Statistical Analysis Plan

Protocol Title:	A Global Phase 1 Study Evaluating the Safety, Tolerability, Pharmacokinetics, and Efficacy of the Half-life Extended Bispecific T-cell Engager AMG 910 in Subjects With Claudin 18.2-Positive Gastric and Gastroesophageal Junction Adenocarcinoma
Short Protocol Title:	A Phase 1 Study of AMG 910 in Subjects with CLDN18.2-Positive Gastric and Gastroesophageal Junction Adenocarcinoma
Protocol Number:	20180292
NCT Number:	NCT04260191
Authors:	[redacted] and [redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive Thousand Oaks, CA 91320
SAP Date:	**Document Version** — **Date** Original (v1.0) — 22 February 2020 **Amendment 1 (v2.0)** — 15 June 2022

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	22Feb2020	N/A
Amendment 1 (v2.0)	15Jun2022	Based on Protocol Amendment 3

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
AUC	Area Under the Curve
BiTE	bispecific T cell engager
BLRM	Bayesian Logistic Regression Model
BMI	Body Mass Index
BOR	Best Overall Response
CI	Confidence Intervals
C_{max}	Maximum Serum Concentration
C_{min}	Minimum Serum Concentration
CPMS	Clinical Pharmacology Modelling and Simulation
CR	Complete Response
CRF	Case Report Form
CRM	Continual Reassessment Method
CT	Computed Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DLRT	Dose Level Review Team
DLT	Dose-limiting toxicities
DMP	Data Management Plan
DOR	Duration of response
ECG	Electrocardiogram
GSO-DM	Global Study Operations-Data Management
HLE	Half-Life extended
IV	intravenous
LTFU	Long Term Follow-Up
MABEL	Minimum Anticipated Biological Effect Level
MedDRA	The Medical Dictionary for Regulatory Activities
mg	milligram
MTD	Maximum Tolerated Dose
Claudin 18.2	Claudins are a family of tight junction proteins establishing paracellular barriers which control flow of molecules between cells
NE	Not Evaluable
OS	Overall survival
PD	Pharmacodynamic
PFS	Progression-free survival

Abbreviation or Term	Definition/Explanation
PK	Pharmacokinetics
PR	Partial Response
RECIST	Response Evaluation Criteria in Solid Tumors
RP2D	Recommended phase 2 dose
SAP	Statistical Analysis Plan
SD	Stable Disease
SFU	Safety follow-up
$t_{1/2}$	Half-life
TP	Time-point
TTR	Time To Response
WHO-DRUG	World Health Organization Drug

1. Introduction

AMG 910 is a novel half-life extended (HLE) bispecific T cell engager (BiTE®) molecule designed to direct T cells towards Claudin 18 **isoform 2** (CLDN18.2) - expressing cells. The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol for study 20180292, AMG 910 dated 02 March 2021. The scope of this plan includes the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints/ Estimand(s)

Objectives	Endpoints
Primary	
• To evaluate the safety and tolerability of AMG 910 in adult subjects • To determine the maximum tolerated dose (MTD) and/or recommended phase 2 dose (RP2D)	• Dose-limiting toxicities (DLT) • Treatment-emergent adverse events • Treatment-related adverse events • Changes in vital signs, electrocardiogram (ECG), and clinical laboratory tests

Objectives	Endpoints
Secondary	
• To characterize the PK of AMG 910	• PK parameters for AMG 910 following **short-term** intravenous (IV) **and extended IV (eIV)** administration including but not limited to maximum serum concentration (C_{max}), minimum serum concentration (C_{min}), area under the concentration-time curve (AUC) over the dosing interval, accumulation following multiple dosing, and, if feasible, half-life ($t_{1/2}$)
• To evaluate preliminary anti-tumor activity of AMG 910	• Objective response (OR) per Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 and iRECIST • Duration of response (DOR) • Time to progression

	• Progression-free survival (PFS), 6-month and 1-year PFS • Overall survival (OS), 1 and 2-year OS

Objectives	Endpoints
Exploratory	
[redacted]	

2.2 Hypotheses and/or Estimations

AMG 910 will demonstrate acceptable safety and tolerability in subjects with gastric or GEJ cancer at 1 or more dose levels with at least 1 dose level showing evidence of anti-tumor activity.

3. Study Overview

3.1 Study Design

This is an open-label, ascending, multiple dose, phase 1 study evaluating AMG 910 in subjects with CLDN18.2-positive gastric and GEJ adenocarcinoma. The study will consist of 2 phases:

- Dose-exploration phase
- Dose-expansion phase

3.1.1 Dose Exploration Phase

The dose-exploration phase of the study will estimate the MTD of AMG 910 using a **Bayesian logistic regression model (BLRM)** based on enrollment of 34 gastric cancer subjects. A RP2D may be identified based on emerging safety, efficacy, and pharmacodynamics data prior to reaching an MTD.

AMG 910 will be administered as a short-term IV infusion (approximately 60 minutes) **or as extended IV (eIV) infusion over 96 hours (in cycle 1 week 1) over six 28-day cycles following the doses and dosing schedule outlined in Table 3-1**.

- For cycles 1 and 2, twice a week at days 1 and 3 of each week in a 28-day cycle, **starting with cycle 1 day 8 dose** and

- For cycles 3 to 6, weekly, i.e., days 1, 8, 15, and 22 in a 28-day cycle

Because of the observation of cytokine release syndrome (CRS) events in cohort 1, the dosing of AMG 910 and the dose escalation schedule will be adapted. In cycle 1 week 1, the administration is to start with an approximately 96-hour eIV infusion followed by short-term weekly or twice weekly infusions (approximately 60 minutes) at the planned target doses of AMG 910 of 6.5, 15, 30, 60, 150, 300, 600, 1000, and 2000 µg from cycle 1 week 2 onwards.

For the new dosing schedule with 96-hour eIV infusion of AMG 910, day 1 and day 3 doses of cycle 1 week 1 will be each infused over 48 hours instead of 60 minutes resulting in a total infusion time of 96 hours. It is hypothesized that the eIV infusion approach over a 96-hour period may reduce the intensity and/or frequency of the symptoms associated with CRS associated with the same target dose of AMG 910 when infused twice over a 60-minute duration. The lowest AMG 910 daily dose for eIV administration will be 3.25 µg/day. The calculation of a next higher eIV dose will be based on the following calculation:

Daily eIV dose = total AMG 910 dose, resulting from the twice weekly dosing at the target dose, divided by number of eIV infusion days.

There will be treatment-free intervals of two weeks after cycles 2 and 4. The treatment-free intervals may be extended up to three weeks to allow for recovery in case of adverse events. The treatment-free interval after cycle 4 may also be extended to three weeks for logistical reasons. The maximum treatment duration will be 6 cycles. Subjects who have shown no disease progression and no worsening of performance status at the end of cycle 6 may continue treatment for an additional two cycles if there are no unresolved grade 3 or 4 toxicities and if the investigator and the Amgen Medical Monitor agree that continued treatment may likely provide benefit for the subject.

The DLRT may recommend enrolling more subjects for safety reasons at any point during the study and **the DLRT may explore 2 different AMG 910 dosing schedules in cohorts running in parallel.**

Dose exploration will be conducted in 2 stages: single-subject cohort(s) followed by multiple-subject cohort(s) (3 to 4 subjects per cohort) (Refer: Table 3-1).

Single subject cohort:

Dose exploration will begin with single-subject cohort(s) and the starting dose is expected to have the first dose C_{max} below MABEL. Serum concentrations to be achieved with first 3 to 4 dose level cohorts are expected to be lower than those at which pharmacodynamic activity translating into treatment-related adverse events or efficacy is predicted to be observed. Therefore, the first two cohorts are pre-planned single subject cohorts. If a subject in the first single-subject cohort does not experience a grade 2 or higher adverse event at least possibly related to AMG 910 or DLT during the initial 28 days of treatment (DLT window for single-subject cohorts), then the next subject may be dosed at the next higher dose level(s) per Table 3-1. Planned doses per cohort level ranging from **2.5** µg to **2000** µg. Switching from a single subject cohort to a multiple-subject cohort will occur with cohort 3 or sooner if specific safety criteria are observed as described below during the DLT window:

- DLT or
- Grade ≥ 2 adverse events based on CTCAE version 5.0, at least possibly related to AMG 910, with the exception of lymphopenia or lymphocyte count decreased.

The DLRT will convene to review the safety data and recommend the appropriate dose to be implemented. The recommended dose will not exceed the levels recommended by the BLRM in the dose escalation portion and the next pre-defined dose level. The DLRT may also convene ad hoc meetings any time to review safety data if deemed necessary.

Multiple subject cohort:

Dose escalation in multiple-subject cohorts (3- to 4 subjects per cohort) will begin with single subject cohort and the dose level(s) per Table 3-1. The DLRT will convene to review the safety data and recommend the appropriate dose to be implemented. The recommended dose will not exceed the levels recommended by the BLRM in the dose escalation portion and the next dose level listed in Table 3-1. The DLRT may recommend enrolling more subjects for safety reasons at any point during the study.

For dose escalation with eIV dosing schedule, the DLRT will in addition to the considerations above evaluate the eIV administration and the short-term IV administrations separately for the occurrence of first-dose effects (e.g., CRS and TLS). The DLRT will review the safety data and may recommend to only escalate the dose for either the eIV administration or the short-term infusion administrations for the next cohort. The recommended dose will not exceed the levels recommended by the BLRM in the dose escalation portion and the next dose level

listed in Table 3-1. For subsequent dose escalation decisions, the dose for the eIV administration and the short-term IV administrations may be again escalated in parallel while maintaining the relative dose difference or the dose for the not escalated administration may be kept constant.

For multiple-subject cohorts, the DLRT will not recommend dose escalation decisions until each of the following occurs:

- All subjects in the cohort have been followed for safety events for a minimum of 28 days from the first dose of AMG 910 (except if a subject experienced a DLT) or have been followed for a minimum of 72 hours after the 7th dose with treatment interruption not constituting a DLT
- At least 3 of the subjects can be evaluated for DLT assessment

Determination of whether a subject is evaluable for DLT assessment will be made in accordance with the following rules:

- Subjects who receive study treatment (at least 80% of the planned dose for cycle 1) and remain on study through the DLT assessment window will be considered DLT-evaluable.
- Subjects who discontinue from the study prior to completing the DLT assessment window for reasons other than a DLRT will be considered non-evaluable for dose-escalation decisions and MTD determination and may be replaced by an additional subject at that same dose level.

Dose escalation/de-escalation decisions in multiple-subject cohorts will be guided by the BLRM model of dose toxicity. The MTD for BLRM is the dose level predicted to have the highest probability of a DLT rate within the target interval of 20% to 33%, subject to overdose control. To control the risk of overdose, the MTD must have less than a 40% predicted probability of overdosing (DLT rate > 33%). If late onset adverse events (i.e., after the 28-day DLT period) occur or late findings from endoscopic evaluations during a cohort, the DLRT may adaptively re-consider the doses evaluated within a cohort for subsequent dosing and/or possibly trigger a de-escalation or withholding of additional doses in subsequent cohorts. Based on the BLRM model and after reviewing all available safety and tolerability data, the DLRT may recommend dosing at intermediate dose levels between the planned dose levels listed in Table 3-1.

3.1.2 Dose Expansion Phase

Following the dose-exploration phase, a dose-expansion phase will be conducted on 36 additional subjects approximately to confirm safety, PK, and pharmacodynamics at the MTD or RP2D and to obtain further safety and efficacy data and enable correlative biomarker analysis. The dose to be evaluated will be at or below the MTD estimated in the dose exploration cohorts. A final estimate of the MTD and RP2D will be evaluated and confirmed utilizing all DLT-evaluable subjects from the dose exploration and the dose expansion cohorts. This final estimate of the MTD and RP2D will not exceed the MTD identified in the dose escalation portion of the study and will not exceed the MTD identified in the dose escalation portion of the study by the BLRM. Efficacy data will be reviewed in the dose expansion part of the study after the first 15 subjects are enrolled and have had the opportunity to receive at least 8 weeks of treatment and received the first response assessment (with recruitment ongoing).

A final BLRM estimate of the MTD will use all data from dose exploration and dose expansion.

3.1.3 Extension of Treatment Duration

Subjects who have shown no disease progression and no worsening of performance status at the end of cycle 6 may continue treatment for an additional two cycles if there are no unresolved treatment related grade 3 or 4 toxicities and if the investigator and the Amgen Medical Monitor agree that continued treatment may likely provide benefit for the subject. Further extensions of the treatment are to be agreed in the same manner at the end of each second cycle of extended treatment. Treatment schedule and schedule of activities will be conducted as described for cycle 6.

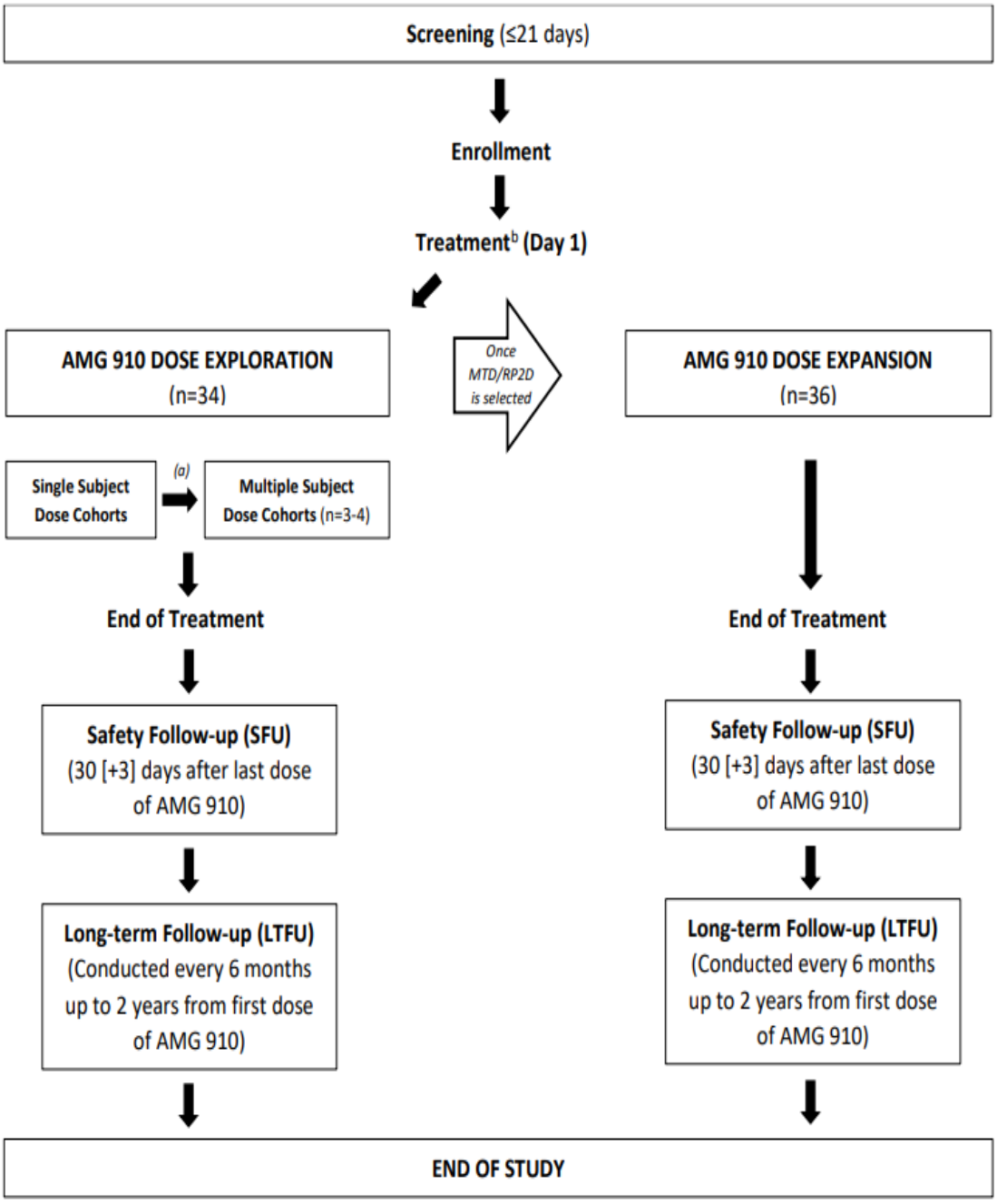
Subjects who remain on the treatment for more than 2 years do not need to have additional long-term follow-up. They will complete the study 30 (+3) days following discontinuing treatment.

3.1.4 Re-treatment of Subjects

Subjects who have derived sustained benefit from treatment, meaning at least stable disease and no worsening of performance status, for at least 3 months after the end of treatment but experience relapse or progression thereafter, and have not received any other systemic anti-cancer therapy since their last AMG 910 treatment, may be retreated in a separate retreatment group **for a total of up to 2 rounds of retreatment (up to 6 cycles each)**. Retreatment will be administered at the highest dose level and currently used schedule that has been assessed as safe by the DLRT at the time of retreatment

initiation. Treatment schedule and schedule of activities will be conducted as described for cycles 1 through 6 in the dose exploration and expansion cohorts.

Figure 3-1. Study Schema



LTFU = long-term follow-up; MTD = maximum tolerated dose; RP2D = recommended phase 2 dose

SFU = safety follow-up

[a] Multiple-subject cohorts will enroll at a dose level where specific safety and/or efficacy criteria were met in a single-subject cohort or per planned dose level from Table 3-1. For subsequent dose level cohorts, 3-4 total subjects will be enrolled.

[b] **For cohorts 1b onwards, AMG 910 is administered as extended IV infusion over 96 hours in cycle 1 week 1 and as short-term IV infusions weekly (or twice weekly) starting cycle 1 day 8 onwards in cycle 1 through 6 of 28-day cycles with scheduled 2-week breaks after cycles 2 and 4.**

Table 3-1. Planned Doses per Dose Cohort Level

Cohort number	Number of subjects[a]	Dose (μg) **Route**
(-1)	**(3 to 4)**	**2.5 IV**
1	**3 to 4**	6.5[b] **IV**
1b	**3 to 4**	**6.5[b] eIV**
2b	**3 to 4**	**15 eIV**
3b	3 to 4	**30 eIV**
4b	3 to 4	**60 eIV**
5b	3 to 4	**150 eIV**
6b	3 to 4	**300 eIV**
7b	3 to 4	**600 eIV**
8b	3 to 4	1000 **eIV**
9b	3 to 4	**2000 eIV**

Routes of administration: eIV = extended intravenous; IV = **short-term** intravenous.

[a] Switching from a single subject cohort to a multiple-subject cohort will occur with cohort 3 or sooner if specific safety criteria are observed as described in Section 6.2.1.1.1 in the protocol.

[b] If a dose reduction is needed according to instructions provided in Table 6-3 as in the protocol, treatment will be continued at a dose of 2.5 μg as specified in Section 6.2.2.3.5 as in the protocol.

Note: The "b" cohort numbers are used to differentiate cohorts 1 through 10 denoted in previous versions of the protocol

3.2 Sample Size

It is anticipated that approximately 70 subjects will be enrolled in this study. Up to 34 subjects will be enrolled in the dose-exploration cohorts and up to 36 additional subjects will be enrolled in the dose-expansion cohort.

The sample size in the dose-escalation phase is based on practical consideration and is consistent with conventional oncology studies with the objective to estimate the MTD. With 3 subjects in a cohort, there is a 27% to 70% probability of observing at least 1 DLT if the true DLT rate is 10% to 33% and with 4 subjects in a cohort, there is a 34% to 80% probability.

In the dose-expansion cohort, a subject number of 36 will provide an 84% probability of observing at least 1 adverse event with 5% incidence rate. An exact 95% binomial CI will be provided for overall response rate. With the 36 subjects and 19% overall response rate, the expected 95% CI would be 8% to 36%.

4. Covariates and Subgroups

4.1 Planned Covariates

Not applicable.

4.2 Subgroups

Not applicable.

5. Definitions

Adverse events

An adverse event **(AE)** is any untoward medical occurrence in a clinical study subject irrespective of a causal relationship with the study treatment.

Age at enrollment

Subject age at enrollment will be determined using the age in years reported in the clinical database.

Area under the Curve (AUC)

The area under the **serum** drug concentration-time curve (AUC) reflects the actual body exposure to drug after administration of a dose of the drug.

Baseline

For any variable, unless otherwise specified, the baseline is the last non-missing assessment taken prior to the first administration of AMG 910.

For parameters/assessments not scheduled to be performed (or scheduled but not performed) on the same day as the first administration of AMG 910, the baseline value is the value from the screening period measured closest to the day of first administration of AMG 910.

Baseline ECG values in triplicate

The mean of values in a triplicate should be calculated for Baseline. Baseline triplicate value is the mean of all pre-dose assessments. For all post- baseline ECG, the mean value for measurements taken at the same assessment will be calculated and used in the analysis.

When an ECG is missing within a triplicate, all available data will be averaged for that time point.

Best Overall Response (BOR) - using RECIST v1.1 criteria

BOR is defined as the best response in the following order: CR, PR, SD, PD, or NE, where CR and PR require confirmation by a repeat consecutive assessment at least 4 weeks after the first detection of radiographical response or progressive disease. Confirmation of SD, minimum 6 weeks duration from the baseline scan is required. The details of derivation of BOR is described in Appendix B.

Complete response (CR) per RECIST v1.1 criteria

Disappearance of all target lesions, non-target lesions and normalization of tumor markers. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to <10 mm.

Objective Response Rate (ORR) per RECIST v1.1 criteria

Objective response rate is defined as the proportion of patients with a BOR of CR or PR.

Partial Response (PR) per RECIST v1.1 criteria

Partial Response (PR) is defined as at least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters. Non target lesions must be non-Progressive Disease (non-PD).

BMI

Body Mass Index should be calculated using the following formula:

$$BMI\ (kg/m^2) = weight\ (kg) / [height\ (cm)/100]^2$$

Change from Baseline

Change from Baseline is the arithmetic difference between post-dose assessments and Baseline value.

Change (absolute) from Baseline = (Post-baseline Value – Baseline Value)

Change (percent) from Baseline = [(Post-baseline Value – Baseline Value) /Baseline Value] x 100

C_{max}

Maximum observed drug concentration.

C_{min}

Minimum observed drug concentration.

Dose Limiting Toxicity

Dose limiting toxicities are defined as any adverse events at least possibly related to AMG 910 with an onset within the first 28 days following first dose for single-subject cohorts and within the safety observation period for multiple-subject cohorts as defined in section 6.2.1.1.3 of the protocol.

Additional information on DLTs are given in the section 6.2.1.1.2 of the protocol.

End of study (End of trial)

It is defined as the date when the last subject across all sites is assessed or receives an intervention for evaluation in the study (i.e., last subject last visit), following any additional parts in the study (e.g., long-term follow-up), as applicable.

Fridericia-corrected QT intervals (QTcF)

The Fridericia correction will be calculated from the investigator reported QT (msec) and RR interval (msec), as follows:

$$QTcF=QT/(RR/1000)^{1/3}$$

Investigational product (IP)

The term 'investigational product' is used in reference to AMG 910.

Last investigational product dose date

The last IP date for each subject is defined as the latest date IP administered.

Long term follow-up (LTFU)

Subjects will be followed for every 6 months up to 2 years from the first dose of AMG 910 for all subjects who have not withdrawn consent by clinic visit, telephone or chart review to assess for survival and/or the commencement of subsequent cancer therapy. **Subjects who remain on the treatment for more than 2 years do not need to have additional long-term follow-up (LTFU). They will complete the study 30 (+3) days following discontinuing treatment.**

Maximum Tolerated Dose (MTD)

A final estimate of the MTD will be made based on a Bayesian Logistic Regression Model (BLRM) utilizing all DLT-evaluable subjects from the dose exploration and dose expansion cohorts. The MTD for BLRM is the dose level predicted to have the highest probability of a DLT rate within the target interval of 20% to 33%, subject to overdose control. To control the risk of overdose, the MTD must have less than a 40% predicted probability of overdosing (DLT rate >33%).

Primary completion

The primary completion date is the date when data for the primary endpoint are last collected for the purposes of conducting the primary analysis. It is the date when the last subject is assessed or receives AMG 910 for the final collection of data for the primary endpoint(s), for the purposes of conducting the primary analysis, whether the study concluded as planned or was terminated early.

Safety follow-up

It is defined as the subject visit at least 30 (+3) days after last dose of AMG 910 is received to assess the risk of delayed adverse events.

Serious adverse event

A Serious Adverse Event is defined as any untoward medical occurrence that, meets at least 1 of the following serious criteria:

- Results in death (fatal).
- Immediately life-threatening.
- Requires in-patient hospitalization or prolongation of existing hospitalization.
- Results in persistent or significant disability/incapacity.
- Other medically important serious event.
- Is a congenital anomaly/birth defect.

Study day

Post-study day: study day= (date – date of Study Day 1) + 1.

Pre-study day: study day= (date – date of Study Day 1).

Study day 1

It is defined as the first day that AMG 910 is administered to the subject.

Treatment emergent adverse event

A treatment - emergent adverse event **(TEAE)** is any adverse event starting on or after the first administration of investigational product, as determined by the flag indicating if the adverse event started prior to the first dose on the Events CRF, and up to and including 30 days after the last IP dose date. The severity of each adverse event will be graded using the CTCAE version 5.0. Adverse events will be coded using MedDRA.

Treatment related adverse event

A treatment-related AE is any treatment-emergent AE that per investigator review has a reasonable possibility of being caused by AMG 910.

6. Analysis Sets

The following sub-sections describe the analysis sets to be used.

6.1 Safety Analysis Set

Safety analysis set defined as all subjects that are enrolled and receive at least 1 dose of AMG 910. As per the protocol, all endpoints, unless noted otherwise, will be conducted on the safety analysis set. Hence, full analysis set, primary analysis set are same as safety analysis set.

6.2 Pharmacokinetic/Pharmacodynamic Analyses Set(s)

The PK Analysis Set will contain all subjects who have received at least 1 dose of the investigational product and have at least 1 PK sample collected. A subject with insufficient number of data points to meet analysis requirements, or where significant protocol deviations have affected the data, or if key dosing or sampling information is missing will not be used to perform PK analysis.

6.3 Study Specific Analyses Set(s)

6.3.1 DLT Analyses Set(s)

Dose limiting toxicity will be performed on the DLT Analysis set. DLT Analysis **set** is defined as all subjects that are enrolled and receive at least 1 dose of AMG 910 with an evaluable DLT endpoint.

The DLT endpoint is evaluable if either: 1) the subject experiences a DLT, or 2) the subject does not experience a DLT within the DLT window in effect for the dose level cohort of the subject.

A non-evaluable subject may be replaced by an additional subject at the same dose level for dose-escalation and MTD determination.

6.3.2 RECIST Evaluable Analyses Set(s)

The analysis of objective response will be conducted on the RECIST Evaluable Analysis Set. It is defined as all subjects that are enrolled and receive at least 1 dose of AMG 910 with at least one tumor lesion that is measurable in contrast-enhanced computed tomography (CT) as defined by RECIST 1.1 **at baseline**.

Objective response will be performed through Tumor burden assessments based on RECIST 1.1 criteria.

7. Planned Analyses

7.1 Primary Analysis

Because the sponsor decided to early terminate this study, primary analysis and final analysis will be combined as one final analysis as part of study close-up activities.

7.2 Final Analysis

The final analysis will occur when target enrollment is complete for both phases and all subjects have ended the study. Data will be subject to on-going checks for integrity, completeness and accuracy in accordance with the Data Management Plan with the expectation that all outstanding data issues are resolved ahead of the snapshot. The data will be locked to prevent further changes, and a snapshot of the locked database will be used in the analysis.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database.

8.3 Handling of Missing and Incomplete Data

Incomplete adverse event, **and death** dates missing data will be imputed as described in Appendix A.

8.4 Detection of Bias

Lack of protocol compliance and the potential for biased statistical analyses will be examined by assessing the incidence of important protocol deviations in each cohort. The clinical study team will identify and document the criteria for important protocol deviations.

8.5 Outliers

Pharmacokinetic (PK) **serum** concentration data will be evaluated for outliers by visual inspection, and decisions to re-assay individual samples will be made in accordance with standard pharmacokinetic evaluation practice.

8.6 Distributional Characteristics

Where appropriate, the assumptions underlying the proposed statistical methodologies will be assessed. If required data transformations or alternative non-parametric methods of analyses will be utilized.

8.7 Validation of Statistical Analyses

Programs will be developed and maintained and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

9. Statistical Methods of Analysis

9.1 General Considerations

Unless otherwise specific, all described analyses will be conducted on the Safety Analysis Set. Descriptive statistics will be provided for selected demographics, safety, PK, pharmacodynamic, and biomarker data by dose, dose schedule, and time as appropriate. Descriptive statistics on continuous data will include means, medians, standard deviations, and ranges, while categorical data will be summarized using frequency counts and percentages. Graphical summaries of the data may also be presented.

Exact method proposed by Clopper-Pearson will be used for the estimation of confidence intervals (CI) for proportion.

9.2 Subject Accountability

The number and percent of subjects who were enrolled, screened but not enrolled (screen failures), enrolled but not treated, received investigational product, discontinued from AMG 910 (including reasons for discontinuing), completed study, discontinued the study (including reasons for discontinuing) will be summarized. Key study dates for the first subject enrolled, last subject enrolled, and last subject's end of study will be presented.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the study. The final IPD list is used to produce the list of Subjects with IPDs. In addition, a separate listing of all inclusion and exclusion deviations will be provided.

9.4 Demographic and Baseline Characteristics

Demographic data collection including sex, age, race, and ethnicity will be collected in order to study their possible association with subject safety and treatment effectiveness. Descriptive statistics for selected demographics, baseline characteristics, safety, PK, PD will be provided by dose, dose schedule, and time as appropriate. Descriptive statistics on

continuous data will include mean, median, standard deviation, and range, while categorical data will be summarized using frequency count and percentage. Graphical summaries of the data may also be presented.

9.5 Efficacy Analyses

9.5.1 Analyses of Primary Efficacy Endpoint(s)/Estimand(s)

No efficacy parameter is considered in primary endpoints.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

The number and percent of subjects with best overall response of complete response, partial response, stable disease, and progressive disease as determined will be tabulated. The proportion of subjects with an objective response (per RECIST 1.1) and 95% CI will be tabulated by planned dose level using the RECIST Evaluable Analysis Set.

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)

Not applicable.

9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoints

Unless otherwise specified, statistical analyses on safety endpoints will be done using subjects from the Safety Analysis Set, which includes subjects that are enrolled and received at least 1 dose of AMG 910.

The analysis of DLTs will be conducted on the DLT Analysis Set. The probability of a subject having a DLT by planned dose level will be estimated using a two-parameter BLRM model.

9.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version [2**5.0**] or later will be used to code all events categorized as adverse events to a system organ class and a preferred term.

Subject incidence of all treatment-emergent adverse events, serious adverse events, adverse events leading to **drug interruption**, withdrawal of investigational product, fatal adverse events and subject incidence of AEs by PT will be listed, summarized and tabulated by system organ class and preferred term in decreasing frequency order.

Summaries of treatment-emergent, serious adverse events will be tabulated by system organ class, preferred term, grade (3 or higher) and dose/cohort.

In addition, separate descriptive tables will be tabulated for all treatment emergent AE, all treatment related AEs, all serious AEs, related serious AEs, treatment emergent AEs

by grade, all fatal AEs, will be tabulated and presented by cohort and dose. **Partial / missing date imputation rules are given in Appendix A.**

9.6.3 Laboratory Test Results

Summaries of the absolute value and/or changes from baseline at each scheduled assessment will be provided for selected laboratory parameters. Shifts in grades of safety laboratory values from baseline for selected laboratory values will also be provided.

9.6.4 Vital Signs

Summary statistics of systolic/diastolic blood pressure, heart rate, respiratory rate, and temperature will be provided for baseline and each scheduled post-baseline assessment. Depending on the size and scope of changes, summaries of changes from baseline over time may be provided.

9.6.5 Physical Measurements

Summary of physical measurement data will be provided for weight, height, BMI.

9.6.6 Electrocardiogram

Summaries over time and/or changes from baseline over time will be provided for all ECG parameters. Subjects' maximum change from baseline in QTc will be categorized and the number and percentage of subjects in each cohort will be summarized. Subjects' maximum post baseline values will also be categorized and the number and percentage of subjects in each cohort will be summarized.

[redacted]

9.6.8 Exposure to Investigational Product

Investigational product exposure will be summarized for number of cycles started, number of doses, cumulative dose and average dose administered. Summaries of the of subjects with dose modifications, interruptions and reasons for modifications, interruptions will be provided. In addition, a listing of the unique manufacturing lot numbers, and a listing of the subjects administered each manufacturing lot number will be provided.

9.6.9 Exposure to Non-Investigational Product

There are no non-investigational products in this study.

9.7 Other Analyses

For AMG 910, **PK** parameters, including but not limited to C_{max}, minimum serum concentration (C_{min}), area under the concentration-time curve (AUC) over the dosing interval, accumulation following multiple dosing, and, if feasible, half-life ($t_{1/2}$), will be determined from the time concentration profile using standard non-compartmental approaches and considering the profile over the complete sampling interval. Based on the review of the data, analyses to describe the relationship between AMG 910 exposure and either pharmacodynamic effect and/or clinical outcome may also be performed.

9.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

Descriptive statistics for pharmacokinetic parameters, including but not limited to maximum serum concentration (C_{max}), minimum serum concentration (C_{min}), area under the concentration-time curve (AUC) over the dosing interval, accumulation following multiple dosing, and, if feasible, half-life ($t_{1/2}$), will be provided from the time concentration profile using standard non-compartmental approaches and considering the profile over the complete sampling interval. Individual concentration-time data will be tabulated and presented graphically.

Nominal sampling times will be used for individual concentration-time plots and tables. Actual dose administered, and actual sampling times will be used for the calculation of PK parameters for each subject. The reasons for excluding any sample from the analyses will be provided.

Summary of PK concentration over time and PK parameters will be provided. Mean concentration-time profiles for each dose will be provided. Based on the review of the data, analyses to describe the relationship between AMG 910 exposure and either pharmacodynamic effect and/or clinical outcome may also be performed. Above analyses will be conducted by Amgen Clinical Pharmacology Modeling and Simulation (CPMS).

10. Changes From Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

11. References

Eisenhauer EA, P Therasse, J Bogaerts, LH Schwartz, D Sargent, R Ford, et al., 2009, New Response Evaluation Criteria in Solid Tumours: Revised RECIST guideline (version 1.1), Eur J Cancer, 45(2):228-247.

Seymour L, Bogaerts J, Perrone A, et al, on behalf of the RECIST working group. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol 2017; 18: e143–52

Fuchs, C. S., Doi, T., Jang, R. W., Muro, K., Satoh, T., Machado, M., Garrido, M. (2018). Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. *JAMA oncology, 4*(5), e180013-e180013.

Bang, Y. J., Van Cutsem, E., Feyereislova, A., Chung, H. C., Shen, L., Sawaki, A., Aprile, G. (2010). Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *The lancet, 376*(9742), 687-697.

Bergstrom, K. S., & Xia, L. (2013). Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology, 23*(9), 1026-1037.

Brookmeyer, R., & Crowley, J. (1982). A confidence interval for the median survival time. *Biometrics*, 29-41.

Saber, H., Del Valle, P., Ricks, T. K., & Leighton, J. K. (2017). An FDA oncology analysis of CD3 bispecific constructs and first-in-human dose selection. Regulatory toxicology and pharmacology, 90, 144-152.

12. Prioritization of Analyses

No analysis is prioritized.

13. Data Not Covered by This Plan

Not applicable.

14. Appendices

14.1 Appendix A. Handling of Dates, Incomplete Dates and Missing Dates

Imputation Rules for Partial or Missing Start Dates

The reference date for the following rules is the date of first dose of AMG 910.

Start Date		Stop Date						
		Complete: *yyyymmdd*		Partial: *yyyymm*		Partial: *yyyy*		Missing
		< 1st dose	≥ 1st dose	< 1st dose *yyyymm*	≥ 1st dose *yyyymm*	< 1st dose *yyyy*	≥ 1st dose *yyyy*	
Partial: *yyyymm*	= 1st dose *yyyymm*	2	1	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	2	2	2	2	2
Partial: *yyyy*	= 1st dose *yyyy*	3	1	3	1	n/a	1	1
	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1=Impute the date of first dose; 2=Impute the first of the month; 3=Impute January 1 of the year; 4=Impute January 1 of the stop year
Note: For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the partial month or first day of year if month is also missing.

Imputation Rules for Partial or Missing Stop Dates

Initial imputation

- If the month and year are present, impute the last day of that month.
- If only the year is present, impute December 31 of that year.
- If the stop date is entirely missing, assume the event or medication is ongoing.

If the imputed stop date is before the start date, set stop date to missing.

If the imputed stop date is after the death date, impute as death date.

Imputation Rules for Partial or Missing Death Dates

If death year and month are available but day is missing:

- If yyyymm for the date last known to be alive equals yyyymm for death date, set death date to the day after the date last known to be alive.
- If yyyymm for the date last known to be alive is less than the yyyymm for death date, set death date to the first day of the death month.

- If yyyymm for the date last known to be alive is greater than yyyymm for death date, assume death date is in error, do not impute and censor the subject survival time.

If month and day are missing and year of death is known:

- If yyyy for the date last known to be alive equals yyyy for death date, set death date to the day after last known to be alive date.
- If yyyy for the date last known to be alive is less than yyyy for death date, set death date to the first day of the death year.
- If yyyy for the date last known to be alive is greater than yyyy for death date, assume death date is in error, do not impute and censor the subject survival time.

If a death date is totally missing:

Do not impute and censor the subject survival time.

14.2 Appendix B. RECIST (v1.1) Criteria

BOR will need to be derived based on timepoint response collected on RECIST 1.1 CRF. Table 14-1 provides the BOR determination per RECIST 1.1 for trials where response confirmation is required. That means in the case of confirming CR or PR, a confirmation is required by a repeat assessment at least 4 weeks after the first detection and assessments of NE may occur between timepoint T1 and timepoint T2. For example, BL, CR, NE, NE, NE, CR – then CR at post baseline 1 is confirmed at post baseline 5. A BOR determined by Table 14-1 is considered a Confirmed_BOR. At interim analysis, an unconfirmed response rate may be reported in addition to the confirmed response rate. The unconfirmed rate includes subjects who achieved a PR or CR and are awaiting a confirmative scan (i.e., unconfirmed responders). If these subjects have disease progression or death before the confirmative scan, then they are no longer considered as unconfirmed responder awaiting confirmative scan. Timepoint response "NON-CR/NON-PD" will be considered as "SD" in the algorithm of confirmation. Interim_BOR is defined to include these subjects in addition to Confirmed_BOR. Table 14-2 outlines the steps to derive Confirmed_BOR and Interim_BOR given the timepoint assessments.

Table 14-1. BOR per RECIST 1.1

Criterion	Timepoint T1 Response	T1 ≥ 35* days after Baseline?	Timepoint T2 Response	T2 ≥ 35* days after Baseline?	T2 ≥ 28 days after T1?	BOR at time point T1
C1	CR	Yes	CR	-	Yes	CR
C2			CR	-	No	SD
C3			PR, SD	-	-	Query data**
C4			PD	-	-	SD
C5			NE, No further evaluations			SD
C6		No	CR	-	Yes	CR

C7			CR	Yes	No	SD
C8			PR, SD	-	-	Query data**
C9			PD	-	-	PD
C10			NE, No further evaluations			NE
C11	PR	Yes	CR, PR	-	Yes	PR
C12			CR, PR	-	No	SD
C13			SD	-	-	SD
C14			PD	-	-	SD
C15			NE, No further evaluations			SD
C16		No	CR, PR	-	Yes	PR
C17			CR, PR	Yes	No	SD
C18			SD	Yes	-	SD
C19			PD	-	-	PD
C20			NE, No further evaluations			NE
C21	SD	Yes	CR, PR, SD, PD, NE, no more evaluation			SD
C22		No	CR, PR, SD	Yes	-	SD
C23			CR, PR, SD	No	-	NE
C24			PD	-	-	PD
C25			NE, No further evaluations			NE
C26	PD		-			PD
C27	NE	-	NE, No further evaluations			NE
C28		-	CR, PR, SD	Yes	-	SD
C29		-	CR, PR, SD	No	-	NE
C30		-	PD	-	-	PD

CR= complete response, PR= partial response, SD= stable disease, PD= progressive disease, and NE= not evaluable.

*CR and PR need to be confirmed, which requires a repeat assessment at least 4 weeks after the first detection. Notice an unlimited number of intermittent assessments of NE can occur between the initial response and the confirmation. For example, CR, NE, NE, NE, CR – then CR at post baseline 1 is confirmed at post baseline 5. Both first detection and repeat assessment can be unscheduled visit and the 4 weeks apart rule needs to be met. For SD, a duration of at least 6 weeks after baseline scan is required. The threshold of **35** days accounts for the earliest available assessment by RECIST 1.1, as for tumor assessment per scheduling of imaging assessment.

**If a CR is truly met at timepoint T1, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes progressive disease at that time point (i.e., disease must have re-appeared after CR). Best overall response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at timepoint T1. Under this circumstances, the original CR should be changed to PR and the best response is PR.

Table 14-2. BOR Derivation Steps for GSP

Step 1. Derive Confirmed_BOR at each visit
Derive *Confirmed_BOR* at each visit: i) At Visit 1: Using the Visit 1 scan result and refer to Table 14-1 to derive *Confirmed_BOR*. ii) At Visit 2 onward:

(a) If current visit is CR or PR: Find last scan that is not NE or SD and at least 28 days before current visit. Derive *BOR_temp* as below:

Last scan	Current	*BOR_temp*
CR	CR	CR
PR	CR	PR
PR	PR	PR

(b) If none of above fits, find last scan that is not NE, reference Table 14-1 to derive *BOR_temp*.

iii) Current visit Confirmed_BOR = best of (BOR_temp, last visit confirmed_BOR). Use rule CR > PR > SD > PD > NE

Step 2. Derive Interim_BOR at last visit prior to analysis data cutoff

For subjects who discontinue tumor assessment (ie, had PD, death, next therapy, withdraw consent, or end of study):	For subjects with potential for more assessment (ie, no PD, death, next therapy, withdraw consent, or end of study), unconfirmed PR/CR can be considered responders.
Assign *Interim_BOR* = *Confirmed_BOR*	• If any scan is CR and no PD/death, then *Interim_BOR* = CR. • If any scan is PR and no CR or PD/death, then *Interim_BOR* = PR. • Else, *Interim_BOR* = *Confirmed_BOR* at the latest scan.