of subjects, odds ratio and its 95% CI.

All subgroup analyses will be unstratified.

Due to potential small number of patients will be eligible for DOR analysis, no subgroup analyses will be conducted for DOR.

5.6. Additional Sensitivity Analyses

The primary analyses for PFS are based on the investigator assessment of radiological assessments only. If PFS_{INV} is statistically significant, a sensitivity analysis of PFS based on the BICR assessment (PFS_{BICR}) is described Section 5.3. As an additional sensitivity analysis, PFS will be re-assessed by using the Investigator's radiological and clinical assessments ($PFSA_{INV}$). In this analysis, $PFSA_{INV}$ will be defined as the time from the date of the randomization until the date of PD by radiological or clinical assessment or death from any cause, whichever occurs first, as determined by the Investigator. If PD is noted in both clinical and radiological assessments, the first date where PD is noted will be used for analysis.

Table 5: $PFSA_{INV}$ Definitions

Situation	Date of PFS Event or Censoring	Outcome
No baseline tumor assessments or post-baseline tumor assessments, and patient did not die within 105 days of randomization and no clinical progression noted	Date of randomization	Censored
No baseline tumor assessments or post-baseline tumor assessments, and patient died within 105 days of randomization and no clinical progression noted	Date of death	Death
Death	Date of death	Death
Radiological or clinical progression	The earliest of: Date of first radiological assessment indicating progression (ie, OR = PD on RECIST response eCRF), or Date of first instance of clinical progression.	Progression
New anti-cancer therapy prior to progression or death (including palliative radiotherapy during study treatment)	Date of last radiological assessment prior to the start of the new anticancer therapy	Censored
No death or progression	Date of last radiological assessment	Censored
Progression (radiological or clinical) or death after missing 2 or more consecutive radiological assessments (PD or death date - last radiological assessment date + 1 ≥ 105 days or >= 231 days if the assessment schedule has changed from every 6 weeks [+-1 week] to every 12 weeks [+- 3 weeks] after Week 36 or after discontinuing from study treatment per local requirements)	Date of last RECIST assessment	Censored

Includes Investigator assessed radiological and clinical progression.
eCRF = electronic case report form; OR = overall response; PD = disease progression; PFS = progression-free survival; RECIST = Response Evaluation Criteria in Solid Tumors.

Additionally, for PFS_{INV}, rather than censoring patients who progressed or died after missing 2 or more consecutive radiological assessments, the date of the PFS event will be backdated to the date of first missed assessment. This analysis will be called $PFSB_{INV}$. Table 6 provides the rules for $PFSB_{INV}$.

Table 6: PFSB $_{INV}$ Definitions

Situation	Date of PFS Event or Censoring	Outcome
No baseline tumor assessments or post-baseline radiological assessments, and patient did not die within 105 days of randomization	Date of randomization	Censored
No baseline tumor assessments or post-baseline radiological assessments, and patient died within 105 days of randomization	Date of death	Death
Death	Date of death	Death
Progression	Date of first radiological assessment indicating progression (ie, OR = PD).	Progression
New anti-cancer therapy prior to PD or death (including palliative radiotherapy during study treatment)	Date of last radiological assessment prior to the start of the new anticancer therapy	Censored
No death or progression	Date of last radiological assessment	Censored
PD or death after missing 2 or more consecutive radiological assessments (PD or death date - last radiological assessment date +1 ≥ 105 days or >= 231 days if the assessment schedule has changed from every 6 weeks [+-1 week] to every 12 weeks [+- 3 weeks] after Week 36 or after discontinuing from study treatment per local requirements).	Date of first missed radiological assessment (ie, date of the last adequate assessment + 42/84 days [ie, 6/12 weeks]).	Progression

Includes radiographic progression only.
OR = overall response; PD = progressive disease or disease progression; PFS = progression-free survival.

Additionally, PFSR$_{INV}$ will be analyzed by not censoring patients receiving palliative radiotherapy during study treatment while everything else is identical to PFS$_{INV}$. Table 7 provides the rules for PFSR$_{INV}$.

Table 7: PFSR$_{INV}$ Definitions

Situation	Date of PFS Event or Censoring	Outcome
No baseline tumor assessments or post-baseline radiological assessments, and patient did not die within 105 days of randomization	Date of randomization	Censored
No baseline tumor assessments or post-baseline radiological assessments, and patient died within 105 days of randomization	Date of death	Death
Death	Date of death	Death
Radiological progression	Date of first radiological assessment indicating progression (ie, OR = PD)	Progression
New anti-cancer therapy prior to PD or death (not including palliative radiotherapy during study treatment)	Date of last radiological assessment prior to the start of the new anticancer therapy	Censored

Table 7: $PFSR_{INV}$ Definitions (Continued)

Situation	Date of PFS Event or Censoring	Outcome
No death or progression	Date of last radiological assessment	Censored
PD or death after missing 2 or more consecutive radiological assessments (PD or death date - last radiological assessment date + 1 ≥ 105 days or >= 231 days if the assessment schedule has changed from every 6 weeks [+-1 week] to every 12 weeks [+- 3 weeks] after Week 36 or after discontinuing from study treatment per local requirements)	Date of last adequate radiological assessment showing no PD	Censored

Includes radiographic progression only.
OR = overall response; PD = progressive disease or disease progression; PFS = progression-free survival.

If the primary endpoint of PFS_{INV} is statistically significant, the following sensitivity analyses will also be performed:

- The concordance/discordance rate between the PFS outcomes for the BICR assessment and the INV assessment will be summarized.
- The concordance/discordance rate between the BICR Best Overall Response and the INV Best Overall Response will be summarized.

5.7. Strategies for Pooling Stratification Factors

This study uses 2 factors to stratify patients for randomization, with a total of 9 stratification levels. At the final analysis, there may be a small number of patients in 1 or more of the stratification levels. The following strategy for pooling stratification levels will be used if 1 or more of the stratification levels contains fewer than 12 patients (Table 8).

After pooling, each stratum is expected to have at least 1 event at the final analysis.

Table 8: Pooling of Stratification Factors

Priors	IC Chemo	If n < 12 Pool With
1	PLD	Next stratum
1	Pac	Previous stratum
1	Topo	Previous stratum
2	PLD	Next stratum
2	Pac	Previous stratum
2	Topo	Previous stratum
3	PLD	Next stratum
3	Pac	Previous stratum
3	Topo	Previous stratum

IC = Investigator's choice; Pac = paclitaxel; PLD = pegylated liposomal doxorubicin; Topo = topotecan.

5.8. Other Efficacy-related Summaries

New anti-cancer therapy data will be summarized for the ITT population by randomized arm as follows:

- Number and percentage of patients receiving new anti-cancer therapy.
- Type of new anti-cancer therapy.
  - Taxanes.
  - Topotecan.
  - Platinum compounds.
  - Anthracyclines.
  - Chemotherapy (other).
  - Bevacizumab.
  - Doxorubincin\PLD
  - Hormonal therapy.
  - Kinase inhibitors.
  - Mirvetuximab soravtansine.
  - PARP inhibitors.
  - PD-1/PD-L1 inhibitors.
  - Immunotherapy
  - Investigational/blinded therapy
  - Other.

Listings of efficacy-related data for will include the following:

- All lesion assessments (target lesion, non-target lesion, new lesion).
- New anti-cancer therapy.
- Investigator's RECIST assessments.
- (BICR RECIST assessments).
- CA-125 results.
- Derived parameters for CA-125 response, BOR, PFS, DOR, OS, and PFS2.
- Censoring for time-to-event variables.

6. SAFETY ANALYSES

The main safety summary tables will use data from the Safety population. Patients will be analyzed according to the actual study drug received.

Listings will be provided for patients in the Safety population. The actual study drug received will be displayed on the listing.

6.1. Exposure

Summary tables will be provided with the following information.

Exposure to MIRV, paclitaxel, topotecan, and PLD will be summarized in tables with descriptive statistics for the number of doses received, the number of cycles received, duration of dosing (weeks and months), total cumulative dose (mg), absolute dose intensity (mg/kg/dose for IMGN853, calculated as total cumulative dose [mg])/number of valid drug administration records /AIBW (kg); mg/m^2/dose for remaining investigational products, calculated as total cumulative dose (mg)/number of valid drug administration records/BSA (m^2)), and relative dose intensity (percentage of planned, calculated as [absolute dose intensity/6] × 100 for IMGN853, [absolute dose intensity/80] × 100 for paclitaxel, [absolute dose intensity/40] × 100 for PLD, [absolute dose intensity/1.25] × 100 for topotecan 1.25 mg/m^2, [absolute dose intensity/4] × 100 for topotecan 4 mg/m^2).

Duration of dosing in weeks are calculated as the following:

MIRV: (last dose date – first dose date + 21)/7

Pac: (last dose date – first dose date + 7)/7

PLD: (last dose date – first dose date + 28)/7

Top (Q4W): (last dose date – first dose date + 14)/7

Top (Q3W): (last dose date – first dose date + 16)/7

If a patient is allowed to change Top from Q4W to Q3W in the middle of study, then dose density calculation needs to break into two parts.

The number of patients who received MIRV will also be summarized by drug lot numbers.

A listing will be provided with the information from all study drug administration eCRFs over the treatment period.

6.2. AEs

AEs will be documented on the AE eCRF and monitored continuously throughout the study from the time of informed consent until 30 days after the patient's last study drug or until the event has resolved, stabilized, or returned to baseline. AEs attributed to study procedures, including those events that occur prior to the first dose, should also be documented on the AE eCRF.

AE data are available to ImmunoGen from 2 sources, the eCRFs and the serious adverse event (SAE) forms. While reconciliation will be performed, the production of data summaries and listings will be based on the data collected on the eCRF.

Pre-treatment AEs are defined as AEs with an onset date prior to the first dose of study drug. TEAEs are defined as AEs with an onset date on or after the first dose of study drug, and within 30 days of the last dose of study drug or prior to the start of a new anti-cancer treatment, whichever occurs first. Medical history conditions that exist before the initiation of study drug but worsen in severity during the study will also be recorded on the AE eCRF as an AE and will be included as treatment-emergent in the summary tables and listings.

The adverse events will be coded using MedDRA (Version 24.0 or later), associating lower-level terms with PT and SOC by the primary hierarchy. The tables will display the counts and percentages of patients who reported at least 1 TEAE in each SOC represented in the AE data. Within each SOC, the tables will display the counts and percentages of patients reporting at least 1 TEAE as designated by the PT.

AEs are graded using Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0. AE summaries may include summaries for all AEs and by the maximum CTCAE grade for the item being summarized (ie, SOC or PT). In these cases, the outputs will include a row for All Grades as well as rows for the 5 potential CTCAE grades, Grade 1 (Mild), Grade 2 (Moderate), Grade 3 (Severe), Grade 4 (Life threatening or disabling), or Grade 5 (Death). AEs missing a severity grade will not be included in the Grade 1-5 rows of the tables. An AE reported by a patient more than once will be represented in the most severe category.

AEs that are definitely, probably or possibly related to the study drug will be considered as related to the study drug. AEs with missing or unknown relationship to the study drug are considered as related to the study drug. AEs with the closest relatedness to the study drug will be used for summaries.

The following AE summary tables will be produced:

- An overall summary of safety and the number of patients who died during the study or within 30 days of last dose.
- All TEAEs by SOC and PT.
- Serious TEAEs by SOC and PT.
- Non-Serious TEAEs (ie, TEAEs excluding SAEs) by SOC and PT.
- Grade 3 or higher TEAEs by SOC and PT.
- TEAEs leading to drug withdrawal. This subset includes TEAEs with an action taken of drug permanently discontinued.
- Related TEAEs leading to drug withdrawal. This subset includes TEAEs related to study drug with an action taken of drug permanently discontinued.
- TEAEs related to study drug by SOC and PT. An AE reported by a patient more than once will be included in this table if at least 1 of the drug association grades is 1 of the grades listed here.
- Serious, related TEAEs by SOC and PT. An AE reported by a patient more than once will be included in this table if at least 1 of the drug association grades is 1 of the grades listed here.

- TEAEs leading to dose modifications (dose reduction, dose not given or dose delay).
  - TEAEs leading to dose reduction.
  - TEAEs leading to dose delay.
  - TEAEs leading to dose reduction, dose not given or dose delay.
- Related TEAEs leading to dose modification (dose reduction, dose not given or dose delay).
  - Related TEAEs leading to dose reduction.
  - Related TEAEs leading to dose delay.
  - Related TEAEs leading to dose reduction, dose not given, or dose delay.
- Deaths on study treatment or within 30 days of the last dose. This table includes all deaths during study treatment or within 30 days of the last dose, regardless of cause of death.
- TEAEs leading to death. This table includes all TEAEs with CTCAE Grade 5.
- Related TEAEs leading to death. This table includes all TEAEs related to study drug with CTCAE Grade 5.

The following listings will be produced:

- All AEs, sorted chronologically by patient. This listing includes SOC, PT, onset and end dates, and other relevant information.
- Serious TEAEs, sorted chronologically within patient.
- TEAEs leading to drug withdrawal. This subset includes TEAEs with an action taken of drug permanently discontinued.
- TEAEs resulting in death. This listing includes TEAEs with a CTCAE grade of Grade 5 (death).
- survival status.

The following groupings of TEAEs will be generated as part of the focused analysis of safety:

- Ocular TEAEs.
- Peripheral neuropathy TEAEs.
- Pneumonitis.

Additionally, a table will be produced which contains the following for each of the focused safety analysis groups:

- The number of patients with at least 1 TEAE in each group, presented by treatment arm by PT and maximum CTCAE grade.
- Time to first onset of each group of TEAEs.
- Action taken with study drug with respect to each group of TEAEs.

Actions include:

- Dose permanently discontinued.
- Dose reduced.
- Dose delayed/not given.
- Infusion interruption.

A subject could have multiple actions above.

- Worst outcome for TEAEs of interest
  - Dose permanently discontinued.
  - Dose reduced.
  - Dose delayed/not given.
  - Infusion interruption.
  - No action taken.

  Severity of action is based on order above. A subject is only counted once in the worst category.
- Resolution status.

For the focused safety analysis, the following listings will be produced:

- Ocular TEAEs.
- Peripheral neuropathy TEAEs.
- Pneumonitis TEAE.

6.2.1. Fresh Biopsy Patient Adverse Events

Summary of adverse events, not necessarily TEAEs, experienced within 7 days of biopsy will be generated by SOC, PT and CTCAE grade. A listing of AEs experienced within 7 days of biopsy will also be generated for patients who have undergone fresh biopsy.

6.2.2. Infusion-related Reactions

Infusion related reactions will be analyzed using Standard MedDRA Query (SMQ) with the keyword of hypersensitivity (narrow) and PT terms 'flushing', 'erythema', 'erythema of eyelid'. A summary table of all TEAEs will be generated by SOC, PT, and CTCAE grade. A listing will also be generated.

6.3. Clinical Laboratory Results

Laboratory test results (including hematology, coagulation, serum chemistry, and urinalysis) and abnormal laboratory values will be presented in data listings.

CTCAE Version 5.0 laboratory grades will also be presented. CTCAE grades will be derived based on laboratory results and will not factor in clinical evaluations.

Shift tables summarizing the changes from baseline in severity of laboratory grades will be provided for laboratory parameters graded according to CTCAE Version 5.0

Clinically significant values in liver function tests (LFTs) will be summarized by the following categories, using the maximum value while on study drug. The denominator for the summaries will be the number of patients who had at least 1 non-missing value during treatment. The categories for each test are not mutually exclusive:

- Aspartate aminotransferase (AST).
  - > 3 × ULN.
  - > 5 × ULN.
  - > 10 × ULN.
  - > 20 × ULN.
- Alanine Aminotransferase (ALT).
  - > 3 × ULN.
  - > 5 × ULN.
  - > 10 × ULN.
  - > 20 × ULN.
- AST or ALT.
  - > 3 × ULN.
  - > 5 × ULN.
  - > 10 × ULN.
  - > 20 × ULN.
- Total bilirubin (TBL).
  - > 1.5 × ULN.
  - > 2 × ULN.
- Alkaline phosphatase (ALP).
  - > 1.5 × ULN.
- (AST or ALT) and TBL (concurrent).
  - AST or ALT > 3 × ULN and TBL > 1.5 × ULN.
  - AST or ALT > 3 × ULN and TBL > 2 × ULN.
- (AST or ALT) and ALP and TBL (concurrent).
  - AST or ALT > 3 × ULN and ALP < 2 × ULN and TBL > 2 × ULN.

Here, concurrent means that all the associated LFTs must be from the same visit. In addition, worst postbaseline grade for LFTs will be summarized. Because baseline values of LFTs are

already taken into account in CTCAE V5.0, shift table from baseline to worst postbaseline grade will not be generated for LFTs.

Results from pregnancy tests will be provided in listings.

6.4. Vital Signs

Vital signs (including temperature, pulse rate, systolic blood pressure, diastolic blood pressure, and respiratory rate) will be collected throughout the study.

Vital signs results will be classified into 4 or 5 categories (low, borderline low, normal, borderline high, or high) according to Table 9 below. Shift tables based on this classification will be summarized from baseline to the last post-baseline visit value.

Table 9: Classification of Vital Signs

Vital Sign	Low	Borderline Low	Normal	Borderline High	High
Heart Rate (beats per minute)	< 50	50-59	60-90	91-99	≥ 100
Systolic BP (mmHg)	< 80	-	80-120	121-139	≥ 140
Diastolic BP (mmHg)	< 60	-	60-80	81-89	≥ 90
Respiratory Rate (breaths per minute)	< 12	-	12-20	21-24	≥ 25
Temperature (°C)	< 35.0	35.0-36.4	36.5-37.2	37.3-37.9	≥ 38.0

6.5. ECGs

All ECG results will be presented in data listings. If a different correction for QT is captured in the eCRF, that QTc will be reported in the listing for that patient with QTcF.

6.6. Concomitant Medications

All medications and supportive therapy taken within 4 weeks prior to Cycle 1 Day 1 and through 30 days after last study treatment must be recorded on the appropriate eCRF. The identity of all medications, dosage, route of administration, frequency, duration of administration, and indication for use will be recorded in the appropriate sections of the eCRF.

Prior medications are defined as medications with a stop date prior to the first dose of study drug.

Concomitant medications are defined as medications which are taken during the course of study treatment and within 30 days of the last dose of study drug, as follows:

- Medications started before the first dose of study drug, but with a stop date after the first dose of study drug and within 30 days of the last dose of study drug will be considered concomitant medications.
- Medications started before the first dose of study drug that are ongoing will be considered concomitant medications.

- Medications started after the first dose of study drug and within 30 days of the last dose of study drug or before the start of a new anti-cancer treatment, whichever occurs first, are considered concomitant medications.
- Medications started before the first dose of study drug, but with a stop date after the first dose of study drug and within 30 days of the last dose of study drug or prior to the start of a new anti-cancer treatment, whichever occurs first, will be considered concomitant medications.

Prior and concomitant medications will be coded using the September 2019 or later version of World Health Organization drug dictionary (WHO Drug). Summary tables will be provided for prior and concomitant medications.

Summary tables will be organized to display the anatomical main class of each coded medication (ATC Level 1 term) and, within that, the pharmacological subgroup (ATC Level 3 term) of the coded medication. The summary table will display number and percentage of patients who reported using at least 1 medication in each represented pharmacological subgroup. If a patient has more than 1 medication in the subgroup, the patient will be counted only once.

A complete listing of medications will be generated by patient. The listing will indicate which medications are prior and which are concomitant. The listing will display entries from the concomitant medications form, ordered within patient by start date. The listing will display the recorded term from the eCRF and the WHO Drug anatomical main class (ATC Level 1 term) and pharmacological subgroup (ATC Level 3 term).

6.7. Concomitant Procedures

All procedures within 4 weeks of Cycle 1 Day 1 and through 30 days after last study treatment must be recorded on the appropriate eCRF.

Prior procedures are defined as occurring before the first dose of study drug (by procedure date).

Concomitant procedures are defined as procedures with a procedure date on or after the first dose of study drug, and within 30 days of the last dose of study drug or prior to the start of a new anti-cancer treatment, whichever occurs first.

Prior and concomitant procedures will be coded using MedDRA (Version 24), associating lower-level terms with PT and SOC by the primary hierarchy. Summary tables will be provided for prior and concomitant procedures. The tables will display number and percentage of patients who reported at least 1 procedure in each SOC represented in the eCRF data. Within each SOC, the tables will display number and percentage of patients reporting at least 1 concomitant procedure as designated by PT.

A complete listing of procedures will be generated. The listing will indicate which procedures are prior and which are concomitant. The listing will display entries from the concomitant procedures form, ordered within patient by date of procedure. The listing will display the recorded term from the eCRF and the SOC and PT.

6.8. Ophthalmic Examinations

Ophthalmic Examinations are collected at the Screening, End-of-Treatment, and 30-Day Follow-up Visits. Results of the ophthalmic examinations will be presented in data listings.

Worst decline in both eye best corrected visual acuity (BCVA) grade will be summarized. Shift of both eye BCVA from baseline to worst post baseline will be summarized using the following categories: more than two lines of improvement, two lines improvement, one line of improvement, less than one line change in visual acuity, one line worsening, two lines worsening, more than two lines of worsening.

Intraocular pressure elevations will be summarized by the number of subjects with 7 mmHg elevations or greater post baseline.

6.9. Ocular Symptom Assessments

Results of the ocular assessments will be presented in data listings.

6.10. Corticosteroid and Lubricating Eye Drop Compliance

All compliance information collected on the eCRF will be presented in data listings.

6.11. Transfusions

All transfusions recorded on the eCRF will be presented in data listings.

6.12. ECOG PS

ECOG PS results will be presented in data listings. Baseline ECOG PS will be summarized in disease characteristics table.

6.13. ECHO/MUGA

Echocardiogram (ECHO)/multiple gated acquisition (MUGA) scans are taken only for patients receiving PLD. All ECHO/MUGA results will be presented in data listings.

7. PRO

Analyses for PRO, including quality of life and healthcare resource utilization, will be covered by a separate, independent analysis plan by the PRO vendor, in collaboration with ImmunoGen.

8. IMMUNOGENICITY

Anti–MIRV seroconversion refers to the development of detectable antibodies that bind to MIRV and is based upon positive results in both screening and confirmatory assays. Patients will be classified as follows:

- Patients who had a baseline-negative ADA result who developed ADAs at any time after initial administration of drug.

- Treatment-enhanced ADAs: Patients who had a baseline positive ADA result in whom the assay signal was enhanced (greater than baseline titer by ≥ 4 folds) at any time after initial drug administration.
- Treatment-unaffected ADA: Patients who had a baseline-positive ADA result in whom the assay signal was not enhanced (not greater than baseline titer by ≥ 4 folds) at any time after initial drug administration.
- Seronegative: a patient who tests negative at all visits.

If a titer is below LLQ, it will be imputed as half of LLQ.

Efficacy endpoints PFS and ORR will be analyzed by the above ADA status group.

TEAE by SOC and PT will be summarized for the above ADA status group.

9. BIOMARKERS

The exploratory biomarker analyses for this study will be covered by a separate, independent analysis plan.

10. PHARMACOKINETICS

Plasma concentration data collected in this study will be listed and summarized in the PK analysis population. Mean and 95% CI, standard deviation, geometric mean, coefficient of variation (CV, expressed as percentage), median, Q1, Q3, min, and max will be reported by visit. Data points that are below the lower limit of quantitation will be excluded from the analyses.

11. REVISION HISTORY

Table 10: Revision History

Version	Effective Date	Comments
1	10Oct 2019	Initial document
2.0	23Mar 2020	1. Added PK analysis population 2. Removed reference to iSAP 3. Specified comparison of RMST at 12 months between treatment arms 4. Removed location of metastasis 5. Changed CTCAE version to 5.0 6. Clarified how to summarize LFT CTCAE grades due to changes from v4.03 to 5.0
3.0	14Feb 2023	1. Clarification on longitudinal period population 2. Clarification on censoring rule for missing 2 or more consecutive radiological assessments. 3. Updated on BCVA and intraocular pressure elevations analysis based on FDA feedback. 4. Clarification on immunogenicity categories.

12. REFERENCES

N/A