Official Title of Study:

A Phase 3, Randomized, Open Label Trial of Lenalidomide/dexamethasone With or Without Elotuzumab in Subjects with Previously Untreated Multiple Myeloma

NCT Number: NCT01335399

Document Date (Date in which document was last revised): November 04, 2016

STATISTICAL ANALYSIS PLAN

A PHASE 3, RANDOMIZED, OPEN LABEL TRIAL OF LENALIDOMIDE/DEXAMETHASONE WITH OR WITHOUT ELOTUZUMAB IN SUBJECTS WITH PREVIOUSLY UNTREATED MULTIPLE MYELOMA

PROTOCOL CA204006

FINAL v2.1

Approved v1.0 930048245 4.0

TABLE OF CONTENTS

LIST OF TABLES

[redacted]

2 STUDY DESCRIPTION

2.1 Study Design

This is a Phase 3, open-label, multi-center trial investigating lenalidomide/low-dose dexamethasone with and without elotuzumab in subjects with newly diagnosed, untreated MM. The study population will consist of subjects who:

- are newly diagnosed with symptomatic MM
- have not received any prior systematic anti-myeloma therapy AND
- have measurable disease AND
- are not candidates for high-dose therapy plus stem-cell transplantation (SCT) because of age (≥ 65 years) or coexisting conditions
- Have met all the eligibility criteria (outlined in Protocol Section 3.3).

Eligible subjects will be randomized in a 1:1 ratio to receive either lenalidomide/low-dose dexamethasone (Ld) or elotuzumab/lenalidomide/low-dose dexamethasone (E-Ld). The randomization will be stratified as described in Section 2.2. Approximately 750 subjects will be randomized.

Subjects should receive Ld with or without elotuzumab in 28-day cycles until disease progression, unacceptable toxicity, or the subject meets other criteria for discontinuation of study drug (see Protocol Section 3.5), whichever occurs earlier.

Tumor assessments will be conducted every four weeks relative to the first dose of study medication, until disease progression, death or withdrawal of consent. Subjects who discontinue study medication for reasons other than progression e.g. for toxicity, should continue to have tumor assessments until progression, even if they have gone on to receive subsequent anti-myeloma therapy. Subjects will be followed at least every 16 weeks after disease progression for survival and subsequent myeloma therapies. The primary analyses of efficacy endpoints (using European Group for Blood and Bone Marrow Transplant [EBMT] criteria) will be based on a blinded review of tumor assessments by an IRC.

Crossover between the arms is not permitted.

Blood samples for pharmacokinetic (PK) and immunogenicity analysis will be collected in all subjects who received elotuzumab, prior to elotuzumab infusion on Day 1 of Cycles 1 through 4, 6, 9, 12, 15 and 18, every 3rd cycle after cycle 18 (21, 24, 27, etc..) as well as at the end of the study or discontinuation of treatment, and 30-day and 60-day follow-up visits. In addition, serial PK samples including 30 minutes post end of infusion and two hours post end of infusion will be drawn in Cycles 1, 2 and 3.

For full details of the schedule and timing of assessments see Protocol Section 5.

2.2 Treatment Assignment

After a subject's eligibility for randomization is established and informed consent has been obtained, the subject's treatment will be assigned via an interactive voice response system (IVRS). Subjects will be randomized to Ld or E-Ld in a 1:1 ratio. The randomization will be stratified by:

- Stage of disease (International Staging System (ISS) 1 - 2 versus 3)
- Age (< 75 years old versus ≥ 75 years old)
- Eastern Cooperative Oncology Group (ECOG) Performance status (0 versus 1 - 2)

and will be carried out via permuted blocks within each stratum.

2.3 Blinding and Unblinding

This is an open label trial. However, in order to maintain equipoise, the study team, including biostatisticians and programmers, will not have access to the randomization codes until the study is completed. Only the DMC and the reporting independent statistical group for the DMC, which is external to BMS, and completely independent from the study team, will be provided with the randomization information prior to the completion of the study.

The formal interim analysis of PFS will be reviewed by the DMC after at least 70% of the target events per IRC have occurred. The DMC will then notify the Sponsor to unblind the PFS data if the pre-specified criterion for statistical significance has been met.

The primary analysis of PFS and tumor response will be based on disease assessments made by the IRC. The assessments will be made retrospectively and the IRC will be blinded to the treatment arms.

2.4 Independent Data Monitoring Committee and Other External Committees

Two independent committees will be established for this study: (i) a DMC and (ii) an IRC.

An independent DMC will be established before the first subject has been treated and will monitor safety throughout the study. The first comprehensive safety review meeting will take place at the same time when 120 subjects have been randomized and followed up for at least 2 months in study [redacted] Subsequent comprehensive safety review meetings will take place every eight months after that. In addition, the DMC will be provided with serious adverse event (SAE) tables every four months. The DMC will also review the efficacy and safety data at the interim analyses to make formal decisions on stopping the trial for superiority (see Section 0). A separate charter will provide further guidance and describe the activities of this committee, including details on the statistical stopping rules.

An IRC will be responsible for reviewing all tumor assessment data to determine the best response and date of progression, based on the EBMT criteria. IRC-determined response and progression will be used in the primary analyses of PFS and ORR. A separate charter describes the activities of this committee.

2.5 Interim Analyses

The independent DMC will review the interim analysis of PFS data after at least 338 (70%) progression events (per IRC) have been observed and all subjects have been randomized. This formal comparison of PFS will allow for early stopping for superiority. The stopping boundaries will be derived based on the O'Brien and Fleming α spending function and will depend on the actual number of events observed at the time of the interim analysis. However, if there were exactly 338 events, the DMC could stop the study for superiority if the p-value were ≤ 0.0148. The significance level for the final look after 482 events would then be 0.0455.

If more than 338 events have occurred by the time of the interim analysis, different stopping boundaries will be calculated in East v5.1. It should be noted that survival follow-up would continue even if the study were stopped at the interim analysis of PFS for superiority.

In addition, an interim analysis of OS will be conducted at the time both the PFS and ORR analyses are positive. The nominal significance levels for the interim and final survival comparisons will be determined by an O'Brien and Fleming alpha spending function.

3 OBJECTIVES

3.1 Primary

The primary objective of this study is to compare the PFS of E-Ld versus Ld in subjects with newly diagnosed, untreated MM.

3.2 Secondary

The secondary objectives are:

- To compare the objective response rate (ORR) between treatment arms
- To compare the overall survival (OS) between treatment arms.
- To assess PFS rates at 1, 2, 3, 4, and 5 years
- To compare the change from baseline of the means score of pain severity and the change

from baseline of the mean score of pain interference using the Brief Pain Inventory - Short Form (BPI-SF) between treatment arms.

[redacted]

[redacted]

4 ENDPOINTS

The primary endpoint is PFS based on IRC tumor assessments using the primary definition of PFS (see section 4.1.1.3). Supportive analyses of PFS are:

[redacted]

The secondary endpoints are:

- ORR. The primary analysis will be based on IRC tumor assessments using the EBMT assessment criteria. Supportive analysis of ORR will be based on investigator assessment of best response.
- OS
- Patient-reported Outcomes (BPI-SF)
- PFS rates at 1, 2, 3, 4, and 5 years

4.1 Endpoint Definitions

4.1.1 Progression-Free Survival (PFS)

4.1.1.1 Definition of an Adequate Tumor Assessment

In the analysis of PFS, subjects who do not progress are censored. A non-progressing subject can be censored on the date of a tumor assessment only if there is sufficient information to rule out progression. An "adequate" tumor assessment visit for ruling out progression will require the following information:

- Serum monoclonal paraprotein results, if measurable at baseline, and

- Urine monoclonal paraprotein results, if measurable at baseline

4.1.1.2 Date of Progression or Censoring When Different Components of a Per Time Point Tumor Assessment Conducted at Different Times

As different tumor measurements may be conducted on different days, for instance, the blood draw for serum M-protein may be on a different date than 24-hour urine, the IRC and investigators were instructed to report the earliest date of the measurements associated with that time point for progression. In contrast, if tumor measurements are done on different dates and the subject is being censored, instructions were to report the latest date of the measurements associated with that time point.

4.1.1.3 Primary Definition of PFS

Primary Definition of PFS: PFS, under the primary definition, is the time from randomization to the date of the first documented tumor progression or to death due to any cause, provided progression or death does not occur more than 10 weeks (two or more assessment visits) after the last adequate tumor assessment. Clinical deterioration will not be considered progression.

The following censoring rules will be applied to the primary definition of PFS:

- Subjects who receive subsequent systemic anti-myeloma therapy prior to documented progression will be censored at the date of the last adequate tumor assessment prior to or on the initiation of the new therapy.
- Subjects who have an event (documented progression or death) > 10 weeks (2 assessment visits) after their last tumor assessment will be censored at their last adequate tumor assessment prior to the event.
- Subjects who neither receive subsequent therapy prior to progression nor have a progression event (including death) will be censored at their last adequate tumor assessment.
- In addition, subjects who do not have any post-baseline tumor assessments and who do not die within 10 weeks of randomization will be censored on the date of randomization.

In all cases, if there are no adequate assessments for censoring then the subject is censored on the date of randomization.

4.1.1.4 Secondary Definition of PFS – ITT Definition

Intent-to-Treat (ITT) Definition of PFS: PFS will also be analyzed applying an ITT definition that utilizes all of the data on each randomly assigned subject at the time of the PFS analysis. PFS under the ITT definition will be defined as the time from randomization to the date of the first documented tumor progression or death due to any cause. Clinical deterioration will not be considered progression. A subject who neither progresses nor dies will be censored on the date of their last adequate tumor assessment. A subject who does not have any post-baseline tumor assessments and who has not died will be censored on the date at which they were randomized.

4.1.2 Objective Response Rate (ORR)

Best overall response is determined as the best assessment based on all on-study efficacy data for the subject. Subjects with objective response are subjects with a best overall response on-study of partial response (PR) or better based on EBMT assessment criteria.

ORR is defined as the proportion of subjects with objective response among all randomized subjects.

4.1.3 Overall Survival

Overall survival is defined as the time from randomization to the date of death from any cause. If a subject has not died, their survival time will be censored at the date of last contact ("last known alive date"). A subject will be censored at the date of randomization if they were randomized but had no follow-up.

5 SAMPLE SIZE AND POWER

The number of events and power for PFS were calculated assuming an exponential distribution for each arm. One interim analysis of PFS is planned, and will allow for early stopping for superiority. An O'Brien and Fleming α spending function will be employed to determine the stopping boundaries.

Approximately 750 subjects will be randomized to the two treatment groups in a 1:1 ratio in order to obtain at least 482 progression events i.e. documented progressions or deaths. Four hundred and eighty-two events ensures that a two-sided 5% level sequential test procedure with one interim analysis will have 90.5% power if the median PFS times in the control and experimental arms were 25 and 33.75 months respectively; i.e. if the hazard ratio (HR) of the experimental arm to the control arm is 0.74.

The interim analysis of PFS will be conducted after 70% of the total events (i.e., at least 338 of the planned 482 events) have been observed.

Assuming an accrual rate of 24 subjects per month, it will take approximately 60 months to obtain the required number of events; 31 months of accrual and 29 months for PFS follow-up. At the time of final PFS analysis, an observed HR of 0.84 or less (which corresponds to a 4.8 months or greater improvement in median PFS (25 vs 29.8 months), would result in a statistically significant difference between the two treatment arms.

Survival, a key secondary endpoint, will be formally compared between the two arms at the time both ORR and PFS analyses are positive (whether that occurs after 70% of PFS events or all 482 events) and again after at least 354 deaths have been observed, which is expected to occur 7 years (84 months) after the initiation of the study. An O'Brien and Fleming α-spending function will be used to obtain the nominal significance level for the interim and final analysis of survival. Three hundred and fifty-four deaths ensure that a two-sided, α=0.05 level test procedure with one interim analysis will have 80% power to detect a 8.7 percentage point improvement in the 4-year survival rate[1] (from 59% to 67.7%, i.e. a HR= 0.74).

Power calculations were performed using East v5.1.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

There are three periods in this study: screening, on-treatment and post-treatment follow-up.

Screening: Most screening procedures must be done no more than 14 days prior to randomization. Some exceptions include efficacy assessments and ECGs, which can be done up to 28 days prior to randomization. See Section 5 of the Protocol for full details of study procedures and timings. For analyses purposes data collected within 60 days of randomization and prior to first dose will fall into the screening/baseline period. Site personnel should make every effort to initiate study drug treatment within three days after randomization.

On-Treatment Period: The on-treatment period has three phases: the first two cycles, cycles 3-18, and cycle 19 and beyond. All cycles are 28 days in duration. The first two cycles are characterized by weekly visits, safety assessments, and, in the experimental arm, elotuzumab infusions. Cycles 3-18 are characterized by fortnightly visits, safety assessments and, in the experimental arm, elotuzumab infusions. Cycles 19 and beyond, are characterized by once every 4 weeks visit, safety assessments and, in the experimental arm, elotuzumab infusions. Subjects should have myeloma urine and serum laboratory assessments every four weeks while on treatment until they progress and should have imaging for extramedullary plasmacytomas, if indicated, and at the time of CR/sCR assessment. The timing of the tumor assessments will be independent of dosing. Study treatment ends when the subject progresses, experiences unacceptable toxicity, or withdraws consent.

Follow-up Period: Subjects who discontinue study therapy prior to progression must continue to undergo tumor assessments on the same schedule used in the on-treatment period, regardless of whether they are receiving new anti-myeloma therapy. The only exception to this is if the subject withdraws consent for all study procedures or loses the ability to consent freely.

Follow-up for survival and subsequent myeloma therapy will be conducted every 16 weeks until the subject dies or the study ends.

6.2 Treatment Regimens

Subjects will be randomized to one of two treatment arms, E-Ld or Ld. Subjects will receive these treatments in 28-day cycles until disease progression, unacceptable toxicity or withdrawal of consent.

On Arm A (E-Ld) subjects receive:

- Elotuzumab: 10 mg/kg intravenous (IV) on Days 1, 8, 15 and 22 of the cycle during Cycles 1 and 2; Days 1 and 15 during Cycles 3-18; and 20mg/kg intravenous (IV) on Day 1 of the cycle during subsequent cycles.
- Lenalidomide: 25 mg orally (po) on Days 1-21 of each cycle.
- Dexamethasone: Administered on Days 1, 8, 15 and 22 of cycle. 28 mg po + 8 mg IV whenever given prior to elotuzumab; 40 mg po on weeks on which no elotuzumab is given.

On Arm B (Ld) subjects receive:

- Lenalidomide: 25 mg po on Days 1-21 of each cycle.
- Dexamethasone: 40 mg po on Days 1, 8, 15 and 22 of each cycle.

6.3 Populations for Analyses

The following subject populations will be used for the statistical analysis:

All enrolled subjects: All subjects who gave signed informed consent and who were entered in the IVRS.

Randomized subjects: All enrolled subjects who were randomized.

Treated subjects: This population includes all randomized subjects who received at least one dose of study medication (lenalidomide, dexamethasone or elotuzumab).

The analysis of baseline characteristics, efficacy and subject-reported outcomes will be carried out on the "Randomized subjects" population, with subjects grouped according to the treatment arm to which they were randomized. The analysis of extent of exposure and safety will be based on the "Treated subjects" population, with subjects grouped according to the treatment received, where treatment received is defined as the treatment arm to which they were randomized, unless they received the wrong treatment throughout the study.

7 STATISTICAL ANALYSES

7.1 General Methods

Continuous variables will be summarized using descriptive statistics; i.e. number of non-missing observations (n), mean, standard deviation (STD), median, minimum, maximum, first quartile and third quartile. Categorical variables will be summarized by frequencies and percentages. Minimum and maximum values will be reported to the precision as the data collected; mean, median and quartiles will be rounded to one decimal place more than the data collected; STD will be rounded to one decimal place more than the summary statistic (e.g. mean) for which it is calculated, and percentages will be rounded to one decimal place. Percentages greater than zero but less than 0.1 will be presented as "<0.1".

The Kaplan-Meier (KM) product limit method will be used to estimate the distribution and median of each time-to-event endpoint in which censoring is involved. The Brookmeyer and Crowley method[9] will be used to compute a 95% confidence interval (CI) for the median of each time-to-event endpoint using the log-log transformation of the survivor function S(t).

The stratified log-rank test will be performed to compare time to event distributions (PFS or OS) between treatment groups. Stratification factors will be obtained from IVRS randomization dataset. The stratified log-rank test will be based on the 'proc lifetest' procedure in SAS 9.2, in which the 'strata statement' specifies the stratification factors with 'group option' equal to randomized group.

A stratified Cox proportional hazards model will be used to compute an estimate of, and CI for, the hazard ratio of the investigational to the control arm, for each endpoint. Medians and hazard ratios for time-to-event variables will be rounded to two decimal places.

The main analyses of tumor response and progression-free survival will be conducted using the best overall response, date of response, and date of progression assigned by the IRC.

Laboratory results, adverse events, and other symptoms will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), Version 3.0, except where CTCAE grades are not available. Individual laboratory values will be presented in both the conventional US units and the International System of Units (SI). Adverse events will be

categorized using the most current version of Medical Dictionary for Regulatory Activities (MedDRA), by system organ class and preferred term. Prior therapies will be summarized using the most current version of the World Health Organization (WHO) drug dictionary.

Statistical analyses will be carried out in SAS (Statistical Analysis System, SAS Institute, North Carolina, USA), unless otherwise indicated.

7.2 Study Conduct

7.2.1 Accrual Patterns

Tables summarizing accrual by center, country, region (North America, European Union, and Rest of World) overall and by treatment group will be generated (see appendix 2 for a list of countries in each region). Subject accrual will also be summarized by the randomization stratification factors per IVRS (Stage of disease [ISS 1 – 2 versus 3], Age [< 75 years old versus ≥ 75 years old], ECOG Performance status [0 versus 1 – 2]), overall and by treatment group. In addition, this summary will also be presented based on the stratification information obtained from the baseline CRF pages. A cross tabulation of IVRS vs. baseline stratification factors will also be summarized.

7.2.2 Protocol Violations

A relevant protocol deviation is defined as a deviation from the protocol which is programmed in the database and which could potentially affect the interpretability of the study results.

The number and percentage of subjects with any relevant protocol deviation and each specific deviation will be presented overall and by treatment group.

The following eligibility deviations are considered relevant in this study and will be summarized for all randomized subjects:

1. Any prior systemic anti-myeloma therapy
2. Non-measurable disease. This occurs when none of the following three conditions are met:
    1. IgG, IgA, or IgM M-protein ≥ 0.5 g/dL
    2. Serum IgD M-protein ≥ 0.05 g/dL
    3. M-protein ≥ 200 mg in 24-hour urine.
3. No baseline tumor assessment. This occurs when there are no tumor assessments at all (laboratory assessments) on or prior to first day of dosing. See section 4.1.1.1, for definition of adequate tumor assessment.

Also, the following on-treatment deviations are considered relevant in this study and will be summarized for all treated subjects:

4. Non-protocol specified systemic anti-myeloma therapy prior to discontinuation of study therapy
5. Received non-assigned treatment regimen throughout the study.

6. Subjects continuing to receive study therapy after 10 weeks of first documented progression per investigator (as progression should be confirmed).

If assessments based on the central laboratory is unavailable, but the local laboratory results are available, then those will be used for evaluating deviations 2 and 3.

A by subject listing of relevant protocol deviations will be provided.

7.3 Study Population

7.3.1 Subject Disposition

A frequency table of enrolled subjects, broken down by whether or not they were randomized, and the reasons for not being randomized, will be produced.

Subject disposition during the treatment period will be summarized, overall and by treatment group. The number and percentage of subjects randomized, treated, still on treatment and discontinued from the treatment will be presented.

By-subject listings will also be produced to accompany the tables on enrollment and the subject disposition table.

In addition, a by-subject listing indicating whether the subject was included in each of the analysis populations will be provided.

7.3.2 Demographic and Subject Characteristics

Demographic and baseline characteristics will be summarized, overall and by treatment group. The following parameters will be summarized; age at the time of informed consent (years), age category at time of informed consent (<65 years, ≥65 years), and (<75 years, ≥75 years), gender (Male, Female), race (White, Black/African American, American Indian/Alaska Native, Asian, Native Hawaiian/Other Pacific Islander, Other), ethnicity – for US subjects only (Hispanic/Latino, Not Hispanic/Latino), weight (kg) and ECOG performance status (0,1,2) .

An accompanying by-subject listing of demographic and subject characteristics will be presented.

7.3.3 Disease Characteristics at Baseline

Baseline disease characteristics will be summarized, overall and by treatment group. The following parameters will be presented to summarize baseline disease:

- Serum M-protein (g/dL)
- Urine M-protein (mg/24 hours)
- Myeloma type (IgG, IgA, IgD, IgM, Light chain disease)
- Number of lytic bone lesions (0, 1-3, >3)
- Myeloma risk category
  - High risk: ISS stage II or III **and** t(4;14) or del(17p) abnormality

  - Low risk: ISS stage I or II **and** absence of t(4;14), del(17p) and 1q21 abnormalities and age < 55 years
  - Standard risk: any subjects not meeting the definition of high or low risk.
  - Not Evaluable: Subjects having missing data preventing the classification in the other 3 categories
- Individual FISH /Cytogenetic abnormalities (del 17p, t(14; 16), t(4; 14), and del 13)
- β2 microglobulin (mg/L) (<3.5, ≥3.5-5.5, ≥5.5)
- Albumin (g/L) (<3.5, ≥3.5)
- LDH (<300 IU/L, ≥300 IU/L)
- Soft tissue plasmacytomas (Yes, No)
- ISS Stage at enrollment (I, II, III)
- Time from disease diagnosis to randomization (months)

An accompanying by-subject listing of disease characteristics at baseline will be presented.

7.3.4 General Medical History

The number and percentage of subjects with any relevant medical history and by body system will be presented, overall and by treatment group.

By subject listings of medical history will be provided. A by subject listing of pregnancy test data will also be provided.

7.3.5 Baseline Safety Laboratory Tests

Baseline safety laboratory evaluations will be the last available samples taken on or before Cycle 1, Day 1 and within 60 days of randomization.

Baseline safety laboratory values will be presented by CTC severity grade, by treatment group. Separate tables will be generated for hematology (hemoglobin, WBC, ANC [neutrophils plus bands], ALC [absolute lymphocyte count] and platelets, liver function (ALT, AST, alkaline phosphatase, albumin, direct bilirubin and total bilirubin) and renal function/electrolytes (sodium, potassium, bicarbonate, calcium, glucose, and creatinine, creatinine clearance < 60 ml/min, ≥ 60 ml/min). Sodium, potassium, calcium, and glucose will be presented separately, based on their high and low values.

CTC grades will be derived as part of the analysis data set programming using version 3.0.

7.4 Extent of Exposure

7.4.1 Study Therapy

Lenalidomide, oral dexamethasone and IV dexamethasone are given at fixed doses and are not adjusted for body surface area or weight. The elotuzumab dose, in contrast, is adjusted for weight. Throughout this SAP, a subject's elotuzumab dose level at a particular time point will refer to the actual dose, in mg/kg, rather than their planned dose. A subject's actual elotuzumab dose level will be computed by dividing their total dose delivered, in mg, as recorded on the "Record of Study Medication - Elotuzumab" eCRF page, by their latest pre-dose weight, in kg, on Day 1 of that cycle.

7.4.1.1 Duration of Study Therapy

The number of cycles of treatment received by subjects will be summarized (using n, mean, STD, median, min, max, q1 and q3) by treatment group.

The duration of each treatment (elotuzumab, dexamethasone [oral, IV, and oral and IV combined] and lenalidomide), in months, will be calculated as:

$$(date\ of\ last\ dose\ of\ the\ drug - date\ of\ first\ dose\ of\ the\ drug + 1)/30.4375$$

Duration of each of the study treatments will be summarized by treatment group.

7.4.1.2 Dose Modifications

Elotuzumab:

- **Dose Reduction**

Reduction of Elotuzumab dosing is not permitted as per the protocol.

- **Dose Delay**

A delay for Elotuzumab will be computed based on records from the dosing CRF page. A delay in Elotuzumab dosing will be defined as the number of days from the start date of the previous infusion to the start date of the current infusion being > 8 days and ≤10 days (if in cycles 1, 2 or first infusion of cycle 3) or > 15 days and ≤ 21 days (from second infusion of cycle 3 - cycle 19) or > 29 days and ≤ 35 days (from cycle 20 onwards).

For courses with a delay of Elotuzumab infusion, the time of delay in infusion will be summarized in the categories: 2 - 3 days, 4 - 7 days.

The number and percentage of subjects with any dose delay, with 1, 2, 3, or ≥ 4 reported delays will be summarized. For subjects with a delay, their reason for delay (from the dose modification page of the CRF) will be summarized in the groups: "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error" and "other". For subjects without an available reason for delay, the category "unknown" will be used.

A by-subject listing for dose delay will be presented for all treated subjects. This listing will include all reported delays per CRF, regardless of whether the subject had a computed delay based on the actual dosing dates.

- **Dose Omission**

An omission for Elotuzumab will be calculated based on records from the dosing CRF page. If the interval between the two dosing dates for Elotuzumab is >10 days (in cycles 1, 2 or first infusion of cycle 3) or >21 days (from second infusion of cycles 3-19), or > 35 days from cycle 20 and beyond, the dose will be considered to have been omitted.

For courses with an omission of Elotuzumab infusion, the time between the previous and current infusion will be summarized in the categories: 11 - 14, 15 - 21, 22 - 27, 28 - 35, ≥ 36 days.

The number and percentage of subjects with a dose omission, and with 1, 2, 3, or ≥4 omissions will be summarized. For subjects with an omission, their reason for omission (from the dose modification page of the CRF) will be summarized in the groups: "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error" and "other". For subjects without an available reason for omission, the category "unknown" will be used.

A by-subject listing for dose omission will be presented for all treated subjects. This listing will include all reported omissions per CRF, regardless of whether the subject had a computed omission based on the actual dosing dates.

- **Dose Interruption**

The number and percentage of subjects with any interruption in their elotuzumab infusion, with 1, 2, 3, or ≥ 4 interruptions, with an interruption due to "infusion reaction", "infusion administration issues" and "other" will be presented. In addition, duration of interruptions (in minutes) will be summarized via descriptive statistics.

A by-subject listing will be generated for elotuzumab IV interruptions. This listing will include reason for interruption, whether the infusion was resumed, and duration of the interruption.

- **IV Rate Reduction**

The number and percentage of subjects with any elotuzumab IV rate reduction, with 1, 2, 3, or ≥ 4 rate reductions, and with rate reductions due to "infusion reaction," "infusion administration issues," and "other" will be presented.

A by-subject listing will be generated for elotuzumab IV rate reductions.

Lenalidomide:

- **Dose Reduction**

Reduction of Lenalidomide will be computed based on the actual dose received. In any study day (excluding, cycle 1, day 1), the drug will have a calculated reduction compared to the previous day, if the actual level of the administered dose is below the actual level of the administered dose in the previous instance. The information for this analysis will be derived programmatically, using the total daily dose on the "Record of Study Medication - lenalidomide" eCRF page.

The daily dose levels are defined as follows:

- Dose level 0 or full dose (25 mg): ≥ 20.0
- Dose level -1 (15 mg): 12.5 mg to < 20.0

- Dose level -2 or starting dose for moderate renal impairment (10 mg): 8.75 to < 12.5 mg
- Dose level -3 for subjects with severe renal impairment (15 mg every 48 hours): 6.25 to < 8.75 mg
- Dose level -4 for subjects with End Stage Renal Disease (ESRD) (5 mg): < 6.25 mg

The number and percentage of subjects with a dose reduction, and the lowest dose level achieved per subject (-1, -2, -3, ≤ -4), will be presented by treatment group. This table will be broken out by whether the subject started at the standard dose level or at the 10 mg dose level.

The reason for dose reduction as reported by the investigator will be tabulated for all instances with a calculated reduction based on the dose modification for Lenalidomide page. A category "unknown" will be defined for all calculated reductions with no reason reported by the investigator.

A by-subject listing of dose reductions will be generated. This will include all reported reductions per investigator, regardless of whether it met the requirements for a calculated reduction.

- **Dose Interruption**

In any cycle, the Lenalidomide dose will be considered as interrupted if there is a gap of two or more days between two dosing dates within that cycle where the subject did not receive the study drug. This will be calculated based on the information from the dose modification CRF for Lenalidomide.

The number and percentage of subjects with dose interruption, with 1, 2, 3, or ≥ 4 reported interruptions will be presented. For subjects with a computed interruption, their reason for interruption will be summarized from the dose modification CRF. For subjects without an available reason for interruption, the category "unknown" will be used.

Dexamethasone (IV) and Oral:

For IV dexamethasone the following summaries will be provided:

- The number of subjects with any dose delay, with 1, 2, 3, or ≥ 4 reported delays, and with a reason for delay of "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error" and "other".
- The number of subjects with any reported dose omission, with 1, 2, 3, or ≥ 4 reported omissions, and with a reported reason for omission of "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error" and "other".
- The number of subjects who discontinued the study drug, and reason for the discontinuation (hematologic toxicity, non-hematologic toxicity, AE, dosing error, or other).

- The number of subjects with any reduction in the dose of the drug, and with a reason for the reduction of "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error", and "other").

For PO dexamethasone the following summaries will be produced:

- The number of subjects with any dose modification, with 1, 2, 3, or ≥ 4 reported modifications, and with a reason for the modification of "hematologic toxicity", "non-hematologic toxicity", "AE", "dosing error" and "other").

By subject listings of record of study medication for each of IV dexamethasone and PO dexamethasone will be provided.

7.4.1.3 Cumulative Dose, Dose Intensity, and Relative Dose Intensity

In order to compute the cumulative dose of dexamethasone, IV dexamethasone will have to be converted to its equivalent oral dose. This will be done by treating each mg of IV dexamethasone as 1.32 mg of oral dexamethasone, i.e. making use of the fact that the mean bioavailability of oral dexamethasone was estimated to be 0.76[10]. A subject's cumulative dose of dexamethasone will be defined as the sum of all dexamethasone doses, oral and IV, converted to oral equivalent. The cumulative dose administered will be summarized for each of these drugs, by treatment group.

Elotuzumab:

A subject's cumulative dose of elotuzumab is measured in mg/kg, and is defined as the sum of their elotuzumab doses (mg/kg) (actual, not planned) over all infusions.

Elotuzumab dose intensity (mg/kg/week) per subject will be calculated separately for Cycles 1 and 2, Cycles 3-18 and Cycle 19 and beyond, since elotuzumab planned dose intensity differ in those cycles.

Cycle 1 and 2

$$7 \times \left(\frac{cumulative\ dose\ of\ elotuzumab\ during\ the\ first\ 2\ cycles}{\left(\min(date\ of\ first\ dose\ of\ elotuzumab\ in\ the\ last\ cycle\ among\ the\ first\ 2\ cycles+28,\ \ death\ date) - date\ of\ first\ dose\ of\ elotuzumab\ in\ Cycle\ 1\right)} \right)$$

Cycle 3-18:

$$7 \times \left(\frac{cumulative\ dose\ of\ elotuzumab\ \ during\ Cycle\ 3 - Cycle\ 18}{\left(\min(\ date\ of\ first\ dose\ of elotuzumab\ in\ last\ cycle\ among\ 18\ cycles + 28, death\ date) - date\ of\ first\ dose\ of\ elotuzumab\ in\ Cycle\ 3\right)} \right)$$

Cycle 19 and beyond:

$$7 \times \left(\frac{cumulative\ dose\ of\ elotuzumab\ starting\ from\ Cycle\ 19}{\left(\min(\ date\ of\ first\ dose\ of\ elotuzumab\ in\ last\ cycle\ \ + 28, death\ date) - date\ of\ first\ dose\ of\ elotuzumab\ in\ Cycle\ 19\right)} \right)$$

The relative dose intensity will be calculated as

Dose Intensity/Planned Dose Intensity (PDI)

Planned dose per week for elotuzumab is 10mg/kg/week for the first 2 cycles, 5mg/kg/week for cycles 3 and beyond. The PDI per subject for elotuzumab will be computed by averaging the planned doses per week over the treatment duration:

$$Elotuzumab\ PDI\ ((mg/kg/wk) = \frac{(\text{DOSE1} * 8 + \text{DOSE2} * (DUR - 8))}{DUR}$$

where DUR is defined as (min (date of death, date of first dose in last cycle in timepoint + 28) - date of first dose in timepoint) / 7.

Dose intensity and relative dose intensity of elotuzumab will be summarized for the E-Ld group only.

Lenalidomide:

A subject's cumulative dose of lenalidomide is defined as the sum of each dose taken, as recorded in their medication diary.

Dose intensity and relative dose intensity for lenalidomide (mg/week) will be calculated separately for subjects who began the study at the standard daily dose, 25mg/day and those who, because of moderate renal impairment, began the study at < 25mg/day (at 10mg/day dose will be assumed when computing the RDI as this is the recommended dose for subject with moderate renal disease per protocol).

The lenalidomide dose intensity, per subject, will be calculated as:

$$7X\left(\frac{cumulative\ dose\ of\ lenalidomide}{\left(\begin{array}{c}min(date\ of\ 1st\ dose\ of\ lenalidomide\ in\ last\ cycle\ \ in\ which\ lenalidomide\ was\ administered\\ +28, death\ date) - date\ of\ first\ dose\ of\ lenalidomide\end{array}\right)}\right)$$

The relative dose intensity of lenalidomide is the dose intensity per week divided by the planned dose intensity per week, 131.25mg (or 52.5mg for subjects who begin therapy at < 25mg (10mg/day per protocol)).

Dexamethasone:

The dose intensity (mg/week) of dexamethasone, per subject, will be calculated separately for the period encompassing Cycles 1 and 2, Cycle 3-18 and Cycle 19 and beyond since dexamethasone dosing is different in these 3 periods.

Cycle 1 and 2:

$$7 \times \left(\frac{cumulative\ dose\ of\ dexamethasone\ during\ the\ first\ two\ cycles}{\left(\begin{array}{c}\min(date\ of\ first\ dose\ of\ dexamethasone\ in\ last\ cycle\ among\ first\ 2\ cycles+28, death\ date) -\\ date\ of\ first\ dose\ of\ dexamethasone\ in\ first\ cycle\end{array}\right)}\right)$$

Cycle 3-18:

$$7 \times \left(\frac{cumulative\ dose\ of\ dexamethasone\ during\ Cycle\ 3 - Cycle\ 18}{\left(\begin{array}{c}\min(\ date\ of\ first\ dose\ of\ dexamethasone\ in\ last\ cycle\ among\ 18\ cycles + 28, death\ date) -\\ date\ of\ first\ dose\ of\ dexamethasone\ in\ Cycle\ 3\end{array}\right)}\right)$$

Cycle 19 and beyond:

$$7 \times \left(\frac{cumulative\ dose\ of\ dexamethasone\ starting\ from\ Cycle\ 19}{\left(\begin{array}{c} \min(\ date\ of\ first\ dose\ of\ dexamethasone\ in\ last\ cycle\ + 28, death\ date) - \\ date\ of\ first\ dose\ of\ dexamethasone\ in\ Cycle\ 19 \end{array} \right)} \right)$$

The relative dose intensity of dexamethasone is the dose intensity per week divided by the planned dose intensity per week. The planned dose of dexamethasone in the experimental arm on elotuzumab dosing days, 8mg IV dexamethasone plus 28mg oral dexamethasone, is equivalent to 38.56mg of oral dexamethasone. The planned dose in the experimental arm when dexamethasone is not being administered with elotuzumab is 40mg. Thus, the planned weekly dose of dexamethasone in the experimental arm for Cycles 1 and 2 will be equivalent to 38.56mg oral dexamethasone, the planned weekly dose of dexamethasone in the experimental arm for Cycle 3 - Cycle 18, during which IV dexamethasone is given on Days 1 and 15 only, will be equivalent to 39.28mg of oral dexamethasone and the planned weekly dose of dexamethasone in the experimental arm for Cycle 19 and beyond, during which IV dexamethasone is given on Days 1 only, will be equivalent to 39.64mg of oral dexamethasone . The planned weekly dose of dexamethasone in the control arm is 40mg for all cycles.

In the dose intensity formulas, if the patient took a dose on his death date then 1 day will be added to the denominator.

Descriptive statistics (n, mean, STD, median, min, max, q1 and q3) will be presented for the cumulative dose, and dose intensity of each agent, by treatment group. The number and percentage of subjects whose relative dose intensity falls into the following categories will be presented, by treatment group: ≥ 90 %, 80% to < 90%, 70% to < 80%, 60% to < 70%, < 60%.

7.4.2 Premedication Other than Dexamethasone for Hypersensitivity Reactions

A by-subject listing of pre-medication for elotuzumab, other than dexamethasone, will be provided. This listing will be generated from the "Pre-medication for Elotuzumab (other than Dexamethasone)" eCRF module and pre-medications will be coded using the BMS WHO drug dictionary.

7.4.3 Concomitant Medication

Concomitant medications are medications, other than study medication or pre-medications for elotuzumab recorded on the "Pre-medication for Elotuzumab" eCRF page, which are taken by subjects any time on-study, no earlier than the first day of study drug and no later than 60 days after the last dose of study drug. Concomitant medications will be coded using the BMS WHO drug dictionary. If there is clear evidence to suggest that the medication stopped prior to the first dose of study medication, the medication will be assumed to be 'Prior' only.

The number and percentage of subjects taking any concomitant medication and each medication (Anatomical Therapeutic Chemical [ATC] classification system drug name) will be summarized, overall and by treatment group.

A by subject listing of concomitant medication will be provided.

7.4.4 Discontinuation of Study Therapy

The number and percentage of subjects who have discontinued all study treatment and reason for discontinuation will be summarized by treatment group and overall using subject status eCRF page from end of treatment. This summary, unlike other safety analyses, will include all randomized subjects and will be grouped by treatment group as randomized. This is done in order to give a full accounting of all subjects who are off study treatment, including those who were randomized but never treated.

In addition, for subjects in the Elotuzumab arm, who discontinued one drug, while continuing at least one of the other two study drugs, their reason for treatment discontinuation will be summarized based on the dose modification CRF. Subjects will be counted in this summary if there is evidence from the dosing CRF that the subject received the combination therapy in at least one cycle, followed by additional cycles, where one of the drugs in the combination was discontinued.

7.4.5 Subsequent Anti-Myeloma Therapies

The number and percentage of subjects with any subsequent systemic anti-myeloma therapy (2nd line), reason for first subsequent systemic anti-myeloma therapy regimen (Documented progression of disease, Clinical deterioration without documented progression, Maintenance therapy without disease progression or clinical deterioration, Consolidation without documented disease progression or clinical deterioration, or Other) and all subsequent systemic anti-myeloma therapy agents given categorized using the BMS WHO drug dictionary will be presented by treatment group**.**

The number and percentage of subjects with any second subsequent systemic anti-myeloma therapy (3rd line) and reason for second subsequent systemic anti-myeloma therapy regimen (Documented progression of disease, Clinical deterioration without documented progression, Maintenance therapy without disease progression or clinical deterioration, Consolidation without documented disease progression or clinical deterioration, or Other) will be presented by treatment group.

The number and percentage of subjects with any subsequent surgery, type of first subsequent surgery (Kyphoplasty, Orthopedic surgery, Debulking surgery or Other) will be presented by treatment group.

The number and percentage of subjects with any subsequent radiation therapy will be presented by treatment group.

A by-subject listing of all subsequent surgery, radiation therapy, and systemic anti-myeloma therapy will be provided.

7.5 Efficacy

Efficacy analyses (PFS, ORR and OS) will be conducted on the population of all randomized subjects, grouped by arm assigned at randomization, unless otherwise noted.

The main analyses of PFS and ORR will be based on IRC evaluation. Unless stated otherwise, whenever a stratified analysis is specified, the stratifications factors will be those used in the randomization (per IVRS), that is:

- Stage of disease (ISS 1 – 2 versus 3)
- Age (< 75 years old versus ≥ 75 years old)
- ECOG Performance status (0 versus 1 - 2)

All p-values reported will be two-sided. Confidence intervals (CI) for the PFS and OS endpoints will be based on the nominal significance level adjusted for the interim analyses to preserve the overall type I error rate. Two-sided 95% CI will also be provided. CIs for the other endpoints will be at the two-sided 95% level. The p-values presented in the clinical study report will be rounded to the fourth decimal place. Point estimates and confidence bounds for efficacy variables will be rounded to the second decimal place.

7.5.1 Primary Analysis of Progression-Free Survival

The primary analysis of PFS will be to compare the two treatment arms via a two-sided stratified log-rank test. The primary definition of PFS using IRC assessments will be used in this analysis. The stratification factors will be the same as those used in the randomization. The two-sided log-rank p-value will be reported. An O'Brien and Fleming α spending function will be employed to determine the significance levels and stopping boundaries at the interim and final looks (see Section 5 for more details).

Further analysis of PFS will include estimation of the hazard ratio and estimation of the PFS distribution in each treatment group.

The estimate of the PFS hazard ratio, of E-Ld to Ld, will be calculated using a stratified Cox proportional hazards model, with the stratification factors used at randomization and treatment as the sole covariate. Ties will be handled using the Breslow method[11]. Two-sided, 100 * (1-α) (α adjusted for the interim analysis) and 95% CI for the hazard ratio will also be presented.

The PFS function for each treatment arm will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each arm will be computed via the Brookmeyer Crowley method. Estimates for 1, 2, 3, 4 and 5-years PFS rates when appropriate will be presented along with their associated 95% CIs. These estimates will come from the KM curve and their standard errors (SEs), for use in constructing CIs, will be computed using Greenwood's formula[12]and log-log transformation applied on the survivor function.

The method of Gail and Simon[13] will be used to test for a qualitative interaction between treatment and strata, stage of disease (ISS 1 – 2 versus 3), age (< 75 years old versus ≥ 75 years old) and ECOG Performance status (0 versus 1 – 2). This test will be conducted at the $\alpha = 0.10$ level.

The proportional hazards assumption will be assessed via the following hazard rate model, which contains a time dependent covariate:

$$\lambda(t,z) = \lambda_i(t)e^{(\{b_1+b_2\times[log(t)]\}\times Z)}, \qquad i = \{1-8\}$$

where i=1-8 corresponds to each of the eight levels the stratum can take, and Z is the treatment indicator, which is equal to 1 for the experimental combination arm and 0 for the control arm. The transformation of time that will be used is log (t). The null hypothesis, that the proportional hazards assumption is valid, i.e., that $b_2=0$, will be tested against the alternative hypothesis that $b_2 \neq 0$ using a Wald statistic.

[redacted]

[redacted]

7.5.5 Analysis of Objective Response

The number and percentage of subjects in each category of best overall response per IRC (stringent complete response [sCR], complete response [CR], very good partial response [VGPR], partial response [PR], minor response [MR], stable disease [SD], progressive disease [PD], or unable to determine [UD]) according to the IRC will be presented, by treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson[14]) will be presented, by treatment group. Also, the number and percentage of subjects in each category of best overall response according to the investigator tumor assessment will be presented by treatment group.

The main analysis of ORR will be based on the IRC assessment and will compare the two treatment groups using a CMH test stratified by the same factors used in the analysis of PFS. An estimate of the treatment odds ratio and corresponding two-sided 95% CI will be presented.

The same analysis will be repeated for the best response per IRC but modified for subsequent therapies. For this analysis only subjects with responses before, or on start of subsequent therapy date will be considered responders. The IRC when doing the blinded assessment of response and progression does not have data on subsequent therapy and discontinuation status. For this reason the IRC best response modified for subsequent therapy will be derived programmatically.

A 2-sided, 95% confidence interval for the difference of ORR between treatment arms will be computed for all randomized subjects by the method of DerSimonian and Laird [15] using a fixed-effects model (setting Δ^2 equal to zero), adjusting for the stratification factors. The weighted response rate difference and 95% CI can be determined using the following formula:

$$\hat{\theta} = \frac{\sum_{i=1}^{12} \hat{\theta}_i w_i}{\sum_{i=1}^{12} w_i} \sim N(\theta, 1/\sum_{i=1}^{12} w_i)$$

where $\hat{\theta}_i$ is the response rate difference of the i[th] stratum and $w_i = 1/var(\hat{\theta}_i)$.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of IRC best response versus the investigator best response will be presented, by treatment group. These measures of concordance will be provided:

- Number of subjects with same response divided by number of randomized subjects
- Number of subjects with same response evaluation: responders (PR or better), non-responder divided by number of randomized subjects

[redacted]

7.5.9 Analyses of Overall Survival

Overall survival will be compared between the treatment arms at the interim and final analyses, using a two-sided stratified log-rank test. The stratification factors will be those used in the analysis of PFS.

An O'Brien and Fleming α-spending function will be employed to determine the nominal significance levels for the interim and final analyses.

All analyses performed for PFS (detailed in section 7.5.1) will be repeated for OS. Supportive analyses 1, 2, 4 and 5 of PFS (detailed in section 7.5.2) as well as the subset analyses (detailed in section 7.5.4) will also be repeated for OS. These will be conducted only at the time of the final analysis of OS.

Estimates for 1, 2, 3, 4, and 5-year OS rates will be presented along with their associated 95% CIs when appropriate follow-up time has occurred. These estimates will come from the KM curve and their standard errors (SEs) used in constructing CIs, will be computed using Greenwood's formula[17] and the log-log transformation applied on the survivor function.

7.5.10 Currentness of PFS and OS Data

Time from last tumor assessment to data cut-off in months will be summarized by treatment arm and overall for all randomized subjects. Subjects who have a PFS event will be considered as current for this analysis. The ITT definition of PFS as per IRC will be used for this summary.

In addition Kaplan-Meier plots of time from randomization to post-baseline tumor assessment will be produced by treatment arm for the first twelve assessments.

Currentness of OS data will be summarized in months, by computing the time from "last known alive" date to data cut-off date. Subjects who have a death event will be considered as current for this analysis.

By-subject listings will also be produced to accompany the subject time from last tumor assessment table.

7.5.11 Hierarchy for Testing Secondary Endpoints

A hierarchical procedure will be used to preserve a two-sided family-wise type I error rate of 5% for testing one primary (PFS) and two key secondary endpoints (ORR and OS) in this study.

If there is a statistically significant improvement in PFS, ORR will be compared between treatment arms at the two-sided 5% alpha level. If, and only if, the comparison of ORR is statistically significant, then OS will be compared between treatment arms using a 5% sequential test procedure.

The two QOL secondary endpoints will not be part of the hierarchical testing and will be tested independently of the significance of the other endpoints. Each of the two comparisons will be made at the two-sided type I error of 2.5%.

7.6 Safety

Safety summaries will be based on the treated subject population, grouped by treatment regimen received.

7.6.1 Adverse Events

AEs will be categorized using the most recent version of the MedDRA, by system organ class (SOC) and preferred term. The severity of AEs will be graded using the NCI CTCAE (v 3.0).

On-study AEs are defined as non-serious and serious AEs with an onset date on or after the first dose until 60 days after the last dose. See Section 8.4, Imputing AE Onset Dates, for a discussion of imputation rules for incomplete or missing AE onset dates. If the relationship to study drug is missing, then the AE will be assumed to be related to study drug.

Unless specified otherwise, AEs will be counted only once within each SOC and preferred term, according to their worst CTC grade.

Tables will be sorted by SOC and preferred term, with SOCs ordered by decreasing frequency overall and then alphabetically. Preferred terms will be sorted within SOCs by descending frequency overall and then alphabetically. The sorting will be done based on the total column when arms are presented side-by-side.

Frequency tables of the worst grade of on-study AE will be presented. One summary table will present AEs broken out by individual grade, 1, 2, 3, 4 or 5 along with an Any Grade category. Another one will present AEs by grouped grades as follows "Any, Grade 3-4 and Grade 5." These summaries will be repeated for on-study drug-related AEs.

A by-subject listing all AEs will be presented.

In addition the following will be presented by treatment arm:

- Frequency of AE before and after cycle 18 for subjects on the Elotuzumab arm
- Exposure-adjusted AE incidence rates (including multiple occurrences of unique events) will be calculated for each SOC and preferred term.

Exposure-adjusted incidence rate per 100 person-years will be used and will be calculated as:

$$100 \times \frac{Total\ number\ of\ unique\ AEs}{\left(\frac{subject\ date\ of\ last\ dose\ of\ study\ drug - subject\ date\ of\ first\ dose\ of\ study\ drug + 60 + 1}{365.25}\right)}$$

and will be displayed along with a count of events. For these additional tables, AEs can be counted multiple times within each SOC and preferred term.

A by-subject listing of unique AEs will be provided.

7.6.2 Serious Adverse Events and Adverse Event Leading to Discontinuation

Summaries of worst grade of on-study SAE, both by individual grade (1, 2, 3, 4, or 5, along with an Any Grade category) and by grade grouped as "Any, Grade 3-4 and Grade 5" will also be presented. These summaries will be repeated for:

- On-study drug-related SAEs
- On-study AEs leading to discontinuation.
- Drug related AEs leading to discontinuation

By-subject listings of SAEs and AEs leading to study drug discontinuation will be produced.

7.6.3 Safety by subgroups

Summaries of all adverse events with ≥ 5 % frequency, adverse events leading to discontinuation, serious adverse events (any grade, grade 3-4), and death within 60 days of last dose will be presented for the levels of the factors listed below.

In addition summaries by the same factors will be produced for SPM and infusion reaction.

- Age:

  - < 65
  - < 65 up to < 75
  - ≥ 75 years
- Gender (male, female)
- Race
- Region (North America, European Union, and Rest of World)

7.6.4 Adverse Events of Special Interest

Infusion reaction is a known elotuzumab toxicity. It will be based on investigator assessment and will be defined as any non-serious or serious adverse event judged by the investigator to be infusion related and which also occurs on the day of or the day after the elotuzumab infusion. These summaries will be presented:

- Frequency tables of the worst grade of on-study investigator infusion reaction (any grade and grade 1 through 5) by treatment group. The same summary will be presented with grades grouped as Any, Grade 3-4 and Grade 5.
- Frequency of infusion reaction by infusion rate.
- Frequency of infusion reaction by cycle.

By-subject listings of infusion reactions will be produced.

Multiple myeloma is associated with immune dysfunction and the natural course of the disease includes increased infection risk. In addition, elotuzumab may inhibit some cellular components of the immune system. Therefore, a thorough characterization of infections will be presented. Infections will be based on the SOC 'infections and infestations' and opportunistic infections will be based on clinically pre-defined PT terms:

- Time to onset and duration of first infection
- Frequency of infections including opportunistic infections by treatment group.
- Absolute lymphocyte count at time of first infection
- Kinetics of lymphocyte reduction

7.6.5 Second Primary Malignancies

A summary table and a by-subject listing of second primary malignancies (SPM) will be provided. In addition, exposure-adjusted SPM incidence rates (including multiple occurrences of unique events) will be presented by treatment group and will be calculated as follows:

$$100 \times \frac{Total\ number\ of\ unique\ SPM\ events}{\left(\frac{subject\ death\ date\ or\ last\ known\ alive\ date - subject\ date\ of\ first\ dose\ of\ study\ drug + 1}{365.25}\right)}$$

Information on SPM will be obtained from the on-treatment eCRF page for secondary malignancies and the "Survival Status" eCRF page.

7.6.6 Deaths

The number and percentage of deaths and the investigator-reported cause of death will be presented, by treatment group and overall. This will be summarized for all deaths, and for those reported on study treatment or within 60 days of discontinuing study treatment.

7.6.7 Clinical Laboratory Evaluations

On-treatment laboratory tests for safety are defined as those that occur after first dose of any study therapy until 60 days after last dose of any study therapy.

The number and percentage of subjects with each worst severity grade for on-study hematology parameters (hemoglobin, WBC, ANC, ALC and platelets) will be presented, by treatment group. Grades will be categorized as Grade 1, Grade 2, Grade 3, Grade 4, Any Grade and Grade 3-4. Subjects will be counted only once for each parameter, according to their worst post baseline CTC grade. The percentage of subjects with each worst severity grade will be calculated out of the number of subjects with on-study assessment for each lab parameter. Subjects with no on treatment assessment for a lab parameter will be reported in the "Not Reported" category.

This summary will be repeated for:

- Liver parameters (ALT, AST, alkaline phosphatase, albumin and total bilirubin) with available CTC grades
- Renal/electrolyte parameters (sodium, potassium, bicarbonate, calcium, glucose and creatinine) with available CTC grades

Sodium, potassium, calcium, and glucose will be presented separately, based on their high and low values.

For blood urea nitrogen (BUN), direct bilirubin and total protein the worst category on-treatment will be presented. Results will be categorized as; below upper normal limits, above upper normal limits or not reported.

For reporting purposes, urea will be converted to BUN, using the conversion factor: urea (mmol/L) / 0.357 = BUN (mg/dL).

Subjects experiencing any potential drug induced liver injury (DILI) will be summarized by treatment group as follows:

- A summary of the number and percentage of subjects with (AST or ALT > 3 × upper limit of normal (ULN)) and (Total bilirubin > 2 × ULN and ALP < 2 × ULN) will be presented.

If any potential cases of DILI are identified then clinical review will be conducted to ensure no other immediate apparent possible causes of AT elevation and hyperbilirubinemia are present, including but not limited to, viral hepatitis, pre-existing chronic or acute liver disease, or the administration of other drugs known to be hepatotoxic.

Note: The timing for total bilirubin rising to >2xULN needs to be concomitant with or within 30 days after ALT or AST rising to >3xULN. ALP needs to be <2xULN at the time of or within 2 weeks before ALT or AST rising to > 3xULN or TBILI>2xULN. This means a normal ALP value (<= 2xULN) must be present within 2 weeks of either the ALT/AST criteria or the TBILI criteria.

A by subject listing of all laboratory data will be provided. A separate listing will be provided for all subjects potentially experiencing DILI.

7.6.8 Electrocardiograms

A by subject listing of electrocardiogram (ECG) data collected at screening will be presented.

7.6.9 Vital Signs, Physical Measurements and Physical Examination

Vital signs parameters (body temperature, respiratory rate, seated diastolic and systolic blood pressure and heart rate) will be presented by visit and time point in a subject listing.

[redacted]

[redacted]

8 CONVENTIONS

8.1 Baseline Definition

Baseline evaluations will be those performed within 60 days of randomization and prior to first dosing date. When an assessment is repeated multiple times within the screening period, the baseline evaluation will be the one closest to the first dosing date.

Laboratory tests and procedures (ECG, physical measurement and 2D echocardiogram MUGA, skeletal survey and plasmacytoma) done on the first date of dosing will be assumed to have occurred prior to dosing and therefore baseline evaluation for those parameters will be those prior or on the first dosing date.

8.2 Age Definition

Age (years) will be calculated as:

$$\frac{date\ of\ informed\ consent - date\ of\ birth + 1}{365.25}$$

8.3 Time Since Initial Diagnosis to Randomization Definition

Time from disease diagnosis to randomization (months) will be calculated as:

$$\frac{date\ of\ randomization - date\ of\ initial\ diagnosis + 1}{30.4375}$$

For the date of cancer diagnosis, the following conventions will be used for imputing partial dates:

- If only the day of the month is missing, the 15th of the month will be used to replace the missing day.
- If the day and month are missing or a date is completely missing, it will be considered as missing.

8.4 Duration and Study Day Definition

In instances in which study period between two dates are to be calculated (for example, duration of response, PFS and OS), the convention to be used is as follows: later date – earlier date + 1 day.

Study day is calculated as assessment date – first dose date + 1 day, if the assessment is taken on or after the first dose day. If the assessment is taken prior to the first dose day, study day will be calculated as assessment date – first dose date. Date of first dose is defined as Study Day 1.

8.5 Day Conversion of Date Imputation

Conversion from days to months or years:

- 1 year = 365.25 days
- 1 month = 30.4375 days

Imputation for partial or missing progression dates:

- If only the day is missing, the 1st of the month will be used to replace the missing day.
- If the day and month are missing or a date is completely missing, the date of progression will be moved back to the last complete tumor assessment date.

In both the cases given above, the imputed date will still be considered an event.

Imputation for partial or missing death dates:

- If only the day is missing, the later of the last known alive date and the 1st of the month will be used to replace the missing day.
- If both the day and the month are missing, the later of the last know alive date and Jan 1st will be used to replace the missing information.

In both the cases given above, the imputed date will still be considered an event.

For partial dates of start of subsequent anti-myeloma therapy, the following conventions will be used:

- When the day is missing, the alternative therapy will be assumed to start on the first day of the given month if this day is later than the last dosing date. Otherwise the alternative therapy will be assumed to start on the day following the last dosing date.
- When the day and the month are missing, the alternative therapy will be assumed to start on the first day of the given year if this day is later than the last dosing date. Otherwise the alternative therapy will be assumed to start on the day following the last dosing date.

8.6 AE, Laboratory Results and Concomitant Medication

Safety data will be handled according to the BMS safety data conventions (described in "Analysis of Safety Data - Reference to CT SOP 109"). This document includes descriptions on how to analyze AE data as well as how to handle partial dates, missing dates, and unknown end dates when analyzing safety data.

The following dictionaries will be used to code medical terms and to derive toxicity grades:

- Adverse events and other symptoms will be graded according to the NCI CTC Version 3.0 and categorized according to the latest version of MedDRA at the time of the database lock.
- Laboratory results will be classified according to the CTC Version 3.0 grading system.
- All medications will be coded as per the latest version of the WHO Drug dictionary at the time of analysis.
- Tables and listings for laboratory results will be available in SI and US units.

8.7 Topline Analyses

Outputs covering the following topics will be included in the top-line analyses:

Study conduct:

- Relevant protocol deviations (Table and Listing)

Subject population:

- End of Treatment Summary
- Accrual by Region
- Accrual by stratification factors at randomization
- Demographic Characteristics Summary
- IVRS Stratification Factors Summary
- Baseline Myeloma Characteristics Summary (SI units)
- Other Baseline Laboratorty Measurements Summary
- Baseline Myeloma (Cytogenetics and FISH) Risk Factors Summary
- Time from Diagnosis to Randomization Summary
- Baseline Physical Measurements Summary

Extent of exposure:

- Number of Cycles Summary

- Elotuzumab Dose Intensity Summary
- Dexamethasone Dose Intensity Summary
- Lenalidomide Dose Intensity Summary
- Elotuzumab Delay Summary
- Elotuzumab Omission Summary
- Elotuzumab Infusion Interruption Summary
- Lenalidomide Reduction Summary
- Lenalidomide Interruption Summary

Progression free survival:

- Currentness of Follow-Up for PFS (IRC, ITT definition) Summary
- PFS (IRC, Primary Def.) Analysis
- Kaplan-Meier plot of PFS (IRC, Primary Def.)
- PFS (IRC, Primary Def.) Analysis Adjusting for Different Baseline Covariates
- PFS (IRC, Primary def.) Hazard Ratio and 95% CI in Subsets (Forest plot)
- PFS (IRC, ITT Analysis)
- Kaplan-Meier plot of PFS (IRC, ITT)
- PFS (IRC, ITT) Analysis Adjusting for Different Baseline Covariates
- PFS (IRC,ITT) Hazard Ratio and 95% CI in Subsets (Forest plot)
- Kaplan-Meier plot of PFS (Inv., Primary Def.)
- Kaplan-Meier plot of PFS (Inv., ITT.)
- Concordance of PFS between the IRC and investigator
- Concordance and Timing of PFS per IRC and Investigator
- Assessment of Censoring Distributions with Reverse Kaplan Meier Plot of PFS (IRC, Primary Def.)
- Assessment of Censoring Distributions with Reverse Kaplan Meier Plot of PFS (IRC, ITT Def.)

Tumor response:

- Best Overall Response (IRC)
- Best Overall Response (Inv)
- Best Overall Response (IRC) Modified for Subsequent Therapy
- Best Overall Response (IRC) Odds Ratio and 95% CI in Subsets (Figure)
- Cross Tabulation of Modified Best Overall Response per IRC and Investigator
- Time to First Response (IRC) Summary
- Kaplan-Meier Plot of Duration of Response (IRC)

Overall survival (only when interim PFS and ORR analyses are positive):

- Currentness of Follow-Up for Overall Survival Summary Overall Survival Analysis
- Overall Survival Analysis Adjusting for Different Baseline Covariates
- Kaplan-Meier plot of overall survival
- Overall survival Hazard Ratio and 95% CI in Subsets (Forest plot)

Safety:

- Hematologic Laboratory Test Results Summary of Worst Toxicity Grade - SI Units
- Renal and Liver Laboratory Test Results Summary of Worst Toxicity Grade - SI Units
- Chemistry Laboratory Test Results Summary of Worst Toxicity Grade - SI Units
- Other Laboratory Test Results Summary (BUN, DBILI, TPRO)
- Potential Drug Induced Liver Injury Cases Summary - SI Units
- Adverse Event Summary by CTC Grade Combined
- Pre and Post Cycle 19 Adverse Event Summary by CTC Grade Combined
- Serious Adverse Event Summary by CTC Grade Combined
- Adverse Event Leading to Discontinuation Summary by CTC Grade Combined
- Infusion Reaction Summary by Common Terminology Criteria
- Second Primary Malignancy Summary
- Deaths
- Deaths within 60 days of last study drug

Specific tables corresponding to these outputs will be identified in a separate Data Presentation Plan.

APPENDIX 1: LIST OF ABBREVIATIONS

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
BMS	Bristol Myers Squibb
BPI - SF	Brief Pain Inventory – Short Form
CI(s)	Confidence Interval(s)
CMH	Cochran Mantel Haenszel
CR	Complete Response
CRF	Case Report Form
CT	Computerized Tomography
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
EBMT	European Group for Blood and Bone Marrow Transplant
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EMEA	European Agency for the Evaluation of Medicinal Products
EORTC	European Organization for Research and Treatment of Cancer
ESRD	End Stage Renal Disease
FDA	Food and Drug Administration
HAHA	Human Anti-Human Antibody
HRQOL	Health Related Quality of Life
IMiD	Immune Modulatory Drugs
IRC	Independent Review Committee
ITT	Intent-to-Treat
IV	Intravenous
IVRS	Interactive Voice Response System
Ld	Lenalidomide, (low-dose) Dexamethasone
E-Ld	Elotuzumab, Lenalidomide, (low-dose) Dexamethasone

LS	Least Squares
MedDRA	Medical Dictionary for Regulatory Activities
MM	Multiple Myeloma
MR	Minor Response
MRI	Magnetic Resonance Imaging
MUGA	Multi Gated Acquisition
NCI	National Cancer Institute
ORR	Objective Response Rate
OS	Overall Survival
PD	Progression
PFS	Progression-Free Survival
PFS-2	Progression-Free Survival After Next Line of Treatment
PK	Pharmacokinetic
PO	Oral
PPK	Population Pharmacokinetics
PR	Partial Response
QLQ-C30	Quality of Life Questionnaire – Core
QLQ-MY20	Quality of Life Questionnaire – Myeloma Specific Module
QoL	Quality of Life
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
sCR	Stringent Complete Response
SD	Stable Disease
SE	Standard Error
sMICA	Soluble Major Histocompatibility Complex Class I-related Chain A
SI	International System of Units (SI)
SOC	System Organ Class
STD	Standard Deviation

UD	Unable to Determine
ULN	Upper Limit of Normal
VGPR	Very Good Partial Response
WHO	World Health Organization

APPENDIX 2: LIST OF COUNTRIES BY REGIONS

Regions will be defined according the following country groupings. This list will be updated at the time of the analysis if subjects are enrolled in additional countries.

Region	Countries
North America (NA)	Canada, Mexico, Puerto Rico, USA
Europe (EU)	Austria, Belgium, Croatia, Czech Republic, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Poland, Portugal, Romania, Russia, Spain, Sweden, Switzerland, UK
Rest of the World (ROW)	Australia, China, Israel, South Korea, Malaysia, Philippines, South Africa, Singapore, Taiwan, Thailand, Turkey, Argentina, Brazil, Chile, Peru

9 DOCUMENT HISTORY

Version Number	Author(s)	Description
Version 1.0	[redacted]	
Version 2.0	[redacted]	[redacted]
Version 2.1	[redacted]	[redacted]