Title: Phase 3, Randomized, Placebo-Controlled, Double-Blind Study of Oral Ixazomib Maintenance Therapy After Initial Therapy in Patients With Newly Diagnosed Multiple Myeloma Not Treated With Stem Cell Transplantation

NCT Number: NCT02312258

SAP Approve Date: 20 September 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

A Phase 3, Randomized, Placebo-Controlled, Double-Blind Study of Oral Ixazomib Maintenance Therapy After Initial Therapy in Patients With Newly Diagnosed Multiple Myeloma not Treated With Stem Cell Transplantation

Protocol #: C16021

SAP Version:
Final 1.0

Date of Statistical Analysis Plan:
20 September 2019

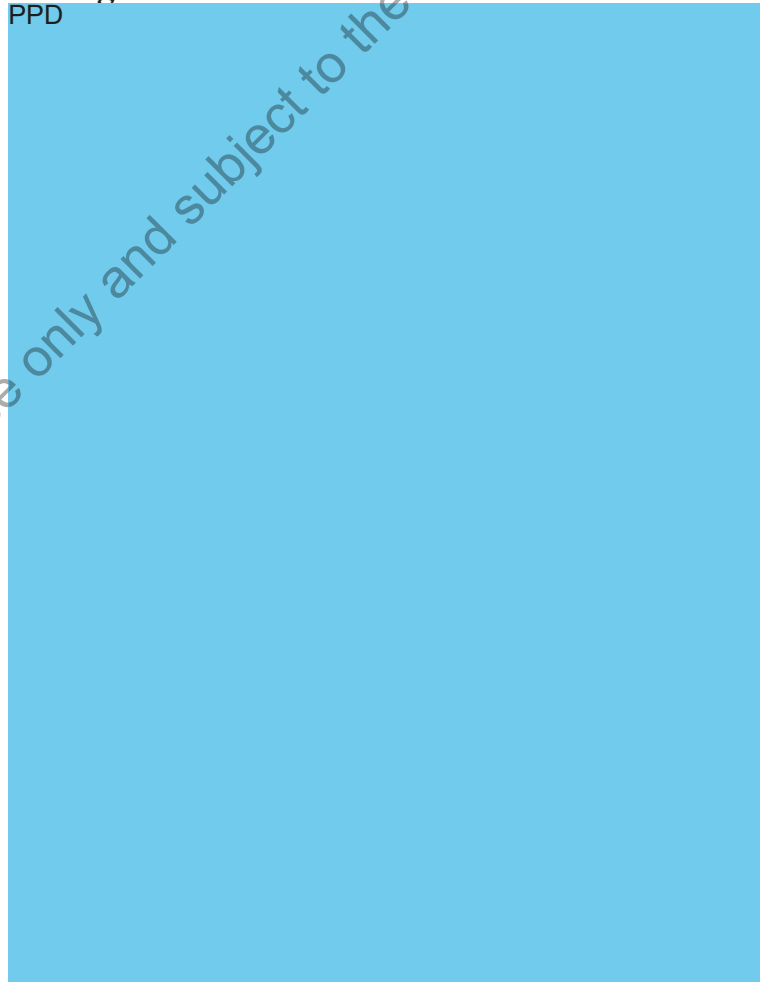## Approval Signatures

PPD

_____
Date

_____
Date

_____
Date

_____
Date

_____
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS AND GLOSSARY OF TERMS

Abbreviation	Term
AE	adverse event
AECI	adverse event clinical importance
AESI	adverse event special interest
ALC	absolute lymphocyte count
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BM	bone marrow
BMA	bone marrow aspirate
BUN	blood urea nitrogen
CDF	cumulative distribution function
CI	confidence interval
CO_2	carbon dioxide
CR	complete response
del	deletion
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC QLQ-MY20	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Multiple Myeloma Module
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
EOT	End of Treatment (visit)
EQ-5D	EuroQol 5-Dimensional Health Questionnaire
EQ-5D-5L	5-Level classification system of the EuroQol 5-Dimensional Health Questionnaire
FA	final analysis
FLC	free light chain
HR	hazard ratio
HRQOL	health related quality of life
HU	health utilization
IA	interim analysis
IDMC	independent data monitoring committee
IMiD	immunomodulating drugs
IMWG	International Myeloma Working Group
IPCW	Inverse Probability of Censoring Weighted
IRC	independent review committee
ISC	independent statistical center

Abbreviation	Term
ISS	International Staging System
ITT	intent-to-treat
IXRS	interactive web/voice response system
LDH	lactate dehydrogenase
LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
MID	minimally important difference
MM	multiple myeloma
MRD	minimal residual disease
MSM	marginal structural model
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NDMM	newly diagnosed multiple myeloma
NEC	not elsewhere classified
NFκB	nuclear factor kappa-light-chain-enhancer of activated B cells
NK	natural killer (cells)
NPM	new primary malignancy
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PFS2	time from randomization to objective disease progression on next-line treatment or death from any cause
PK	pharmacokinetic(s)
PN	peripheral neuropathy
PP	per-protocol
PR	partial response
PRO	patient-reported outcome
RISS	Revised International Staging System
SAE	serious adverse event
SAP	statistical analysis plan
SCT	stem cell transplant/therapy
SD	stable disease
TEAE	treatment-emergent adverse event
TTNT	time to next-line therapy
TTP	time to progression
VGPR	very good partial response
VZV	varicella-zoster virus
WBC	white blood cell (count)
WHO	World Health Organization

1. INTRODUCTION

In general, the purpose of the statistical analysis plan (SAP) is to provide a framework that addresses the protocol objectives in a statistically rigorous fashion, with minimized bias and analytical deficiencies. Specifically, this plan has the following purpose:

To prospectively (a priori) outline the types of analyses and data presentations that will address the study objectives outlined in the protocol, and to explain in detail how the data will be handled and analyzed, adhering to commonly accepted standards and practices of biostatistical analysis in the pharmaceutical industry.

1.1 Study Design

This is a phase 3, randomized, double-blind, placebo-controlled, multicenter study in patients with newly diagnosed multiple myeloma (NDMM) not receiving stem cell transplantation (SCT). Patients who have not undergone SCT may not have done so because of frailty due to advanced age (eg, ≥ 65 years) or comorbidity or because they declined SCT for other reasons.

Patients must have received initial therapy, for 6 to 12 months, according to standard of care before study enrollment and have been treated to a major response category (partial response [PR] or better) that is judged to be their best response by the investigator/treating physician. Partial response, very good partial response (VGPR), or complete response (CR) must be documented at screening, and patients must have met all additional inclusion/exclusion criteria. Eligible and consenting patients are to be randomized no later than 60 days after the last dose of initial therapy. Randomization will occur in a 3:2 ratio of ixazomib or matching placebo. Approximately 700 patients are planned to be enrolled in this study.

There are 4 stratification factors: initial therapy (proteasome-inhibitor–containing or not), International Staging System (ISS) status before initial therapy (stage I or II vs stage III), age (< 75 vs ≥ 75 years) at randomization, and response to initial therapy as measured during screening (CR or VGPR vs PR).

Patients will receive blinded ixazomib or matching placebo capsules (both hereafter referred to as "study drug") orally on Days 1, 8, and 15 of every 28-day cycle. The starting dose will be 3 mg of study drug, which—if tolerated during the first 4 cycles—will be escalated to 4 mg beginning with Cycle 5 Day 1. The treatment period will be approximately 24 months (equivalent to 26 cycles [if no cycle delays], to the nearest complete cycle) or until patients experience PD or unacceptable toxicity, whichever occurs first.

The primary endpoint of progression-free survival (PFS) will be supported by prespecified evidence of clinical benefit as measured by the key and other secondary endpoints. There will be 2 interim analyses (IAs) and 1 final analysis (FA) in the study. The first IA will be the FA for PFS for statistical testing purposes and the first IA for OS. The second IA and the FA are for OS. An independent data monitoring committee (IDMC) will review safety and efficacy data at the IAs and safety data at regularly scheduled meetings. An independent review committee (IRC) will assess disease response or progressive disease (PD) according to the International Myeloma Working Group (IMWG) uniform response criteria, version 2011 [1].

1.2   Study Objectives

The primary objective is:

- To determine the effect of ixazomib maintenance therapy on PFS, defined as the time from randomization to PD or death from any cause, compared with placebo, in patients with NDMM who have had a major response — defined as CR, VGPR, or PR—to initial therapy and who have not undergone SCT

The key secondary objective is:

- To determine the effect of ixazomib maintenance therapy on overall survival (OS) compared with placebo

Other secondary objectives are:

- To determine the effect of ixazomib maintenance therapy on improving best response for patients who enroll in the study at PR or VGPR and on maintaining best overall response for patients who enroll in the study at CR

- To determine the effect of ixazomib maintenance therapy on time to progression (TTP)

- To determine the effect of ixazomib maintenance therapy on progression-free survival 2 (PFS2), defined as the time from randomization to objective disease progression on next-line treatment or death from any cause

- To determine the effect of ixazomib maintenance therapy on the time to next-line therapy (TTNT)

- To determine the effect of ixazomib maintenance therapy on the time to end of next-line therapy

- To determine the effect of ixazomib maintenance therapy on duration of next-line therapy

- To assess the incidence of new primary malignancy in patients receiving ixazomib maintenance therapy

- To evaluate the frequency of conversion from minimal residual disease (MRD) positive to MRD negative, or the maintenance of MRD negativity, using 8-color flow cytometry

- To assess the correlation between MRD status (detected using 8-color flow cytometry) and PFS and OS, using bone marrow aspirates.

- To determine the effects of ixazomib maintenance therapy on PFS and OS in high-risk cytogenetic patient groups characterized by individual or multiple cytogenetic abnormalities including, but not limited to, del17, t(4;14), and t(14;16).

- To determine the long-term safety and tolerability of ixazomib maintenance therapy

- To assess health-related quality of life (HRQOL) as measured by the global health domain of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) in patients who receive ixazomib maintenance therapy

- To assess the correlation between frailty status and PFS and OS in patients receiving ixazomib maintenance therapy

- To collect pharmacokinetic (PK) data to contribute to population PK and exposure-response (safety/efficacy) analysis

- To evaluate the resolution and improvement of peripheral neuropathy (PN), if it occurs, in patients receiving ixazomib maintenance therapy

The exploratory objectives are:

CCI

2. POPULATIONS FOR ANALYSIS

2.1 Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all patients who are randomized with post randomization data available. 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 (PRO) analyses.

2.2 Safety Population

The safety population is defined as all patients who receive at least 1 dose of ixazomib or placebo. Patients will be analyzed according to the treatment they actually receive. Patients who receive any dose of ixazomib will be included in the ixazomib arm, and patients who only receive placebo will be included in the placebo arm, regardless of their randomized treatment.

The safety population will be used for all safety related analyses such as adverse events (AE), concomitant medications, laboratory tests, and vital signs.

2.3 Per-Protocol (PP) population

The PP population consists of all ITT patients who do not violate the terms of the protocol in a way that would affect the study outcome significantly. All decisions to exclude patients from the PP population will be made by the Takeda Project Clinician or designee prior to unblinding the study for IAs or FA purpose.

The PP population will be used as a sensitivity analysis of the ITT population for the primary efficacy endpoint PFS if more than 5% patients are excluded from this analysis.

3. HYPOTHESES AND DECISION RULES

3.1 Statistical Hypotheses

There is one primary endpoint in this study.

The null and alternative hypotheses for PFS are:

H_0: PFS in ixazomib arm = PFS in placebo arm

H_a: PFS in ixazomib arm > PFS in placebo arm

There is one key secondary efficacy endpoint in this study.

The null and alternative hypotheses for OS are:

H_0: OS in ixazomib arm = OS in placebo arm

H_a: OS in ixazomib arm > OS in placebo arm

3.2 Statistical Decision Rules

A closed sequential-testing procedure will be used to test the primary endpoint of PFS and the key secondary endpoint of OS, with the following testing order:

1. At the first IA—PFS as assessed by the IRC in the ITT population (primary endpoint) and PFS as assessed by the IRC in 3 prespecified subgroups: a) ISS stage III; b) patients aged ≥75 years; and c) patients with a response of CR or VGPR to initial therapy;

2.   OS (key secondary endpoint) at the IAs or FA.

OS will be tested at the IAs or FA at the significance level determined by the O'Brien-Fleming alpha spending function (the Lan-DeMets method [2]). The initial alpha allocation and propagation of alpha across ITT PFS, Subgroup PFS, and ITT OS is detailed in the graph below. Because this graph leads to a consonant testing procedure, family-wise type I error rate is strongly controlled for all three endpoints (Bretz and Maurer, 2009[3]).
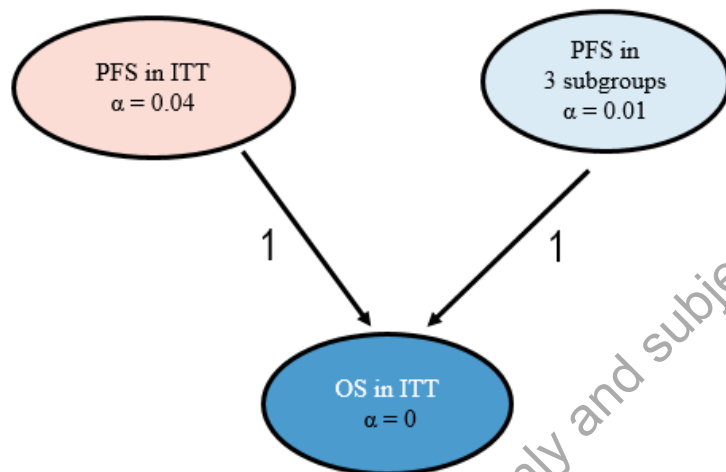
The testing procedure will proceed as follows. First, $\alpha= 0.04$ and $\alpha= 0.01$ will be assigned to test PFS in the ITT population and 3 pre-specified subgroups, respectively. If PFS in ITT test is significant, then $\alpha= 0.04$ originally assigned to PFS in ITT will be propagated (i.e. relocated) to OS test in ITT. Similarly, if PFS is tested significantly in all 3 subgroups, then $\alpha= 0.01$ originally assigned to PFS in subgroups will be propagated to test OS in ITT population. Although the graphical approach allows OS to be tested at $\alpha= 0.01$ in the event that PFS in all 3 subgroups is significant, but PFS in the ITT population is not, the study may be stopped due to the fact that PFS in ITT is not significant.

All other efficacy endpoints will be tested at a 2-sided alpha level of 0.05.

4.   INTERIM ANALYSIS

4.1   Interim Analysis

There are 2 planned IAs. The first IA will be performed when approximately 392 of the IRC-assessed PFS events have occurred or approximately 10 months after the last patient has been enrolled, whichever occurs later. This IA is expected to occur approximately 50

months after the first patient is enrolled. This is the primary analysis and the only analysis of PFS for statistical testing purposes and the first IA for OS.

At the first IA, PFS will be tested in both the ITT population and in 3 prespecified subgroups. The subgroup testing strategy includes 2 major components: a) preservation of the ability to detect the overall treatment effect using a reduced overall significance level of alpha = 0.04, which will be used for the ITT population; and b) test of treatment effect for the 3 prespecified subgroups: 1) ISS stage III, 2) patients aged $\geq$75 years, and 3) CR or VGPR to initial therapy. Subgroup testing will be conducted using the remaining alpha = 0.01 and the Hochberg procedure for multiplicity correction (refer to the Section 3.2 for proof of strong control of the Type I error rate). Because the size of the treatment effect may be substantially greater in a prespecified subgroup than in the overall study population, analysis of patients in each subgroup at a stringent significance level may still provide a statistically significant outcome.

After PFS is tested at the first IA, central efficacy and investigator assessments of disease response for protocol purposes will be discontinued (except for investigator assessment of PFS2).

If the test for PFS is not statistically significant in any population (the ITT or any of the 3 subgroups), there will be no formal hypothesis conducted afterwards.

The second IA will be performed when approximately 206 deaths (approximately 70% of the total expected 295 deaths, the minimal number of events for the OS final analysis) have been observed.

For the testing of OS, alpha spending for the first and second IAs will always be based on the observed events (information fraction) and a total of 295 death events with a different adjustment of the critical value at OS FA testing (Cui-Hung-Wang test statistics [4] used for the primary analysis of OS at FA) based on the following scenarios:

- If PFS in the ITT population is significant, and PFS in at least 1 subgroup is not significant, then OS in the ITT population will be tested using a total alpha of 0.04.

- If PFS in the ITT is significant and PFS in all 3 subgroups is significant, then OS in the ITT population will be tested with a total alpha of 0.05.

- If PFS in the ITT population is not significant, and PFS in at least 1 subgroup is significant, then no formal ITT OS testing will be conducted.

The family-wise error rate for the 4 null hypotheses for PFS and the 1 hypothesis for OS for the overall study population is controlled using a prespecified, 2-sided 0.05 level of

significance. The proof of strong control of the type I error rate for testing PFS and OS in the ITT population and PFS in the subgroup populations is shown in Section 3.2. Because of the closed sequential testing property, the family-wise error rate is strongly controlled for both the primary endpoint and the key secondary endpoint.

Based on OS results in the second IA, the planned number of OS events needed to trigger the FA may be increased if the observed treatment effect is promising, but not large enough to yield the likely conclusion of statistical significance at the end of the study using the originally planned number of OS events. It is also possible for the entire study design to remain unchanged as a result of the IAs. The Cui-Hung-Wang test statistic [4] will be used in the FA of OS to protect the type I error.

4.2   Independent Data Monitoring Committee

An independent data monitoring committee (IDMC) supported by an independent statistician from an independent statistical center (ISC) will review safety and efficacy data at planned IAs. The IDMC will provide a recommendation regarding study continuation based on the safety and efficacy parameters. If the study is terminated early on the basis of the IDMC recommendation, the sponsor will notify the appropriate regulatory authorities.

In addition, the IDMC will periodically review safety data at regularly scheduled meetings prespecified in the IDMC charter. The first formal safety review will occur after approximately 60 subjects (36 in the ixazomib arm and 24 in the placebo arm) have been randomized and received at least 1 cycle of study treatment. Subsequently, periodic safety reviews will also occur as prespecified in the IDMC charter.

Study accrual will not be interrupted because of the scheduled safety reviews. The IDMC or ixazomib study team may request an ad hoc meeting for any reason, including a significant unexpected safety event, unplanned unblinding of study results, follow-up of an observation during a planned IDMC meeting, or a report external to the study, such as publication of study results from a competing product. At each review, subject incidence rates of AEs (including all serious AEs [SAEs], treatment-related AEs, serious treatment-related events, and events resulting in the discontinuation of study drug) will be tabulated. Listings and/or narratives of "on-study" deaths and other serious and significant AEs, including any early withdrawals due to AEs, will be provided. Records of all meetings will be archived. The IDMC will communicate major safety concerns and recommendations regarding study modification or termination to the sponsor. Further details will be provided in the IDMC charter. At the 2nd IA if OS significance is not claimed, the conditional power based on OS will be calculated. During the closed session of the IDMC meeting at the 2nd IA, the IDMC

will compare the conditional power for OS based on the interim results with the prespecified effect size adaptation rules, and recommend to the sponsor executive committee the final adaptation decision. This recommendation will be documented in the IDMC closed meeting minutes.

4.3   Independent Review Committee

An independent review committee (IRC), blinded to treatment arm assignments and investigator determination of response, will review all disease evaluation data between screening and PD (including PFS follow-up period; does not apply to PFS2 assessment) from the study and determine disease status (response and progression), according to IMWG criteria [1] as specified in the IRC charter. IRC will not review disease evaluation data during next line therapy for those patients who have reached PD2. Data from the IRC will not be provided back to the investigator during the conduct of the study.

5.   STATISTICAL METHODOLOGY

In general, summary tabulations will be presented by treatment arm and will be displayed by the number of observations, mean, and standard deviation, median, minimum, and maximum for continuous variables, and the number and percentage per category for categorical data. The Kaplan-Meier (KM) survival curves and 25th, 50th (median), and 75th percentiles will be provided along with their 2-sided 95% confidence intervals (CIs) for time-to-event data.

5.1   Sample Size Justification

The primary objective of this study is to determine whether ixazomib improves PFS as compared with placebo. To obtain an adequate statistical power for OS, however, the study will not be stopped after the PFS analysis even if a significant PFS is observed.

A total of approximately 700 patients will need to be randomized in a 3:2 ratio into the 2 treatment arms, assuming an average enrollment rate of approximately 9 patients/month for the first 12 months and approximately 20 patients/month thereafter. The total sample size is calculated on the basis of maintaining 80% power to test the OS. The study is also adequately powered to test PFS. There are 2 planned IAs and 1 FA. The first IA will be the FA (and the only analysis) for PFS for statistical testing purposes. If the test for PFS is significant at the first IA, then OS will be tested at this first IA and at the subsequent IA, and at the FA if needed.

The first IA for OS (FA for PFS) will be performed when approximately 392 IRC-assessed PFS events have been observed or approximately 10 months after the last patient has been enrolled, whichever occurs later. The PFS will be tested in the ITT population with a 2-sided alpha = 0.04. In addition, PFS will be tested in parallel with 2-sided alpha = 0.01, using the Hochberg testing approach, in 3 prespecified subgroups: 1) ISS stage III, 2) patients aged ≥75 years, and 3) patients who had a CR or VGPR to initial therapy.

With 392 IRC-assessed PFS events, the study will have 90% power to detect a hazard ratio for PFS of 0.71 (median PFS of 11 months for control vs 15.5 months with treatment) using a 2-sided log-rank test at a 2-sided alpha level of 0.04 and assuming a drop-out rate of approximately 20% at Month 20. This will be the FA for PFS for statistical testing purposes, with the opportunity to claim PFS benefit. If the test for PFS is not statistically significant in any population (in the ITT population or any of the 3 subgroups), the study will be deemed unsuccessful, and no further testing will be conducted.

If the test for PFS is significant at the first IA, OS will be tested. If the OS results are statistically significant at either the first or second IA, the study can be stopped early, and this OS analysis will be the FA for formal hypothesis testing of OS. Otherwise, determination of whether the final number of OS events might increase will occur at the second IA.

The total event size calculation for OS is based on the adaptive sample size reassessment approach, which, in this study, is an adaptive event size reassessment approach. The minimum event size of 295 death events is based on an optimistic assumption of a hazard ratio of 0.71 (ie, median OS of 70 months for the ixazomib arm vs 50 months for the placebo arm, for a 41% improvement with ixazomib), with 80% power at a 2-sided level of significance of 0.04. The O'Brien-Fleming alpha spending function (the Lan-Demets method) is used to calculate the significance boundary on the basis of the observed number of death events at each IA, with a total of 295 OS events for the FA.

The second IA for OS will be performed when approximately 206 death events have been observed. If OS significance is not claimed, the conditional power based on OS will be calculated. If the conditional power falls in the promising zone, the event size will be determined according to a prespecified event size adaptation rule, with an event cap of approximately 393 death events. No futility analysis will be performed in the study.

The event size adaptation rule is a prespecified stepwise function to avoid the problem of back calculation resulting from an event size corresponding to either barely promising or highly promising interim results. The event size adaptation rule will be designed by the

sponsor's independent design statistician and approved by the sponsor's head of biostatistics. Neither the independent design statistician nor the head of biostatistics is involved in the conduct of this study.

The adaptation rules will be outlined in a separate document and will not be accessible to the sponsor's study team until completion of the study. The rules will be available only to the sponsor's independent design statistician, the sponsor's head of biostatistics, the IDMC, and the statistics representative on the sponsor's executive committee (if different from the sponsor's head of biostatistics).

5.2   Randomization and Stratification

The randomization scheme will be generated by an independent statistician who is not on the study team. Before dosing, a randomization number will be assigned to each patient. The randomization assignment will be implemented by an interactive voice/web response system (IXRS).

Eligible patients will be randomized in a 3:2 ratio to receive ixazomib or placebo treatment arms, stratified by initial therapy (proteasome inhibitor-containing or not), ISS status before initial therapy (stage I or II vs stage III), age (< 75 vs ≥ 75 years) at randomization, and best response to initial therapy (CR or VGPR vs PR) assessed by investigators

5.3   Blinding and Unblinding

This is a double-blind study: all study personnel including the investigators, site personnel, study clinicians, and the sponsor will be blinded to the treatment assignments. Only the ISC and IDMC will have access to unblinded individual patient level data in the electronic data capture system. The periodic safety analyses will be generated for the IDMC by an ISC. The formal IA analyses will also be conducted by an ISC for the IDMC. One unblinded submission working group might be formed to prepare a submission package and/or work with Agencies for submission purposes at IA.

Refer to Section 4.2 for the roles and responsibilities of the IDMC.

5.4   Data Handling

5.4.1   Methods for Handling Missing Data

All available efficacy and safety 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. For PRO data, primarily missing data imputation will be based on published instrument specific methods. Other missing data imputation methods such as last observation carried forward (LOCF) and the multiple imputation method may be explored as sensitivity analyses for PRO data.

5.4.1.1 Missing/Partial Dates in Screening Visit

The following rules apply to dates recorded during the Screening visits:

- If only the day component is missing, the first day of the month will be used if the year and the month are the same as those for the first dose of study drug. Otherwise, the 15th will be used.

- If only a year is present, and it is the same as the year of the first dose of study drug, the 15th of January will be used unless it is later than the first dose, in which case the date of the first of January will be used, unless other data indicate that the date is earlier.

- If only a year is present, and it is not the same as the year of the first dose of study drug, the 15th of June will be used, unless other data indicate that the date is earlier.

5.4.1.2 Missing/Partial Dates in Adverse Events/Concomitant Therapies/Subsequent Therapies

5.4.1.2.1 Missing/Partial Dates in Adverse Events

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 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 same as the start date.

Adverse events with stop dates that are completely or partially missing will be imputed 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 are missing, then impute date to 31st of December, in the year of last dose.

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.

5.4.1.2.2   Missing/Partial Dates in Concomitant Therapies

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 date occurs prior to start date, then keep the imputed date same as the start date.

5.4.1.2.3   Missing/partial dates in subsequent therapies

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

- When month and year are present and the day of the month is missing,

  - If the onset month and year are the same as the month and year of last dose with study drug, the day of last dose + 1 will be imputed.

  - If the onset month and year are not the same as the month and year of last dose with study drug, the first day of the month is imputed.

- When only a year is present,

  - If the onset year is the same as the year of last dose with study drug, the date of last dose + 1 will be imputed.

  - If the onset year is not the same as the year of last dose with study drug, the first day of the year is imputed.

- If no components of the onset date are present the date of last dose + 1 will be imputed.

5.4.2 Definition of Reference Values

Unless otherwise specified, the reference for assessment during this study will be based on the value collected at the time closest to, but before, the start of study drug administration.

The reference for response assessment during this study will be based on the value collected at the time of initial diagnosis. For the purpose of assessing PD, the disease nadir will be considered as study entry (or sometime later as appropriate).

5.4.3 Windowing of Visits

All data will be categorized based on 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).

5.4.4 Justification of Pooling

All data from all sites will be pooled. Study center or treatment-by-center interaction will not be included in any statistical analysis.

5.4.5 Withdrawals, Dropouts, Loss to Follow-up

Time to event parameters will be censored if patients withdraw, drop out, or are lost to follow-up before documentation of the events (PD/death). Rules for censoring are detailed in Section 5.8.

5.5 Patient Disposition

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 participating in any follow-up, patients with dose escalation at C5D1, 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.

5.6 Demographics and Baseline Disease Characteristics

5.6.1 Demographics

Demographics will be summarized by treatment groups in a descriptive fashion in the ITT population. Baseline demographic data to be evaluated will include age, sex, race, ethnicity,

height, weight, and other parameters as appropriate. Patient enrollment by region and country will also be summarized by treatment groups.

5.6.2 Medical History

General medical history and prior medications will be listed for all patients by treatment groups.

Prior induction regimens will be summarized by PI containing, IMiD containing, corticosteroids containing, akaylator containing, monoclonal antibody, as appropriate.

The duration of prior induction regimens will be summarized by treatment groups.

5.6.3 Disease Status at Initial Diagnosis

Efficacy data including serum M-protein, urine M-protein, and serum involved FLC, serum FLC ratio will be summarized for ITT population. Other characteristics include type of myeloma, β_2 - macroglobulin, albumin, Durie-Salmon stage, lactate dehydrogenase (LDH), cytogenetics, international staging system stage (ISS), revised international staging system stage (RISS), lytic bone, and extramedullary disease. Time from initial diagnosis to first dose of study treatment, and time from the first dose of induction to study entry will be summarized for all patients.

5.6.4 Disease Status at Study Entry

Disease characteristics at study entry includes, but are not limited to, Eastern Cooperative Oncology Group (ECOG) performance status, serum M-protein, urine M-protein, serum involved FLC and its ratio, serum creatinine and its category (≤ 2, > 2 mg/dL), 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, lytic bone lesions, extramedullary disease, and comorbidity status will be summarized for all patients.

A patient's type of myeloma is determined by heavy chain type (IgG, IgA, IgM, IgD, IgE, and other) 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:

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

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

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

Distribution of stratification factors will also be summarized.

5.6.4.1   Extent of disease at study entry

The following categories of extent of disease at study entry will be summarized: bone marrow aspirate (number of patients, % plasma cells), bone marrow biopsy (number of patients, % plasma cells, marrow cellularity, Kappa/Lambda ratio), combined % plasma cells in bone marrow aspiration and biopsy, and plasmacytomas.

5.6.5   Bone Marrow Cytogenetic at Initial Diagnosis

High risk cytogenetic categories are defined as (1) del17 group: patients with del17 alone (2) Cytogenetic high-risk group: patients with any of the following cytogenetic abnormalities: del17, t(4;14) , t(14;16). The standard risk group in the high-risk category is defined as patients for whom the test del17, t(4;14) , t(14;16) are normal. "Unclassifiable" is defined as patients who do not have cytogenetic data that can be categorized to high risk or standard risk corresponding to high risk group, either because of missing, unknown or indeterminate results. (3) Cytogenetic Expanded high-risk group: patients with any of the following abnormalities: del17, t(4;14) , t(14;16), or ampl 1q. The standard risk group corresponding to expanded high risk group is defined as patients for whom del17, t(4;14) , t(14;16) and ampl 1q are normal. "Unclassifiable" is defined as patients who do not have cytogenetic data that can be categorized to expanded high risk or standard risk corresponding to expanded high risk group, either because of missing, unknown or indeterminate results.

The percentage of each category will be summarized.

5.7 Treatments and Medications

5.7.1 Concomitant Medications

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 End of Treatment (EOT) 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. A by-patient listing will also be presented for concomitant medications.

5.7.2 Study Treatments

Following the Screening period, eligible patients will be randomized to receive ixazomib or placebo in a double-blind fashion with the randomization ratio of 3:2, respectively.

Ixazomib Arm: Patients will receive ixazomib citrate on Days 1, 8, and 15 of a 28-day cycle.

Placebo Arm: Patients will receive placebo capsule on Days 1, 8, and 15 of a 28-day cycle.

In both arms, a starting dose of 3 mg ixazomib or matched placebo will be used for all patients through Cycle 4. Upon evaluation of toxicities at the completion of Cycle 4, if during the most recent 2 cycles (Cycle 3 and 4), there have been no nonhematologic AEs $\geq$ Grade 2 related to study drug, no dose interruptions related to study drug toxicities, and no delays of greater than 1 week in starting a cycle due to study drug toxicities, the dose may be escalated to 4 mg at Cycle 5 Day 1. Patients who have had any dose reductions will not get dose escalated. If dose escalation was inadvertently missed at Cycle 5, escalation may be performed with permission from the Millennium project clinician or designee.

Patients will receive study treatment for a maximum duration of approximately 24 months (26 cycles), or until documented disease progression or intolerable toxicities, whichever comes first.

The number of patients who have been escalated at Cycle 5 Day 1, and number of patients who didn't escalate at Cycle 5 Day 1 will be summarized. For those who didn't escalate, their reasons of no escalation will be displayed.

5.7.2.1 Duration of Follow-up

The duration of OS follow-up is defined as time from the randomization date to the death or last known visit. If a subject is alive, this patient is treated as an event in OS follow-up and the duration equals to the date of a subject is last known to be alive – randomization date + 1. If a subject is dead, this patient is censored for OS follow-up and the duration equals to date of death - randomization date + 1. The Kaplan Meier (K-M) approach will be used to calculate the median duration of follow-up.

5.7.2.2 Extent of Exposure

A summary of drug exposure to Ixazomib/placebo will be characterized by number of treated cycles, numbers and percentages of patients who had ≥1, ≥2 , …, and = 26 treated cycles, total amount of dose taken, total number of dose taken, and relative dose intensity (%), by each treatment group in the safety population. Aggregate summary of numbers and percentages of patients who had 1-4, 5-8, 9-12, 13-16, 17-20, 21-24, 25-26 treated cycles will also be presented in the same table. Extent of exposure (days), which is calculated as (last dose date of study drug – first dose date of study drug + 1), will also be presented.

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

Relative dose intensity (%) is defined as 100 * (Total amount of dose taken) / (Total prescribed dose of treated cycles). Total prescribed dose of treated cycles is calculated as: for patients who were escalated at or after C5D1, it equals number of prescribed doses per cycles * dose prescribed at enrollment (3mg) * 4 cycles + dose prescribed at C5D1 (4mg) * number of prescribed doses per cycle * (number of treated cycles - 4). For patients who were not treated more than 4 cycles, it equals dose prescribed at enrollment (3mg) * number of prescribed doses per cycle * number of treated cycles.

Relative dose intensity will also be displayed as <50%, 50% - <= 80%, 80% - < 100%, = 100%, and > 100%. The duration of treatment at 4 mg 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. In addition, relative dose intensity will be calculated for those escalated to 4 mg at Cycle 5, counted only for doses starting from Cycle 5.

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

5.7.2.3 Treatment Modifications

Dose modification on each study drug due to adverse event will be summarized by Cycle 1 - 26, Cycles 1- 4, 5 - 8, 9- 12, 13 - 16, 17 -20, 21 -24, and 25 - 26 and total for each treatment group in the safety population. Action on drug will be summarized using the similar manner.

5.8 Efficacy Analyses

All efficacy evaluations will be conducted using the ITT population unless otherwise specified.

5.8.1 Primary Efficacy Endpoint

There is 1 primary endpoint, PFS, which is defined as the time from the date of randomization to the date of first documentation of PD or death due to any cause, whichever occurs first. Patients without documentation of PD will be censored at the date of last response assessment. The details regarding the handling of missing assessments and censoring for PFS analysis are presented in Table 5-1.

Table 5-1   Censoring Rules for PFS Primary Analysis Based on FDA Guidance

Situation	Date of Progression or Censoring	Outcome
No randomization and/or no post randomization assessment, no subsequent anticancer therapy after study treatment, no death	Date of randomization	Censored
Disease progression documented prior to randomization**	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.

**Patients will be considered major protocol deviation, and excluded from PP population.

5.8.1.1 Primary Efficacy Endpoint

PFS will be analyzed when approximately 392 IRC-assessed PFS events have occurred or approximately 10 months after the last patient has been enrolled, whichever occurs later. A 2-sided, stratified log-rank test will be used to compare the treatment groups with respect to PFS at a 2-sided alpha level of 0.04 for ITT population, and a 2-sided alpha level of 0.01 in 3 prespecified subgroups (see Section 3.2) using Hochberg testing approach. In addition, an unadjusted stratified Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect using the stratification factors. The K-M survival curves and K-M median PFS (if estimable), along with their 2-sided 95% CIs, will also be provided for each treatment group. PFS assessed by IRC in ITT population will be the primary analysis.

Sensitivity analyses for PFS include:

1. PFS assessed by investigator will be analyzed in the ITT population.

2. PFS assessed by IRC will be analyzed in the PP population if more than 5% patients are excluded from this analysis.

3. PFS assessed by IRC using the missing assessment and censoring rules based on EMA guidance with two combined alterations from FDA guidance as presented in Table 5-2.

Table 5-2   Handling of missing assessment and censoring for PFS Sensitivity Analysis based on EMA guidance

Situation	Date of Progression or Censoring	Outcome
Alternate antineoplastic therapy 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

PFS assessed by IRC evaluations to which different censoring mechanisms have been applied will be analyzed in the ITT population, for example, not censoring for patients who discontinue treatment and go on alternative antineoplastic therapy. Sensitivity analyses will be performed on the basis of one alteration at a time, not on combined alterations unless specified otherwise. Additional sensitivity analysis for PFS might be conducted on treating start date of alternate antineoplastic therapy as events.

In addition, a stepwise Cox model may be implemented to identify potential predictive factors using relevant demographic or diagnostic covariates, with the entry level fixed at 0.25 and a stay level fixed at 0.10. Besides treatment and the stratification factors, the model

may include the following significant covariates including, but not limited to, treatment arm, age; race (white; non-white); ECOG score at study entry (0 or1, 2), cytogenetic test (high risk, other), , ISS (I, II or III), revised ISS (I or II, III), , frailty status, MRD status at study entry, etc.

The plan of subgroups for PFS is presented in the Table 5-3 below with a few identified key subgroups:

Table 5-3   List of subgroups

Subgroup	Definition of Group
Age	< 65 years, $\geq$ 65 and < 75 years, $\geq$ 75 years
Age	< 75 years, $\geq$ 75 years (as one stratification factor)
Pre-induction ISS stage	I, II, III
Pre-induction ISS stage	I or II, III (as one stratification factor)
Revised ISS stage at initial diagnosis	I, II, III, unclassifiable
Sex	male, female
Race	white, Asian, other
Region	APAC, EMEA, other
Best response to initial therapy	CR or VGPR, PR (as one stratification factor)
Response at study entry	CR, VGPR, PR
Cytogenetic risk	high risk group vs. standard risk group corresponding to high risk group vs. unclassifiable group; expanded high risk group vs. standard risk group corresponding to expanded high risk group vs. unclassifiable group
Induction regimen	PI exposed, PI naive (as one stratification factor)
Induction regimen	IMiD exposed, IMiD naive
Frailty Status	fit, unfit, frail
MRD status at study entry	Known positive, known negative, unknown
Patient ineligibility categories to SCT *	age, comorbidity, other

* Patient ineligibility to SCT will be categorized by hierarchical order. First identify age, then identify comorbidity among the rest, and then group all the rest as other.

Additional exploratory analysis may be performed if deemed necessary.

5.8.2   Key Secondary Efficacy Endpoint

There is 1 key secondary endpoint: OS.

Overall survival is defined as the time from the date of randomization to the date of death. Patients without documentation of death at the time of analysis will be censored at the date last known to be alive. Overall survival will be analyzed based on the ITT population.

A 2-sided, stratified log-rank test will be used to compare the treatment groups with respect to OS. The test significance level at the IA and FA is decided by the O'Brien-Fleming alpha spending function (the Lan-DeMets method [4]). In addition, an unadjusted stratified Cox model will be used to estimate the hazard ratio and its 95% CIs for the treatment effect using the stratification factors. The 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.

In addition, a stratified stepwise Cox regression model may be used to further evaluate the treatment effects on OS after adjusting for some prognostic factors. Besides treatment and the stratification factors, the model may include the following significant covariates including, but not limited to, treatment arm, age; race (white; non-white); ECOG score at study entry (0 or1, 2), cytogenetic test (high risk, other), ISS (I, II or III), revised ISS (I or II, III), , frailty status, MRD status at study entry, etc.

To adjust for the potential effects of subsequent therapies after patients discontinued study treatment, the following two methods will be used:

- Marginal Structural Models (MSM) by Robins, Hernan and Brumback (2000) [5]

- Inverse Probability of Censoring Weighted (IPCW) method by Robins and Finkelstein (2000) [6]

In the MSM and IPCW analyses, in order to derive weights adjusting for the time-fixed and time-varying confounding effects due to taking alternative therapies, the covariates affecting disease progression, post-progression treatment, and OS endpoint will be used. Potential time-fixed covariates at study entries are region (APAC, EMEA, other), age (<75, ≥ 75), race(white, non-white), ECOG score (0 or 1, 2), induction therapy(PI exposed , PI naive), induction therapy (IMiD exposed, IMiD navie), response to prior therapy (CR or VGPR, PR), , presence of extramedullary plasmacytomas (yes, no), presence of lytic bone lesions (yes, no), hemoglobin, platelets, creatinine clearance, albumin, corrected calcium, LDH and MRD status. Potential time-fixed covariates at initial diagnosis are type of myeloma (IgA, other), ISS (I or II, III), RISS (I or II, III or unclassifiable), cytogenetic abnormalities (high risk, others), , MRD status at study entry (known positive, known negative, unknown), frailty status at study entry(fit, unfit, frail), patient eligibility to SCT categories . Time-varying covariates include duration of exposure, disease progression status at each study visit, hemoglobin value at each study visit and progression/relapse visit, platelets value at each study visit and progression/relapse, M-protein value at each study visit and progression/relapse. The final criteria for selected covariates would need to be statistically have a p-value of less than or equal to 0.1 in the multivariate logistic regression models for

weight calculations. If there are more than 5% missing in the baseline covariate, then this covariate will be dropped from the weighting calculation and final OS model. For both MSM and IPCW analyses, logistic regression models on repeated measurements will be used to approximate the Cox models in the weight derivations from which stabilized weights will be derived per subject per observation. Adjusted K-M curves will also be presented along with hazard ratios (HRs), 95% confidence intervals for HRs, and adjusted p-values based on MSM and IPCW approaches. SAS proc PHREG procedure with counting process type of data input, which takes multiple observations per subject, will be used as the final Cox model for OS for both MSM and IPCW approaches, where robust variance will be used to accommodate covariance introduced by correlated longitudinal observations within each subjects and other extra variabilities due to departure from model assumptions.

Subgroup analyses will be performed for OS following table 5-3.

5.8.3   Other Secondary Efficacy Endpoints and Analyses

Other secondary efficacy parameters include best response achieved or maintained before PD or to subsequent therapy, time to progression, PFS2, time to start of the next line of therapy, time to end of the next line of therapy, duration of the next line of therapy, OS and PFS in high-risk population, conversion of MRD positive to MRD negative, maintenance of MRD negativity, correlation between MRD status and PFS/OS, and correlation between frailty status and PFS/ OS.

Disease response-related endpoints prior to PD will be analyzed using IRC-assessed responses.

Other efficacy analyses such as PFS or OS from start of induction therapy, may be performed as needed.

<u>Best Response achieved or maintained prior to PD or subsequent therapy</u>

The time frame for a response is determined from the start of the study treatment until confirmed PD or subsequent therapy, whichever comes earlier.

The percentage of response (PR, VGPR, CR, or sCR) will be determined relative to the ITT population.

For patients entered into study with PR response, the percentage of maintaining PR as best confirmed response, and convert to VGPR/CR as best confirmed response will be displayed. For patients entered into study with VGPR response, the percentage of maintaining VGPR as best confirmed response, and convert to CR as best confirmed response will be displayed.

For patients entered into study with CR response, the percentage of maintaining CR as best confirmed response will be displayed.

For patients entered into study with CR response, the percentage of maintaining CR at 12 and 24 months will be displayed; duration of CR will be summarized descriptively using the Kaplan-Meier method.

The IRC-assessed response data will be used for the analysis. Investigator-assessed response data will be used for the sensitivity analysis.

Time to Progression (TTP)

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 analysis will be censored at the date of last response assessment. Patients who take alternative antineoplastic therapy prior to progression, or die during treatment will also be censored at the date of last response assessment. Time to progression will be analyzed based on the ITT population using the similar method as PFS. The subgroup analysis of TTP will be analyzed following Table 5-3.

Progression-Free Survival 2 (PFS2)

Progression-free survival 2 is defined as the time from the date of randomization to the date of first documentation of PD (as assessed by investigator) on the next-line antineoplastic therapy or death due to any cause, whichever occurs first.

Progression-free survival 2 will be analyzed based on the ITT population as assessed by investigator using the similar method as PFS

The details of the handling of missing assessments and censoring are presented in Table 5-4 and Table 5-5.

Table 5-4   Censoring for PFS2 For Those Who have Received Second line Therapy following Study Treatment

Situation	Date of Progression or Censoring	Outcome
Documented death or disease progression during second line therapy	Date of death/disease progression	Event
No documented death or disease progression during second line therapy	Date of last visit date	Censored
Lost to follow-up, withdraw consent before any documented death or disease progression during second line therapy	Date of last visit date	Censored
Start of third line therapy prior to the disease progression during second line therapy	Date of starting the third line therapy	Censored

If a patient has no response assessment during second line therapy, it will be censored at first dose of second line therapy.

Table 5-5   Censoring for PFS2 for Those Who have not received Second Line of Therapy

Situation	Date of Progression or Censoring	Outcome
No documented death	Date of last visit	Censored
Death	Date of death	Event

In addition, one sensitivity analysis for PFS2 might be conducted on treating start date of 3rd line therapy as events if the patients have not experience PD on the second line yet.

Time to start of the next line of therapy

Time to start of the next line of therapy is defined as the time from the date of randomization to the date of the first dose of the next line of antineoplastic therapy, for any reason.

Time to start of next line therapy will be analyzed based on the ITT population using the similar method as PFS. Patients who have not started the second line therapy will be censored at date of last known to be alive.

Time to End of Next-Line Therapy

Time to end of next-line therapy is defined as the time from the date of randomization to the date of last dose of next antineoplastic therapy following study treatment or death due to any cause, whichever occurs first.

Time to end of next-line therapy will be analyzed based on the ITT population using the similar method as PFS. Patients who have not completed the next line of therapy will be censored at date of last known to be alive.

Duration of Next-Line Therapy

Duration of next line therapy is defined as the time from the date of first dose of next line therapy to the date of the last dose of the next antineoplastic therapy or death due to any cause, whichever occurs first.

Duration of next line therapy will be analyzed on those patients who actually received next line therapy following the study treatment using the ITT population. Patients who are still on treatment on the next line of therapy will be censored at date of last known to be alive. Duration of next line therapy will be summarized using the Kaplan-Meier method.

PFS and PS in Cytogenetic Risk population

PFS and OS will be analyzed in high risk group using a similar method as those in the ITT population.

- Cytogenetic high risk group defined as patients carrying any of the following cytogenetic abnormalities: del17, t(4;14), or t(14;16)

- Cytogenetic expanded high-risk group defined as patients carrying any of the following cytogenetic abnormalities: del17, t(4;14), t(14;16) or ampl 1q

In addition, PFS and OS analyses will be summarized by individual cytogenetic abnormalities del17, t(4;14), t(14;16) and ampl 1q if data permit.

Correlation Between Frailty Status and Progression-Free Survival and Overall Survival

Patients' frailty status will be assessed on the basis of 4 components: age, the Charlson Comorbidity Scoring System, the Katz Index of Independence in Activities of Daily Living, and the Lawton Instrumental Activities of Daily Living Scale.

Specifically, ages of < 75, 75 to 80, and > 80 years correspond to frailty scores of 0, 1, and 2, respectively. Charlson Comorbidity Scoring System scores of ≤ 1 and ≥ 2 correspond to frailty scores of 0 and 1, respectively. Katz Index of Independence in Activities of Daily Living scores of > 4 and ≤ 4 correspond to frailty scores of 0 and 1, respectively. Instrumental Activities of Daily Living Scale scores of > 5 and ≤ 5 correspond to frailty scores of 0 and 1, respectively. The sum of the 4 frailty scores equals the total frailty score.

A total frailty score of 0 corresponds to a frailty status of fit; a total score of 1, to unfit; and a total score of 2 or more, to frail.

An unadjusted stratified Cox model including frailty status (fit, unfit, frail) and treatment will be used to estimate the hazard ratio and 95% CIs for the treatment effect and frailty status using the stratification factors for both PFS and OS. A status of "fit" will be compared with a status of "unfit" or "frail."

5.9 Pharmacokinetic and Biomarker Analysis

5.9.1 Pharmacokinetic Analyses

Plasma concentration-time data will be presented in listings and summarized by time point in tables.

Pharmacokinetic (PK) data collected in this study will 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.

5.9.2 Biomarker Analysis

CCI

5.9.3 Minimal Residual Disease Analysis

Minimal residual disease (MRD) will be assessed at study entry, at Cycle 12/13 (after approximately 12 months of treatment) and at EOT (approximately 24 months) in all the VGPR and CR patients independent of study arm, unless already done with the most recent 2 cycles.

MRD status by response at study entry for each treatment arm and overall population will be summarized. MRD status by response at C12/13, EOT for VGPR and CR patients for each treatment arm and overall population will be summarized. PFS and OS by MRD subgroup analyses will be analyzed according to MRD subgroups listed below.

- For overall population: MRD+ vs. MRD- at study entry

- For patients who were MRD- at study entry: treatment vs. control.

- For patients who were MRD+ at study entry and converted to MRD – at any time point post study entry: treatment vs. control.

- For patients who were MRD+ at study entry and still maintain MRD+ at any time point post study entry: treatment vs. control.

The rate of maintaining MRD negativity at Cycle 12/13, EOT and any time point post study entry for patients who were MRD negative at study entry will be compared between treatment and control arms. The rate of maintaining MRD negativity at Cycle 12/13, EOT and any time point post study entry for patients who were CR and MRD negative at study entry will be compared between treatment and control arms. The rate of maintaining MRD negativity at Cycle 12/13 and EOT for patients who were MRD negative at study entry may be compared between treatment and control arms in selected subgroups if data permit.

The rate of converting to MRD negative by 3-month interval and any time point post study entry for patients who were MRD positive at study entry will be made between treatment and control arms. The rate of converting to MRD negative at any time point post study entry for patients who were CR and MRD positive at study entry will be made between treatment and control arms. The rate of converting to MRD negative at any time point post study entry for patients who were MRD positive at study entry may be made between treatment and control arms in selected subgroups if data permit.

Time from MRD negative to MRD positive, PD or death will be summarized using the Kaplan-Meier method.

Time from MRD positive at study entry to first MRD negative post study entry will be summarized using the Kaplan-Meier method.

Limit of Detection of MRD will be summarized for all MRD evaluable patients.

MRD analysis on maintenance of MRD negativity, conversion to MRD negativity and correlation with PFS and OS based on different threshold (10^{-4}, 10^{-5}, 10^{-6}) may be performed if data permit.

Additional sensitivity analysis will be performed for MRD by assuming patients are MRD positive when the MRD status is missing or not evaluable.

5.10   Analyses of Patient-Reported Outcomes and Health Economics

5.10.1   Patient-Reported Outcomes (PROs)

Descriptive Statistics

Health-related QOL will be assessed using the cancer-specific EORTC QLQ-C30
(Table 5-6) and EORTC QLQ-MY20 (Table 5-7). The EORTC QLQ-C30 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 has 20 items across 2 functional subscales and 2 symptoms scales.

Raw scores from both the EORTC QLQ-C30 and MY20 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.

PRO assessments using the EORTC QLQ-C30 and the EORTC QLQ-MY20 will be
analyzed using patients with PRO measurements at study entry and at least one post study
entry measurement in the ITT population.

The descriptive statistics of actual value and change from study entry 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. Specifically, two different groupings of
observations may be used:

1.) based on cycle visit, and

2.) based on 4-week intervals from study entry.

Table 5-6   Definition of Subscale Scores of EORTC QLQ-C30

Subscale	Individual Items
Physical functioning	1-5
Role functioning	6-7
Emotional functioning	21-24
Cognitive functioning	20, 25
Social functioning	26-27
Quality of life	29-30
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

Abbreviations: EORTC QLQ-C30 = European Organization for Research and
Treatment of Cancer Quality of Life Questionnaire Core 30.

Table 5-7   Definition of Subscale Scores of EORTC QLQ-MY20

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

Abbreviations: EORTC QLQ-MY20 = European Organization for Research and
Treatment of Cancer Quality of Life Questionnaire Multiple Myeloma Module.

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 either a
stable score or an improvement in score from study entry based on minimally important
differences (MIDs) of 10 (primary analyses) [7] - [9] and 5 (sensitivity analyses) [10] will be
summarized by treatment group over time. Specifically, patients with a change from study
entry for the better of ≥MID will be classified as "improved". Those with no change in score
from study entry or a change in score within MID will be classified as "stable".

Analysis based on Linear mixed effects model by incorporating covariates

For the summary score and each subscale score of EORTC QLQ-C30 as well as each subscale score of QLQ-MY20, the change in score from study entry to each scheduled treatment cycle visit will be analyzed using repeated measures linear mixed models. These models will include the following covariates: treatment group, score at study entry, stratification factors, visits, and interactions between treatment group and visits. The interaction term between score at study entry and visit may also be considered as a covariate. The estimated mean change in score from study entry with 95% CIs for each treatment group will be provided at each treatment cycle visit. In addition, the mean difference in the changes from study entry between the treatment groups with 95% CIs and p-values will be provided at each treatment cycle visit. Analyses based on the 4-week interval timescale may be considered.

Change from study entry scores using cumulative distribution function (CDF) figures

The change in score from study entry to Cycle 26 (or last visit prior to Cycle 26) will be presented using cumulative distribution function (CDF) figures. The x-axis represents the changes in score (range: -100 to 100) and the y-axis represents the cumulative percentage of patients with a given change in score.

Missing Data

Details of scoring and initial handling of missing data are included in the EORTC QLQ-C30 and QLQ-MY-20 scoring guidelines.

Missing data patterns will be examined. As sensitivity analyses, different imputation methods for missing data including Last Observation Carry Forward (LOCF) and imputing death as worst possible score, and pattern mixture model may be performed if appropriate after examining missing data patterns.

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

5.10.2 Health Economics Analysis Using Medical Resource Utilization and Utility

EQ-5D-5L item scores as well as EQ-5D-5L Visual Analogue Scale (VAS) will be summarized in descriptive statistics for treatment arms over time in a manner consistent with the analyses above.

Compliance for EQ-5D-5L will also be summarized by number of expected and number and percentage of received by treatment group over time.

HU data will be summarized in descriptive statistics of medical encounters (number and rates of encounters, reasons for encounters, and length of stay). Categories of interest include ICU visits, non-ICU visits, emergency department visits, and outpatient visits.

Further modeling may be performed separately at post hoc analyses.

5.11 Safety Analyses

Safety will be evaluated by the incidence of AEs, severity and type of AEs, and by changes from baseline in the patient's vital signs, weight, and clinical laboratory results using the safety population. Exposure to the study drug regimen and reasons for discontinuation will be tabulated.

5.11.1 Adverse Events

5.11.1.1 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). All AEs will be presented in a by-patient listing. Treatment-emergent AEs (TEAEs) 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.

Adverse events will be tabulated according to MedDRA by system organ class, high level term, and preferred term 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 (ie, those events reported by ≥ 10% of patients in either treatment group)

- Serious AEs (SAEs)

- Drug-related SAEs

- AE resulting in study drug discontinuation

- AEs in Cycle 1-4 that prevented dose escalation at Cycle 5

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.

Treatment-emergent AEs will also be summarized by the National Cancer Institute Common Terminology Criteria for Adverse Events (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) of new primary malignancy and AEs of clinical importance (AECI) of peripheral neuropathy, rash, encephalopathy, liver impairment, hypotension, heart failure, arrhythmias, myocardial infarction, thrombocytopenia, neutropenia, gastrointestinal, and renal impairment will be presented.

Incidence of New Primary Malignancies

Two types of incidence rates will be calculated for the safety population based on the new primary malignancy (NPM) 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.

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.

Time to Resolution and Improvement of Peripheral Neuropathy (PN) Events

Peripheral neuropathy (PN) is defined as the treatment emergent AEs including neuropathy peripheral, peripheral motor neuropathy, peripheral sensory neuropathy, and peripheral sensorimotor neuropathy.

A PN event is considered as resolved if its final outcome is resolved with no subsequent PN event of the same preferred term occurring on the resolution date or the day before and after. A PN event is considered as improved if the event improves from the maximum grade (meaning that all the grades recorded after the maximum grade are less than the maximum grade).

Time to resolution and time to improvement are to be defined for each PN event. Time to resolution is defined as the time from the initial onset date (inclusive) to the resolution date for resolved events. Time to improvement is defined as the time from the initial onset date (inclusive) of the maximum grade to the first onset date that the toxicity grade is below the maximum grade with no higher grade thereafter, or the resolution date, whichever occurs first.

Time to improvement and time to resolution of PN events will be summarized by outcome (improvement or resolution) using the Kaplan-Meier method. The K-M survival curve and K-M medians (if estimable), along with their 2-sided 95% CIs, will be presented. This analysis is event based, thus 1 subject could contribute multiple observations if the subject has more than 1 PN event.

The analysis may be conducted for patients with any PN events or those with ≥ 2 PN events or those with ≥ 3 PN events if data permit.

5.11.1.2 Serious Adverse Events

The number and percentage of patients experiencing at least 1 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).

5.11.1.3   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).

5.11.1.4   Adverse Events Resulting in Discontinuation of Study Drug

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

5.11.2   Laboratory Data

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 of (x-1)% (mainly for $< 5\%$ for CR) will be used.

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, absolute neutrophil count (ANC), platelets counts, lymphocytes and leukocytes

- Serum chemistry: blood urea nitrogen (BUN), creatinine, total bilirubin, urate, lactate dehydrogenase (LDH), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), glucose, corrected calcium, sodium, potassium, chloride, carbon dioxide (CO_2), magnesium, and phosphate

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 CTCAE grade. Parameters to be tabulated will include:

- Hematology: ANC, hemoglobin, platelets, lymphocytes and leukocytes

- Serum chemistry: ALT, AST, ALP, creatinine, total bilirubin, corrected calcium, magnesium, potassium, sodium, and phosphate

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/SGPT, alkaline phosphatase, and total bilirubin).

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

5.11.3   Electrocardiograms

Descriptive statistics for the actual values and changes from values at study entry in Electrocardiograms (ECGs) will be listed by time point.

QTc interval will be calculated using Bazett's correction and Fridericia's correction, if necessary. The formulas are:

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

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

$$\text{where } RR = 60 / \text{heart rate (bpm)}$$

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 ≥ 500 msec) 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.

Electrocardiogram abnormalities will be presented in a data listing.

5.11.4   Vital Signs

The actual values of vital sign parameters including temperature, blood pressure, heart rate, respiratory 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.

5.11.5   Eastern Cooperative Oncology Group (ECOG) Performance Status

Eastern Cooperative Oncology Group Performance Status and shifts from study entry to post study entry assessment over time, and ECOG score frequency table over time will be summarized. Shifts from study entry to the worst post study entry score will be tabulated by treatment arm.

5.11.6 Other Safety Assessments

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. Tables will be provided with a summary of the patient incidence of all TEAEs of clinical importance by PT, severity, and seriousness for each analysis set within each category of TEAEs of clinical importance.

6. CHANGES TO PLANNED ANALYSES FROM PROTOCOL

Reference materials for this statistical plan include Clinical Study Protocol C16021 amendment 2 (Protocol amendment dated 14 June 2016) and amendment 5 (Protocol amendment dated 28 September 2018).

7. PROGRAMMING CONSIDERATIONS

7.1 Statistical Software

SAS version 9.2 (or higher) will be used for all analyses.

7.2 Rules and Definitions

Patient populations are defined in Section 2.

Values at study entry are defined in Section 5.4.2.

8. 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. Lan K, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70:659-63.

3. Bretz F, Maurer W, Brannath W and Posch M. A graphical approach to sequentially rejective multiple test procedures. Statist. Med. 2009; 28: 586-604

4. Cui L, Hung HM, and Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics* 1999; 55:853–857.

5. Robins JM, Hernan M, Brumback B. Marginal strucural models and causal inference in epidemiology. Epidemiology 2000; 11(5): 550-560.

6. 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.

7. 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.

8. 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.

9. Kvam 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.

10. 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.

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
PPD	Biostatistics Approval	24-Sep-2019 15:22 UTC
	Biostatistics Approval	24-Sep-2019 15:23 UTC
	Clinical Science Approval	24-Sep-2019 15:44 UTC
	Pharmacovigilance Approval	24-Sep-2019 15:59 UTC
	Clinical Science Approval	25-Sep-2019 00:26 UTC