NCT02631876

ImmunoGen, Inc.

Protocol #: 0403

FORWARD 1: A Randomized, Open Label Phase 3 Study to Evaluate the Safety and Efficacy of Mirvetuximab Soravtansine (IMGN853) Versus Investigator's Choice of Chemotherapy in Women with Folate Receptor α–positive Advanced Epithelial Ovarian Cancer, Primary Peritoneal Cancer or Fallopian Tube Cancer

Statistical Analysis Plan

Version 4.0

Document last saved: 3-Mar-2020 1:56 PM

Prepared for:

[redacted]

ImmunoGen, Inc.

830 Winter Street

Waltham, MA 02451

Prepared by:

[redacted]

Novella Clinical

1700 Perimeter Park Dr.

Morrisville, NC 27560

Approvals

[redacted] Date

[redacted] ImmunoGen, Inc.

[redacted] Date

[redacted] ImmunoGen, Inc.

[redacted] Date

[redacted] Novella Clinical

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or Specialist Term	Explanation
ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse event of special interest
AIBW	Adjusted ideal body weight
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIRC	Blinded Independent Review Committee
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
FcγR	Fc gamma receptor
FRα	Folate receptor alpha
FOSI	FACT-ovarian symptom index
GCIG	Gynecologic Cancer Intergroup
IC	Investigator's choice/selected standard-of-care chemotherapy
IDMC	Independent Data Monitoring Committee

Abbreviation or Specialist Term	Explanation
IHC	Immunohistochemistry
IMGN	ImmunoGen
INV	Investigator
ITT	Intent-to-treat
MDR1	Multi-drug resistant gene
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum tolerated dose
MUGA	Multigated acquisition scan
ORR	Objective response rate
OS	Overall survival
Pac	Paclitaxel
PD	Progressive disease
PFS	Progression-free survival
PFS2	Time to second disease progression
PgP	P-glycoprotein
PK	Pharmacokinetics
PLD	Pegylated liposomal doxorubicin
PR	Partial response
PR interval	Pulse rate interval
PRO	Patient reported outcomes
PT	Preferred term
Q3W	Every 3 weeks
Q4W	Every 4 weeks
QoLP	Primary endpoint for quality of life
QRS	Part of electrocardiographic wave representing ventricular depolarization
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

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

1. INTRODUCTION

1.1. Background

The Forward 1 study IMGN853-0403 was initiated during the dose-expansion stage of study IMGN853-0401, a first-in-human study of single-agent mirvetuximab soravtansine (also known as IMGN853).

During the dose-escalation stage of study IMGN853-0401, two dosing schedules were evaluated: Schedule A (IMGN853 administered on Day 1, with cycles repeating every 21 days [Q3W]) and Schedule B (IMGN853 administered on Days 1, 8, and 15, with cycles repeating every 28 days [modified weekly]). Dose expansion began once the Schedule A maximum tolerated dose (MTD) was identified and a dose was selected for further exploration. Schedule B evaluation was stopped prior to identifying the MTD.

Study IMGN853-0401 identified the recommended Phase 2 dose of single-agent mirvetuximab soravtansine in patients with epithelial ovarian cancer (EOC) and other folate receptor alpha (FRα)-positive tumors to be 6.0 mg/kg adjusted ideal body weight (AIBW) administered on Day 1 of a 21-day cycle.

Study IMGN853-0403 was initially opened as a Phase 2 study to further evaluate the 2 dosing schedules of mirvetuximab soravtansine (Q3W and Q4W). Following emerging data from study IMGN853-0401, IMGN853-0403 was placed on hold and redesigned to be a randomized, open-label, Phase 3 study to evaluate the safety and efficacy of mirvetuximab soravtansine (IMGN853) versus Investigator's choice (IC) of chemotherapy in women with FRα–positive advanced EOC, primary peritoneal cancer, or fallopian tube cancer.

The protocol for study IMGN853-0403 describes the general approach to the analysis of study data. This statistical analysis plan (SAP) describes the additional detail needed to complete such an analysis. Table, figure, and listing shells will be supplied in an accompanying document.

This SAP is based on Amendment 8A of Protocol 0403.

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	18 August 2015	
Amendment 1	01 September 2015	Secondary objectives for Stage 1 were revised.
Amendment 2	23 September 2015	No changes that require accommodation in this analysis plan.
Amendment 3	20 October 2015	Modification of the CA-125 assessment schedule
Amendment 4	01 February 2016	No changes that require accommodation in this analysis plan.
Amendment 5	17 August 2016	The trial has been changed from a Phase 2 study to a Phase 3 study, and limited to women with

Version	Approval Date	Salient Changes (if any)*
		FRα-positive advanced EOC, primary peritoneal cancer, or fallopian tube cancer. The title of the study was revised to reflect these changes, and Exclusion Criterion 1 for male patients was added. Stage 1 of the study has been closed for enrollment. The dosing schedule for Arm 1 (IMGN853) of Stage 2 was updated to 6mg/kg AIBW Q3W. Chemotherapy agents for the IC (Arm 2) chemotherapy were updated to include only Pac, PLD, and Topo. Randomization stratification factors have been revised to include number of prior lines of therapy, FRα levels, and IC chemotherapy, while BRCA mutation status has been excluded. Central radiology reviews by a BIRC were added to assess tumor response in addition to the Investigator's radiology assessment. The primary endpoint for Stage 2 was revised. PFS as assessed by the BIRC is the primary endpoint. The sample size has been updated to reflect the change in the primary endpoint. An interim futility analysis has been added with rules for study termination. The secondary objectives for Stage 2 were updated to include ORR, DOR by BIRC and Investigator, and PFS as assessed by the Investigator. Soluable FRα level in blood was added as an exploratory endpoint. Patients on IC chemotherapy will have the option to cross over to the IMGN853 arm after BIRC-confirmed PD. A window for start of treatment from randomization was added; Cycle 1 Day 1 should be within 7 calendar days of randomization. PRO questionnaires to be used in the study have been revised to include EORTC QLQ-C30, EORTC QLQ-OV28, eight-item FOSI, and EQ-5D-5L questionnaires instead of the FACT-O questionnaire. Statistical tests for PROs were added.

Version	Approval Date	Salient Changes (if any)*
		Collection of TEAEs has been revised to events collected from the time of informed consent until 30 days after the patient's last study treatment. Pneumonitis was added as an AESI for IMGN853. Requirement for post-dose ECGs as well as triplicate ECGs were removed. Requirement for pulse oximetry assessment was removed.
Amendment 6	19 September 2016	No changes that require accommodation in this analysis plan.
Amendment 7	04 November 2016	Optional crossover from IC chemotherapy to IMGN853 after BIRC-confirmed PD was removed. PFS2 was added as an exploratory endpoint
Amendment 8	08 May 2017	No changes that require accommodation in this analysis plan.
Amendment 8A (North America)	25 September 2017	The protocol was amended to add the possibility for the collection of urine samples from patients receiving IMGN853 at designated sites in North America. The exploratory objectives of the protocol have been updated. However, the analysis of the urine samples will be conducted under the separate PK analysis plan.

*Changes expected to require accommodation in analysis plan.
AESI = adverse event of special interest; AIBW = adjusted ideal body weight; BIRC = Blinded Independent Review Committee; BRCA = breast cancer susceptibility gene; CA-125 = cancer antigen 125; DOR = duration of response; ECG = electrocardiogram; EOC = epithelial ovarian cancer; EORTC = European Organisation for Research and Treatment of Cancer; FOSI = FACT-ovarian symptom index; FRα = folate receptor alpha; IC = Investigator's choice; ORR = overall response rate; Pac = paclitaxel; PD = progressive disease; PFS = progression-free survival; PFS2 = time to second disease progression; PK = pharmacokinetic(s); PLD = pegylated liposomal doxorubicin; PRO = patient-reported outcome; QLQ = Quality of Life Questionnaire; TEAE = treatment-emergent adverse event; Topo = topotecan.

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

The Phase 3 portion of the study is the main focus of this SAP and will be refered to as 'study', information included that specifically relates to a Stage 1 topic is designated as such.

2. STUDY DESIGN

2.1. Protocol Objectives

2.1.1. Primary Objective

- To compare the progression-free survival (PFS) of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC), as assessed by the Blinded Independent Review Committee (BIRC) in the Intent-to-treat (ITT) population (defined in Section 3.12.2) and in the high-FRα subgroup (≥ 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 IMGN853 versus selected standard-of-care chemotherapy (IC).
  - Primary analysis of ORR will be based on BIRC assessments. ORR based on Investigator's assessment will be analyzed as sensitivity analysis.
- To compare the overall survival (OS) of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC).
- To compare the primary patient reported outcomes (PRO) endpoint using QLQ-OV28 assessments from patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC) as described in Protocol Section 11.7.

2.1.2.2. Other Secondary Objectives:

- To compare the safety and tolerability of IMGN853 with that of selected standard-of-care chemotherapy (IC).
- To compare the duration of response (DOR) of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC).
  - Primary analysis of DOR will be based on BIRC assessments. DOR based on Investigator's assessment will be analyzed as sensitivity analysis.
- To compare the cancer antigen 125 (CA-125) response rate per Gynecologic Cancer Intergroup (GCIG) CA-125 criteria of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC).
- To compare the PFS of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC), as assessed by the Investigator.
- To evaluate the pharmacokinetics (PK) of IMGN853.
- To assess the immunogenicity of IMGN853 (anti-drug antibodies, ADA).
- To assess PRO using the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ)-C30, EORTC

QLQ-OV28, EQ-5D-5L, and eight-item FACT-ovarian symptom index (FOSI) questionnaires.

2.1.3. Exploratory Objectives

- To evaluate potential biomarkers in blood and tumor tissue that might predict response to IMGN853.
- To compare the time to second disease progression (PFS2) of patients randomized to IMGN853 versus selected standard-of-care chemotherapy (IC).
- To identify DM4 metabolites in urine.

2.2. Study Endpoints

2.2.1. Primary Endpoint

- PFS: the time from the date of randomization until the time of death or progressive disease (PD), as assessed by the BIRC (PFS_{BIRC}):
  - In all patients randomized to the study.
  - In patients with high FRα level (≥ 75% of tumor staining at ≥ 2+ intensity).

2.2.2. Secondary Endpoints

2.2.2.1. Key Secondary Endpoints:

- ORR per Response Evaluation Criteria in Solid Tumors (RECIST) Version 1.1 criteria as assessed by BIRC.
- OS: the time from date of randomization until the date of death.
- Primary PRO endpoint using the QLQ-OV28 questionnaire as described in Protocol Section 11.7.

2.2.2.2. Other Secondary Endpoints:

- Treatment-emergent adverse events (TEAEs) and laboratory test results, physical examination, electrocardiograms (ECGs), or vital signs.
- Gynolocologic Cancer Intergoup CA-125 criteria clinical response rate.
- Time-to-event endpoints:
  - PFS as assessed by Investigator.
  - DOR: the time from first objective response (complete response[CR]/partial response[PR]) to the time of PD among those who have achieved a PR or CR), as assessed by BIRC.
  - DOR as assessed by the Investigator.
- PK parameters of IMGN853.

- Immunogenicity of IMGN853: ADA.
- Patient scores on EORTC QLQ-C30, EORTC QLQ-OV28, EQ-5D-5L, and 8-item FOSI questionnaires.

2.2.3. Exploratory Endpoints

- Evaluate association of breast cancer susceptibility gene (BRCA) mutation status in tumor tissue and FRα expression level with anti-tumor activity of IMGN853.
- Evaluate the association of anti-tumor activity and/or safety with the following:
  - Mutational status and other genomic alterations in tumor samples.
  - Activation status of oncogenic pathways in tumor samples.
  - Expression of drug transporters such as multi-drug resistant gene 1 (MDR1) (ie, P-glycoprotein [PgP]) as well as other proteins that may influence anti-tumor activity or safety.
  - Blood-based biomarkers (such as soluble FRα).
  - Genotyping of Fc gamma receptor (FcγR).
- PFS2, as assessed by the Investigator: the time from randomization to the time of second PD or death.
- Qualitative analysis of DM4 metabolites in urine.

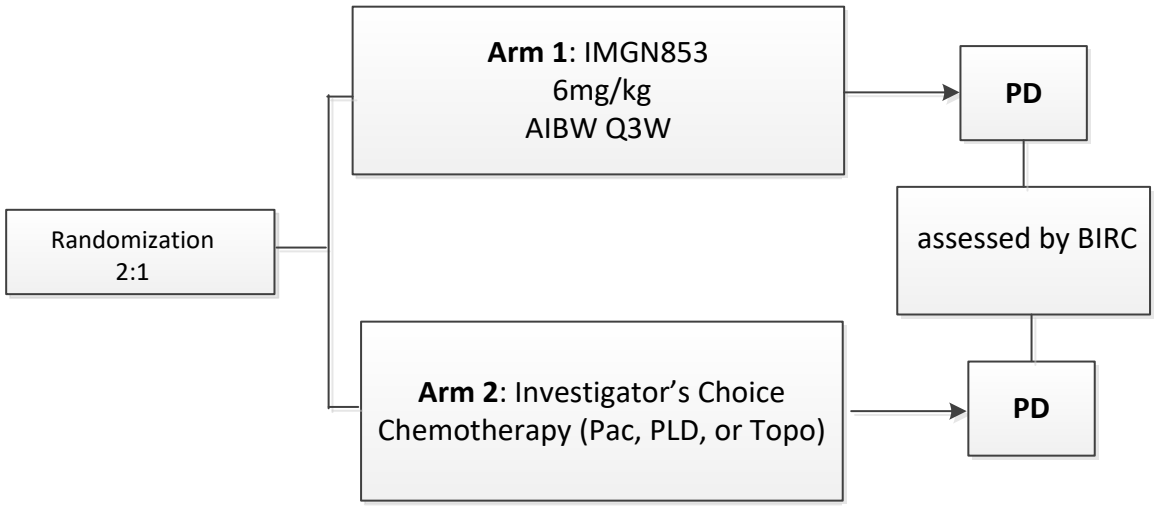
2.3. Study Overview

This study is designed to compare the efficacy of IMGN853 to that of selected standard-of-care chemotherapy (IC) in women with platinum-resistant advanced EOC, primary peritoneal cancer, or fallopian tube cancer.

Patients will be stratified by number of prior lines of treatment (1 or 2 vs 3), FRα levels (≥ 75% tumor staining at ≥ 2+ intensity [high] vs ≥ 50% and < 75% tumor staining at ≥ 2+ intensity [medium]), and IC chemotherapy (paclitaxel [Pac], pegylated liposomal doxorubicin (PLD), or topotecan [Topo]). Patients will be randomized 2:1 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



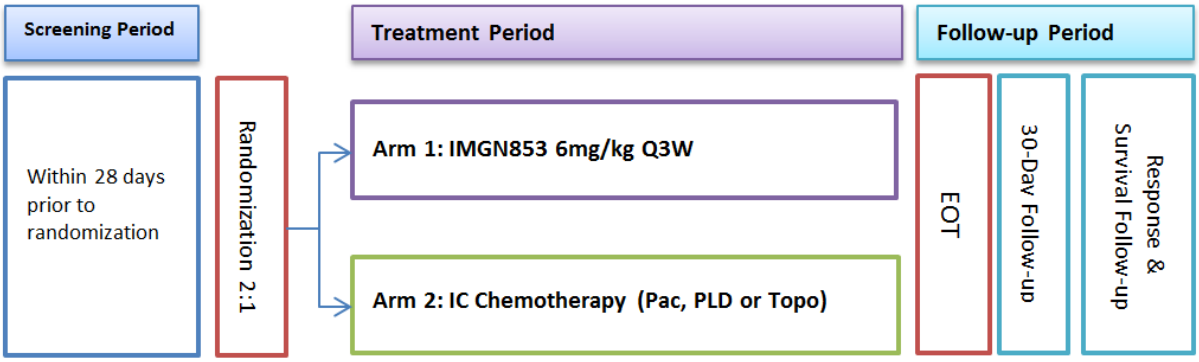
AIBW = adjusted ideal body weight; BIRC = Blinded Independent Review Committee; Pac = paclitaxel; PD = disease progression; PLD = pegylated liposomal doxorubicin; Q3W = every 3 weeks; Topo: topotecan.

Patients will continue to receive study treatment until they present with PD per RECIST Version 1.1, (as assessed by the BIRC), develop unacceptable toxicity, or withdraw consent, whichever comes first, or until the Sponsor terminates the study.

Tumor assessments, including radiological assessments by computerized tomography or magnetic resonance imaging scans, will be performed at screening and subsequently every 6 weeks (± 1 week) for the first 36 weeks then every 12 weeks (± 1 week) until PD per RECIST Version 1.1, death, or the initiation of subsequent anti-cancer therapy, whichever occurs first. All patients will be followed every 3 months (± 2 weeks) for survival until death, patient is lost to follow-up, patient withdraws consent for survival, or until End of Study, whichever comes first.

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.

2.3.1. 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 according to RECIST Version 1.1.
- Received at least 1 but no more than 3 prior systemic lines of anti-cancer therapy.
- Confirmation of FRα positivity by immunohistochemistry (IHC) (≥ 50% of tumor staining at ≥ 2+ intensity) 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.

2.3.2. Power and Sample Size

2.3.2.1. Primary Endpoints

The study has 2 primary endpoints, PFS as assessed by BIRC in all randomized patients and PFS as assessed by BIRC in the FRα high-expression subgroup. The study is designed to test the null hypothesis that the survival function for PFS is the same between the IMGN853 arm and the IC chemotherapy arm versus the alternative hypothesis that the survival function for PFS is different between the IMGN853 and IC chemotherapy arms. The Hochberg procedure will be used to control the overall type I error (see Section 3.11 Hypothesis Testing and Multiple Comparisons for further details).

Approximately 333 patients will be randomized 2:1 (222 to the IMGN853 arm and 111 to the IC arm) over a period of approximately 21 months across approximately 130 study centers globally. The final analysis will be conducted when at least 236 PFS events are observed.

An interim futility analysis will be conducted when at least 80 PFS events have been observed. The study will be terminated for futility at interim analysis if the observed hazard ratio [IMGN853 to IC chemotherapy] is greater than 1 in all randomized patients and in the FRα high-expression subgroup.

The study will have 91% power to detect a hazard ratio of 0.583 in the FRα high-expression subgroup and 96% power in all randomized patients at a study-wise alpha level of 5%; the study will have a 39% probability of stopping for futility at interim analysis under the null hypothesis.

Sample size and power were determined by simulations using SAS® software with the following assumptions:

- Median PFS for the IC arm is 3.5 months.
- Median PFS for the IMGN853 arm is 6 months.
- Exponential distribution for both event and censoring processes.
- Ratio of FRα high to FRα medium is 2:1.
- Annual censoring rate is 20% in both arms.

2.3.2.2. Power for Key Secondary Endpoints

ORR as assessed by BIRC

With a sample size of 333 (222 in the IMGN853 arm and 111 in the IC chemotherapy arm), the study will have a power of 89% to detect a 17% difference in ORR (35% in IMGN853 vs 18% in IC chemotherapy).

OS

At the final analysis of PFS, the study will have 29% power to detect an OS hazard ratio of 0.778, assuming the median OS for IC chemotherapy is 14 months.

Primary PRO endpoint

Number of patients achieving at least a 15% improvement on the QLQ-OV28 abdominal/gastointestinal 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 IMGN853 versus 10% in IC chemotherapy).

2.3.3. Treatment Randomization and Stratification

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

- ***Arm 1:*** IMGN853 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 or 2 vs 3).
- FRα levels (high, defined as ≥ 75% of tumor staining at ≥ 2+ intensity vs medium, defined as ≥ 50% and < 75% at ≥ 2+ intensity).
- IC chemotherapy (Pac, PLD, or Topo).

The required FRα expression level is ≥ 50% of tumor staining at ≥ 2+ intensity. A stratification cutoff of ≥ 75% at ≥ 2+ will also be implemented to evaluate anti-tumor activity of IMGN853 based on higher levels of FRα expression.

Cycle 1 Day 1 should occur within 7 calendar days from randomization.

2.3.4. Blinding and the BIRC

This is an open-label study. Treatment assignment will not be blinded. A BIRC will be used to provide an independent assessment of radiographic tumor assessments. The BIRC 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, clinical scientists, medical monitors, study statisticians, and study statistical programmers will not have access to the RECIST tumor assessments for either the Investigator's assessment captured in the EDC system or the BIRC data. Data managers and clinical trial managers will have access to the tumor assessment data.

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

2.3.4.1. BIRC

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 BIRC.

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

All time points other than the Investigator-determined PD will be assessed by BIRC in a non-expedited manner. Results of these assessments will not be communicated to the sites.

For time points with Investigator-determined PD, the BIRC will perform an expedited review. While the images are assessed by the BIRC, the patient should continue on the study treatment as planned if it is clinically acceptable to do so. If the BIRC confirms PD, the patient must discontinue study treatment. If the BIRC does not confirm PD, the patient should continue receiving the study treatment unless there is a medical need (ie, rapid PD or clinical deterioration) that requires an immediate change in therapy. Although disease progression may be determined by the Investigator based upon clinical deterioration, every

effort should be made to document PD using radiographic methods. The basis for determination of disease progression per clinical deterioration should be documented in the eCRFs.

The central imaging vendor will ensure that the central radiologists remain blinded to the local assessment from the Investigator and other unblinding information. This and all other imaging procedures will be documented in an independent review charter agreed upon between ImmunoGen and the imaging vendor before initiation of any BIRC reviews. BIRC assessment of PD will be used for the primary endpoint analysis.

2.3.5. Assessment Schedule

See protocol Appendix A, Appendix B, and Appendix C for study schedules of assessments.

2.4. Interventions

2.4.1. 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. Patient entry into the eCRF system and randomization are completed through the Y Prime system. The data are then automatically mapped into the eCRF system, and will be available through the eCRF data extract.

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 adverse event (AE) onset and resolution dates, concomitant medication start and stop dates, and concomitant procedure, 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.

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

In the primary efficacy analysis, no covariate adjustment will be made. 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

An interim futility analysis will be conducted when at least 80 PFS events as assessed by BIRC have occurred. The study will be terminated for futility at interim analysis if the observed hazard ratio [IMGN853 to IC chemotherapy] is greater than 1 in all randomized patients as well as in the FRα high-expression subgroup. No alpha spending is planned for this futility.

Under the enrollment assumptions (see Appendix 2: Monte-Carlo Simulation Report), the interim analysis will occur when approximately 160-180 patients have been randomized. This is expected to occur approximately 1 year after the first patient is randomized.

The interim analysis of efficacy endpoints will follow the analysis methods described in the interim analysis plan.

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 333 patients for full enrollment (ie, the study is not stopped for futility at the interim analysis), the final analysis will be conducted when at least 236 PFS events have been observed in the ITT population (as determined by the BIRC).

3.10. Test Sizes

Any tested hypotheses will be tested against 2-sided alternatives, using procedures that provide an expected probability of Type I error (α) of 0.05.

3.11. Hypothesis Testing and Multiple Comparisons

3.11.1. Primary Endpoints

The 2 null hypotheses for primary endpoints are:

- H_{01}: the survival function for PFS_{BIRC} is the same between the IMGN853 arm and the IC chemotherapy arm in the ITT population, and
- H_{02}: the survival function for PFS_{BIRC} is the same between the IMGN853 arm and the IC chemotherapy arm in the FRα high-expression subgroup.

And the alternative hypotheses are:

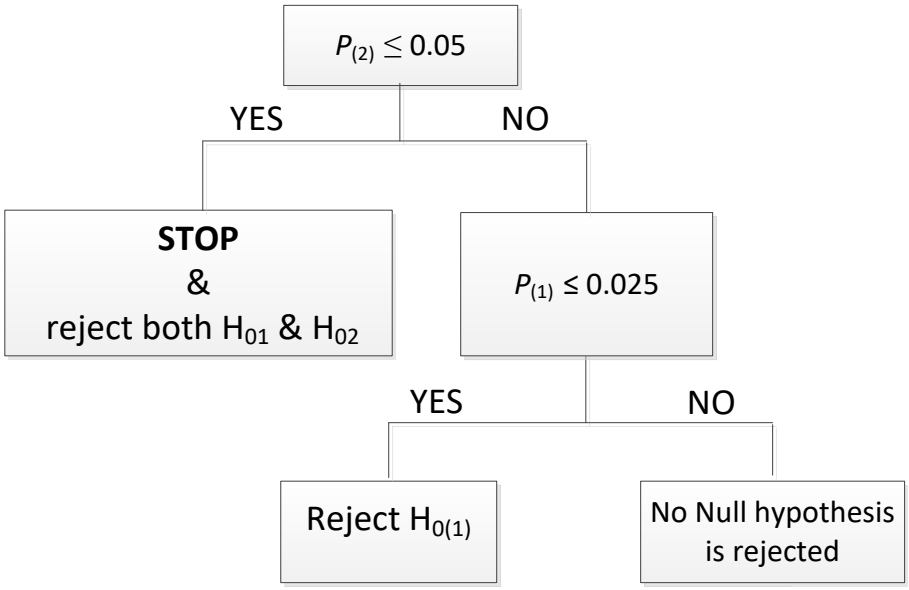
- H_{a1}: the survival function for PFS_{BIRC} is different between the IMGN853 arm and the IC chemotherapy arm in the ITT population, and
- H_{a2}: the survival function for PFS_{BIRC} is different between the IMGN853 arm and the IC chemotherapy arm in the FRα high-expression subgroup.

The Hochberg procedure will be used to control the overall type I error. Assume the log-rank p-values for the 2 primary endpoints are $P_{(1)}$ and $P_{(2)}$ ($P_{(1)} \leq P_{(2)}$). Let $H_{0(1)}$ be the null hypothesis corresponding to $P_{(1)}$ and $H_{0(2)}$ the hypothesis corresponding to $P_{(2)}$. The Hochberg will first test $P_{(2)}$ against 0.05.

- If $P_{(2)} \leq 0.05$, then reject both H_{01} and H_{02} and claim statistical significance in both endpoints.
- If $P_{(2)} > 0.05$, then $P_{(1)}$ will be tested against 0.025.
  - If $P_{(1)} \leq 0.025$, then reject $H_{0(1)}$ and claim statistical significance for that endpoint.
  - If $P_{(1)} > 0.025$, then no null hypothesis will be rejected.

This Hochberg procedure is illustrated in Figure 3.

Figure 3: Hochberg Procedure



3.11.2. Key Secondary Endpoints

For the 3 key secondary endpoints, a hierarchical testing procedure will be applied to control the overall type I error only if both null hypotheses for the primary endpoints are rejected using the Hochberg method. The order of the hierarchical testing is as follows:

- ORR_{BIRC}
- OS
- Primary PRO endpoint as described in Protocol Section 11.7 (primary endpoint for quality of life [QoLP])

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

- H_{03}: the ORR_{BIRC} is the same between the IMGN853 arm and the IC chemotherapy arm in the ITT population.

And the alternative hypothesis is as follows:

- H_{a3}: the ORR_{BIRC} is different between IMGN853 arm and the IC chemotherapy arm in the ITT population.

The null hypothesis for OS is as follows:

- H_{04}: the survival function for OS is the same between the IMGN853 arm and the IC chemotherapy arm in the ITT population.

And the alternative hypothesis is as follows:

- H_{a4}: the survival function for OS is different between the IMGN853 arm and the IC chemothearpy arm in the ITT population.

The null hypothesis for QoLP is as follows:

- H_{05}: the QoLP is the same between the IMGN853 arm and the IC chemotherapy arm in the ITT population.

And the alternative hypothesis is as follows:

- H_{a5}: the QoLP is different between the IMGN853 arm and the IC chemotherapy arm in the ITT population.

This procedure first tests H_{03} at a 2-sided alpha level of 0.05. If H_{03} is rejected favoring the IMGN853 arm, statistical significance will be claimed for ORR_{BIRC}, and H_{04} will be tested at a 2-sided alpha level of 0.05. If H_{04} is rejected favoring the IMGN853 arm, statistical significance will be claimed for OS, and H_{05} will be tested at a 2-sided alpha level of 0.05.

The procedure will stop at the first test not significant at a 2-side alpha level of 0.05 and no claims will be made for that endpoint and subsequent endpoints.

3.12. Analysis Populations

Five 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			
	Baseline	Patient Disposition	Efficacy	Safety
Screened	X			
ITT	X	X	X	
Per Protocol			X	
Response Evaluable			ORR only	
CA-125 Evaluable			CA-125 only	
Safety				X

CA-125 = cancer intigen 125; ITT = intent-to-treat; ORR = objective response rate.

3.12.1. Screened

The Screened population includes all patients entered into the eCRF system who have signed an informed consent.

3.12.2. ITT Population

The ITT population is defined as all patients randomized in the Phase 3 portion of the study (Amendment 6 and onwards).

3.12.3. Per Protocol Population

Per Protocol Population is defined as all patients randomized in the Phase 3 portion of the study who have received at least 1 dose of the study drug, excluding patients with protocol deviations in the following categories: patients -

- whose tumors have FRα low expression (<50% tumor staining at ≥ 2+ intensity), and/or
- who have received four (4) or more prior lines of cancer therapy, and/or
- who are not secondary platinum resistant.

3.12.4. Response-Evaluable Population

The Response-Evaluable population is defined as all patients randomized in the Phase 3 portion of study who have received at least 1 dose of study drug, undergone baseline and at least 1 post-baseline tumor assessment, or who died within 105 days of randomization.

3.12.5. CA-125-Evaluable Population

The CA-125-Evaluable population is defined as all patients randomized in the Phase 3 portion of the study whose pretreatment sample 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.

3.12.6. Safety Population

All enrolled patients who received at least 1 dose of IMGN853 or IC chemotherapy will be included in the Safety population.

3.13. Data Display Characteristics

Data displays produced for this study will include summary tables, data listings, and figures. Data listings will be produced for all recorded data as described in Section 4, Section 5, Section 6, and Section 8. Summary tables will be produced as specified in Section 4, Section 5, Section 6, and Section 8. Additional data listings will be produced for outcome measures that involve extensive procedures to derive the analyzed outcomes. Figures will be produced when specified in Section 4, Section 5, and Section 6.

Stage 1 data will be provided in listings only. As there are only 4 patients in Stage 1, no tables or figures will be provided. Tables, listings, and figures will be provided for the Phase 3 portion of the study. Stage 1 and Phase 3 data will be presented separately.

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, unless described otherwise in Section 4, Section 5, and Section 6, and Section 8. For safety

analyses, the columns of the summary tables will be IMGN853 6mg/kg Q3W, Pac, PLD, Topo Q3W or Q4W, and Total Investigator's Choice. For efficacy analyses, the columns of the summary tables will be Arm 1: IMGN853 6mg/kg Q3W and Arm 2: Investigator's Choice.

The summary statistics displayed will be a function of the type of data associated with the summarized assessment. Unless stated otherwise in relevant sections, Section 4, Section 5, and Section 6, 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). For Stage 1, listings of the following information will be provided for the Safety population by the treatment actually received (actual treatment).

If the primary efficacy endpoint (PFS_{BIRC}) is only statistically significant in patients with FRα high expression but not in the ITT, the following subgroup tables will be prepared for patients with FRα high expression.

- Demographics and baseline disease characteristics
- Prior cancer therapy

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 Y Prime 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 and childbearing potential yes/no
- Ethnicity
- Race

All demography 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. Note that AIBW and BSA are collected depending on the treatment group to which a patient is assigned. Thus, baseline AIBW and BSA will only be summarized for the treatment groups for which this data is available.

4.1.3. Medical History

Medical history will be coded using Medical Dictionary for Regulatory Affairs (MedDRA) (Version 18.1 or later), 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:

- Primary diagnosis (epithelial ovarian cancer, fallopian tube cancer, primary peritoneal cancer, other).
- FRα status
- Histology
- Disease stage at initial diagnosis
- Histologic grade
- Prior radiotherapy (yes/no)
- Prior systemic therapy
  - The number of prior systemic therapies is defined as the number of distinct treatment regimens
    - Adjuvant ± Neoadjuvant will be considered 1 regimen
    - Maintenance therapy will be considered part of the preceding regimen
    - PARP inhibitor therapy will be considered a prior treatment unless given as maintenance therapy
    - Hormonal therapy will not be considered prior systemic treatment (included in this category are tamoxifen, anastrozole, exemestane, megasterol, and letrozole)
- Prior surgery (yes/no)
- Prior exposure to bevacizumab (yes/no)
- Any BRCA mutations mutations (yes[BRCA1, BRCA2]/no)

4.2. Patient Disposition

4.2.1. Screened and Enrolled Patients

The number and percentage of patients who were screened and enrolled (randomized) will be summarized by overall, geographic region (North America and Europe), 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.

A listing of End-of-Treatment and End-of-Study information will be provided for the ITT population. For Stage 1, a listing of End-of-Treatment and End-of-Study information will be provided for the Safety population.

A hierarchical table of the populations in which screened patients were included will summarize the relationship to the analysis populations. The number of screened patients will be provided in the Total column. Given population definitions, the ITT population will be treated as the starting, parent population. The number and percentage of ITT patients included in the Safety population and Per-protocol population 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 summary of the number of patients with any protocol deviation will be provided by treatment group for the ITT population.

Protocol deviations will also be summarized by the following categories, as collected in the protocol deviation log:

- Those who entered the study even though they did not satisfy the entry criteria.
- Those who developed withdrawal criteria during the study but were not withdrawn.
- Those who received the wrong treatment or incorrect dose.
- Those who received an excluded concomitant treatment.

The summary will be grouped first by deviation categories and then by site within each deviation category.

A listing from the protocol deviation log will be provided, grouped by site for the ITT population. The Stage 1 listing will use the Safety population.

5. EFFICACY ANALYSES

The main 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. Best Overall Response

Best overall response (BOR) for a patient is the best response designation (as assessed by either the BIRC or the Investigator, 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 randomisation). The confirmatory scan is valid following treatment discontinuation as long as the patient has not started a new anti-cancer therapy.

5.1.2. 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.

5.1.3. 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 best overall response 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.

Patients who started a new anti-cancer therapy prior to documented PD or death will be censored at the last radiological assessment prior to initiation of new anti-cancer therapy. Patients who did not have PD or death will be censored at the date of their last radiological assessment. Patients who had PD or death after missing 2 or more consecutive radiological assessments (PD or death date - last radiological assessment date +1 ≥ 105 days) will be censored at the date of their last radiological assessment. When an analysis cutoff date is implemented, only data (deaths or radiological assessments) occurring on or before the cutoff date will be used for analysis.

5.1.4. ORR

The ORR will be calculated as the number of patients with a BOR of CR or PR divided by the number of patients in the ITT population. Patients without at least 1 post-baseline RECIST assessment will be treated as non-responders (ie, these patients will contribute to the denominator, but not the numerator).

5.1.5. 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 whose date of death or the last date at which the patient was known to be alive (eg, after the data cutoff date) will be censored at the analysis cutoff date.

5.1.6. 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 BIRC or the Investigator (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[a]

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

Situation	Date of PFS Event or Censoring	Outcome
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)	Date of last adequate radiological assessment showing no PD	Censored

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

5.1.7. RECIST-Related Endpoints – BIRC or Investigator's Assessment

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

5.1.8. 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 confirmed CR or PR.

5.1.9. 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 long-term follow-up form.
- Date of death.
- Date of end of next-line treatment.
  - For next-line treatment, any therapy other than radiotherapy will be considered.
- Any event of PD reported during the long-term follow-up evaluation will be considered, including PD in patients who did not receive subsequent next-line treatment.

- Patients who did not experience PD (BIRC) during the study period but who had PD documented on the 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 long-term 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 BIRC
  - In all patients randomized on the study (ie, $PFS_{BIRC,All}$)
  - In patients with high FRα level (≥ 75% of tumor staining at ≥ 2+ intensity) (ie, $PFS_{BIRC,High}$).

The distribution of $PFS_{BIRC,All}$ and $PFS_{BIRC,High}$ 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 $PFS_{BIRC,All}$ and $PFS_{BIRC,High}$ will also be reported at 3 month intervals (eg, 3 months, 6 months, etc.). The RMST will be compared between treatments stratified by the randomization stratification factors (Zucker DM 1998). As a sensitivity analysis, the RMST from an unstratified analysis will also be provided. Stratified analysis using values collected on the eCRFs will be conducted as sensitivity analysis.

For $PFS_{BIRC,All}$, the strata will be FRα expression level (high vs medium), IC chemotherapy (Pac vs PLD vs Topo), and the number of prior lines of therapy (1 or 2 vs 3). For $PFS_{BIRC,High}$, the strata will be IC chemotherapy (Pac vs PLD vs Topo), and the number of prior lines of therapy (1 or 2 vs 3).

The hazard ratio for $PFS_{BIRC,All}$ and $PFS_{BIRC,High}$ 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 hazard ratio (IMGN853 arm vs IC chemotherapy arm) will be reported using the maximum likelihood estimate along with 95% CI. The IC chemotherapy 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 an IMGN853 benefit over IC chemotherapy.

Kaplan-Meier estimates of PFS and RMST for PFS at 3, 6, 9, and 12 months will also be compared between the IMGN853 and IC chemotherapy arms.

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

For diagnostic purposes, the proportional hazard assumption will be checked using appropriate approaches; including comparing the hazard ratios from stratified and unstratified Cox models, as well as checking graphs (ie, Kaplan-Meier plots: plot of ln(-ln(s(time))) vs time or ln(time), and plots of Scheonfeld residuals vs time). A summary of these checks will be provided to the Sponsor.

5.3. Secondary Efficacy Analyses

The protocol describes the following key secondary efficacy endpoints:

- ORR per RECIST Version 1.1 criteria based on the BIRC assessment (ORR_{BIRC}).
- OS.
- Primary PRO endpoint as described in the PRO analysis appendix.

Additionally, the following efficacy endpoints will also be analyzed:

- PFS as assessed by the Investigator (PFS_{INV}).
- ORR per RECIST Version 1.1 as assessed by the Investigator (ORR_{INV}).
- DOR based on BIRC assessment (DOR_{BIRC}).
- DOR based on Investigator's assessment (DOR_{INV}).
- GCIG CA-125 criteria clinical response rate.
- TTR based on BIRC assessment (TTR_{BIRC}).
- TTR based on Investigator's assessment (TTR_{INV}).

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

As a sensitivity analysis, ORR_{BIRC} and OS will also be summarized on the Per-protocol population.

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

DOR will be summarized in patients with confirmed CR or PR only.

ORR and GCIG CA-125 response will be analyzed using a Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors to test for differences between the IMGN853 arm and the IC chemotherapy arm. P-value and 95% CIs for the proportion of patients with ORR or GCIG CA-125 clinical response will be provided.

The distribution of time-to-event variables (PFS, DOR, and OS) will be summarized using the Kaplan-Meier method. PFS, DOR, and OS 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% CIs 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.

The hazard ratio for the time-to-event variables (PFS, DOR, and 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 hazard ratio (IMGN853 arm vs IC chemotherapy arm) will be reported using the maximum likelihood estimate along with 95% CI. The IC chemotherapy 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 an IMGN853 benefit over IC chemotherapy.

For diagnostic purposes, the proportional hazard assumption will be checked using appropriate approaches; including comparing the hazard ratios from stratified and unstratified Cox models, as well as checking graphs (ie, Kaplan-Meier plots: plot of ln(-ln(s(time))) vs time or ln(time), and plots of Scheonfeld residuals vs time). A summary of these checks will be provided to the Sponsor.

TTR will be summarized using descriptive statistics for patients who achieved a confirmed response of CR or PR as assessed by BIRC or the Investigator (TTR_{BIRC} and TTR_{INV}).

The analysis of the primary PRO endpoint will be described in a separate document to the SAP.

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

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 IMGN853.
- Evaluate the association of anti-tumor activity and/or safety with the following:
  - Mutational status and other genomic alterations in tumor samples.
  - Activation status of oncogenic pathways in tumor samples.

  - Expression and polymorphism of drug transporters such as MDR1 (ie, PgP) and other proteins that may influence anti-tumor activity or safety.
  - Soluble FRα level in blood samples.
  - Genotyping of FcγR.
- Evaluate PFS2. PFS2 analysis will be conducted using the same method as in PFS analysis.

5.5. Efficacy Analysis on Subgroups of Patients

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

- BRCA status (positive vs negative/unknown).
- FRα levels (≥ 75% tumor staining at ≥ 2+ intensity vs < 75% tumor staining at ≥ 2+ intensity).
- Age (< 65 year vs ≥ 65 years).
- Baseline Eastern Cooperative Oncology Group Performance State (ECOG PS) (0 vs 1).
- Prior exposure to bevacizumab (yes vs no).
- Number of prior lines of therapy (1 or 2 vs 3).
- Type of IC chemotherapy.
- Location of metastasis (visceral vs non-visceral).
- Race (white vs non-white).

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 (IMGN853 to IC chemo) and its 95% CI.

The forest plot of time to event variables will display the number of subjects, events, hazard ratio (IMGN853 to IC chemo) and its 95% CI.

The summaries for ORR include the ORR and its 95% CI, and difference in ORR (IMGN853 – IC chemo) and its 95% CI.

The forest plot of ORR will display the number of subjects, difference in ORR (IMGN853 – IC chemo) and its 95% CI.

5.6. Additional Sensitivity Analyses

As additional sensitivity analyses, $PFS_{BIRC,All}$ will also be analyzed by:

- Stratifying only by FRα level.
- Stratifying by FRα level, and using the number of prior therapies and IC chemotherapy as covariates.

As additional sensitivity analyses, $PFS_{BIRC,High}$ will also be analyzed by:

- Using the number of prior therapies and IC chemotherapy as covariates.

The primary analyses for PFS are based on the BIRC assessment of radiological assessments only. A sensitivity analysis of PFS based on the Investigator assessment (PFS_{INV}) 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[a]

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 (ie, on RECIST response eCRF, the earliest date 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)	Date of last RECIST assessment	Censored

[a] 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_{BIRC}, 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 $PFSB_{BIRC}$. Table 6 provides the rules for $PFSB_{BIRC}$.

Table 6: $PFSB_{BIRC}$ Definitions[a]

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).	Date of first missed radiological assessment (ie, date of the last adequate assessment + 42 days [ie, 6 weeks]).	Progression

[a] Includes BIRC assessed radiographic progression only.

OR = overall response; PD = disease progression; PFS = progression-free survival.

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

Table 7: $PFSR_{BIRC}$ Definitions[a]

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
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)	Date of last adequate radiological assessment showing no PD	Censored

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

The following sensitivity analyses will also be performed:

- The concordance/discordance rate between the PFS outcomes for the BIRC assessment and the INV assessment will be summarized.
- The concordance/discordance rate between the BIRC Best Overall Response and the INV Best Overall Response will be summarized.
- In cases where the BIRC and INV assessment are different, PFS will be re-calculated using several methods of censoring:

- Method 1 (PFSC): INV PFS of patients with at least 1 scan submitted for BIRC review.
  - The rules for PFSC are the same as those described in Table 4
  - except only patients with at least 1 scan submitted for BIRC review are included in the analysis.
- Method 2 (PFSD): If PFS_{BIRC} was censored but PFS_{INV} had an event, replace PFS_{BIRC} with PFS_{INV}.
  - Table 4 describes the rules for PFS_{BIRC} and PFS_{INV}.
  - PFS_{BIRC} and PFS_{INV} will be calculated prior to calculating PFSD.
  - PFSD will be equal to PFS_{BIRC} except in cases where PFS_{BIRC} is censored, but PFS_{INV} shows an event. In those cases, PFS4 will be equal to PFS_{INV}.
- Method 3 (PFSE): If both PFS_{BIRC} and PFS_{INV} had an event, but the INV date is earlier than the BIRC date, replace BIRC PFS date with INV PFS date.
  - Table 4 describes the rules for PFS_{BIRC} and PFS_{INV}.
  - PFS_{BIRC} and PFS_{INV} will be calculated prior to calculating PFSE.
  - PFSE will be equal to PFS_{BIRC} except in cases where PFS_{BIRC} and PFS_{INV} both indicate an event and the PFS_{INV} event date is earlier than the PFS_{BIRC} event date. In those cases, PFSE will be equal to PFS_{INV}.

5.7. Strategies for Pooling Stratification Factors

This study uses 3 factors to stratify patients for randomization, with a total of 12 stratification levels. At the final analysis, levels may have 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

FRα	Priors	IC Chemotherapy	If n < 12 Pool With
high	1,2	PLD	Next stratum
high	1,2	Pac	Previous stratum
high	1,2	Topo	Previous stratum
high	3	PLD	Next stratum
high	3	Pac	Previous stratum
high	3	Topo	Previous stratum
medium	1,2	PLD	Next stratum
medium	1,2	Pac	Previous stratum
medium	1,2	Topo	Previous stratum
medium	3	PLD	Next stratum
medium	3	Pac	Previous stratum
medium	3	Topo	Previous stratum

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

5.7.1. Rationale for pooling approach

The FRα levels will not be pooled, and the number of prior lines of therapy is an important prognostic indicator. Therefore, the approach above is to pool patients by the type of IC chemotherapy in consideration prior to randomization.

5.8. Prognostic and Predictive Value of FRα Expression on PFS_{BIRC}

Prognostic value of FRα expression on PFS_{BIRC} will be analyzed for patients in the IC chemo arm. Hazard ratio (FRα high to FRα medium) and its 95% CI will be reported.

Predictive value of FRα expression on PFS_{BIRC} will be analyzed for the ITT population using a Cox proportional hazard regression model with treatment, FRα expression, and their interaction as covariates. The hazard ratio (IMGN853 to IC chemo) will be reported for patients with FRα high expression and FRα medium expression separately. The interaction hazard ratio and its 95% CI will also be reported. The null hypothesis that the hazard ratio (IMGN853 to IC chemo) is the same between patients with FRα high expression and patients with FRα medium expression will be tested using the interaction term.

5.9. 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.

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

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

Listings of efficacy-related data for Stage 1 will be presented using the Safety population and actual treatment received, as follows:

- All lesion assessments (target lesion, non-target lesion, new lesion).
- New anti-cancer therapy.
- Investigator's RECIST assessments.
- CA-125 results.

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.

If the primary efficacy endpoint (PFS_{BIRC}) is only statistically significant in patients with FRα high expression but not in the ITT, the following subgroup tables will also be prepared for patients with FRα high expression.

- Exposure
- Treatment-emergent adverse events (TEAEs)
- Serious TEAEs
- Lab shift tables
- Clinically significant values in liver function tests

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 IMGN853, 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), total cumulative dose (mg), absolute dose intensity (mg/kg/dose for IMGN853, calculated as total cumulative dose [mg])/number of valid drug administration records (performed infusions or held infusions)/AIBW (kg); mg/m^2/dose for remaining investigational products, calculated as total cumulative dose (mg)/number of valid drug administration records (performed infusions or held infusions)/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).

The number of infusions with dose decreased, infusions interrupted, and dose delayed will also be summarized by study drug. The number of infusions with the rate of infusion decreased will also be summarized for the IMGN853 study-drug groups.

The number of patients who received IMGN853 pivotal lots, non-pivotal lots, and different lot numbers will be summarized.

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

For Stage 1, a separate listing will be provided with the information from all study drug administration eCRFs.

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 18.1 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 4.03. 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.

The following AE summary tables will be produced:

- An overall summary of safety will summarize the numbers of patients with TEAEs of each grade and the number of patients who died during the study or within 30 days of last dose.
- All TEAE.s
- Serious TEAEs.
- Non-Serious TEAEs (ie, TEAEs excluding SAEs).
- Grade 3 or higher TEAEs.
- TEAEs leading to drug withdrawal. This subset includes TEAEs with an action taken of drug permanently discontinued.

- TEAEs related to study drug. This table will include TEAEs with a drug relationship of possibly related, probably related, or definitely related. The table will include TEAEs with missing drug relationships. 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. This subset includes all serious TEAEs with a drug relationship of possibly related, probably related, or definitely related. The table will include serious TEAEs with missing drug relationships. 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.
- Deaths on study or within 30 days of the last dose.

The following AE listings will be produced:

- All pre-treatment AEs will be listed.
- 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 related to study drug. This listing will include TEAEs with a drug relationship of possibly related, probably related, or definitely related. This listing will include TEAEs with missing drug relationships.
- Serious TEAEs related to study drug. This listing will include serious TEAEs with a drug relationship of possibly related, probably related, or definitely related. This listing will include serious TEAEs with missing drug relationships.
- TEAEs resulting in death. This listing includes TEAEs with a CTCAE grade of Grade 5 (death) or TEAEs with an outcome of fatal.
- All deaths, with an indication of whether the death occurred on study or within 30 days of the last dose.

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

- Ocular TEAEs
  - A list of PTs for ocular AEs will be provided and finalized by the Sponsor before the final database lock.
- Peripheral neuropathy TEAES
  - A list of PTs for peripheral neuropathy AEs will be provided and finalized by the Sponsor before the final database lock.
- Pneumonitis

For each of the focused safety analysis groups, the following summary tables will be produced:

- TEAEs by SOC, PT, and maximum CTCAE grade.
- Related TEAEs by SOC, PT, and maximum CTCAE grade.
- Serious TEAEs by SOC, PT, and maximum CTCAE grade.
- Related serious TEAEs by SOC, PT, and maximum CTCAE grade.

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.
- Time to first onset of each group of TEAEs.
- Action taken with study drug with respect to each group of TEAEs.

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

- Ocular TEAEs.
- Peripheral neuropathy TEAEs.
- Pneumonitits TEAE.

For Stage 1, the following AE listings will be produced:

- All pre-treatment AEs.
- All AEs, sorted chronologically within 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 related to study drug. This listing will include TEAEs with a drug relationship of possibly related, probably related, or definitely related. The listing will include TEAEs with missing drug relationships.
- Serious TEAEs related to study drug. This listing will include serious TEAEs with a drug relationship of possibly related, probably related, or definitely related. This listing will include serious TEAEs with missing drug relationships.
- TEAEs resulting in death. This listing includes TEAEs with CTCAE Grade 5 (death) or TEAEs with an outcome of fatal.
- All deaths.

6.2.1. Fresh Biopsy Patient TEAE

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.

6.2.2. Infusion-related Reactions

A summary table of infusion-related reactions, occurring 1 day or within 3 days of any infusion will be generated by SOC, PT, and CTCAE grade. A listing of infusion-related reactions occurring within 3 days after any infusion will also be generated. A list of the PTs for infusion related reactions will be determined by the Sponsor prior to final database lock.

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 4.03 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 4.03. Grade 3 or higher laboratory values will be summarized based on the worst grade observed on study.

Clinically significant values in liver function tests 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)
  - > 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

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, respiratory rate, and weight) will be presented in data listings.

Summaries of actual values and changes from baseline will be presented by treatment group for each assessment time point, beginning with the Cycle 1 Day 1 Completion of Infusion time point.

6.5. ECGs

ECG results (rhythm, heart rate, PR interval, RR interval, QRS interval, QT interval, QTcF interval, and classification of within normal limits, abnormal, not clinically significant, and abnormal, clinically significant 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.

The number and percentage of patients in the following non-mutually exclusive categories will be summarized by visit:

- QTc > 450 ms
- QTc > 480 ms
- QTc > 500 ms
- QTc change from baseline > 30 ms
- QTc change from baseline > 60 ms

For the purpose of QTc analysis, derived values using QT and RR will be used if QTcF or QTcB are not collected on the eCRF:

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

$$QTcB = \frac{QT}{\sqrt{RR}}$$

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:

- For Stage 1, medications started after the first dose of study drug and within 30 days of the last dose of study drug 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 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.
- Medications started before the first dose of study drug that are ongoing will be considered concomitant medications.

Prior and concomitant medications will be coded using the September 2015 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, as follows:

- For Stage 1, procedures occurring on or after the first dose of study drug, and within 30 days of the last dose of study drug (by procedure date).
- Procedures are defined as concomitant 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 18.1 or later), 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.

A summary table will be provided of the overall exam result (within normal limits, abnormal, not clinically significant, and abnormal, clinically significant) by treatment group and visit.

6.9. Ocular Symptom Assessments

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

A summary table will be provided of the ocular symptom assessment (within normal limits, abnormal, not clinically significant, and abnormal, clinically significant) by treatment group and visit.

6.10. Corticosteriod 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. Transfusions will be coded using MedDRA Version 18.1 or later.

6.12. Physical Examination

Physical examination results will be presented in data listings.

6.13. Pulmonary Function Tests

Pulmonary function test results will be presented in data listings.

6.14. Pulse Oximetry

Pulse oximetry data were collected only for patients in Stage 1, and will be presented in a data listing.

6.15. ECOG PS

ECOG PS results will be presented in data listings.

6.16. ECHOCARDIOGRAM/MULTIPLE GATED ACQUISITION SCANS

Echocardiogram (ECHO)/multiple gated acquisition (MUGA) scans are taken only for patients receiving PLD. All ECHO/MUGA results will be presented in data listings. Summaries of actual values and changes from baseline will be presented for the PLD treatment group for each assessment time point, beginning with Cycle 4 Day 1.

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

The ADA analyses, including the impact of immunogenicity on PK, will be covered by a separate, independent analysis plan (see Section 10).

Impact of immunogenicity on safety and efficacy will be explored. A by patient TEAE listing for those patients determined seropositive for IMGN853 antibodies will be produced. Summary of TEAEs experienced within 1 and 3 days of ADA+ will be generated for those patients who sero-convert following the first dose of mirvetuximab soravtansine. Efficacy will be analysed by seroconversion status for patients treated with IMGN853.

9. BIOMARKERS

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

10. PHARMACOKINETICS

The PK analyses will be covered by a separate, independent analysis plan. The PK analyses will be performed using the plasma collected for PK purposes.

REFERENCES

Zucker DM, Restricted mean life with covariates: modification and extenstion of a useful survival analysis method, J Am Stat Assoc 93: 702-709, 1998

APPENDIX 1: TEMPLATE SAS® CODE

This appendix provides template SAS® code (Version 9.3 or higher) that will be used for analysis.

Template code for the ORR, DCR, and GCIG CA-125 response estimates along with the 95% CIs:

```
proc freq data=AD_RESP;
by TRTP;
tables AVAL / binomial alpha=0.05;
run;
```

Template code for the stratified CMH test:

```
proc freq data=AD_RESP;
tables STRATUM*TRTP*AVAL / cmh;
run;
```

Template code for the unstratified chi-squared test:

```
proc freq data=AD_RESP;
tables TRTP*AVAL / chisq;
run;
```

Template code for median time-to-event estimates along with 95% confidence intervals:

```
proc lifetest data=AD_TTE;
time AVAL*CNSR(1);
strata TRTP;
run;
```

Template code for the stratified log-rank test:

```
proc lifetest data=AD_TTE;
time AVAL*CNSR(1);
strata STRATUM / group=TRTP;
run;
```

Template code for the unstratified log-rank test:

```
proc lifetest data=AD_TTE;
time AVAL*CNSR(1);
strata TRTP;
run;
```

Template code for the stratified Cox proportional hazards model:

```
proc phreg data=AD_TTE;
        class trtp;
model AVAL*CNSR(1) = TRTP / risklimits;
strata STRATUM;
run;
```

Template code for the unstratified Cox proportional hazards model:

```
proc phreg data=AD_TTE;
model AVAL*CNSR(1) = TRTP / risklimits;
run;
```

where:

AD_RESP is the name of the ADaM dataset used for response analysis

AD_TTE is the name of the ADaM dataset used for time-to-event analysis

TRTP is a variable containing the treatment used for analysis.

AVAL is a variable containing the value for analysis (eg, objective response or time to death)

CNSR is a variable indicating whether the AVAL is censored (1=censored, 0=non-censored)

STRATUM is a variable containing all levels of the randomization stratification factors

APPENDIX 2: MONTE-CARLO SIMULATION REPORT

This appendix describes the assumptions, algorithms, and results of Monte-Carlo simulation studies evaluating the type I error and power of this study under different scenarios.

Randomization and Stratification

The study will randomize 333 patients over a period of 21 months with a randomization ratio of 2:1 (222 to the IMGN853 arm and 111 to the IC chemotherapy arm). The randomization will be stratified by the following factors:

- FRα expression level by IHC (≥ 75% tumor staining at ≥ 2+ intensity vs ≥ 50% and < 75% tumor staining at ≥ 2+ intensity).
- Number of prior lines of therapy (1 or 2 vs 3).
- IC chemotherapy (Pac, PLD or Topo).

Interim Futility Analysis

The study will have an interim futility analysis when 80 PFS events have occurred in the ITT population. If the hazard ratio (HR, IMGN853 relative to IC chemotherapy) estimate for PFS is greater than 1 in both the ITT population and the FRα high expression subgroup, the study will be stopped for futility. Otherwise, the study will continue to full enrollment of 333 patients in the ITT population.

Primary Efficacy Endpoints

The final analysis will occur when at least 236 PFS events have occurred in the ITT population and 2 primary efficacy endpoints will be analyzed as follows:

- PFS in the ITT population.
- PFS in the FRα high-expression subgroup.

Hypothesis Testing

The null hypotheses are as follows:

- H_{01}: the survival function for PFS is the same between the IMGN853 arm and the IC chemotherapy arm in the ITT population, and
- H_{02}: the survival function for PFS is the same between the IMGN853 arm and the IC chemotherapy arm in the FRα high-expression subgroup.

And the alternative hypotheses are as follows:

- H_{a1}: the survival function for PFS is different between the IMGN853 arm and the IC chemotherapy arm in the ITT population, and
- H_{a2}: the survival function for PFS is different between the IMGN853 arm and the IC chemotherapy arm in the FRα high-expression subgroup.

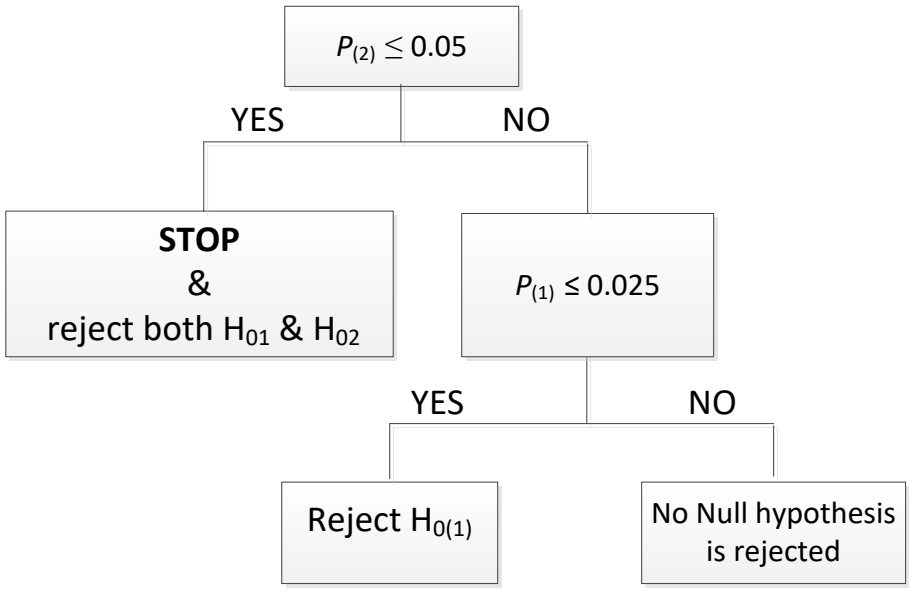
Each endpoint will be tested using a stratified log-rank test (stratified by the stratification factors used in randomization).

Multiple Testing Procedure

The Hochberg procedure will be used to control the study-wise type I error. Assume the log-rank p-values for the 2 endpoints are $P_{(1)}$ and $P_{(2)}$ ($P_{(1)} \leq P_{(2)}$). Let $H_{0(1)}$ be the null hypothesis corresponding to $P_{(1)}$ and $H_{0(2)}$ the hypothesis corresponding to $P_{(2)}$. The Hochberg will first test $P_{(2)}$ against 0.05.

- If $P_{(2)} \leq 0.05$, then reject both H_{01} and H_{02} and claim statistical significance in both endpoints.
- If $P_{(2)} > 0.05$, then $P_{(1)}$ will be tested against 0.025.
  - If $P_{(1)} \leq 0.025$, then reject $H_{0(1)}$ and claim statistical significance for that endpoint.
  - If $P_{(1)} > 0.025$, then no null hypothesis will be rejected.

This Hochberg procedure is illustrated as follows:



Simulation Assumptions

The following assumptions are used in the Monte-Carlo simulation studies.

- Monthly enrollment rate increases linearly during first 6 months, followed by uniform enrollment during months 7-21.
- PFS follows exponential distribution for the IC chemotherapy arm, FRα medium-expression subgroup in the IMGN853 arm, and FRα high-expression subgroup in the IMGN853 arm, respectively.
- Censoring process follows an exponential distribution with 1-year censoring rate of 20% for both arms.

- Ratio of number of patients between FRα high expression and medium expression is 2:1 or 1:1.
- Median PFS for the IC chemotherapy arm is 3.5 months.
  - Number of trials under each scenario is 10000.

Simulation Results

Under different scenarios, the probability of stopping for futility at interim and the probability of rejecting H_{01} (for the ITT population) and H_{02} (for FRα high-expression subgroup) are summarized in Table 9. If the FRα high/medium ratio is 2:1, there is 39% probability of stopping for futility at interim if both null hypotheses H_{01} and H_{02} are true and the overall type I error is controlled at 4.0%. Using the Hochberg procedure, the study will have a 96% and 91% power to detect a median PFS of 6.0 months in the IMGN853 arm (HR = 0.583) in the ITT population and FRα high-expression subgroup, respectively. If the FRα high/medium ratio is 1:1, there is 37% probability of stopping for futility at interim if both null hypotheses H_{01} and H_{02} are true and the overall type I error is controlled at 3.6%. Using the Hochberg procedure, the study will have a 96% and 81% power to detect a median PFS of 6.0 months in the IMGN853 arm (HR = 0.583) in the ITT population and FRα high-expression subgroup, respectively.

Table 9: Summary of Monte-Carlo Simulations

True Median PFS (months)			FRα High/ Medium Ratio	Probability of Stopping for Futility at Interim Analysis	Probability of Rejecting	
IC Chemo	IMGN853 (FRα medium)	IMGN853 (FRα high)			H_{01}	H_{02}
3.5	3.5	3.5	2:1	39%	2.1%	1.9%
3.5	3.5	6.0	2:1	2.1%	75%	85%
3.5	4.5	6.0	2:1	1.2%	89%	89%
3.5	5.0	6.0	2:1	0.8%	93%	90%
3.5	5.5	6.0	2:1	0.6%	95%	90%
3.5	6.0	6.0	2:1	0.5%	96%	91%
3.5	3.5	3.5	1:1	37%	1.9%	1.7%
3.5	3.5	6.0	1:1	3.4%	51%	72%
3.5	4.5	6.0	1:1	1.8%	80%	77%
3.5	5.0	6.0	1:1	1.2%	88%	79%
3.5	5.5	6.0	1:1	0.8%	93%	80%
3.5	6.0	6.0	1:1	0.5%	96%	81%

Chemo = chemotherapy; FRα = folate receptor alpha; IC = Investigator's choice; PFS = progression-free survival.