Statistical Analysis Plan

Protocol Title:	A Phase 1b Study Evaluating the Safety, Tolerability, Pharmacokinetics, and Efficacy of AMG 160 in Subjects with Non-Small Cell Lung Cancer
Short Protocol Title:	A Phase 1b Study of AMG 160 in Subjects with Non-Small Cell Lung Cancer
Protocol Number:	20180273
NCT Number:	NCT04822298
Authors:	[redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive Thousand Oaks, CA 91320, USA
SAP Date:	**Document Version** — **Date** Original (v1.0) — 21 September 2021 **Amendment (v2.0)** — **11 October 2022**

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	21SEP2021	Original SAP
Amendment (v2.0)	**11OCT2022**	**Changes related to protocol amendment 3 dated 19 October 2021 and protocol amendment 4 dated 25 March 2022.** **Updated the Definitions for TEAE, TRAE, Clinical progression, time to radiographic progression, subsequent therapy, Baseline triplicate, .** **Added Recist 1.1 evaluable analysis set.** **Updated Appendix C, added BOR derivation, added interim analysis response definitions.** **Updated Appendix B and Appendix D** **Added Appendix E.**

Table of Contents

List of Table

List of Figures

List of Appendices

[redacted]

List of Abbreviations

Abbreviation or Term	Definition/Explanation
AUC	area under the concentration-time curve
C_{max}	maximum serum concentration
C_{min}	minimum serum concentration
CI	confidence intervals
CR	complete response
CRS	cytokine release syndrome
CTCAE	Common Terminology Criteria for Adverse Events
CT	computed tomography
DLRM	dose level review meeting
DLRT	dose level review team
DLT	dose-limiting toxicities
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EOI	end of investigational product
FIH	first in human
^{68}Ga	^{68}Gallium
ICANS	immune-effector cell associated neurologic syndrome
IV	Intravenous
mCRPC	metastatic castration-resistant prostate cancer
MRI	magnetic resonance imaging
MTD	maximum tolerated dose
mTPI	modified toxicity probability interval
NSCLC	non-small cell lung cancer
ORR	objective response rate
OS	overall survival
[redacted]	
PD	disease progression
PD1	programmed cell death protein 1
PDL1	programmed death-ligand 1
PET	positron emission tomography
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	partial response

Abbreviation or Term	Definition/Explanation
PSMA	prostate-specific membrane antigen
RECIST	Response Evaluation Criteria in Solid Tumors
RP2D	recommended phase 2 dose
SAP	statistical analysis plan
SC	subcutaneous
SD	stable disease
t1/2	half-life
TPI	toxicity probability intervals

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol amendment 4 for study 20180273, AMG 160 dated 25 March 2022. The scope of this plan includes the interim analysis, the primary analysis and 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/Estimands

• Objectives	• Endpoints
• Primary	
• To evaluate the safety and tolerability of AMG 160 • To evaluate the maximum tolerated dose (MTD) or the recommended phase 2 dose (RP2D)	• Dose-limiting toxicities (DLT) • Treatment-emergent adverse events • Treatment-related adverse events • Changes in vital signs, electrocardiograms (ECGs), and clinical laboratory tests
• Secondary	
• To characterize the pharmacokinetics (PK) of AMG 160	• PK parameters for AMG 160 following intravenous (IV) administration including but not limited to, maximum serum concentration (Cmax), minimum serum concentration (Cmin), area under the concentration-time curve (AUC) over the dosing interval, accumulation, and half-life (t1/2)
• To evaluate the preliminary anti-tumor activity of AMG 160	• Objective response (OR) per modified Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 • Overall Survival • Progression-free survival (radiographic, clinical) • Time to response • Time to progression (radiographic, clinical) • Duration of response • Time to subsequent therapy

Exploratory

[redacted]

2.2 Hypotheses and/or Estimations

No formal statistical hypotheses will be tested.

3. Study Overview

3.1 Study Design

This is an open label phase 1b study evaluating the safety, tolerability, pharmacokinetics (PK), and efficacy of AMG 160 monotherapy in subjects with relapsed, refractory non-small cell lung cancer (NSCLC). There are 2 parts in the study,

- Dose-exploration
- Dose-expansion

Dose-exploration

The dose-exploration part of the study will estimate the MTD/RP2D of AMG 160 in NSCLC where the AMG 160 monotherapy will be administered as a short-term IV infusion every 2 weeks after target dose is reached in a 28-day cycle as monotherapy therapy. The starting dose of AMG 160 will be 1 dose level below the RP2D or otherwise the highest dose level deemed safe and tolerable as determined in the ongoing first in human (FIH) study of AMG 160 (Study 20180101) in subjects with mCRPC.

To mitigate the risk of CRS [redacted] Based on the MTD or RP2D and the associated dosing schedule

selected in Study 20180101, one of the following [redacted] schedules will be implemented: [redacted] infusion (see Table 6-1 in protocol for further details).

Dose Level Determination

Dose exploration will begin with 2 to 4 subjects. The study DLT window is 28 days. To be DLT evaluable, subjects must receive all planned cycle 1 doses of AMG 160. After a minimum of 2 subjects enrolled at a certain dose level are DLT evaluable, a DLRT meeting will be convened. Depending on observed safety data, the following may occur:

1. additional enrollment to dose Level 1; or
2. dose de-escalation to 1 dose below the starting dose; or
3. alternative dosing schedules, including further dose escalation, as per the ongoing mCRPC trial (Study 20180101), may be explored based on emerging safety and PK data.

Rules for dose-escalation/de-escalation are derived using a modified toxicity probability interval (mTPI) model with a target toxicity probability of 0.28.

Dose level decisions are made based on this posterior probability of toxicity, using 3 toxicity probability intervals (TPI):

- under-dosing TPI: DLT rate from 0 to $<$ 23%
- target TPI: DLT rate from 23% to 33%, inclusive
- over-dosing: DLT rate $>$ 33%

For the current dose level and after adjusting for the width of the under-dosing TPI, if the DLT rate is most likely in the under-dosing TPI then the recommendation is to dose escalate or stay at the current level. If the DLT rate is most likely in the target TPI then the recommendation is to stay at the current level. If the DLT rate is most likely in the over-dosing TPI then the recommendation is to de-escalate. If DLTs are observed, the MTD is the dose level with a DLT toxicity rate closest to 28%. The mTPI Trial monitoring table is described in Table 3.1

Table 3.1 . mTPI Trial Monitoring Table

Number of Subjects at current dose

# of DLTs	2	3	4	5	6	7	8	9	10
0	E	E	E	E	E	E	E	E	E
1	S	S	S	S	S	E	E	E	E
2	DU	D	S	S	S	S	S	S	S
3		DU	DU	D	D	S	S	S	S
4			DU	DU	DU	DU	D	S	S
5				DU	DU	DU	DU	DU	D
6					DU	DU	DU	DU	DU
7						DU	DU	DU	DU
8							DU	DU	DU
9								DU	DU
10									DU

E: Escalate to the next higher dose, if feasible; Otherwise, stay at the same dose; S: Stay at the same dose; D: De-escalate to the previous lower dose; DU: De-escalate to the previous lower dose and the current dose will never be used again in the trial

Dose Expansion

If 1 or more objective responses (OR) are observed (per modified RECIST 1.1 criteria) in dose exploration, additional subjects (up to 40) will be enrolled in dose expansion of the study at the MTD/RP2D identified in dose exploration. In dose expansion, 2 cohorts will be opened:

- Cohort 1 (up to 30 subjects with non-squamous NSCLC) and
- Cohort 2 (up to 10 subjects with squamous NSCLC).

Based on emerging safety and tolerability data in Cohort 2, the dose of AMG 160 could be lowered 1 dose level below the starting dose identified in dose exploration if necessary. Dose modification may occur as a result of interim analysis as described in Section 7.1.

[redacted]

3.2 Sample Size

Up to 50 subjects will be enrolled in the study (approximately 10 subjects in a dose exploration phase and up to 40 subjects in a dose expansion phase). **Additional subjects (up to 20) may be enrolled in 1 or more monotherapy dose levels that**

have been shown to be safe and tolerable (defined as backfill enrollment). This backfill enrollment may be done to better estimate the RP2D and better characterize the safety, efficacy, PK, and [redacted] of AMG 160 monotherapy and may be concurrent with dose escalation to identify the MTD.

An initial 2 to 4 subjects will be enrolled and treated with Dose Level 1 (1 dose level below RP2D of AMG 160 Study 20180101). Rules for dose-escalation/de-escalation are derived using an mTPI model with a target toxicity probability of 0.28**.** If the dose is not tolerated (based on target DLT rate > 28%), the dose level will be reduced 1 level below the starting dose. Alternative dosing schedules, including further dose escalation as per the ongoing mCRPC trial (Study 20180101), may be explored based on emerging safety and PK data. The sample size in the dose exploration phase is based on practical consideration and is consistent with conventional oncology studies with the objective to estimate the MTD.

With 2 subjects in a cohort, there is a 19% to 55% 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. With 6 subjects in a cohort, there is a 47% to 91% probability of observing at least 1 DLT if the true DLT rate is 10% to 33% and with 10 subjects in a cohort, there is a 65% to 98% probability.

In the dose expansion phase, a subject number of 40 will provide a 87% 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 40 subjects and 20% overall response rate, the expected 95% CI would be 9% to 36%.

4. Covariates and Subgroups

4.1 Planned Covariates

The relationship between covariates and efficacy endpoints may be explored if appropriate **and will be finalized in the Statistical Analysis plan before the database lock.**

[redacted]

5. Definitions

5.1 Study Endpoints

5.1.1 Primary Endpoints Definitions

Dose-limiting Toxicities (DLTs):

Investigators determine whether a TEAE qualifies as a DLT per the definition described in protocol section 6.2.1.3. For an adverse event to qualify as DLT, the start date must be within 28 days from the date of first dose of AMG 160.

Fridericia – corrected QT Interval (QTcF):

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

(QTcF=QT/(RR/1000)^0.33)

Treatment-Emergent Adverse Event (TEAE):

Events categorized as Adverse Events (AEs) starting on or after first dose of investigational product (AMG 160) as determined by the flag indicating if the AE started prior to the first dose on the Events eCRF and including 30 days after the last dose of investigational product (AMG 160) or End of Study date or start of new anti-cancer therapy, whichever is earlier.

Treatment-Related Adverse Event (TRAE):

A treatment-related adverse event is any treatment-emergent adverse event that per investigator review has a reasonable possibility of being caused by the investigational product (AMG 160) determined by the flag indicating an event may have been caused by the investigational product (AMG 160) on the events eCRF. In the unlikely event that the flag is missing, the TEAE will be considered treatment-related.

5.1.2 Secondary Endpoints Definitions

Objective Response (OR) per RECIST 1.1:

Objective Response is defined as a best overall response (Appendix C) of either Complete Response (CR) or Partial Response (PR) per RECIST 1.1. Radiographic response (Complete Response, Partial Response) requires confirmation by a repeat, **evaluable** scan at least 28 days after the first documentation of response.

Subjects who do not experience a PR/CR or do not have any follow-up tumor assessments will be regarded as non-responders.

Overall Survival (OS):

Overall survival is defined as the time from the date of Study Day 1 until death due to any cause:

OS time (Days) = date of death - Study Day 1 + 1.

Any subject not known to have died at the time of analysis will be censored based on the last recorded date on which the subject was alive.

Duration of Response (DOR):

The duration of response is defined as time from first evidence of CR or PR to disease progression or death due to any cause, whichever occurs first. DOR will be calculated only for subjects who achieve a best overall response of PR or CR.

DOR (months) = (PD / death date - response start date + 1) ×12 / 365.25

The time of the initial response will be defined as the earliest of the dates contributing towards the first visit response of PR or CR.

- For confirmed responders whose PR or CR are confirmed in two **evaluable** assessment**s**, the first PR/CR assessment among the two **evaluable** assessment**s** will be used to calculate DOR.
- For unconfirmed responders awaiting follow-up confirmation scans, the PR or CR assessment that is awaiting confirmation scan will be used to calculate DOR.

Censoring rules for Duration of Response (DOR):

- Data for responders who are alive and without disease progression or new anti-cancer therapy are censored at the time of last **evaluable** tumor assessment **by CT/MRI scan**.
- Data for responders who are alive and without disease progression but who start new anti-cancer therapy are censored at the time of last **evaluable** tumor assessment **by CT/MRI scan** before the start of new anti-cancer therapy.
- Data for responders who die or have disease progression after two or more missed evaluable tumor assessments are censored at the time of last evaluable tumor assessments **by CT/MRI scan** before the two or more missed evaluable tumor assessments.

Details about the censoring rules are included in Appendix E.

Radiographic Progression:

Disease progression per RECIST 1.1.occurs.

Radiographic Progression-free Survival (radiographic PFS):

Radiographic PFS is defined as the interval from Study Day 1 to the earlier of a radiographic progression or death from any cause; otherwise **PFS is censored based on criteria in Appendix E.**

Clinical Progression:

Clinical disease progression is based on clinical symptoms and are not associated with radiologic progression.

This is recorded by the field 'disease progression is equal to clinical progression' on the EOIP eCRF page.

Details about the censoring rules are included in Appendix E.

Clinical Progression-free Survival (clinical PFS):

Clinical PFS is the time from Study Day 1 to clinical disease progression or death from any cause.

Details about the censoring rules are included in Appendix E.

Time to Response (TTR):

TTR is defined as the time from the date of Study Day 1 to the date of the first documented CR or PR that is subsequently confirmed per RECIST1.1. For a subject who did not respond, time to response will be censored at the latest evaluable radiological assessment date. If a subject has no tumor evaluation in the study, time to response will be censored at the date of Study Day 1.

The time of the initial response will be defined as the earliest of the dates contributing towards the first visit response of PR or CR.

- For confirmed responders whose PR or CR are confirmed in two **evaluable** assessment**s**, the first PR/CR assessment among the two **evaluable** assessment**s** will be used to calculate TTR.
- For unconfirmed responders awaiting follow-up confirmation scans, the PR or CR assessment that is awaiting confirmation scan will be used to calculate TTR.

Time to Radiographic Progression:

Time to radiographic progression is defined as the interval from Study Day 1 to radiographic progression; otherwise, time to radiographic progression is censored based on the criteria in Appendix E.

Time to Clinical Progression:

Time to clinical progression is defined as the interval from Study Day 1 to clinical progression; otherwise, time to clinical progression is **based on the criteriain Appendix E.**

Subsequent Therapy:

Subsequent therapy is defined as subjects receiving anti-cancer therapies intended to treat NSCLC or any other anti-cancer therapies prior to end of study.

Time to Subsequent Therapy:

The time from Study Day 1 to the time a patient starts/receives the subsequent cancer therapy/subsequent therapy; otherwise time to subsequent therapy is censored at the last known date of any of the study assessment prior to initiating the subsequent cancer therapy/subsequent therapy.

5.2 Study Time Points

Enrollment Date:

A subject is considered as enrolled when the investigator decides that the subject has met all eligibility criteria. The investigator should document this decision and date in the subject's medical record and in the enrollment case report form.

Study Day 1:

Study Day 1 is defined as the date of the first dose of the AMG 160 administered to the subject.

End of Study for Individual Subject:

End of study for individual subject is defined as the date of the subject last completed a protocol-specified procedure. The date will be recorded on the End of Study eCRF page.

End of Treatment:

End of treatment is defined as the date the decision was made to end investigational product (AMG 160) as recorded on the End of Investigational Page eCRF page.

Safety follow-up Visit:

Safety follow-up visit is defined as upon permanent discontinuation from the study treatment for any reason, a safety follow-up visit will be performed approximately 30 (+3) days after the end of the last dose of AMG 160, or prior to initiation of other therapy, whichever occurs first.

Long-term follow-up Visit:

Long-term follow-up visit will be conducted every 6 months up to 3 years from the first dose of AMG 160 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.

5.3 Demographics and Baseline Related Definitions

Age at enrollment:

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

Baseline:

The baseline is defined as the last non-missing value on or prior to the pre-dose assessments of AMG 160 on cycle 1 day 1.

Baseline for Triplicate ECG:

Baseline for ECG is defined as mean of the 3 triplicate ECG results. If fewer than 3 triplicate ECG results are available, the mean of available triplicate should be calculated. For all post-baseline ECG, the mean of one triplicate ECG results 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 timepoint.

Triplicate ECGs (to be only performed if the dose being evaluated is higher than the MTD established in subjects with mCRPC [Study 20180101])".

Change from Baseline:

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

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

Investigational Product:

The term investigational product is used in reference of AMG 160 (acapatamab).

Percent Change from Baseline:

Percent change from baseline is the arithmetic difference between post-baseline and baseline divided by baseline values times 100.

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

6. Analysis Sets

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

6.1 Safety Analysis Set

The analysis of all endpoints, unless noted otherwise, will be conducted on the Safety Analysis Set defined as all subjects that are enrolled and receive at least 1 dose of AMG 160.

6.2 RECIST 1.1 Evaluable Analysis Set

RECIST 1.1 Evaluable Analysis Set is defined as all subjects that are enrolled, receive at least 1 dose of investigational product (AMG 160) and have measurable baseline disease per RECIST 1.1 per protocol and have the opportunity to be followed for at least 9 weeks starting from Study Day 1. Subjects who stopped disease assessment prior to 9 weeks will be included in this analysis set if the data cutoff is at least 9 weeks after their Study Day 1.

6.3 DLT Analysis Set

The analysis of DLT will be conducted on the DLT Analysis Set defined as all subjects that are enrolled and receive at least 1 dose of AMG 160 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 after receiving all planned doses within the 28-day DLT window in cycle 1.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

Safety data will be reviewed on an ongoing basis.

During dose exploration and formally during dose level review meetings (DLRMs), Amgen, in consultation with the site investigators, will review all available cumulative data by cohort prior to making dose escalation or dose de-escalation recommendations.

Adverse events and DLTs observed in all subjects will be evaluated continually and fully integrated into all DLRMs and considered in all enrollment and dosing decisions. During dose expansion, Amgen will conduct evaluations of the ongoing grade 4 or higher treatment-related adverse event rate to assess if the threshold for possible early trial termination has been reached. If this threshold is met, enrollment to dose expansion will be halted pending review of safety data by the DLRT. After receiving the DLRT recommendation, Amgen will choose to take one of the following actions.

1) Terminate the trial
2) Amend the protocol to potentially improve the benefit/risk for subjects (eg, increase safety monitoring, modify dose/schedule, mandate [redacted])
3) Continue dose expansion without any changes

The stopping rules use a Bayesian approach proposed by Thall et al, 1995 to terminate the study if the posterior probability that the grade 4 or higher treatment-related adverse event rate is greater than 20% is > 80%. The stopping boundaries assume a prior distribution of Beta (0.40, 1.60) are presented in Table 7-1 and the operating characteristics with pre-specified batch size of 10 new subjects per batch are presented in Table 7-2. The operating characteristics in Table 7-2 provide the probability of stopping the trial early for given hypothetical true rate of grade 4 or higher treatment-related adverse events, whereas the stopping criteria in Table 7-2 are based on situations where the empirical evidence would result in a posterior probability of ≥ 80% that the true grade 4 or higher treatment-related adverse event rate is ≥ 20%.

Table 7-1. Stopping Boundary for Dose Expansion With Posterior Probability of 80% and Grade 4 or Higher Treatment-related Adverse Event Limit of 20%

Number of Subjects	Stop Study if Observing These Many Grade 4 or Higher Treatment-related Adverse Events
10	≥ 4
20	≥ 6
30	≥ 9
40	Dose Expansion Complete

Table 7-2. Operating Characteristics With Batch Size of 10 Subjects

True Grade 4 or Higher Treatment-related Adverse Event Rate	Probability of Early Stopping of Dose Expansion	Average Dose Expansion Sample Size
0.10	2.0%	39.5
0.15	9.1%	37.6

0.20	23.2%	33.9
0.25	42.2%	28.8
0.30	61.6%	23.4

A formal interim analysis of available safety and efficacy data will occur when the first 10 evaluable subjects enrolled have had the opportunity to complete 4 months on study. This interim analysis will estimate MTD, support the determination of RP2D, conduct futility analysis of objective response rate (ORR), and support the evaluation of benefit/risk profile of AMG 160 as a monotherapy.

When deemed necessary, the scope of this interim analysis may include other endpoints mentioned in Section 2.1 if applicable. However, analysis of DOR and time-to-event endpoints will not be conducted unless there are more than 10 events.

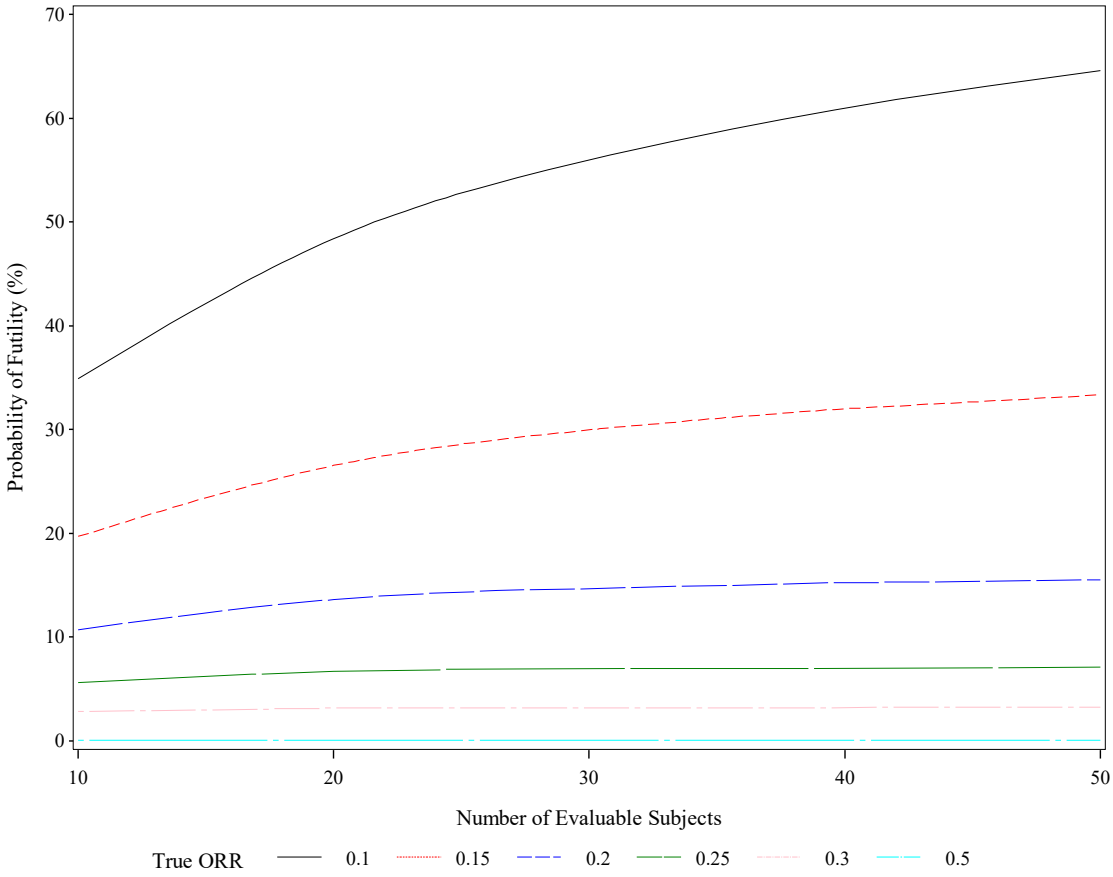
A futility analysis of ORR will be performed using Bayesian predictive probability (Lee and Liu, 2008) when the first 10 evaluable subjects reach month 4. If the predictive probability of observing ORR response > 15% at the end of the study with 50 evaluable subjects is less than 30%, the study may stop early (pT = 0.25). Equivalently, if we observe ≥ 1 objective responder in the first 10 subjects, an additional subjects (up to 40) will be enrolled into the expansion cohort. The predictive probability is calculated with a noninformative prior, Jeffreys prior, Beta(0.5, 0.5).

Similarly, futility analysis will also be conducted during the expansion cohort. The stopping boundaries are presented in Table 7-3. The operating characteristics in Figure 7-1 provide the probability of stopping the trial early for given hypothetical true rate of ORR.

Table 7-3. Objective Response Rate Stopping Boundary

Number of Evaluable Subjects Each Month 4	Stop Study if Observed Objective Response Rate (ORR) Response
10	< 1
20	< 2
30	< 3
40	< 4
50	Study Complete

Figure 7-1. The Probability of Stopping The Trial Early



Data will be subject to ongoing checks for integrity, completeness and accuracy in accordance with the Data Management Plan, however there is no requirement to resolve outstanding data issues ahead of the snapshot. The data will be taken “as-is” on the day of snapshot for futility analysis. No locking is performed and the “as-is” snapshot of the database will be used in the futility analysis. Data will be subject to ongoing 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 for the formal interim analysis. The data will be locked to prevent further changes, and a snapshot of the locked database will be used in the formal interim analysis.

7.1.1 Stopping Rules for Treatment With Siltuximab

Amgen will conduct evaluations of the treatment and outcome of the CRS events treated with siltuximab on an ongoing basis to assess if the threshold for pausing siltuximab treatment has been reached as outlined in the table below. If these stopping rules are met, an ad hoc DLRM will be triggered to review safety data and available PK, [redacted], and efficacy data. If recommended by DLRT, the use of siltuximab will resume. The stopping rules to trigger an adhoc DLRM to

review siltuximab treatment use a Bayesian approach proposed by Thall et al, 1995; an adhoc DLRM will be triggered if the posterior probability that the CRS progression to grade 3 rate is greater than 30% is > 80% or the posterior probability that the CRS progression to grade 4 rate is greater than 10% is > 80%; or observation of any grade 5 CRS after the event has been treated with siltuximab. The stopping boundaries presented below assume a prior distribution of Beta (0.6, 1.4) for progression to grade 3 CRS and a prior distribution frequently if necessary to address emerging safety concerns. If the triggered adhoc DLRM coincides with the regular DLRM, they may be combined.

Number of subjects treated with siltuximab	Trigger DLRM if severity of any CRS event treated with siltuximab progresses to Grade 5	
	Or this number of subjects with severity of CRS progressed to Grade 3 after being treated with siltuximab	Or this number of subjects with severity of CRS progressed to Grade 4 after being treated with siltuximab
5	≥ 3	≥ 2
10	≥ 5	≥ 3
15	≥ 7	≥ 3
20	≥ 8	≥ 4
25	≥ 10	≥ 5
30	≥ 12	≥ 5
35	≥ 13	≥ 6
40	≥ 15	≥ 6

7.2 Primary Analysis

The primary analysis for the dose-exploration phase will occur when target enrollment is complete and each subject either completes 6 months on the study or withdraws from the study.

The primary analysis for time to event endpoints (eg, duration of response, PFS, and 1-year overall survival [OS]) will occur when target enrollment is complete for the dose expansion phase and each subject either completes 1 year on study or withdraws from the study.

Data will be subject to ongoing 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.

7.3 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 ongoing 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

Subjects may miss specific data points for a variety of causes. In general, data could be missing due to a subject's early withdrawal from the study, a missed visit, or inability to evaluate an endpoint at a particular point in time.

In general, the safety analysis set will be used without any imputation for missing data for the primary and secondary endpoints. [redacted]

[redacted]

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. The clinical study team will identify and document the criteria for important protocol deviations.

8.5 Outliers

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

Descriptive statistics will be used to identify potential outliers in key variables. Suspected outliers will be included in the analyses unless there is sufficient scientific justification to exclude them.

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, 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

Descriptive statistics will be provided for selected demographics, safety, PK, efficacy, [redacted] 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.

Confidence intervals (CI) for proportions will be estimated using an exact method proposed by Clopper-Pearson (Clopper and Pearson, 1934). Kaplan-Meier methods will be used to estimate the median and percentiles for time to event endpoints with CI calculated using the Brookmeyer and Crowley (Brookmeyer and Crowley, 1982) method. Kaplan-Meier methods will be used to estimate landmarks for time to event endpoints (eg, 1-year overall survival [OS]) with the Greenwood formula (Kalbfleisch and Prentice, 1980) used to estimate the standard error used in CI calculation.

Greenwood formula will be used for landmark analysis, and for other time-to event analysis the Brookmeyer and Crowley formula will be used. We will be only doing the landmark analysis for OS.

9.2 Subject Accountability

The number and percent of subjects who were enrolled and received the investigational product, discontinued from AMG 160 (including reasons for discontinuing), completed study, discontinued the study (including reasons for discontinuing) will be summarized.

Key study dates for the first subject enrolled, the last subject enrolled, and last subject's end of study will be presented.

A subject listing and summary noting inclusion in each analysis subset will be provided for all subjects enrolled. A subject listing noting AMG 160 administration start and end time, reason for discontinuation of treatment, and reason for discontinuing the study will be provided. A list of subjects screened but not enrolled (screen failures) may be provided.

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 course of the study. Eligibility deviations are defined in the protocol.

The final IPD list is used to produce the summary of IPDs table and list the subjects with IPDs. In addition, a separate listing of all inclusion and exclusion deviations will be

provided. Important protocol deviations thought to potentially impact the safety of subjects or the interpretation of the analyses will be listed and tabulated using incidence and percentages by deviation type.

The number of subjects reporting Protocol Deviation due to COVID-19 will be summarized in a table. A Protocol Deviation listing of subjects impacted due to COVID-19 will also be provided.

9.4 Demographic and Baseline Characteristics

The following descriptive summaries of demographics and baseline characteristics will be summarized.

Demographics:

- Age (years) at enrollment (continuous summary statistics)
- Age categories (number and percent of subjects 18-64, 65-74, 75-84, >=85 years)
- Sex (number and percentages of males and females)
- Race (number and percentages of subjects in each category or combinations)
- Ethnicity (number and percentages of subjects in each category)

Baseline Characteristics:

- Height and Weight (continuous summary statistics)
- ECOG performance status (frequency and percentages)

9.5 Efficacy Analyses

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

Not applicable to this study.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

For the secondary efficacy endpoints mentioned below, we will consider the endpoints to be derived and presented separately for the Central review and the Investigator's review except for the overall survival, clinical progression-free survival, time to clinical progression and time to subsequent therapy.

Table 9-1. Secondary Efficacy Endpoint Summary Table

Endpoint	Statistical Analysis	Analysis Set
Objective response per modified Response Evaluation Criteria in	The proportion of subjects with an objective response per modified RECIST 1.1 along with the corresponding	**RECIST 1.1 Evaluable Analysis Set**

Endpoint	Statistical Analysis	Analysis Set
Solid Tumors (RECIST) 1.1	exact 95% CI will be calculated using the Clopper-Pearson method (Clopper and Pearson, 1934) Waterfall plots for the maximum percent decrease from baseline in sums of longest diameters will also be provided.	Applicable for all subjects with measurable baseline disease and at least one post-baseline evaluable tumor response assessment
Overall Survival (OS)	OS will be estimated and the median and other percentiles, as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using Greenwood's formula to estimate the standard error of the landmark estimates, see (Kalbfleisch and Prentice, 1980).	Safety Analysis Set
Radiographic Progression Free Survival (PFS) Clinical Progression Free Survival (clinical PFS)	The median and other percentiles as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using the Brookmeyer and Crowley (Brookmeyer and Crowley, 1982) method.	Safety Analysis Set
Time to Response (TTR)	The median and other percentiles as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using the Brookmeyer and Crowley (Brookmeyer and Crowley, 1982) method.	**RECIST 1.1 Evaluable Analysis Set**
Time to Progression (TTP) Radiographic Time to Progression (TTP) Clinical	The median and other percentiles as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using the Brookmeyer and Crowley (Brookmeyer and Crowley, 1982) method.	Safety Analysis Set
Duration of Response (DOR)	The median and other percentiles as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using the Brookmeyer and	**RECIST 1.1 Evaluable Analysis Set**

Endpoint	Statistical Analysis	Analysis Set
	Crowley (Brookmeyer and Crowley, 1982) method.	
Time to subsequent therapy	The median and other percentiles as appropriate, will be presented using the Kaplan-Meier curve and 95% CI using the Brookmeyer and Crowley (Brookmeyer and Crowley, 1982) method.	Safety Analysis Set

9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoint(s)

Table 9-3. Primary Safety Endpoint Summary Table

Endpoint	Statistical Analysis	Analysis Set
Dose Limiting Toxicities (DLTs)	Subject incidence of DLT will be tabulated by planned dose level. A table of DLT will be provided. The probability of subject having a DLT by	DLT Analysis Set

	dose level will be estimated using a modified Toxicity Probability Interval (mTPI) 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 and grade. The severity of each adverse event will be graded using CTCAE version 5.0 with the exception of CRS, TLS, and ICANS.

CRS must be graded using the criteria referenced by Lee et al, 2014, TLS must be graded using the criteria referenced by Coiffier et al, 2008, and ICANS must be graded using the criteria referenced by Lee et al, 2019.

The subject incidence of all treatment-emergent adverse events, treatment-related adverse events, serious adverse events, adverse events leading or discontinuation of investigational product or other protocol required therapies, and fatal adverse events will be summarized and tabulated by system organ class and preferred term in alphabetical order.

Subject incidence of all CRS, TLS, and ICANS events will be summarized.

Subject incidence of events of interest (standardized MedDRA queries and/or Amgen Medical Queries) will also be summarized according to their categories, preferred term, and worst severity grade.

Treatment emergent adverse events occurring on or after the COVID-19 infection will be summarized.

A subgroup analysis of safety with CRS outcome and PK will be performed for subjects who were administered siltuximab.

9.6.3 Laboratory Test Results

Clinical chemistry, hematology, and urinalysis data will be reviewed for each subject. Depending on the size and scope of changes in laboratory data the analyses of safety laboratory endpoints will include summary statistics over time and/or changes from baseline over time may be provided.

Shifts in grades of safety laboratory values from baseline for selected laboratory values may also be provided.

9.6.4 Vital Signs and ECOG

Vital signs data will be reviewed for each subject. The analyses of vital signs will include summary statistics over time and/or changes from baseline over time may be provided.

Shifts in scores for ECOG performance status scores between the baseline and each assessed time point will be tabulated and ECOG performance status scores will be summarized at each assessed time points.

9.6.5 Physical Measurements

Physical measurements will be reviewed and listed for each subject.

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 QTcF 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. All on-study ECG data will be listed and select parameters of interest plotted.

[redacted]

9.6.8 Exposure to Investigational Product

Details of AMG 160 administration will be listed for every subject. A listing of unique manufacturing lot numbers and a listing of the subjects administered each manufacturing lot number will be provided.

Descriptive statistics of the total dose (mg), duration of usage, number of cycles, cumulative dose, number and percentage of subjects with dose modifications, reasons for modifications will be produced to describe the exposure to investigational product by treatment group.

9.6.9 Exposure to Non Investigational Product

Subject-level data may be provided instead of the summary if the subject incidence is low or single dose is given. The number of days on non-investigational product, the daily dose, and the proportion of subjects receiving each dose level will be summarized using descriptive statistics.

Compliance with non-investigational product will summarized for each subject using descriptive statistics.

9.6.10 Exposure to Other Protocol-required Therapy

Descriptive statistics of total dose (mg/kg), total dose (mg), average daily dose, cumulative dose, and number of cycles will be produced to describe the exposure to each of the other protocol-required therapy.

9.6.11 Exposure to Concomitant Medication

Number and proportion of subjects receiving therapies of interest will be summarized by preferred term or category for each treatment group as coded by the World Health Organization Drug dictionary.

9.7 Other Analyses

The other analyses included in the study are described as below. These exploratory analyses may not be part of the clinical study report.

9.7.1 Analyses of Pharmacokinetic or Pharmacogenetic Endpoints

Serum concentrations of AMG 160 will be determined using a validated assay. PK parameters will include but are not limited to maximum observed concentration (C_{max}), minimum observed concentration (C_{min}) and area under the concentration-time curve over the dosing interval [AUC] and if feasible half-life ($t_{1/2}$). Pharmacokinetic parameters will be estimated using standard non-compartmental approaches based on the PK analysis set and summarized by dose level using descriptive statistics but not limited to means, standard deviations, medians, minimums, and maximums.

Above analysis will be conducted by Amgen Clinical Pharmacology Modeling and Simulation (CPMS).

[redacted]

10. Changes From Protocol-specified Analyses

RECIST 1.1 Evaluable Analysis Set is added for secondary efficacy endpoint evaluation.

Nomenclature referring to rPFS has been removed as rPFS is a specific endpoint for prostate cancer. For this study, the terminology radiographic PFS will be used to assess the RECIST 1.1 endpoints referring to radiographic progression and radiographic progression-free survival.

11. Literature Citations / References

Bailis J, Deegen P, Thomas O, et al. Preclinical evaluation of AMG 160, a next-generation bispecific T cell engager (BiTE) targeting the prostate-specific membrane antibody PSMA for metastatic castration-resistant prostate cancer (mCRPC). *J Clin Oncol*. 2019;37:(suppl 7S; abstr 301).

Brahmer JR, Govindan R, Anders RA, et al. The Society for Immunotherapy of Cancer Consensus Statement on Immunotherapy for the Treatment of Non-Small Cell Lung Cancer (NSCLC). *J Immunother Cancer*. 2018;6:75.

Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. *Biometrics*. 1982;38(1):29-41.

Clopper CJ, Person ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometricka*. 1934;26:404-413.

Coiffier B, Altman A, Pui CH, Younes A, Cairo MS. Guidelines for the management of pediatric and adult tumor lysis syndrome: an evidence-based review. *J Clin Oncol*. 2008;26:2767-2778.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247.

Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. John Wiley & Sons, New York; 1980.

Lee DW, Gardner R, Porter DL, et al. Current concepts in the diagnosis and management of cytokine release syndrome. *Blood*. 2014;124:188-195.

Lee DW, Santomasso BD, Locke FL, et al. ASTCT Consensus Grading for Cytokine Release Syndrome and Neurologic Toxicity Associated with Immune Effector Cells. *Biol Blood Marrow Transplant*. 2019;25(4):625-638.

Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5:649–655.

Shetty D, Loh H, Bui C, et al. Elevated ^{68}Ga Prostate-Specific Membrane Antigen Activity in Metastatic Non-Small Cell Lung Cancer. *Clin Nucl Med*. 2016;41:414-416.

Thall PF, Simon RM, Estey EH. Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes. *Statistics in Medicine*. 1995;14:357-379.

12. Prioritization of Analyses

There is no prioritization of analyses.

13. Data Not Covered by This Plan

Exploratory data include in this plan may be analyzed later or may be analyzed by a different Amgen department.

14. Appendices

[redacted]

Appendix B. Code Fragments

95% Confidence Interval by Clopper Pearson Method:

```
proc freq data=<data> ;

tables <variable>*<variable> / binomial (exact);

run;
```

Kaplan Meier Method **and 95% Confidence Interval by Brookmeyer and Crowley Method:**

```
proc lifetest data=<data> method = km conftype = linear alpha = 0.05;

time T * status (0) ;

run ;
```

Where T is time to event, status is censoring variable and 0 is flagged for censored observation.

Appendix C. RECIST (v1.1) Criteria

The protocol requires that a PD must be confirmed. The analysis of modified RECIST 1.1 BOR endpoints will follow standard RECIST 1.1 convention which does not require confirmation of PD.

Objective response (OR) per Response Evaluation Criteria in Solid Tumors (RECIST) 1.1

Timepoint Response by Investigator:

Timepoint response by Investigator is the overall tumor response collected on the Overall Response modified RECIST 1.1 CRF at each tumor assessment.

Confirmed BOR by Investigator:

Best Overall Response (BOR) by Investigator is derived from the timepoint response by investigator using RECIST 1.1 criteria. At each time point, BOR will be derived based upon the evaluated time points up to and including the current assessment. The following rules will apply to BOR:

- CR is better than PR; PR is better than SD; SD is better than PD; PD is better than NE.
- For a BOR of SD, a duration of SD ≥ 49 days since study day 1 is required.
- For a BOR of CR and PR, confirmation is required **by a repeat assessment at least** ≥ 28 days after the initially observed assessment of CR or PR
    - A CR must be confirmed by CR, a PR can be confirmed by a PR or CR.
    - **For confirmation of PR or CR,** an unlimited number of intermittent assessments of NE can occur between the initial response and the confirmation. For example, BL, CR, NE, NE, NE, CR – the CR at post-baseline 1st **assessment** is confirmed at post-baseline 5th **assessment**.
    - **For confirmation of PR, an unlimited number of intermittent assessments of SD can occur between the initial PR and the confirmation. It is reasonable to consider a patient with time point response of PR-SD-SD-PR as a confirmed response i.e., confirmed PR.**

Table 14.1 provides the BOR determination per RECIST 1.1 for trials where confirmation of response is required. A BOR determined by Table 14.1 is considered as confirmed

BOR. Table 14.2 outlines the steps to derive Confirmed_BOR (step 1) given investigator's timepoint assessments.

Interim BOR by Investigator:

At interim futility analysis, due to lack of sufficient follow-up time for confirmation scan, study team may choose to report both confirmed responders and unconfirmed responders (subjects had an initial PR or CR and still has potential for future confirmative scans). Interim BOR is defined to include unconfirmed responders in addition to those who achieve confirmed BOR.

Timepoint Response by Central Review:

Timepoint response by Central Review is the Overall response at each tumor assessment as assessed by Central radiographic reader(s). This data is usually non-CTDB dataset transferred from Central Review vendor.

Confirmed BOR by Central Review:

Confirmed BOR by Central Review will be derived from the Timepoint Response by Central Review in the same way as Confirmed BOR by Investigator.

Interim BOR by Central Review:

Interim BOR by Central Review will be derived from Timepoint response by Central Review in the same way as Interim BOR by Investigator.

Best Overall Response (BOR):

The best overall response is the best observed disease response recorded from the start of the study treatment until disease progression/recurrence, the start of new anticancer therapy or the end of study. Confirmed BOR by Investigator, Interim BOR by Investigator, Confirmed BOR by Central Review, Interim BOR by Central Review have been separately defined above.

The followings steps would be followed to derive ORR as per RECIST 1.1.

Table 14.1. BOR per RECIST 1.1 where confirmation of CR/PR is required

Criterion	Time point T1 Response	T1 ≥ 49 days after Baseline?	Time point T2 Response	T2 ≥ 49 days after Baseline?	T2 ≥ 28 days after T1?	Confirmation
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.
*If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best 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 the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

Note: The UE response collected on eCRF will be mapped to NE for all the derivations and analysis.

Deriving Objective Response Rate (ORR) per RECIST 1.1

ORR is defined as the proportion of subjects with Objective Response while on study as defined by RECIST 1.1. All subjects that do not meet the criteria for objective response by the analysis cut-off date will be considered as non-responders.

ORR = (Number of Subjects with Objective Response) / Total number of subjects in the analysis set.

Note: Study allows NE between CR and its confirmation scan (CR-NE-CR), and NE/SD between PR and its confirmation scan (PR-SD-PR or PR-NE-PR).

Note: We will be deriving two sets of ORR and **presenting**, namely Objective Response Rate by Central Review and Objective Response Rate by Investigator Review.

Unconfirmed Responses for Interim Analysis

For a study where CR/PR confirmation is required, due to lack of sufficient follow-up time at interim analysis, both confirmed responders and unconfirmed responders will be reported. Subjects with any objective response (i.e., CR or PR) that is not confirmed as defined in Section 5.1.2 will be summarized based on below definitions.

Interim response of unconfirmed CR awaiting confirmatory scan:

A subject will have an interim response of unconfirmed CR awaiting confirmatory scan if and only if each of the following requirements is met: 1) the confirmed best overall response is not a confirmed CR and 2) the last evaluable tumor assessment prior to data cutoff (DCO) is a CR and 3) none of the following events are recorded (end of study date, death date, new anti-cancer therapy, end of radiographic follow up date, progression date).

Interim response of unconfirmed PR awaiting confirmatory scan:

A subject will have an interim response of unconfirmed PR awaiting confirmatory scan if and only if each of the following requirements is met: 1) the confirmed best overall response is not a confirmed CR and not a confirmed PR and 2) the last evaluable non-SD tumor assessment prior to data cutoff (DCO) is a PR and 3) none of the following events are recorded (end of study date, death date, new anti-cancer therapy, end of radiographic follow up date, progression date).

Interim response of unconfirmed response awaiting confirmatory scan:

A subject will have an interim unconfirmed response awaiting confirmatory scan if the subject has either an unconfirmed PR or an unconfirmed CR, awaiting a confirmatory scan.

Interim response of unconfirmed CR and not confirmed at next evaluable scan:

A subject will have an interim response of unconfirmed CR and not confirmed at next evaluable scan if and only if each of the following requirements is met: 1) the confirmed best overall response is not a confirmed CR and 2) any evaluable tumor assessment prior to data cutoff (DCO) is a CR and 3) either of the following occurs:

- **A CR is recorded at the last evaluable tumor assessment and any of the following events are recorded (end of study date, death date, new anti-cancer therapy, end of radiographic follow up date, progression date) or**
- **For any CR recorded prior to the last evaluable tumor assessment, the subsequent evaluable tumor assessment is never a CR.**

Interim response of unconfirmed PR and not confirmed at next evaluable scan:

A subject will have an interim response of unconfirmed PR and not confirmed at next evaluable scan if and only if each of the following requirements is met: 1) the confirmed best overall response is not a confirmed CR and not a confirmed PR and 2) the subject does not have an interim response of unconfirmed CR and 3) any evaluable tumor assessment prior to data cutoff (DCO) is a PR and 4) either of the following occurs:

- **A PR is recorded at the last evaluable non-SD tumor assessment and any of the following events are recorded (end of study date, death date, new anti-cancer therapy, end of radiographic follow up date, progression date) or**
- **For any PR recorded prior to the last evaluable non-SD tumor assessment, the subsequent evaluable tumor assessment is never a PR or CR.**

Interim response of unconfirmed response and not confirmed at next scan:

A subject will have an interim unconfirmed response and not confirmed at next scan if the subject is classified as either unconfirmed PR or an unconfirmed CR, not confirmed at next scan.

Appendix D. Additional Definition

Information described below will be reported from eCRF

Partial Response (PR):

Partial response (PR) is defined as at least a 30% decrease in the sum of longest diameters of target lesions, taking as reference the baseline sum diameters. There can be no appearance of new lesions.

Complete disappearance of all index lesions with presence of non-index lesions. Confirmation scan required **by a repeat** assessment no less than 4 weeks from the date of the first documented response.

Complete Response (CR):

The 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. There can be no appearance of new lesions.

Confirmation scan required **by a repeat** assessment no less than 4 weeks from the date of the first documented response.

Progressive Disease (PD):

Progressive disease (PD) is defined as radiologic detection of ≥ 20% increase in tumor burden relative to nadir and at least 5 mm absolute increase or unequivocal progression of non-index lesions and confirmation of new lesions.

Stable Disease (SD):

Stable disease (SD) is defined as neither sufficient shrinkage of CR to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum of diameters since the treatment started.

Unable to evaluate (UE):

Any lesion present at baseline which was not assessed or was unable to be evaluated leading to an inability to determine the status of that particular tumor for that time point.

Measurable baseline disease:

Measurable baseline disease is defined as “Lesions must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of 10 mm by CT scan or MRI. To be measurable a lymph node just

be >=15mm in short axis when assessed by CT scan or MRI. All tumor measurements must be recorded in millimeters.

Evaluable tumor assessment:

Either of the following assessment must be performed for an evaluable tumor assessment (on or before subsequent anti-cancer therapy):

- **CT/MRI scan to assess tumor response per RECIST 1.1**

The results must be adequate to define a visit response, i.e., a visit overall response other than not evaluable (NE) for RECIST 1.1.

Appendix E. Censoring Rules

Censoring rules for PFS or TTP (radiographic):

Situation up to DCO/EOS	Date of Event or Censoring	Outcome
No evaluable post-baseline or on-study disease assessment; On study without PD or death	Date of the first dose of investigational product (AMG 160)	Censored
PD prior to new anti-cancer therapy	Date of PD	Event
No PD, but death recorded without new anti-cancer therapy	PFS: Date of death; TTP: Date of last evaluable radiographic tumor assessment	PFS: Event; TTP: Censored
Start of new anti-cancer therapy prior to PD or death, or prior to any other disease assessment if there is no PD or death recorded	Date of last evaluable radiographic tumor assessment prior to the start of new anti-cancer therapy	Censored
No PD or death, no new anti-cancer therapy	Date of last evaluable radiographic tumor assessment	Censored
Death or PD immediately after more than one consecutively missed tumor assessment**	Date of last evaluable radiographic tumor assessment prior to the missing assessments*	Censored

DCO: data cutoff; EOS: end of study; PD: progressive disease

*** This supersedes the previous rules that result in Event at date of PD or death**

****Implementation details: censoring with death or PD immediately after more than one consecutively missed tumor assessment**

Scenarios of PD or Death	Censoring Rule	X Days in Censoring Rule
While on treatment	If PD or death is X days after the last scan, censor at the last scan	119

Censoring rules for Duration Of Response (CT/MRI assessment):

Situation up to DCO/EOS	Date of Event or Censoring	Outcome
PD prior to new anti-cancer therapy	Date of PD	Event
No PD, but death recorded without new anti-cancer therapy	Date of death	Event
Start of new anti-cancer therapy prior to PD or death, or prior to any other disease assessment if there is no PD or death recorded	Date of last evaluable tumor assessment by CT/MRI prior to the start of new anti-cancer therapy	Censored
No PD or death, no new anti-cancer therapy	Date of last evaluable tumor assessment by CT/MRI	Censored
Death or PD immediately after more than one consecutively missed tumor assessment**	Date of last evaluable tumor assessment by CT/MRI prior to the missing assessments*	Censored

DCO: data cutoff; EOS: end of study; PD: progressive disease

*** This supersedes the previous rules that result in Event at date of PD or death**

****Implementation details: censoring with death or PD immediately after more than one consecutively missed tumor assessment**

Scenarios of PD or Death	Censoring Rule	X Days in Censoring Rule
While on treatment	**If PD or death is X days after the last scan, censor at the last scan**	**119**

Censoring rules for Clinical Progression-free Survival:

Situation up to DCO/EOS	Date of Event or Censoring	Outcome
No evaluable baseline or on-study disease assessment; on study withoutdisease progression or death recorded.	**Date of the first dose of investigational product (AMG 160)**	**Censored**
Clinical progression	**Date of clinical progression**	**Event**
Death	**Date of death**	**Event**
All other scenarios	**Date of last assessment**	**Censored**

DCO: data cutoff; EOS: end of study;

Censoring rules for Time to Clinical Progression:

Situation up to DCO/EOS	Date of Event or Censoring	Outcome
No evaluable baseline or on-study disease assessment; on study without disease progression or death recorded.	**Date of the first dose of investigational product (AMG 160)**	**Censored**

Clinical progression	**Date of clinical progression**	**Event;**
All other scenarios	**Date of last assessment**	**Censored**

DCO: data cutoff; EOS: end of study;