Statistical Analysis Plan

Protocol Title:	A Phase 1b Multicenter, Open-label, Expansion Study to Assess the Safety and Efficacy of AMG 420 as Monotherapy in Subjects With Relapsed and/or Refractory Multiple Myeloma
Short Protocol Title:	A Phase 1b Study to Assess AMG 420 Monotherapy in Subjects with Relapsed and/or Refractory Multiple Myeloma
Protocol Number:	20160370
NCT Number:	NCT03836053
Authors:	[redacted] Biostatistician II, IQVIA [redacted] Biostatistics Director, Amgen Inc
Sponsor:	Amgen Inc One Amgen Center Drive Thousand Oaks, CA 91320
SAP Date:	**Document Version** — **Date** **Amendment 1 (v[2.0])** — 12 June 2020

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	11JAN2019	NA
Amendment 1 (v2.0)	12Jun2020	• Removal of Phase 1b Part 2 (combination cohort) and Phase 2 text globally

Table of Contents

Table of Contents 3

1. Introduction 9

2. Objectives, Endpoints and Hypotheses 9
   - 2.1 Objectives and Endpoints 9
   - 2.2 Hypotheses and/or Estimations 11

3. Study Overview 11
   - 3.1 Study Design 11
   - 3.2 Sample Size 13
   - 3.3 Adaptive Design 13

4. Covariates and Subgroups 13
   - 4.1 Planned Covariates 13
   - 4.2 Subgroups 13

5. Definitions 14

6. Analysis Sets 20
   - 6.1 Safety Analysis Set 20
   - 6.2 Dose Limiting Toxicity Evaluable Analysis Set 20
   - 6.3 Pharmacokinetic/Pharmacodynamic Analyses Set(s) 20
   - 6.4 Retreatment Analysis Set 21

7. Planned Analyses 21
   - 7.1 Interim Analysis and Early Stopping Guidelines 21
     - 7.1.1 Interim Analyses 21
     - 7.1.2 Dose Level Review Team 21
   - 7.2 Primary Analysis 21
   - 7.3 Final Analysis 21

8. Data Screening and Acceptance 21
   - 8.1 General Principles 21
   - 8.2 Data Handling and Electronic Transfer of Data 21
   - 8.3 Handling of Missing and Incomplete Data 21
   - 8.4 Detection of Bias 22
   - 8.5 Outliers 22
   - 8.6 Distributional Characteristics 22
   - 8.7 Validation of Statistical Analyses 22

9. Statistical Methods of Analysis 22
   - 9.1 General Considerations 22
   - 9.2 Subject Accountability 23
   - 9.3 Important Protocol Deviations 23
   - 9.4 Demographic and Baseline Characteristics 23

9.5 Efficacy Analyses 24
9.5.1 Analyses of Key Secondary Efficacy Endpoint(s) 24
9.5.1.1 Overall Response Rate 24
9.5.1.2 Duration of Response 24
9.5.1.3 MRD Negativity at CR 24
9.5.2 Analyses of Secondary Efficacy Endpoint(s) 24
9.5.2.1 Progression-free Survival 24
9.5.2.2 Overall Survival 24
9.5.2.3 Time to Response 25
9.5.3 Analyses of Exploratory Efficacy Endpoint(s) 25
9.6 Safety Analyses 25
9.6.1 Analyses of Safety Endpoint(s) 25
9.6.2 Adverse Events 25
9.6.3 Laboratory Test Results 27
9.6.4 Vital Signs 27
9.6.5 Physical Measurements 28
9.6.6 Electrocardiogram 28
[redacted] 28
9.6.8 Exposure to Investigational Product 28
9.6.9 Exposure to Concomitant Medication 28
9.7 Other Analyses 29
9.7.1 PK Analyses 29
9.7.2 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints 29
9.7.3 Analyses of Clinical Outcome Assessments 29
9.7.4 Analyses of Health Economic Endpoints 29
9.7.5 Analyses of Biomarker Endpoints 29

10. Changes From Protocol-specified Analyses 29

11. Literature Citations / References 30

12. Prioritization of Analyses 31

13. Data Not Covered by This Plan 31

14. Appendices 32
Appendix A. Reference Values/Toxicity Grades 33
Appendix B. Guidance for Response Confirmation 35
Appendix C. Clinical Outcome Assessment Forms/Instruments 37
Appendix D. Health Economic Forms/Instruments 38
Appendix E. Details of PK or PK/PD Methods for Modeling 39

List of Tables

Table 1. Rules for Holding Enrollment in Phase 1b 12
Table 2. Conventions for Censoring for PFS and DOR 17

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
APRIL	a proliferation-inducing ligand
AST	aspartate aminotransferase
AUC_{last}	area under the concentration-time curve from time 0 to the last measurable concentration
BAFF	B-cell activating factor
BAFF-R	B-cell activating factor receptor
[redacted]	[redacted]
BiTE®	bispecific T-cell engager
BM	bone marrow
BOR	best overall response
CAR	chimeric antigen receptor
cIV	continuous intravenous
C_{max}	maximum concentration
CR	complete response
C_{ss}	steady-state concentration
CTCAE	Common Terminology Criteria for Adverse Events
DLRM	dose level review meeting
DLRT	dose level review team
DLT	dose limiting toxicity
DOR	duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture
EM	Extramedullary
EOI	end of infusion
EOS	end of study
EOT	end of treatment
FDA	Food and Drug Administration
FIH	first-in-human
FISH	fluorescent in situ hybridization

Abbreviation or Term	Definition/Explanation
FLC	free light chain
IMiD	Immunomodulator
IMWG	International Myeloma Working Group
ISS	International Staging System
IV	intravenous(ly)
KM	Kaplan-Meier
LTFU	long-term follow-up
MR	minimal response
MRD	minimal residual disease
MTD	maximum tolerated dose
mTPI	modified toxicity probability interval
[redacted]	[redacted]
ORR	overall response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PI	proteasome inhibitor
PK	pharmacokinetic(s)
PR	partial response
PR interval	PR interval is measured from the beginning of the P wave to the beginning of the QRS complex in the heart's electrical cycle as measured by ECG or partial response
QRS interval	QRS interval is the interval between the Q wave and the S wave in the heart's electrical cycle as measured by ECG; represents the time it takes for the depolarization of the ventricles
QTc interval	QT interval corrected for heart rate using accepted methodology
RRMM	Relapsed and/or refractory multiple myeloma
[redacted]	[redacted]
SC	subcutaneous(ly)
sCR	stringent complete response
sFLC	serum free light chain
SFU	safety follow-up
SPEP	serum protein electrophoresis
$t_{1/2}$	half-life
TBL	total bilirubin
TFI	treatment-free interval
TLS	tumor lysis syndrome

Abbreviation or Term	Definition/Explanation
TNF	tumor necrosis factor
TPI	toxicity probability interval
T_{reg} cells	regulatory T cells
UIFE	urine immunofixation
ULN	upper limit of normal
UPEP	urine protein electrophoresis
US	United States
VGPR	very good partial response

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 20160370, AMG 420 dated **20 May 2020**. The scope of this plan includes the interim analysis, primary analysis and 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

Objectives	Endpoints
Phase 1b	
Primary	
• Establish the safety and tolerability of AMG 420 at dose levels of 400 μg/day and 600 μg/day in subjects with relapsed and/or refractory multiple myeloma (RRMM)	• Dose-limiting toxicities (DLTs), treatment-emergent adverse events, treatment-related adverse events and changes in vital signs, electrocardiograms (ECGs), and clinical laboratory tests
Key Secondary	
• Estimate overall response rate (ORR) and duration of response (DOR) of AMG 420 in subjects with RRMM	• ORR • DOR
• Evaluate the rate of minimal residual disease (MRD)-negativity at the time of CR	• MRD negativity at the time of CR
Secondary	
• Establish the safety and tolerability of AMG 420 in subjects with extramedullary relapsed multiple myeloma	• DLTs, treatment-emergent adverse events, treatment-related adverse events, and changes in vital signs, ECGs, and clinical laboratory tests
• Characterize the pharmacokinetics (PK) of AMG 420 when administered as 4-week continuous intravenous (cIV) infusion	• AMG 420 PK parameters including, but not limited to, half-life ($t_{1/2}$), clearance, and apparent C_{ss}
• Evaluate other measures of anti-myeloma activity of AMG 420 in subjects with RRMM:	• Efficacy parameters according to International Myeloma Working Group (IMWG) response criteria, per **investigator** assessment:

Objectives	Endpoints
Secondary (Continued)	
– Time to response – Progression-free survival (PFS) – Overall survival (OS) – Best overall response (BOR)	– Time to response – PFS – OS – BOR
Exploratory for Phase1b	
[redacted]	[redacted]
Enter Exploratory Objectives Endpoints	

2.2 Hypotheses and/or Estimations

The ORR according to IMWG response criteria along with 95% CI will be calculated per **investigator** assessment for subjects with RRMM. The lower bound of the binomial 2-sided 95% confidence interval for ORR not lower than 35% is considered clinically meaningful.

3. Study Overview

3.1 Study Design

This is a phase 1b, multicenter, non-randomized, open-label expansion study. AMG 420 will be evaluated in adult subjects with RRMM.

The study will consist of up to a 21-day screening, a treatment, safety follow-up (SFU), and long-term follow-up (LTFU) period. In the treatment period, treatment will be continued until progressive disease (PD) or relapse as defined by IMWG response criteria, unacceptable safety events, next anti-multiple myeloma treatment, or other reason for permanent treatment discontinuation. All subjects will be assessed for anti-myeloma activity according to the IMWG response criteria. An SFU visit will be conducted 30 (+3) days after last administration of AMG 420, and an LTFU period that will begin after the SFU visit is completed. The total duration of the LTFU will be up to 5 years from the first dose of AMG 420 as described in Study Schema of protocol section 2.1.

Phase 1b (Monotherapy Dose Confirmation)

In phase 1b, approximately 10 subjects will be treated at the 400 μg/day dose (comprising cohort 1). Approximately 10 subjects will be treated at the 600 μg/day dose level (cohort 2). Each subject in phase 1b will have at least a 30-day SFU.

The dose level review team (DLRT) will **convene and** review all available safety, laboratory, and pharmacokinetic (PK) data at the following scheduled time points:

- After at least 10 subjects in cohort 1 (400 μg/day dose level) in phase 1b have had the opportunity to complete the 4-week infusion period of the first treatment cycle: On the basis of this review, the DLRT will make a recommendation regarding the initiation of the 600 μg/day dose level (cohort 2) in phase 1b.
- After at least 3 subjects with extramedullary (EM) relapsed disease in cohort 1 (400 μg/day dose level) have had the opportunity to complete the 4-week infusion period of the first treatment cycle: On the basis of this review, the DLRT will make a recommendation regarding the appropriateness of continued enrollment of subjects with EM relapsed disease to cohort 2 in phase 1b **(600 μg/day dose level)**. Please note that enrollment of subjects with EM relapsed disease to cohort 2 in phase 1b will not be allowed until this safety review of the first 3 subjects with EM disease is completed by the DLRT.
- After the first 3 subjects have been dosed with AMG 420 at 600 μg/day for the full 4-week infusion cycle.

The safety data will be reviewed on an ongoing basis and enrollment may be terminated at any time. To further reduce the risk of overdose to ongoing and future subjects, protocol-defined enrollment holding rules will be used. These rules are based on a modified Toxicity Probability Interval (mTPI), where the target Toxicity Probability Interval (TPI) is a dose-limiting toxicity (DLT) probability from 20% to 33% and the overdose TPI is > 33% DLT probability. Enrollment to the ongoing dose level/schedule will be paused per the following rules during enrollment into phase 1b cohort 1 and 2.

Table 1. Rules for Holding Enrollment in Phase 1b

Number of DLTs	≥ 2	≥ 3	≥ 4
Number of Treated Subjects	≤ 4	≤ 6	≤ 8

If the rules for pausing enrollment are met, no further enrollment to the respective dose level/schedule will be allowed until after a careful review of all safety data by the DLRT. Additionally, enrollment will not resume until the protocol has been amended to include additional safeguards for subject safety (eg, modifications to dose level and schedule).

Subjects with EM disease will be able to enroll into cohort 2 **after** the safety evaluation is complete for 3 subjects with EM disease treated at 400 μg/day (cohort 1).

An **investigator** will review efficacy data and assess outcomes in accordance with IMWG criteria. **Investigator** assessments will be used for efficacy analysis.

3.2 Sample Size

Approximately 20 subjects will be enrolled in this study.

3.3 Adaptive Design

The rules for halting enrollment are based on an mTPI design. The mTPI models the probability of toxicity for each dose level using a Bayesian model where each dose level has the same prior on the probability of toxicity, a Beta (1,1). When subjects are treated at the current dose level, the posterior probability of toxicity is updated using the observed data from this level. Enrollment recommendations are made based on this posterior probability of toxicity, using three toxicity probability intervals (TPI).

- Under-dosing TPI: DLT rate from 0 to < 20%
- Target TPI: DLT rate from 20% to 33%
- Over-dosing TPI: DLT rate > 33%

For the current dose level and after adjusting for the width of each TPI interval, if the DLT rate is most likely in the under-dosing TPI or most likely in the target TPI then the recommendation is to continue enrollment at the current level. If the DLT rate is most likely in the over-dosing TPI then the recommendation is to halt enrollment.

4. Covariates and Subgroups

4.1 Planned Covariates

There are no planned covariates in this study.

4.2 Subgroups

The baseline variables for the subgroup analyses may include:

- baseline demographics and characteristics:
  - age (< 65, 65-74, ≥ 75 years)
  - sex (female, male)
  - race (White, Asian, other)
- baseline disease characteristics:
  - revised International Staging System (ISS) stage (Stage I, Stages II or III)
  - BM percent of plasma cells (< 50%, ≥ 50%)
  - [redacted]
  - [redacted]
  - extramedullary disease (yes, no). Until safety evaluation is complete for 3 subjects with EM disease, imaging evaluation at screening will be used to document EM disease status.
- number of prior lines of therapy (only 3, > 3)

- prior allogeneic HSCT (yes, no)
- refractory to
  - IMiD (yes, no)
  - PI (yes, no)
  - IMiD + PI (yes, no)
  - daratumumab or other CD 38-targeting antibody (yes, no)

Subgroup analysis will not be performed due to each sub group may not have justifiable count as less number of sample size.

5. Definitions

Adverse Event (AE)

An adverse event is defined as any untoward medical occurrence in a clinical trial subject. The event does not necessarily have a causal relationship with study treatment. The definition of adverse events includes worsening of a preexisting medical condition. A preexisting condition that has not worsened during the study, or involves an intervention, is not considered an adverse event.

Age at Enrollment

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

Baseline

For any variable, unless otherwise specified the baseline is the last assessment taken prior to the first administration of AMG 420. For parameters/assessments not scheduled to be performed (or scheduled but not performed) on the same day as the first administration of AMG 420, the baseline value is the value from the screening period measured closest to the day of first administration of AMG 420.

Baseline ECG Values in Triplicate

For the three sets of triplicate pre-dose ECG, the mean of values in a triplicate should be calculated before taking the mean of the three sets of averages. 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.

Death Date

For subjects who die during the study, the death date will be recorded on the end of study CRF in the date subject ended study field. Incomplete death dates where only the day of death date is missing will be imputed using the following rules:

Day 1 of the month will be used if year and month indicate that death happened in different month from last known alive date;

One day after last known alive date will be used if death happened in the same month as last known alive date.

The imputed death date will be used in calculation of duration of response, progression-free survival and overall survival.

Dose-limiting Toxicities (DLTs)

Investigators determine whether a TEAE qualifies as a DLT per the definition described in protocol Section 7.4.1. For an adverse event to qualify as a DLT, the start date must be within 28 days from the date of the first dose of investigational product.

Best Overall Response

Best overall response for a subject is the best observed post baseline disease response per the International Myeloma Working Group Uniform Response Criteria for Multiple Myeloma (IMWG-URC). **The detailed algorithm is provided in** Appendix B**.**

Overall Response Rate (ORR)

ORR refers to the proportion of subjects with an objective response. Objective response refers to any of the following responses per the International Myeloma Working Group Uniform Response Criteria for Multiple Myeloma (IMWG-URC): (sCR, CR, VGPR or PR) Per IMWG-URC, any response requires 2 consecutive assessments made any time before starting any new therapy. Subjects who do not experience a sCR/CR/VGPR/PR or do not have any follow-up tumor assessments will be regarded as non-responders.

Duration of Response (DOR)

Duration of response (DOR) will be calculated for subjects who achieve a PR or better. For such subjects, the duration of overall response is defined as the time (months) from first documentation of response to disease progression or death due to any cause. Dates of progression and censoring will be determined as described for the analysis of PFS.

DOR (in months) = (Earliest date of disease progression, death, censor - Date of first observation of PR or better before confirmation + 1)/30.4

Subjects who have not ended their response at the time of analysis will have DOR censored on the date of their latest evaluable assessment.

Time to response

Time to response applies for subjects who achieved best overall response of PR, VGPR, CR or sCR. Time to response is calculated as the number of months (one mont = 30.4 days) from the first administration of AMG 420 to the first objective assessment of response (ie, date of the first detection of response) as per IMWG-URC.

TTR (in months) = (Date of objective response – Date of first dose + 1)/30.4

MRD Negativity (MRD-)

MRD- is defined as a response of CR or sCR accompanied by any of the following types of MRD negativity.

- Sustained MRD-
- Sequencing MRD-
- Imaging plus MRD-

Subjects who do not experience a MRD negativity or do not have any evaluable follow-up assessments will be regarded as MRD non-responders.

Progression-free survival (PFS)

Progression free survival (PFS) is defined as number of months between start of treatment and first evidence of documented disease progression or death (due to any cause), whichever occurs first.

PFS (in months) = (Earliest date of disease progression, death, censor – Date of first dose + 1)/30.4

Progression-free survival will be right-censored for subjects who meet one of the following conditions: 1) no post baseline disease assessments, 2) non-protocol systemic anticancer treatment started before documentation of PD or death, 3) death or disease progression after more than 1 missed disease assessment visit or 4) alive and does not have documentation of PD before a data analysis cutoff date. For such subjects, the

primary analysis of PFS will be right-censored according to the conventions described in Table 2.

Table 2. Conventions for Censoring for PFS and DOR

Situation	Date of Progression or Censoring	Outcome
No post baseline disease assessments**	Date of first dose	Censored
New anticancer therapy started before documentation of PD or death	Date of last evaluable disease assessment prior to start of new anticancer therapy	Censored
Death or PD immediately after more than 1 consecutively missed disease assessment visit*	Date of last evaluable disease assessment visit before the first missed visit	Censored
Alive and without PD documentation (including lost to follow-up without PD)	Date of last evaluable disease assessment	Censored
Death or PD between planned disease assessments	Date of death or first disease assessment showing PD, whichever occurs first	Progressed
Death before first disease assessment	Date of death	Progressed

* If death or PD is more than 84 days after previous disease assessment, or Study Day 1 if there is no previous disease assessment.
** Only apply to PFS.

Overall survival (OS)

Overall survival is defined as the interval from Study Day 1 until death due to any cause; otherwise, OS is censored on the last recorded date on which the subject was alive.

Last Evaluable Disease Assessment Date

For responses assessed by investigator, last disease assessment date is the last evaluable response assessment date collected using IMWG Response CRF (except the response assessment date of Inevaluable) before confirmed PD, on or before new anti-cancer therapies or procedures.

Last Known Alive Date

For subjects not known to have died, their last date known to be alive will be determined as the latest date associated with clinic visits before data cutoff date including, for example, but not limited to the following:

- Date of Enrollment on Subject Enrollment CRF
- Date First Taken, Date Last Taken on Concomitant Medications CRF

- Date Performed on ECOG Performance Status, Vital Signs, Electrocardiogram, Procedures CRFs
- Admission Date, Discharge Date on Hospitalizations CRF
- Date of Examination on Physical Measurement CRF
- Date Collected on Reproductive Status and Pregnancy Test (Local Lab), Chemistry (Local Lab), Hematology (Local Lab), Coagulation (Local Lab), Urinalysis (local Lab), Immunology (Local lab) CRFs and in central lab data
- Date of Assessment on IMWG Response for Time Point, Peripheral Neuropathy Assessment CRFs
- Start Date and End Date on Investigational Product Administration CRF
- Date Started and Date Ended or Resulted in Death on Events CRF
- Start date, Stop date on Anti-Myeloma / Anti-Cancer Therapies, Other Protocol Required Therapy (GLUCOCORTICOID), Other Protocol Required Therapy (Tocilizumad), Hematopoietic Stem Cell Transplantation Autologous, Hematopoietic Stem Cell Transplantation Allogenic CRFs
- Subject Status Date if status is Alive on Survival Status CRF
- Assessment Date on Skeletal Survey and Plasmacytoma Assessment CRF
- **End of study date if the subject's primary reason for ending study is not "Lost to follow-up"**

Fridericia-corrected QT Interval (QTcF)

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

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

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

Investigational Product (IP)

The term 'investigational product' is used in reference to AMG 420.

Primary Completion:

The primary completion date is defined as the date when the last subject is assessed or receives an intervention 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 in the protocol or was terminated early.

If the study concludes prior to the primary completion date originally planned in the protocol (ie, early termination of the study), then the primary completion will be the date when the last subject is assessed or receives an intervention for evaluation in the study (ie, last subject last visit).

End of Study

The end of study (EOS) date is defined as the date when the last subject across all sites is assessed or receives an intervention for evaluation in the study (ie, last subject last visit), following any additional parts in the study (eg, LTFU), as applicable.

Lost to Follow-up

A subject will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. The following actions must be taken if a subject fails to return to the clinic for a required study visit:

- The site must attempt to contact the subject and reschedule the missed visit as soon as possible and counsel the subject on the importance of maintaining the assigned visit schedule and ascertain whether or not the subject wishes to and/or is able to continue in the study.
- In cases in which the subject is deemed lost to follow-up, the investigator or designee must make every effort to regain contact with the subject (where possible, 3 telephone calls and, if necessary, a certified letter to the subject's last known mailing address or local equivalent methods). These contact attempts are to be documented in the subject's medical record.
- If the subject continues to be unreachable, he/she will be considered to have withdrawn from the study with a primary reason of lost to follow-up.
- For subjects who are lost to follow-up, the investigator can search publicly available records where permitted to ascertain survival status. This ensures that the data set(s) produced as an outcome of the study is/are as comprehensive as possible.

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 protocol-specified investigational product is administered to the subject.

Treatment-emergent Adverse Event (TEAE)

A treatment-emergent adverse event 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 33 days after the last IP dose date or EOS date, whichever occurs first. The severity of each adverse event will be graded using the CTCAE version 5.0 with exception of CRS which will be graded according to adopted grading system in Lee et al, 2014 and TLS that will be graded according to Cairo-Bishop criteria (refer to protocol Appendix 12). Adverse events will be coded using MedDRA.

Safety Follow-up

SFU visit will be conducted 30 (+3) days after last administration of AMG 420.

Long-term Follow-up

LTFU period begin after the SFU visit is completed. The total duration of the LTFU will be up to 5 years from the first dose of AMG 420, or until subject death, whichever is first."

6. Analysis Sets

6.1 Safety Analysis Set

Safety Analysis Set defined as all subjects that are enrolled and receive at least 1 dose of AMG 420.

6.2 Dose Limiting Toxicity Evaluable Analysis Set

The analysis of DLT will be restricted to DLT-evaluable subjects. **DLT evaluable analysis set includes subjects who complete the DLT evaluable period or experienced a DLT anytime during the DLT evaluable period**. A subject will be not DLT evaluable if he/she drops out before completion of the DLT-evaluable period for reasons other than a DLT prior to completing the 4 weeks of AMG420 treatment.

The DLT window may also be extended to assess events starting within the window in case the DLT definition is time dependent.

6.3 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. These subjects will be evaluated for PK analysis unless the number of data points required for analysis is not enough, or significant protocol deviations have affected the data, or if key dosing or sampling information is missing.

6.4 Retreatment Analysis Set

Retreatment analysis set defined as subjects that are enrolled and receive at least one dose of AMG 420 during the second course of treatment after attaining CR and subsequently have disease progression while off-therapy.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

7.1.1 Interim Analyses

No specific interim analysis planned.

7.1.2 Dose Level Review Team

DLRT will be responsible for monitoring safety and making recommendations regarding termination of enrollment and dose level/schedule escalation/changes. See Protocol Section 12.3 for further details of DLRT responsibilities.

7.2 Primary Analysis

The primary analysis will occur when the target enrollment is complete and each subject has had the chance to be treated for at least 6 months.

7.3 Final Analysis

A final analysis is planned after all subjects have ended the study.

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. The database will be subject to edit checks outlined in the data management plan (DMP) by Amgen Clinical Data Management (CDM) department. Data inconsistencies and suspicious values will be reviewed and resolved before the database is locked.

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

The following imputation for missing or incomplete data will be performed if required:

Incomplete adverse event and concomitant medication dates missing data will be imputed as described in Appendix A.

Non-pharmacokinetic measurements (eg, biomarker data) that are below the lower limit of quantification will be considered equal to half of the lower limit of quantification for all analyses unless specified otherwise.

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

Outlier data will not be excluded unless scientifically justified.

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

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. Time to event endpoints will be summarized with Kaplan-Meier (KM) proportions at select time points, KM quartiles (when estimable), the number of

subjects with events, the number of subjects censored, and censoring reasons. Duration of follow-up for time to event endpoints will be estimated using the reverse Kaplan Meier method (Schemper and Smith, 1996). Point estimates for efficacy endpoints will be accompanied by 2-sided 95% confidence intervals including KM quartiles (Klein and Moeschberger, 1997), KM proportions (Kalbfleisch and Prentice, 1980), and binomial proportions (Clopper and Pearson, 1934). Graphical summaries of the data may also be presented.

9.2 Subject Accountability

The following subject disposition information will be summarized as follows

- Number of screened subjects
- Number of subjects in Safety Analysis Set
- Number (%) of subjects who discontinued treatment
- Number (%) of subjects who discontinued the study
- Primary reason for ending treatment
- Primary reason for study discontinuation

The percentages are based on the number of subjects **enrolled in the study**.

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.

9.4 Demographic and Baseline Characteristics

The following descriptive summaries of the demographic and baseline characteristics will be produced: Demographic (ie, age, age groups [< 65, 65-74 and >= 75], sex, race, ethnicity), prior line of therapies (median, range) and baseline characteristics (height, weight, Eastern Cooperative Oncology Group (ECOG) Performance Status will be summarized using descriptive statistics. If multiple races have been reported for a subject, the subject will be categorized as multiple.

9.5 Efficacy Analyses

9.5.1 Analyses of Key Secondary Efficacy Endpoint(s)

9.5.1.1 Overall Response Rate

The number (%) of subjects with a best overall response of sCR, CR, VGPR, PR, MR, SD, PD or IE will be summarized.

Overall response rate is defined as the proportion of subjects for whom the best overall response is sCR, CR, VGPR, or PR as determined by International Myeloma Working Group Uniform Response Criteria (IMWG-URC). The ORR along with the associated 95% exact binomial confidence interval (Clopper-Pearson method) will be determined.

9.5.1.2 Duration of Response

Kaplan-Meier methods will be used to estimate the distribution of DOR. And the median and other quartiles in addition to the corresponding two-sided 95% confidence intervals will be calculated.

A Kaplan-Meier figure showing the estimated DOR distribution will be provided.

9.5.1.3 MRD Negativity at CR

MRD negativity analysis will be performed for all subjects included in the safety analysis set. Number of MRD[-] responders at CR along with exact 2-sided 95% will be provided by using Clopper & Pearson method

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

9.5.2.1 Progression-free Survival

PFS analysis will be performed for all subjects included in the safety analysis set. The Kaplan-Meier curve will be generated. PFS rate at 3-month intervals and the median and other quartiles in addition to the corresponding two-sided 95% confidence intervals will be provided. Disease progression determined by using IMWG-URC will be used for the analysis.

9.5.2.2 Overall Survival

OS analysis will be performed for all subjects included in the safety analysis set. **The Kaplan-Meier curve will be generated**. Overall survival rate at 3-month intervals, the median and other quartiles in addition to the corresponding two-sided 95% confidence intervals will be calculated.

9.5.2.3 Time to Response

Time to response will be calculated for subjects who achieved best overall response of PR, VGPR, CR or sCR. Time to response is calculated as the number of months from the first administration of AMG 420 to the first objective assessment of response (ie, date of the first detection of response) as per IMWG-URC.

Time to response will be descriptively analyzed, summary statistics of the number, mean, standard deviation, median, and range will be reported.

Listings will be produced for all subjects indicating the MRD negativity, OS, PFS, time to response, and duration of response.

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)

9.6 Safety Analyses

9.6.1 Analyses of Safety Endpoint(s)

Statistical analyses on safety endpoints will be done using subjects from the Safety Analysis Set **(except the analysis of the DLT endpoint, which will be based on subjects from the DLT evaluable set)**, which includes subjects that are enrolled and received at least 1 dose of AMG 420.

9.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 23.0 or later will be used to code all adverse events to a system organ class and a preferred term. The severity of each adverse event will be graded using CTCAE version 5.0 criteria with exception of CRS which will be graded according to adopted grading system in Lee et al, 2014 and TLS graded according to Cairo-Bishop criteria (refer to protocol Appendix 12). Where appropriate the tables will also be presented by worst grade.

Treatment-emergent adverse events will be summarized based on the number (%) of subjects experiencing events by MedDRA SOC and PT. The denominator for the percentage will be based on the number of subjects in the safety analysis set.

The subject incidence of adverse events will be summarized for

- **TEAE**
  - **All treatment-emergent AEs**
  - **Grade 3 or higher treatment-emergent AEs**
  - **Treatment-emergent SAEs**
  - **Treatment-emergent AEs leading to interruption of investigational product**
  - **Treatment-emergent AEs leading to withdrawal of investigational product**
  - **Fatal treatment-emergent AEs**
- **Treatment-related AEs**
  - **All treatment-related AEs**
  - **Grade 3 or higher treatment-related AEs**
  - **Treatment-related SAEs**
  - **Treatment-related AEs leading to interruption of investigational product**
  - **Treatment-related AEs leading to withdrawal of investigational product**
  - **Fatal treatment-related AEs**

The summaries that are displayed by SOC, PT and grade will be ordered by descending order of incidence of SOC, PT and grade with in each SOC for the total group.

Then summary of TEAEs, summary of TEAEs, TESAEs, Treatment-related TEAEs, and Treatment-related SAEs by SOC, PT and Grade will be generated for extramedullary subjects.

The above adverse event tables will not be created if two or fewer subjects experience the adverse event.

Dose Limiting Toxicities

The analysis of the probability of DLTs will include data from DLT-evaluable subjects (see Protocol Section 7.4.1 for definition of DLT-evaluable). The primary analysis will only include DLTs that occur during the DLT evaluation period. If DLTs occur outside the DLT evaluation period, a sensitivity analysis will be performed where DLTs occurring outside of the specified time interval are included. The number and percentage of subjects reporting DLTs will be evaluated overall and by AMG 420 dose level.

9.6.3 Laboratory Test Results

Chemistry, Hematology and Coagulation

Actual values and change from baseline for selected chemistry and hematology parameters will be summarized descriptively for each scheduled visit based on the safety analysis set.

Chemistry	Hematology	Coagulation
Total bilirubin	ANC	PT or INR
ALT (SGPT)	Hemoglobin	APTT
AST (SGOT)	Platelets	
Alkaline phosphatase	WBC	
Creatinine	Lymphocyte	
Phosphate		
Calcium (Corrected)		
Potassium		
Sodium		
Albumin		
Glucose		
Amylase		
Lipase		

Shift tables will be presented for the baseline toxicity grade by the worst on-study toxicity grade in a table form for selected parameters mentioned above if applicable. CTCAE version 5.0 will be used for toxicity grading for the above mentioned parameters excluding Total bilirubin, ALT (SGPT), AST (SGOT) and Alkaline phosphatase (CTCAE version 4.03 will be used for toxicity grading).

Potential Hy's law cases may be summarized.

9.6.4 Vital Signs

The analyses of vital signs include summary statistics over time and/or changes from baseline over time.

9.6.5 Physical Measurements

The analyses of physical measurements will include summary statistics at baseline.

9.6.6 Electrocardiogram

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

[redacted]

9.6.8 Exposure to Investigational Product

The average dose administration (μg), cumulative dose (μg), number of cycles, duration of treatment, number and percentage of subjects with dose modifications, reasons for modification will be summarized using descriptive statistics. Subject-level data may be provided instead of the summary if the subject incidence is low or single dose is given.

Descriptive statistics will be produced to describe the exposure to investigational product by treatment group.

Extent of exposure to study treatments will be summarized for Safety population

- Duration of treatment (days)
  - Defined as the duration (in weeks) from the date of the first dose to the date of the last dose of study drug + 1
- Total number of treatment cycles started
  - Defined as the total number of treatment cycles in which at least one dose of study drug is administered.
- Cumulative dose received of AMG420 across all cycles, defined as the cumulative dose of AMG 420.
- Average dose of AMG420, defined as the total dose received divided by the number of days on treatment administered .

9.6.9 Exposure to Concomitant Medication

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

9.7 Other Analyses

9.7.1 PK Analyses

The PK parameters of AMG 420 including, but not limited to, average C_{ss}, clearance, and half-life ($t_{1/2}$) for serum AMG 420 will be estimated using non-compartmental methods and summarized by dose level using means, geometric means, standard

deviations, coefficients of variation, medians, minimums, and maximums. Individual concentration-time profiles will be summarized by dose level. Plasma AMG 420 concentrations at each time point along with PK parameter values may be listed for each subject. Summary statistics will be computed for each sampling time and parameter as appropriate. The relationship between AMG 420 exposure and efficacy/safety may be conducted.

9.7.2 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

PK/PD analyses will be provided by the Department of Clinical Pharmacology, Modeling & Simulation (CPMS). Actual collection times of PK samples will be recorded to assess the plasma concentration. The PK parameter estimates for AMG420 will be summarized. Exploratory analyses may be performed to evaluate the relationship between the estimated PK parameters and selected safety, biomarker, or clinical effect endpoints. In addition, the population modeling program may be used to fit a nonlinear mixed effects model to estimate PK parameters. Details regarding the analyses will be provided in a separate analysis plan by CPMS.

9.7.3 Analyses of Clinical Outcome Assessments

Analysis of clinical outcome assessments is not planned for this study.

9.7.4 Analyses of Health Economic Endpoints

Analyses of Health Economic Endpoints are not planned for this study.

9.7.5 Analyses of Biomarker Endpoints

The analysis of Biomarker will be described and performed by Clinical Biomarker Group.

10. Changes From Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

11. Literature Citations / References

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika.* 1934;26(4):404-413.

Ji Y, Liu P, Li Y, Bekele BN. A modified toxicity probability interval method for dose-finding trials. *Clin Trials.* 2010;7(6):653-663.

Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, Hoboken, New Jersey: John Wiley & Sons; 1980.

Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag; 1997.

Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol.* 2016;17(8):e328-346.

Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials.* 1996;17(4):343–46.

Lee JJ, Liu DD. A predictive probability design for phase II cancer clinical trials. *Clinical Trials*. 2008;5(2):93-106.

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

12. Prioritization of Analyses

No Priority of outputs planned for this study

13. Data Not Covered by This Plan

The analysis of Biomarkers (except MRD) is not covered in this plan.

14. Appendices

Appendix A. Reference Values/Toxicity Grades

The Common Terminology Criteria for Adverse Events (CTCAE) v5.0 is available at the following link:

http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

Technical Detail and Supplemental Information Regarding Statistical Procedures and Programs

Imputation Rules for Partial or Missing Stop Dates

If the month and year are present, impute the last day of the 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.

Imputation Rules for Partial or Missing Start Dates

		Stop Date						
		Complete: yyyymmdd		**Partial: yyyymm**		**Partial: yyyy**		
Start Date		**<1st Dose**	**≥1st Dose**	**<1st Dose yyyymm**	**≥1st Dose yyyymm**	**<1st Dose yyyy**	**≥1st Dose yyyy**	**Missing**
Partial: yyyymm	**=1st Dose yyyymm**	N/A	1	**N/A**	1	N/A	1	1
	≠ 1st Dose yyyymm		2	**2**	2	2	2	2
Partial: yyyy	**=1st Dose yyyy**	N/A	1	N/A	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 day 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.

Note: If the start date imputation leads to a start date that is after the stop date, then do not impute the start date.

Imputation rules for stop date:

For partial stop date mmyyyy, impute the last day of the month.

For partial stop date yyyy, impute December 31 of the year.

For completely missing stop date, do not impute.

If the stop date imputation leads to a stop date that is after the death date, then impute the stop date as the death date.

If the stop date imputation leads to a stop date that is before the start date, then there is a data error and do not impute the stop date. (ie. set the stop date as missing).

Code Fragments:

Provisional Code Fragments for calculating a confidence interval using the Clopper Pearson Method. The following example SAS code will be utilized for the response rate analysis providing the proportion of subjects responding to treatment with corresponding 80% confidence intervals.

```
data propci (keep = ns p low_ci upp_ci);
n=xx; * total n within the treatment group;
do ns= xx; *number of responders;
p=ns/n; * response rate;
q=1-p;
lowF=FINV(0.1, 2*ns, 2*(n-ns+1)); /* use for 2-sided 80% CI */
UppF=FINV(1-0.1, 2*(ns+1), 2*(n-ns)); /* use for 2-sided 80% CI */
low_ci = 1 / (1+(n-ns+1) / (ns*lowf)); * lower CI for response rate;
upp_ci = 1 / (1+(n-ns) / ((ns+1)*uppf)); *upper CI for response rate;
if p=1 then upp_ci=1;
if p=0 then low_ci =0;
output;
end;run;
```

Appendix B. Guidance for Response Confirmation

For responses assessed by investigator, the following table and notes will be used to confirm the responses:

Confirmation of Response Assessed by Investigator at Each Visit (applicable to MR or better and PD)

"Has response been confirmed" Entered on IMWG Response CRF for the current assessment (Yes/No)	**'Has response been confirmed' Entered on IMWG Response CRF for the next assessment (Yes/No)**	**Response Category of the current and the next assessments**	**Confirmation of the current assessment**
Yes	-	-	Confirmed
No	Yes	If the current assessment is MR or better, and the next assessment is no worse than the current one, keep the current response status as is.	Confirmed
		If the current assessment is MR or better, and the next assessment is worse than the current one but also is MR or better, select the response status same as the next assessment.	Confirmed
		If the current assessment is PD, and the next assessment is also PD, keep the current response status as is.	Confirmed
		If the current assessment is PD, but the next assessment is not PD, keep the current response status as is.	Unconfirmed
	No	Leave the response as is.	Unconfirmed

Note:

Response assessments done after initiation of new anti-cancer therapy will be excluded. But if new anti-cancer therapy is started before confirmed progressive disease, additional testing during subsequent therapy can be used to confirm progressive disease (PD), as long as this testing is within 35 days from the last disease assessment on study.

Stable disease (SD) means the response does not fulfil criteria for a confirmed MR, PR, VGPR, CR, sCR, or PD. An assessment of SD does not need to be confirmed, however it should be remains same at least 6 weeks. A response of SD is not upgraded until there is a confirmed response of sCR, CR, VGPR, PR, MR or PD.

For subjects that do not qualify for a confirmed response or stable disease (SD), the BOR will be not evaluable

Appendix C. Clinical Outcome Assessment Forms/Instruments

Not applicable

Appendix D. Health Economic Forms/Instruments

Not applicable

Appendix E. Details of PK or PK/PD Methods for Modeling

CPMS group will be responsible for PK/PD modeling.