STATISTICAL ANALYSIS PLAN

Study Title:	A Phase 2 Study of Magrolimab Combination Therapy in Patients With Head and Neck Squamous Cell Carcinoma
Name of Test Drug:	Magrolimab
Study Number:	GS-US-548-5916
Protocol Version (Date):	Amendment 6 (01 November 2023)
Analysis Plan Version:	2.0
Analysis Plan Date:	21 October 2024
Analysis Plan Author(s):	PPD
	PPD

CONFIDENTIAL AND PROPRIETARY INFORMATION

TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN 1
TABLE OF CONTENTS 2
LIST OF IN-TEXT TABLES 4
LIST OF IN-TEXT FIGURES 4
LIST OF ABBREVIATIONS 5
1. INTRODUCTION 7
1.1. Study Objectives and Endpoints 7
1.2. Study Design 8
1.2.1. Safety Run-in 1 10
1.2.2. Safety Run-in 2 11
1.2.3. Pre-expansion Safety Run-in for Magrolimab + Pembrolizumab 11
1.2.4. DLT Assessment Period 11
1.2.5. Safety Review Team (SRT) 11
1.2.6. Phase 2 Cohort 1 12
1.2.7. Phase 2 Cohort 2 12
1.2.8. Phase 2 Cohort 3 12
1.2.9. Diagnostic Tissue Testing 12
1.3. Sample Size and Power 13
2. TYPE OF ANALYSIS 14
2.1. Interim Analyses 14
2.1.1. Dose Determination Analysis 14
2.1.2. Informal Interim Analysis 14
2.2. Primary Analysis 14
2.3. Final Analysis 14
3. GENERAL CONSIDERATIONS FOR DATA ANALYSES 15
3.1. Analysis Sets 15
3.1.1. All Enrolled Analysis Set 15
3.1.2. Intent-to-Treat Analysis Set 15
3.1.3. Modified Intent-to-Treat Analysis Set 15
3.1.4. Safety Analysis Set 15
3.1.5. Dose-Limiting Toxicity (DLT)-Evaluable Analysis Set 15
3.1.6. Pharmacokinetic Analysis Set 16
3.1.7. Immunogenicity Analysis Set 16
3.2. Subject Grouping 16
3.2.1. Subject Grouping for Each Analysis Set 16
3.2.2. Subject Grouping for Each Type of Analysis 17
3.3. Strata and Covariates 18
3.4. Examination of Subject Subgroups 18
3.5. Adjustment for Multiplicity 18
3.6. Missing Data and Outliers 19
3.6.1. Missing Data 19
3.6.2. Outliers 19
3.7. Data Handling Conventions and Transformations 19
3.8. Analysis Visit Windows 21
3.8.1. Definition of Study Day 21
3.8.2. Analysis Visit Windows 21

3.8.3. Selection of Data in the Event of Multiple Records in an Analysis Visit Window 22
3.9. Assessment of COVID-19 Impact 22
4. PROTOCOL DEVIATIONS 24
5. SUBJECT INFORMATION 25
5.1. Patient Enrollment and Disposition 25
5.2. Extent of Study Treatment Exposure 26
5.2.1. Duration of Exposure to Study Drug 26
5.2.2. Relative Dose Intensity 26
5.3. Demographics and Baseline characteristics 26
5.4. Prior Anti-cancer Therapy 27
5.5. Prior and On Study Radiotherapy 27
5.6. Prior and On Study Surgeries and Procedures 27
5.7. Medical History 27
5.8. Prior and Concomitant Medications 27
5.8.1. Prior Medications 27
5.8.2. Concomitant Medications 28
5.9. On Study Treatment and Post Study Treatment Anti-cancer Therapies 28
6. EFFICACY ANALYSES 29
6.1. Primary Efficacy Endpoints 29
6.1.1. Definition of the Primary Efficacy Endpoints 29
6.1.2. Analysis of the Primary Efficacy Endpoint 29
6.1.3. Sensitivity Analysis of the Primary Efficacy Endpoint 30
6.2. Secondary Efficacy Endpoints 31
6.2.1. Definition of Secondary Efficacy Endpoints 31
6.2.2. Analysis Methods for Secondary Efficacy Endpoints 33
6.3. Other Efficacy Definition and Analyses 33
6.3.1. Other Definition Related to Efficacy 33
6.3.2. Other Analyses Methods Related to Efficacy 34
6.4. Changes From Protocol-Specified Efficacy Analyses 34
7. SAFETY ANALYSES 35
7.1. Adverse Events and Deaths 35
7.1.1. Adverse Event Dictionary 35
7.1.2. Adverse Event Severity 35
7.1.3. Relationship of Adverse Events to Study Treatment 35
7.1.4. Serious Adverse Events 35
7.1.5. Treatment-Emergent Adverse Events 35
7.1.6. Summaries of Adverse Events and Deaths 36
7.1.7. Additional Analysis of Adverse Events 38
7.2. Laboratory Evaluations 38
7.2.1. Summaries of Numeric Laboratory Results 38
7.2.2. Graded Laboratory Values 39
7.2.3. Liver-related Laboratory Evaluations 40
7.3. Body Weight, Height, and Vital Signs 41
7.4. Electrocardiogram Results 41
7.5. Other Safety Measures 41
7.6. Changes From Protocol-Specified Safety Analyses 41
8. CHANGES FROM PROTOCOL-SPECIFIED ANALYSES 42
9. PHARMACOKINETIC (PK) AND IMMUNOGENICITY ANALYSES 43

9.1. PK and Anti-drug-antibody (ADA) Sample Collection 43
9.2. PK Analyses 43
9.3. Immunogenicity analysis 43
9.3.1. Definition of Terminology 43
9.3.2. Evaluation of Immunogenicity Data 44
10. BIOMARKER ANALYSIS 45
11. REFERENCES 46
12. SOFTWARE 47
13. SAP REVISION 48
14. APPENDICES 49

Appendix 1. RECIST 1.1-based Assessments - Overall visit response 50
Appendix 2. Overview of EORTC QLQ-C30 and Questionnaire 51

LIST OF IN-TEXT TABLES

Table 3-1. Analysis Visit Windows for Lab By-visit Summaries 21
Table 3-2. Analysis Visit Windows for By-visit Summaries of PRO assessments 22
Table 6-1. Censoring rule for primary and sensitivity analysis of PFS 30
Table 7-1. Treatment Emergent Adverse Events of Clinical Importance 38

LIST OF IN-TEXT FIGURES

Figure 1-1. Study Schema 10

LIST OF ABBREVIATIONS

5-FU	5-fluorouracil
ADA	antidrug antibody
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BICR	blinded independent central review
BMI	body mass index
BOR	best overall response
CI	confidence interval
CPS	combined positive score
CR	complete response
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DLT	dose-limiting toxicity
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
FAS	Full Analysis Set
HNSCC	Head and Neck Squamous Cell Carcinoma
HPV	Human papillomavirus
HR	hazard ratio
HRQoL	Health-related Quality of Life
ITT	intent to treat
IRT	interactive response technology
KM	Kaplan-Meier
LOQ	limit of quantitation
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MST	MedDRA Search Term
NA	Not applicable
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
ORR	objective response rate

OS	overall survival
PD	progression of disease
PD-L1	programmed cell death ligand 1
PFS	progression-free survival
PK	pharmacokinetics
PR	partial response
PRO	Patient reported outcome
PS	Patient Safety
PT	preferred term
Q1, Q3	first quartile, third quartile
RECIST	response evaluation criteria in solid tumors
RP2D	recommended Phase 2 dose
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SE	standard error
SMQ	Standardised MedDRA Queries
StD	standard deviation
SOC	system organ class
SRT	safety review team
TE	treatment-emergent
TEAE	treatment-emergent adverse event
TTR	time to response
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 informal interim, primary and final analyses of Study GS-US-548-5916. This SAP is based on the study protocol amendment 6 dated 01 November 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: • To evaluate the safety, tolerability, and recommended Phase 2 dose (RP2D) of magrolimab in combination with the following: — Pembrolizumab + platinum + 5-fluorouracil (5-FU) — Docetaxel Phase 2 Cohorts: • To evaluate the progression-free survival (PFS) with magrolimab in combination with pembrolizumab + platinum + 5-FU versus pembrolizumab + platinum + 5-FU as determined by investigator assessment (Phase 2 Cohort 1) • To evaluate the efficacy of magrolimab in combination with pembrolizumab, if this optional cohort is opened, and magrolimab in combination with docetaxel as determined by the investigator-assessed objective response rate (ORR) (Phase 2 Cohorts 2 and 3)	Safety Run-in: • Incidence of adverse events (AEs) and laboratory abnormalities defined as dose-limiting toxicities (DLTs) according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0 Phase 2 Cohorts: • PFS, defined as the time from the date of randomization until the earliest date of documented disease progression, as determined by investigator assessment, or death from any cause, whichever occurs first (Phase 2 Cohort 1, Arm A versus Arm B) • ORR, defined as the proportion of patients who achieve a complete response (CR) or partial response (PR) as measured by response evaluation criteria in solid tumors (RECIST), Version 1.1, as determined by investigator assessment (Phase 2 Cohorts 2 and 3)
Secondary Objectives	**Secondary Endpoints**
Safety Run-in: • To evaluate the pharmacokinetics (PK) and immunogenicity of magrolimab in combination with anticancer therapies Phase 2 Cohorts: • To evaluate PFS for magrolimab + zimberelimab + platinum + 5-FU versus pembrolizumab + platinum + 5-FU as determined by investigator assessment (Phase 2 Cohorts 1) • To evaluate ORR as determined by investigator assessment (Phase 2 Cohort 1)	Safety Run-in: • Magrolimab concentration versus time and antidrug antibodies (ADAs) to magrolimab Phase 2 Cohorts: • PFS, as determined by investigator assessment or death from any cause, whichever occurs first (Phase 2 Cohort 1, Arm C versus concurrent Arm B) • ORR, as determined by investigator assessment (Phase 2 Cohort 1)

• To evaluate PFS by investigator assessment (Phase 2 Cohorts 2 and 3) • To evaluate additional measures of efficacy, including duration of response (DOR) and overall survival (OS) • To evaluate the PK and immunogenicity of magrolimab in combination with anticancer therapies • To evaluate patient-reported outcomes (PROs)/quality-of-life measures	• PFS from date of dose initiation (Phase 2 Cohort 2 and Phase 2 Cohort 3) until 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 • DOR, defined as time from first documentation of CR or PR to the earliest date of documented disease progression or death from any cause, whichever occurs first • OS, defined as time from date of randomization (Phase 2 Cohort 1) or date of dose initiation (Phase 2 Cohort 2 and Phase 2 Cohort 3) to death from any cause • Magrolimab concentration versus time and ADAs to magrolimab • PRO assessment (European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core - Questionnaire [EORTC QLQ-C30], European Organisation for Research and Treatment of Cancer Quality of Life - Head and Neck Module [EORTC QLQH& N35], and 5 level – EuroQol - 5 dimensions questionnaire [EQ-5D- 5L]) scores
CCI	

1.2. Study Design

This is a Phase 2, open-label, multicenter study evaluating magrolimab in combination with pembrolizumab + platinum + 5-FU chemotherapy; magrolimab in combination with pembrolizumab; magrolimab in combination with zimberelimab + platinum + 5-FU chemotherapy in patients with untreated metastatic or unresectable, locally recurrent HNSCC; and magrolimab in combination with docetaxel in patients with locally advanced/metastatic HNSCC (mHNSCC) who were previously treated with at least 1 and no more than 2 lines of prior systemic therapy.

This study will consist of the following 2 safety run-in evaluations:

- Safety Run-in 1: magrolimab + pembrolizumab + platinum + 5-FU in patients with untreated metastatic or unresectable, locally recurrent HNSCC regardless of programmed cell death ligand 1 (PD-L1) status

- Safety Run-in 2: magrolimab + docetaxel in patients with locally advanced/mHNSCC regardless of PD-L1 status who were previously treated with at least 1 and no more than 2 lines of prior systemic therapy

Additionally, a pre-expansion safety run-in evaluation of magrolimab + pembrolizumab in patients with untreated metastatic or unresectable, locally recurrent HNSCC with a PD-L1 combined positive score (CPS) ≥ 1 may be conducted at the sponsor's discretion prior to the initiation of the optional Phase 2 cohort of magrolimab + pembrolizumab.

Once the safety review team (SRT) reviews the safety for patients in each run-in and the sponsor determines the RP2D for that cohort, the following Phase 2 cohorts will be conducted:

- Phase 2 Cohort 1: a randomized, open-label cohort of magrolimab + pembrolizumab + platinum + 5-FU (Arm A) versus pembrolizumab + platinum + 5-FU (Arm B) versus a delayed open arm magrolimab + zimberelimab + platinum + 5-FU (Arm C) in patients with untreated metastatic or unresectable, locally recurrent HNSCC regardless of PD-L1 status.

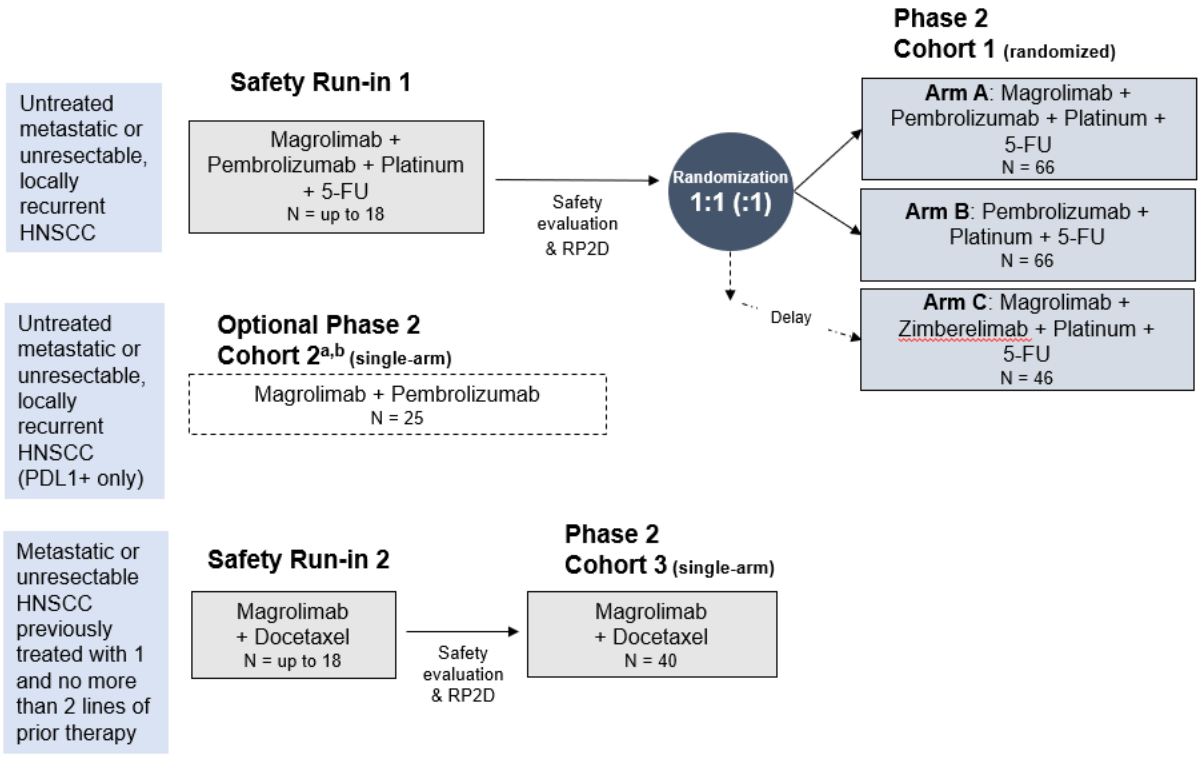
- Optional Phase 2 Cohort 2: a cohort of magrolimab + pembrolizumab in patients with untreated metastatic or unresectable, locally recurrent HNSCC with a PD-L1 CPS ≥ 1

- Phase 2 Cohort 3: a cohort of magrolimab + docetaxel in patients with locally advanced/mHNSCC regardless of PD-L1 status who were previously treated with at least 1 and no more than 2 lines of prior systemic therapy

Approximately 200 and up to 279 patients may be enrolled in the study, with at least 12 and up to 36 patients in total in Safety Run-ins 1 and 2, and optional Pre-expansion Safety Run-in for Cohort 2; approximately 178 patients in Phase 2 Cohort 1; approximately 25 patients in optional Phase 2 Cohort 2; and approximately 40 patients in Phase 2 Cohort 3. Cycle length is 21 days for all cohorts. All patients will continue on study treatment unless they meet study treatment discontinuation criteria.

The study schema is provided in Figure 1-1.

Figure 1-1. Study Schema



5-FU = 5-fluorouracil; HNSCC = head and neck squamous cell carcinoma; PD-L1 = programmed cell death ligand 1; RP2D = recommended Phase 2 dose

a Optional cohort to be opened at the discretion of the sponsor.

b A pre-expansion safety run-in evaluation of magrolimab + pembrolizumab may be conducted at the sponsor's discretion prior to the initiation of this optional cohort.

1.2.1. Safety Run-in 1

A dose-limiting toxicity (DLT) evaluation period of 1 cycle (21 days) will occur. After 6 evaluable patients have completed the DLT evaluation period, a decision will be made on further expansion or dose de-escalation.

Even though 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. Dose de-escalation decisions will be made as follows:

- If no more than 2 of 6 DLT-evaluable patients experience a DLT in Cycle 1, enrollment into Phase 2 Cohorts 1 and 2 will begin at this dose level.

- If 3 or more (> 34%) DLT-evaluable patients experience a DLT at any time, another 6 patients will be enrolled at a lower dose and will be evaluated in the same manner to define the recommended dose for the combination regimen.

1.2.2. Safety Run-in 2

Safety Run-in 2 will open for enrollment at the same time as Safety Run-in 1. A dose-limiting toxicity (DLT) evaluation period of 1 cycle (21 days) will occur. After 6 evaluable patients have completed the DLT evaluation period, a decision will be made on further expansion or dose de-escalation. The same design and DLT rules will apply to Safety Run-in 2.

1.2.3. Pre-expansion Safety Run-in for Magrolimab + Pembrolizumab

At the discretion of the sponsor, 6 patients may be initially enrolled to receive magrolimab + pembrolizumab in this pre-expansion run-in evaluation. Dose evaluation and possible de-escalation will be performed as described for Safety Run-in 1.

1.2.4. DLT Assessment Period

The DLT assessment period will be the first cycle (21 days) and applies to each safety run-in. Patients will be considered evaluable for assessment of DLTs if either of the following criteria is met during the DLT assessment period:

- The patient experiences a DLT at any time after initiation of the first infusion of magrolimab.
- The patient does not experience a DLT and completes at least 2 infusions of magrolimab and at least 1 dose of pembrolizumab, 1 dose of platinum, and 1 dose of 5-FU for Safety Run-in 1; at least 1 dose of docetaxel for Safety Run-in 2; and at least 1 dose of pembrolizumab for the pre-expansion safety run-in for magrolimab + pembrolizumab (if applicable).

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 evaluations will be replaced.

The DLT definition is provided in the study protocol.

Patients enrolled in the safety run-in evaluations will continue treatment until unacceptable toxicity or disease progression, whichever occurs first, and will not change their magrolimab dose level after the RP2D is determined.

1.2.5. Safety Review Team (SRT)

An SRT will be established to assess safety of the patients in various cohorts. The SRT will include at least 1 investigator, the Gilead Sciences (Gilead) medical monitor, and the Gilead Patient Safety (PS) physician. Others may be invited to participate as members of the SRT if additional expertise is desired (i.e., representatives from Clinical Operations, Biostatistics, Clinical Pharmacology, and Biomarker Sciences, as applicable). The medical monitor serves as the chair of the SRT. An SRT charter (or similar document) will be agreed on by all SRT members prior to the first SRT meeting. The data reviewed at the SRT meeting to make dose de-escalation decisions will be defined in the SRT charter (or similar document). The quality control checks performed on the data reviewed and used for making decisions will be described in the SRT charter (or similar document).

1.2.6. Phase 2 Cohort 1

Once the SRT reviews Safety Run-in 1 and the sponsor determines the RP2D for magrolimab + pembrolizumab + 5-FU + platinum, a Phase 2, randomized, open-label cohort with a 2 treatment-group design will open for enrollment. Approximately 132 patients with untreated metastatic or unresectable, locally recurrent HNSCC regardless of PD-L1 status will be randomized in a 1:1 ratio to receive either magrolimab + pembrolizumab + platinum + 5-FU (Arm A) or pembrolizumab + platinum + 5-FU (Arm B). The primary efficacy assessment will be PFS by investigator assessment with the primary analysis to occur after 93 events. Stratification factors for randomization include the following:

- PD-L1 expression (CPS ≥ 1 versus CPS < 1)
- p16 HPV status (positive versus negative)

Once the Phase 2 Cohort 1 enrolls 20 patients in each Arm A and Arm B, a third arm will open: magrolimab + zimberelimab + platinum + 5-FU (Arm C; n = 46). Randomization will continue 1:1:1 across all 3 arms, with the same stratification factors.

1.2.7. Phase 2 Cohort 2

Dose finding for magrolimab + pembrolizumab alone (without chemotherapy) may be performed by the sponsor, as needed. An optional Phase 2 Cohort 2 will open for enrollment once the RP2D for magrolimab + pembrolizumab is determined. Approximately 25 patients with untreated metastatic or unresectable, locally recurrent HNSCC with a PD-L1 CPS ≥ 1 will be enrolled to receive magrolimab + pembrolizumab. The primary efficacy assessment will be investigator-assessed ORR.

1.2.8. Phase 2 Cohort 3

Once the SRT reviews Safety Run-in 2 and the sponsor determines the RP2D for magrolimab + docetaxel, the Phase 2 Cohort 3 will open for enrollment. Approximately 40 patients with locally advanced/mHNSCC regardless of PD-L1 status who were previously treated with at least 1 and no more than 2 lines of prior systemic therapy will be enrolled to receive magrolimab + docetaxel. The primary efficacy assessment will be investigator-assessed ORR.

1.2.9. Diagnostic Tissue Testing

Tumor expression of PD-L1 will be evaluated prospectively using an assay approved by the FDA for detection of PD-L1 in HNSCC tissues.

Testing for p16 HPV positivity will be performed by immunohistochemistry analysis using the CINtec® Histology (p16) assay (Ventana Medical Systems, Inc., Tucson, AZ). A central laboratory may be used if testing as specified is not available locally. Oral cavity, hypopharynx, and larynx cancers are not required to undergo HPV testing by immunohistochemistry, as by convention they are assumed to be HPV negative.

1.3. Sample Size and Power

For Phase 2 Cohort 1 Arm A versus Arm B, a total of 93 PFS events provides 75% power at a 1-sided alpha of 0.15 to detect a hazard ratio (HR) of 0.7 (assuming median PFS of 7 months compared with a control group median PFS of 4.9 months) using an unstratified log-rank test. Assuming an accrual period of 15 months, a minimum follow-up time of 5 months, and a 10% annual drop-out rate, 66 patients per treatment group (Arm A versus Arm B) would be required to obtain 93 events. Once the Phase 2 Cohort 1 enrolls 20 patients in each Arm A and Arm B, Arm C will open. For Arm C versus Arm B, a total of 61 PFS events provides 64% power at a 1-sided alpha of 0.15 to detect an HR of 0.7 (assuming a median PFS of 7 months compared with a control group median PFS of 4.9 months) using an unstratified log-rank test. Assuming an approximate accrual of 11 months, a minimum follow-up time of 5 months, and a 10% annual drop-out rate, 46 patients per treatment arm (Arm C versus concurrent Arm B) would be required to obtain 61 events. The control group assumption is based on pembrolizumab + platinum + 5-FU efficacy in a historical study (KEYNOTE-048) {Burtness 2019}. The power calculation was performed using EAST 6.5.

For Phase 2 Cohort 2, no formal sample size calculation has been performed (25 patients).

For Phase 2 Cohort 3, a sample size of 40 patients provides 83% power at a 1-sided alpha of 0.15 to detect an ORR of 18% compared with a null ORR of 7.9% using a chi-squared test. The null ORR is based on historical taxane efficacy data in the second-line setting {Cohen 2019, Ferris 2016}. The power calculation was performed using nQuery 8.0.

2. TYPE OF ANALYSIS

2.1. Interim Analyses

2.1.1. Dose Determination Analysis

For the purposes of making the dose de-escalation decisions for the safety run-in evaluations, dose determination analyses of relevant safety data focusing on DLTs and overall safety profile was conducted by the sponsor after all patients have completed 1 dosing cycle (21 days), as defined in Section 1.2. Dose Determination Analyses have been completed and are not within the scope of this SAP.

2.1.2. Informal Interim Analysis

For the purpose of potential phase 3 planning, interim efficacy and safety analyses are planned at approximately 13 weeks after 20 subjects are treated with magrolimab at the RP2D in Safety Run-In 1 or Phase 2 Cohort 1, and at approximately 13 weeks after 20 subjects are treated with magrolimab at the RP2D in Safety Run-In 2 or Phase 2 Cohort 3.

Additional informal interim analyses may be conducted to support decision making concerning the current clinical study, the sponsor's clinical development projects in general or in case of any safety concerns.

2.2. Primary Analysis

For the primary analysis of each cohort, outstanding data queries will have been resolved or adjudicated as unresolvable, and the data will have been cleaned and finalized for the analysis.

For Phase 2 Cohort 1, the primary analysis of PFS will be conducted after 93 PFS events occur in Arm A and Arm B.

For Phase 2 Cohorts 2 and 3, the primary analysis of ORR will be conducted 6 months after the last patient is enrolled.

Due to the termination of the program, the primary analysis has been terminated.

2.3. Final Analysis

The final analysis will be performed after all patients have completed the study or discontinue 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.

3.1. Analysis Sets

For each analysis set, the number and percentage of patients eligible for inclusion, as well as the number and percentage of patients who were excluded and the reasons for their exclusion, will be summarized by treatment group.

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 Analysis Set

Intent-to-treat (ITT) Analysis Set includes all patients who were randomized in Phase 2 Cohort 1 of the study.

This is the primary analysis set for efficacy analyses in Phase 2 Cohort 1.

3.1.3. Modified Intent-to-Treat Analysis Set

The Modified Intent-to-Treat (mITT) Analysis Set includes all patients who took at least 1 dose of study drug. This is the primary analysis set for efficacy analyses for Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3.

3.1.4. Safety Analysis Set

The Safety Analysis Set includes all patients who took at least 1 dose of any study drug. Analysis will be conducted according to the actual treatment received. This is the primary analysis set for safety analyses, except for DLT analyses.

3.1.5. Dose-Limiting Toxicity (DLT)-Evaluable Analysis Set

For the safety run-in evaluations, the primary analysis set for the DLT analysis is the DLT-Evaluable Analysis Set, defined as all patients in the safety run-in evaluations who meet either of the following criteria during the DLT assessment period:

- The patient experiences a DLT at any time after initiation of the first infusion of magrolimab.

- The patient does not experience a DLT and completes at least 2 infusions of magrolimab and at least 1 dose of pembrolizumab, 1 dose of platinum, and 1 dose of 5-FU for Safety Run-in 1; at least 1 dose of docetaxel for Safety Run-in 2; and at least 1 dose of pembrolizumab for the pre-expansion safety run-in for magrolimab + pembrolizumab (if applicable).

Patients who are not evaluable for DLT assessment in the safety run-in evaluations would be replaced. For patients who are replaced but received at least 1 dose of any study drug, they would be included in the Safety Analysis Set and not in the DLT-Evaluable Analysis Set.

3.1.6. Pharmacokinetic Analysis Set

The Pharmacokinetic (PK) analysis for magrolimab will be conducted on the Magrolimab PK Analysis Set, defined as all patients who receive any amount of magrolimab and have at least 1 evaluable post-treatment serum concentration of magrolimab.

The PK analysis for zimberelimab may be conducted on the Zimberelimab PK Analysis Set when appropriate, defined as all patients who receive any amount of zimberelimab and have at least 1 evaluable post-treatment serum concentration of zimberelimab.

3.1.7. Immunogenicity Analysis Set

The immunogenicity analysis for magrolimab will be conducted on the Magrolimab Immunogenicity Analysis Set, defined as all patients who receive any amount of magrolimab and have at least 1 evaluable anti-magrolimab antibody test result.

The immunogenicity analysis for zimberelimab may be conducted on the Zimberelimab Immunogenicity Analysis Set when appropriate, defined as all patients who receive any amount of zimberelimab and have at least 1 evaluable anti-zimberelimab antibody test result.

3.2. Subject Grouping

3.2.1. Subject Grouping for Each Analysis Set

For 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 All Enrolled Analysis Set and mITT Analysis Set, patients will be grouped according to the treatment to which they were enrolled. For analyses based on the Safety Analysis Set and DLT-Evaluable Analysis Set, patients will be grouped according to the actual treatment received. The actual treatment received will differ from the randomized treatment only when their actual treatment differs from randomized treatment for the entire treatment duration.

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

3.2.2. Subject Grouping for Each Type of Analysis

3.2.2.1. Informal Interim Analysis

For patients who are in Safety Run-In 1 and Phase 2 Cohort 1, the informal interim analysis (IA) is planned at approximately 13 weeks after 20 subjects are treated with magrolimab at the RP2D. The efficacy endpoints will be analyzed by treatment group using pooled mITT analysis set from Safety Run-In 1 and ITT analysis set from Phase 2 Cohort 1. Patients will be grouped according to the treatment to which they were enrolled/randomized as following:

- Safety Run-In 1 + Phase 2 Cohort 1 Arm A (magrolimab + pembrolizumab + platinum + 5-FU)
- Phase 2 Cohort 1 Arm B (pembrolizumab + platinum + 5-FU)
- Phase 2 Cohort 1 Arm C (magrolimab + zimberelimab + platinum + 5-FU)

The safety endpoints will be analyzed by treatment group using safety analysis set from Safety Run-In 1 and Phase 2 Cohort 1. Patients will be grouped according to the actual treatment received as following:

- Safety Run-In 1 + Phase 2 Cohort 1 Arm A (magrolimab + pembrolizumab + platinum + 5-FU)
- Phase 2 Cohort 1 Arm B (pembrolizumab + platinum + 5-FU)
- Phase 2 Cohort 1 Arm C (magrolimab + zimberelimab + platinum + 5-FU)

For patients who are in in Safety Run-In 2 and Phase 2 Cohort 3, the informal IA is planned at at approximately 13 weeks after 20 subjects are treated with magrolimab at the RP2D. The efficacy endpoints will be analyzed using mITT analysis set. The safety endpoints will be analyzed using safety analysis set. Patients will be pooled from Safety Run-In 2 and Phase 2 Cohort 3.

3.2.2.2. Primary Analysis

For Phase 2 Cohort 1, the primary analysis of PFS will be conducted after 93 PFS events occur in Arm A and Arm B. Efficacy for Safety Run-In 1 will be listed only. Efficacy endpoints for Phase 2 Cohort 1 will be summarized by treatment group using ITT analysis set. Safety and other analyses will be summarized for Safety Run-In 1 and Phase 2 Cohort 1 by phase and treatment group.

For Phase 2 Cohort 3, the primary analysis of ORR will be conducted 6 months after the last patient is enrolled. Efficacy for Safety Run-In 2 will be listed only. Efficacy endpoints for Phase 2 Cohort 3 will be summarized using mITT analysis set. Safety and other analyses will be summarized for Safety Run-In 2 and Phase 2 Cohort 3 by phase.

Due to the termination of the program, the primary analysis has been terminated.

3.2.2.3. Final Analysis

For analyses to be performed at the final analysis, same grouping strategy as the primary analysis will be used.

3.3. Strata and Covariates

In Phase 2 Cohort 1, approximately 132 patients will be randomized via the interactive response technology (IRT) in a 1:1 ratio using a stratified randomization schedule to receive either magrolimab + pembrolizumab + platinum + 5-FU (Arm A) or pembrolizumab + platinum + 5-FU (Arm B). Stratification factors for randomization include the following:

- PD-L1 expression (CPS $\geq$ 1 versus CPS < 1)
- p16 HPV status (positive versus negative)

Once the Phase 2 Cohort 1 enrolls 20 patients in each Arm A and Arm B, a third arm will open: magrolimab + zimberelimab + platinum + 5-FU (Arm C; n = 46). Randomization will continue 1:1:1 across all 3 arms, with the same stratification factors.

If there are discrepancies in stratification factor values between the IRT and the clinical database, the values recorded in the IRT will be used for the primary analyses. Efficacy endpoints will be evaluated using stratification factors as covariates or stratification variables for analyses, as specified in Section 6.

In the situation where there is insufficient information in a stratum, pooling of the stratum with the adjacent stratum for stratified analyses will be considered.

Stratified analyses are not applicable to Safety Run-in Cohorts or Phase 2 Cohorts 2 and 3 of the study.

3.4. Examination of Subject Subgroups

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

3.5. Adjustment for Multiplicity

CCI [redacted]

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 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 Day 1 visit date will be used instead. If an enrolled or randomized subject was not dosed with any study drug, the enrollment or randomization date will be used instead of the Day 1 visit date. For screen failures, the date the first informed consent was signed will be used for the age derivation. Age required for longitudinal and temporal calculations and analyses (e.g., 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 lower LOQ at the same precision level of the originally reported value will be used to calculate descriptive statistics if the datum is reported in the form of "x" (where x is considered the lower 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 upper LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the upper LOQ). Values with decimal points will follow the same logic as above.
- The lower or upper 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 lower or upper LOQ, respectively).

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 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 day of first dose of any study drug administration. For Phase 2 Cohort 1, if the subject is randomized but not dosed, the randomization date will be study day 1. For Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3, if the subject is enrolled but not dosed, the enrollment 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 nominal 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	53
Week 9	64	54	74
Week xx[d]	(xx)*7 + 1	(xx)*7 - 9	(xx)*7 + 11

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

The analysis windows for PRO are provided in Table 3-2.

Table 3-2. Analysis Visit Windows for By-visit Summaries of PRO assessments

Analysis Visit	Nominal Study Day	Visit Window Study Day	
		Lower Limit	Upper Limit
Baseline		(none)	1
Week 3	22	2	32
Week 6	43	33	53
Week 9	64	54	74
Week xx[a]	(xx)*7 + 1	(xx)*7 - 9	(xx)*7 + 11

a xx >= 12.

3.8.3. Selection of Data in the Event of Multiple Records in an Analysis Visit Window

If multiple valid, nonmissing measurements exist in an analysis window, 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 if available) will be selected, unless specified differently. For Phase 2 Cohort 1, if the subject is randomized but not dosed, the last nonmissing value on or prior to the randomization date will be selected as baseline. For Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3, if the subject is enrolled but not dosed, the last nonmissing value on or prior to the enrollment date will be selected as baseline. 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 lowest 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 may be handled in the analysis:

- Study treatment discontinuation due to COVID-19
- Protocol deviations due to COVID-19
- Adverse events due to COVID-19
- Death 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 cohort and treatment group regardless of whether they were exempted by the sponsor or not, based on the 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 cohort and treatment group for the All Enrolled Analysis Set.

5. SUBJECT INFORMATION

Generally, listings and disposition table will be based on All Enrolled Analysis Set. Summary of treatment exposure, prior medications, concomitant medications and post-treatment anti-cancer therapies will use Safety Analysis Set. The remaining table summaries in this section will be based on ITT Analyses Set for Phase 2 Cohort 1 and mITT Analysis Set for Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3, unless specified otherwise.

5.1. Patient Enrollment and Disposition

Key study dates, including first patient screened, first patient enrolled or randomized, last patient enrolled or randomized, 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 group for each country, investigator and overall.

A similar enrollment table will be provided for Phase 2 Cohort 1 by randomization stratum. If there are discrepancies in the value used for stratification assignment between the IRT and the clinical database, the value collected in the IRT will be used for the summary.

A summary of patient disposition will be provided by cohort and treatment group. 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:

- Safety Analysis Set
- Continuing study treatment (for each study drug)
- Discontinued study drug (for each study drug) with reasons for treatment discontinuation for the corresponding drug
- On-going in study
- 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 specific analysis set corresponding to that column.

A by-subject listing of reasons for study drug and study discontinuation will be provided by subject identification (ID) number in ascending order to support the above summary tables.

5.2. Extent of Study Treatment Exposure

The study drugs including Magrolimab, Pembrolizumab, Zimberelimab, Cisplatin, Carboplatin, 5-Fluorouracil and Docetaxel. Extent of exposure to study treatment will be summarized using descriptive statistics for total duration of exposure to each study drug, total number of infusions for each study drug, total number of cycles received for each study drug, total cumulative dosage administered for Magrolimab, relative dose intensity (%) of Magrolimab, and number (%) of patients with dose modifications (i.e., infusion interruption, dose delayed or not administered) and reasons for each study drug by cohort and treatment group.

5.2.1. Duration of Exposure to Study Drug

Total duration of exposure to each study drug 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 (e.g., 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, 3 weeks, 6 weeks, 9 weeks, 12 weeks, 15 weeks, 18 weeks, 21 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 below formula for each study drug:

$$\text{Relative Dose Intensity (\%)} = \frac{\text{Total Amount of Study Drug Administered}}{\text{Total Amount Study Drug Expected to be Administered}} \times 100$$

Descriptive statistics for the relative dose intensity with the number and percentage of patients belonging to relative dose intensity categories (eg, $< 75\%$, ≥ 75 to $< 90\%$, $\geq 90\%$) will be provided by cohort and treatment group for the Safety Analysis Set.

A by-subject listing of each study drug administration will be provided by cohort and treatment group, subject ID number (in ascending order), and visit (in chronological order).

5.3. Demographics and Baseline characteristics

Demographic summaries will include age, sex, race and ethnicity. Baseline data will include a summary of body weight, height, body mass index (BMI), body surface area (BSA), ECOG performance status, and randomization stratification factors (i.e., PD-L1 expression and p16 status for Phase 2 Cohort 1). The analysis will be performed using ITT Analysis Set for Phase 2 Cohort 1, and mITT analysis set for Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3. By-subject listing will be generated by cohort, treatment group and subject ID number.

5.4. Prior Anti-cancer Therapy

Prior Adjuvant/Neoadjuvant therapy (Yes/No) will be summarized for Safety Run-in 1, Phase 2 Cohort 1 and the optional Phase 2 Cohort 2 (if opened).

Prior Metastatic therapy (Yes/No), number of prior regimens in the metastatic setting, and best response for the last prior therapy in the metastatic setting before enrollment date will be summarized for Safety Run-in 2 and Phase 2 Cohort 3.Prior anti-cancer therapy will be coded using the World Health Organization (WHO) Drug. Dictionary and summarized by Preferred Drug Name.

A by-patient listing including information collected in prior anti-cancer therapy eCRF and all derived information above will be provided.

5.5. Prior and On Study Radiotherapy

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

Medical history will be collected at screening for disease-specific and general conditions (i.e., conditions not specific to the disease being studied). General medical history will be coded using the current version of Medical Dictionary for Regulatory Activities (MedDRA). Disease-specific medical history will be summarized by cohort and treatment group for each condition.

A by-patient listing for general medical history and the coded terms including SOC and PT 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.

5.8.1. Prior Medications

Prior medications are defined as any medications taken before a patient took the first study drug.

Prior medications will be summarized by Anatomical Therapeutic Chemical (ATC) drug class Level 2 preferred name using the number and percentage of subjects by cohort and treatment group. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary

will be ordered alphabetically by ATC medical class and then by preferred term in order of descending overall frequency within each ATC medical class. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to the first dosing date of study drug will be included in the prior medication summary regardless of when the stop date is. If a partial start date is entered the medication will be considered prior unless the month and year (if day is missing) or year (if day and month are missing) of the start date are after the first dosing date. Medications with a completely missing start date will be included in the prior medication summary, unless otherwise specified. Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

5.8.2. Concomitant Medications

Concomitant medications are defined as medications taken while a patient took study treatment. Use of concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) drug class Level 2 preferred name using the number and percentage of patients by cohort and treatment group. A patient reporting the same medication more than once will be counted only once when calculating the number and percentage of patients who received that medication. Medications may appear under multiple ATC drug classes. The summary will be ordered alphabetically by ATC medical class and then by preferred term in descending overall frequency within each ATC medical class. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medications with a start date prior to or on the first dosing date of study treatment and continued to be taken after the first dosing date, or started after the first dosing date but prior to or on the 30 days after the last dosing date of study treatment (or 90 days after the last dosing date of zimberelimab) will be considered concomitant medications. Medications started and stopped on the same day as the first dosing date or 30 days after the last dosing date of study treatment (or 90 days after the last dosing date of zimberelimab) will also be considered concomitant. Medications with a stop date prior to the date of first dosing date of study treatment or a start date after the last dosing date of study treatment plus 30 days (or last dosing date of zimberelimab plus 90 days) will be excluded from the concomitant medication summary. Medications with partially or completely missing start and stop dates will be included in the concomitant medication summary, unless the partial missing date suggested otherwise. Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

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 study treatment anti-cancer therapies including flag for on-study treatment and post study treatment therapies will be provided in a by-patient listing sorted by patient ID number and administration date in chronological order. Note that anti-cancer therapies (other than study drugs) are not allowed during the study drug treatment per protocol. The analysis will be performed for Safety Analysis Set.

6. EFFICACY ANALYSES

Generally, summary of efficacy endpoints in this section will be based on ITT Analyses Set for Phase 2 Cohort 1 and mITT Analysis Set for Safety Run-in Cohorts and Phase 2 Cohorts 2 and 3, unless specified otherwise. Listings will be based on All Enrolled Analysis Set.

6.1. Primary Efficacy Endpoints

6.1.1. Definition of the Primary Efficacy Endpoints

The primary efficacy endpoint of this study including:

- Progression-free survival (PFS) by investigator assessment for Phase 2 Cohort 1 Arm A versus Arm B
- Objective response rate (ORR) by investigator assessment for Phase 2 Cohorts 2 and 3

Progression-Free Survival (PFS) in Phase 2 Cohort 1

For Phase 2 Cohort 1, PFS is defined as the interval from randomization date to the earlier date of the first documentation of objective disease progression or death from any cause. Disease progression is determined based on RECIST, Version 1.1.

Objective Response Rate (ORR)

ORR is defined as the proportion of patients who achieve the best overall response (BoR) of complete response (CR) or partial response (PR) that is confirmed at least 4 weeks after initial documentation of response. Tumor response assessments after the date of patients receiving new anticancer therapy will be excluded from the analysis. The response definition of each response category is based on RECIST, Version 1.1. Patients who do not have baseline or on-study response status assessment or received new anticancer therapy prior to achieving CR or PR, will be considered as non-responders.

6.1.2. Analysis of the Primary Efficacy Endpoint

For Phase 2 Cohort 1 Arm A versus Arm B, the primary analysis of 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 stratified by the stratification factors as recorded in the IRT will be used to compare treatment difference in PFS by investigator assessment.

In addition, the treatment effect will be estimated by HR along with its 2-sided 95% CI using the Cox proportional hazards regression model stratified by the stratification factors at randomization.

For Phase 2 Cohorts 2 and 3, the ORR by investigator assessment along with its 2-sided 95% exact confidence interval (CI) based on Clopper-Pearson method will be estimated. Patients who do not have sufficient baseline or on-study tumor assessment to characterize response will be counted as nonresponders.

6.1.3. Sensitivity Analysis of the Primary Efficacy Endpoint

In order to evaluate the robustness of the primary endpoint PFS by investigator assessment for Phase 2 Cohort 1 Arm A versus Arm B, sensitivity analyses will be performed with a different set of censoring rules. The censoring rules for primary and sensitivity analyses are summarized in Table 6-1.

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

Situation	Primary Analysis	Sensitivity Analysis 1
Documented Progressive Disease (PD)	Progressed at date of earliest sign of PD	Same as "Primary Analysis"
Death before first PD	Progressed at date of death	Same as "Primary Analysis"
Death within 2 disease assessments window and no baseline or post-baseline assessment	Progressed at date of death	Same as "Primary Analysis"
Alive and progression-free	Censored at date of last evaluable disease assessment	Same as "Primary Analysis"
Initiated subsequent anticancer therapy prior to PD or death	Censored at date of last evaluable assessment on or prior to subsequent anticancer therapy	Progressed at date of PD or death
PD or death immediately after ≥ 2 consecutively missed or not evaluable (NE) disease assessments	Censored at date of last evaluable assessment prior to missed or NE assessments	Progressed at date of documented PD or death
No disease assessment at baseline or post-baseline assessment	Censored at date of randomization (Phase 2 Cohort 1) or first dosing date (Phase 2 Cohorts 2 and 3)	Same as "Primary Analysis"

Given the scheduled visit assessment scheme (i.e., every 6 weeks (± 7 days) until 36 weeks then every 9 weeks (± 7 days) thereafter) the definition of 2 missed visits will change. The interval of 2 missing visits will be 2 x 6 weeks + 14 days if the two missing visits occur before scheduled frequency of assessments changes and 2 x 9 weeks + 14 days if the two missing visits occur after scheduled frequency of assessments changes. If the two missed visits occur over the period when the scheduled frequency of assessments changes from every 6 weeks to every 9 weeks this will be: 6 weeks for an early assessment + 9 weeks for a late assessment + 14 days.

6.2. Secondary Efficacy Endpoints

6.2.1. Definition of Secondary Efficacy Endpoints

The secondary efficacy endpoint of this study including:

- PFS by investigator assessment for Phase 2 Cohort 1 Arm C versus concurrent Arm B
- PFS for Phase 2 Cohorts by investigator assessment
- ORR by investigator assessment for Phase 2 Cohorts
- Duration of response (DOR) for Phase 2 Cohorts
- Overall Survival (OS) for Phase 2 Cohorts
- PRO assessment scores for Phase 2 Cohorts

Progression-Free Survival (PFS) in Phase 2 Cohorts 2 and 3

For Phase 2 Cohorts 2 and 3, PFS is defined as the interval from the first dosing date of any study treatment to the earlier date of the first documentation of objective disease progression or death from any cause. Disease progression is determined based on RECIST, version 1.1.

Overall Survival (OS) for Phase 2 Cohort 1

For Phase 2 Cohort 1, OS is defined as the interval from randomization to death from any cause. For patients who were not known to have died at the time of the analysis, OS data will be censored at the last date that they were known to be alive.

Overall Survival (OS) for Phase 2 Cohorts 2 and 3

For Phase 2 Cohorts 2 and 3, OS is defined as the interval from first dosing date of study treatment to death from any cause. For patients who were not known to have died at the time of the analysis, OS data will be censored at the last date that they were known to be alive.

Duration of response (DOR)

DOR is defined for confirmed responders as duration of time from the date of initial response to the date of first documentation of disease progression or the date of death due to any cause, whichever occurs first. Date of initial response is the date of first response achieved and then confirmed by a subsequent disease assessment conducted on or prior to the initiation of the next line of anticancer therapy. DOR will follow the same censoring rule as PFS primary analysis.

PRO endpoints measured by EORTC QLQ-C30 and QLQ-H&N35

The EORTC QLQ-C30 includes 30 separate questions (items) resulting in 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) {Fayers 2001} (See details in Appendix 2). The recall period is 1 week (past week). It will take about 11 minutes to complete.

The EORTC QLQ-H&N35 is a 35-item measure that supplements the EORTC QLQ-C30 and assesses symptoms, impacts, and treatment symptoms associated with head and neck cancer. The instrument consists of 18 domains comprising of seven multi-item symptom scales (pain, swallowing, sense, speech, social eating, social contact, and sexuality) and 11 single-item symptom scales (teeth, open mouth, dry mouth, saliva, coughing, feeling ill, pain killers, supplements, feeding tube, weight loss, and weight gain). Each item is rated on a 4-point verbal rating scale and has a recall period of the "past week".

According to the EORTC 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): $RS = (I_1 + I_2 + \ldots + I_n)/n$.

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, Janssen 2013}. 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.

6.2.2. Analysis Methods for Secondary Efficacy Endpoints

The analysis of secondary endpoints PFS (by investigator assessment for Phase 2 Cohort 1 Arm C versus concurrent Arm B and by investigator assessment for Phase 2 Cohorts 1-3), OS and DOR will be performed using the Kaplan-Meier method. Median, Q1, and Q3 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 be provided.

The ORR by investigator assessment along with its 2-sided 95% exact confidence interval (CI) based on Clopper-Pearson method will be provided.

PRO endpoints

The descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores at each post-baseline visit for each scale (or domain) of EORTC QLQ-C30 and EORTC QLQ-H&N35 by cohort and treatment group.

The frequency of observed scores at each scheduled visit will be summarized for each dimension of EQ-5D by cohort and treatment group. The descriptive statistics will be used to summarize the observed scores at each scheduled visit and the change from baseline scores for EQ-VAS at each post-baseline visit by cohort and treatment group.

6.3. Other Efficacy Definition and Analyses

6.3.1. Other Definition Related to Efficacy

The PFS, ORR, DOR, OS for Safety Run-In Cohorts have the same definitions as those for Phase 2 Cohorts 2 and 3.

Best overall response (BoR)

Best overall response response (BoR) is calculated based on the overall visit responses from each RECIST assessment including unscheduled visits. It is the best response a patient has had from the start of treatment until objective documentation of PD (per RECIST v1.1), or patient withdrawn from the study, or patient started new anticancer therapy, whichever occurs first. Categorization of BoR will be based on RECIST v1.1 using the following response categories: CR, PR, SD, PD and NE.

A BoR of CR or PR must be confirmed. A BOR of CR/PR requires confirmation no less than 4 weeks (28 days) after the first response criterion was met and with no evidence of progression between the initial and CR/PR confirmation visit. Patients who responded with an unconfirmed

CR/PR at the time of data cutoff will be reported under the SD category provided the minimum criteria for SD duration are met, otherwise this will be reported under the NE category.

For determination of a BoR of SD, the SD should be recorded at least 42 days after randomization date (Phase 2 Cohort 1) or the first dosing date (Safety Run-in Cohorts or Phase 2 Cohorts 2 and 3). For the determination of BoR, an overall visit response with "Non-CR/non-PD" is considered as SD.

6.3.2. Other Analyses Methods Related to Efficacy

The PFS, ORR, DOR, OS for Safety Run-In Cohorts may be summarized by pooling with their corresponding Phase 2 Cohorts or listed only.

Best objective response (BoR)

For each treatment arm, best objective response (BoR) will be summarised by n (%) for each category (CR, PR, SD, PD and NE). No formal statistical analyses are planned for BoR.

6.4. Changes From Protocol-Specified Efficacy Analyses

There are no deviations from the protocol-specified Efficacy analyses.

7. SAFETY ANALYSES

Safety analysis will be performed in the Safety Analysis Set. Listing will be based on the All Enrolled Analysis Set. 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 sponsor. 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 Global Patient Safety before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

A Treatment-emergent adverse events (TEAE) is defined as any AE that begins on or after the date of the first dose of study drug up to and including last dose date of study drug plus 30 days (or last dose date of zimberelimab plus 90 days) and prior to the day of initiation of subsequent anti-cancer therapy. 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 is a missing or incomplete date for the start date or stop date of an AE, the most conservative approach is 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) 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) of the date corresponding to the end date of treatment-emergent period.

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

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 group.

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, and cohort and treatment group:

- TEAEs
- TEAEs with Grade 3 or 4

- TE treatment-related AEs for any study drug and for each study drug
- TE treatment-related AEs with Grade 3 or 4 for any study drug and for each study drug
- TE SAEs
- TE treatment-related SAEs for any study drug and for each study drug
- TEAEs leading to dose reduction of for any study drug and for each study drug
- TEAEs leading to dose delay or interruption of for any study drug and for each study drug
- TEAEs leading to discontinuation of any study drug and for each study drug
- TEAEs leading to death
- TEAEs leading to death related to magrolimab

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 and TEAEs leading to death will be summarized by PT only 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 study drug
- All AEs leading to dose reduction of study drug
- TEAEs leading to dose delay or interruption of study drug

A summary (number and percentage of patients) of deaths will be provided for each cohort and treatment group. 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. Treatment Emergent Adverse Events (TEAE) of Clinical Importance

The TEAE of clinical importance are defined in Table 7-1. Number and percentage of subjects with the following TEAE of clinical importance will be summarized by PT for each cohort and treatment group.

Table 7-1. 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 each laboratory test specified in the 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

Plots of lab parameters may include (but not limited to) hemoglobin, hematocrit, 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 last dose date of study drug plus 30 days (or last dose date of zimberelimab plus 90 days) and prior to the day of initiation of subsequent anti-cancer therapy. 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. Treatment-Emergent Marked Laboratory Abnormalities

Treatment-emergent marked laboratory abnormalities are defined as values that increase from baseline by at least 3 toxicity grades at any postbaseline time point up to and including last dose date of study drug plus 30 days (or last dose date of zimberelimab plus 90 days) and prior to the day of initiation of subsequent anti-cancer therapy. If the relevant postbaseline lab (with at least 3 toxicity grades increasing from baseline) is assessed on or before the last dose date, the laboratory abnormality is considered as treatment-emergent marked laboratory abnormality, regardless of the initiation of subsequent anti-cancer therapy. If the relevant baseline laboratory value is missing, any Grade 3 or 4 values observed within the timeframe specified above will be considered treatment-emergent marked abnormalities.

7.2.2.3. Summaries of Laboratory Abnormalities

The following summaries (number and percentage of patients) for treatment-emergent (TE) 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:

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

For all summaries of laboratory abnormalities, the denominator is the number of patients in the safety analysis set with nonmissing postbaseline values during the treatment-emergent period.

By-patients listings of all laboratory test results that were collected throughout the study for the laboratory test of interest will be provided for each cohort and treatment group, with flags for treatment-emergent laboratory abnormalities and treatment-emergent marked laboratory abnormalities. By-patient listing for post-infusion hematology will also be provided.

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 subjects 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 during the treatment-emergent period. For individual laboratory tests, subjects will be counted once based on the most severe postbaseline values. For both the composite endpoint of AST or ALT and total bilirubin, subjects will be counted once when the criteria are met at the same postbaseline visit date. The denominator is the number of subjects in the Safety Analysis Set who have at least one postbaseline visit, at which all the relevant tests are nonmissing at the same postbaseline visit date.

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

7.3. Body Weight, Height, and Vital Signs

Descriptive statistics will be provided by cohort and treatment group for body weight, height BMI, blood pressure, heart rate, temperature, and respiratory rate as follows:

- Baseline values
- Values at each postbaseline time point
- Change and percentage change from baseline at each postbaseline time point

A baseline value will be defined as the last available value collected on or prior to the first dosing date of study drug (and prior to first dosing time if available). For subjects who were not dosed, the last available value collected on or prior to the randomization date or enrollment date will be selected as baseline.

A by-subject listing of vital signs will be provided by subject ID number and time point in chronological order. Body weight, height, and BMI will be included in the vital signs listing, if space permits. If not, they will be provided separately.

7.4. Electrocardiogram Results

A by-subject listing of ECG results will be provided by subject ID number in ascending order for each cohort and treatment group.

7.5. Other Safety Measures

A by-subject listing will be provided by subject ID number in ascending order for ECOG performance status.

A by-subject listing will be provided by subject ID number in ascending order for pregnancy test results.

7.6. Changes From Protocol-Specified Safety Analyses

There are no deviations from the protocol-specified safety analyses.

8. CHANGES FROM PROTOCOL-SPECIFIED ANALYSES

Due to the termination of the program, the primary analysis has been terminated.

9. PHARMACOKINETIC (PK) AND IMMUNOGENICITY ANALYSES

9.1. PK and Anti-drug-antibody (ADA) Sample Collection

Blood samples for PK assessment will be collected at predose at multiple time points and at 1 hour after the end of infusion (EOI) in Cycle 3 Day 1 from patients who received magrolimab according to the schedule of assessments in protocol Appendix Table 1. Blood samples for PK assessment will be collected at predose and at the EOI on Cycle 1 Day 1 and Cycle 3 Day 1 from patients who received zimberelimab according to the schedule of assessments in protocol Appendix Table 1.

Peripheral blood for immunogenicity assessments for ADAs against magrolimab and zimberelimab will be collected as described in the schedule of assessments in protocol Appendix 2. When collected on the day of study drug dosing, the ADA blood sample must be collected at the same time as the predose PK sample.

9.2. 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 cohort. Boxplot plots of serum concentration by visit will be generated for magrolimab.

Serum concentrations for zimberelimab may be listed and summarized using descriptive statistics by sampling time point.

9.3. Immunogenicity analysis

9.3.1. Definition of Terminology

Subjects Evaluable for ADA Prevalence: subjects who have at least one reportable ADA result at baseline or post-baseline.

Subjects Evaluable for ADA Incidence: subjects who have non-missing baseline ADA result and at least one reportable ADA result at post-baseline.

ADA Prevalence: the proportion of subjects who had at least one positive ADA sample (baseline or post-baseline) based on the Immunogenicity Analysis Set.

Treatment-Induced ADA Rate: the proportion of subjects who had negative baseline ADA sample and at least one positive post-treatment ADA sample based on subjects who had both non-missing baseline and at least one post-treatment ADA result reported (i.e. ADA Incidence Analysis Set).

Treatment-Boosted ADA Rate: the proportion of subjects 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 based on the ADA Incidence Analysis Set.

ADA Incidence (treatment-emergent ADA): the proportion of subjects who had treatment-induced or treatment-boosted ADA based on subjects who had non-missing baseline ADA sample and at least one post-treatment ADA result reported in Immunogenicity Analysis Set.

Persistent ADA is defined as:

a) 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

b) Treatment-Induced ADA detected in the last sampling time point of the treatment study period.

ADA Persistence Rate: the proportion of subjects who had persistent ADA based on the ADA Incidence Analysis Set.

Transient ADA is defined as:
Treatment-Induced ADA that does not meet the definition of persistent ADA. The proportion of subjects who had transient ADA is based on the subjects evaluable for ADA incidence.

Neutralizing antibody (NAb) Incidence: the proportion of subjects who had at least one positive neutralizing antibody result reported based on the treatment-emergent ADA (treatment-induced or treatment-boosted ADA) among the subjects evaluable for ADA incidence.

9.3.2. Evaluation of Immunogenicity Data

The rate and magnitude of magrolimab anti-drug antibody (ADA) prevalence, incidence, persistence, and transience will be summarized for the Immunogenicity Analysis Set. Neutralizing antibody occurrence rate will also be summarized.

10. BIOMARKER ANALYSIS

Biomarker analysis will be described in a separate Biomarker Analysis Plan.

11. REFERENCES

Burtness B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro G, Jr., et al. Pembrolizumab Alone or With Chemotherapy Versus Cetuximab With Chemotherapy for Recurrent or Metastatic Squamous Cell Carcinoma of the Head and Neck (KEYNOTE-048): A Randomised, Open-Label, Phase 3 Study. Lancet 2019;394:1915-28.

Cohen EEW, Soulieres D, Le Tourneau C, Dinis J, Licitra L, Ahn MJ, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. Lancet 2019;393 (10167):156-67.

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:

Fayers PM. Interpreting quality of life data: population-based reference data for the EORTC QLQ-C30. Eur J Cancer 2001;37 (11):1331-4.

Ferris RL, Blumenschein G, Jr., Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med 2016;375 (19):1856-67.

Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013;22 (7):1717-27.

12. SOFTWARE

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

nQuery Advisor(R) Version 8.0. Statistical Solutions, Cork, Ireland.

EAST Version 6.5, Cytel Inc., MA, USA

13. SAP REVISION

Revision Date (DD MMM YYYY)	Section	Summary of Revision	Reason for Revision
10 Oct 2024	1.1, 6.1.1, 6.1.2, 6.1.3, 6.2.1, 6.2.2,	Replacing “by independent central review” by “as determined by investigator assessment”	For consistency with PA6, all instances of independent central review should be removed (replaced with investigator assessment) as we plan to close our central imaging vendor, Clario.
10 Oct 2024	7.1.7.1	Replacing “TEAE of Special Interest” by “TEAT of Clinical Importance	For consistency with other magrolimab solid tumor studies.
10 Oct 2024	7.1.7.1	Adding “Severe Neutropenia”, “Serious Infections”, “Pneumonitis”, and “Myocardial infarction” in TEAE of Clinical Importance.	Adding TEAE of Clinical Importance per Patient Safety.
10 Oct 2024	7.1.6	Replacing “Grade 3 or higher” by “Grade 3 or 4” in the summaries of adverse events	Following the new standarards per Patient Safety.
10 Oct 2024	7.1.6	Adding “TEAEs leading to death related to magrolimab” in the summaries of adverse events.	Information needed for Investigator's Brochure per Patient Safety.
10 Oct 2024	7.2.1	Adding “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” in the summaries of laboratory results.	For consistency with other magrolimab solid tumor studies.
18 Oct 2024	2.3	A sentence added to indicate that the primay analysis has been terminated.	The primary analysis is determincated due to program decision.
18 Oct 2024	3.2.2.2	A sentence added to indicate that the primay analysis has been terminated.	The primary analysis is determincated due to program decision.
18 Oct 2024	8	A sentence added to indicate that the primay analysis has been terminated.	The primary analysis is determincated due to program decision.

14. APPENDICES

Appendix 1. RECIST 1.1-based Assessments - Overall visit response
Appendix 2. Overview of EORTC QLQ-C30 and Questionnaire

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 arecombined 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	NE

CR Complete response, NA Not applicable, PR Partial response, SD Stable disease, PD Progression of disease, NE Not evaluable.

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

SAP_GS-US-548-5916_v2.0_RDMS_approval

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Early Development Lead (EDL) eSigned	21-Oct-2024 19:28:30
PPD	Biostatistics eSigned	21-Oct-2024 23:16:42