STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 2, Randomized, Open-Label Study Evaluating the Safety and Efficacy of Magrolimab in Combination With Bevacizumab and FOLFIRI Versus Bevacizumab and FOLFIRI in Previously Treated Advanced Inoperable Metastatic Colorectal Cancer (mCRC)
Name of Test Drug:	Magrolimab in Combination With Bevacizumab and FOLFIRI
Study Number:	GS-US-587-6156
Protocol Version (Date):	Amendment 5 (31 October 2023)
Analysis Plan Version:	1.0
Analysis Plan Date:	16 July 2024
Analysis Plan Author(s):	PPD

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF ABBREVIATIONS

ADA	anti-drug antibody
AE	adverse event
AECI	adverse events of clinical importance
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate transaminase
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
BOR	best overall response
BSA	body surface area
CBR	clinical benefit rate
CI	confidence interval
CR	complete response
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EOT	end of treatment
KM	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified intent to treat
NA	not applicable
NCI	National Cancer Institute
NE	not evaluable
NTLs	non-target lesions
ORR	objective response rate
OS	overall survival
PD	disease progression
PFS	progression-free survival
PK	pharmacokinetics
PR	partial response
PT	preferred term
Q1, Q3	first quartile, third quartile
RECIST	Response Evaluation Criteria in Solid Tumors (version 1.1)
SAE	serious adverse event
SAP	statistical analysis plan

SD	stable disease
SE	standard error
SMQ	Standard MedDRA Query
StD	standard deviation
SOC	system organ class
TLs	target lesions
ULN	upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and defines key elements including variable definitions for the final analysis of data of Study GS-US-587-6156 in support of the clinical study report (CSR). This SAP is based on the study protocol Amendment 5 dated 31 October 2023. Any changes made after the finalization of the SAP will be documented in the CSR.

Analysis methods specified in this document take precedence over those described in protocol should there be any difference.

1.1. Study Objectives and Endpoints

Primary Objectives	Primary Endpoints
Safety Run-in Cohort: • To evaluate the safety, tolerability, and recommended Phase 2 dose (RP2D) of magrolimab in combination with bevacizumab and FOLFIRI in previously treated patients with advanced inoperable mCRC Randomized Cohort: • To evaluate the efficacy of magrolimab in combination with bevacizumab and FOLFIRI in mCRC as determined by PFS by investigator assessment	Safety Run-in Cohort: • Incidence of dose-limiting toxicities (DLTs), adverse events (AEs), and laboratory abnormalities according to National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0 Randomized Cohort: • PFS, defined as the time from the date of randomization until the earliest date of documented disease progression as determined by investigator assessment using Response Evaluation Criteria in Solid Tumors Version 1.1 (RECIST V1.1), or death from any cause, whichever occurs first
Secondary Objectives	**Secondary Endpoints**
Randomized Cohort: • To evaluate objective response rate (ORR) by investigator assessment • To evaluate additional measures of efficacy of magrolimab in combination with bevacizumab and FOLFIRI, including duration of response (DOR) and OS • To evaluate patient-reported outcomes (PRO)/quality-of-life measures for the Randomized Cohort in mCRC with magrolimab in combination with bevacizumab and FOLFIRI Safety Run-in and Randomized Cohort: • To evaluate the PK and immunogenicity of magrolimab in combination with bevacizumab and FOLFIRI	Randomized Cohort: • Confirmed ORR, defined as the proportion of patients with complete response (CR) or partial response (PR) on 2 assessments, at least 28 days apart, as determined by investigator assessment using RECIST V1.1 • DOR, defined as time from first documentation of CR or PR to the earliest date of documented disease progression as determined by investigator assessment, per RECIST V1.1, or death from any cause, whichever occurs first • OS, defined as time from date of randomization to death from any cause • PRO assessments (European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Core Questionnaire [EORTC-QLQ-C30], the 5-level EuroQol 5 dimensions questionnaire [EQ-5D-5L]) scores, and Functional Assessment of Cancer Therapy [FACT] Colorectal Symptom Index [FCSI]

	Safety Run-in and Randomized Cohorts: • Magrolimab concentration versus time and antidrug antibodies (ADA) to magrolimab
Exploratory Objectives	**Exploratory Endpoints**
CCI	

1.2. Study Design

1.2.1. Study Design Overview

This is a Phase 2, randomized, open-label, multicenter study to evaluate magrolimab in combination with bevacizumab and FOLFIRI in previously treated patients with advanced inoperable mCRC CCI

This study will consist of the following 2 cohorts:

- **Safety Run-in Cohort:** magrolimab in combination with bevacizumab and FOLFIRI in previously treated patients with advanced inoperable mCRC

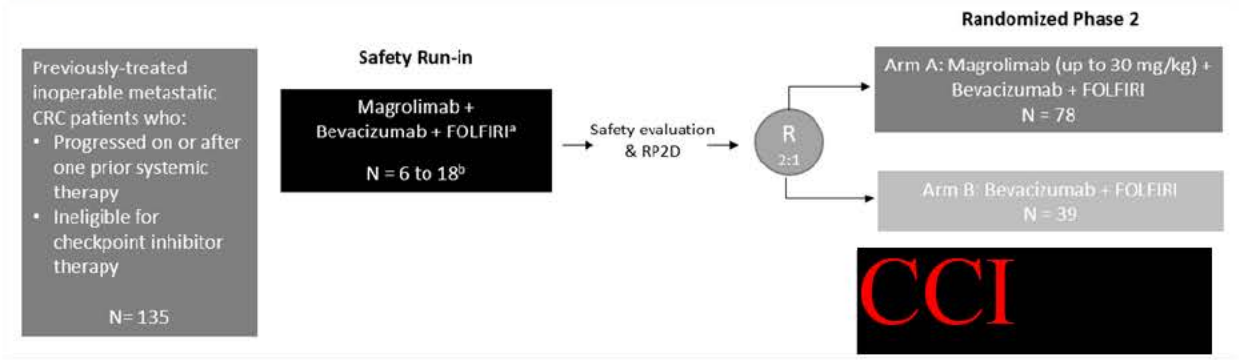
After completion of the Safety Run-in Cohort, the Randomized Cohort will be open to enrollment.

- **Randomized Cohort:** magrolimab in combination with bevacizumab and FOLFIRI (Experimental Arm A) versus bevacizumab and FOLFIRI (Control Arm B) in previously treated patients with advanced inoperable mCRC.

Approximately 135 patients may be enrolled in the study, with approximately 6 to 18 patients in Safety Run-in Cohort (additional patients could be enrolled in this cohort or in dose de-escalation cohorts) and approximately 117 patients in the Randomized Cohort Arms A and B.

The study schema is presented in Figure 1-1.

Figure 1-1. Study Schema



5-FU = 5-fluorouracil; CRC = colorectal cancer; FOLFIRI = 5-fluorouracil, irinotecan, and leucovorin; CCI [redacted]; RP2D = recommended Phase 2 dose; CCI [redacted]

a FOLFIRI: irinotecan 180 mg/m^2, leucovorin 400 mg/m^2, fluorouracil 400 mg/m^2

b Additional patients may be enrolled in the safety run-in or in dose de-escalation cohorts

1.2.2. Safety Run-in Cohort

Initially, approximately 6 patients will be enrolled in the Safety Run-in Cohort at a starting dose level. A DLT assessment period of 28 days will occur.

Although no dose-dependent toxicities have been observed with magrolimab, in order to preserve the efficacious doses of the combination partner drugs, dose de-escalation will take place for magrolimab as follows:

- CCI [redacted]

[redacted]

1.2.3. Randomized Cohort

Once the Safety Run-in Cohort is completed and the RP2D for magrolimab in combination with bevacizumab and FOLFIRI is determined, the sponsor will open the Randomized Cohort.

In this open-label, randomized, 2-arm study, patients with mCRC will be randomized in a 2:1 ratio to receive either magrolimab in combination with bevacizumab and FOLFIRI (Experimental Arm A) or bevacizumab and FOLFIRI (Control Arm B).

CCI [redacted]

1.2.4. Dosage and Administration of Study Treatment

The study treatment dosing regimen for the safety run-in and randomized cohorts is presented in Table 1-1 and Table 1-2.

Table 1-1. Safety Run-in Cohort Dosing and Dose De-escalation Regimen

Drug/Dose/Route	Dose Schedule (Day per 28-Day Cycle)		
	Cycle 1	Cycle 2	Cycle 3+
Bevacizumab 5 mg/kg IV every 2 weeks[a]	Days 1, 15	Days 1, 15	Days 1, 15
FOLFIRI IV[b] Irinotecan 180 mg/m^2 over 30-90 minutes on first day of dose administration Leucovorin 400 mg/m^2 over 2 hours on first day of dose administration[c] Fluorouracil 400 mg/m^2 bolus on first day of dose administration, followed by 2400 mg/m^2 over 46 hours, continuous infusion	Days 1, 15	Days 1, 15	Days 1, 15
Magrolimab Administration			
Magrolimab 1 mg/kg IV (3 hours ± 30 min)	Day 1		
Magrolimab starting dose level 30 mg/kg IV (2 hours ± 30 min)	QW beginning at Day 8 visit and the next 6 doses (Cycle 1 Days 8, 15, and 22, Cycle 2 Days 1, 8, 15, and 22)		
Magrolimab 30 mg/kg IV (2 hours ± 30 min)	Q2W beginning 1 week after the last weekly 30 mg/kg dose (starting Cycle 3 Day 1 onward)		
Magrolimab de-escalation Level −1 20 mg/kg IV (2 hours ± 30 min)	QW beginning at Day 8 visit and the next 6 doses (Cycle 1 Days 8, 15, and 22, Cycle 2 Days 1, 8, 15, and 22) Q2W beginning 1 week after the last weekly 20 mg/kg dose (starting Cycle 3 Day 1 onward)		
Magrolimab de-escalation Level −2 15 mg/kg IV (2 hours ± 30 min)	QW beginning at Day 8 and the next 6 doses (Cycle 1 Days 8, 15, and 22, Cycle 2 Days 1, 8, 15, and 22) Q2W beginning 1 week after the last weekly 15 mg/kg dose (starting Cycle 3 Day 1 onward)		

FOLFIRI = 5-fluorouracil, irinotecan, and leucovorin; IV = intravenous; QW = every week; Q2W = every 2 weeks

a Bevacizumab should be administered per standard of care and/or institutional guidelines. CCI [redacted]

b FOLFIRI should be administered per standard of care and/or institutional guidelines. Recommendations include administering the first dose of irinotecan over 30-90 minutes. CCI [redacted]

c Levoleucovorin 200 mg/m^2 may be used if leucovorin is unavailable. Generics for leucovorin or levoleucovorin are also permitted. Leucovorin and levoleucovorin should be administered per standard of care and/or institutional guidelines. Recommendations include administering the first dose over 2 hours. Different leucovorin doses may be used if recommended by regional or institutional guidelines.

Table 1-2. Randomized Cohort Dosing Regimen

Drug/Dose/Route	Dose Schedule (Day per 28-Day Cycle)		
	Cycle 1	Cycle 2	Cycle 3+
Bevacizumab 5 mg/kg IV every 2 weeks[a]	Days 1, 15	Days 1, 15	Days 1, 15
FOLFIRI IV[b] Irinotecan 180 mg/m^2 over 30-90 minutes on first day of dose administration Leucovorin 400 mg/m^2 over 2 hours on first day of dose administration[c] Fluorouracil 400 mg/m^2 bolus on first day of dose administration, followed by 2400 mg/m^2 over 46 hours, continuous infusion	Days 1, 15	Days 1, 15	Days 1, 15
Magrolimab Administration			
Magrolimab 1 mg/kg IV (3 hours ± 30 min)	Day 1		
Magrolimab RP2D IV (2 hours ± 30 min)	QW beginning at Day 8 visit and the next 6 doses (Cycle 1 Days 8, 15, and 22, Cycle 2 Days 1, 8, 15, and 22)		
Magrolimab RP2D IV (2 hours ± 30 min)	Q2W beginning 1 week after the last weekly RP2D (starting Cycle 3 Day 1 onward)		

FOLFIRI = 5-fluorouracil, irinotecan, and leucovorin; IV = intravenous; QW = every week; Q2W = every 2 weeks;
CCI

a Bevacizumab should be administered per standard of care and/or institutional guidelines. CCI

b FOLFIRI should be administered per standard of care and/or institutional guidelines. Recommendations include administering the first dose of irinotecan over 30-90 minutes. CCI

c Levoleucovorin 200 mg/m^2 may be used if leucovorin is unavailable. Generics for leucovorin or levoleucovorin are also permitted. Leucovorin and levoleucovorin should be administered per standard of care and/or institutional guidelines. Recommendations include administering the first dose over 2 hours. Different leucovorin doses may be used if recommended by regional or institutional guidelines.

1.3. Sample Size and Power

CCI

Gilead had decided to terminate the study. By the time Gilead made the decision, 67 patients had been randomized to phase 2 randomized cohort, which is less than the planned sample size of 117 patients.

2. TYPE OF ANALYSIS

2.1. Dose Determination Analysis

CCI [redacted]

2.2. Treatment-related Toxicity Monitoring Analysis

Treatment-related toxicity will be monitored by a Gilead Data Review Committee (GDRC) at a preset frequency with the stopping boundary shown in Table 2-1 after the first 15 patients from Arm A are treated at the dose level for the Randomized Phase 2 Cohort with at least 1 cycle of follow up, and thereafter when safety data from 40 Arm A patients, from the first 3 cycles become available. The frequency of GDRC review will be outlined in the GDRC charter. This is a Pocock-type boundary {Ivanova 2005} that yields the probability of crossing the boundary of at most 80% when the rate of Grade 4 or higher treatment-related AEs is 33%, or the rate of treatment-related deaths is 15%.

Table 2-1. Stopping Boundary Due to Toxicity (Phase 2 Randomized Cohort Arm A)

	N=15	N=40
Grade 4/5 treatment-related TEAEs	≥ 6	≥ 15
Treatment-related deaths	≥ 3	≥ 7

TEAE = treatment-emergent adverse event

2.3. Final Analysis

The final analysis will be performed after all patients have completed the study or discontinued early, outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized.

3. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of patients in each category will be presented; for continuous variables, the number of patients (n), mean, standard deviation (StD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

All statistical tests will be 2-sided and performed at the 5% significance level unless otherwise specified.

3.1. Analysis Sets

For each analysis set, the number and percentage of patients eligible for inclusion will be summarized by treatment group and cohorts.

A listing of reasons for exclusion from analysis sets will be provided by patient.

3.1.1. All Enrolled Analysis Set

All Enrolled Analysis Set includes all patients who received a study subject identification number in the study after screening.

3.1.2. Intent-to-Treat (ITT) Analysis Set

Intent-to-treat (ITT) Analysis Set includes all patients who were randomized in the study. This is the primary analysis set for efficacy analsyis for Randomized Phase 2 Cohort.

3.1.3. Modified ITT (mITT) Analysis Set

The mITT Analysis Set includes all enrolled patients who took at least 1 dose of any study drug.

3.1.4. Safety Analysis Set

The Safety Analysis Set includes all patients who took at least 1 dose of any study drug. This is the primary analysis set for safety analyses, except for DLT analysis.

3.1.5. Dose-Limiting Toxicity (DLT) Analysis Set

For the Safety Run-in Cohort, the primary analysis set for the DLT analysis is the DLT-Evaluable Analysis Set, defined as all patients who meet one of the following criteria in the DLT-evaluable period (defined as the first 28 days):

- Patient experienced a DLT at any time after initiation of the first infusion of magrolimab.
- Patient did not experience a DLT and completes at least 3 infusions of magrolimab and at least 2 doses of bevacizumab and FOLFIRI in the Safety Run-in Cohort.

If a patient experiences a DLT during the DLT-assessment period, the patient will discontinue treatment. Patients who are not evaluable for DLT assessment in the Safety Run-in Cohort will be replaced.

3.1.6. Pharmacokinetic Analysis Set

The PK Analysis Set, defined as all patients who received any amount of magrolimab and have at least 1 measurable posttreatment serum concentration of magrolimab.This is the primary analysis set for all PK analyses.

3.1.7. Immunogenicity Analysis Set

Immunogenicity Analysis Set will include all patients who received any amount of magrolimab and have at least 1 evaluable anti-drug antibody test result.

CCI [redacted]

[redacted]

3.2. Subject Grouping

For efficacy analyses based on the ITT Analysis Set, patients will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set, patients will be grouped according to the cohort they were enrolled in and the actual treatment received as follows:

- Safety Run-in cohort
- Phase 2 Randomized Cohort Arm A
- Phase 2 Randomized Cohort Arm B
- Safety Run-in cohort + Phase 2 Randomized Cohort Arm A

The actual treatment received will differ from the randomized or planned treatment only when their actual treatment differs from randomized or planned treatment for the entire treatment duration.

For the PK Analysis Set, patients will be grouped according to the actual treatment they received. The same applies to the Immunogenicity Analysis Set.

3.3. Strata and Covariates

Due to the early termination of the program, efficacy endpoints will not be evaluated using stratification factors as covariates or stratification variables for analyses.

3.4. Examination of Subject Subgroups

There are no prespecified patient subgroupings for efficacy or safety analyses.

3.5. Adjustment for Multiplicity

This study is exploratory/proof-of-concept in nature and therefore no formal adjustments for multiplicity were made.

3.6. Missing Data and Outliers

3.6.1. Missing Data

In general, missing data will not be imputed unless methods for handling missing data are specified.

3.6.2. Outliers

Outliers will be identified during the data management and data analysis process, but no sensitivity analyses will be conducted. All data will be included in the data analysis.

3.7. Data Handling Conventions and Transformations

The following conventions will be used for the imputation of date of birth when it is partially missing or not collected:

- If only year of birth is collected, then "01 July" will be imputed as the day and month of birth
- If year of birth is missing, then date of birth will not be imputed

In general, age collected at the Study Day 1 (in years) will be used for analyses and presented in listings. If age at Day 1 is not available for a subject, then age derived based on date of birth and the date of Study Day 1 will be used instead. For screen failures or patients who are not enrolled or randomized, the date the first informed consent was signed will be used for the age derivation. Age required for longitudinal and temporal calculations and analyses (eg, estimates of creatinine clearance, age at date of AE) will be based on age derived from date of birth and the date of the measurement or event, unless otherwise specified.

Non-PK data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1,etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used to calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of"≤ x" or "x" (where x is considered the LOQ).

If methods based on the assumption that the data are normally distributed are not adequate, analyses may be performed on transformed data or nonparametric analysis methods may be used, as appropriate.

PK Data Handling

Natural logarithmic transformation will be used for analyzing non-BLQ concentrations and PK parameters. Concentration values that are below the limit of quantitation (BLQ) will be presented as "BLQ" in the concentration data listing. Values that are BLQ will be treated as 0 at predose and postdose time points for summary purposes. The number of samples will be summarized to reflect the actual number of samples assessed at that time point.

At predose, if all concentration values are BLQ, then the mean, and order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as 0 and the rest of the summary statistics (ie, SD and CV) will be missing. If any values are non-BLQ, then the number of samples, order statistics, and all summary statistics will be displayed.

At any given postdose time point, if more than one-third of the participants have a concentration value of BLQ, then only the number of samples and order statistics will be displayed; otherwise, order statistics and summary statistics will be displayed.

The following conventions will be used for the presentation of order statistics for postdose time points:

- If at least 1 participant has a concentration value of BLQ for the time point, the minimum value will be displayed as "BLQ."

- If more than 25% of the participants have a concentration data value of BLQ for a given time point, the minimum and Q1 values will be displayed as "BLQ."

- If more than 50% of the participants have a concentration data value of BLQ for a given time point, the minimum, Q1, and median values will be displayed as "BLQ."

- If more than 75% of the participants have a concentration data value of BLQ for a given time point, the minimum, Q1, median, and Q3 values will be displayed as "BLQ."

- If all participants have concentration data values of BLQ for a given time point, all order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as "BLQ."

PK parameters that are BLQ will be excluded before log transformation or statistical model fitting and displayed as described above.

3.8. Analysis Visit Windows

3.8.1. Definition of Study Day

Study day will be calculated from the first dosing date of any study drug, which is the date of the first dose of magrolimab or the combination drug, whichever occurs first and derived as follows:

- For postdose study days: Assessment Date – First Dosing Date + 1
- For days prior to the first dose: Assessment Date – First Dosing Date

Therefore, Study Day 1 is the date of first dosing of any study drug. If the subject is enrolled or randomized but not dosed, the enrollment or randomization date will be Study Day 1.

3.8.2. Analysis Visit Windows

Subject visits might not occur on protocol-specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows.

The analysis windows for lab assessments are provided in Table 3-1. Post-infusion lab assessments for hemoglobin and hematocrit will use nonimal visit and will be excluded when applying Table 3-1.

Table 3-1. Analysis Visit Windows for Lab By-visit Summaries

Analysis Visit	Nominal Study Day	Visit Window Study Day	
		Lower Limit	Upper Limit
Baseline		(none)	1[a]
Day 2[b]	2	1[c]	2
Week 1	8	3	11
Week 2	15	12	18
Week 3	22	19	25
Week 4	29	26	32
Week 5	36	33	39
Week 6	43	40	46
Week 7	50	47	53
Week 8	57	54	63
Week 10	71	64	77
Week xx[d]	(xx)*7 + 1	(xx)*7 - 6	(xx)*7 + 7

a Prior to first dose date time
b Day 2 visit not applicable to chemistry lab assessments. Chemistry labs assessed post first dose date time through study day 2 will be assigned to "Week 1" visit
c Post first dose date time
d xx >= 12

PRO will use nonimal visit for by-visit summaries of PRO assessments.

3.8.3. Selection of Data in the Event of Multiple Records in the Same Visit

If multiple valid, nonmissing measurements exist in the same visit, records will be chosen based on the following rules if a single value is needed:

- For baseline, the last nonmissing value on or prior to the first dosing date of study drug (and prior to first dosing time) will be selected, unless specified differently. If there are multiple records with the same time or no time recorded on the same day, the baseline value will be the average of the measurements for continuous data, or the measurement with the best severity for categorical data.

- For postbaseline values:

  — The record closest to the nominal study day for that visit will be selected.

  — If there are 2 records that are equidistant from the nominal day, the later record will be selected.

  — If there is more than 1 record on the selected day, the average will be taken for continuous data and the worst severity will be taken for categorical data, unless otherwise specified.

3.9. Assessment of COVID-19 Impact

This study was ongoing during the novel coronavirus (COVID-19) pandemic which has an impact on the study conduct. Some patients were unable to attend onsite visits due to shelter in place guidelines, site closures, or other reasons. The following special situations due to COVID-19 will be handled in the analysis:

- Study treatment or study discontinuation due to COVID-19

- Protocol deviations due to COVID-19

- Missed and Virtual Visits due to COVID-19

- Adverse events due to COVID-19

4. PROTOCOL DEVIATIONS

Patients who did not meet at least one eligibility criterion for study entry, but enrolled in the study will be summarized by treatment group and cohorts regardless of whether they were exempted by the sponsor or not, based on All Enrolled Analysis Set. The summary will also present the number and percentage of patients who did not meet specific criteria.

Protocol deviations occurring after patients entered the study are documented during routine monitoring. The number and percentage of patients with important protocol deviations by deviation reason (e.g., eligibility criteria, informed consent) will be summarized by treatment group and cohorts based on All Enrolled Analysis Set. A by-subject listing will be provided for those patients with important protocol deviation.

5. SUBJECT INFORMATION

Generally, disposition tables and all listings will be based on All Enrolled Analysis Set. Safety Analysis Set will be used for the summary of treatment exposure, unless otherwise specified.

5.1. Patient Enrollment and Disposition

Key study dates, including first patient screened, first patient enrolled, last patient enrolled, last patient last visit for the primary endpoint, and last patient last visit for the clinical study report will be provided.

A summary of patient enrollment will be provided by cohort and treatment, for each country, investigator and overall.

A similar enrollment table will be provided by randomization stratification group for Phase 2 Randomized Cohort only.

A summary of patient disposition will be provided by cohort and treatment. This summary will present the number of patients enrolled or randomized, the number of patients enrolled or randomized but not dosed, and the number of patients in each of the categories listed below:

- Intent-to Treat (ITT) Analysis Set
- Modified ITT (mITT) Analysis Set
- Safety Analysis Set
- Discontinued each study drug (magrolimab, bevacizumab and FOLFIRI) with reasons for treatment discontinuation for the correscponding drug
- Discontinued the study with reasons for discontinuation of study

For the status of study drug and study completion and reasons for discontinuation, the number and percentage of patients in each category will be provided. The denominator for the percentage calculation will be the total number of patients in the All Enrolled Analyses Set corresponding to that column.

In addition, the following by-subject listings will be provided:

- Reasons for study drug discontinuation
- Reasons for study discontinuation

5.2. Extent of Study Treatment Exposure

Extent of exposure to study treatment will be summarized using descriptive statistics for total duration of exposure to each study drug (magrolimab, bevacizumab and FOLFIRI), total number of infusions for each study drug, total number of cycles received for each study drug for each study drug, relative dose intensity (%) of Magrolimab, and number (%) of patients with dose modifications (i.e. infusion interuption, dose delayed or not administered) and reasons by cohort and treatment.

5.2.1. Duration of Exposure to Study Drug

Total duration of exposure to each study drug (magrolimab, bevacizumab and FOLFIRI) will be defined for a subject as last dosing date minus first dosing date plus 1 day, regardless of any temporary interruptions in study drug administration, and will be expressed in weeks using up to 1 decimal place (eg, 4.5 weeks).

The total duration of exposure to each study drug will be summarized using descriptive statistics for continuous variables, as well as using the number (i.e., cumulative counts) and percentage of patients exposed for at least the following time periods: 1 day, 4 weeks, 8 weeks, 12 weeks, 16 weeks, 20 weeks, and 24 weeks, etc.

5.2.2. Relative Dose Intensity

Relative dose intensity is the percentage of the total amount of study drug administered relative to the total amount of study drug expected to be administered during a patient's actual on-treatment period based on the study drug regimen.

Relative dose intensity will be summarized by formulas below for magrolimab.

$$Relative\ dose\ intensity\ (\%) = \left(\frac{Cumulative\ dosage\ received\ \left({}^{mg}/_{kg}\right)}{Total\ planned\ dosage\ of\ specific\ drug\left({}^{mg}/_{kg}\right)} \right) \times 100$$

Descriptive statistics for the relative dose intensity with the number and percentage of patients belonging to relative dose intensity categories (eg, < 75%, ≥ 75 to < 90%,≥ 90%) will be provided.

Listings of study drug exposure will be provided.

5.3. Demographics and Baseline characteristics

Patient demographics (i.e., age, sex at birth, race, ethnicity) and baseline characteristics including, but not limited to, body weight [in kg], height [in cm], body mass index [BMI; in kg/m^2], Body Surface Area [BSA; in m^2] will be summarized with descriptive statistics by cohort and treatment. The analysis will be performed for All Enrolled Analysis Set.

5.4. Prior Anti-cancer Therapy

Prior anti-cancer therapy will be coded using the World Health Organization (WHO) Drug Dictionary.

A by-patient listing prior anti-cancer therapy will be provided.

5.5. Prior and On Study Radiation Therapy

A by-patient listing including information collected in eCRF and the flag for prior and on study radiation therapy will be provided.

5.6. Prior and On Study Surgeries and Procedures

A by-patient listing including information collected in eCRF and the flag for prior and on study surgery and procedure will be provided.

5.7. Medical History

General medical history data will be collected at screening. Medical history will be coded using the current version of Medical Dictionary for Regulatory Activities (MedDRA).

A by-patient listing of medical history will be provided.

5.8. Prior and Concomitant Medications

Medications collected at screening and during the study will be coded using the current version of the World Health Organization (WHO) Drug dictionary.

All prior and concomitant medications (other than per-protocol study drugs) will be provided in a by-patient listing.

5.9. On Study Treatment and Post Study Treatment Anti-cancer Therapies

All on study treatment and post treatment anti-cancer therapies including flag for on-study treatment and post study treatment therapies will be provided in a by-patient listing.

In case when the date of initiation of new anti-cancer therapy other than the study treatment is incomplete and needs to be imputed, the following algorithm will be followed:

If the day is missing but the month and year are available, then the imputed day will be first day of the month or the day of last dose +1 if the month and year of new anti-cancer therapy and the month and year of last dose are the same.

If the day and month are missing but year is available, then the imputed day and month will be 01Jan or the date of last dose +1 if the year of new anti-cancer therapy and the year of last dose are the same.

6. EFFICACY ANALYSES

Generally, summary of efficacy endpoints in this section will be based on ITT Analyses Set for Phase 2 Randomized Cohort, unless otherwise specified. Listings will be provided for All Enrolled Analysis Set. Safety run-in cohort 1 subjects will only be listed for efficacy related endpoints.

6.1. Primary Efficacy Endpoints

6.1.1. Definition of the Primary Efficacy Endpoints

For Phase 2 Randomized Cohort, the primary efficacy endpoint is progression-free survival (PFS) by investigator assessment, defined as the time from the date of randomization until the earliest date of documented disease progression as determined by investigator assessment, as determined based on response evaluation criteria in solid tumors (RECIST), Version 1.1, or death from any cause.

6.1.2. Analysis of the Primary Efficacy Endpoint

For Phase 2 Randomized Cohort, PFS by investigator assessment will be analyzed using Kaplan-Meier (KM) methods. The KM estimate of the survival function will be computed, and the results will be presented using KM curves by treatment group. The median, Q1, Q3 will be provided along with the corresponding 95% CI calculated by the Brookmeyer and Crowley method with log-log transformation. A log-rank test will be used to compare treatment difference in PFS by investigator with 2-sided P-value provided.

In addition, the treatment effect will be estimated by Hazard Ratio (HR) along with corresponding 2-sided 95% CI using the Cox proportional hazards regression model.

The censoring rules are summarized in Table 6-1.

Table 6-1. Censoring rule for primary and sensitivity analysis of PFS

Situation	Primary Analysis
Alive and no disease assessment at baseline or post-baseline assessment*	Censored at date of randomization (or first dosing date for non-randomized cohort)
Alive and progression-free	Censored at date of last evaluable disease assessment
Documented Progressive Disease (PD)	Progressed at date of earliest sign of PD
Death before first PD	Progressed at date of death
Death within 2 disease assessments window	Progressed at date of death
PD or death after ≥ 2 consecutively missed or not evaluable (NE) disease assessments	Censored at date of last evaluable assessment prior to missed or NE assessments, or date of randomization if there is no adequate post-baseline assessment before PD or Death

Situation	Primary Analysis
No PD or death before or on initiation of subsequent anticancer therapy	Censored at date of last evaluable assessment on or prior to subsequent anticancer therapy

*Baseline tumor assessment should be performed at screening visit before the start of study drug, unless otherwise specified in the protocol.

Given the scheduled visit assessment scheme for randomized cohort is every 8 weeks (± 7 days) during the study, the definition of 2 missed visits are defined as 2 x (8 weeks + 7) days.

6.1.3. Sensitivity Analysis of the Primary Efficacy Endpoint

There is no sensitivity analysis planned.

6.2. Secondary Efficacy Endpoints

6.2.1. Definition of Secondary Efficacy Endpoints

Objective Response Rate (ORR):

ORR is defined as the proportion of patients who achieve complete response or partial response that is confirmed at least 4 weeks after initial documentation of response, as determined by investigator assessment per RECIST, Version 1.1.

Overall Survival (OS)

The OS is measured from the date of randomization to the date of death from any cause. Those who are not observed to die during the study will be censored at their last known alive date.

Duration of response (DOR)

DOR is defined as time from first documentation of complete response or partial response to the earliest date of documented disease progression as determined by investigator assessment, per RECIST Version 1.1, or death from any cause, whichever occurs first.

PRO endpoints measured by EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 items, which address 15 HRQoL domains: 5 functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), 1 global health status scale, 3 symptom scales (fatigue, nausea and vomiting, and pain) and 6 single items (dyspnea, insomnia, loss of appetite, constipation, diarrhea, and financial difficulties) (See details in Appendix 2). The recall period is 1 week (past week). It will take about 11 minutes to complete.

According to the EORTC QLQ-C30 Scoring Manual, scores for each scale (or domain) should be calculated if responses are given to at least 50% of the items in that particular scale; otherwise, it should be considered as missing.

For all scales, the raw score (RS) is defined as the mean of the non-missing component items (I_i): RS = (I_1 + I_2 +…+I_n)/n.

Functional scales: $S = (1 - \frac{RS-1}{range}) \times 100$

Symptomscales and Global health status/QoL: $S = (\frac{RS-1}{range}) \times 100$

All of the scales and single-item measures range in score from 0 to 100. A high scale score represents a higher response level. Thus a high score for a functional scale represents a high / healthy level of functioning, a high score for the global health status / QoL represents a high QoL, but a high score for a symptom scale / item represents a high level of symptomatology / problems.

PRO endpoints measured by EQ-5D-5L

The EQ-5D-5L is an instrument for use as a measure of health outcome {EuroQol Research Foundation 2017}. The EQ-5D-5L consists of 2 sections: the EQ-5D descriptive system and the EQ visual analogue scale (EQ-VAS).

The descriptive system comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The patient is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the 5 dimensions. This decision results in a 1-digit number that expresses the level selected for that dimension. The digits for the 5 dimensions can be combined into a 5-digit number that describes the patient's health state.

The EQ-VAS records the patient's self-rated health on a vertical VAS, where the end points are labeled "the best health you can imagine" and "the worst health you can imagine." The EQ-VAS can be used as a quantitative measure of health outcome that reflects the patient's own judgment.

PRO endpoints measured by FACT Colorectal Symptom Index

The FCSI is a well-validated PRO instrument as a set of brief, clinically relevant, colorectal cancer symptoms for assessing symptomatic response. It comprises the most important symptoms associated with colorectal cancer, including energy, pain, weight, diarrhea, nausea, swelling or cramps in the stomach area, appetite, ability to enjoy life, and overall quality of life. Its minimum important difference (MID) range from 1.5 to 3 has been published {Colwell 2010}.

The FCSI questionnaire is provided in Appendix 4.

6.2.2. Analysis Methods for Secondary Efficacy Endpoints

Objective Response Rate (ORR)

Confirmed ORR by investigator assessment along with the 95% CI will be estimated based on the Clopper-Pearson method for Phase 2 Randomized Cohort. The chi-square test may be used to compare treatment difference in ORR. Odds ratios and corresponding 95% CIs will also be presented.

Overall Survival (OS)

The analysis of overall survival (OS) will be performed using the Kaplan-Meier method for Phase 2 Randomized Cohort in the ITT analysis set. Median, Q1, and Q3 of the DOR will be derived based on KM estimates along with the corresponding 95% CI by the Brookmeyer and Crowley method with log-log transformation. Kaplan-Meier curves will also be provided.

Duration of response (DOR)

The analysis of duration of response (DOR) will be performed using the Kaplan-Meier method for Phase 2 Randomized Cohort in the ITT analysis set. Median, Q1, and Q3 of the DOR will be derived based on KM estimates along with the corresponding 95% CI by the Brookmeyer and Crowley method with log-log transformation. Kaplan-Meier curves will also be provided.

Change from baseline in PRO endpoints

The descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores for each domain of EORTC QLQ-C30 at each post-baseline visit by treatment group in the ITT Analysis Set for Phase 2 Randomized Cohort and mITT Analysis Set for safety run-in cohort.

The descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores for each dimension of EQ-5D-5L and EQ-VAS at each post-baseline visit by treatment group in the ITT Analysis Set for Phase 2 Randomized Cohort and mITT Analysis Set for safety run-in cohort.

The descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores for FCSI at each post-baseline visit by treatment group in the ITT Analysis Set for Phase 2 Randomized Cohort and mITT Analysis Set for safety run-in cohort.

CCI [redacted]

6.4. Changes From Protocol-Specified Efficacy Analyses

Due to the early termination of the program, efficacy endpoints will not be evaluated using stratification factors as covariates or stratification variables for stratified analysis model. For PRO endpoints, only descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores.

7. SAFETY ANALYSES

Safety analysis will be performed in the Safety Analysis Set. Listing will be based on the All Enrolled analysis set unless specified otherwise. Analysis of DLT was performed using the DLT evaluable analysis set for Safety Run-in evaluations.

7.1. Adverse Events and Deaths

7.1.1. Adverse Event Dictionary

Clinical and laboratory adverse events (AEs) will be coded using the current version of Medical Dictionary for Regulatory Activities (MedDRA).

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to CTCAE Version 5.0.

7.1.3. Relationship of Adverse Events to Study Treatment

Related AEs are those for which the investigator selected "Related" on the AE CRF to the question of "Related to Study Treatment" for each study drug and overall. Relatedness will always default to the investigator's choice, not that of the medical monitor. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-patient data listings will show the relationship as missing.

7.1.4. Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if the AEs met the definitions of SAEs that were specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Drug Safety and Public Health Department before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any AEs with an onset date on or after the date of the first dose of study treatment up to 30 days after the date of the last dose of study treatment, or the day before initiation of subsequent anti-cancer therapy, whichever comes first.

If the AE onset date is on or before the last dose date, the AE is considered as TEAE, regardless of the initiation of subsequent anti-cancer therapy.

7.1.5.2. Missing or incomplete Dates

If there was a missing or incomplete date for the start date or stop date of an AE, the most conservative approach was used for analysis.

If the onset date of the AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset is the same as or after the month and year (or year alone if month is not recorded) of the first dosing date of study drug, and
- The AE onset date is the same as or before the month and year (or year alone if month is not recorded) of the date corresponding to the cutoff date of TEAE period, which is defined as the 30 days after the study drug last dose date or the day before the initiation of new anticancer therapy (whichever is earlier)

An AE with completely missing onset and stop dates, or with the onset date missing and the stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent.

In case when the AE onset date is incomplete and needs to be imputed, the following algorithm will be followed:

- If the day is missing but the month and year are available, then the imputed day will be the first dosing date if they have the same month and year, or the first day of the month otherwise.
- If the day and month are missing but year is available, then the imputed day and month will be the first dosing date if they have the same year, or 01Jan otherwise.

7.1.6. Summaries of Adverse Events and Deaths

7.1.6.1. Adverse Events and Deaths

A brief, high-level summary of the number of percentage of patients who experienced at least 1 TEAE in the categories described below will be provided for each cohort and treatment.

The number and percentage of patients who experienced at least 1 TEAE will be provided and summarized by SOC, PT for each cohort and treatment group.

For the AE categories described below, summaries will be provided by SOC, PT:

- TEAEs
- TEAEs with Grade 3 or 4
- TE treatment-related AEs for Magrolimab and for study drug other than Magrolimab
- TE treatment-related AEs with Grade 3 or 4 for Magrolimab and for study drug other than Magrolimab
- TE SAEs
- TE treatment-related SAEs for Magrolimab and for study drug other than Magrolimab
- TEAEs leading to dose interruption of Magrolimab and study drug other than Magrolimab
- TEAEs leading to dose reduction of Magrolimab and study drug other than Magrolimab
- TEAEs leading to discontinuation of Magrolimab and study drug other than Magrolimab
- TEAEs leading to death

Multiple events will be counted only once per subject in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and then by PT within each SOC in the descending order of overall frequency. For summaries by severity, the most severe severity will be used for those AEs that occurred more than once in a given subject during the study.

In addition to the above summary tables, all TEAEs, TEAEs of Grade 3 or 4, TE SAEs, TE treatment-related SAEs for Magrolimab and TEAEs leading to death will be summarized by PT in descending order of overall frequencies.

In addition, data listings will be provided for the following:

- All AEs, indicating whether the event is treatment emergent
- All SAEs
- All Deaths
- All SAEs leading to death
- All AEs with severity of Grade 3 or 4
- All AEs leading to discontinuation of Magrolimab
- All AEs leading to dose delay or interruption of Magrolimab

A summary (number and percentage of patients) of deaths will be provided by cohort and treatment. Summary will include the following categories:

- All deaths
- Deaths within 30 days of the last dosing of study treatment
- Deaths beyond 30 days of the last dosing of study treatment

7.1.7. Additional Analysis of Adverse Events

7.1.7.1. Dose Limiting Toxicity (DLT)

In Safety run-in Cohort, a summary of DLT by PT among the DLT evaluable patients will be provided. DLT-type AEs will also be summarized in the same way, where DLT-type AEs are toxicities that met protocol specified criteria of DLTs and occurred beyond the protocol-specified DLT assessment period (onset date beyond the 28 days period).

By-patient listing of the DLT AEs and DLT-type AEs will also be provided.

7.1.7.2. Treatment-Emergent Adverse Events (TEAE) of Clinical Importance

Number and percentage of participants with the following AEs of Clinical Importance (AECI) will be summarized by PT.

Treatment Emergent Adverse Events of Clinical Importance

TEAE of Clinical Importance	Search Strategy
Anaemia	Medical Search Term (MST) Anemia Extravascular transient hemolysis
Infusion related reactions	Standardized MedDRA Queries (SMQ) Hypersensitivity (narrow) + within one day of latest infusion of any study drug
Severe Neutropenia	PT: Grade 3+ Febrile neutropenia, Grade 3+ Neutrophils count decreased, Grade 3+ Neutropenia
Serious Infections	SAE within SOC: Infections and infestations
Transfusion reactions due to magrolimab interference with RBC typing	MST Transfusion reactions due to magrolimab interference with RBC typing
Thromboembolic events	SMQ Embolic and thrombotic events (broad)
Pneumonitis	SMQ Interstitial lung disease (broad)
Myocardial Infarction	SMQ Myocardial infarction (broad)

7.2. Laboratory Evaluations

Laboratory data collected during the study will be analyzed and summarized using both quantitative and qualitative methods. Summaries of laboratory data will be provided for the Safety Analysis Set and will include data collected during the treatment-emergent period. The analysis will be based on values reported in conventional units.

7.2.1. Summaries of Numeric Laboratory Results

Descriptive statistics will be provided by cohort and treatment group for selected laboratory test specified in the study protocol as follows:

- Baseline values
- Postbaseline maximum value
- Change and percentage change from baseline to postbaseline maximum value
- Postbaseline minimum value
- Change and percentage change from baseline to postbaseline minimum value
- Values at each postbaseline time point
- Change and percentage change from baseline at each postbaseline time point

Median (Q1, Q3) of change from baseline will be plotted for the lab parameters including (but not limited to) hemoglobin, platelet, and absolute neutrophil counts over time.

7.2.2. Graded Laboratory Values

The CTCAE Version 5.0 will be used to assign toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (i.e., increased, decreased) will be presented separately.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase at least 1 toxicity grade from baseline at any postbaseline time point, up to and including the date of last dose of study drug plus 30 days or the day before initiation of new anticancer therapy, whichever is earlier. If the relevant postbaseline lab is assessed on or before the last dose date, the laboratory abnormality is considered as treatment-emergent laboratory abnormality, regardless of the initiation of subsequent anti-cancer therapy. If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent.

7.2.2.2. Summaries of Laboratory Abnormalities

The following summaries (number and percentage of patients) for treatment-emergent laboratory abnormalities will be provided by lab test for each cohort and treatment group; patients will be categorized according to the most severe postbaseline abnormality grade for a given lab test:

- Laboratory abnormalities (Grade 1 to 4 separately)
- Grade 3 or 4 laboratory abnormalities

For all summaries of laboratory abnormalities, the denominator is the number of patients in the the number of patients with nonmissing postbaseline values up to 30 days after the last dosing date or the day before initiation of new anticancer therapy, whichever is earlier.

By-patients listings of treatment-emergent Grade 3 or 4 laboratory abnormalities will be provided by subject ID number and visit in chronological order. This listing will include all test results that were collected throughout the study for the lab test of interest, with all applicable severity grades displayed.

7.2.3. Liver-related Laboratory Evaluations

Liver-related abnormalities after the first dose of any study drug will be examined and summarized for each cohort and treatment group using the number and percentage of patients who were reported to have the following laboratory test values for postbaseline measurements:

- Aspartate aminotransferase (AST): > 3 times of the upper limit of reference range (ULN)
- Alanine aminotransferase (ALT): > 3 x ULN
- AST or ALT: > 3 x ULN
- Total bilirubin: > 2 x ULN
- AST or ALT > 3 x ULN and total bilirubin > 2 x ULN
- AST or ALT > 3 x ULN and total bilirubin > 2 x ULN and alkaline phosphatase (ALP) < 2 x ULN

The summary will include data from all postbaseline visits up to 30 days after the last dose of study drug or the day before initiation of new anticancer therapy, whichever is earlier. For individual laboratory tests, patients will be counted once based on the most severe postbaseline values. For the composite endpoints of AST or ALT and total bilirubin and ALP, patients will be counted once when the criteria are met at the same postbaseline visit date. The denominator is the number of patients in the Safety Analysis Set who have at least one postbaseline visit, at which all relevant tests are nonmissing at the same postbaseline visit.

A listing of patients who met at least 1 of the above criteria will be provided.

7.2.4. Shifts Relative to the Baseline Value

Shift tables will be presented by showing change in severity grade from baseline to the worst grade postbaseline for hematology and chemistry tests.

7.3. Pregnancy Test

A by-patient listing of the pregnancy test results as collected in CRF will be provided.

7.4. Changes From Protocol-Specified Safety Analyses

Body weight, vital sign and physical examination will not be summarized.

8. CHANGES FROM PROTOCOL-SPECIFIED ANALYSES

Due to the early termination of the program, efficacy endpoints will not be evaluated using stratification factors as covariates or stratification variables for stratified analysis model. Only descriptive statistics will be summarized for PRO endpoints. Body weight, vital sign and physical examination will not be summarized.

9. PHARMACOKINETIC (PK) AND IMMUNOGENICITY ANALYSES

9.1. PK Analyses

The PK Analysis Set will be used for summaries of PK concentration of magrolimab versus time. Serum concentrations will be listed and summarized for magrolimab using descriptive statistics by sampling time point and treatment. Box plots of serum concentration versus time will be generated.

9.2. Immunogenicity analysis

Serum samples for antidrug antibody (ADA) assessments will be conducted utilizing a tiered approach (screen, confirmatory and titer), and ADA data will be collected at scheduled visits shown in the protocol. ADA result from each sample will be reported as either positive or negative. If the sample is positive, the ADA titer will be reported as well. In addition, the presence of neutralizing antibody (nAb) will be tested for all ADA positive samples using a ligand-binding assay. The nAb results will be reported as positive or negative.

Patients for ADA prevalence are defined as patients who had at least one positive ADA sample at any time, including baseline and post-baseline, in Immunogenicity Analysis Set.

Patients for ADA incidence are defined as evaluable patients with at least one non-missing baseline ADA sample and at least one post-treatment ADA sample in Immunogenicity Analysis Set that showed treatment-induced ADA or treatment-boosted ADA.

Patients for nAb prevalence are defined as patients who had at least one positive nAB sample at any time, including baseline and post-baseline, in Immunogenicity Analysis Set

Among patients for ADA incidence, patients who had Treatment-Induced ADA or Treatment-Boosted ADA are categorized as ADA positive by incidence, where:

- Treatment-Induced ADA is defined as patients who had negative baseline ADA sample and at least one positive post-treatment ADA sample.
- Treatment-Boosted ADA is defined as patients who had positive baseline ADA sample and at least one positive post-treatment ADA sample and the (max titer of the post-treatment ADA) / (titer of baseline ADA) ≥ 4.

The remaining patients among patient for ADA incidence are ADA negative by incidence.

Transient ADA is defined as:

- Treatment-Induced ADA detected only at one sampling time point during the treatment or follow-up observation period (excluding the last sampling time point, which ought to be considered persistent unless shown to be undetectable at a later time)

  or

- Treatment-Induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive sample (irrespective of any negative samples in between) are separated by a period less than 16 weeks, and the patient's last sampling time point is ADA-negative.

ADA Transience Rate: the proportion of patients who had transient among patients for ADA incidence.

Persistent ADA is defined as:

- Treatment-Induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive sample (irrespective of any negative samples in between) are separated by a period of 16 weeks or longer.

  or

- Treatment-Induced ADA incidence (i.e. Treatment-induced ADA) only in the last sampling time point of the treatment study period or at a sampling time point with less than 16 weeks before an ADA-negative last sample.

ADA Persistence Rate: the proportion of patients who had persistent ADA among patients for ADA incidence.

The following ADA categories will be summarized:

- ADA prevalence
- ADA positive post-baseline and positive at baseline
- ADA not detected post-baseline and positive at baseline
- ADA incidence
- Treatment-induced ADA
- Treatment-boosted ADA
- Persistent ADA
- Transient ADA
- nAb prevalence

ADA titer and nAb data will be listed for samples confirmed positive for the presence of ADA to study drug. The effect of ADA on PK, safety and efficacy may be examined by descriptive summaries if data allow.

10. BIOMARKER ANALYSIS

CCI [redacted]

11. REFERENCES

Colwell H, Mathias S, Turner M, Lu J, Wright N, Peeters M, et al. Psychometric Evaluation of the FACT Colorectal Cancer Symptom Index (FCSI-9): Reliability, Validity, Responsiveness, and Clinical Meaningfulness. The Oncologist 2010.

EuroQol Research Foundation. EQ-5D-5L: About. Available at: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/. Accessed: 10 August 2020. Last Updated: 18 April. 2017:

Ivanova A, Qaqish BF, Schell MJ. Continuous toxicity monitoring in phase II trials in oncology. Biometrics 2005;61 (2):540-5.

12. SOFTWARE

SAS® Software Version 9.4. SAS Institute Inc., Cary, NC, USA.

EAST Version 6.5, Cytel Inc., MA, USA.

13. APPENDICES

Appendix 1. RECIST 1.1-based Assessments - Overall visit response

The RECIST tumour response data will be used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients will be assigned a RECIST 1.1 visit response of CR, PR, SD or PD, using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a patient has had a tumour assessment which cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE). For patients with no disease at baseline (i.e. no TLs and no NTLs), evaluation of overall visit responses will be based on absence/presence of new lesions. If no TLs and no NTLs are recorded at a visit, both the TL and NTL visit response will be recorded as NA and the overall visit response will be no evidence of disease (NED). If a new lesion is observed then the overall visit response will be PD. Appendix Table 1 summarizes overall visit response given the visit responses from TL and NTL are combined with new lesion.

Appendix Table 1. Overall Visit Response

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

CR Complete response, NA Not applicable, PR Partial response, SD Stable disease, PD Progression of disease, NE Not evaluable, NED No evidence of disease.

Appendix 2. Overview of EORTC QLQ-C30 and Questionnaire

The EORTC QLQ-C30 is a well-validated questionnaire commonly used in oncology trials. It consists of 30 items, which address 15 HRQoL domains: five multi-item functional scales, three multi-item symptom scales, a global health status/QoL scale, and six single-item symptom scales (Appendix Table 2).

Appendix Table 2. Overview of EORTC QLQ-C30 (Version 3)

EORTC QLQ-C30 Domains	Number of Items	Item Range	Item Numbers (Version 3)
Global health status/QoL	2	1–7	29, 30
Functional Domains			
Physical functioning	5	1–4	1–5
Role functioning	2	1–4	6, 7
Emotional functioning	4	1–4	21–24
Cognitive functioning	2	1–4	20, 25
Social functioning	2	1–4	26, 27
Symptom and Financial Difficulty Domains			
Fatigue	3	1–4	10, 12, 18
Nausea and vomiting	2	1–4	14, 15
Pain	2	1–4	9, 19
Dyspnea	1	1–4	8
Insomnia	1	1–4	11
Appetite loss	1	1–4	13
Constipation	1	1–4	16
Diarrhea	1	1–4	17
Financial difficulties	1	1–4	28

Abbreviations: EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Core 30 Questionnaire; QoL = quality of life

ENGLISH

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in your initials:
Your birthdate (Day, Month, Year):
Today's date (Day, Month, Year): 31

	Not at All	A Little	Quite a Bit	Very Much
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2. Do you have any trouble taking a long walk?	1	2	3	4
3. Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4. Do you need to stay in bed or a chair during the day?	1	2	3	4
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:

	Not at All	A Little	Quite a Bit	Very Much
6. Were you limited in doing either your work or other daily activities?	1	2	3	4
7. Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8. Were you short of breath?	1	2	3	4
9. Have you had pain?	1	2	3	4
10. Did you need to rest?	1	2	3	4
11. Have you had trouble sleeping?	1	2	3	4
12. Have you felt weak?	1	2	3	4
13. Have you lacked appetite?	1	2	3	4
14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4
16. Have you been constipated?	1	2	3	4

Please go on to the next page

ENGLISH

During the past week:	Not at All	A Little	Quite a Bit	Very Much
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

Appendix 3. 5-Level EuroQol 5 Dimensions Questionnaire (EQ-5D-5L)

Health Questionnaire

English version for the UK

Under each heading, please tick the ONE box that best describes your health TODAY.

MOBILITY

I have no problems in walking about ❑
I have slight problems in walking about ❑
I have moderate problems in walking about ❑
I have severe problems in walking about ❑
I am unable to walk about ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN / DISCOMFORT

I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

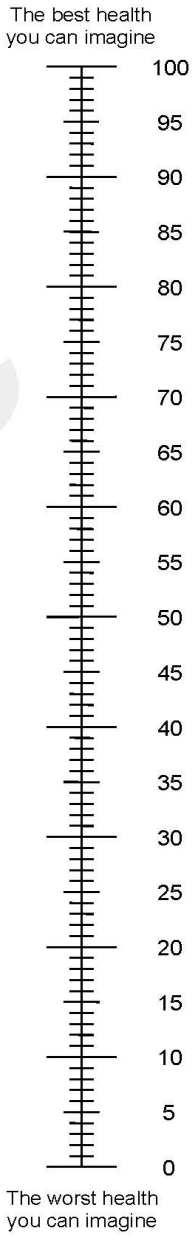
ANXIETY / DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

2

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



Appendix 4. FACT Colorectal Symptom Index (FCSI)

Below is a list of statements that other people with your illness have said are important.
Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

		Not at all	A little bit	Some-what	Quite a bit	Very much
GP1	I have a lack of energy	0	1	2	3	4
GP4	I have pain	0	1	2	3	4
C2	I am losing weight	0	1	2	3	4
C5	I have diarrhea (diarrhoea)	0	1	2	3	4
GP2	I have nausea	0	1	2	3	4
C1	I have swelling or cramps in my stomach area	0	1	2	3	4
C6	I have a good appetite	0	1	2	3	4
GF3	I am able to enjoy life	0	1	2	3	4
GF7	I am content with the quality of my life right now	0	1	2	3	4

GS-US-587-6156 SAP

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Biostatistics eSigned	16-Jul-2024 23:57:58
PPD	Global Development Lead (GDL) eSigned	17-Jul-2024 05:38:44