Title: A Phase 2, Randomized, Open-Label Study Comparing Oral Ixazomib/Dexamethasone and Oral Pomalidomide/Dexamethasone in Relapsed and/or Refractory Multiple Myeloma

NCT Number: NCT03170882

SAP Approve Date: 11 November 2019

Certain information within this Statistical Analysis Plan has been redacted (ie, specific content is masked irreversibly from view with a black/blue bar) to protect either personally identifiable (PPD) information or company confidential information (CCI).

This may include, but is not limited to, redaction of the following:

- Named persons or organizations associated with the study.
- Proprietary information, such as scales or coding systems, which are considered confidential information under prior agreements with license holder.
- Other information as needed to protect confidentiality of Takeda or partners, personal information, or to otherwise protect the integrity of the clinical study.

STATISTICAL ANALYSIS PLAN

STUDY NUMBER: C16029

A Phase 2, Randomized, Open-Label Study Comparing Oral Ixazomib/Dexamethasone and Oral Pomalidomide/Dexamethasone in Relapsed and/or Refractory Multiple Myeloma

PHASE 2

Version: Final 1.0

Date: 11 November 2019

Prepared by:

PPD

Based on:

Protocol Version: Amendment 5

Protocol Date: 16 September 2019

1.1 Approval Signatures

Electronic signatures can be found on the last page of this document.

2.0 TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

3.0 LIST OF ABBREVIATIONS

AE	adverse event
ANC	absolute neutrophil count
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CR	complete response
Del	deletion
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture
EOT	End of Treatment (visit)
EQ-5D	EuroQol 5-Dimensional Health Questionnaire
FA	final analysis
FDA	United States Food and Drug Administration
GCP	Good Clinical Practice
HR	Hazard Ratio
HU	health utilization
IA	interim analysis
IDMC	independent data monitoring committee
IMiD	immunomodulatory drug
IMWG	International Myeloma Working Group
ITT	intent-to-treat
ixa+dex	ixazomib+dexamethasone (investigational study therapy, given in Arm A)
K-M	Kaplan-Meier
MedDRA	Medical Dictionary for Regulatory Activities
MID	minimally important difference
NCI	National Cancer Institute
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
ORR	overall response rate
OS	overall survival
PD	progressive disease (disease progression)
PFS	progression-free survival
PK	pharmacokinetic(s)
PN	peripheral neuropathy
pom+dex	pomalidomide+dexamethasone (control study therapy; given in Arm B)
PR	partial response
PRO	patient-reported outcome
PSMB1	Proteasome subunit beta type-1
QLQ	Quality of Life Questionnaire (EORTC)

QOL	quality of life
QTc	rate-corrected QT interval (millisec) of electrocardiograph
RRMM	relapsed and/or refractory multiple myeloma
SAE	serious adverse event
SAP	statistical analysis plan
sCR	stringent complete response
SD	stable disease
T	translocation
TTP	time to progression
VGPR	very good partial response
WHO	World Health Organization

4.0 OBJECTIVES

4.1 Primary Objective

The primary objective is to compare the effect of ixazomib+dexamethasone (ixa+dex) versus pomalidomide+dexamethasone (pom+dex) on progression-free survival (PFS) in Relapsed and/or Refractory Multiple Myeloma (RRMM) patients who have received at least 2 prior lines of therapy, including lenalidomide and a proteasome inhibitor, and are refractory to lenalidomide but not refractory to proteasome inhibitors.

4.2 Secondary Objectives

Secondary objectives are:

- To compare overall survival (OS) in patients treated with ixa+dex versus pom+dex.
- To compare duration of response, overall response rate (ORR), time to response, and time to progression with ixa+dex versus pom+dex.
- To obtain health-related quality of life (QoL) data related to physical functioning of patients treated with ixa+dex versus pom+dex.
- To assess health-related QoL by additional function and symptom domains of the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ) Core 30 (C30) instrument and by the EORTC QLQ multiple myeloma module (MY20) and EuroQol 5-Dimensional Health Questionnaire (EQ-5D-5L) instruments in patients receiving ixa+dex versus those receiving pom+dex.
- To evaluate health utilization (HU) by patients receiving ixa+dex versus those receiving pom+dex.
- To collect plasma concentration-time data for ixazomib to contribute to population pharmacokinetic (PK) characterization of ixazomib and to conduct exposure-response analyses for patients receiving ixa+dex.

4.3 Safety Objective

The safety objective is to compare the safety/tolerability of ixa+dex to that of pom+dex.

4.4 Exploratory Objective

CCI [redacted]

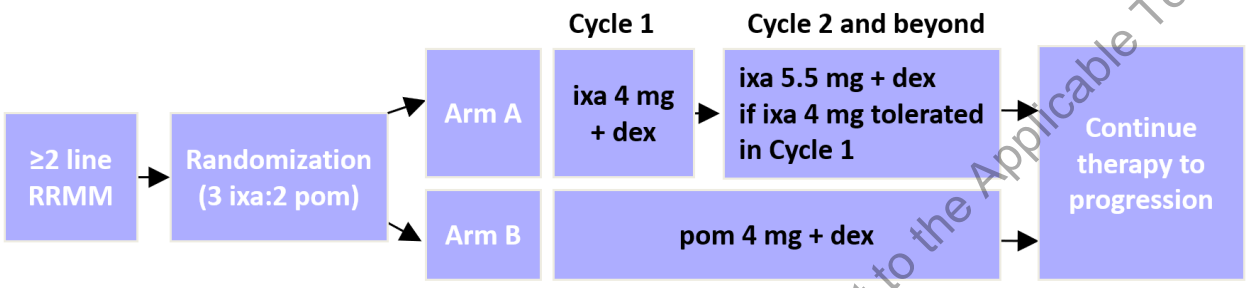
4.5 Study Design

This is a prospective open-label, randomized, 2-arm, multicenter phase 2 study of patients with RRMM who have received at least 2 prior lines of therapy. The 3 stratification factors are International Staging System stage (I or II vs III at study entry), prior lines of therapy (2 vs 3 or more), and age (<65 vs ≥65 years). Patients will be randomized to receive ixa+dex (Arm A) or

pom+dex (Arm B) in a 3:2 ratio until first confirmed progressive disease (PD) or unacceptable toxicities.

The study design is illustrated in Figure 4.a.

Figure 4.a Study Schema, from Randomization Through End of Therapy



5.0 ANALYSIS ENDPOINTS

5.1 Primary Endpoint

The primary endpoint is PFS, defined as the time from randomization to the first occurrence of confirmed PD, as evaluated by the investigator, according to International Myeloma Working Group (IMWG) criteria, or death from any cause, whichever occurs first.

5.2 Secondary Endpoints

Secondary endpoints are:

- OS, measured as the time from randomization to death from any cause.
- ORR, defined as partial response (PR), very good PR (VGPR), or complete response (CR), as evaluated by an investigator according to IMWG criteria.
- Duration of response, defined as the time from the first documentation of PR or better to first documentation of PD.
- Time to response, defined as the time from randomization to the first documentation of PR or better.
- Time to progression (TTP), defined as the time from randomization to first documentation of PD.
- Health-related QOL as measured by the physical domain of the EORTC QLQ-C30.
- Health-related QOL as measured by other domains of the EORTC QLQ-C30, by the EORTC QLQ-MY20, and by the 5-level classification system of the EQ-5D-5L.
- HU as measured by the number and duration of medical encounters.

5.3 Safety Endpoint

The safety endpoint is the safety/tolerability of ixa+dex versus pom+dex.

5.4 Exploratory Endpoint

CCI

[redacted]

6.0 DETERMINATION OF SAMPLE SIZE

Approximately 120 patients will be enrolled. The primary endpoint is PFS with assumption of an HR of 0.62 (median PFS, 7.3 months for ixa+dex vs 4.5 months for pom+dex). The analysis of PFS will be performed on the basis of approximately 81 PFS events in total (for 80% power at a 2-sided 0.20 level of significance).

The PFS assumption for pom+dex is based on the PFS of 3.9 months (with a 95% CI of 3.5-4.6) for the subgroup of patients in the NIMBUS study with 2 prior lines who were lenalidomide refractory, with a higher value of 4.5 months chosen to be on the conservative side. The PFS assumption for ixa+dex is based on the PFS of 9.6 months in a similar population (2 prior lines, lenalidomide refractory), calculated on the basis of data from the [redacted] study as discussed in the protocol, Because this latter study had a smaller population and the PFS had a wider 95% CI (4.6-16.3) than the NIMBUS study, a more conservative value of 7.3 months was chosen.

7.0 METHODS OF ANALYSIS AND PRESENTATION

7.1 General Principles

All statistical analyses will be conducted using SAS® Version 9.2, or higher.

Where appropriate, variables will be summarized descriptively by study visit. For the categorical variables, the count and proportions of each possible value will be tabulated by treatment group. The denominator for the proportion will be based on the number of subjects who provided non-missing responses to the categorical variable. For continuous variables, the number of subjects with non-missing values, mean, median, SD, minimum, and maximum values will be presented by treatment. The Kaplan-Meier survival curves and 25th, 50th (median), and 75th percentiles will be provided along with their 95% CIs for time-to-event data.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations (SDs) will be presented to 2 more decimal places than the recorded data. Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate.

All confidence intervals, statistical tests, and resulting P-values will be reported as 2-sided and will be assessed at α=0.05 significance level unless otherwise stated. P-values will be rounded to 3 decimal places prior to assessment of statistical significance.

Baseline values are defined as the last observed value before the first dose of study medication.

7.1.1 Study Definitions

7.1.2 Definition of Study Days

Study Day 1 is defined as the date on which a subject is administered their first dose of the medication. Other study days are defined relative to the Study Day 1 with Day 1 being Study Day 1 and Day -1 being the day prior to Study Day 1.

7.1.3 Definition of Study Visit Windows

All data will be categorized on the basis of the scheduled visit at which they are collected. These visit designators are predefined values that appear as part of the visit tab in the electronic case report form (eCRF).

7.1.4 Conventions for Missing Adverse Event Dates

Adverse events with start dates that are completely or partially missing will be imputed as follows:

- If month and year are known but day is missing.
  - If month and year are the same as month and year of first dose date, then impute to first dose date.

  - If month and year are different than month and year of first dose date, then impute to first date of the month.
- If year is known but day and month are missing.
  - If year is same as year of 1st dose date, then 1st dose date will be used instead.
  - If year is different than year of 1st dose date, then 1st of January of the year will be imputed.
  - If all is missing, then it is imputed with 1st dose date.

Imputing missing AE start date is mandatory. After the imputation, all imputed dates are checked against the stop dates to ensure the start date does not occur after stop date. If the imputed start date occurs after stop date, then keep the imputed date same as the stop date. Adverse events with stop dates that are partially missing will be imputed as follows:

- If month and year are known but day is missing, the last day of the month will be imputed
- If year is known, but day and month are missing,
  - If YYYY < year of last dose, then 31st of December will be imputed.
  - If YYYY = year of last dose, then 31st of December will be imputed.
  - If YYYY > year of last dose, then 1st of January will be imputed.

Imputing missing AE stop date is not mandatory if AE is regarded as ongoing. However if it is to be done, the rules are outlined above. If subject dies, then use death date for AE stop date.

After the imputation, all imputed dates are checked against the start dates to ensure the stop date does not occur before start date. If the imputed stop date occurs prior to start date, then keep the imputed date the same as the start date.

7.1.5 Conventions for Missing Concomitant Medication Dates

Concomitant therapies with start dates that are completely or partially missing will be analyzed as follows:

- If month and year are known, but day is missing, then impute day to first of the month.
  - If year is known, but day and month are missing, then 1st of January of the year will be imputed.
- If all is missing, then impute date to Date of Birth (DOB).
  - If DOB is not available but age is available, then estimate DOB by using screening date and age (age = screening date – DOB).

Concomitant therapies with stop dates that are completely or partially missing will be analyzed as follows:

- If “ongoing” is checked, no imputation is necessary.

- If month and year are known but day is missing, the last day of the month will be imputed.
- If year is known, but day and month are missing,
  - If YYYY < year of last dose, then 31st of December will be imputed.
  - If YYYY = year of last dose, then 31st of December will be imputed.
  - If YYYY > year of last dose, then 1st of January will be imputed.
- If all is missing, then impute date to 31st of December in the year of last dose.

Imputing missing concomitant therapies is optional. However if it is to be done, the rules are outlined above. If subject dies, then use death date for concomitant therapies stop date. After the imputation, all imputed dates are checked against the start dates to ensure stop date does not occur before start date. If the imputed stop dates occurs prior to start date, then keep the imputed date same as the start date.

The imputation rules for missing/partial start dates of subsequent therapies are as follows:

- When month and year are present and the day of the month is missing,
  - If the month and year of the start date are the same as the month and year of treatment termination, the day of treatment termination is imputed, whichever is earliest.
  - If the start month and year are not the same as the month and year of treatment termination, the first day of the month is imputed.
- When only a year is present, or no components of the start date are present, the date will not be imputed.

7.2 Analysis Sets

7.2.1 Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all patients who are randomized. Patients will be analyzed according to the treatment they were randomized to receive, regardless of any errors of dosing.

The ITT population will be used for the primary, secondary efficacy analyses, and resource utilization and patient reported outcome analysis.

7.2.2 Safety Population

The safety population is defined as all patients who receive at least 1 dose of any study drug. Patients will be analyzed according to the treatment they received. That is, those patients who are randomized to the ixazomib arm but receive the regimen in the control arm will be included in the control arm; those patients who are randomized to the control arm but receive the regimen in the ixazomib arm will be included in the active arm for safety analyses. More specifically, patients who received any dose of Ixazomib will be included in the Ixazomib arm, and patients

who did not receive any dose of Ixazomib will be included in the control arm, regardless of their randomized treatment.

Safety population will be used for all safety related analyses such as adverse events (AEs), concomitant medication, laboratory tests, and vital signs.

7.3 Disposition of Subjects

Patient disposition includes the number and percentage of patients for the following categories: patients in each of the study populations, patients discontinued from the treatment, primary reason to discontinue from the treatment, patients on-going on treatment, patients participated in PFS follow-up/OS follow-up, patients discontinued from the study, and primary reason to discontinue from the study. All percentages will be based on the number of patients in the ITT population.

A listing will present data concerning patient disposition.

7.4 Demographic and Other Baseline Characteristics

The ITT population will be used for the summaries by treatment groups in a descriptive fashion.

Demographics

Demographic data at study entry to be evaluated will include age, sex, race, ethnicity, height, weight, and body surface area. Patient enrollment by region and country will also be summarized by treatment group and control group.

Disease Specific History

Characteristics include type of myeloma, Durie-Salmon stage, ISS, evidence of lytic bone and extramedullary disease will be summarized.

Prior therapies will be summarized including the numbers and percentages of patients who had prior systemic therapy, prior transplant, prior surgery, and prior radiation therapy. This table will also include regimens and the number of lines of prior systemic therapies, best hematologic responses to prior systemic therapy, relapsed and/or refractory status to prior systemic therapy, and months since progression from last prior systemic therapy and months since diagnosis. A by-patient listing will also be presented for prior systemic therapy.

Months from diagnosis to the randomization date for each treatment is calculated by

$$\frac{\text{randomization date - date of diagnosis}}{365.25/12}$$

Baseline disease characteristics

Baseline characteristics includes, but are not limited to, Eastern Cooperative Oncology Group (ECOG) performance status, type of myeloma, serum M-protein, urine M-protein, serum involved FLC and its ratio, β_2 - microglobulin by category (i.e, <3.5, ≥3.5 and < 5.5, ≥5.5 mg/L), serum creatinine and its category (≤2, >2 mg/dL), international staging system, creatinine

clearance by category (ie, <30, ≥30 and <60, ≥60 and <90, ≥90 mL/min), lactate dehydrogenase, serum albumin by category (ie, <35, ≥35 g/L), corrected calcium, hemoglobin will be summarized.

A patient's type of myeloma is determined by heavy chain type (IgG, IgA, IgM, IgD, IgE, biclonal, other and unknown) and light chain type (kappa, lambda, and biclonal). In descriptive summaries, Myeloma type will be summarized separately for the heavy chain patients (according to IgG, IgA, IgM, IgD, IgE, biclonal, other) and for the light chain patients (according to kappa or lambda or biclonal).

Creatinine clearance is to be calculated using the Cockcroft-Gault formulas as follows:

For male patients:

$$\text{creatinine clearance} = \frac{(140 - \text{Age[years]}) \times \text{weight[kg]}}{72 \times (\text{serum creatinine[mg/dL]})}$$

For female patients:

$$\text{creatinine clearance} = 0.85 \times \frac{(140 - \text{Age[years]}) \times \text{weight[kg]}}{72 \times (\text{serum creatinine[mg/dL]})}$$

Distribution of stratification factors will also be summarized.

Bone marrow evaluation and Extramedullary Disease Assessment at study entry

The following categories of extent of disease at study entry will be summarized: bone marrow aspirate and bone marrow biopsy (number of patients, % plasma cells, marrow cellularity, Kappa/Lambda ratio), combined % plasma cells in bone marrow aspiration and biopsy (the higher value will be used if both available), skeletal survey results, number and percentage of lytic bone lesions present, number of extramedullary plasmacytoma present, and type of extramedullary plasmacytoma.

7.5 Medical History and Concurrent Medical Conditions

General medical history such as cardiac risk, peripheral neuropathy at study entry will be summarized by treatment groups.

7.6 Medication History and Concomitant Medications

The prior medications will be listed for all patients by treatment groups. Prior regimens will be summarized by PI containing, IMiD containing, corticosteroids containing, cyclophosphamide containing, liposomal doxorubicin containing, platinum containing, akaylator containing, monoclonal containing, busulfan containing, and others.

The concomitant medications will be coded by preferred term using the World Health Organization (WHO) Drug Dictionary. The number and percentage of patients taking concomitant medications from the first dose through 30 days after the last dose of study medication will be tabulated by Anatomical Therapeutic Chemical (ATC) classification

pharmacological subgroup and WHO drug generic term for each treatment group in the safety population. Concomitant medication of antibacterials by indication, concomitant medication of antimetics, and prophylaxis in relation to herpes zoster will be summarized. By-patient listing will also be presented for concomitant medications.

Concomitant procedures will not be coded.

7.7 Study Drug Exposure and Compliance

A summary of drug exposure to Ixazomib/Pomalidomide will be characterized by number of treated cycles, numbers and percentages of patients who had ≥1, ≥2, …, and ≥18 treated cycles, total amount of dose taken, total number of dose taken, extend pf exposure (days), relative dose intensity (%), by each treatment group in the safety population. Summary of numbers and percentages of patients using aggregate treated cycles, such as 1 to 3, 4 to 6, 7 to 9, 10 to 12, 13 to 18, and ≥19 treated cycles, will also be presented in the same table.

Exposure to dexamethasone will be characterized similar to exposure to Ixazomib/Pomalidomide by treatment group.

Extent of exposure (days) is calculated as (last dose date of study drug – first dose date of study drug + 1).

A treated cycle is defined as a cycle in which the patient received any amount of any study drug regimen.

Relative dose intensity (%) is defined as 100 * (Total amount of dose taken) / (Total prescribed dose of treated cycles). Total prescribed dose of treated cycles for Ixazomib is calculated as: for all patients after C2D1, it equals number of prescribed doses per cycles * dose prescribed at enrollment (4 mg) * 1 cycle + dose prescribed at C2D1 (5.5 mg) * number of prescribed doses per cycle * (number of treated cycles - 1).

Relative dose intensity will be displayed as <50%, 50% - <= 80%, 80% - <100%, = 100%, and >100% by treatment group and for all patients.

The duration of treatment at 5.5 mg Ixazomib will be also calculated from the first date when subjects were dosed with 4 mg till either the last dosing date or the first time they had dose reduced, whichever comes earlier.

Dosing data will also be presented in a by-patient listing.

7.8 Efficacy Analysis

All available efficacy data will be included in data listings and tabulations. Data that are potentially spurious or erroneous will be examined according to standard data management operating procedures. In general, missing data will be treated as missing, and no data imputation will be applied, unless otherwise specified.

7.8.1 Primary Efficacy Analysis

The analysis of the primary endpoint, PFS, will be based on the ITT population using investigator-assessed confirmed PD data using unstratified analysis. PFS is defined as the time from randomization to the first occurrence of confirmed PD or death from any cause, whichever occurs first. Patients without documentation of PD will be censored at the date of the last response assessment that is SD or better. Per IMWG criteria, 2 consecutive response assessments are required to document PD.

The null and alternative hypotheses for PFS are:

H_0: PFS in ixa+dex Arm = PFS in pom+dex Arm

H_a: PFS in ixa+dex Arm > PFS in pom+dex Arm

A 2-sided, unstratified log-rank test will be used to compare the treatment groups with respect to PFS at a 2-sided alpha level of 0.05. In addition, an unadjusted unstratified Cox model will be used to estimate the HR and its 95% CIs for the treatment effect. The Kaplan-Meier (K-M) survival curves and K-M medians (if estimable), along with their 2-sided 95% CIs, will also be provided for each treatment group.

The details regarding the handling of missing assessment and censoring for PFS analysis are presented in Table 7.a.

Table 7.a Handling of Missing Assessment and Censoring for PFS Primary Analysis Based on FDA Guidance

Situation	Date of Progression or Censoring	Outcome
No baseline and/or no post baseline assessment, no subsequent anticancer therapy after study treatment, no death	Date of Randomization	Censored
Disease progression documented between scheduled visits	Date of documented disease progression	Event
No documented death or disease progression	Date of last adequate assessment*	Censored
Lost to follow-up, withdraw consent before any documented death or disease progression	Date of last adequate assessment*	Censored
Death or progression after more than one missed visit	Date of last adequate assessment*	Censored
Alternate antineoplastic therapy started prior to disease progression	Date of last adequate assessment prior to starting alternate antineoplastic therapy	Censored
Death before first assessment	Date of death	Event
Death between adequate assessment visits	Date of death	Event

*Adequate disease assessment is defined as there is sufficient data to evaluate a patient's disease status.

Sensitivity analyses for PFS will include:

- PFS assessed by stratified analyses in the ITT population.

Subgroup analyses will be performed for PFS relative to baseline stratification factors and demographic data such as sex, race or ethnic group, and age, as appropriate.

Table 7.b Handling of missing assessment and censoring for PFS Sensitivity Analysis based on EMA guidance

Situation	Date of Progression or Censoring	Outcome
Alternate antineoplastic started prior to disease progression	Date of documented disease progression	Event
Death or disease progression after more than one missed visit	Date of death or disease progression	Event

The plan of subgroups for PFS is presented in the Table 7.c below:

Table 7.c List of subgroups

Subgroup	Definition of Group
Age	<65 vs ≥65
Sex	male vs female
Race	white, black-African American, Asian, other
ISS stage	I or II vs. III
Region	Europe vs other region
Lines of prior therapies	2 vs 3 or more
Prior Velcade therapy	Exposed vs naïve
Refractory to last line of prior therapy	yes vs no
Relapsed and/or refractory	Refractory vs Relapse and refractory
Renal function based on baseline creatinine clearance	<60 mL/min, and ≥60 mL/min
Liver Function based on baseline ALT or AST	≥1.5× ULN vs. <1.5 × ULN
ECOG performance status	0 or 1 vs 2

Additional exploratory analyses may be performed if deemed necessary.

7.8.2 Secondary Efficacy Analysis

The primary endpoint of PFS will be supported by prespecified evidence of clinical benefit as measured by other secondary endpoints. Other secondary efficacy parameters are OS, ORR, duration of response, time to response, and TTP. Disease response–related endpoints will be analyzed using investigator-assessed response rate for the ITT population. Unstratified analyses will be conducted.

Overall Survival

The OS will be analyzed on the basis of the ITT population and is defined as the time from randomization to death from any cause. Patients without documentation of death at the time of analysis will be censored at the date last known to be alive. A 2-sided, unstratified log-rank test will be used to compare the treatment and control groups with respect to OS. In addition, an

unadjusted unstratified Cox model will be used to estimate the HR and its 80% and 95% CIs for the treatment effect. OS will be also tested at 2-sided alpha=0.2. The KM survival curves and KM medians (if estimable), along with their 95% CIs, will also be provided for each study group.

Overall Response Rate

The ORR is defined as the proportion of patients who achieved PR or better relative to the ITT population. A logistic regression model will be used to estimate the treatment effect in terms of odds ratio. The odds ratio and its associated 95% CIs will be presented.

Time to Response

Time to response is defined as the time from randomization to the first documentation of PR or better. Time to response will be compared in the ITT population and summarized descriptively for the responders.

Duration of Response

Duration of response is defined as the time from the date of first documentation of a PR or better to the date of first documentation of PD for responders. Responders without documentation of PD will be censored at the date of their last response assessment that is SD or better. Duration of response will be summarized descriptively using the K-M method.

Time to Progression

Time to progression is defined as the time from the date of randomization to the date of first documentation of PD. Patients without documentation of PD at the time of the analysis will be censored at the date of their last response assessment that is SD or better. Time to progression will be analyzed in the ITT population using a similar method as PFS.

7.9 Pharmacokinetic/Pharmacodynamic Analysis/Biomarker Analysis

7.9.1 Pharmacokinetic Analysis

Ixazomib plasma concentration-time data will be presented in listings and summarized by time point.

PK data collected in this study may contribute to population PK and exposure/response (safety and efficacy) analyses. These analyses may include data from other ixazomib clinical studies. The analysis plan for the population PK and exposure/response analyses will be separately defined and the results of these analyses will be reported separately.

7.9.2 Biomarker analysis

CCI

[redacted]

CCI

[redacted]

7.10 Other Outcomes Analyses of Patient-Reported Outcomes and Health Economics

7.10.1 Patient Reported Outcomes (PROs)

Health-related QOL will be assessed using the cancer-specific EORTC QLQ-C30 (Table 7-4), which contains 30 items across 5 functional scales, 9 symptom scales and a global health status/QOL scale. Items 1-28 have 4 response levels (not at all, a little, quite a bit, and very much) and items 29 and 30 rely on a 7-point numeric rating scale. A summary score of EORTC QLQ-C30 will be calculated from the mean of 13 of the 15 EORTC QLQ-C30 subscales (the Global health status/Quality of Life scale and the Financial Difficulties scale are not included).

The EORTC QLQ-MY20 (Table 7-5) has 20 items across 2 functional subscales and 2 symptoms scales.

Raw scores are converted into scale scores ranging from 0 to 100. For the functional subscales and the global health status/QOL subscale, higher scores represent better QOL; for the symptom subscales, lower scores represent better QOL.

Descriptive Analyses

Patient-reported outcome (PRO) assessments using the EORTC QLQ-C30 and EORTC QLQ-MY-20 will be analyzed using patients with PRO measurements at baseline and at least one post baseline measurement in the ITT population.

The descriptive statistics of actual value and change from baseline for the subscale scores and summary score of EORTC QLQ-C30 and subscale scores of EORTC QLQ-MY20 will be summarized by treatment group over time. Particular emphasis will be placed on the Physical Functioning subscale of EORTC QLQ-C30.

Table 7.d Definition of Subscale Scores of EORTC QLQ-C30

Subscale	Individual Items
Global health status/ Quality of life	29-30
Functional scales	
Physical functioning	1-5
Role functioning	6-7
Emotional functioning	21-24
Cognitive functioning	20, 25
Social functioning	26-27
Symptom scales/items	
Fatigue	10, 12, 18
Nausea and vomiting	14-15
Pain	9, 19
Dyspnea	8
Insomnia	11
Appetite loss	13
Constipation	16
Diarrhea	17
Financial difficulties	28

Table 7.e Definition of Subscale Scores of EORTC QLQ-MY20

Subscale	Individual Items
Functional scales/items	
Future perspective	18-20
Body image	17
Symptom scales	
Disease symptoms	1-6
Side effects of treatment	7-16

Analysis based on minimally important difference (MID)

For the summary score and each subscale score of EORTC QLQ-C30 as well as each subscale score of EORTC QLQ-MY20, the number and percentage of patients with an improvement in score from baseline based on minimally important differences (MIDs) of 10 (primary analyses)[1,2,3] and 5 (sensitivity analyses)[4] will be summarized by treatment group over time.

[1] Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16:139–44.

Specifically, patients with a change from baseline for the better of ≥MID will be classified as "improved".

Analysis of time to maintained quality of life deterioration

For the summary score and each subscale score of EORTC QLQ-C30 as well as each subscale score of EORTC QLQ-MY20, deterioration will be defined as a worsening in score from study entry ≥10 (primary analyses) and ≥5 (sensitivity analyses). Maintained deterioration will be defined as two consecutive scores that meet the definition of deterioration. Time to maintained QoL deterioration will be defined as the time from the date of randomization to the date of the second of the two consecutive scores that meet the definition for deterioration. Patients without any QoL deterioration or with only one score that meets the definition of deterioration will be censored at the date of last QoL measurement.

A 2-sided stratified log-rank test will be used to compare the treatment groups with respect to time to maintained QoL deterioration. In addition, an unadjusted, stratified Cox model will be used to estimate the hazard ratio and its 95% CI. The Kaplan-Meier survival curves will also be provided for each treatment group.

Analysis based on linear mixed effects models

For the summary score and each subscale score of EORTC QLQ-C30 as well as each subscale score of EORTC QLQ-MY20, the change from baseline score to each scheduled treatment cycle visit will be analyzed using linear mixed models. These models will include the following covariates: treatment group, baseline score, stratification factors (International Staging System stage, prior lines of therapy, and age), visits, and interactions between treatment group and visits. The interaction term between baseline score and visit may also be considered as a covariate. The estimated mean change from baseline score with 95% CIs for each treatment group will be provided at each treatment cycle visit. In addition, the mean difference in the changes from baseline between the treatment groups with 95% CIs and p-values will be provided at each treatment cycle visit.

Missing data

Details of scoring and initial handling of missing data are included in the EORTC QLQ-C30 and EORTC QLQ-MY20 scoring guidelines.

---

[2] Cella D, Pickard AS, Duh MS, et al. Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomized phase III trial. Eur J Cancer. 2012 Feb;48(3):311-23.

[3] Kyam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. *Eur J Haematol.* 2011;87(4):330-337

[4] Stewart AK, Dimopoulos MA, Masszi T, et al. Health-Related Quality-of-Life Results From the Open-Label, Randomized, Phase III ASPIRE Trial Evaluating Carfilzomib, Lenalidomide, and Dexamethasone Versus Lenalidomide and Dexamethasone in Patients With Relapsed Multiple Myeloma. J Clin Oncol. 2016 Nov 10;34(32):3921-3930.

Sensitivity analyses may be conducted to study the impact of missing data.

Compliance for EORTC QLQ-C30 and EORTC QLQ-MY20 will also be summarized by number expected and number and percentage received by treatment group over time.

7.10.2 Health Economics Analysis Using Medical Resource Utilization and Utility

The EQ-5D-5L consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) each rated on 5 levels. The visual analogue scale (VAS) ranges from 0 (worst imaginable health state) to 100 (best imaginable health state).

EQ-5D-5L item scores and VAS scores will be summarized in descriptive statistics for treatment arms over time.

Compliance of EQ-5D-5L will also be summarized by treatment group over time.

HU data will be summarized in descriptive statistics of medical encounters (length of stay, inpatient, outpatient, and reason), number of missing days from work (by patient and care-giver) for treatment arms.

7.11 Safety Analysis

7.11.1 Adverse Events

Adverse events

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA). All AEs will be presented in a by-patient listing. Treatment-emergent AEs are AEs that occur after administration of the first dose of any study drug and through 30 days after the last dose of any study drug.

AEs will be tabulated according to the MedDRA by system organ class, high level terms and preferred terms and will include the following categories:

- Treatment-emergent AEs.
- Drug-related treatment-emergent AEs.
- Grade 3 or higher treatment-emergent AEs (also report Grade 3 and 4 separately).
- Grade 3 or higher drug-related treatment-emergent AEs (also report Grade 3 and 4 separately).
- The most commonly reported treatment-emergent AEs (i.e., those events reported by ≥10% of patients in either treatment group).
- Serious AEs (SAEs).
- Drug related SAEs.
- AEs that prevented dose escalation at C2.

Patients with the same AE more than once will have that event counted only once within each body system, once within each high level term, and once within each preferred term.

Drug-related treatment-emergent AEs will also be summarized by the National Cancer Institute Common Toxicity Criteria (NCI CTCAE) version 4.03. Patients with the same AE more than once will have the maximum intensity of that event counted within each body system, once within each high level term, and once within each preferred term.

The most commonly reported treatment-emergent AEs (ie, those events reported by ≥10% of any treatment arm) will be tabulated by preferred term. Patients with the same AE more than once will have that event counted only once within each preferred term.

An overall summary AE table will include numbers and percentages of patients who had at least one AE, drug-related AE, grade 3 or higher AE (also grade 3 and 4 AE respectively), grade 3 or higher drug-related AE (also grade 3 and 4 drug-related AE respectively), SAEs, drug-related SAE, AE resulting in discontinuation, and on-study deaths. On-study death is defined as the death that occurs between the first dose of any study drug and within 30 days of the last dose of any study drug.

All concomitant medications collected from screening through the study period will be classified to preferred terms according to the World Health Organization (WHO) drug dictionary.

Additionally, by-patient listings and summary tables of the AE of special interest (AESI) and AEs of clinical importance (AECI) will be presented including New primary malignancy, Peripheral neuropathy, Rash, Encephalopathy, Liver impairment, Hypotension, Heart failure, Arrhythmias, Myocardial infarction, Thrombocytopenia, Neutropenia, Gastrointestinal, and Renal impairment.

Incidence of New Primary Malignancies (NPM)

Two types of incidence rates will be calculated for the safety population based on the new primary malignancy assessment:

- Incidence proportions, defined as the percentage of the subjects reporting any new primary malignancy in the safety population with available information.
- Incidence rates, defined by the number of the subjects reporting any new primary malignancy divided by the total duration of follow-up (patient-years = pt-yrs) in the safety population with available information up to the onset of new primary malignancies.

For incidence proportions, the relative risks, defined as the ratio of incidence proportions between the 2 randomized treatment groups, were provided along with their 95% CIs. For incidence rates, the relative risks, along with their 95% CIs, will be calculated using an exponential regression model for lifetime data (assuming constant hazards).

Due to the distinct nature of hematologic and nonhematologic neoplasms, as well as the emerging signals of new primary malignancies for immunomodulating agents, analyses of new primary malignancies may be performed separately for hematologic and nonhematologic malignancies.

Serious Adverse Events

The number and percentage of patients experiencing at least one treatment-emergent SAE will be summarized by MedDRA primary system organ class, high level term, and preferred term. Drug-related SAEs will be summarized similarly.

In addition, a by-patient listing of the SAEs will be presented (the patient listing will contain all SAEs regardless of treatment-emergent AE status).

Deaths

A by-patient listing of the deaths will be presented. All deaths occurring on-study will be displayed (regardless of treatment-emergent AE status).

Adverse Events Resulting in Discontinuation of Study Drug

A by-patient listing of treatment-emergent AEs resulting in discontinuation of study drug will be presented.

7.11.2 Clinical Laboratory Evaluations

For the purposes of summarization in both the tables and listings, all laboratory values will be converted to standardized units. If a lab value is reported using a non-numeric qualifier (eg, less than (<) a certain value, or greater than (>) a certain value), the given numeric value will be used in the summary statistics, ignoring the non-numeric qualifier. However, for the bone marrow plasma cell percentage, the convention as (x-1) % (mainly for < 5% for CR) will be used.

Laboratory test results from the central laboratory will be used when they are available. Laboratory test results from local laboratory will only be used when no central laboratory test results exist at the same scheduled sample collection time point.

If a patient has repeated laboratory values for a given time point, the value from the last evaluation will be used.

Laboratory test results will be summarized according to the scheduled sample collection time point. Change from baseline will also be presented. Unscheduled laboratory test results will be listed and included in laboratory shift tables. The parameters to be analyzed are as follows:

- Hematology: hemoglobin, neutrophils (ANC), platelets counts, and leukocytes with differential.
- Serum chemistry: blood urea nitrogen (BUN), creatinine, total bilirubin, urate, lactate dehydrogenase (LDH), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), β2-microglobulin, calcium, potassium, C-reactive protein, magnesium, and gamma glutamyl transferase (GCT).

Shift tables will be constructed for laboratory parameters to tabulate changes in NCI CTCAE for toxicity (version 4.03) from study entry to post study entry worst CTC grade. Parameters to be tabulated will include:

- Hematology: ANC, hemoglobin, platelets,

- Serum chemistry: ALT, AST, ALP, creatinine, total bilirubin, calcium, magnesium, potassium.

Summary statistics will also be presented for shift from urinalysis values at study entry.

- Urinalysis: Turbidity and color, pH, specific gravity, protein, ketones, bilirubin, occult blood, nitrite, urobilinogen, glucose, leukocytes, microscopic analysis.

Mean laboratory values and box plots over time for key lab parameters will be produced, including but not limited to ANC, platelets, and liver function tests (ALT/SGPT, AST/SGOT, alkaline phosphatase, and total bilirubin).

By-patient listings to be presented include hematology, serum chemistry, urinalysis, urine total protein, and urine creatinine.

7.11.3 Vital Signs

The actual values of vital sign parameters including temperature, blood pressure, heart rate, and body weight, will be summarized over time for each treatment arm. Change from baseline will also be presented.

A by-patient listing will also be presented.

7.11.4 12-Lead ECGs

Descriptive statistics for the actual values and changes from baseline in ECGs will be listed by time point.

Rate-corrected QT interval (millisec) of electrocardiograph (QTc) interval will be calculated using Bazett's correction and Fridericia's correction, if necessary. The formulas are:

$$QTc\ (Bazett) = QT / (RR^{0.5})$$

$$QTc\ (Fridericia) = QT / (RR^{0.33})$$

where RR = 60 / heart rate (bpm)

In addition, a categorical analysis of QTc intervals will be performed for each time point. The number and percentage of patients in each QTc interval (< 450 msec, 450-480 msec, > 480-<500 msec, and ≥ 500msec) will be summarized at baseline and each of the subsequent time points. Categories of changes from baseline (≥ 30 msec and ≥ 60 msec) will be summarized as well. Maximum QTc intervals and maximum changes from baseline will also be summarized similarly in a separate display.

ECG abnormalities will be presented in a data listing.

7.11.5 Other Observations Related to Safety

Eastern Cooperative Oncology Group performance status and shifts from baseline to post-baseline assessment over time, and ECOG score frequency table over time will be summarized. Shifts from baseline to the worst post-baseline score will be tabulated by treatment arm.

Pregnancy testing results will be presented in a by-patient listing.

Additional safety analyses may be performed to most clearly enumerate rates of toxicities and to further define the safety profile of Ixazomib, e.g. analyses of TEAEs of clinical importance.

Type I Error Control

Type I error is controlled by the primary endpoint, because only 1 test is planned for primary endpoint of PFS at a 2-sided level of 0.20.

7.12 Changes in the Statistical Analysis Plan

In protocol amendment 05, even though it was specified that the unstratified analysis will be implemented for PFS and OS, in the description, it mentioned that unadjusted stratified Cox model will be used to estimate the HR. It is clarified in the SAP that unstratified log-rank test as well as unadjusted unstratified Cox model will be used for both PFS and OS analyses.

8.0 REFERENCES

1. Durie BG, Harousseau JL, Miguel JS, Blade J, Barlogie B, Anderson K, et al. International uniform response criteria for multiple myeloma. Leukemia 2006; 20(9): 1467-73.
2. Robins JM and Finkelstein DM. Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests. Biometrics 2000; 56: 779-788.
3. Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics – Theory and Methods 1991; 20: 2609–2631.
4. D'Agostino R.B., Lee M.L., Belanger A.J., Cupples L.A., Anderson K., and Kannel W.B. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Statistics in Medicine, 9(12): 1501–1515, 1990.
5. Xie J. and Liu C. (2005). Adjusted Kaplan-Meier Estimator and Log-Rank Test with Inverse Probability of Treatment Weighting for Survival Data. Statistics in Medicine, 10;26(10):2276, 2007.

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
PPD	Biostatistics Approval	13-Nov-2019 16:18 UTC